THE 5TH INTERNATIONAL CONFERENCE ON INTEGRATED MODELING AND ANALYSIS IN APPLIED CONTROL AND AUTOMATION

SEPTEMBER 12-14 2011
ROME, ITALY

EDITED BY
AGOSTINO BRUZZONE
GENEVIÈVE DAUPHIN-TANGUY
SERGIO JUNCO
MIQUEL ANGEL PIERA

THE 5TH INTERNATIONAL CONFERENCE ON INTEGRATED MODELING AND ANALYSIS IN APPLIED CONTROL AND AUTOMATION

SEPTEMBER 12-14 2011
ROME, ITALY

ORGANIZED BY

DIPTEM – UNIVERSITY OF GENOA

LIOPHANT SIMULATION

SIMULATION TEAM

IMCS – INTERNATIONAL MEDITERRANEAN & LATIN AMERICAN COUNCIL OF SIMULATION

MECHANICAL DEPARTMENT, UNIVERSITY OF CALABRIA

MSC-LES, MODELING & SIMULATION CENTER, LABORATORY OF ENTERPRISE SOLUTIONS

MODELING AND SIMULATION CENTER OF EXCELLENCE (MSCOE)

MISS LATVIAN CENTER – RIGA TECHNICAL UNIVERSITY

LOGISIM

LSIS – LABORATOIRE DES SCIENCES DE L'INFORMATION ET DES SYSTEMES

MISS – UNIVERSITY OF PERUGIA

MISS – BRASILIAN CENTER, LAMCE-COPPE-UFRJ

MISS - MCLEOD INSTITUTE OF SIMULATION SCIENCES

M&SNET - MCLEOD MODELING AND SIMULATION NETWORK

LSS   LATVIAN SIMULATION SOCIETY

ESISA   ECOLE SUPERIEURE D'INGENIERIE EN SCIENCES APPLIQUEES

FCEIA   FACULTAD DE CIENCIAS EXACTAS. INGEGNERIA Y AGRIMENSURA

ULL   UNIVERSITY OF LA LAGUNA

CIFASIS   CIFASIS: CONICET-UNR-UPCAM

INSTICC   INSTICC - INSTITUTE FOR SYSTEMS AND TECHNOLOGIES OF INFORMATION, CONTROL AND COMMUNICATION

I3M 2011 INDUSTRIAL SPONSORS

PRESAGIS   PRESAGIS

CAE   CAE

CALTEK   CAL-TEK

MAST   MAST

AEgis TECHNOLOGIES   AEGIS TECHNOLOGIES

I3M 2011 MEDIA PARTNERS

MS&T   MILITARY SIMULATION & TRAINING MAGAZINE

Euromerci   EURO MERCI

EDITORS

AGOSTINO BRUZZONE
MISS-DIPTEM, UNIVERSITY OF GENOA, ITALY
agostino@itim.unige.it

GENEVIÈVE DAUPHIN-TANGUY
ECOLE CENTRALE DE LILLE, FRANCE
genevieve.dauphin-tanguy@ec-lille.fr

SERGIO JUNCO
UNIVERSIDAD NACIONAL DE ROSARIO, ARGENTINA
sjunco@fceia.unr.edu.ar

MIQUEL ANGEL PIERA
AUTONOMOUS UNIVERSITY OF BARCELONA, SPAIN
miquelangel.piera@uab.cat

THE INTERNATIONAL MEDITERRANEAN AND LATIN AMERICAN MODELING MULTICONFERENCE, I3M 2011

GENERAL CO-CHAIRS
AGOSTINO BRUZZONE, *MISS DIPTEM, UNIVERSITY OF GENOA, ITALY*
MIQUEL ANGEL PIERA, *AUTONOMOUS UNIVERSITY OF BARCELONA, SPAIN*

PROGRAM CHAIR
FRANCESCO LONGO, *MSC-LES, UNIVERSITY OF CALABRIA, ITALY*

THE 5TH INTERNATIONAL CONFERENCE ON INTEGRATED MODELING AND ANALYSIS IN APPLIED CONTROL AND AUTOMATION, IMAACA 2011

CONFERENCE CHAIRS
SERGIO JUNCO, *UNIVERSIDAD NACIONAL DE ROSARIO, ARGENTINA*
GENEVIÈVE DAUPHIN-TANGUY, *ECOLE CENTRALE DE LILLE, FRANCE*

IMAACA 2011 International Program Committee

Tracks and Workshop Chairs

CONFERENCE CHAIRS' MESSAGE

WELCOME TO IMAACA 2011!

This year the wonderful framework of the IMAACA conference is the Eternal City: we are honored to welcome all the IMAACA authors as well as all the I3M 2011 participants to Rome.

Based on past experiences (Fes, Morocco; Buenos Aires, Argentina; Marseille, France; and Genoa Italy) we are even more positive about the success of IMAACA 2011. The very high scientific quality of selected papers and the invaluable work carried out (over roughly one year) by the International Program Committee will bring experts together for the purpose of presenting and discussing all type of application relevant control and automation problems emphasizing the role of model analysis as integral part of the complete design of the automated system. In fact an accurate review of the selected papers reveals that most recent and innovative concepts, methods, techniques, and tools, conceived in order to support an integrative interplay of modeling, identification, simulation, system analysis and control theory in all the stages of system design, are considered at IMAACA 2011.

In addition IMAACA will be co-located with the I3M 2011 multi-conference that includes EMSS, MAS, HMS and DHSS therefore providing the authors with the opportunity to attend multiple sessions and have a direct look to the most advanced and innovative international research in the Modeling & Simulation area.

Finally, we would like to wish to all the attendees a fruitful and enjoyable IMAACA 2011 in Rome.

Sergio Junco,
Universidad Nacional
de Rosario, Argentina

Geneviève Dauphin-Tanguy,
Ecole Centrale de Lille, France

ACKNOWLEDGEMENTS

LOCAL ORGANIZATION COMMITTEE

AGOSTINO G. BRUZZONE, *MISS-DIPTEM, UNIVERSITY OF GENOA, ITALY*
ENRICO BOCCA, *SIMULATION TEAM, ITALY*
FRANCESCO LONGO, *MSC-LES, UNIVERSITY OF CALABRIA, ITALY*
FRANCESCA MADEO, *UNIVERSITY OF GENOA, ITALY*
MARINA MASSEI, *LIOPHANT SIMULATION, ITALY*
LETIZIA NICOLETTI, *CAL-TEK, ITALY*
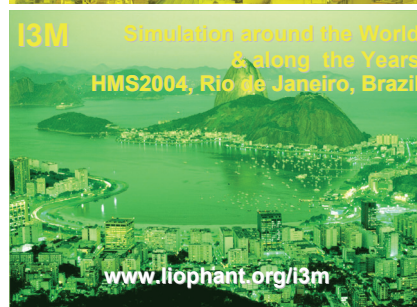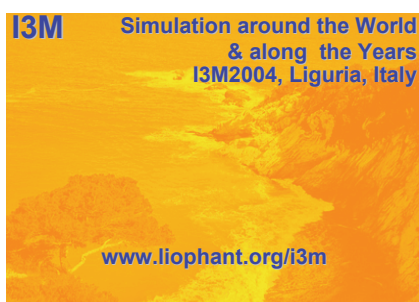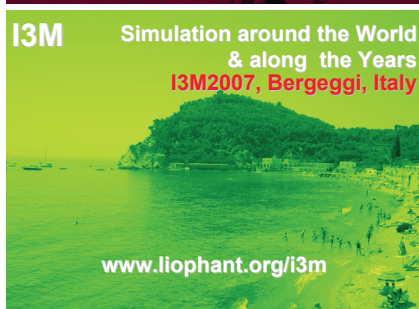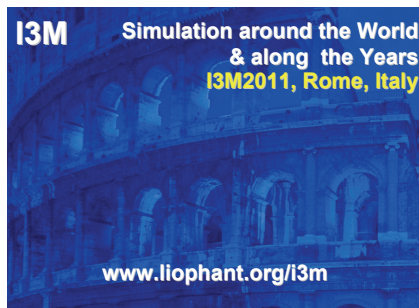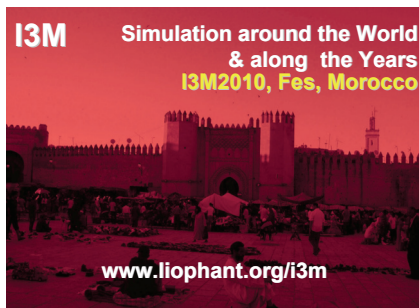FEDERICO TARONE, *SIMULATION TEAM, ITALY*
ALBERTO TREMORI, *SIMULATION TEAM, ITALY*

This International Conference is part of the I3M Multiconference: the Congress leading *Simulation around the World and Along the Years*

Index

ANALYSIS OF CONTROLLED SEMI-MARKOV QUEUEING MODELS

Kashtanov V.A, Kondrashova E.V.

Department of Operation research
Moscow State Institute of Electronics and Mathematics
Bolshoi Trekhsvyatitelskij per., 3/12, 109028, Moscow, Russia

Elizavetakondr@gmail.com

ABSTRACT
The present paper is devoted to the research of the controlled semi-markov queueing model M/M/n*/N*. Research of income functional is carried out at control of several parameters. The control is based on the system structure. In the article the income functional is constructed on the trajectories of the controlled semi-markov process. The main problem is to search an optimal control strategy for the given model. The given research algorithm also can be used for various models at control with both discrete and continuous model parameters. Algorithmization of the problem is presented.

Keywords: optimization, semi-markov controlled process, control strategy, income functional

1. INTRODUCTION
The main part of research works of queueing theory showed that the model structure is considered invariable, and the model characteristics are investigated at the fixed initial parameters and functions. The interarrival time distribution, the service time distribution, the number of servers, the system capacity are considered to be fixed. In the present paper we change the model structure and investigate the system model using variables of controls.

The purpose of the queueing model theory is to present recommendations on maintenance for high efficiency system functioning.

Notice that the functioning increase can be reached using control parameters. There can be several control parameters in the model. There may be the interarrival time distribution, the service time distribution, the number of servers, the system capacity and others. We are interested in several control parameters, that means the controls set expansion.

One of the main results of the controlled semi-markov processes theory is strategy determination that gives the maximum value for the functional. The main theory of semi-markov processes can be found in the following works (Bellman R.A. 1957, Jewell W. S., 1967).

It is important to note that we can obtain new results using investigation for not only stationary characteristics (stationary queue length, loss probability, etc.), but using structure model changing.

Stochastic character of the interarrival time and the service time generates a stochastic process in the queueing model. For the controlled stochastic process the problem of a choice of optimum control strategy is given. In the present paper a controlled semi-markov process is used. On the trajectories of the controlled semi-markov process the income functional is constructed.

The result of the functional research is a choice of optimal control strategies for the given model.
Similar works linked with the control in queueing systems (Banik A.D. 2009, Swishchuk A. 1999, Kashtanov V.A. 2010), one of the advantages of the given work is simultaneous control using several parameters of system.

To calculate parameters of the system functioning quality and to choose optimum control, the following algorithm is used:

* Construct semi-markov process,
* Determine structure of control,
* Enter constants describing incomes and costs,
* Calculate additional characteristics,
* Calculate income functional and search for optimum control.

2. PROBLEM STATEMENT
The queueing model can be represented using Kendall's notation M/M/n*/N*.

For the given model:

* the interarrival time distribution is a Poisson distribution with parameter λ,
* the service time distribution is an exponential distribution with parameter μ (service intensity),
* the number of servers channels is a parameter of control, n* is a maximum possible value,
* the maximum number of places in the queue (system capacity) is also a parameter of control. N* is a maximum possible value.

The present paper is devoted to the research of the controlled semi-markov queueing model using model structure control.

Describe the system in details and enter some assumptions:

- Maximum possible number of servers is n, each of which can serve simultaneously only one requirement. The number of servers is a control parameter and can change from 0 up to n.
- Maximum possible number of places in the queue is N. It is a parameter of control and can change from 0 up to N.
- Markov moments are the moments of the service termination and the arrival epoch. The control strategy is to be chosen at the markov moments only.
- Each state of the system can be described using pair of parameters (i, j), where
- i - the numbers which are served, $0 \le i \le n$.
- j - the numbers which are in the queue, $0 \le j \le N$.
- It is not rational to increase the number of empty/free places in the queue more, than by one at every markov moment. Therefore in the store never will be more than one empty seat in the queue.
- Instead of addition of an extra seat being in state (0,0) it is possible to add extra servers (if it is more profitable, expense for the free working server is less than an expenses for an additional places in the queue.
- We assume, that the queueing model is directed on service of the greatest possible number of requirements, hence, we assume, that it is more profitable to send the requirement to server (if such opportunity is represented). Thus, when from the store the requirement sends on service, the number of places in the queue decreases for one unit.
- If one of the servers has finished service and there are requirements in the queue model passes from set (i, j) to set (i, j-1).
- When one of channels has finished service and in the store there are no requirements, the system passes from set (i, 0) in set (i-1,0).
- Note, that the control is carried out at markov moments. The transition probability depends on the made decision (strategy).

3. RESEARCH ALGORITHM

To calculate parameters of the system functioning quality and to choose optimum control, the algorithm investigated by Kashtanov is used.

For construction of the controlled semi-markov process it is necessary to define the parameters: markov moments, the states set, the set of controls, a semi-markov kernel. The next step is to determine:

- transition probabilities of the embedded markov chain,
- conditional mathematical expectations of the saved up income provided that process is in state (i,j) and through time t will pass in state (i',j'),
- mathematical expectations of the saved up income for the full period when process is in state (i,j).
- Then, a parameter value of the functioning quality (the income functional) is calculated.

Parameter of functioning quality is the income functional. The income functional is defined as follows:

$$S = \lim_{t \to \infty} \frac{S_i(t)}{t} = \frac{\sum_{i \in E} s_i \pi_i}{\sum_{i \in E} m_i \pi_i} \tag{1}$$

where

s_i - mathematical expectation of the saved up income for the full period when process is in state i,

m_i - mathematical expectation of the saved up income for the full period when process is in state i,

π_i - stationary probabilities of the embedded markov chain.

4. CONSTRUCTION OF THE CONTROLLED SEMI-MARKOV PROCESS

Semi-markov process $\xi(t)$ is described with a 2-dimensional markov chain (ξ_n, θ_n), $n \ge 0, \xi_n \in E, \theta_n \in R^+ = [0, \infty)$.

The markov chain (ξ_n, θ_n) is defined with transition probabilities (semi-markov kernel):

$$P\{\xi_{n+1} = j, \theta_{n+1} < t | \xi_n = i, \theta_n = \tau\} =$$
$$P\{\xi_{n+1} = j, \theta_{n+1} < t | \xi_n = i\} = Q_{ij}(t) \tag{2}$$

where $n \ge 1, i, j \in E, t \in [0, \infty)$ and using initial distribution

$$p_i = P\{\xi_0 = i\} \ge 0, i \in E, \sum_{i \in E} p_i = 1,$$

assume $P\{\theta_0 = 0\} = 1$.

A 3-dimensional markov chain (controlled semi-markov process) is described as follows:

(ξ_n, θ_n, u_n), $n \ge 0, \xi_n \in E, \theta_n \in R^+ = [0, \infty), u_n \in U$.

Where $E = \{1, 2, ..., N\}$ - the final states' set, θ_n - a component is identified as time duration. On the set $R^+ = [0, \infty)$ σ-algebra is stated. U -is a control set with σ-algebra A of a subset of this set.

Markov chain (ξ_n, θ_n, u_n) is defined using transition probabilities:

$$P\{\xi_{n+1} = j, \theta_{n+1} < t, u_{n+1} \in B | \xi_n = i, \theta_n = \tau, u_n = u\},$$

$$i, j \in E, t, \tau \in R^+, B \in A, u \in U$$

Assume that:

$$P\{\xi_{n+1} = j, \theta_{n+1} < t, u_{n+1} \in B | \xi_n = i, \theta_n = \tau, u_n = u\} =$$
$$= P\{\xi_{n+1} = j, \theta_{n+1} < t, u_{n+1} \in B | \xi_n = i\} \qquad (3)$$

and

$$p_i = P\{\xi_0 = i\} \geq 0, i \in E, \sum_{i \in E} p_i = 1, P\{\theta_0 = 0\} = 1.$$

The probability
$$Q_{ij}(m,t,B) = \{\xi_{n+1} = j, \theta_{n+1} < t, u_{n+1} \in B | \xi_n = i\}$$ generates on the set (U_i, A_i) probabilities measures:

$$G_i(B) = P\{u_{n+1} \in B / \xi_n = i\} = \sum_{i \in E} Q_{ij}(m,B), i \in E, B \in A_i.$$

Then $Q_{ij}(m,t,B) \leq G_i(B)$, so:

$$Q_{ij}(m,t,B) = \int_B Q_{ij}(m,t,u)G_i(du)$$

where $Q_{ij}(m,t,u) = P\{\xi_{n+1} = j, \theta_{n+1} < t | \xi_n = i, u_{n+1} = u\}$

$$Q_{ij}(m,t,B) =$$
$$\int_B P\{\xi_{n+1} = j, \theta_{n+1} < t | \xi_n = i, u_{n+1} = u\}G_i(du) \qquad (4)$$

Each system state can be described by pair of parameters (i, j), where

i - the number of requirements which are on service $0 \leq i \leq n$.

j - the number of requirements which are on places in the queue $0 \leq j \leq N$.

Define the control set. We can control choosing the number of servers and the number of additional/extra places in the queue.

The control set of the servers quantity is $U = \{0,1,.....,u\}$, where:

0 - we do not add servers,

1 - we make decision to add 1 extra server.

…..

u - we add u additional servers, where $1 \leq u \leq n-i$.

The control set of the extra places is $V = \{0,1\}$, where:

$v = 0$ – we do not add places,

$v = 1$ – we add one place for expectation in the queue.

Note, that at the markov moments in each state the desicion {u,v} can be made, that depends on the concrete system state.

In state (0,0) there are no requirements either on service, or in the queue in the system.

In the given state it is possible to make a decision on addition of an extra place (or inclusion of the server).

Hence, in state (0,0) following controls U = {0,1}, V = {0,1} are possible. It is meaningful to add only one seat in the queue or to include one server/device.

In state (i,0), $0 < i \leq n$ the following controls are possible U={0,1},V={0,1}.

In state (0,j), $0 < j \leq N$, there are no requirements on service, there are j requirements in the queue. So in state (0,j) the controls U={0,u},V={0,1} are possible. And $0 \leq u \leq \min(j,n)$.

In state (i,j), $0 < i < n, 0 < j < N$, there are i requirements on service, there are j requirements in the queue. In state (i,j) the controls U={0,u},V={0,1} are possible. And $0 \leq u \leq \min(j, n-i)$.

In state (i,N), $i < n$, are i requirements on service, there are N requirements in the queue (maximum possible). So in (i,N) the controls U={0,u},V={0} are possible. And $0 \leq u \leq \min(N, n-i)$.

In state (n,j), $0 < j < N$, there are n requirements on service (maximum possible), there are j requirements in the queue. The controls U={0},V={0,1} are possible.

In state (n,N) there are n requirements on service (maximum possible), there are N requirements in the queue (maximum possible). So in state (n,N) the controls U={0},V={0} are possible.

Construct semi-markov kernel for the given system.

The interarrival time distribution is a Poisson distribution with parameter λ, $F(t) = 1 - e^{-\lambda t}$.

The service time distribution is an exponential distribution with parameter μ (service intensity), $G(t) = 1 - e^{-\mu t}$.

Semi-markov kernel $Q_{(i,j)(i',j')}(t,u,v)$ is a probability of that the semi-markov process will pass in state (i',j') and the time before this transition will not surpass t provided that process is in state (i,j) and in this state a decision from the set of controls (u,v) is made.

In state (0,0) $U_{(0,0)} = \{0,1\}$, $V_{(0,0)} = \{0,1\}$.

$$Q_{(0,0)(0,1)}(t,0,1) = Q_{(0,0)(1,0)}(t,1,0) = Q_{(0,0)(0,0)}(t,0,0) =$$
$$F(t) = 1 - e^{-\lambda t}$$

In state (i,0), $0 < i \leq n$ $U_{(i,0)} = \{0\}$, $V_{(i,0)} = \{0,1\}$.

$$Q_{(i,0)(i-1,0)}(t,0,1) = Q_{(i,0)(i-1,0)}(t,0,0) = \frac{i\mu}{\lambda + i\mu}(1 - e^{-t(\lambda + i\mu)})$$

3

$$Q_{(i,0)(i,1)}(t,0,1) = Q_{(i,0)(i,0)}(t,0,0) = \frac{\lambda}{\lambda + i\mu}(1 - e^{-t(\lambda + i\mu)})$$
.

In state (0,j) $U_{(i,0)} = \overline{\{0, \min(j,n)\}}$, $V_{(i,0)} = \{0,1\}$

If $u = j, j \leq n$ then:

$$Q_{(0,j)(j,1)}(t,j,1) = Q_{(0,j)(j,0)}(t,j,0)$$
$$= \frac{\lambda}{\lambda + j\mu}(1 - e^{-t(\lambda + j\mu)})$$
$$Q_{(0,j)(j-1,0)}(t,j,0) = Q_{(0,j)(j-1,0)}(t,j,1)$$
$$= \frac{j\mu}{\lambda + j\mu}(1 - e^{-t(\lambda + j\mu)})$$

Else:

$$Q_{(0,j)(u,j-u+1)}(t,u,1) = Q_{(0,j)(u,j-u)}(t,u,0) =$$
$$\frac{\lambda}{\lambda + u\mu}(1 - e^{-t(\lambda + u\mu)})$$
$$Q_{(0,j)(u,j-u-1)}(t,u,0) = Q_{(0,j)(u,j-u-1)}(t,u,1) =$$
$$\frac{u\mu}{\lambda + u\mu}(1 - e^{-t(\lambda + u\mu)})$$

In state (i,j), $0 < i < n, 0 < j < N$, $U_{(i,j)} = \overline{\{0, \min(j, n-i)\}}$, $V_{(i,0)} = \{0,1\}$, $u \in U_{(i,j)}$.

If $u = j, j \leq n - i$ then:

$$Q_{(i,j)(i,1)}(t,j,1) = Q_{(i,j)(i+j,0)}(t,j,0)$$
$$= \frac{\lambda}{\lambda + (i+j)\mu}(1 - e^{-t(\lambda + (i+j)\mu)})$$
$$Q_{(i,j)(i+j-1,0)}(t,j,0) = Q_{(i,j)(i+j-1,0)}(t,j,1)$$
$$= \frac{(i+j)\mu}{\lambda + (i+j)\mu}(1 - e^{-t(\lambda + (i+j)\mu)})$$

Else:

$$Q_{(i,j)(i,j-u+1)}(t,u,1) = Q_{(i,j)(i+u,j-u)}(t,u,0)$$
$$= \frac{\lambda}{\lambda + (i+u)\mu}(1 - e^{-t(\lambda + (u+a)\mu)})$$
$$Q_{(i,j)(i+u,j-u-1)}(t,u,0) = Q_{(i,j)(i+u,j-u-1)}(t,u,1)$$
$$= \frac{(i+u)\mu}{\lambda + (i+u)\mu}(1 - e^{-t(\lambda + (i+u)\mu)})$$

In state (i,N), $i < n$, $U_{(i,N)} = \overline{\{0, \min(N, n-i)\}}$, $V_{(i,N)} = \{0,1\}$ $u \in U_{(i,N)}$.

And if $V_{(i,N)} = \{1\}$, then $U_{(i,N)} = \overline{\{1, \min(N, n-i)\}}$.

If $u = N, N \leq n - i$ then:

$$Q_{(i,N)(i,1)}(t,N,1) = Q_{(i,N)(i+N,0)}(t,N,0)$$
$$= \frac{\lambda}{\lambda + (i+N)\mu}(1 - e^{-t(\lambda + (i+N)\mu)})$$
$$Q_{(i,N)(i+N-1,0)}(t,N,1) = Q_{(i,N)(i+N-1,0)}(t,N,0)$$
$$= \frac{(i+N)\mu}{\lambda + (i+N)\mu}(1 - e^{-t(\lambda + (i+N)\mu)})$$

Else:

$$Q_{(i,N)(i,N-u+1)}(t,u,1) = Q_{(i,N)(i+u,N-u)}(t,u,0)$$
$$= \frac{\lambda}{\lambda + (i+u)\mu}(1 - e^{-t(\lambda + (u+a)\mu)})$$
$$Q_{(i,N)(i+u,N-u-1)}(t,u,0) = Q_{(i,j)(i+u,N-u-1)}(t,u,1)$$
$$= \frac{(i+u)\mu}{\lambda + (i+u)\mu}(1 - e^{-t(\lambda + (i+u)\mu)})$$

In state (n,j), $0 < j < N$, $U_{(n,j)} = \{0\}$, $V_{(n,j)} = \{0,1\}$.

$$Q_{(n,j)(n,j)}(t,0,0) = Q_{(n,j)(n,j+1)}(t,0,1) =$$
$$\frac{\lambda}{\lambda + n\mu}(1 - e^{-t(\lambda + n\mu)})$$
$$Q_{(n,j)(n,j-1)}(t,0,0) = Q_{(n,j)(n,j-1)}(t,0,1) =$$
$$\frac{n\mu}{\lambda + n\mu}(1 - e^{-t(\lambda + n\mu)})$$

In state (n,N), $U_{(n,N)} = \{0\}$, $V_{(n,j)} = \{0\}$.

$$Q_{(n,N)(n,N)}(t,0,0) = \frac{\lambda}{\lambda + n\mu}(1 - e^{-t(\lambda + n\mu)})$$
$$Q_{(n,N)(n,N-1)}(t,0,0) = \frac{n\mu}{\lambda + n\mu}(1 - e^{-t(\lambda + n\mu)})$$

Notice, that cases of states (0, j), (n, j), (i, N), (n, N) are described by formulas for a case (i, j), considering the controls sets for each concrete state.

5. ADDITIONAL CALCULATIONS
For construction of a conditional mathematical expectation of the saved up income $R_{(i,j)(i',j')}(t,u,v)$ provided that process is in state (i,j) and through time t will pass in state (i',j') and in this state a decision from the set of controls (u,v) is made, we shall enter the constants describing incomes and costs:

c_1 - income received per service of one number;

$-c_2$ - payment per time unit of server work device during the service time of the customer ;

$-c_3$ - payment per time unit in case the number is in the queue;

$-c_4$ - payment per time unit in case extra place in the queue is not used;

$-c_5$ - payment per one lost number;

$-c_6$ - payment per time unit in case server is working, but not used (without service).

Then for all possible states $R_{(i,j)(i',j')}(t,u,v)$ can be described.

Conditional mathematical expectation of the saved up income $R_{(i,j)(i',j')}(t,u,v)$ provided that process is in state $(0,0)$ can be described:

$$R_{(0,0)(0,1)}(t,0,1) = R_{(0,0)(0,0)}(t,0,1) = -c_4 t$$
$$R_{(0,0)(1,0)}(t,1,0) = R_{(0,0)(0,0)}(t,1,0) = -c_6 t$$
$$R_{(0,0)(0,0)}(t,0,0) = -c_5$$

In state (i,j):

$$R_{(i,j)(i+u,j-u)}(t,u,0) = -(c_2(i+u)+c_3(j-u))t - c_5$$
$$R_{(i,j)(i+u,j-u+1)}(t,u,1) = -(c_2(i+u)+c_3(j-u)+c_4)t$$
$$R_{(i,j)(i+u,j-u-1)}(t,u,0) = -(c_2(i+u)+c_3(j-u))t + c_1$$
$$R_{(i,j)(i+u,j-u-1)}(t,u,1) = -(c_2(i+u)+c_3(j-u)+c_4)t + c_1$$

If $u = j, j \leq n - i$ then:

$$R_{(i,j)(i+j,0)}(t,j,0) = -c_2(i+j)t - c_5$$
$$R_{(i,j)(i+j,1)}(t,j,1) = -(c_2(i+j)+c_4)t$$
$$R_{(i,j)(i+j-1,0)}(t,j,0) = -c_2(i+j)t + c_1$$
$$R_{(i,j)(i+j-1,0)}(t,j,1) = -(c_2(i+j)+c_4)t + c_1$$

We can define conditional mathematical expectation of the saved up income for all possible system states (see semi-markov kernel).

Calculate $m_{(i,j)}$ mathematical expectations of the saved up income for the full period when process is in state (i,j).

$$m_{(i,j)} = \int_0^\infty \left[1 - \sum_{(i',j')\in E} Q_{(i,j)(i',j')}(t) \right] dt \tag{5}$$

$$Q_{(i,j)(i',j')}(t) = \sum_{u,v} Q_{(i,j)(i',j')}(t,\{u,v\}) p_{u,v}^{[i,j]} \tag{6}$$

$$m_{(i,j)} = \int_0^\infty \left[1 - \sum_{(i',j')\in E} \sum_{u,v} Q_{(i,j)(i',j')}(t,\{u,v\}) p_{u,v}^{[i,j]} \right] dt \tag{7}$$

Where $p_{u,v}^{[i,j]}$ is a probability that in state (i,j) we choose strategy (u,v) from the given control sets.

For example, in state (0,0), U={0,1}, V={0,1}.

$$m_{(0,0)} = \int_0^\infty \left[1 - \sum_{(i',j')\in E} \sum_{u,v} Q_{(0,0)(i',j')}(t,\{u,v\}) p_{u,v}^{[0,0]} \right] dt$$

$$m_{(0,0)}(0,0) = m_{(0,0)}(0,1) = m_{(0,0)}(1,0) = \frac{1}{\lambda}$$

In state (i,j), $0 < i < n, 0 < j < N$, U={0,u}, V={0,1}, $0 \leq u \leq \min(j, n-i)$.

If $u = j, j \leq n - i$ then:

$$m_{(i,j)}(j,0) = m_{(i,j)}(j,1) = \frac{1}{\lambda + (i+j)\mu}$$

Else:

$$m_{(i,j)}(u,0) = m_{(i,j)}(u,1) = \frac{1}{\lambda + (i+u)\mu}$$

Notice, that cases of states (0, j), (n, j), (i, N), (n, N) are described by formulas for a case (i, j), considering the controls sets for each concrete state.

Provide additional calculations. Calculate variables $s_{(i,j)}$. Variables $s_{(i,j)}$ are mathematical expectations of the saved up income for the full period when process is in state (i,j).

$$s_{(i,j)}(t,\{u,v\}) = \sum_{(i,j)\in E} \int_0^t R_{(i,j)(i',j')}(x,\{u,v\}) dQ_{(i,j)(i',j')}(x,\{u,v\}) +$$
$$+ \sum_{(i,j)\in E} \int_t^\infty R_{(i,j)(i'j')}(x,t,\{u,v\}) dQ_{(i,j)(i',j')}(x,\{u,v\}) \tag{8}$$

$$s_{(i,j)}(t) = \sum_{u,v} s_{(i,j)}(t,\{u,v\}) p_{u,v}^{[i,j]} \tag{9}$$

Consider $s_{(i,j)} = \lim_{t \to \infty} s_{(i,j)}(t)$.

$$s_{(i,j)} = \sum_{u,v} \sum_{(i',j')\in E} \int_0^\infty R_{(i,j)(i'j')}(x,\{u,v\}) dQ_{(i,j)(i',j')}(x,\{u,v\}) p_{u,v}^{[i,j]} \tag{10}$$

Taking into consideration degenerated distributions (see next page) we have:

$$s_{(i,j)} = \sum_{(i',j')\in E} \int_0^\infty R_{(i,j)(i',j')}(x,\{u^{[i,j]},v^{[i,j]}\}) dQ_{(i,j)(i',j')}(x,\{u^{[i,j]},v^{[i,j]}\}) \tag{11}$$

For example, in state (i,N):

$$s_{(i,N)}(u,1) = \frac{-(c_2(i+u)+c_3(j-u)+c_4)+c_1(i+u)\mu}{\lambda+(i+u)\mu}$$

$$s_{(i,N)}(u,0) = \frac{-(c_2(i+u)+c_3(N-u))-c_5\lambda+c_1(i+u)\mu}{\lambda+(i+u)\mu}$$

Notice, that cases of states (0, j), (n, j), (i, N), (n, N) are described by formulas for a case (i, j), considering the controls sets for each concrete state.

Stationary probabilities of the embedded Markov chain are defined from the equations:

$$\begin{cases} \pi_l = \sum_{k \in E} \pi_k p_{kl} \\ \sum_{k \in E} \pi_k = 1 \end{cases} \tag{12}$$

where $p_{ij} = \lim_{t \longrightarrow \infty} Q_{ij}(t, \{u, v\})$

Taking into consideration that model state is described using pair of parameters, the system can be defined:

$$\pi_{(i',j')} = \sum_{(i,j) \in E} \pi_{(i,j)} p_{(i,j)(i',j')}, \quad \sum_{(i',j') \in E} \pi_{(i',j')} = 1 \tag{13}$$

where

$$p_{(i,j)(i',j')} = \lim_{t \longrightarrow \infty} Q_{(i,j)(i',j')}(t, \{u, v\}).$$

Use theorem for controlled semi-markov processes for the income functional calculation (Kashtanov and Medvedev 2002) and find S (1).

Note, that calculations are to be automated because of complexity.

Using the theorem about the accumulation functional fractional-linearity concerning the distributions defining structure of the accumulation (see Appendix), and the theorem about a maximum of fractional-linear functional, we use the basic conclusion: it is possible to search for optimum strategy in the set of determined strategies of control (Kashtanov and Medvedev 2002). So all results of the calculations are obtained substituting degenerated distributions.

We receive the functional, depending on the variables which are responsible for control $S((v_{00}, .. v_{ij} .., v_{nN}), u_{00}, .. u_{ij} ..., u_{nN})$.

Further it is necessary to define a maximum of the functional S depending on u_{ij}, v_{ij} (see set of controls). Then we receive the solve of the problem. That is the strategy $((v*_{00}, ..., v*_{nN}), u*_{00}, ..., u*_{nN})$, at which the maximal income is received. Hence, when the process is in state (i, j) we take the decision $\{u*_{ij}, v*_{ij}\} \in U$. The choice of the decision defines the most effective system work.

6. CALCULATION EXAMPLE

The calculation results have shown the opportunities of control expansion, in other words using several parameters of control leads to more effective model functioning. For example for model M/G*/1/2* we can use discrete and continuos controls (Kashtanov and Kondrashova 2010).

There are three states in the system.

State {0} – there are no numbers in the system;

State {1} – one number is being served, there are no numbers in the queue;

State {2} – there are one number is on service, there is one number in the queue.

In state 1 we can add 0,1,2 additional places in the queue. In state 2 we can add 0 or 1 additional places in the queue. We receive, that the quantity of every possible strategy combinations for choice v equals six. u –service duration in state, $u \in [0, \infty)$ can be chosen in each state (control parameter).

In state i we choose the service time and we make the decision on creation of additional/extra places in the queue (except state i=0). Hence, we use two parameters of control from the given set of controls: {u, v}.

Write down the strategy pairs for states 1 and 2 : (0,0); (0,1); (1,0); (1,1); (2,0); (2,1). Last strategy pair concerns to a classical case of the problem when all places are used and we control only with the service time. And the control of service duration is also used (τ_1, τ_2 – service duration in 1 and 2 state, that take maximum value for income functional at fixed v).

Table 1: Calculation example

Strategy	Maximum value S	Value τ_1	Value τ_2
(0,0)	8	0	-
(0,1)	8	0	-
(1,0)	2	0	-
(1,1)	2	0	-
(2,0)	0,24418416	1,05	0
(2,1)	1,16934403	1,367	0,674
(0,0)	1	0	-
(0,1)	1	0	-
(1,0)	-0,743454243	0,879	-
(1,1)	-0,743454243	0,879	-
(2,0)	1,54211326	1,387	0
(2,1)	2,28817567	1,482	0,448

7. CONCLUSION

In the paper the research of the controlled semi-markov queueing model M/M/n*/N* is carried out. The control is based on the system structure. In the paper the construction of the controlled semi-markov process and the construction of the income functional on its trajectories are used.

The formula for the income functional calculation for the given system is received. The calculation example for the model is demonstrated. Algorithmization of the problem is presented.

The calculation results have shown the opportunities of control expansion, in other words using several parameters of control leads to more effective system functioning.

APPENDIX A
Theorem. If there is a controlled semi-markov process, for which the embedded Markov chain is ergodic, at

least, one of the distributions $Q_{ij}(t)$ is non-trellised and the average income S_i is final at any $i \in E$, then:

$$S = \lim_{t \to \infty} \frac{S_i(t)}{t} = \frac{\sum_{i \in E} s_i \pi_i}{\sum_{i \in E} m_i \pi_i}, \qquad (14)$$

where π_i are stationary probabilities of conditions' distribution of the embedded Markov chain.

APPENDIX B
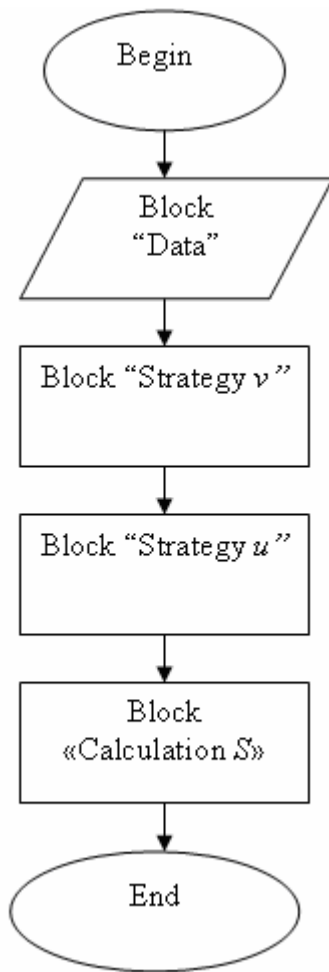
Figure 1: Block-diagram of the automated decision for M/M/n*/N*

Block "Data": input n, N, λ, μ, cost constants c_1, c_2, c_3, c_4, c_5, c_6.

Block "Strategy v": generating of strategy massive for v for every possible combinations.

Block "Strategy u": generating of strategy massive for u for every possible combinations.

Block "Calculation S": additional calculations for necessary characteristics for each strategy combination (cycle). Choice of optimum functional value and optimum strategy.

REFERENCES

Banik A.D., 2009. Queueing analysis and optimal control of BMAP/G$^{(a,b)/}$1/N and BMAP/MSP$^{(a,b)/}$1/N systems. *Computers & Industrial Engineering 57,748–761*

Bellman R.A., 1957. Markovian Decision Process. *J. Math. And Mech.*, 6,p. 679.

Jewell W. S., 1967. Markov-renewal programming. *Operation Res* 11, 938-971.

Kashtanov V.A., 2010. Controlled semi-markov processes in modeling of the reliability and redundancy maintenance of queueing systems. *Computer Modelling and New Technologies*, Vol.14, No.1, 26–30

Kashtanov V. A., Medvedev A.I., 2002. The reliability theory of complex systems (theory and practise). *Moscow: European quality center.*

Kashtanov V.A., Kondrashova E.V., 2010. Controlled semi-markov queueing model. *Proceedings of Third International Conference on Accelerated Life testing, Reliability-based Analysis and Design*, pp. 243-248. May 19-21, Clermont-Ferrand, France.

Swishchuk A., 1999. Stochastic stability and optimal control of semi-markov risk processes in insurance mathematics. *Intern. Mathem. Center, Institute of Mathem.*, National Acad. of Scien. of Ukraine, Kiev, Ukraine.

MODELING OF FLEXURAL VIBRATION OF ROTATING SHAFT WITH IN-SPAN DISCRETE EXTERNAL DAMPING

Vikas Rastogi [a], Amalendu Mukherjee [b], Anirban Dasgupta[c]

[a] Mechanical Engineering Department
Sant Longowal Institute of Engineering and Technology
Longowal-148106, Punjab, India
[b] [c] Mechanical Engineering Department
Indian Institute of Technology, Kharagpur, 721302, India

[a]rastogivikas@yahoo.com, [b]amalendu@mech.iitkgp.ernet.in , [c]anir@mech.iitkgp.ernet.in

ABSTRACT
This paper deals with the modeling of one-dimensional continuous systems with in-span discrete external damping. The mathematical formulation for this system is similar to an internally and externally damped rotor driven through a dissipative coupling; however the umbra-Lagrangian density contributed by external damping is different. Using such formulation, the invariance of the umbra-Lagrangian density is obtained through an extension of Noether's theorem. The rotor shaft is modeled as a Rayleigh beam. The dynamic analysis of the rotor shaft is obtained and validated through simulation studies. Results show an interesting phenomenon of limiting behaviour of the rotor shaft with internal damping beyond certain threshold speeds, which are obtained theoretically and affirmed by simulation. The results show that regenerative energy in the rotor shaft due to internal damping is dissipated through the discrete damper as well as the dissipative coupling between drive and the rotor shaft. In such case, the excitation frequency is more, the shaft speed will not increase but the slip between drive and shaft will increase due to loading of drive.

Keywords: Umbra-Lagrangian density, Noether's theorem, Flexural vibration, In-span external damping

1. INTRODUCTION
The methods of Lagrange and Hamilton based on the variational principle on fields are employed to describe continuous systems. There are few direct methods, which approximate the continuous system by considering finite or discrete particles and then examining the changes in the equations describing the motion as the continuous limit is approached. The general relationship between one-parameter continuous symmetries and conserved quantities in field theory has been discussed by Boyer (1967). Cantrijn and Sarlet (1981) had introduced a direct method for associating conserved quantities with each dynamical symmetry group of a Lagrangian system. Some other useful results related with the symmetries aspects of higher Lagrangian and Hamiltonian formulism are discussed in

papers of Katzin and Levine (1976) and Damianou and Sophocleous (2004). To extend the scope of Lagrangian-Hamiltonian mechanics, a new proposal of umbra-time was made by Mukherjee (1994). A brief and promising commentary of this kind of extension has been given by Brown (2007). The detailed theory and applications of this extended Lagrangian-Hamiltonian mechanics are presented in various references (Mukherjee et al. 2006, Mukherjee et al. 2007). Recently, Mukherjee et. al.(2009) has applied the extended Lagrangian- Hamiltonian mechanics for one dimensional continuous systems with gyroscopic coupling and non-conservative fields.

The discrete continuous modeling of rotor system was presented by Szoic (2000). In his paper, dynamical investigations of rotor shaft systems are performed by means of the discrete continuous mechanical models. In these systems, rotating cylindrical shaft is represented as continuous systems, whereas bearings are assumed as discrete elements. Krenk (2004) has shown the complex mode analysis of cables and beams problems involving concentrated viscous dampers.

The basic aim of this paper is mainly focused on extending the Lagrangian-Hamiltonian mechanics for discrete-continuous systems. In the extended Lagrangian-Hamiltonian mechanics, umbra-Lagrangian density has been used to describe the motion of the continuous system rather than the umbra-Lagrangian itself. The invariance of umbra-Lagrangian density is obtained through an extension of Noether's theorem over manifolds. The case study considered in this paper shows an interesting phenomenon of limiting behaviour of the rotor shaft with internal damping beyond certain threshold speeds of instability, which are obtained theoretically as well as numerically. The effect of discrete external damping is examined, and entrainment of whirling speeds at natural undamped modes is observed.

2. METHODOLOGY

2.1. Noether's Theorem for Continuous Systems

Constants of motion or conserved quantities can be found for continuous systems by applying Noether's theorem. The Noether's theorem (Noether, 1918) states that under certain conditions there exists a set of integrals of motion or dynamical invariants that characterize a field or a system of fields. In fact, Noether proved her theorem for fields, which is based on invariance of the Lagrangian with respect to a certain group of continuous transformations. Here, the symmetries of the Lagrangian density are applied rather than the Lagrangian itself.

2.2 Extended Formulation of Noether's Theorem for Umbra-Lagrangian Density

Extended formulation of Noether's theorem may be obtained by as the methodology provided in Ref. (Mukherjee et al., 2009). However, some basic concepts are being provided in reference (Mukherjee et al., 2006, 2007, 2009) for ready reference of the readers. The extended Noether's equation may be written as,

$$D_t Z_1 + D_x Z_2 + \lim_{\eta \to t} P_r^3 V_t(\mathcal{L}) = 0 \cdot \tag{1}$$

where,

$$Z_1 = \lim_{\eta \to t} \left[\sum_{i=1}^{2} \xi_i(\eta) \left\{ \frac{\partial \mathcal{L}}{\partial \dot{u}_i(\eta,x)} - \frac{d}{dx} \left(\frac{\partial \mathcal{L}}{\partial \dot{u}_{ix}(\eta,x)} \right) + \frac{d^2}{dx^2} \left(\frac{\partial \mathcal{L}}{\partial \dot{u}_{ixx}(\eta,x)} \right) \right\} \right] \tag{2}$$

$$Z_2 = \lim_{\eta \to t} \left[\sum_{i=1}^{2} \left(\xi_i(\eta) \frac{\partial \mathcal{L}}{\partial u_{ix}(\eta,x)} + \xi_i'(\eta) \left(\frac{\partial \mathcal{L}}{\partial u_{ixx}(\eta,x)} \right) - \xi_i(\eta) \frac{d}{dx} \left(\frac{\partial \mathcal{L}}{\partial u_{ixx}(\eta,x)} \right) \right) \right] + $$
$$\lim_{\eta \to t} \left[\sum_{i=1}^{2} \left(\dot{\xi}_i(\eta) \frac{\partial \mathcal{L}}{\partial \dot{u}_{ix}(\eta,x)} + \dot{\xi}_i'(\eta) \left(\frac{\partial \mathcal{L}}{\partial \dot{u}_{ixx}(\eta,x)} \right) - \dot{\xi}_i(\eta) \frac{d}{dx} \left(\frac{\partial \mathcal{L}}{\partial \dot{u}_{ixx}(\eta,x)} \right) \right) \right] \tag{3}$$

Equation (1) is called an extended Noether's field equation for the umbra-Lagrangian. Z_1 may be assumed as local density, Z_2 as current or flux density (often termed as Noether's current density) and the last additional term is called the modulatory convection term, which is the contribution of nonconservative and gyroscopic actions, may be assumed as local rate of production.

3. ANALYSIS OF ROTATING SHAFT WITH IN-SPAN EXTERNAL DAMPING

In this case study, a rotor shaft with internal and in-span discrete external damping driven by a constant speed source through a dissipative coupling is considered as shown in Fig.1. The umbra-Lagrangian of the system is may be written adopting the procedure of Reference (Mukherjee et al., 2009).

Figure 1: Continuous Shaft with Internal and In-span Concentrated External Damping driven by a Constant Speed Source through Dissipative Coupling.

$$\mathcal{L} = \int_{x_0}^{x_1} \left[\sum_{i=1}^{2} \left\{ \frac{1}{2} \rho \left(\frac{\partial u_i(\eta,x)}{\partial \eta} \right)^2 - \frac{1}{2} EI \left(\frac{\partial^2 u_i(\eta,x)}{\partial x^2} \right)^2 - \frac{1}{2} I_d \left(\frac{\partial^2 u_i(\eta,x)}{\partial \eta \partial x} \right)^2 \right\} \right.$$
$$\left. - \mu_i \left\{ \left(\frac{\partial^3 u_1(t,x)}{\partial t \partial x^2} + \dot{\theta}(t) \frac{\partial^2 u_2(t,x)}{\partial x^2} \right) \frac{\partial^2 u_1(\eta,x)}{\partial x^2} + \left(\frac{\partial^3 u_2(t,x)}{\partial t \partial x^2} - \dot{\theta}(t) \frac{\partial^2 u_1(t,x)}{\partial x^2} \right) \frac{\partial^2 u_2(\eta,x)}{\partial x^2} \right\} \right] dx$$
$$\int_{x_0}^{x_1} R_a \delta(x - \lambda)(u_1 \dot{u}_2 - u_2 \dot{u}_1) dx$$
$$- \left[\mu_i \int_{x_0}^{x_1} \left\{ \frac{\partial^2 u_1(t,x)}{\partial x^2} \frac{\partial^3 u_2(t,x)}{\partial t \partial x^2} - \frac{\partial^2 u_2(t,x)}{\partial x^2} \frac{\partial^3 u_1(t,x)}{\partial t \partial x^2} \right\} dx + R_c(\dot{\theta}(t) - \Omega) \right] \theta(\eta)$$
$$+ \frac{1}{2} J \dot{\theta}^2(\eta) \tag{4}$$

where R_c is the dissipative coupling, R_a is the damping coefficient of in-span external damper, and Ω is the excitation frequency. In Eq. (4), the term contributed by external damping is very significant. For this system with in-span discrete external damping, the term, which needs special consideration, is

$$\int_0^L R_a \delta(x - \lambda)(u_1 \dot{u}_2 - u_2 \dot{u}_1) dx, \text{ where } \lambda \in (0, L). \tag{5}$$

In this case, the following form of motion is assumed

$$u_1 = \sum_{n=1}^{\infty} A_n \cos \Omega_n t \sin \frac{n\pi}{L} x \qquad \text{and,}$$
$$u_2 = \sum_{n=1}^{\infty} A_n \sin \Omega_n t \sin \frac{n\pi}{L} x, \tag{6}$$

where A_n are slowly varying functions of time.

Substitution of Eq. (6) into Eq. (5) yields

$$R_a \left(\Omega_n \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} A_n A_m \sin \frac{n\pi}{L} \lambda \, \sin \frac{m\pi}{L} \lambda \right), \tag{7}$$

with the assumption that A_n vary slowly with time.

3.1 Extended Noether's theorem for rotating shaft

The modified Noether's rate equation with in span concentrated damping may be expressed as

$$\frac{L}{2}\sum_{n=1}^{\infty} A_n \left[\begin{array}{c} \dot{A}_n + \frac{1}{2\rho}\left\{ \mu_i \frac{n^4\pi^4}{L^4} - \frac{\mu_i\, n^4\pi^4}{\Omega_n\, L^4}\dot{\theta}(t)\right\}A_n \\ + \frac{R_a}{\rho\, L}\sum_{k=1}^{\infty} A_k \sin\frac{n\pi}{L}\lambda\,\sin\frac{k\pi}{L}\lambda \end{array} \right] = 0 . \quad (8)$$

Considering independent variations in $A_n + \delta A_n$, the variational equation may be written as (dropping the factor $\frac{L}{2}$)

$$\sum_{n=1}^{n} \delta A_n \left[\begin{array}{c} \dot{A}_n + \frac{1}{2\rho}\left\{ \mu_i \frac{n^4\pi^4}{L^4} - \frac{\mu_i\, n^4\pi^4}{\Omega_n\, L^4}\dot{\theta}(t)\right\}A_n \\ + \frac{R_a}{\rho\, L}\sum_{k=1}^{\infty} A_k \sin\frac{n\pi}{L}\lambda\,\sin\frac{k\pi}{L}\lambda \end{array} \right]$$

$$+ \sum_{n=1}^{n} A_n \delta \left[\begin{array}{c} \dot{A}_n + \frac{1}{2\rho}\left\{ \mu_i \frac{n^4\pi^4}{L^4} - \frac{\mu_i\, n^4\pi^4}{\Omega_n\, L^4}\dot{\theta}(t)\right\}A_n \\ + \frac{R_a}{\rho\, L}\sum_{k=1}^{\infty} A_k \sin\frac{n\pi}{L}\lambda\,\sin\frac{k\pi}{L}\lambda \end{array} \right] = 0.$$

$$(9)$$

This is for the symmetry being valid for neighbouring paths, which needs the following conditions to be satisfied

$$\dot{A}_n + \frac{1}{2\rho}\left\{ \mu_i \frac{n^4\pi^4}{L^4} - \frac{\mu_i\, n^4\pi^4}{\Omega_n\, L^4}\dot{\theta}(t)\right\}A_n$$
$$+ \frac{R_a}{\rho\, L}\sum_{k=1}^{\infty} A_k \sin\frac{n\pi}{L}\lambda\,\sin\frac{k\pi}{L}\lambda = 0 \quad (10)$$

The condition for entraining the nth mode is $A_n \to$ finite limit and $A_k \to 0$, if $k \neq n$ for $t \to \infty$, obtained as

$$\frac{1}{2\rho}\left\{ \mu_i \frac{n^4\pi^4}{L^4} - \frac{\mu_i\, n^4\pi^4}{\Omega_n\, L^4}\dot{\theta}(t)\right\} + \frac{R_a}{\rho\, L}\sin^2\frac{n\pi}{L}\lambda = 0 , \quad (11)$$

The value of $\dot{\theta}(t)$ will be obtained from Eq. (11) and written as

$$\dot{\theta}(t) = \Omega_n \left[1 + \frac{2R_a L^3}{\mu_i n^4 \pi^4}\sin^2\frac{n\pi}{L}\lambda \right]. \quad (12)$$

3.2 Umbra-Hamiltonian density of the system

Umbra-Hamiltonian density for this system may be written as

$$\mathcal{H} = \int_{x_0}^{x_1} \left[\begin{array}{c} \sum_{i=1}^{2}\left\{ \begin{array}{c} \frac{1}{2\rho}P_i^2(\eta,x) + \frac{1}{2}EI\left(\frac{\partial^2 u_i(\eta,x)}{\partial x^2}\right)^2 \\ + \frac{1}{2}I_d\left(\frac{\partial^2 u_i(\eta,x)}{\partial\eta\,\partial x}\right)^2 \end{array}\right\} \\ + \mu_i\left\{ \begin{array}{c} \left(\frac{\partial^3 u_1(t,x)}{\partial t\,\partial x} + \dot{\theta}(t)\frac{\partial^2 u_2(t,x)}{\partial x^2}\right)\frac{\partial^2 u_1(\eta,x)}{\partial x^2} \\ + \left(\frac{\partial^3 u_2(t,x)}{\partial t\,\partial x} - \dot{\theta}(t)\frac{\partial^2 u_1(t,x)}{\partial x^2}\right)\frac{\partial^2 u_2(\eta,x)}{\partial x^2} \end{array}\right\} \end{array} \right] dx$$

$$+ \int_{x_0}^{x_1} R_a(x-\lambda)(u_1\dot{u}_2 - u_2\dot{u}_1)dx$$

$$+ \left[\begin{array}{c} \mu_i \int_{x_0}^{x_1}\left\{ \frac{\partial^2 u_1(t,x)}{\partial x^2}\frac{\partial^3 u_2(t,x)}{\partial t\,\partial x^2} - \frac{\partial^2 u_2(t,x)}{\partial x^2}\frac{\partial^3 u_1(t,x)}{\partial t\,\partial x^2}\right\}dx \\ + R_c\left(\dot{\theta}(t) - \Omega\right) \end{array} \right]\theta(\eta)$$

$$+ \frac{1}{2J_d}P_\theta^2 ,$$

$$(13)$$

where $P_i(\eta,x) = \rho\dfrac{\partial u_i(\eta,x)}{\partial\eta}$ and $P_\theta(\eta) = J_d\dot{\theta}(\eta)$.

As discussed in reference (Mukherjee et al., 2009), one has the similar theorems for umbra-Hamiltonian density, expressed as

$$\underset{\eta\to t}{Lim}\frac{\partial\mathcal{H}^*}{\partial\eta} = 0 \Rightarrow \frac{d\mathcal{H}_i^*}{dt} = -\underset{\eta\to t}{Lim}\frac{d\mathcal{H}_e^*}{\partial\eta} .$$

Now considering exterior umbra-Hamiltonian density, one may have

$$\underset{\eta\to t}{Lim}\frac{\partial\mathcal{H}_e}{\partial\eta} = \int_{x_0}^{x_1}\left[\begin{array}{c} R_a(x-\lambda)(u_1\dot{u}_2 - u_2\dot{u}_1)dx \\ + \mu_i\left[\begin{array}{c} \left\{\frac{\partial^3 u_1(t,x)}{\partial t\,\partial x^2} + \dot{\theta}(t)\frac{\partial^2 u_2(t,x)}{\partial x^2}\right\}\frac{\partial^3 u_1(t,x)}{\partial t\,\partial x^2} \\ + \left\{\frac{\partial^3 u_2(t,x)}{\partial t\,\partial x^2} - \dot{\theta}(t)\frac{\partial^2 u_1(t,x)}{\partial x^2}\right\}\frac{\partial^3 u_2(t,x)}{\partial t\,\partial x^2} \end{array}\right]dx \end{array} \right]$$
$$+ \left[\begin{array}{c} \mu_i \int_{x_0}^{x_1}\left\{ \begin{array}{c} \frac{\partial^2 u_1(t,x)}{\partial x^2}\frac{\partial^3 u_2(t,x)}{\partial t\,\partial x^2} \\ - \frac{\partial^2 u_2(t,x)}{\partial x^2}\frac{\partial^3 u_1(t,x)}{\partial t\,\partial x^2} \end{array}\right\}dx \\ + R_c\left(\dot{\theta}(t) - \Omega\right) \end{array} \right]\dot{\theta}(t) = 0.$$

$$(14)$$

Considering end conditions of continuous shaft as pin-pin, and substituting Eq. (6) in Eq. (14), one obtains the following two terms distinguished as $\{M\}$ and $\{N\}$,

$$Lim_{\eta \to t} \frac{\partial \mathcal{H}_e}{\partial \eta} = R_a \Omega_n \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} A_n A_m \sin \frac{n\pi}{L} x \sin \frac{m\pi}{L} x$$

$$+ \mu_i \frac{L}{2} \left[\begin{array}{c} \sum_{n=1}^{\infty} A_n^2 \Omega_n^2 \left(\frac{n\pi}{L} \right)^4 \\ + \dot{\theta}(t) \sum_{n=1}^{\infty} A_n^2 \Omega_n \left(\frac{n\pi}{L} \right)^4 + \sum_{n=1}^{\infty} \dot{A}_n^2 \left(\frac{n\pi}{L} \right)^4 \end{array} \right]$$

$$----------Term\{M\}--------$$

$$+ \left[\mu_i \left\{ \int_0^L \sum_{n=1}^{\infty} A_n^2 \Omega_n \sin^2 \frac{n\pi x}{L} \left(\frac{n\pi}{L} \right)^4 \right\} dx + R_c \left(\dot{\theta}(t) - \Omega \right) \right] \dot{\theta}(t).$$

$$-----------Term\{N\}-----------$$

Equating term {M} to zero, one obtains

$$\dot{A}_n^2 + A_n^2 \Omega_n \left[\begin{array}{c} \dfrac{\mu_i \Omega_n n^4 \pi^4}{L^4} - \dfrac{\mu_i n^4 \pi^4}{L^4 \Omega_n} \dot{\theta}(t) \\ + \dfrac{2 R_a}{L} \sin^2 \dfrac{n\pi}{L} \lambda \end{array} \right] = 0 .$$

Considering independent variations in $A_n + \delta A_n$, the variational equation may be written similar to Eq. (9) after dropping the factor $L/2$. Repeating the same steps from Eq. (10) to (11), one obtains the value of $\dot{\theta}(t)$ as

$$\dot{\theta}(t) = \Omega_n \left[1 + \frac{2 R_a L^3}{\mu_i n^4 \pi^4} \sin^2 \frac{n\pi}{L} \lambda \right],$$

which makes the term {M} zero and also matches with Eq. (12).

Equating term {N} to zero, one obtains

$$\mu_i \sum_{n=1}^{\infty} A_n^2 \Omega_n \left(\frac{n\pi}{L} \right)^4 \int_0^L \phi_n^2 dx + R_c \left(\dot{\theta}(t) - \Omega \right) = 0 . \qquad (10)$$

If the shaft's speed is latched at $\dot{\theta}(t) = \Omega_n \left[1 + \dfrac{2 R_a L^3}{\mu_i n^4 \pi^4} \sin^2 \dfrac{n\pi}{L} \lambda \right]$, then all the

amplitudes go to zero except A_n, thus; the amplitude of this shaft is obtained as

$$A_n = \sqrt{ \frac{2 R_c \left(\Omega - \underset{n \in I^+}{Min} \Omega_n \left[1 + \dfrac{2 R_a L^3}{\mu_i n^4 \pi^4} \sin^2 \dfrac{n\pi}{L} \lambda \right] \right)}{\mu_i \dfrac{n^4 \pi^4}{L^3} \Omega_n} } ,$$

or

$$A_n = \sqrt{ \frac{2 R_c \left(\Omega - \underset{n \in I^+}{Min} \Omega_n \left[1 + \dfrac{2 R_a L^3}{\mu_i n^4 \pi^4} \sin^2 \dfrac{n\pi}{L} \lambda \right] \right)}{\zeta_n \Omega_n} } ,$$

$$(11)$$

where $\zeta_n = \mu_i \dfrac{n^4 \pi^4}{L^3}$. The next section presents the

modeling of rotating shaft with in span concentrated damping with simulation results.

4. Simulation of rotating shaft with in- span discrete external damping

The physical system with in-span discrete external damping driven through dissipative coupling is shown in Fig.1. The boundary conditions of the rotor shaft are taken as pin-pin. The coupling in the system is absolutely flexible in transverse and bending but torsionally rigid. In this analysis, torsional vibration is not considered. The bond graph technique is being used as a modeling tool and bond graph model of the system is created using object-oriented reusable capsules with in-span discrete external damping.

4.1 Simulation results

The bond graph model (Mukherjee et al. 2006, Karnopp et al., 1990) of the rotor shaft with in span concentrated external damping is simulated on SYMBOLS-Shakti (Mukherjee and Samantaray, 2000), in order to visualize the complex modes of the system. The integrated bond graph model with object-oriented sub models (Mukherjee, Karmakar and Samantaray, 2006) of shaft and hub elements is shown in Fig. 2.

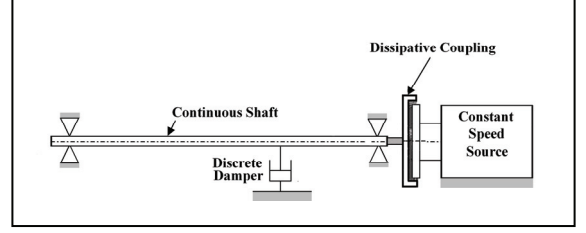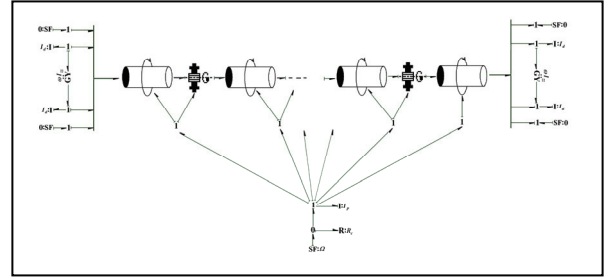Figure 2: Bond graph Model of Integrated System of Rotor Shaft with Internal and In-span Concentrated External Damping driven by a Constant Speed Source through Dissipative Coupling.

However, the external damping in sub models of shaft has been considered as discrete or finite. In this case, the effects of concentrated damper are closely linked to the complex character of modes. The shaft is rotated by a constant speed source. The simulation rig consists of a hollow rotating shaft with 10 reticules and two ends are well supported on a self-aligning bearing. The dimensional data are as follows: L_{beam} =5 m, R_o =0.02 m, R_i =0.01 m. The material and bearing properties are as follows: E =104.5 e^9, ρ =4420, R_c =0.002, μ_i =2.0 e^{-4}, R_a =0.020 (for Mode 1). The damping coefficient of dissipative coupling is taken as $R_c = 0.002$. It has been assumed that in-span discrete damper permits motion to the rotor and at the same time leads to substantial energy dissipation.

Initially, simulation is done for the parameters given in Table-1 with discrete external damping, $R_a = 0.020$ and excitation frequency, Ω =5 Hz or 31.41 rad/s. To initiate the simulation, an initial momentum of 0.001 Kg- m^2 was given. This is done to reduce the simulation time. Fig. 3 shows that the trajectories of the rotating shaft reach the limiting orbit.

This is due to the loading of the source. The another feature is that the angular speed of the shaft gets entrained at 22.392 rad/s, the first threshold speed of instability, $\dot{\theta}_{1^{st}\,mode}$ and matches closely with the calculated value of shaft spinning speed (entrained) $\dot{\theta}(t)$ (Table-1), which is equal to 22.266 rad/s.
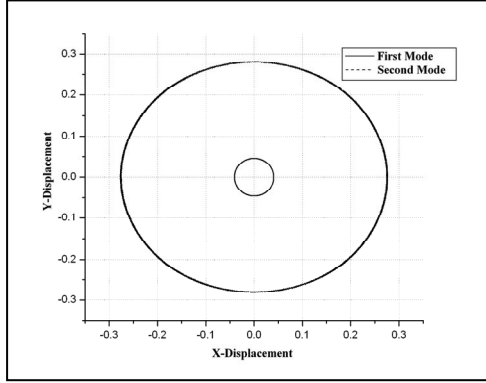
Figure3: Limiting Orbit of Shaft with In-span Concentrated External Damping at First Two Undamped Natural Modes

Table-1: Simulation Parameters

L_{beam}	Length of beam	5 m
N_{elem}	Number of elements	10
E	Modulus of elasticity	104.5 e9 N/m^2
R_i	Internal Diameter	0.01 m
R_o	External Diameter	0.02 m
ρ	Material density	4420 kg/m^3
R_c	Dissipative coupling coefficient	0.002
Ω	Excitation frequency	5 Hz for first mode 30 Hz for second mode
μ_i	Internal Damping Coefficient	1.0 e-4 Ns/m
μ_{ex}	Discrete External Damping Coefficient	0.020 for first mode 6.0 for second mode
$\dot{\theta}(t)$	Shaft Spinning Speed (entrained) Calculated	22.392rad/s, 95.405rad/s, For Mode 1,2, respectively

Table-2: Calculation of Natural Frequency

First mode natural frequency	3.414 Hz	Second mode natural frequency	13.66 Hz

Now, in span concentrated external damping, R_a of the shaft is increased to 6.0 and excitation frequency, Ω to 30 Hz or 188.327 rad/sec, the next threshold speed of the shaft is obtained. Fig.3 shows the trajectories of the same rotating shaft and angular speed of the shaft gets entrained at 95.490 rad/sec, the second threshold speed of instability, $\dot{\theta}_{2^{nd}\,mode}$ and matches very close to the calculated value of shaft spinning speed as given in Table-1. The interesting phenomenon observed from this simulation is that the one obtains the natural undamped modes in this case. The result may also be analyzed through FFT analyzer tool as shown in Fig.4, where first two natural frequencies for natural modes are superimposed, which matches exactly with the calculated natural modes as given in Table-2.
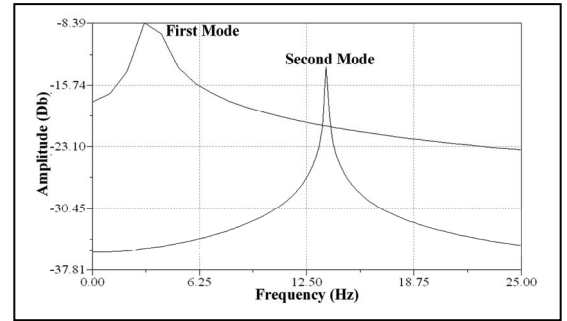
Figure 4: Superimposed Frequency Responses at First Two Undamped Natural Modes.

The results may also be visualized through animation, where two natural modes are shown in Fig. 5(a)- 5(b) through animated frames.
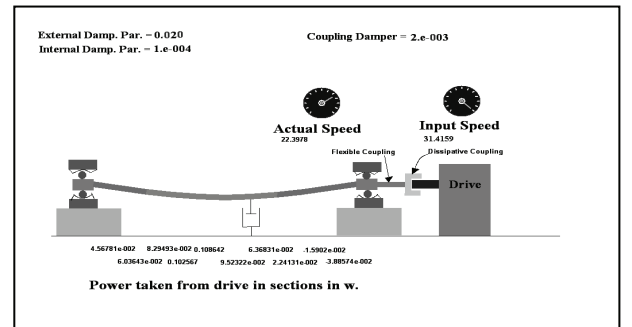
Figure 5(a): First Natural Mode of Rotor Shaft with In span Concentrated External Damping R_a=0.020, μ_i=1e-4, Input Speed=31.41 rad/s, Actual Speed=22.392rad/s.
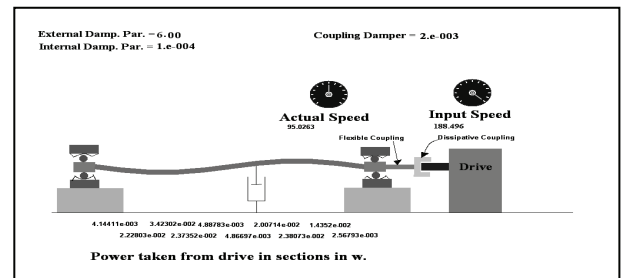
Figure 5(b): Second Natural Mode of Rotor Shaft with In-span Concentrated External Damping R_a=6.00, μ_i=1e-4, Input Speed=188.496 rad/s, Actual Speed=95.490rad/s.

The animation shows that regenerative energy in the shaft due to internal damping is dissipated through the discrete damper and the dissipative coupling between drive and the rotor shaft. If excitation frequency is more, then the shaft speed will not increase but the slip between drive and shaft will increase due to loading of drive.

4. Conclusions

In this paper, interesting case study of a rotor shaft with internal damping and in-span external damping driven through a dissipative coupling has been presented. The dynamic behaviour has been obtained through extended Noether's theorem and umbra-Hamiltonian theoretically as well as numerically. The study has examined the various aspects of limiting dynamics of this rotor and the same results are validated through simulation. Further, the study examined that the regenerative energy in the shaft due to internal damping is dissipated through the in-span damper and the dissipative coupling. Limiting dynamics basically occurred due to the power imported by internal damping from the shaft spin, which is balanced by dissipation of power by in-span external damper, dissipation in the coupling and a part of action of internal damping, acts as a external damping. The animation results of the system have revealed the entrainment of the whirl speed at different natural frequencies.

Nomenclature:

A_n = Amplitude of nth mode of the rotor
EI = Rigidity of the continuous rotor
H^* = Umbra-Hamiltonian of the system
I_d = Rotary inertia of the rotor
I_p = Inertia of the beam through principal axis
L = Lagrangian of the system
L^* = Umbra-Lagrangian of the system
R_c = Damping coefficient of dissipative coupling
R_a = In-span discrete damper
V = Infinitesimal generator of rotational SO (2) group
V_t = Real time component of infinitesimal generator
V_η = Umbra time component of infinitesimal generator
a = Cross sectional area of the rotor
n = Mode number
$p(\eta)$ = Umbra-time momentum
$p(t)$ = Real-time momentum
$q(t)$ = Generalized displacement in real time
$q(\eta)$ = Generalized displacement in umbra-time
$\dot{q}(t)$ = Generalized velocity in real time
$\dot{q}(\eta)$ = Generalized velocity in umbra-time

s = Angle or linear variables for rotational transformation or linear transformation.
$x_i(\ )$ = Linear displacements in real time or umbra-time, where $i = 1...n$
$\dot{x}_i(\ )$ = Linear velocity in real time or umbra-time, where $i = 1...n$
t = Real-time in s.
Ω = Excitation frequency in rad/s.
Ω_n = Natural frequency of the rotor shaft in rad/s.
η = Umbra-time in s.
ω = constant angular velocity
$\theta(\ )$ = Angular displacement in umbra-time or real time in rad.
$\dot{\theta}(\ )$ = Angular velocity of the shaft in umbra-time or real time in rad/s.
$\mathcal{L}$ = Umbra-Lagrangian density
ρ = Mass density of rotor shaft
μ_a = External damping of the beam
μ_i = Internal damping of the beam
γ^* = Damping ratio
$u_i(t)$ = Real displacement coordinates of beam
$u_i(\eta)$ = Umbra- displacement coordinates of beam
$\mathcal{H}$ = Umbra-Hamiltonian density
$\mathcal{H}_i$, $\mathcal{H}_e$ = Interior and exterior umbra-Hamiltonian density.

Reference:

Arnold, V.I., 1974. *Mathematical Methods of Classical Mechanics*, Springer-Verlag, New York–Heidelberg-Berlin, 88.

Boyer, T. H., 1967. Continuous Symmetries and Conserved Currents. *Annals of Physics, 42*(3), 445-466.

Brown, F.T., 2007. Engineering System Dynamics, 2ed, CRC, Taylor & Francis.

Damianou P.A. and C. Sophocleous, 2004. Classification of Noether Symmetries for Lagrangian with Three Degree of Freedom, *Nonlinear Dynamics*, 36, 3.

Hassani, S., 1999. *Mathematical Physics*, Springer-Verlag, New York-Heidelberg-Berlin, 936.

Katzin G.H. and J. Levine, 1976. A Gauge Invariant Formulation of Time-dependent Dynamical Symmetry Mappings and Associated Constants of Motion for Lagrangian Particle Mechanics. *Int. J. Math. Phys.* 17(7), 1345.

Krenk, S., 2004. Complex Modes and Frequencies in Damped structural vibrations, *Journal of Sound and Vibration*, 270, 981-996.

Karnopp, D.C., R.C. Rosenberg, and D.L. Margolis, 1990. *System Dynamics: A Unified Approach*, John-Wiley and Sons Inc., USA .

Mukherjee, A., 1994. Junction Structures of Bondgraph Theory from Analytical Mechanics Viewpoint, *Proc. of CISS-1st Conference of International Simulation Societies*. Zurich, Switzerland, 661.

Mukherjee A. and A. K. Samantaray, 2000. *SYMBOLS-2000 User's Manual*, High-Tech Consultants, STEP Indian Institute of Technology, Kharagpur, India.

Mukherjee, A., V. Rastogi, and A. Dasgupta, 2006. A Methodology for finding Invariants of Motion for Asymmetric Systems with Gauge-transformed Umbra Lagrangian Generated by Bond Graphs, *Simulation*, 82(4), 207-226.

Mukherjee, A. and R. Karmakar, and A. Samantaray, 2006. *Modelling and Simulation of engineering System through Bond Graph*, Narosa publishing House, New Delhi, reprinted by CRC press for North America and by Alpha Science for Europe.

Mukherjee, A., V. Rastogi, and A. Dasgupta, 2007. A Study of a Bi-Symmetric Electro-mechanical System through Umbra Lagrangian Generated by Bond Graphs, and Noether's Theorem, *Simulation*, 83(9), 611-630.

Mukherjee, A., V. Rastogi, and A. Dasgupta, 2009 Extension of Lagrangian-Hamiltonian Mechanics for Continuous Systems-Investigation of Dynamics of a One dimensional internally Damped Rotor driven through a dissipative coupling, *Nonlinear Dynamics*, 58 (1), 107-127.

Szoic, T., 2000. On the discrete Continuous Modelling of Rotor system for the analysis of Coupled Lateral Vibration, *International Journal of Rotating Machinery*, 6 (2), 135-149.

Sarlet W. and F. Cantrijn, 1981. Generalizations of Noether's theorem in Classical Mechanics. *SIAM Rev.*, 23(4), 467.

Noether, E., 1918. Invariante Variationsprobleme, *Ges. Wiss. Gottington*. 2, 235.

Olver, P., 1986. *Application of Lie groups to Differential Equations*, Springer-Verlag.

AUTHORS BIOGRAPHY

Vikas Rastogi: He obtained his Master of Technology degree from NIT, Kurukshetra in 1995. In the same year, he joined as a lecturer in Sant Longowal Institute of Engineering & Technology, SLIET, Punjab. He has completed his PhD from IIT, Kharagpur in 2005. He is presently working as an Associate Professor in SLIET, Longowal (Deemed to be University). His research areas are Lagrangian –Hamiltonian mechanics for general class of systems, study of symmetries for discrete and continuous systems, Modeling and Simulation for dynamical systems.

Amalendu Mukherjee: He joined as a lecturer in Mechanical Engineering Department in 1970. In the year 1980, he was awarded Alexander Von Humboldt fellowship for postdoctoral work in Rotor dynamics at Clusthal-Zellerfeld, West Germany. On his return, he undertook the studies on unified approach system modelling through Bond graph theory. He then initiated the creation of family of Software now called SYMBOLS-Shakti for modelling and simulation. SYMBOLS Shakti is being used by several industries, Universities in India and abroad. Presently, he is working in field of Classical mechanics and Study of Symmetries through bond graphs. He is the author of a book as well as Video Course on Bond Graph Modelling of Dynamical Systems.

Anirvan Dasgupta: He received his Ph. D. degree from Indian Institute of Technology, Kanpur. He is presently working as Professor in Mechanical Engineering Department at IIT, Kharagpur. In 2002, he was awarded Alexander Von Humboldt fellowship for postdoctoral work in Germany. His present areas of work are Non-holonomic motion planning, space and under water robot dynamics, and dynamics of continuous systems.

A RECURSIVE NONLINEAR SYSTEM IDENTIFICATION METHOD BASED ON BINARY MEASUREMENTS

Laurent Bourgois[a], Jérôme Juillard[b]

SUPELEC E3S
Plateau du Moulon
3 rue Joliot-Curie
91192 Gif-sur-Yvette Cedex, France

[a]laurent.bourgois@supelec.fr, [b]jerome.juillard@supelec.fr

ABSTRACT

An online approach to nonlinear system identification based on binary observations is presented in this paper. This recursive method is a nonlinear extension of the LMS-like (least-mean-square) identification method using binary observations. It can be applied in the case of weakly nonlinear Duffing oscillator coupled with a linear system characterized by a finite impulse response. It is then possible to estimate simultaneously both impulse response and Duffing coefficients, knowing only the system input and the sign of the system output. The impulse response is identified up to a positive multiplicative constant. The proposed method is compared in terms of convergence speed and estimation quality with the usual LMS approach, which is not based on binary observations.

Keywords: nonlinear system identification, self-test, binary data processing, microsystems

1. INTRODUCTION

Microfabrication of electronic components such as micro-electro-mechanical systems (MEMS) or nano-electro-mechanical systems (NEMS) has known an increasing interest over the past two decades. The most notable innovation emanating from these systems is the possibility to massively integrate sensors with self-test features on the same piece of silicon. Indeed, it is well-known that, as characteristic dimensions become smaller, the dispersions afflicting electronic devices tend to become larger. Typical sources of dispersions and uncertainties are variations in the fabrication process or environmental disturbances such as temperature, pressure and humidity fluctuations. As a result, it is usually impossible to guarantee *a priori* that a given device will work properly. Moreover, expensive tests must then be run after fabrication to ensure that only suitable devices are commercialized. An alternative consists in implementing self-test (and self-tuning) features such as parameter estimation routines, so that devices can adapt to changing conditions.

However, traditional identification methods (Walter and Pronzato 1997; Ljung 1999) are often tricky to "straightforwardly" adapt from macroscopic scale to microscopic scale. Their integration requires the implementation of a high-resolution analog-to-digital converter (ADC), which results in longer design times as well as larger silicon areas. Thus, parameter estimation routines based on binary observations are very attractive because they only involve the integration of a 1-bit ADC. Some important contributions that keep the added cost of testing as small as possible are available in the literature.

In (Wigren 1998), Wigren has developed an LMS approach to the problem of online parameter estimation from quantized observations. The principle is to estimate the gradient of the least-square criterion by approximating the quantizer. Under some hypothesis, it is possible to guarantee the asymptotic convergence of this method to the nominal parameters. In (Negreiros, Carro, and Susin 2003), the authors have suggested using a white Gaussian input to excite the unknown linear system and to estimate the power spectral density (PSD) of the binary output. From this estimated PSD, the modulus of the unknown system transfer function can be analytically derived. However, it is not possible to obtain any information concerning the phase of this transfer function. This limitation has been overcome by deriving an analytical relationship between the impulse response coefficients of the system and the cross-covariance of its binary input and output. Although this approach is fairly simple to implement, it relies on the mixing properties of the linear system, which may not be guarantee *a priori*. Recently, a basic identification method using binary observations (BIMBO) has been introduced in (Colinet and Juillard 2010). The theoretical framework of this offline WLS-like (weighted-least-square) approach is based on the minimization of a criterion, where the parameter-dependent weights are chosen in order to smooth out the discontinuities of the unweighted least-square criterion. It is then possible to guarantee the consistency of this approach even in the presence of measurement noise, provided that the signal at the input of the quantizer is Gaussian and centered. Furthermore, the estimation quality of BIMBO has been investigated in the sense of

correlation coefficient between the nominal and the estimated system parameters. An alternative WLS criterion, which is easier to implement in the context of microelectronics, has also been presented in (Juillard, Jafari, and Colinet 2009). This approach is as efficient as the previous one without measurement noise, but leads to a systematic error otherwise. Finally, an online LMS-like identification method based on binary observations (LIMBO) has been derived from the past two WLS approaches in (Jafari, Juillard, and Colinet 2010). Simulations have provided similar results than those obtained with the Wigren's method in terms of convergence speed and estimation quality, and those with a lesser computational complexity.

Unfortunately, the methods listed above deal with linear systems, while in many engineering applications, and especially in microfabricated devices, the dynamic may significantly be affected by nonlinear effects, which must be accounted for in order to robustly model the system. In (Zhao, Wang, Yin, and Zhang 2007), the authors have studied identification of Wiener and Hammerstein systems, which are particular nonlinear structures, with binary-valued output observations. In this paper, we propose to extend LIMBO to nonlinear systems. For that, we consider a nonlinear Duffing oscillator that is coupled with a linear system characterized by a finite impulse response. The convergence of this recursive method is illustrated by numerical simulations and our results are compared with those obtained by the conventional LMS algorithm (*i.e.* without quantization).

The structure of the article is the following. In section 2, the nonlinear system and its model are introduced. In section 3, the nonlinear LIMBO algorithm is derived. In section 4, the proposed method is compared with a traditional online method, which is not based on binary observations, in terms of convergence speed and estimation quality. Finally, concluding remarks and perspectives are given in section 5.

2. FRAMEWORK AND NOTATIONS
Let us consider a nonlinear system illustrated in figure 1 below.
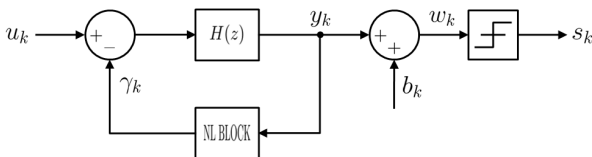
Figure 1: Block diagram of the system model

The first branch of this block diagram corresponds to a discrete-time linear time-invariant system H. We assume that this transfer function has a finite impulse response of length L, *i.e.* the impulse response can be represented by a column vector $\boldsymbol{\theta} = (\theta_l)_{l=1}^L$. A cubic nonlinearity (the so-called Duffing nonlinearity) is then introduced at the level of the negative feedback branch such that $\gamma_k = \alpha y_k^3$ with $\alpha \in \mathbb{R}_+$. Obviously, the

subscript indices k denotes the discrete time. Let $\boldsymbol{b}$ be an unknown additive measurement noise and let $\boldsymbol{w} = \boldsymbol{y} + \boldsymbol{b}$ be the noisy output. The system output is measured via a 1-bit ADC such that only the sign $s_k = S(w_k)$ is known. Here, the function S of a real number x is defined as follows:

$$S(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ -1 & \text{otherwise} \end{cases} \tag{1}$$

By supposing the system weakly nonlinear, the following approximation can be done (Schoukens, Nemeth, Crama, Rolain, and Pintelon 2003):

$$
\begin{aligned}
y_k &= h_k * (u_k - \alpha y_k^3) \\
&= h_k * \left(u_k - \alpha \big(h_k * (u_k - \alpha y_k^3) \big)^3 \right) \\
&\simeq h_k * (u_k - \alpha (h_k * u_k)^3)
\end{aligned}
\tag{2}
$$

This approximation remains valid as long as the nonlinear term αy_k^3 remains smaller or of the same order of magnitude as u_k. The new block diagram of the system model is then graphically illustrated in figure 2 below.
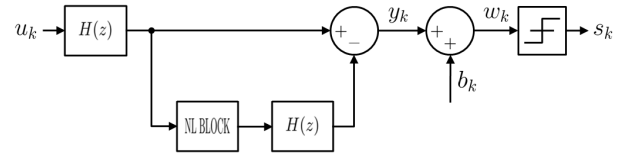
Figure 2: New block diagram of the system model

Consequently, the scalar value of the system output at time k is given by:

$$y_k = \boldsymbol{\theta}^T \boldsymbol{\varphi}_{k,L} - \alpha \boldsymbol{\theta}^T \boldsymbol{\psi}_{k,L} \tag{3}$$

In relation (3), $\boldsymbol{\varphi}_{k,L} = (u_l)_{l=k}^{k-L+1}$ is the regression column vector of dimension L at time k and $\boldsymbol{\psi}_{k,L}$ is defined by:

$$\boldsymbol{\psi}_{k,L} = \left(\left(\boldsymbol{\theta}^T \boldsymbol{\varphi}_{l,L} \right)^3 \right)_{l=k}^{k-L+1} \tag{4}$$

Our purpose is to develop a recursive estimation method to find good estimates of both the parameter vector $\boldsymbol{\theta}$ and the Duffing coefficient α, starting from N observations of the binary output $\boldsymbol{s}$ and knowing the input $\boldsymbol{u}$. Let $\widehat{\boldsymbol{\theta}}_k$ (respectively, $\hat{\alpha}_k$) be the estimated parameter vector (respectively, Duffing coefficient) at time k. Let us also introduce $\hat{y}_k$ the estimated system output at time k and $\hat{s}_k = S(\hat{y}_k)$.

3. PROPOSED LMS APPROACH
Under its original form (Jafari, Juillard, and Colinet 2010), LIMBO has been carried out in order to estimate online the parameters of a linear system from binary

observations. From now on, since only the sign s_k of the system output is available at time k, the authors have judiciously defined the following instantaneous error:

$$\varepsilon_k = |s_k - \hat{s}_k| \hat{y}_k \tag{5}$$

The general LMS algorithm is used to adjust the system parameters by minimizing this error using the following criterion:

$$\varepsilon_k^2 = (s_k - \hat{s}_k)^2 \hat{y}_k^2 \tag{6}$$

This suitable formulation has been specified to ensure the derivability with respect to $\widehat{\boldsymbol{\theta}}_k$. We adopt the same criterion to deal with nonlinear constraints. Indeed, ε_k^2 is clearly differentiable with respect to $\hat{y}_k$, which is also differentiable with respect to $\widehat{\boldsymbol{\theta}}_k$ and $\hat{\alpha}_k$. Consequently, the criterion defined in (6) is differentiable with respect to the system parameters, and we can write:

$$
\begin{aligned}
\widehat{\boldsymbol{\theta}}_{k+1} &= \widehat{\boldsymbol{\theta}}_k - \frac{1}{2}\mu \frac{\partial \varepsilon_k^2}{\partial \widehat{\boldsymbol{\theta}}_k} \\
&= \widehat{\boldsymbol{\theta}}_k - \frac{1}{2}\mu \frac{\partial \varepsilon_k^2}{\partial \hat{y}_k} \frac{\partial \hat{y}_k}{\partial \widehat{\boldsymbol{\theta}}_k} \\
&= \widehat{\boldsymbol{\theta}}_k - \mu(s_k - \hat{s}_k)^2 \hat{y}_k \frac{\partial \hat{y}_k}{\partial \widehat{\boldsymbol{\theta}}_k}
\end{aligned} \tag{7}
$$

In the same manner, the following relation can be established:

$$\hat{\alpha}_{k+1} = \hat{\alpha}_k - \lambda(s_k - \hat{s}_k)^2 \hat{y}_k \frac{\partial \hat{y}_k}{\partial \hat{\alpha}_k} \tag{8}$$

In relations (7) and (8), μ and λ correspond to the LMS step-size parameters that guarantee stability and control the speed of convergence. The selection of these parameters is therefore very critical. We propose a procedure to determine an appropriate adaptive step-size. We start by assuming that $\mu = \lambda$. The idea then consists in finding μ_k such that $\hat{y}_k|_{(\widehat{\boldsymbol{\theta}}_{k+1}, \hat{\alpha}_{k+1})} = 0$. Unfortunately, no analytical solution of this problem is currently available (except in the trivial linear case). Nevertheless, a numerical solution can be obtained without much difficulty directly by applying the widely-used secant method, which is known to be a fast iterative method. In counterpart, an extra division is required.

The derivative with respect to $\widehat{\boldsymbol{\theta}}_k$ which appears in relation (7) can be expressed by:

$$\frac{\partial \hat{y}_k}{\partial \widehat{\boldsymbol{\theta}}_k} = \boldsymbol{\varphi}_{k,L} - \hat{\alpha}_k \left(\boldsymbol{\psi}_{k,L} + \widehat{\boldsymbol{\theta}}_k^T \frac{\partial \boldsymbol{\psi}_{k,L}}{\partial \widehat{\boldsymbol{\theta}}_k} \right) \tag{9}$$

By using the expression of $\boldsymbol{\psi}_{k,L}$ introduced in relation (4), we have:

$$\frac{\partial \boldsymbol{\psi}_{k,L}}{\partial \widehat{\boldsymbol{\theta}}_k} = 3 \left(\boldsymbol{\varphi}_{l,L} \left(\widehat{\boldsymbol{\theta}}_k^T \boldsymbol{\varphi}_{l,L} \right)^2 \right)_{l=k}^{k-L+1} \tag{10}$$

This yields the following relation:

$$\widehat{\boldsymbol{\theta}}_k^T \frac{\partial \boldsymbol{\psi}_{k,L}}{\partial \widehat{\boldsymbol{\theta}}_k} = 3 \left(\left(\widehat{\boldsymbol{\theta}}_k^T \boldsymbol{\varphi}_{l,L} \right)^3 \right)_{l=k}^{k-L+1} = 3\boldsymbol{\psi}_{k,L} \tag{11}$$

Thus, the derivative with respect to $\widehat{\boldsymbol{\theta}}_k$ is finally obtained by introducing (11) into (9). As a result, we find:

$$\widehat{\boldsymbol{\theta}}_{k+1} = \widehat{\boldsymbol{\theta}}_k - \mu_k(s_k - \hat{s}_k)^2 \hat{y}_k \left(\boldsymbol{\varphi}_{k,L} - 4\hat{\alpha}_k \boldsymbol{\psi}_{k,L} \right) \tag{12}$$

By following the same reasoning, the derivative with respect to $\hat{\alpha}_k$ which appears in relation (8) can also be easily expressed by:

$$\frac{\partial \hat{y}_k}{\partial \hat{\alpha}_k} = -\widehat{\boldsymbol{\theta}}_k^T \boldsymbol{\psi}_{k,L} \tag{13}$$

And we obtain the following equation:

$$\hat{\alpha}_{k+1} = \hat{\alpha}_k + \mu_k(s_k - \hat{s}_k)^2 \hat{y}_k \widehat{\boldsymbol{\theta}}_k^T \boldsymbol{\psi}_{k,L} \tag{14}$$

Algorithm 1 summarizes the main steps of the method described above.

Algorithm 1: Nonlinear LIMBO

LIMBO NL
Require: $\boldsymbol{u}, \boldsymbol{s}, L, N$
1:   $\widehat{\boldsymbol{\chi}}_1 \leftarrow [1 \quad 0 \quad \dots \quad 0]^T$
2:   $\widehat{\boldsymbol{\theta}}_1 \leftarrow \frac{\widehat{\boldsymbol{\chi}}_1}{\|\widehat{\boldsymbol{\chi}}_1\|_2}$
3:   $\hat{\eta}_1 \leftarrow 0$
4:   $\hat{\alpha}_1 \leftarrow 0$
5:   **for** $k = 1$ to N **do**
6:     $\boldsymbol{\varphi}_{k,L} \leftarrow (u_l)_{l=k}^{k-L+1}$
7:     $\boldsymbol{\psi}_{k,L} \leftarrow \left(\left(\widehat{\boldsymbol{\theta}}_k^T \boldsymbol{\varphi}_{l,L} \right)^3 \right)_{l=k}^{k-L+1}$
8:     $\hat{y}_k \leftarrow \widehat{\boldsymbol{\theta}}_k^T \boldsymbol{\varphi}_{k,L} - \hat{\alpha}_k \widehat{\boldsymbol{\theta}}_k^T \boldsymbol{\psi}_{k,L}$
9:     $\hat{s}_k \leftarrow S(\hat{y}_k)$
10:   $\widehat{\boldsymbol{\chi}}_{k+1} \leftarrow \widehat{\boldsymbol{\theta}}_k - \mu_k(s_k - \hat{s}_k)^2 \hat{y}_k \left(\boldsymbol{\varphi}_{k,L} - 4\hat{\alpha}_k \boldsymbol{\psi}_{k,L} \right)$
11:   $\widehat{\boldsymbol{\theta}}_{k+1} \leftarrow \frac{\widehat{\boldsymbol{\chi}}_{k+1}}{\|\widehat{\boldsymbol{\chi}}_{k+1}\|_2}$
12:   $\hat{\eta}_{k+1} \leftarrow \hat{\alpha}_k + \mu_k(s_k - \hat{s}_k)^2 \hat{y}_k \widehat{\boldsymbol{\theta}}_k^T \boldsymbol{\psi}_{k,L}$
13:   $\hat{\alpha}_{k+1} \leftarrow \hat{\eta}_{k+1} \|\widehat{\boldsymbol{\chi}}_{k+1}\|_2^3$
14: **end for**
15: **return** $\widehat{\boldsymbol{\theta}}_{k+1}, \hat{\alpha}_{k+1}$

In algorithm 1, the normalization step on line 11 ensures that the norm of $\hat{\boldsymbol{\theta}}_k$ remains constant and equal to 1. As a result, line 13 is added to maintain the homogeneity. We then guarantee the stability of our method (in the stability limits of the system). In counterpart, it is not possible to obtain any information about the amplitude of the impulse response. This identifiability problem has also been encountered in the linear case.

Finally, the full operating model is graphically illustrated in figure 3 below.
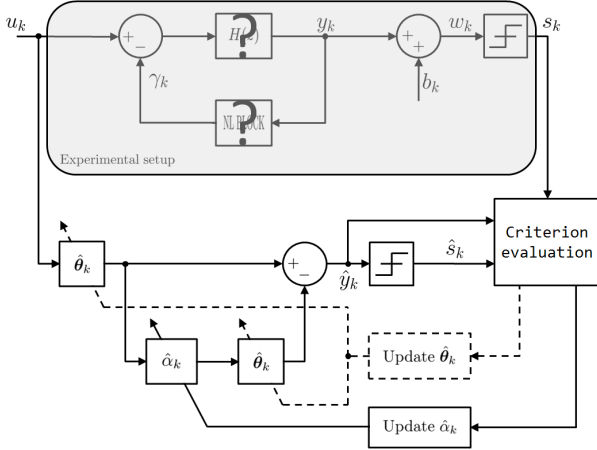
Figure 3: Block diagram of nonlinear LIMBO

4. RESULTS AND DISCUSSION

In this section, the results obtained with nonlinear LIMBO are compared in terms of convergence speed and estimation quality with those obtained by applying a typical LMS procedure. Let us underline that contrary to our approach, the standard LMS method is not based on quantized output measurements. The objective of this work is to compare the performances of our method with a widely-used one, which does not suffer from a lack of *a priori* information.

The input signal is a Gaussian white noise with zero mean and unit standard deviation. We consider an impulse response of length $L = 50$ and the Duffing coefficient is set to $\alpha = 0.01$. The identification procedure detailed in the previous section is applied starting from $N = 10^5$ observations of the binary output. The quality of the online estimation $\hat{\boldsymbol{\theta}}_k$ is defined as $1 - v_k$ where v_k is the cosine of the angle made by $\hat{\boldsymbol{\theta}}_k$ and $\boldsymbol{\theta}$. Since both vectors are normalized, we have $v_k = \boldsymbol{\theta}^T \hat{\boldsymbol{\theta}}_k$ and the following equivalence relation:

$$\lim_{k \to \infty} (1 - v_k) = 0 \Leftrightarrow \lim_{k \to \infty} v_k = 1$$

$$\Leftrightarrow \lim_{k \to \infty} \hat{\boldsymbol{\theta}}_k = \boldsymbol{\theta} \tag{15}$$

Concerning the impulse response, both methods present encouraging results in terms of estimation quality and convergence speed, in absence of measurement noise. Indeed, the fifty coefficients of the

column vector $\boldsymbol{\theta}$, *i.e.* the entire impulse response, have been rapidly and successfully estimated. Without surprise, Duffing coefficient identification also yields reasonable results for both methods in terms of estimation quality, but with a notable advantage for the nonlinear LMS approach in terms of convergence speed. This significant difference, which is an immediate consequence of quantized data, is shown in figure 4 below.
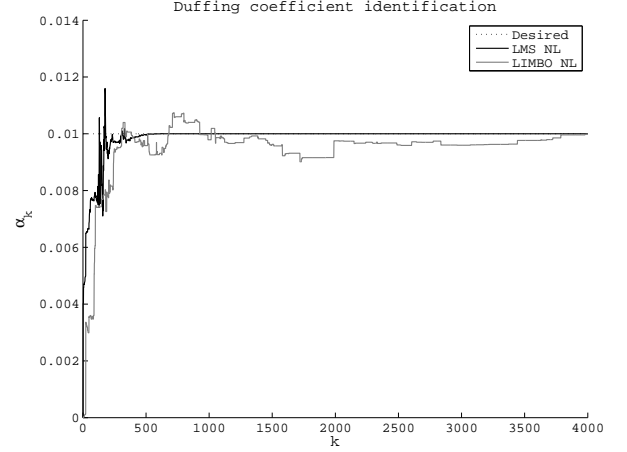
Figure 4: Comparison of nonlinear LMS and LIMBO methods for Duffing coefficient identification

The same behavior is distinctly observable in figure 5, which displays the quality of the online estimation. Indeed, LIMBO stops converging after having reached an error level approximately equal to 10^{-6}, while the LMS approach converges to the nominal parameters within the limits of finite machine precision.
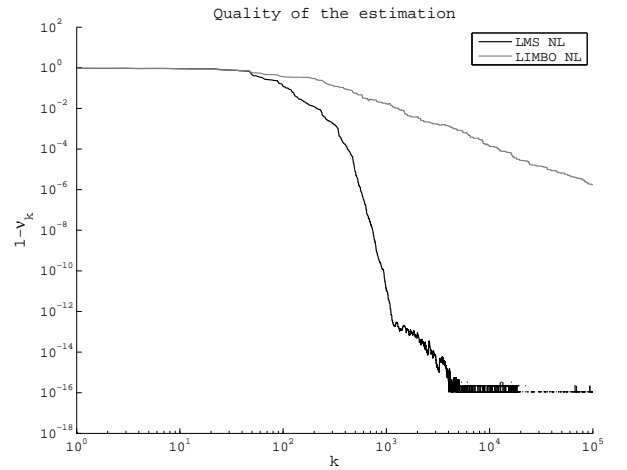
Figure 5: Comparison of nonlinear LMS and LIMBO methods in terms of convergence speed and estimation quality (SNR = ∞ dB)

In order to perturb the data, we consider an additive Gaussian noise such that the output signal-to-noise ratio (SNR) is equal to 20 dB. The estimation quality is graphically illustrated in figure 6 below. In this experiment, the nonlinear LMS (respectively,

LIMBO) approach stops converging after having reached an error level approximately equal to 10^{-4} (respectively, 10^{-3}). Although measurement noise has induced significant performance degradation, the estimation quality remains quite appreciable. Once again, the standard LMS method presents the best results in terms of convergence speed, but the gap is slightly reduced.
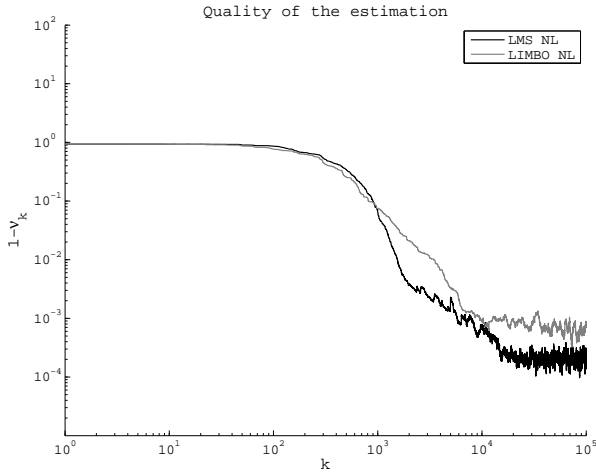
Figure 6: Comparison of nonlinear LMS and LIMBO methods in terms of convergence speed and estimation quality (SNR = 20 dB)

Finally, let us remember that in LIMBO, unknown parameters are updated only if $s_k - \hat{s}_k$ is not null, *i.e.* only if $s_k \neq \hat{s}_k$. This "change of sign" has appeared about 450 times in absence of noise, and about 3700 times with a SNR of 20 dB. Consequently, the LIMBO method has shown quite similar performances than the typical LMS method, especially in the case of perturbed data, and those with a lesser iteration number. However, contrary to the LMS approach, it is not possible to obtain any information concerning the amplitude of the impulse response, since θ is normalized in LIMBO algorithm.

5. CONCLUSION

In this paper, we have extended the LIMBO method, which has been introduced in (Jafari, Juillard, and Colinet 2010), in order to estimate online the parameters of a nonlinear system starting from binary observations. We have focused on the identification of a nonlinear Duffing oscillator that is coupled with a linear system characterized by a finite impulse response. Simulation results, in terms of convergence speed and estimation quality, have been truly admirable without measurement noise, and nearly similar to those obtained by applying a typical LMS procedure, which is not based on binary observations, in the noisy case. Consequently, nonlinear LIMBO is an inexpensive online test method that can be easily implemented on microfabricated devices, since it only requires the integration of a 1-bit ADC.

REFERENCES
Colinet, E., Juillard, J., 2010. A weighted least-squares approach to parameter estimation problems based on binary measurements. *IEEE Transactions on Automatic Control*, 55 (1), 148–152.

Jafari, K., Juillard, J., Colinet, E., 2010. A recursive system identification method based on binary measurements. *Proceedings of the 49th IEEE Conference on Decision and Control*, pp. 1154-1158. December 15-17, Atlanta (Georgia, USA).

Juillard, J., Jafari, K., Colinet, E., 2009. Asymptotic consistency of weighted least-square estimators for parameter estimation problems based on binary measurements. *Proceedings of the 15th IFAC Symposium on System Identification*, pp. 72-77. July 6-8, Saint-Malo (France).

Ljung, L., 1999. *System identification: theory for the user*. 2nd ed. Linkoping University: Prentice Hall PTR.

Negreiros, M., Carro, L., Susin, A., 2003. Ultimate low cost analog BIST. *Proceedings of the 40th annual Design Automation Conference*, pp. 570-573. June 2-6, New York (NY, USA).

Schoukens, J., Nemeth, J., Crama, P., Rolain, Y., Pintelon, R., 2003. Fast approximate identification of nonlinear systems. *Proceedings of the 13th IFAC Symposium on System Identification*, pp. 61-66. August 27-29, Rotterdam (Netherlands).

Walter, E., Pronzato, L., 1997. *Identification of parametric models from experimental data*. London: Springer.

Wigren, T., 1998. Adaptive filtering using quantized output measurements. *IEEE Transactions on Signal Processing*, 46 (12), 3423–3426.

Zhao, Y., Wang, L.Y., Yin, G.G., Zhang, J.-F., 2007. Identification of Wiener systems with binary-valued output observations. *Automatica*, 43 (10), 1752–1765.

AUTHORS BIOGRAPHY
Laurent Bourgois was born in Calais, France, in 1983. He received his M.Sc. degree in mathematical engineering and signal processing from the University of Littoral Côte d'Opale in 2006. He also completed his PhD in computer engineering, automatic control and signal processing from the same institution in 2010. From 2009 to 2010, he taught computer engineering and automatic control courses in parallel as a teaching assistant. In November 2010, he began his postdoctoral research in nonlinear dynamics and parameter estimation methods in the department of Signal Processing and Electronic Systems in Supelec.

Jérôme Juillard was born in Nice, France, in 1973. He graduated from Ecole Centrale Paris in 1995 and obtained his PhD in physical acoustics from University Paris 7 in 1999. He is now a professor in the department of Signal Processing and Electronic Systems in Supelec. He teaches numerical analysis and MEMS design and modeling courses. His research interests include NEMS/MEMS sensors, reduced-order modeling, nonlinear dynamics and parameter estimations methods.

REDUCING COMPUTATION TIME IN OPTIMIZATION PROCEDURE FOR FIXTURE LAYOUT BASED ON SWARMITFIX

Xiong Li[a], Matteo Zoppi[b], Rezia Molfino[c]

[a]DIMEC, University of Genoa
[b] DIMEC, University of Genoa
[c] DIMEC, University of Genoa

[a]lixiong@dimec.unige.it, [b]zoppi@dimec.unige.it, [c]molfino@dimec.unige.it

ABSTRACT
This paper addresses a method to reduce computation time of optimization procedure for fixture layout based on a developing project self-reconfigurable swarm intelligent fixture system (SwarmItFIX) funded by European Commission FP7. The SwarmItFIX combines flexibility, self-reconfigurability, automation, and swarm multi-agent cooperation. Based on a developed optimization procedure which combines genetic algorithm with finite element analysis, the proposed method for reducing its computation time includes the simplified finite element modeling method, dynamic mutation probability and tree-database for adjusting the optimization algorithm. Simulation is finally used to prove the efficiency of the method proposed.

Keywords: SwarmItFIX, flexible manufacture, genetic algorithm, computation time reduction

1. INTRODUCTION
Towards the current trend of life-cycle design, sustainability production, and geometrical complexity, the effect on manufacture equipment is gradually shifting towards more flexibility, reconfigurability, and automation. The same tendency affects the development of fixtures, the devices used for locating and clamping the workpieces. Aiming to such trend, an EU FP7 project--self-reconfigurable swarm intelligent fixture system (SwarmItFIX), target for thin-sheet workpiece, is being developed, Molfino, Zoppi and Zlatanov (2009).

During the development, there is a problem that how to define the support position since for thin-sheet workpiece, the deformation generated by the manufacturing operations and gravity cannot be neglect. An optimization procedure combining genetic algorithm with finite element analysis was developed. However, during the simulation, we found the computation cost is time-consuming. This paper is going to address the solution for reducing the computation time.

An, Choi, and Kim (2003) proposed a hybrid algorithm to reduce the computation time of genetic algorithm. Zulkarnain (2010) adopted the method of reducing computation steps to increase the conputation efficiency. Schonning, Nayfeha, Zarda (2003) utilized the dependency-tracking language to reduce computational time during multidisciplinary design optimization.

From practical composition of the optimization procedure, by absorbing previous researchers' achievements, this paper will reduce the computation time from three aspects: to simplify finite element modeling, to modify genetic algorithm parameters, and to construct tree-type database.

The structure of this paper is like that, after having introduced the characteristics and composition of SwarmItFIX, the optimization procedure combining genetic algorithm and finite element analysis for the purpose of obtaining the optimal fixture layout for flexible workpiece is presented. Based on such optimization procedure, the method to reduce the computation time is proposed. The simulation cases of fixture layout optimization demonstrate the efficiency of this method.

2. SELF-RECONFIGURABLE SWARM INTELLIGENT FIXTURE SYSTEM
As illustrated in Figure 1, the SwarmItFIX consists of a bench and several actuated fixtures (agents), which collaborate, without human interference, to support a thin-sheet workpiece. The bench provides a surface for the agents to move, and lock once they are at the desired location. It will also incorporate the power-supply and communication systems, as well as the means to measure the accurate position of the agents.
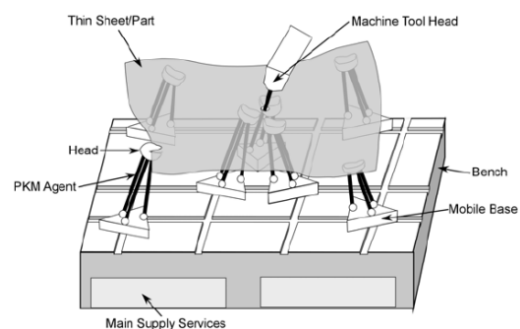
Figure 1: General Scheme Of The SwarmItFIX System

Each agent is composed of one mobile base, one parallel mechanism (PM), and one adjustable head. The mobile base supports the whole agent and hosts all the electronic parts for control. It provides the locomotion of the agent on the bench and communicates with the bench. The PM provides necessary workspace, support, and six degree of freedom for the head. The head directly contacts the sheet panel and has smart material with phase-change capabilities to conform to the local geometry of the workpiece. During the machining process, the fixtures will change positions on the bench, to form different supporting layouts as adapting to the process plan. The developed system towards prototype is shown in Figure 2.

Figure 2: The Developed SwarmItFIX

Beside its primary target application, the aerospace industry (fuselage sections, aerofoils, and other panels), in the future this system can be beneficial to other secondary sectors as shown in Figure 3.
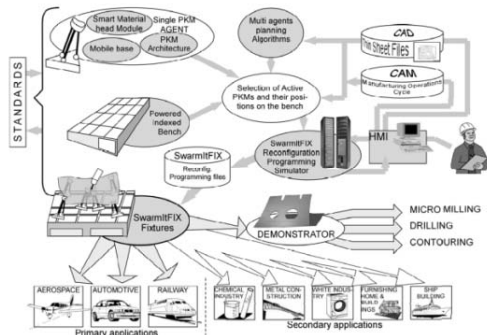
Figure 3: SwarmItFIX System Design and Development: Primary and Secondary Beneficial Sectors

3. OPTIMIZATION PROCEDURE

The optimization algorithm combines FE analysis to the use of a genetic algorithm. Genetic Algorithms (GA) based optimization is radically different from the search optimization methods, including traditional optimization methods and other stochastic methods (Renner, and Ekart 2003; Krishnakumar and Melkote 2000). No derivates is needed and it can escape local minima. Its inductive nature makes it do not need to know any rules of the problems and work by its own internal rule. Thus, GA is ideally suited for the fixture layout optimization problem since only the design

variables are used with no gradient or other auxiliary information (Goldberg 1989).

The fixture layout optimization procedure developed is shown in Figure 4.

In each generation of individuals generated by the GA, the individual (support position) is sent to ABAQUS to calculate the deformation. The values are passed to the genetic algorithm as fitness values until the final iteration condition is met. The final iteration condition is the iteration number reaching the upper limit of the generation number. All the code of genetic algorithm is implemented in MATLAB, and the finite element parameter modeling is implemented with Python, the program language of ABAQUS.
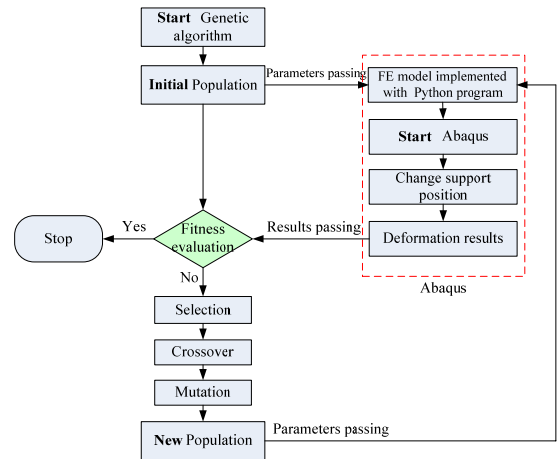
Figure 4: Optimization Procedure

The algorithm of the proposed optimization procedure is shown below.

Algorithm:
Step 1: Set the parameters of genetic algorithm
Step 2: Initialize the population, randomly generate individuals
Step 3: Convert the binary individuals into the support positions of fixtures
Step 4: Pass the converted data to FE program to change the support positions of fixtures
Step 5: Run FEA
Step 6: Return the deformation to GA program
Step 7: Use the deformation for fitness valuation
Step 8: If the fitness satisfy the convergence condition, go to Step 11.
Step 9: Follow the normal genetic algorithm procedure: selection, crossover and mutation
Step 10: Get the new population, go to Step 3
Step 11: Return the final solution and Stop

4. REDUCTION OF COMPUTATION TIME

As noticed on many simulation cases, this optimization procedure is time-consuming. There are three possibilities to reduce the computation time: to simplify FE modeling, to modify GA parameters, and to adjust the optimization procedure.

4.1. Simplification of Finite Elements Modeling

During the FE modeling, the interaction between head and workpiece needs to be set. Usually, such interaction is set as contact model, either point-to-point contact or point-to-surface contact (Satyanarayana and Melkote 2004). These contact models involve nonlinear computation with a significant increase in model complexity compared to linear modelling. However, the heads, as shown in Figure 5, apply to the workpiece a 6-DOF constrain and their stiffness is higher than the bending stiffness of the workpiece. Since heads never detach from the workpiece (what would require modeling with contacts), we can use the linear computation tie constraint of ABAQUS in place of contacts reducing remarkably the complexity of the FE model.
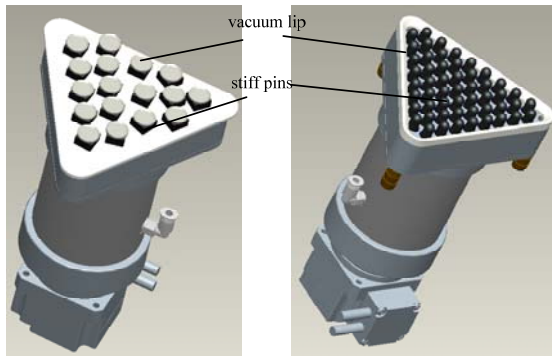
Figure 5: Two Kinds of the Developed Head

Simulation results derived from different typical cases are shown to check feasibility of such simplification as shown in Table 1. We can find that the deformation under contact model is average 7% larger than the one with tie constraint model. Therefore, such simplification is acceptable either from the point of view of practical application and computation results.

Table 1: Comparison of Simulation Results under Different Interactive Model

Workpiece dimension	Maximum deformation under different models	
	Under contact model	Under tie constraint
2D case 2m × 1m × 0.004m	2.475e-3m	2.323e-3m
3D case 0.6m × 0.7m Thickness 0.004m Average curvature 0.0011 mm^{-1}	3.25e-5m	3.12e-5m

One of the simulation cases is shown in Figure 6.
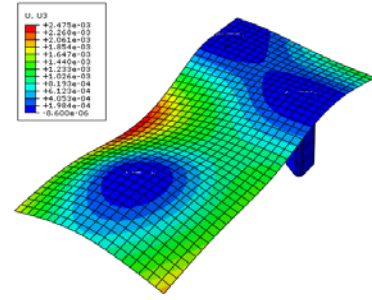
(a) Deformation under Contact Model
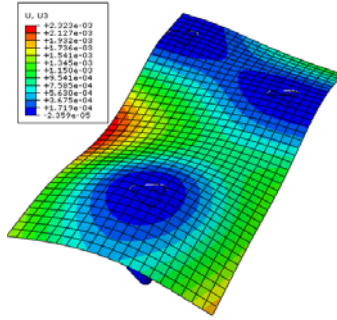
(b) Deformation under Tie Constraint

Figure 6: Optimization Results of 2D Cases under Different FE Modeling

4.2. Modification of Genetic Algorithm Parameters

From the genetic algorithm itself, dynamic mutation probability can be used for speeding up the convergence velocity and allowing large possibility to search global optimal solution.

A compromise has been found by trial and error between velocity of convergence and risk to fall in local minima. The experimental mutation rate obtained is:

$$p_m = p_m - \frac{j^4 - (p_m - \frac{p_c}{k})}{k^4} \qquad (1)$$

Where j is the generation number, k is the iteration number, p_m is the mutation probability and it changes with the generation number until the maximum number of iterations is reached. With the increase of the number of generation, the mutation probability gradually decreases. The relationship between the GA convergence and the variable mutation rate is also consistent with the relationship in (Beasley and Chu 1996).

4.3. Adjusting the Optimization Procedure

Since this optimization procedure is based on random generation, inevitably, there will be some individuals (support positions) identical to the individuals in the previous population pool. In this case, a tree-type database, as shown in Figure 7, can be constructed to reduce the repeated computation.
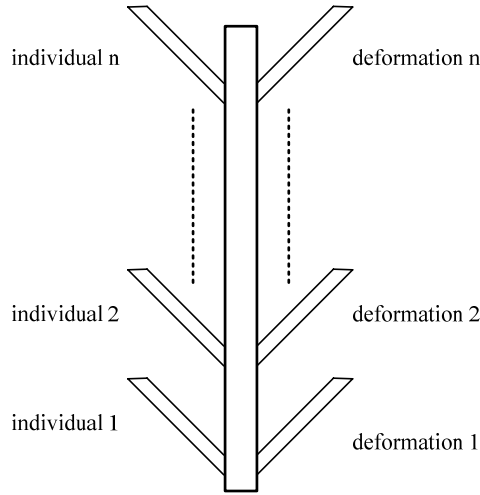
Figure 7: Tree-type Database

In this tree-type database, there is a rule of correspondence between the left branches and the right branches such that there is a deformation value assigned to each individual in the left branches. The "tree" will grow up with the increasing of iteration number. Therefore, before running FEA at the Step-5 of the algorithm, each individual will compare with the data in this tree-type database. If the individual is the same, there is no need to run FEA. It can obtain the deformation directly from the branches and thus save the computation time. Conversely, if there is no individual in the tree to match the individuals which need to be computed, FEA will be run and correspondence results will be saved in new branches. Such adjusting can save the computation time, obviously, especially during the convergence stage, since many individuals are the same with those in the previous population pool.

After all the above modification, the time-efficient optimization algorithm is illustrated as follows.

New Algorithm:
Step 1: Set the parameters of genetic algorithm and empty matrix for tree-database
Step 2: Initialize the population, randomly generate individuals
Step 3: Convert the binary individuals into the support positions of fixtures
Step 4: Pass the converted data to simplified FE program to change the support positions of fixtures
Step 5: Run FEA
Step 6: Return the deformation to GA program
Added Step 1: Keep the individuals and deformation in the tree-database
Step 7: Use the deformation for fitness valuation
Step 8: If the fitness satisfy the convergence condition, go to Step 11.
Step 9: Follow the normal genetic algorithm procedure: selection, crossover and dynamic mutation
Step 10: Get the new population, repeat the operation addressed in Step 3
Added Step 2: Compare the individuals with the individuals in the database. If some individuals is the

same with individuals in the database, get the corresponding deformation, go to Step 7.
Added Step 3: keep the remained individuals, go to Step 4
Step 11: Return the final solution and Stop

5. SIMULATION OF DIFFERENT CASES AND DISCUSSION

For the purpose to check the efficiency of the solution, different 2D and 3D cases are simulated. The hardware specification is: Intel Core(TM) i5 CPU 660@3.33GHz, 3.49GB of RAM. The computation was implemented in Matlab R2009 and ABAQUS V 6.10. The genetic algorithm parameters are set as listed in Table 2.

Table 2 Genetic Algorithm Parameters

Population size	Generation number	Selection probability	Mutation probability
20	50	0.8	0.8

For the comparison of the computation time, for each case considered only the optimization algorithm is different. All the other settings and conditions are the same. Four typical cases are simulated and each of their simulation runs 10 times. Their average computation time is recorded in Table 3.

Table 3 Simulation Cases

Case type	Workpiece dimension	Fixture number	Computation time (hour)	
			Previous algorithm	New algorithm
2D	2m × 1m × 0.004m	3	175h	20h
2D	0.85m × 1.5m × 0.004m	3	160h	18h
3D	1m × 0.9m Thickness 0.004m Average curvature 0.0015 mm^{-1}	3	150h	14.5h
3D	0.6m × 0.7m Thickness 0.004m Average curvature 0.0011 mm^{-1}	2	90h	10h

As shown in Table 3, the new algorithm can reduce the computation time sharply. One of the simulation case results is shown in Figure 8.
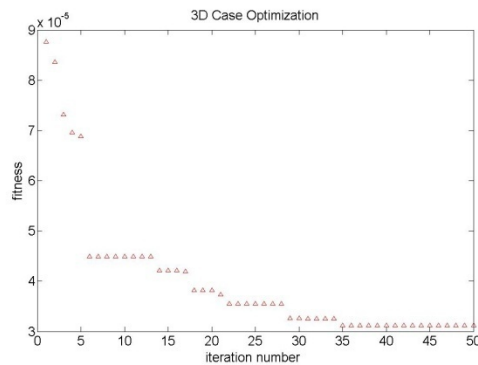
(a) Fitness Results at Each Generation
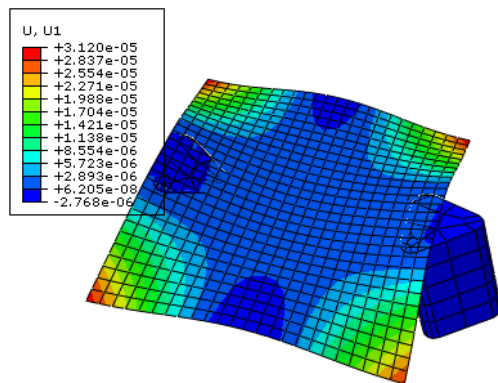
(b) Final Fitness Result Shown in FEM
Figure 12: Optimal SimulationResults of 3D Workpiece

6. CONCLUSION

A developing project about self-reconfigurable swarm intelligent fixture system which combines flexibility, self-reconfigurability, automation, and swarm multi-agent cooperation is addressed. The reduction in computation time for the development of the optimization procedure for optimal fixture layout is proposed. Such solution can be implemented in practice and has both academic and application significance.

ACKNOWLEDGMENTS

REFERENCES

An, J.W. and Choi, J.K. and Kim, N. 2003, A hybrid algorithm to reduce the computation time of genetic algorithm for designing binary phase holograms, *Journal of the optical society of Korea*, Vol.7, No. 4, pp. 264-268

Beasley J.E., Chu P.C. 1996, A genetic algorithm for the set covering problem, *European Journal of Operational Research* 94, pp. 392-404

Goldberg D.E. 1989, *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison-Wesley, Reading, MA

Krishnakumar K., Melkote S.N. 2000, Machining fixture layout optimization, *International Journal of Machine Tools and Manufacture,* 40, pp. 579–598

Molfino R., Zoppi M., Zlatanov D. 2009. Reconfigurable swarm fixtures. *ASME/IFToM International Conference on Reconfigurable Mechanisms and Robots*. Proceedings, June 22–24, 2009, pp. 730–735

Renner G., Ekart A. 2003, Genetic algorithms in computer aided design, *Computer Aided Design*, 35, pp. 709–726

Satyanarayana S., Melkote S.N. 2004, Finite element modeling of fixture-workpiece contacts: single contact modeling and experimental verification, *International Journal of Machine Tools & Manufacture* 44, pp. 903-913

Schonning M.A., Nayfeha J.F., Zarda P.R.. 2003, Utilization of a dependency-tracking language to reduce computational time during multidisciplinary design optimization, *Advances in Engineering Software*, 34, pp. 115–122

Zulkarnain Md. Ali., 2010, Reduce computation steps can increase the efficiency of computation algorithm, *Journal of Computer Science*, 6(10), 1023-1027

AUTHORS BIOGRAPHY
Xiong Li : PhD candidate in DIMEC, University of Genoa

Matteo Zoppi : PhD researcher at University of Genoa. Mechanisms and robots creative design. Involved in many European Projects and Networks (e.g. PICAV, SwarmItFIX, EURON, CLAWAR). ASME member

Rezia Molfino : Full professor at University of Genoa, responsible of PMAR lab Robotics. President of SIRI (Italian Association of Robotics and Automation), National Coordinator at IFR (International Federation of Robotics). Appointed reviewer for international journals in robotics and automation. Author and co-author of about 250 scientific papers. About 20 patents on robotic devices and systems.

ROBUST TRAJECTORY TRACKING PROBLEM: CHEAP CONTROL APPROACH

Vladimir Turetsky [(a)], Valery Y. Glizer [(b)]

[(a)] Department of Mathematics, Ort Braude College, P.O.B. 78, Karmiel 21982, Israel
[(b)] Department of Mathematics, Ort Braude College, P.O.B. 78, Karmiel 21982, Israel

[(a)] E-mail:turetsky1@braude.ac.il, [(b)]E-mail: valery48@braude.ac.il

ABSTRACT

A robust trajectory tracking problem is treated in the framework of a zero-sum linear-quadratic differential game of a general type. For a cheap control version of this game, a novel solvability condition is derived. Condition, guaranteeing that the tracking problem is solved by a cheap control game optimal strategy, is established. A boundedness of the minimizer's control is analyzed. Illustrative pursuit-evasion examples are presented.

Keywords: Trajectory tracking, robust control, linear-quadratic differential game, cheap control

1. INTRODUCTION

The problem of tracking a given trajectory under uncertainties (trajectory planning, path following etc.) is a well-known challenge in aerospace (Ben-Asher et al., 2004; Zhang et al., 2008), underwater vehicles control (Aguiar & Hespanha, 2007; Kiselev, 2009), robotics (Wang et al., 2009; Kowalczyk et al., 2009)and many other applications. Most of the approaches, known in the literature, provide the trajectory tracking asymptotically for time tending to infinity (see e.g. (Aguiar & Hespanha, 2007; Sun et al., 2009; Cheng et al., 2007)). In many practical applications, for example, in aerospace, the tracking should be guaranteed on a finite time interval. In real-life systems, the trajectory should be tracked in the presence of uncertainty and/or disturbance. However, to the best knowledge of the authors, only a small number of the papers considers the tracking problem from this viewpoint (see e.g. (Pei et al., 2003; Mahony & Hamel, 2004)). In (Basar & Bernhard, 1995; Tretyakov & Turetsky, 1995; Ben-Asher et al., 2004), the tracking problem is formulated on a finite time horizon in the framework of a differential game. In this paper, based on such a formulation, the tracking problem is solved by using a cheap control linear-quadratic approach.

Condition for solvability of the linear-quadratic differential game (LQDG) was first formulated by Bernhard (1979; 1980) as lack of conjugate points on a game time interval. This condition, although being necessary and sufficient, cannot be verified directly.

Thus, for the LQDG, the conditions, guaranteeing the lack of the conjugate points, are of a great importance. A number of works, dealing with this issue, can be mentioned. In (Basar & Bernhard, 1995), it was established that the game solution exists if the maximizer's control cost in the performance index is sufficiently large. In (Reid, 1972), the game solvability follows from the invertibility of the solution of some matrix linear differential equation. Due to (Mou & Liberty, 2001), the game solvability is provided by the existence of so-called lower and upper solutions of the matrix Riccati differential equation, associated with the game. These conditions do not provide a direct verification scheme based on the dynamics and cost functional coefficients. This drawback was partially surmounted by the condition, formulated in (Shinar *et al.*, 2008) in terms of eigenvalues of some integral operator in a Hilbert space. This condition can be directly verified, based on the dynamics and cost functional coefficients.

The cheap control problem is an optimal control problem (differential game) with a small control cost (with respect to a state cost) in the cost functional. This problem is of considerable meaning in such topics of control theory as singular optimal control and its regularization (Bell & Jacobson, 1975), limitations of linear optimal regulators and filters (Braslavsky *et al.*, 1999; Kwakernaak & Sivan, 1972), limitations of nonlinear optimal regulators (Seron *et al.*, 1999), high gain control (Kokotovic *et al.*, 1986; Young *et al.*, 1977), inverse control problems (Moylan & Anderson, 1973), guidance problems (Cottrell, 1976; Zarchan, 1994), robust control of systems with disturbances (Turetsky & Glizer, 2004; Turetsky & Glizer, 2007), and some others. Cheap control problems have been investigated extensively for systems with a single decision maker (see e.g. (Kokotovic, 1984) and references therein). More recent results can be found in (Woodyatt *et al.*, 2002; Glizer *et al.*, 2007) and references therein. Cheap controls for differential games have been investigated much less. To the best knowledge of the authors, there are only few works where differential games with cheap control have been studied. In (Starr & Ho, 1969; Glizer, 2000; Turetsky & Glizer, 2004; Turetsky & Glizer, 2007), a finite-horizon game was investigated, while in (Petersen, 1986; Glizer,

2009), an infinite-horizon case was analyzed. In all these works, excepting (Turetsky & Glizer, 2007) the case of the minimizer's cheap control was treated, while in (Turetsky & Glizer, 2007), both the minimizer's and the maximizer's controls were assumed to be cheap.

In the present paper, the general tracking problem is considered. In this problem, a tracking criterion is chosen as a Lebesgue-Stilties integral G of squared discrepancy between the system motion and a given vector function (tracked trajectory), calculated over the mixed discrete-continuous measure. The problem is solved by using an auxiliary LQDG, where the state term of the cost functional is represented by G. Both the minimizer's and the maximizer's controls are cheap. Note that this game is a cheap control version of the LQDG, considered in (Shinar et al., 2008). Novel, cheap control, solvability condition is established. It is shown that, subject to some additional conditions, the optimal cheap control strategy also solves the original tracking problem.

2. TRACKING PROBLEM

2.1. Motivating Guidance Example

Consider a planar engagement between two moving objects (players) - a pursuer and an evader. The schematic view of this engagement is shown in Fig. 1. The X axis of the coordinate system is aligned with the initial line of sight. The origin is collocated with the initial pursuer position. The points (x_p, y_p), (x_e, y_e) are the current coordinates; V_p and V_e are the velocities and a_p, a_e are the lateral accelerations of the pursuer and the evader respectively; φ_p, φ_e are the respective angles between the velocity vectors and the reference line of sight; and $y = y_e - y_p$ is the relative separation normal to the initial line of sight. The line-of-sight angle λ is the angle between the current and initial lines of sight, r is the current range between the objects.
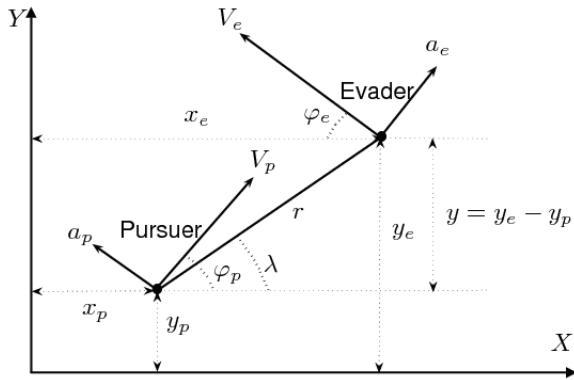
Figure 1: Interception geometry

It is assumed that the dynamics of each object is expressed by a first-order transfer function with the time constants τ_p and τ_e, respectively. The velocities V_p, V_e and the bounds of the lateral acceleration commands a_p^{max}, a_e^{max} of both objects are constant.

If the aspect angles φ_p and φ_e are small during the engagement then (Shinar, 1981) the trajectories of the pursuer and the evader can be linearized with respect to the nominal collision geometry. The final interception time can be easily calculated as $t_f = r_0/(V_p + V_e)$, where r_0 is the initial distance between the objects and the initial time $t_0 = 0$. The linearized model is described by the following differential equation:

$$\dot{x} = Ax + bu + cv, \quad x(0) = x_0, \qquad (1)$$

where the state vector is
$x = (x_1, x_2, x_3, x_4)^T = (y, \dot{y}, a_e, a_p)^T$, the superscript T denotes the transposition,

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & -1/\tau_e & 0 \\ 0 & 0 & 0 & -1/\tau_p \end{bmatrix},$$

$$b = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1/\tau_p \end{bmatrix}, \quad c = \begin{bmatrix} 0 \\ 0 \\ 1/\tau_e \\ 0 \end{bmatrix}, \qquad (2)$$

$$x_0 = (0, x_{20}, 0, 0)^T, \quad x_{20} = V_e \varphi_e(0) - V_p \varphi_p(0). \qquad (3)$$

The controls of the pursuer u and the evader v are the respective lateral acceleration commands.

In this engagement, the pursuer can have different objectives. For example, (1) to intercept the evader, (2) to intercept it with zero relative velocity (rendezvous), (3) to reach some prescribed points during the engagement, (4) to track a prescribed relative separation profile, etc. These objectives are expressed by the following cost functionals:

$$J_1 = x_1^2(t_f), \quad J_2 = x_1^2(t_f) + x_2^2(t_f), \qquad (4)$$

$$J_3 = \sum_{i=1}^{K} (x_1(t_i) - x_{1i})^2, \quad J_4 = \int_0^{t_f} (x_1(t) - y(t))^2 dt, \qquad (5)$$

where (t_i, x_{1i}), $i = 1, \ldots, K$, are given points on the plane (t, x_1); $y(t)$ is the prescribed function. Thus,

the pursuer's objective becomes to guarantee a small value for the quadratic functionals (4) - (5). In the next section, based on these functionals, a general tracking problem is formulated.

2.2. Robust Tracking Problem Formulation

Consider a controlled system

$$\dot{x} = A(t)x + B(t)u + C(t)v, \; x(t_0) = x_0, \; t_0 \leq t \leq t_f, \tag{6}$$

where $x \in R^n$ is the state vector, $u \in R^r$ and $v \in R^s$ are the control and the disturbance, respectively, t_0 and t_f are prescribed initial and final time instants; the matrices $A(t)$, $B(t)$ and $C(t)$ are continuous.

Let $t_i \in (t_0, t_f]$, $i = 1, \ldots, K$, and $(a_j, b_j) \subset [t_0, t_f]$, $j = 1, \ldots, L$, be prescribed time instants and non-intersecting intervals, such that at least one of the conditions $t_K = t_f$, $b_L = t_f$, is satisfied. Let $y(t)$ and $D(t)$, $t \in [t_0, t_f]$, be vector and matrix functions of the dimensions n and $n \times n$, respectively, continuous on each interval $[a_j, b_j]$, $j = 1, \ldots, L$. Let define the cost functional

$$J = G(x(\cdot)) = \sum_{i=1}^{K} \left| D(t_i)(x(t_i) - y(t_i)) \right|^2 +$$
$$\sum_{j=1}^{L} \int_{a_j}^{b_j} \left| D(t)(x(t) - y(t)) \right|^2 dt, \tag{7}$$

where $|\cdot|$ is the Euclidean norm of the vector. Note that the first sum in the right-hand part of (7) is the sum of the intermediary costs, mentioned by Bernhard Bernhard. The functional (7) can be written as a Lebesgue-Stilties integral

$$G(x(\cdot)) = \int_{[t_0, t_f]} \left| D(t)(x(t) - y(t)) \right|^2 dm(t), \tag{8}$$

where the bounded variation function $m(t)$ has the following structure. Let $T = \bigcup_{j=1}^{L}(a_j, b_j)$ and $\zeta(t)$ be an indicator function of T: $\zeta(t) = 1$ for $t \in T$ and $\zeta(t) = 0$ for $t \notin T$. Let $\chi([a, b])$ be the number of the values $t_i \in [a, b]$. Then,

$$m(t) = \int_{t_0}^{t} \zeta(s)ds + \chi([t_0, t]). \tag{9}$$

It is assumed that the disturbance satisfies the integral constraint

$$\int_{t_0}^{t_f} v^2(t)dt \leq \nu, \; t \in [t_0, t_f). \tag{10}$$

Problem. For a given $\nu > 0$ and for any given $\zeta > 0$, to construct a feedback strategy $u_\zeta(t, x)$ such that the inequality

$$J < \zeta, \tag{11}$$

is satisfied for any admissible disturbance $v(t)$.

This *robust tracking problem* is treated by embedding it into an auxiliary linear-quadratic differential game.

3. LINEAR-QUADRATIC DIFFERENTIAL GAME

3.1. Game Formulation

For (6), let formulate a differential game with the cost functional

$$J_{\alpha\beta} = J_{\alpha\beta}(u(\cdot), v(\cdot)) = G(x(\cdot)) +$$
$$\alpha \int_{t_0}^{t_f} |u(t)|^2 \, dt - \beta \int_{t_0}^{t_f} |v(t)|^2 \, dt, \tag{12}$$

where the functional G is given by (7), the constants α, β are positive.

The objective of the first player ($u(t)$) is to minimize (12), while the second player ($v(t)$) maximizes it, by using feedback strategies $u(t, x)$ and $v(t, x)$, respectively. These strategies are defined for $t \in [t_0, t_f]$, $x \in R^n$. Due to (Krasovskii & Subbotin, 1988) it is assumed that the functions $u(t, x)$ and $v(t, x)$ are measurable w.r.t. t for each fixed x and satisfy the Lipschitz condition w.r.t. x uniformly w.r.t. t. Moreover, for $u = u(t, x)$ and any $v(\cdot) \in L_2^s[t_0, t_f]$, the initial value problem (6) has a solution on the entire interval $[t_0, t_f]$, where $L_2^m[t_0, t_f]$ denotes the space of square-integrable functions $f(\cdot) : [t_0, t_f] \rightarrow R^m$. Similarly, for $v = v(t, x)$ and any $u(\cdot) \in L_2^r[t_0, t_f]$, the initial value problem (6) has a solution on the entire interval $[t_0, t_f]$. In the sequel, the sets Y and ς of all such functions $u(t, x)$ and $v(t, x)$ are called the sets of admissible feedback strategies of the minimizer and the

maximizer, respectively. For a given $u(\cdot) \in Y$, the value

$$J_u(u(\cdot); t_0, x_0) = \sup_{v(t) \in L_2^S[t_0, t_f]} J_{\alpha\beta}, \qquad (13)$$

is called the *guaranteed result of* $u(\cdot)$. The strategy $u^0(\cdot)$ is called *optimal* if

$$J_u(u^0(\cdot); t_0, x_0) = \min_{u(\cdot) \in Y} J_u(u(\cdot); t_0, x_0) \cong J_u^0(t_0, x_0). \qquad (14)$$

Similarly the *guaranteed result of the strategy* $v(\cdot) \in \varsigma$ is

$$J_v(v(\cdot); t_0, x_0) = \inf_{u(t) \in L_2^r[t_0, t_f]} J_{\alpha\beta}. \qquad (15)$$

The *optimal* strategy $v^0(\cdot)$ is defined by

$$J_v(v^0(\cdot); t_0, x_0) = \max_{v(\cdot) \in \varsigma} J_v(v(\cdot); t_0, x_0) \cong J_v^0(t_0, x_0). \qquad (16)$$

If

$$J_u^0(t_0, x_0) = J_v^0(t_0, x_0) \cong J^0(t_0, x_0), \qquad (17)$$

then $J^0(t_0, x_0)$ is called the LQDG *value*, and the pair of optimal strategies $(u^0(\cdot), v^0(\cdot))$ is called the LQDG *saddle point*. In this case, the LQDG is called *solvable*, and the triplet $\{J^0(\cdot), u^0(\cdot), v^0(\cdot)\}$ constitutes its solution.

In the sequel, this LQDG is analyzed for the cheap control case

$$\alpha \to 0, \quad \beta \to 0, \quad \alpha/\beta = \mu = const. \qquad (18)$$

3.2. Cheap Control Solvability Condition

It can be directly shown that the solvability condition, established by Shinar et al (2008), is not suitable for a cheap control case (18). In this section, a novel condition is formulated.

For a given $t \in [t_0, t_f]$, the function $m(\cdot)$ generates (see (Balakrishnan, 1976)) the Hilbert space $L_2^n([t, t_f], m)$ of the vector functions $f(\cdot): [t, t_f] \to R^n$, square-integrable over $m(\cdot)$ on the interval $[t, t_f]$. The inner product in this space is defined as

$$\langle f_1(\cdot), f_2(\cdot) \rangle_t = \int_{[t, t_f]} f_1^T(\eta) f_2(\eta) dm(\eta), \qquad (19)$$

yielding the norm

$$\| f(\cdot) \|_t = \langle f(\cdot), f(\cdot) \rangle_t^{1/2} = \left(\int_{[t, t_f]} |f(\eta)|^2 dm(\eta) \right)^{1/2}. \qquad (20)$$

Let introduce the parametric family of the operators $\Phi_{\alpha\mu}(t)$, $t \in [t_0, t_f]$, mapping $L_2^n([t, t_f], m)$ into itself:

$$\Phi_{\alpha\mu}(t) = \frac{1}{\alpha} \Phi_\mu(t), \quad \Phi_\mu(t) \cong \mu \Phi_v(t) - \Phi_u(t), \qquad (21)$$

where

$$\Phi_k(t) f(\cdot) = \int_{[t, t_f]} F_k(t, \eta, \nu) f(\nu) dm(\nu), \eta \in [t, t_f], \ k = u, v, \qquad (22)$$

$$F_k(t, \eta, \nu) = D(\eta) X(\eta, t_f) \left(\int_t^{\min(\eta, \nu)} Q_k(\tau) d\tau \right) X^T(\nu, t_f) D^T(\nu), \quad k = u, v, \qquad (23)$$

$$Q_u(t) = X(t_f, t) B(t) B^T(t) X^T(t_f, t), \qquad (24)$$
$$Q_v(t) = X(t_f, t) C(t) C^T(t) X^T(t_f, t),$$

and $X(t, \tau)$ is the fundamental matrix of the homogenous equation $\dot{x} = A(t)x$. Also denote

$$Q_{\alpha\mu}(\tau) = \frac{1}{\alpha} (\mu Q_v(\tau) - Q_u(\tau)). \qquad (25)$$

Due to (Shinar *et al.*, 2008), for any $t \in [t_0, t_f]$, the operators $\Phi_u(t)$ and $\Phi_v(t)$ are self-adjoint, positive and compact.

Let introduce the matrix, vector and scalar functions $R_{\alpha\mu}(t)$, $r_{\alpha\mu}(t)$ and $\rho_{\alpha\mu}(t)$, satisfying the following impulsive differential equations

$$\frac{dR}{dt} = -R Q_{\alpha\mu}(t) R - \zeta(t) S(t). \qquad (26)$$

$$R(t_f + 0) = 0, \ R(t_i + 0) - R(t_i) = -S(t_i), \qquad (27)$$

$$\frac{dr}{dt} =$$

$$- R(t)Q_{\alpha\mu}(t)r + 2\zeta(t)X^T(t,t_f)D^T(t)y(t), \quad (28)$$

$$r(t_f + 0) = 0,$$

$$r(t_i + 0) - r(t_i) = \quad (29)$$

$$2X^T(t_i, t_f)D^T(t_i)D(t_i)y(t_i),$$

$$\frac{d\rho}{dt} = -\frac{1}{4}r^T(t)Q_{\alpha\mu}(t)r(t) - \quad (30)$$

$$\zeta(t)y^T(t)D^T(t)D(t)y(t),$$

$$\rho(t_f + 0) = 0,$$

$$\rho(t_i + 0) - \rho(t_i) = \quad (31)$$

$$-y^T(t_i)D^T(t_i)D(t_i)y(t_i),$$

where $i = 1, \ldots, K$; $\zeta(t)$ is the indicator function of the set T; $S(t) = X^T(t, t_f)D^T(t)D(t)X(t, t_f)$, $Q_{\alpha\mu}(t)$ is given by (25). Note that this impulsive system consists of the Riccati matrix differential equation (26), the linear vector differential equation (28) and the trivial scalar differential equation (30).

Let $\lambda_{\mu i}(t)$, $i = 1, 2, \ldots$, be the eigenvalues of the operator $\Phi_\mu(t)$.

Theorem 1. *Let for fixed $\mu > 0$,*

$$\sup_{t \in [t_0, t_f)} \sup_i \lambda_{\mu i}(t) \leq 0. \quad (32)$$

Then, for an arbitrary small $\alpha > 0$ and $\beta = \alpha/\mu$, the LQDG (6) -- (12) is solvable. For any position $(t, x) \in [t_0, t_f] \times R^n$, the value and the saddle point of the LQDG are given by

$$J^0(t, x) = J^0_{\alpha\mu}(t, x) =$$

$$x^T X^T(t_f, t)R_{\alpha\mu}(t)X(t_f, t)x + \quad (33)$$

$$r^T_{\alpha\mu}(t)X(t_f, t)x + \rho_{\alpha\mu}(t),$$

$$u^0(t, x) = u^0_{\alpha\mu}(t, x) = -\frac{1}{2\alpha}B^T(t)l^0_{\alpha\mu}(t, x), \quad (34)$$

$$v^0(t, x) = v^0_{\alpha\mu}(t, x) = \frac{1}{2\beta}C^T(t)l^0_{\alpha\mu}(t, x), \quad (35)$$

where

$$l^0_{\alpha\mu}(t, x) \square X^T(t_f, t)(2R_{\alpha\mu}(t)X(t_f, t)x \quad (36)$$

$$+ r_{\alpha\mu}(t)),$$

$R_{\alpha\mu}(t)$, $r_{\alpha\mu}(t)$ and $\rho_{\alpha\mu}(t)$ satisfy (26) - (31).

4. TRACKING PROBLEM SOLUTION

4.1. Tracking by LQDG Cheap Control

In this section, it is shown how the optimal minimizer strategy $u^0_{\alpha\mu}(\cdot)$ can be used for tracking a given trajectory $y(t)$. In this case, the opponent control $v(t)$ is considered as an unknown disturbance from $L^s_2[t_0, t_f]$. The tracking accuracy is evaluated by the functional (7).

Let $x_{\alpha\mu}(t)$ denote the solution of (6) for $u = u^0_{\alpha\mu}(t, x)$ and $v = v(t)$, and $u_{\alpha\mu}(t) \cong u^0_{\alpha\mu}(t, x_{\alpha\mu}(t))$. Let introduce the operator $\Phi_{\mu 0} = -\Phi_\mu(t_0)$, and the function

$$w_0(t) = D(t)(X(t, t_0)x_0 - y(t)), \quad t \in [t_0, t_f]. \quad (37)$$

Let the eigenvalues and the eigenfunctions of $\Phi_{\mu 0}$ be λ_k and $f_k(t)$, $k = 1, 2, \ldots$, $t \in [t_0, t_f]$, and the function $w_0(t)$ be represented by the series

$$w_0(t) = \sum_{k=1}^{\infty} w_k f_k(t). \quad (38)$$

Theorem 2. *Let the inequality (32) hold. If $\lambda_k > 0$ for all $k = 1, 2, \ldots$, and*

$$\sum_{k=1}^{\infty} \frac{w_k^2}{\lambda_i} < \infty, \quad (39)$$

then for any $v(\cdot) \in L^s_2[t_0, t_f]$,

$$\lim_{\alpha \to 0, \beta = \alpha/\mu} G(x_{\alpha\mu}(\cdot)) = 0. \quad (40)$$

Corollary 1. *Let the conditions of Theorem 2 hold and $v(\cdot) \in L^s_2[t_0, t_f]$ satisfies the constraint (10). Then for any $\zeta > 0$ there exists $\alpha = \alpha(\zeta, v)$ such that*

$$G(x_{\alpha\mu}(\cdot)) < \zeta. \quad (41)$$

This corollary means that, subject to the conditions of Theorem 2, the LQDG optimal strategy $u^0_{\alpha(\zeta, v)\mu}(t, x)$ for sufficiently small α solves the robust tracking problem.

4.2. Control Boundedness

In the previous section, it has been shown that, subject to conditions of Theorem 2, the optimal minimizer strategy $u^0_{\alpha\mu}(t, x)$ solves the tracking problem in the sense (40). However, the corresponding time realization

$u_{\alpha\mu}(t)$ can become unbounded. In some real-life problems, such an unboundedness is inconvenient and even unacceptable. Therefore, deriving the conditions, guaranteeing the boundedness of $u_{\alpha\mu}(t)$, is of a considerable importance.

It can be shown, that if the conditions of Theorem 2 are valid, the control time realization is bounded in the sense of $L_2^r[t_0, t_f]$. Next theorem establishes necessary conditions of boundedness in the sense of $C[t_0, t_f]$.

Let $u_\alpha(t, x)$ be a family of admissible minimizer's feedback strategies, where α is a positive parameter, $t \in [t_0, t_f]$, $x \in R^n$. Let $x_\alpha(t)$ denote the solution of (6) for $u = u_\alpha(t, x)$ and arbitrary but fixed $v(\cdot) \in L_2^s[t_0, t_f]$.

Theorem 3. *Let for any* $v(\cdot) \in L_2^s[t_0, t_f]$,

$$\lim_{\alpha \to 0} G(x_\alpha(\cdot)) = 0. \tag{42}$$

If, subject to (42), the control time realization $u_\alpha(t) = u_\alpha(t, x_\alpha(t))$ *is bounded, i.e. there exists* $C > 0$ *such that* $|u_\alpha(t)| \le C$, *for all* $t \in [t_0, t_f]$ *and sufficiently small* $\alpha > 0$ *then: either*

(A)
$$a_1 > t_0, \tag{43}$$
or

(B) $a_1 = t_0$ *and*
$$x_0 - y(t_0) \in \mathrm{Ker}D(t_0). \tag{44}$$

Theorem 3 implies that if the condition
$$a_1 = t_0 \text{ and } x_0 - y(t_0) \notin \mathrm{Ker}D(t_0), \tag{45}$$
is valid and the optimal minimizer strategy $u_{\alpha\mu}^0(\cdot)$ solves the tracking problem, then its time realization $u_{\alpha\mu}(t)$ is necessarily unbounded.

5. EXAMPLES

5.1. Scalar Illustrative Example

Consider the LQDG for the scalar system

$$\dot{x} = u + v, \tag{46}$$
with the pure integral cost functional

$$J_{\alpha\beta} = \int_0^{t_f} (x(t) - y(t))^2\, dt +$$
$$\alpha \int_0^{t_f} u^2(t)dt - \beta \int_0^{t_f} v^2(t)dt, \tag{47}$$

for which $t_0 = 0$, $A(t) \equiv 0$, $B(t) = C(t) = D(t) = X(t, t_0) \equiv 1$, $K = 0$, $L = 1$, $a_1 = t_0 = 0$, $b_1 = t_f$, yielding $Q_u(t) = Q_v(t) \equiv 1$.

By simple algebra, due to (21) -- (22), this implies that

$$\Phi_{\mu 0} f(\cdot) = (1 - \mu)\int_0^{\eta t_f}\int_v f(\xi)d\xi)dv. \tag{48}$$

Assume $\mu < 1$, which guarantees that this operator is positive. Thus, the Sturm-Liouville problem for the operator (48) can be transformed to the boundary value problem

$$\ddot{g} + \gamma^2 g = 0, \ g(0) = 0, \ g(t_f) = 0, \tag{49}$$

where $g(\eta) \cong \int_0^{\eta t_f}\int_v f(\xi)d\xi)dv$, $\gamma^2 \cong (1 - \mu)/\lambda$, λ is the eigenvalue of (48). From (49), the eigenvalues and the eigenfunctions of (48) are

$$\lambda_k = \frac{(1 - \mu)t_f^2}{(\pi/2 + \pi k)^2} > 0,$$
$$f_k(t) = \frac{1 - \mu}{\lambda_k}\sin\sqrt{\frac{1 - \mu}{\lambda_k}}t, \ t \in [0, t_f], \tag{50}$$

$k = 1, 2, \ldots.$

The eigenfunctions $f_k(t)$ satisfy
$$\langle f_k(\cdot), f_m(\cdot)\rangle_{t_0} = 0, \ k \ne m;$$
$$\| f_k(\cdot)\|_{t_0}^2 = \frac{(2k + 1)^4\pi^4}{32 t_f^3}. \tag{51}$$

Let the tracked function be $y(t) = t$, yielding $w_0(t) = x_0 - t$. The coefficients w_k of the series (38) are calculated as

$$w_k = \frac{1}{\| f_k(\cdot)\|_{t_0}^2}\int_0^{t_f}(t_0 - \xi)f_k(\xi)d\xi =$$
$$\frac{16(\pi x_0(2k + 1) - 2(-1)^k t_f)t_f^2}{(2k + 1)^4\pi^4}, \tag{52}$$

implying by (50) that

$$w_k^2 / \lambda_k =$$

$$\frac{256(2(-1)^k t_f - (2k+1)\pi x_0)^2 (\pi/2 + \pi k)^2 t_f^2}{(1-\mu)(2k+1)^8 \pi^8}, \quad (53)$$

which directly yields (39). Thus, the conditions of Theorem 2 are valid, and the cheap control LQDG optimal strategy solves the tracking problem for

$$G(x(\cdot)) = \int_0^{t_f} (x(t)-t)^2 dt .$$

In Fig. 2, the curve $x = t$ and the trajectories of (46) for decreasing values of α, $\mu = 0.5$, $x_0 = 0$, $u = u_{\alpha\mu}^0(t,x)$ and $v(t) = \sin 10t$, are depicted. It is seen that the smaller is α, the better is tracking.
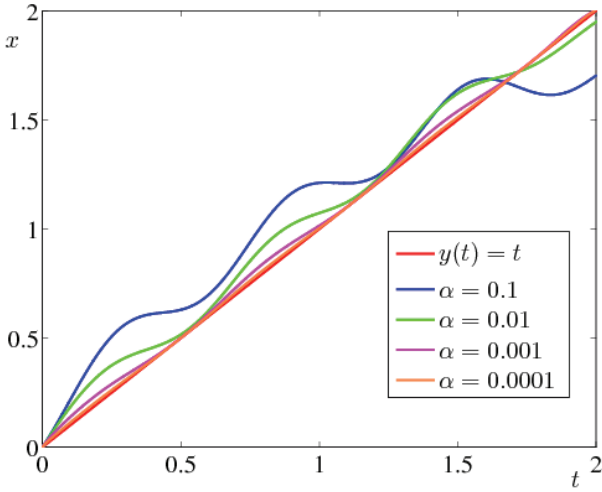
Figure 2: Tracking function $y(t) = t$

The respective control time realizations are presented in Fig. 3. It is seen that in this example the realizations are bounded.
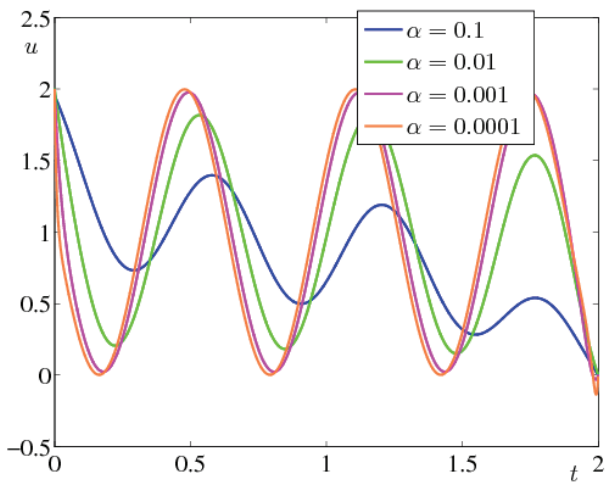
Figure 3: Control time realizations

5.2. Guidance Example

For $t_f = 4$, consider the guidance tracking problem for the system (1) -- (3) with the functional

$$G(x(\cdot)) = [x_1(1)-1]^2 + [x_1(1.5)-1.5]^2 +$$
$$[x_1(1.8)-3]^2 + x_1(4)^2 + x_2(4)^2 + \quad (54)$$
$$\int_2^4 [x_1(t) - t(4-t)]^2,$$

which is the particular case of (7). In Fig. 4 and 5, the tracking results (i.e. the graphs of $x_1(t)$ and $x_2(t)$) for $\tau_p = 0.2$ s, $\tau_e = 0.3$ s, $\alpha = 10^{-6}$, $\mu = 0.5$, $v(t) \equiv 100$ m/s^2. It is seen that the system trajectory (shown in blue) tracks accurately both prescribed points and prescribed function (shown in red).
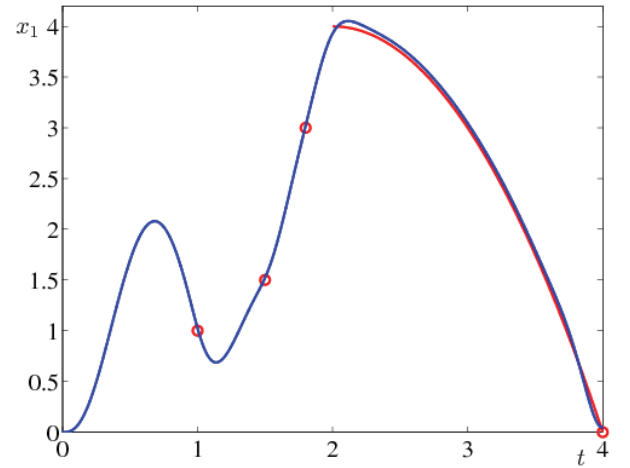
Figure 4: Guidance tracking problem: $x_1(t)$
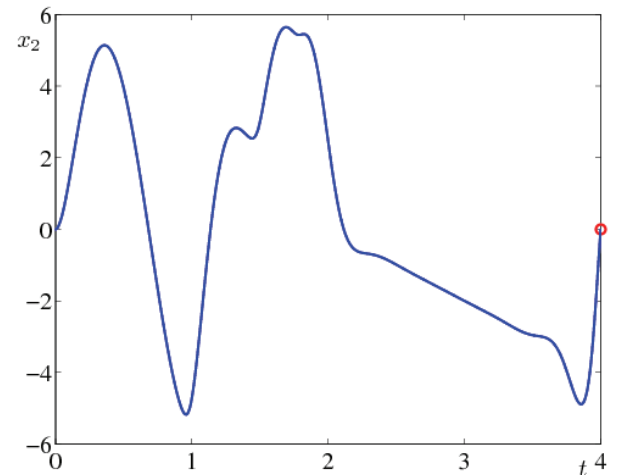
Figure 5: Guidance tracking problem: $x_2(t)$

Remark 1. *In Section 5.1, a relatively simle scalar example was considered. This example admits an analytical solution. For more complex example of Section 5.2, a numerical solution was obtained.*

REFERENCES

Aguiar, A. P., and Hespanha, J. P., 2007. Trajectory-tracking and path-following of underactuated autonomous vehicles with parametric modeling uncertainty. *IEEE Transactions on Automatic Control*, 52, 1362 –1379.

Balakrishnan, A. V., 1976. *Applied functional analysis*. Springer, New York, NY.

Basar, T., and Bernhard, P., 1995. H^∞-*optimal control and related minimax design problems: a dynamic game approach*. Birkhauser, Boston, MA.

Bell, D. J., and Jacobson, D. H., 1975. *Singular optimal control problems*. Academic Press, New York, NY.

Ben-Asher, J. Z., Levinson, S., Shinar, J., and Weiss, H., 2004. Trajectory shaping in linear-quadratic Pursuit-Evasion Games. *Journal of Guidance, Control, and Dynamics*, 27, 1102 – 1105.

Braslavsky, J. H., Seron, M. M., Maine, D. Q., and Kokotovic, P. V., 1999. Limiting performance of optimal linear filters. *Automatica*, 35, 189 – 199.

Cheng, J., Yi, J., and Zhao, D., 2007. Design of a sliding mode controller for trajectory tracking problem of marine vessels. *IET Control Theory and Applications*, 1, 233 – 237.

Cottrell, R.G., 1976. Optimal intercept guidance for short-range tactical missiles. *AIAA Journal*, 14,1210 – 1212.

Glizer, V.Y., 2000. Asymptotic solution of zero-sum linear-quadratic differential game with cheap control for the minimizer. *NoDEA: Nonlinear Differential Equations and Applications*, 7, 231–258.

Glizer, V.Y., 2009. Cheap control for a class of linear systems with state delays. *Journal of Nonlinear and Convex Analysis*, 10, 235 – 259.

Glizer, V.Y., Fridman, L.M., and Turetsky, V., 2007. Cheap suboptimal control of uncertain systems with state delays. *IEEE Transactions on Automatic Control*, 52, 1892 – 1898.

Kiselev, L. V., 2009. Control of the motion of an autonomous underwater robot in the trajectory tracking of physical fields of an ocean. Automation and Remote Control, 70, 692 – 698.

Kokotovic, P. V., 1984. Applications of singular-perturbation techniques to control problems. *SIAM Review*, 26, 501 – 550.

Kokotovic, P. V., Khalil, H. K., and O'Reilly, J., 1986. *Singular perturbation methods in control: analysis and design*. Academic Press, London, 1986.

Kowalczyk, W., Kozlowski, K. R., and Tar, J. K., 2009. Trajectory tracking for formation of mobile robots. In: M. Thoma, F. Allgover, M. Morari, editors, *Robot Motion and Control 2009*, Lecture Notes in Control and Information Sciences, Springer, Berlin: 57 – 66.

Krasovskii, N.N. and Subbotin, A.I., 1988. *Game-theoretical control problems*. Springer Verlag, New York, NY.

Kwakernaak, H., and Sivan, R., 1972. The maximally achievable accuracy of linear optimal regulators and linear optimal filters. *IEEE Transactions on Automatic Control*, 17, 79 – 86.

Mahony, R., and Hamel, T., 2004. Robust trajectory tracking for a scale model autonomous helicopter. *International Journal of Robust and Nonlinear Control*, 14, 1035 – 1059.

Mou, L., and Liberty, S.R., 2001. Estimation of maximal existence intervals for solutions to a riccati equation via an upper-lower solution method. *Proceedings of the 39th Allerton Conference on Communication, Control, and Computing*, Monticello, IL.

Moylan, P. J., and Anderson, B. D. O., 1973. Nonlinear regulator theory on an inverse optimal control problem. *IEEE Transactions on Automatic Control*, 18, 460 – 465.

Pei, X.-Z., Liu, Z.-Y., and Pei, R., 2003. Adaptive stabilization and trajectory tracking of airship with neutral buoyancy. *Acta Automatica Sinica*, 29, 876 – 882.

Petersen, I. R., 1986. Linear-quadratic differential games with cheap control. *Systems & Control Letters*, 8, 181 – 188.

Reid, W.T., 1972. *Riccati differential equations*, volume 86 of Mathematics in Science and Engineering. Academic Press, New York, NY.

Seron,M. M., Braslavsky, J. H., Kokotovic, P. V., and Mayne, D. Q., 1999. Feedback limitations in nonlinear systems: from Bode integrals to cheap control. *IEEE Transactions on Automatic Control*, 44, 829 – 833.

Shinar, J., 1981. Solution techniques for realistic pursuit-evasion games. In: C.T. Leondes ed., *Advances in control and dynamic systems*, Academic Press, New York, NY, 63 – 124.

Shinar, J., Turetsky, V., Glizer, V.Y., and Ianovsky, E., 2008. Solvability of linear-quadratic differential games associated with pursuit-evasion problems. *International Game Theory Review*, 10, 481 – 515.

Starr, A. W., and Ho, Y.C., 1969. Nonzero-sum differential games. *Journal of Optimization Theory and Applications*, 3, 184 – 206.

Sun, Z., Zhang, Z., and Tsao, T.-C., 2009. Trajectory tracking and disturbance rejection for linear time-varying systems: input/output representation. *Systems & Control Letters*, 58, 452 – 460.

Tretyakov, V. E., and Turetsky, V. Ya., 1995. Using an auxiliary differential game to synthesize a pursuit strategy. *Journal of Computer and System Sciences International*, 33, 140 – 145.

Turetsky, V., and Glizer, V.Y., 2004. Robust state-feedback controllability of linear systems to a hyperplane in a class of bounded controls. *Journal of Optimization Theory and Applications*, 123, 639 – 667.

Turetsky, V., and Glizer, V.Y., 2007. Robust solution of a time-variable interception problem: a cheap control approach. *International Game Theory Review*, 9, 637 – 655.

Wang, Z., Goldsmith, P., and Gu, J., 2009. Adaptive trajectory tracking control for Euler-Lagrange systems with application to robot manipulators. *Control and Intelligent Systems*, 37, 46 – 56.

Woodyatt, A.R., Seron, M.M., Freudenberg, J.S., and Middleton, R. H., 2002. Cheap control tracking performance for nonright-invertible systems. *International Journal of Robust and Nonlinear Control*, 12, 1253 – 1273.

Young,K. D., Kokotovic, P. V., and Utkin, V. I., 1997. A singular perturbation analysis of high-gain feedback systems. *IEEE Transactions on Automatic Control*, 22, 931 – 938.

Zarchan, P., 1994. *Tactical and strategic missile guidance*, volume 157 of Progress in Astronautics and Aeronautics. AIAA, Reston, VI, 2nd edition.

Zhang, Y., Qu, W.-D., Xi, Y.-G.., and Cai, Z.-L., 2008. Adaptive stabilization and trajectory tracking of airship with neutral buoyancy. *Acta Automatica Sinica*, 34, 1437 – 1440.

INTEGRATED DAMPING PARAMETERS AND SERVO CONTROLLER DESIGN FOR OPTIMAL H_2 PERFORMANCE IN HARD DISK DRIVES

Tingting Gao[a,b], Weijie Sun[a], Chunling Du[b], Lihua Xie[a]

[a] School of EEE, Nanyang Technological University, Singapore
[b] Data Storage Institute, A*STAR, Singapore

gaot0002@e.ntu.edu.sg, auwjsun@scut.edu.cn, du_chunling@dsi.a-star.edu.sg, elhxie@ntu.edu.sg

ABSTRACT

Achieving the highest achievable tracks per inch (TPI) in hard disk drives (HDDs) is equivalent to minimize the H_2 norm from the disturbances to the position error signals (PES). Thus the H_2 optimal performance is a matter of great significance. This paper presents an integrated method of the plant and controller design sequentially to achieve an optimal H_2 performance for voice coil motor (VCM) plant. The VCM plant is redesigned first to guarantee that the modified plant has a better H_2 performance by using the linear matrix inequality (LMI) based approach. Then the H_2 optimal controllers are designed for the modified plant. Finally, simulation results are presented to validate the effectiveness of the proposed modified plant redesign and the H_2 optimal controller design method. It is found that the modified plant has a better H_2 performance, higher bandwidth and phase margin than the original plant.

Keywords: damping ratio, H_2 performance, integrated design, H_2 optimal controller design

1. INTRODUCTION

Traditionally, people employ the sequential design method in the mechanical control design problem, where the mechanical plant satisfying the requirements of stiffness, strength, weight, etc., is first designed and then the controller design follows. For the controller design step, it is known that the change in the controller topology and/or the application of the advanced controller algorithm can improve the controlled performance of a given plant. In particular, the key role of advanced control techniques in improving precision and accuracy has been well recognized in the field of nanopositioning (Devasia, Eleftheriou, and Moheimani 2007; Du, Xie, Guo, and Teoh 2007; Du, Xie, Lewis, and Wang 2009). However, this sequential design method does not take the advantage of the freedom in plant design to help the mechanical controlled plant achieve a much better overall mechanical control performance. It fails to meet the practical requirements as systems demand higher performance (Joshi 1999), for example the ever increasing high precision control of the HDDs.

In order to tackle this problem, the simultaneous design method which aims to numerically optimize the parameters of both plant and controller is proposed (Hiramoto, and Grigoriadis 2006; Lu, and Skelton 2000; Grigoriadis, and Wu 1997). And it has been known that the system changes such as structural parameters can influence the ultimate performance. Unfortunately, this scheme has its inherent significant limitation since it has been known that such an integrated design method for most of the mechanical plant falls into a bilinear matrix inequality (BMI) problem, which requires iterative algorithm and intractable computation leading to a local optimal solution. The optimal solution depends on the initial conditions and conservative of the adopted iterative algorithm. Besides, this method does not provide us a deep understanding of how the plant parameters have an impact on the plant property.

There is another mechanical control design method proposed by Iwasaki, Hara, and Yamauchi 2003. In this scheme, they consider the plant design to guarantee the existence of a controller that achieves a good closed-loop performance and argue that the mechanical plant which is designed to be positive real in the desired control bandwidth is the so-called good plant. An LMI approach is given to verify this finite frequency positive real (FFPR) condition. Based on the concept of finite frequency property, the condition Π and finite frequency high-gain (FFHG) are further proposed to clarify the so-called easily controllable plants (Kanno, Hara, and Onishi 2007; Hara, Iwasaki, and Shimizu 2002). It seems that all of these three concepts of defining a good plant, i.e., FFPR, condition Π and FFHG, point to a fact that a good or easily controllable plant is the one that is designed to have the performance as close as possible to the open-loop shaping specifications of controller design, whose relationship we have shown in the Figure 1. In fact, this inherent relationship is not hard to intuitively understand since it is obvious that a plant having performance close to the open-loop shaping specifications may make the further controller design easier and even improve the overall controlled performance to a large extent. Although there is not much theoretical justification for these three concepts, they have been proven to be efficient in tackling the placement problem of actuator and sensor

to enhance the controlled performance (Iwasaki, Hara, and Yamauchi 2003; Kanno, Hara, and Onishi 2007; Hara, Iwasaki, and Shimizu 2002). However, when these concepts come to the design problem of the stiffness and damping parameters, they also result in a BMI problem.
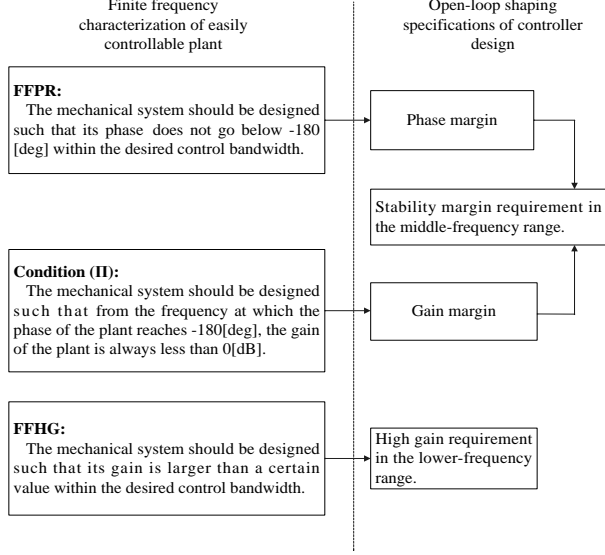
Figure 1: Relationship between Finite Frequency Plant Property and Typical Open-loop Shaping Specifications of Controller Design

Recently, the damping parameter design in structural systems such as the vector second order systems with collocated sensors and actuators has been paid more and more attention (Hiramoto, and Grigoriadis 2008; Mohammadpour, Meisami-Azad, and Grigoriadis 2008; Bai, and Grigoriadis 2005; Bai, Grigoriadis, and Demetriout 2006). Different from the simultaneous design method and the FFPR condition, the damping parameter design of such structural systems can be formulated as a convex optimization problem. This method has advantage over the other integrated design methods in the computational efficiency, especially for large scale structures.

In this paper, we will study the damping parameter design for a class of non-collocated vector second order systems. The static output feedback is used to modify the damping ratios of the plant, then the control design will be discussed. Furthermore, the result will be applied into the VCM plant model in HDD, which shows the efficiency of the proposed method.

2. SYSTEM MODELING
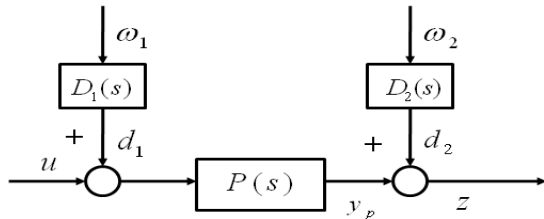
Figure 2: A Dynamic System $P(s)$ with Disturbances

Consider a dynamic system with disturbances as shown in Figure 2, where P is the plant of the dynamic system, d_1 is the input disturbance, d_2 is the output disturbance. D_1 and D_2 are the input and output disturbance models, and ω_1, ω_2 are white noises with zero mean and unit variance. The plant model of the system is easily identified using the curve fitting method. Generally, the plant can be represented as the following transfer function:

$$P(s) = \sum_{i=0}^{m} \frac{k_i}{s^2 + 2\zeta_i \overline{\omega}_i s + \overline{\omega}_i^2}, \tag{1}$$

where $\zeta_i > 0$ is the damping ratio, $\overline{\omega}_i$ is the resonance frequency, k_i is the residue of the resonance mode and m is the number of the resonance modes. This transfer function can be realized by the following vector second order system with non-collocated actuators and sensors

$$
\begin{aligned}
M\ddot{q}(t) + D\dot{q}(t) + Kq(t) &= B_u d_1(t) + B_u u(t), \\
y_p(t) &= C_0 q(t), \\
z(t) &= y_p(t) + d_2(t),
\end{aligned} \tag{2}
$$

where, $q(t)$, $u(t)$, $y_p(t)$ and $z(t)$ are the generalized coordinate vector, control input, plant output and measured output vector, respectively. Matrices M, D and K are symmetric positive definite matrices that represent the structural system mass, damping and stiffness matrices of the mechanical system, respectively. Matrices B_u and C_0 are the distribution matrices on the control input and the control output, respectively.

It is easy to know that (1) can be described in the form of (2), where

$$
M = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}, D = \begin{bmatrix} 2\zeta_0\overline{\omega}_0 & 0 & \cdots & 0 \\ 0 & 2\zeta_1\overline{\omega}_1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 2\zeta_m\overline{\omega}_m \end{bmatrix},
$$

$$
K = \begin{bmatrix} \overline{\omega}_0^2 & 0 & \cdots & 0 \\ 0 & \overline{\omega}_1^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \overline{\omega}_m^2 \end{bmatrix}, B_u = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix},
$$

$$
C_0 = [k_0 \quad k_1 \quad \cdots \quad k_m].
$$

The disturbance models D_1 and D_2 can be modeled based on the power spectrum of the PES. Assume that the input disturbance model D_1 is a constant gain G_D and the output disturbance model has the form

$$D_2(s) = \sum_{i=1}^{r} \frac{k_{d_i}}{s^2 + 2\zeta_{d_i}\omega_{d_i}s + \omega_{d_i}^2}, \tag{3}$$

where $\zeta_{d_i} > 0, \omega_{d_i}$ k_{d_i} and r is the damping ratio, resonance frequency, residue and number of the output disturbance resonance modes, respectively, then d_2 can also be generated by the following vector second order system

$$
\begin{aligned}
M_d\ddot{q}_d(t) + D_d\dot{q}_d(t) + K_d q_d(t) &= B_\omega \omega_2(t), \\
d_2(t) &= C_\omega q_d(t),
\end{aligned} \tag{4}
$$

where q_d is the coordinate vector of output disturbance model. Matrices M_d, D_d and K_d are the symmetric positive definite matrices that represent the mass, damping and stiffness matrices of the output disturbance model, respectively. Matrices B_ω and C_ω are the distribution matrices on the output disturbance input and the output disturbance output, respectively.

Denote $\bar{x}(t) = [q(t) \quad q_d(t) \quad \dot{q}(t) \quad \dot{q}_d(t)]$ and $\omega(t) = [\omega_1(t) \quad \omega_2(t)]^T$, then the state-space representation is

$$\dot{\bar{x}}(t) = \bar{A}\bar{x}(t) + \bar{B}_1\omega(t) + \bar{B}_2u(t),$$
$$z(t) = \bar{C}\bar{x}(t), \tag{5}$$

where $\bar{A} = \begin{bmatrix} 0 & 0 & I & 0 \\ 0 & 0 & 0 & I \\ -M^{-1}K & 0 & -M^{-1}D & 0 \\ 0 & -M_d^{-1}K_d & 0 & -M_d^{-1}D_d \end{bmatrix}$,

$\bar{B}_1 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ M^{-1}B_uG_D & 0 \\ 0 & M_d^{-1}B_\omega \end{bmatrix}$, $\bar{B}_2 = \begin{bmatrix} 0 \\ 0 \\ M^{-1}B_u \\ 0 \end{bmatrix}$, and

$\bar{C} = [C_0 \quad C_\omega \quad 0 \quad 0]$.

Remark 2.1: When $B_u = C_0^T$, the vector second order system corresponds to the one with collocated actuator and sensor placement. In particular, for a single-input-single-output (SISO) system, this would lead to an in-phase property of the mechanical systems in the whole frequency range. The mechanical plant with such an in-phase property has been claimed to be helpful to expect a good controlled performance (Ono and Teramoto 1992), however, it is a rather ideal case since the actuators and sensors are not perfect and such a requirement in the mechanical design may be practically infeasible (Iwasaki, Hara, and Yamauchi 2003) due to the unavoidable limitation on the control bandwidth. The integrated design for such a collocated vector second order system has been extensively studied in several papers, see Hiramoto, and Grigoriadis (2008); Mohammadpour, Meisami-Azad, and Grigoriadis (2008); Bai, and Grigoriadis (2005); Bai, Grigoriadis, and Demetriout (2006) and the references therein. What we consider here is more general than the previous discussed ones in that our formulation allows the case when $B_u \neq C_0^T$ and furthermore we will incorporate the output disturbance model. It is in this sense that we call our problem as integrated design for the non-collocated vector second order systems.

3. INTEGRATED DAMPING PARAMETERS AND CONTROLLER DESIGN FOR OPTIMAL H_2 PERFORMANCE

This section provides an LMI expression for the integrated damping parameters and controller design. The plant is designed first by using the output feedback to change the damping parameters.

Consider the vector second order system (5) with the measured output equation

$$v = C_0\dot{q}(t) + C_\omega\dot{q}_d(t). \tag{6}$$

Our objective is to design the static output feedback control law

$$u_G(t) = -Gv(t), \tag{7}$$

such that the closed-loop system is stable with its H_2 norm satisfying $\|T_{z\omega}\|_2 \leq \gamma$ for a given γ, as shown in Figure 3.
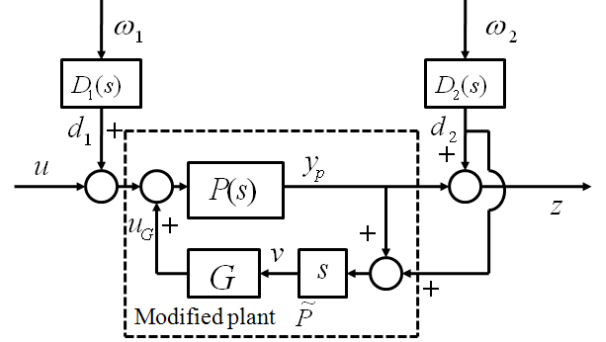
Figure 3: Structure of Vector Second Order System with Output Feedback Loop

It is easy to obtain the closed-loop system

$$\dot{\bar{x}}(t) = A_{cl}\bar{x}(t) + \bar{B}_1\omega(t) + \bar{B}_2u(t),$$
$$z(t) = \bar{C}\bar{x}(t), \tag{8}$$

where,

$$A_{cl} = \begin{bmatrix} 0 & 0 & I & 0 \\ 0 & 0 & 0 & I \\ -M^{-1}K & 0 & -M^{-1}\tilde{D} & -M^{-1}B_uGC_\omega \\ 0 & -M_d^{-1}K_d & 0 & -M_d^{-1}D_d \end{bmatrix},$$
$$\tilde{D} = D + B_uGC_0.$$

It has been shown that an upper bound of the H_2 performance for the collocated vector second order system can be explicitly given using a solution for the linear matrix inequality formulation of the norm analysis conditions (Mohammadpour, Meisami-Azad, and Grigoriadis 2007; Mona, Javad, and Karolos 2009; Mohammadpour, Meisami-Azad, and Grigoriadis 2008). In this paper, we will study the case of the non-collocated vector second order systems and apply the result into the analysis of VCM plant model. For this purpose, we first introduce the following lemmas.

Lemma 3.1 (Scherer, Gahinet, and Chilali 1997): Consider the stable system (5) with no control input, i.e. $u = 0$, then the system can achieve an H_2 norm γ if and only if there exist symmetric positive matrices Q and Z such that

$$\begin{bmatrix} \bar{A}^TQ + Q\bar{A} & Q\bar{B}_1 \\ \bar{B}_1^TQ & -I \end{bmatrix} < 0, \tag{9}$$

$$\begin{bmatrix} Q & \bar{C}^T \\ \bar{C} & Z \end{bmatrix} > 0, \tag{10}$$

$$trace(Z) < \gamma^2. \tag{11}$$

Lemma 3.2 (Boyd 1994): The block matrix

$\begin{bmatrix} S_{11} & S_{12} \\ S_{12}^T & S_{22} \end{bmatrix}$, where S_{11} and S_{22} are symmetric, is positive definite if and only if $S_{11} > 0$ and $S_{22} - S_{12}^T S_{11}^{-1} S_{12} > 0$; or $S_{22} > 0$ and $S_{11} - S_{12} S_{11}^{-1} S_{12}^{-1} > 0$.

Using the above formulation and lemmas, the solution of the integrated damping parameters and output feedback controller design to satisfy higher closed loop H_2 performance is obtained.

Theorem 3.1 For the closed-loop system (8), if the following matrix inequalities are satisfied,

$$\begin{bmatrix} -(\widetilde{D} + \widetilde{D}^T) + \alpha B_0 B_0^T & -B_u G C_\omega \\ -(B_u G C_\omega)^T & -2D_d + \alpha B_\omega B_\omega^T \end{bmatrix} < 0, \quad (12)$$

$$\begin{bmatrix} \alpha K & 0 & C_0^T \\ 0 & \alpha K_d & C_\omega^T \\ C_0 & C_\omega & Z \end{bmatrix} > 0, \quad (13)$$

$$trace(Z) < \gamma^2, \quad (14)$$

then the integrated design problem of designing the output feedback gain G and the damping parameters $\zeta_i, i = 0,1,\cdots, m$ in D is solvable such that the H_2 norm of the closed-loop system (8) satisfies $\|T_{z\omega}\|_2 \le \gamma$ for a given γ.

Proof: From Lemma 3.1, the closed loop system (8) with no input can achieve an H_2 norm γ if and only if there exist symmetric positive matrices Q and Z such that LMI (9)-(11) is feasible.

Choosing the candidate Lyapunov matrix Q as
$$Q = \alpha \begin{bmatrix} K & 0 & 0 & 0 \\ 0 & K_d & 0 & 0 \\ 0 & 0 & M & 0 \\ 0 & 0 & 0 & M_d \end{bmatrix},$$
where α is a positive scalar to be used as an additional degree of freedom in order to reduce the conservativeness of the H_2 norm bound. Then the LMI (9) can be simplified as follows,
$$\begin{bmatrix} -\alpha(\widetilde{D} + \widetilde{D}^T) & -\alpha B_u G C_\omega & \alpha B_0 & 0 \\ -(\alpha B_u G C_\omega)^T & -2\alpha D_d & 0 & \alpha B_\omega \\ \alpha B_0^T & 0 & -I & 0 \\ 0 & \alpha B_\omega^T & 0 & -I \end{bmatrix} < 0.$$

By using the Lemma 3.2, we can obtain
$$\begin{bmatrix} -(\widetilde{D} + \widetilde{D}^T) + \alpha B_0 B_0^T & -B_u G C_\omega \\ -(B_u G C_\omega)^T & -2D_d + \alpha B_\omega B_\omega^T \end{bmatrix} < 0$$
yielding the LMI (12) with respect to α, G and D.

Furthermore, $\begin{bmatrix} Q & \bar{C}^T \\ \bar{C} & Z \end{bmatrix} > 0$ is equivalent to LMI (13) by using Lemma 3.2 again.

Remark 3.1 The integrated damping parameter and output feedback gain design problem can be solved as an H_2 optimization problem as follows:

$$\min_{\alpha,\beta,\zeta_i,G} \gamma^2 \quad (15)$$
Subject to (12)-(14)

Remark 3.2 In fact, the closed-loop system (8) after integrated damping parameters and output feedback gain design in this section can be seen as a modified plant (see Figure 3). Based on the observation in Figure 1, we have concluded that the plant having a good open-loop specification may be expected to ease the control and improve the overall performance. Then the advanced control techniques can be further applied to achieve possible better control performance.

4. APPLICATION IN HDD

In this section, we will apply the result obtained in the above sections into the damping parameters design of the VCM plant in HDDs. It has been shown in the mechanical engineering community that the damping configuration for the head actuator can influence its tracking dynamics (Jiang and Miles 1999).

In order to achieve higher TPI, one would have to minimize track misregistration (TMR) caused by the disturbances. TMR is defined as three times of the standard deviation of the PES (z in Figure 7), i.e. $3\sigma_{PES}$. σ_{PES} is defined as

$$\sigma_{PES} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n} PES(i)^2}, \quad (16)$$

where n is the number of the PES samples.

Let $T_{z\omega}$ denote the transfer function from the disturbances to the PES. When n is large enough, the H_2 norm of $T_{z\omega}$ is given by

$$\|T_{z\omega}\|_2 = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n} PES(i)^2}. \quad (17)$$

Therefore, the control problem to increase TPI can be formulated as the H_2 optimization problem.

Consider the typical servo loop of HDDs as shown in Figure 2, where P is the VCM plant, d_1 is the input disturbance representing all torque disturbances, d_2 is the output disturbance including disturbances due to disk motion, motor vibration, suspension and slider vibrations.

Table 1: Parameters of the Original Plant

mode i	k_i	ζ_i	ω_i (rad/s)
0	1	0.2	$2\pi 100$
1	-0.8	0.02	$2\pi 9700$
2	-0.5	0.03	$2\pi 14000$
3	-1	0.03	$2\pi 20500$

We consider the VCM plant in a 2.5 inch HDD with three resonance modes, i.e. $= 3$, to show the efficiency of the integrated design method. The frequency response of the original plant is shown in Figure 4. The dash line stands for the measured data from the experiment and the solid line represents the modeled frequency response using the resonance

parameters in Table 1. Here, we omit a resonance mode at 7k Hz because of the large damping ratio.
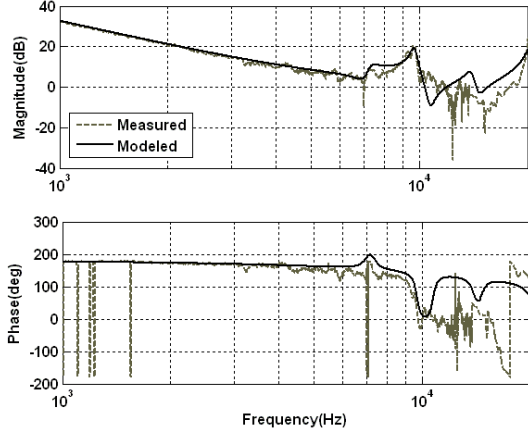
Figure 4: Frequency Response of the VCM Plant
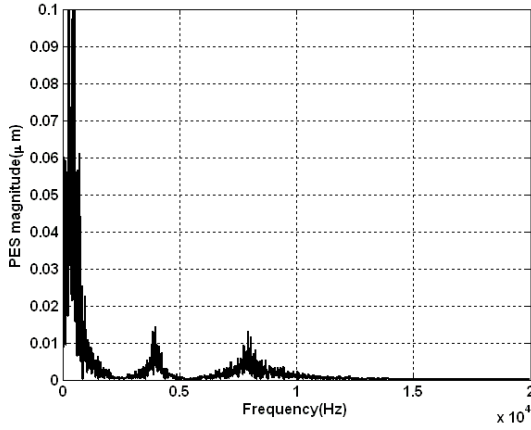
Figure 5: Power Spectrum of PES NRRO without Servo Control

Based on the power spectrum of PES NRRO data without servo control as shown in Figure 5, we can model the input disturbance model D_1 as the constant gain and the output disturbance model D_2 as the structure in Equation (3) with $r = 3$. The parameters of the output disturbance model are shown in Table 2, where three main resonance modes are consider.

Table 2: Parameters of the Output Disturbance

mode i	k_{d_i}	ζ_{d_i}	ω_{d_i} (rad/s)
1	0.04	0.3	$2\pi500$
2	0.005	0.03	$2\pi4000$
3	0.01	0.05	$2\pi8000$

The integrated design is proposed for the HDD VCM plant model to determine the damping ratio ζ_i and feedback gain G such that the H_2 performance of the closed-loop system (8) would satisfy $\|T_{z\omega}\|_2 \leq \gamma$ for some positive scalar γ. Such an LMI problem can be solved by (15).

After obtaining the damping parameters and the feedback gain of the plant, we can get the modified plant model $\tilde{P}(s)$. Here we just take two cases for example, as shown in Figure 6. The modified plant$_1$ and the modified plant$_2$ has the H_2 norm 4.4103nm and 3.9424nm, respectively. On the contrary, the original plant has the H_2 norm of 4.9655 nm. Obviously, the modified plant which is closer to the rigid mode will have a better control performance than the original plant.
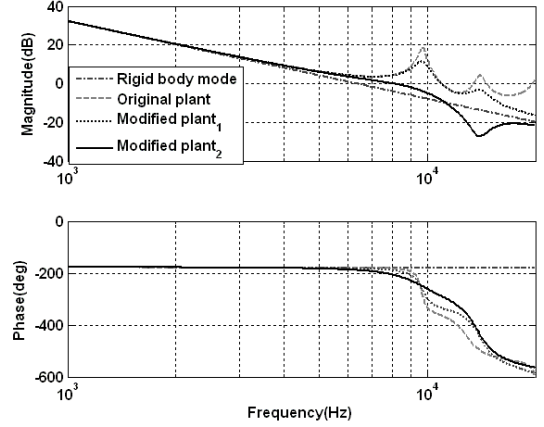
Figure 6: Comparison of the Frequency Responses of the Rigid Body Mode, Original Plant and Modified Plants

A block diagram presentation of a typical HDD servo loop with disturbances and measurement noise is shown in Figure 7. Then the optimal H_2 controller $\tilde{K}$ is designed for the original plant $P(s)$ and the modified plant $\tilde{P}(s)$ using algebraic Riccati equations (Li, Guo, Chen and Lee 2001), respectively.
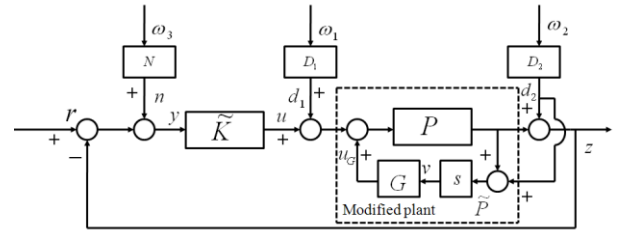
Figure 7: H_2 Control Scheme for HDD Servo Loop with Disturbance Models

Remark 4.1 In fact, the H_2 optimal controller $\tilde{K}$ and the output feedback G can be considered together as the combined controller K_{cc}, as shown in the Figure 8. Let $P: (A_p, B_p, C_p, 0)$ with the state $x_p = [q^T \quad \dot{q}^T]^T$, $D_2: (A_d, B_d, C_d, 0)$ with the state $x_d = \left[q_d^T \quad \dot{q}_d^T\right]^T$, $\tilde{K}: (A_c, B_c, C_c, 0)$ with the state x_c, $v = \tilde{C}_p x_p + \tilde{C}_\omega x_d$ with $\tilde{C}_p = [0 \quad C_0]$, $\tilde{C}_\omega = [0 \quad C_\omega]$. Define the state of the combined controller as $x_{cc} = \left[x_p^T \quad x_d^T \quad x_c^T\right]^T$, then the combined controller K_{cc} is

$$x_{cc}(k+1) = A_{cc}x_{cc}(k) + B_{cc}y(k),$$
$$\tilde{u}(k) = C_{cc}x_{cc}(k),$$

where $A_{cc} = \begin{bmatrix} A_p - B_p G \tilde{C}_p & -B_p G \tilde{C}_\omega & B_p C_c \\ 0 & A_d & 0 \\ 0 & 0 & A_c \end{bmatrix}$,

$B_{cc} = \begin{bmatrix} 0 \\ 0 \\ B_c \end{bmatrix}$, $C_{cc} = \begin{bmatrix} -G\tilde{C}_p & -G\tilde{C}_\omega & C_c \end{bmatrix}$,

$A_p = \begin{bmatrix} 0 & I \\ -M^{-1}K & -M^{-1}D \end{bmatrix}$, $B_p = \begin{bmatrix} 0 \\ M^{-1}B_u \end{bmatrix}$,

$A_d = \begin{bmatrix} 0 & I \\ -M_d^{-1}K_d & -M_d^{-1}D_d \end{bmatrix}$.
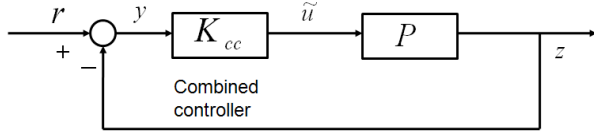
Figure 8: Equivalent Servo Control Systems of Figure 7

Table 3: Control Performance Comparison

Performance	Rigid body mode	Original plant	Modified plant$_1$	Modified plant$_2$
Open loop 0 dB crossover frequency (Hz)	1.82	1.45	1.55	1.74
Gain margin (dB)	7.93	7.93	7.93	7.93
Phase margin (deg)	30.3	28.1	28.3	28.6
H$_2$ performance (σ_z)(nm)	2.4803	2.8222	2.6985	2.5482

The comparison of the control performances is summarized in Table 3. The modified plants have the higher phase margin, higher bandwidth and the same gain margin. The H$_2$ performance of the closed loop systems with modified plants is lower than the original plant. The sensitivity functions are also compared in Figure 9. The modified plants have the higher bandwidth, but the sensitivity hump is almost the same. The power spectrum of PES NRRO data (z in Figure 7) of the original plant and the modified plant$_2$ are compared in Figure 10, from which we can observe that the modified plant$_2$ yields a considerable improvement in the low frequency up to 2k Hz. Above 8k Hz, the modified plant$_2$ leads to a bit small degradation. Overall, the H$_2$ norm of the modified is lower than the original plant, as shown in Table 3.
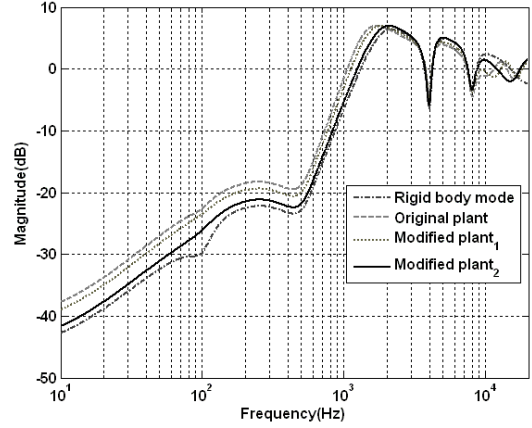
Figure 9: Sensitivity Function Comparison of the Rigid Body Mode, Original Plant and Modified Plants
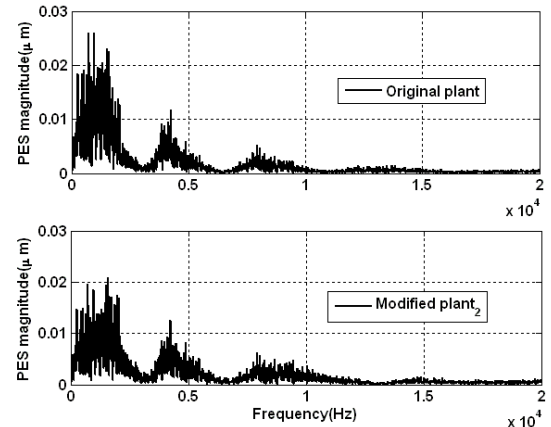
Figure 10: Power Spectrum Comparison of the PES NRRO with the Original Plant and Modified Plant$_2$

5. CONCLUSIONS

We have developed an LMI based approach to design the damping ratios of the plant and the output feedback gain to guarantee that the modified plant has a better H$_2$ optimal performance. The corresponding H$_2$ optimal controllers are designed for each case. A numerical example of the VCM plant in HDD is presented to show the validity and effectiveness of the proposed method. Two modified plants are proposed and compared with the original VCM plant. From the simulation, the proposed modified VCM plants have the lower H$_2$ performance, higher phase margin and bandwidth.

REFERENCES

Bai, Y., Grigoriadis, K. M., 2005. H$_\infty$ collocated control of structural systems: an analytical bound approach. *AIAA Journal of Guidance, Control, and Dynamics*, 28(5): 850-855.

Bai, Y., Grigoriadis, K. M., Demetriou, A. M., 2006. Damping parameter design in structural systems using an analytical H$_\infty$ norm bound approach. *Proceedings of the 45th IEEE Conference on Decision & Control*, 6537- 6542. December 13-15, San Diego, CA, USA.

Boyd, S., Ghaoui L. E., Feron, E. Balakrishnan, V., 1994. *Linear matrix inequalities in system and control theory*. Philadelphia: SIAM.

Devasia, S., Eleftheriou, E., Moheimani, S.O.R., 2007. A survey of control issues in nanopositioning. *IEEE Transactions on Control Systems Technology*, 15(5), 802-823.

Du, C., Xie, L., Guo, G., Teoh, J. N., 2007. A generalized KYP lemma based approach for disturbance rejection in data storage systems. *Automatica*, 43(12), 2112-2118.

Du, C., Xie, L., Lewis, F. L., Wang, Y., 2009. Multi-frequency disturbance rejection via blending control technique for hard disk drives. *Automatica*, 45(10), 2244-2249.

Grigoriadis, K.M., Wu, F., 1997. Integrated H_∞ plant controller design via linear matrix inequalities. *Proceedings of the 36th Conference on Decision & Control*, 789-790. December 10-12, San Diego, CA, USA.

Hara, S., Iwasaki, T., Shimizu, F., 2002. Finite frequency characterization of easily controllable mechanical systems under control effort constraint. *15th world congress of the international federation of automatic control*. July 22-26, Barcelona, Spain.

Hiramoto, K., Grigoriadis, K. M., 2006. Integrated design of structural and control systems with a homotopy like iterative method. *International Journal of Control*, 79(9), 1062-1073.

Hiramoto, K., Grigoriadis, K. M., 2009. Upper bound H_2 and H_∞ control for collocated structural systems. *Structural Control and Health Monitoring*, 16(4), 425-440.

Iwasaki, T., Hara, S., Yamauchi, H., 2003. Dynamical system design from a control perspective: finite frequency positive-realness approach. *IEEE Transaction on Automatic Control*, 48(8), 1337-1354.

Jiang, L., Miles, R. N., 1999. A passive damper for the vibration modes of the head actuator in hard disk drives, *Journal of Sound and Vibration*, 220(4), 683-694.

Joshi, S. S., 1999. The need for a systems perspective in control theory and practice. *IEEE Control Systems Magazine*, 19(6), 56-63.

Kanno, M., Hara, S., Onishi, M., 2007. Characterization of easily controllable plants based on the finite frequency phase/gain property: a magic number $\sqrt{4+2\sqrt{2}}$ in H_∞ loop shaping design. *American Control Conference*, 5816-5821. July 11-13, New York, NY, USA.

Kazuhiko, H., Javad, M., Karolos, M. G., 2009. Integrated design of system parameters, control and sensor actuator placement for symmetric mechanical systems. *48th IEEE Conference on Decision and Control joint with 28th Chinese Control Conference*, 2855-2860. December 16-18, Shanghai, P.R. China.

Li, Z., Guo, G., Chen, B., Lee, T.H., 2001. Optimal track-following design for highest track per inch in hard disk drives. *Journal of Information Storage and Processing Systems*. 3, 27-41.

Lu, J., Skelton, R. E., 2007. Integrating structure and control design to achieve mixed H_2/H_∞ performance, *International Journal of Control*. 73(16), 1449-1462.

Mohammadpour, J., Meisami-Azad, M., Grigoriadis, K. M., 2007. An efficient approach for damping parameter design in collocated structural systems using an H_2 upper bound, *American Control Conference*, 3015-3016. July 11-13, New York, NY, USA.

Mohammadpour, J., Meisami-Azad, M., Grigoriadis, K. M., 2009. Integrated damping parameter and control design in structural systems for H_2 and H_∞ specifications. *Structural and Multidisciplinary Optimization Journal*, 38(4), 377-387.

Mona, M., Javad, M., Karolos, M. G., 2009. Explicit solutions for collocated structural control with guaranteed H_2 norm Performance Specifications. *Smart Materials and Structures*, 18(3), 035004 (9pp).

Ono, K., Teramoto, T., 1992. Design methodology to stabilize the natural modes of vibration of a swing-arm positioning mechanism. *ASME Advances in Information Storage Systems*, 4, 343-359.

Scherer, C., Gahinet, P., Chilali, M., 1997. Multiobjective output feedback control via LMI optimisation. *IEEE Transaction on Automation Control*, 42(7), 896-911.

Skelton, R. E., Iwasaki, T., Grigoriadis, K. M., 1998. *A Unified Algebraic Approach to Linear Control Design*. London: Taylor & Francis.

Smith, M. J., Grigoriadis, K. M., Skelton, R. E., 1991. The optimal mix of passive and active control in structures, *American Control Conference*, 1459-1464. June 26-28, Boston, MA, USA.

TINGTING GAO
received the B.S and M.S degrees in mathematics from Northeastern University, Shenyang, China, in 2005 and 2008, respectively. Since 2008, she is a Ph.D. student with school of Electrical and Electronic Engineering, Nanyang Technological University, Singapore. Her research interests include high precision motion control and integrative approach for controller design for optimal performance of hard disk drive systems.

WEIJIE SUN
received the Ph.D. degree from the South China University of Technology in 2009. He is now a lecturer with the South China University of Technology. He is also with the Joint Research Center for Automation Science and Engineering, Chinese University of Hong Kong and South China University of Technology. His research interests are robust and adaptive control of nonlinear systems, and with applications to MEMS and power electronics.

CHUNLING DU
received the B.E. and M.E. degrees in electrical engineering from Nanjing University of Science and Technology, Nanjing, China, in 1992 and 1995, respectively, and the Ph.D degree in electrical engineering from Nanyang Technological University, Singapore, in 2000. In 1999, she joined the Agency for Science, Technology and Research (A*STAR) Data Storage Institute (DSI), Singapore, where she is currently a Senior Researcher in the Division of Mechatronics and Recording Channel. She is also an Adjunct Assistant Professor at Nanyang Technological University. Her current research interests include optimal and robust control, vibration modeling and control, motion control and multidimensional systems. She has authored or coauthored a few papers and the book H-infinity Control and Filtering of Two-dimensional Systems (Springer, 2002) and Modeling and Control of Vibration in Mechanical Systems (CRC 2010), and holds three patents.

LIHUA XIE
received his B.E. and M.E. degrees in Electrical Engineering from Nanjing University of Science and Technology in 1983 and 1986, respectively, and his Ph.D. degree in Electrical Engineering from the University of Newcastle, Australia, in 1992. Since 1992, he has been with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, where he is currently a Professor and the Director, Centre for Intelligent Machines. He held teaching appointments in the Department of Automatic Control, Nanjing University of Science and Technology from 1986 to 1989. Dr. Xie's current research interests include robust control and estimation, networked control, sensor networks, and multiagent systems. He has served as an Associate Editor of IEEE Transactions on Automatic Control, Automatica, IEEE Transactions on Control Systems Technology, IEEE Transactions on Circuits and Systems-II, International Journal of Control, Automation and Systems, and IET Proceedings on Control Theory and Applications. Dr. Xie is a Fellow of IEEE and a Fellow of IFAC.

Analytical Redundancy Relations from Bond Graphs of Hybrid System Models

W. Borutzky

Bonn-Rhein-Sieg University of Applied Sciences, D-53754 Sankt Augustin, Germany

Abstract—This paper picks up on one of the ways reported in the literature to represent hybrid models of engineering systems by bond graphs with *static* causalities. The representation of a switching device by means of a modulated transformer (MTF) controlled by a Boolean variable in conjunction with a resistor has been used so far to build a model for simulation. In this paper, it is shown that it can also constitute an approach to bond graph based quantitative fault detection and isolation in hybrid system models. Advantages are that Analytical Redundancy Relations (ARRs) do not need to be derived again after a switch state has changed. ARRs obtained from the bond graph are valid for all system modes. Furthermore, no adaption of the standard sequential causality assignment procedure (SCAP) with respect to fault detection and isolation (FDI) is needed.

It is shown that the approach proposed in this paper can produce the same ARRs given by Low et. al. for a network example reproduced in Fig. 1. Moreover, its usefulness is illustrated in a small case study by application to a switching circuit from the realm of power electronic systems. The approach, however, is not limited to FDI of such systems. Analytically checked simulation results give confidence in the approach.

Keywords—Hybrid models, FDI, ARRs, bond graphs with system operation mode independent causalities, power electronic systems, averaged bond graph models.

I. Introduction

Depending on the application, it is justified and convient to model fast state transitions as instantaneous discrete events giving rise to hybrid system models encompassing time continuous state transitions and discrete events. Bond graph representations of such hybrid models have been considered for a long time and various approaches have been reported in the literature. Early proposals have been to represent switches considered non-ideal by means of a modulated transformer controlled by a Boolean variable in conjunction with a resistor accounting for the small ON-resistance of the switch [1–3]. Other approaches also aiming at an *invariant* causality assignment independent from system modes have been the use of sinks of fixed causality switching off degrees of freedom [4], or the use of a Petri net representing system modes and discrete changes between them along with a set of bond graphs with standard elements modelling the time continuous behaviour in each identified system mode [5]. Also, in order to account for ideal switching in a bond graph with time-invariant causalities, so-called switched power junctions (SPJs) have been introduced more recently [6, 7].

Bond graph representations of hybrid models allowing for *variable* causalities are based on (ideal) switches [8–11] (switched bond graphs), or junctions controlled by a local automaton [12, 13]. Bond graphs with such controlled junctions are usually called hybrid bond graphs.

Moreover, early publications such as [14] have given rise to an increasing interest in bond graph model based quantitative fault diagnosis resulting in remarkable achievements during recent years [15–21]. Due to the nature of bond graphs, the focus has been mainly on fault detection and isolation (FDI) in systems represented by time continuous models. In [22], hybrid bond graphs are used for fault diagnosis in systems represented by hybrid models. As switching on or off controlled junctions entails at least a partial reassignment of causalities in a bond graph and affects the generation of Analytical Redundancy Relations (ARRs), a modification of the causality assignment procedure with respect to FDI has been recently proposed by Low and his co-workers [23–25]. They term the result a Diagnostic Hybrid Bond Graph.

In this paper, bond graph based FDI in hybrid models does not start from controlled junctions that entail either a dynamical reassignment of causality or a modification of the SCAP. Instead, conceptual switches are represented by transformers with Boolean modulus in conjunction with a resistor and by applying the unchanged standard Sequential Causality Assignment Procedure (SCAP) to the bond graph. As a result, ARRs derived from the bond graph with invariant causalities hold for all system modes. It is shown that derivation of equations leads to the same ARRs given by Low et. al. for an example [23].

Clearly, the representation of switching devices by a transformer with Boolean modulus and a resistor is not limited to electronically implemented switches and may be used for switching devices in other energy domains as well, e.g. for hydraulic check valves. Furthermore, suitable formulation of derived equations may enable the ON resistance of switches to be set to zero turning them into ideal switches.

The paper is organised as follows. The following section briefly revisits bond graph model based fault detection and isolation. Subsequently, the derivation of ARRs from a bond graph of a hybrid model is considered. The approach is illustrated by application to a simple buck converter example. Simulation results are verified by some analytical evaluation. The conclusion summarises the advantages of the approach.

II. Bond Graph Model-based Fault Detection and Isolation

Fault detection and isolation clearly needs fault indicators. In a bond graph model based approach, the sum of efforts or flows respectively at junctions can provide them. In a bond graph model of a non-faulty system, evaluation of these sums results in values called residuals that are equal or close to zero

due to numerical inaccuracies. If nonlinearities in the model permit, unknown variables in these equations can be replaced by inputs and known variables. That is, the result is a constraint relation between known variables usually termed Analytical Redundancy Relations (ARRs). Output variables of a bond graph model are often indicated by detector elements. In the context of FDI, these sensed variables are considered known variables. Accordingly, the causality of detectors in inverted. Furthermore, in order to be independent of initial values of energy stores, Samantaray et al. suggested to assign derivative causality as preferred causality to energy stores [21] and have termed the resulting bond graph diagnostic bond graph.

If values of the known variables in an ARR have been obtained by measurements of a real process or from another behavioural bond graph model accounting for possible faults in the process, then the residual of the ARR is likely to be different from zero over time due to noise in measurement, to parameter uncertainties or due to the occurrence of a fault. Noise in measurement can be suppressed by appropriate filtering before measured values are used in a diagnostic model. If the residual exceeds certain thresholds, then this event indicates that a fault has occurred in one of the system's components. As more than one system component usually contribute to an ARR, it is not clear in which component the indicated fault has happened. The information of which components are involved in an ARR is called the signature of the residual. Residuals with different signature are called structurally independent. Their number is equal to the number of sensors added to a system [21]. However, the set of ARRs is not unique. The information of which system component contributes to which residual is usually expressed in a Fault Signature Matrix (FSM) [26]. Its diagonal part directly indicates single faults that can be isolated. For isolation of simultaneous faults, parameter estimation by means of least squares optimisation has been used [20]. In this paper, the single fault hypothesis is adopted.

III. DERIVATION OF ARRs FROM A BOND GRAPH OF A HYBRID MODEL

In this paper, switches in hybrid models are represented in a bond graph by means of a transformer modulated by a Boolean variable and a resistor with statically assigned conductance causality as has been initially proposed by Ducreux, Dauphin-Tanguy and Rombaut for bond graph modelling of power electronic circuits [2]. The advantage of this approach is that the hybrid model is represented by one single bond graph and application of the SCAP results in causalities that hold for all physically feasible combinations of switch states. From such a bond graph of a hybrid model, ARRs can be derived in the same way as from a bond graph of a time continuous model.

For illustration, the network example used by Low et. al. in [23] is adopted. It is shown that the approach in this paper leads to the same ARRs given in [23]. Fig. 1 displays the circuit diagram and Fig. 2 an associated bond graph. As can be seen from the bond graph in Fig. 2, detector causalities have been inverted and *derivative* causality has been assigned to energy stores as proposed by Samantaray et al. [20]. although
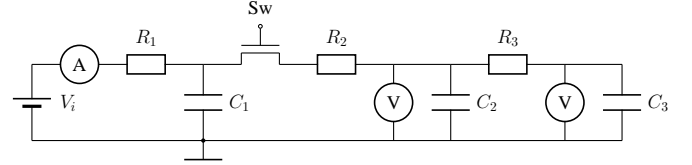
Figure 1. Network with a switch (cf. [23])
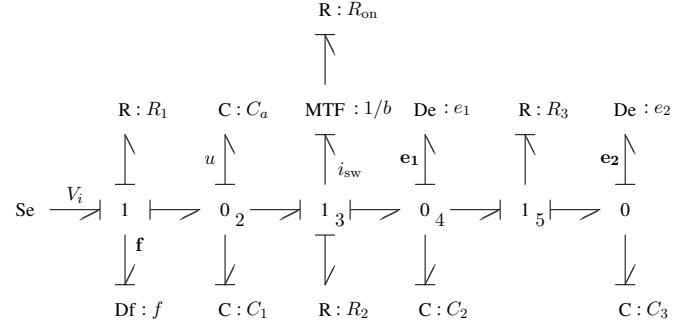
Figure 2. Bond graph of the network in Fig. 1

this is not necessary for the derivation of ARRs expressed in terms of the derivatives of energy storage variables instead of their integral. The auxiliary capacitor $C : C_a$ in integral causality with a small capacitance C_a has been added to resolve the causal conflict at junction 0_2. In its constitutive relation solved for its current, C_a is considered small so that the current vanishes. In the formulation of equations, the parameter of an auxiliary storage element used for resolving a causality conflict at a junction is set to zero. That is, the auxiliary storage elements will not lead to a set of stiff model equations with regard to simulation performed for a numerical evaluation of residuals.

For comparison, Fig. 3 reproduces the diagnostic hybrid bond graph given by Low et al. in [23]. In that diagnostic hybrid bond graph, junction 1_3 is a controlled junction accounting for the connection and the disconnection of circuit nodes by the pass transistor modelled as a switch. The resistor $R : R_{p2}$ is an artificial resistor resolving the causal conflict at junction 0_4 similar to the auxiliary capacitor $C : C_a$ in the bond graph of Fig. 2.

In [23], ARRs are derived from the junctions $0_2, 0_4, 1_5$. In general, ARRs are obtained from the balance equation of those junctions to which a detector has been attached that represent a real sensor. According to the choice made in [23], summation of flows at junction 0_2 of the bond graph in Fig. 2 yields for

Figure 3. Diagnostic hybrid bond graph of the network (cf. [23])

Figure 4. Two tank system (cf. [25])

Figure 5. Bond graph with Boolean controlled MTFs of the two tank system

Figure 6. Schematic of the DC-DC buck converter circuit

the residual r_1

$$r_1 = f - C_1 \dot{u} - i_{\mathrm{sw}} - C_a \dot{u} . \tag{1}$$

As C_a is assumed very small, the term $C_a \dot{u}$ can be neglected in (1). Due to the conductance causality of the ON resistance $R : R_{\mathrm{on}}$ the constitutive relation of the switch takes the form

$$i_{\mathrm{sw}} = \frac{b^2}{R_{\mathrm{on}} + b^2 R_2}(u - e_1) \tag{2}$$

where $b \in \{0, 1\}$. Finally, the voltage u is determined from the sum of efforts at the left 1-junction

$$u = V_i - R_1 f . \tag{3}$$

As a result, the ARR for r_1 reads

$$\begin{aligned} r_1 = & f - C_1 \frac{d}{dt}(V_i - R_1 f) \\ & - \frac{b^2}{R_{\mathrm{on}} + b^2 R_2}(V_i - R_1 f - e_1) . \end{aligned} \tag{4}$$

Likewise, summation of flows at junction 0_4 and of efforts at junction 1_5 respectively and elimination of unknown variables according causal paths in the bond graph gives the ARRs

$$r_2 = i_{\mathrm{sw}} - C_2 \dot{e}_1 - C_3 \dot{e}_2 \tag{5}$$

$$r_3 = e_1 - R_3 C_3 \dot{e}_2 - e_2 . \tag{6}$$

These results are identical to the ones reported in [23] if R_{on} is neglected in the ON mode of the switch.

Clearly, the derivation of ARRs from bond graphs in which elements modelled as switches are represented by a Boolean controlled MFT in conjunction with a resistor in conductance causality is not limited to electronic circuits with switching elements. In [25], Low et. al. consider a hydraulic two tank system in which an ON-OFF controller ensures that the fluid level in the tank connected to the pump does not exceed a certain level. This tank has got a drain to prevent an overflow in case the controller fails (cf. Fig. 4). Fig. 5 displays a bond graph with Boolean controlled MTFs of that system. The residual r_1 derived from the bond graph in Fig. 5 reads

$$r_1 = b_1 Q_p - C_1 \dot{p}_1 - b_2 k_1 \sqrt{b_2 |p_1 - p_2|} - Q \tag{7}$$

where k_1 is a constant, $b_2 = 0$ for $p_1 \leq p_D$ and $b_2 = 1$ for $p_1 > p_D$ (overflow).

IV. CASE STUDY

Beyond the result of the previous section, FDI in a DC-DC buck converter circuit [16] is considered in a small case study in order to show that bond graph modelling based on the representation of switches by modulated transformers and the use of the standard SCAP can well support FDI in hybrid system models and produces correct result.

Fig. 6 shows the circuit schematic of the buck converter. A bond graph of the buck converter is displayed in Fig. 7. Again, the capacitor $C : C_a$ has been attached to junction 0_2 to resolve the causal conflict at that junction caused by the fixed conductance causality of the ON resistors of the switches. Physically, it can be justified by the small capacitance of the circuit node against ground.

Assume that the inductor current i_L and the voltage drop across the capacitor u_C are measured. Accordingly, Fig. 8

Figure 7. Bond graph of the DC-DC buck converter circuit (cf. Garcia, 1997)

Figure 8. Diagnostic bond graph of the DC-DC buck converter circuit

TABLE I
FAULT SIGNATURE MATRIX OF THE BUCK CONVERTER WITH SENSORS
Df : i_L AND De : u_C

Component	Parameter/ Output	r_1	r_2	D_b	I_b
Supply of signal m_1	m_1	1	0	1	0
Switch Q	R_Q	$m_1 = 1$	0	$m_1 = 1$	0
Diode D	R_D	$m_2 = 1$	0	$m_2 = 1$	0
Inductor	L	1	0	1	0
Capacitor	C	0	1	1	0
Resistor	R	0	1	1	0
Sensor of i_L	i_L	1	1	1	0
Sensor of u_C	u_C	1	1	1	0

depicts the diagnostic bond graph. From the diagnostic bond graph, the following two ARRs can be derived

$$1_2: \quad r_1 \;=\; u - L\frac{di_L}{dt} - u_C \tag{8}$$

$$0_2: \quad r_2 \;=\; i_L - C\dot{u}_C - \frac{1}{R}u_C \tag{9}$$

where

$$u \;=\; km_1^2 E - kR_Q i_L \tag{10}$$

and

$$k := \frac{R_D}{m_1^2 R_D + m_2^2 R_Q} \tag{11}$$

with $m_2 = 1 - m_1$.

With these two residuals the structural fault signature matrix in Table I can be set up. Clearly, a fault in the pass transistor or in the diode can only be detected when these elements are active. This is indicated in the first additional column with the heading D_b. As can be seen from the last column with the heading I_b, no fault can be isolated given the two sensors. However, isolability can be improved by adding more sensors.

TABLE II
PARAMETERS OF THE BUCK CONVERTER

Parameter	Value	Units
E	100	V
R_{on}	0.1	Ω
L	50	mH
C	50	μF
R	50	Ω
T	10^{-3}	s
α	0.7	-

Figure 9. Time evolution of the inductor current i_L for the case of a non-faulty operation

A. Non-Faulty Behaviour

Neglecting the capacitance C_a, the following two state equations can be derived from the bond graph in Fig. 7.

$$\frac{di_L}{dt} \;=\; \frac{1}{L}\left[u - u_c\right] \tag{12a}$$

$$\frac{du_C}{dt} \;=\; \frac{1}{C}\left[i_L - \frac{u_C}{R}\right] \tag{12b}$$

Given the parameters listed in Table II, and assuming that $R_Q = R_D = R_{on}$, Fig. 9 shows the time evolution of the current i_L through the inductor and of its mean value i_{La} for the case of a non-faulty operation of the buck converter. In Table II, T denotes the duty cycle of the signal switching the pass transistor Q1 on and off and α the duty ratio. The transistor is on for the period αT, while it is off for the remaining part $(1 - \alpha)T$ of the period.

For $R_Q = R_D = R_{on}$, the dynamic equations of the average inductor current i_{La} and the the average voltage u_{Ca} read

$$\frac{di_{La}}{dt} \;=\; \frac{1}{L}\left[\alpha E - R_{on} i_{La} - u_{Ca}\right] \tag{13a}$$

$$\frac{du_{Ca}}{dt} \;=\; \frac{1}{C}\left[i_{La} - \frac{u_{Ca}}{R}\right]. \tag{13b}$$

Fig. 10 shows a bond graph of the average model.

B. Fault Scenario 1

As one of the possible fault scenarios, consider the case that the signal controlling the switch Q is not properly supplied for
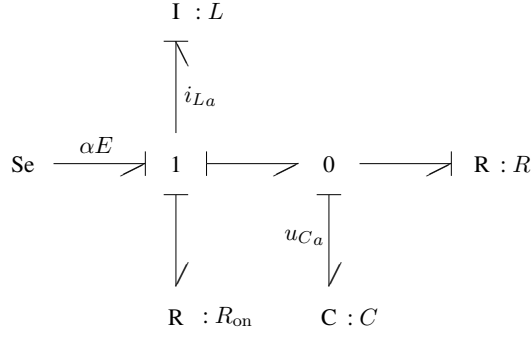
Figure 10. Bond graph of the average model

Figure 11. Faulty time evolution of the inductor current due to a temporarily permanently closed switch
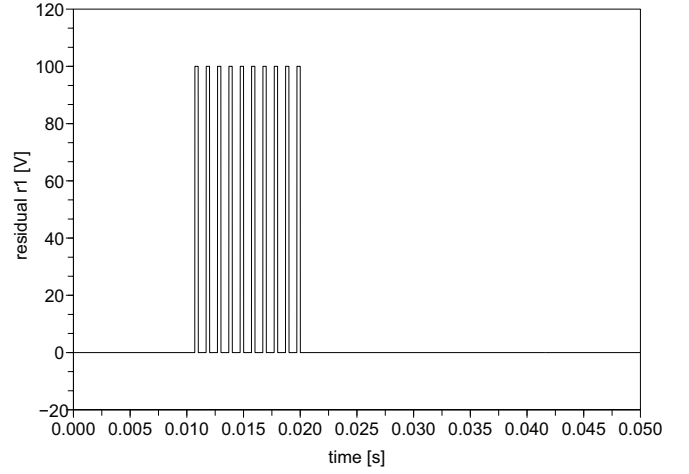

Figure 12. Residual r_1 indicating the temporary permanent closure of the switch
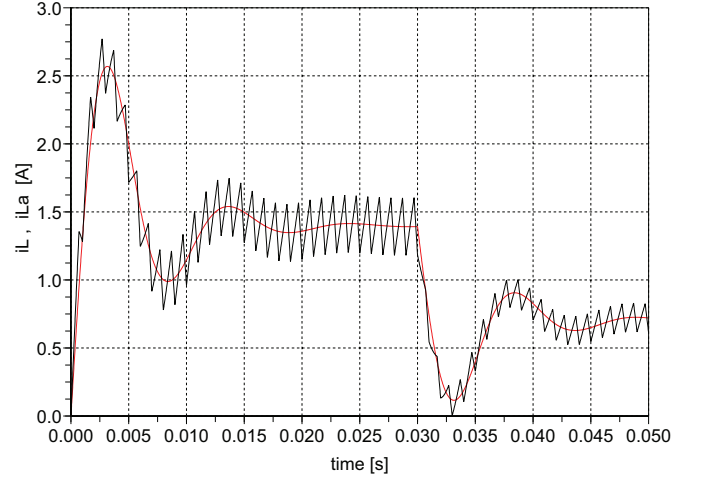
Figure 13. Faulty time evolution of the inductor current in case 2

some period of time. Let the switch be permanently closed for the time interval $[0.01s, 0.02s]$ and assume that this is the only fault. By consequence, the diode is permanently off during that interval. Fig. 11 shows the faulty time history of the current i_L. As to be expected from the fault signature matrix, this single fault is indicated by the time evolution of residual r_1 (cf. Fig. 12). For the residual r_1, an analytical expression can be found and used for verification of the result obtained by simulation. Let

$$\tilde{k} = \frac{\tilde{R}_D}{\tilde{m}_1^2 \tilde{R}_D + \tilde{m}_2^2 \tilde{R}_Q} \qquad (14)$$

with the tilde denoting possibly disturbed variables or parameters. Then, the expression for r_1 reads

$$r_1 = (\tilde{k}\tilde{m}_1^2 - km_1^2)E - (\tilde{k} - k)R_Q i_L . \qquad (15)$$

A permanent closure of the switch Q1 during the time interval under consideration means $\tilde{m}_1 = 1$, while in the non-faulty system model, m_1 switches values between 0 and 1. Thus, for $\tilde{m}_1 = 1, \tilde{m}_2 = 0$ and $m_1 = 1, m_2 = 0$ the residual r_1 becomes zero. For $\tilde{m}_1 = 1, \tilde{m}_2 = 0$ and $m_1 = 0, m_2 = 1$, the expression for r_1 reduces to E. That is, during the time interval $[0.01s, 0.02s]$, the value of r_1 oscillates between zero and $E = 100V$ as displayed in Fig. 12.

C. Fault Scenario 2

In a second fault scenario, the diode is considered to be permanently conducting as of a time instant $t_3 = 0.03s$. That is, $\tilde{m}_1 = m_1$, $\tilde{m}_2 = 1$. Fig. 13 shows the faulty time history of the current $\tilde{i}_L$ along with its mean value for this case. For $t > t_3$, the state equations read

$$L\frac{d\tilde{i}_L}{dt} = \frac{R_D m_1^2}{m_1^2 R_D + R_Q}E - \frac{R_D}{m_1^2 R_D + R_Q}R_Q \tilde{i}_L$$
$$- \tilde{u}_C \qquad (16a)$$
$$C\frac{d\tilde{u}_C}{dt} = \tilde{i}_L - \frac{\tilde{u}_C}{R} . \qquad (16b)$$

From these equations, the mean values $\tilde{i}_{La}$, $\tilde{u}_{Ca}$ for $t \to \infty$ can be computed. For $R_Q = R_D = R_{on}$, these values are $\tilde{u}_{Ca} = 35V$ and $\tilde{i}_{La} = 0.7A$ in accordance with Figures 13 and 14. This fault occurring for $t > t_3$ is indicated by the time evolution of residual r_1 in Fig. 15 as to be expected from the FSM. Again, the simulation result for r_1 can be checked analytically. Let $t > t_3$ and $R_Q = R_D = R_{on}$. In this case,
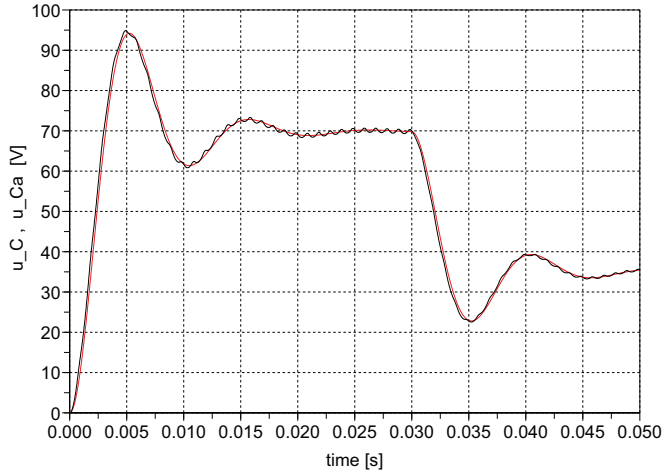
47

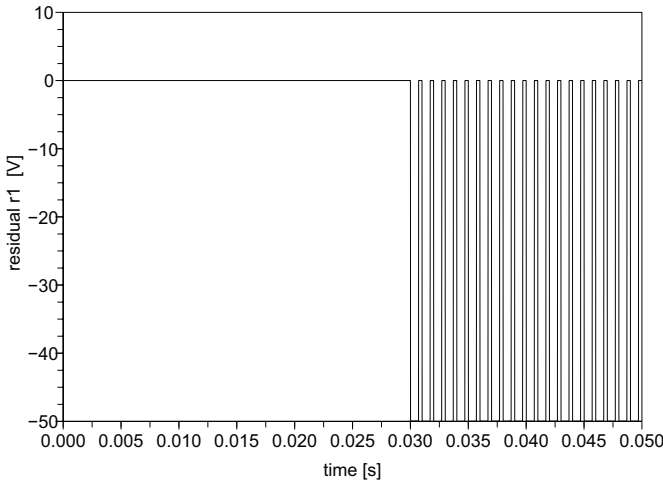Figure 14. Faulty time evolution of the voltage across the capacitor in case 2

Figure 15. Residual r_1 in case 2

(15) then reduces to

$$r_1 = -\frac{m_1^4}{m_1^2+1}E + \frac{m_1^2}{m_1^2+1}R_Q i_L . \qquad (17)$$

As the ON resistance R_{on} is small, R_Q is neglected in (17). As a result, it can be seen that r_1 oscillates between zero ($m_1 = 0$) and the value $-E/2 = -50V$ in case $m_1 = 1$.

V. CONCLUSION

Aiming at a bond graph based derivation of ARRs for hybrid models, this paper picks up on a proposal known for a long time to represent switches by a transformer modulated by a Boolean variable and a resistor in fixed conductance causality accounting for its ON resistance. This representation has been used to come up with a model for the purpose of simulation. The paper demonstrates that this representation can well constitute an approach to bond graph based quantitative FDI of hybrid models offering the following advantages.

- Computational causalities once assigned are independent of system modes. There is no need for adjusting causalities after a change from one system operation mode to

another.

This means that neither model equations nor ARRs need to be derived again after a discrete change of a system mode.

- There is no need for a special version of the SCAP as has been proposed for bond graphs with controlled junctions in [23].
- Existing bond graph software such as SYMBOLS [27] could be used to generate a single set of ARRs. As these ARRs include Boolean variables, there is not one single FSM but a set. Detection and isolation of faults become system mode dependent. (In [23], they are called Global ARRs.)

Causal conflicts at junctions may require auxiliary storage elements with a small parameter value to be attached. However, in the derivation of equations from a diagnostic bond graph with storage elements in preferred derivative causality, the parameter of these auxiliary storage elements can be set to zero so that the additional storage elements will not lead to a set of stiff model equations with regard to simulation performed for a numerical evaluation of residuals. If it is decided to keep the small ON resistance of the switch model, i.e. switching devices are not represented by ideal switches, then small time constants may result.

It is shown that the approach in this paper can come up with the same ARRs given in [23] for a network example. Moreover, the approach has been applied to a switching circuit. Simulation results obtained with Scilab [28] have been analytically checked giving rise to confidence in this approach.

REFERENCES

[1] G. Dauphin-Tanguy and C. Rombaut, "Why a unique causality in the elementary commutation cell bond graph model of a power electronics converter," in *1993 IEEE International Conference on Systems, Man and Cybernetics*, vol. 1, 1993, pp. 257–263.

[2] J.P. Ducreux, G. Dauphin-Tanguy, and C. Rombaut, "Bond Graph Modelling of Commutation Phenomena in Power Electronic Circuits," in *International Conference on Bond Graph Modeling, ICBGM'93, Proc. of the 1993 Western Simulation Multiconference*, J.J. Granda and F.E. Cellier, Eds. SCS Publishing, January 17-20 1993, pp. 132–136, simulation Series, volume 25, no. 2, ISBN: 1-56555-019-6.

[3] J. Garcia-Gomez, "Approche bond graph pour la modélisation des effets thermiques dans les composants de commutation en électronique de puissance," Ph.D. dissertation, Université des Sciences et Technologies de Lille, Lille, France, 1997.

[4] W. Borutzky, "Representing discontinuities by means of sinks of fixed causality," in *1995 International Conference on Bond Graph Modeling, ICBGM'95, Proc. of the 1995 Western Simulation Multiconference*, F. E. Cellier and J. J. Granda, Eds. SCS Publishing, January 15-18 1995, pp. 65–72, simulation Series, Vol. 27, Number 1, ISBN: 1-56555-037-4.

[5] W. Borutzky, J. F. Broenink, and K. C. J. Wijbrans, "Graphical Description of Physical System Models Containing Discontinuities," in *Modelling and Simulation 1993, Proc. of the 1993 European Simulation Multiconference*, A. Pavé, Ed. SCS Publishing, June 7-9 1993, pp. 208–214, lyon, France.

[6] A.C. Umarikar, "Modelling of Switched Mode Power Converters: A Bond Graph Approach," Ph.D. dissertation, Centre for

Electronics Design and Technology, Indian Institute of Science, Bangalore, India, August 2006.

[7] A.C. Umarikar and L. Umanand, "Modelling of switched systems in bond graphs using the concept of switched power junctions," *Journal of the Franklin Institute*, vol. 342, pp. 131–147, 2005.

[8] G.M. Asher, "The Robust Modelling of Variable Topology Circuits Using Bond Graphs," in *International Conference on Bond Graph Modeling, ICBGM'93, Proc. of the 1993 Western Simulation Multiconference*, J.J. Granda and F.E. Cellier, Eds. SCS Publishing, January 17-20 1993, pp. 126–131, simulation Series, volume 25, no. 2, ISBN: 1-56555-019-6.

[9] J. Buisson, H. Cormerais, and P-Y. Richard, "Analysis of the bond graph model of hybrid physical systems with ideal switches," *Proc. of the Institution of Mechanical Engineers Part I: Systems and Control Engineering*, vol. 216(1), pp. 47–63, 2002.

[10] K. Edström, "Switched Bond Graphs: Simulation and Analysis," Ph.D. dissertation, Linköping University, Linköping, Sweden, 1999.

[11] J.E. Strömberg, "A mode switching modelling philosophy," Ph.D. dissertation, Linköping University, Linköping, Sweden, 1994.

[12] P.J. Mosterman, "Hybrid Dynamic Systems: A hybrid bond graph modeling paradigm and its application in diagnosis," Ph.D. dissertation, Vanderbilt University, Nashville, TN, USA, 1997.

[13] I. Roychoudhury, M. Daigle, G. Biswas, X. Koutsoukos, and P.J. Mosterman, "A Method for Efficient Simulation of Hybrid Bond Graphs," in *Proc. of the 2007 International Conference on Bond Graph Modeling and Simulation*, J. J. Granda and F. E. Cellier, Eds., vol. 39 (1), 2007, pp. 177–184.

[14] M. Tagina, J.P. Cassar, G. Dauphin-Tanguy, and M. Staroswiecki, "Monitoring of Systems Modelled by Bond-Graphs," in *ICBGM'95, International Conference on Bond Graph Modeling and Simulation*, ser. Simulation Series, F.E. Cellier and J.J. Granda, Eds., vol. 27(1). Las Vegas, Nevada, USA: SCS Publishing, 15–18 Jan 1995, pp. 275–280.

[15] W. Borutzky, "Bond Graph Model-Based Fault Detection Using Residual Sinks," *Proc. of the Institution of Mechanical Engineers Part I Journal of Systems and Control Engineering*, vol. 223(3), pp. 337–352, 2009.

[16] W. Borutzky, *Bond Graph Methodology – Development and Analysis of Multidisciplinay Dynamic System Models*. Springer-Verlag, London, UK, 2010, ISBN : 978-1-84882-881-0.

[17] W. Borutzky, Ed., *Bond Graph Modelling of Engineering Systems – Theory, Applications and Software Support*. Springer-Verlag, NY, NY, U.S.A, 2011.

[18] M.A. Djeziri, R. Merzouki, B. Ould Bouamama, and G. Dauphin-Tanguy, "Robust Fault Diagnosis by Using Bond Graph Approach," *IEEE/ASME Transactions on Mechatronics*, vol. 12, no. 6, pp. 599–611, December 2007.

[19] S.K. Ghoshal, "Model-based Fault Diagnosis and Accommodation using Analytical Redundancy: A Bond Graph Approach," Ph.D. dissertation, Dept. of Mechanical Engineering, Indian Institute of Technology, Kharagpur, India, 2006.

[20] A.K. Samantaray and B. Ould Bouamama, *Model-based Process Supervision – A Bond Graph Approach*, ser. Advances in Industrial Control. Springer, London, 2008.

[21] A.K. Samantaray, K. Medjaher, B. Ould Bouamama, M. Staroswiecki, and G. Dauphin-Tanguy, "Diagnostic bond graphs for online fault detection and isolation," *Simulation Modelling Practice and Theory*, vol. 14, no. 3, pp. 237–262, 2006.

[22] S. Narasimhan, "Model-based diagnosis of hybrid systems," Ph.D. dissertation, Vanderbilt University, 2002.

[23] C.B. Low and D. Wang and S. Arogeti and J.B. Zhang, "Monitoring ability analysis and qualitative fault diagnosis using hybrid bond graph," in *Proceedings of the 17th World Congress*. The International Federation of Automatic Control, July 6–11 2008, pp. 10 516–10 521, seoul, Korea.

[24] C.B. Low and D. Wang and S. Arogeti and J.B. Zhang, "Causality assignment and model approximation for quantitative hybrid bond graph-based fault diagnosis," in *Proceedings of the 17th World Congress*. The International Federation of Automatic Control, July 6–11 2008, pp. 10 522–10 527, seoul, Korea.

[25] C.B. Low and D. Wang and S. Arogeti and J.B. Zhang, "Causality Assignment and Model Approximation for Hybrid Bond Graph: Fault Diagnosis Perspectives," *IEEE Transactions on Automation Science and Engineering*, vol. 7(3), pp. 570–580, 2010.

[26] B. Ould Bouamama, A.K. Samantaray, M. Staroswiecki, and G. Dauphin-Tanguy, "Derivation of Constraint Relations from Bond Graph Models for Fault Detection and Isolation," in *Proc. of the International Conference on Bond Graph Modeling, ICBGM'03*, J.J. Granda and F.E. Cellier, Eds. Orlando, Florida, USA: SCS Publishing, January 19–23 2003, pp. 104–109, simulation Series, volume 35, no 2, ISBN: 1-56555-257-1.

[27] HighTec Consultants, "SYMBOLS Shakti™ ." [Online]. Available: http://www.htcinfo.com/

[28] Scilab Consortium, "Scilab." [Online]. Available: http://www.scilab.org/

Wolfgang Borutzky is a professor for Modelling and Simulation of Engineering Systems at Bonn-Rhein-Sieg University of Applied Sciences, Germany. He obtained his Diploma Degree in Mathematics in 1979 and his Doctoral Degree in Mechanical Engineering in 1985 both from the Technical University of Braunschweig, Germany. He was a visiting professor at Twente University in Enschede, The Netherlands (1993), at the University of Arizona in Tucson, Arizona, U.S.A (1996) and at École Centrale de Lille, France (2001, 2002). Since 2008 he is also an Associate Professor of Electrical Engineering and Information Technology at the University of Dubrovnik, Croatia.

His main scientific interests include modelling and simulation methodologies for multidisciplinary systems, especially Bond Graph based as well as object oriented modelling; modelling, simulation, control and diagnosis of mechatronic systems, modelling languages, software design for modelling and simulation of continuous, as well as hybrid systems, scientific computing, numerical algorithms and software design for (parallel) continuous system simulation.

Dr. Borutzky has published extensively in major international conferences on Modelling and Simulation and in refereed scientific journals. He is the author of a Springer monograph on Bond Graph Modelling and the editor and a co-author of a Springer compilation text on Bond Graph Modelling of Engineering Systems with contributions from experts in various fields from all over the world. He was also the guest editor and a cou-author of two special journal issues on bond graph modelling.

Since 1990 he has served in many international scientific conferences on Modelling and Simulation in various capacities, in 2005 as Assistant General Chair of the European Conference on Modelling and Simulation (EMCS) in Riga, Latvia and as General Chair of the ECMS 2006 in Sankt Augustin, Germany. In 2009, he was an invited speaker of the IASTED Conference on Modelling, Simulation and Identification (MSI 2009) in Beijing, China and in 2010 he was one of the invited keynote speakers to the Conference on Power Control and Optimisation (PCO 2010) in Kuching, Malaysia. He has also served as an invited external examiner of Ph.D. theses in France, India and Pakistan.

Dr. Borutzky is a member of ASIM, a member of the IASTED Technical Committee on Modelling and Simulation and a senior member of SCS. During the 2004-2006 biennium he served on the SCS Board of Directors. From 2005 to 2007 he served on the Board of the European Council for Modelling and Simulation. He is also active as a member of the Editorial Board of some major modelling and simulation related journals. Currently, he is again an Associate Editor of the journal Simulation: Transactions of the Society for Modeling and Simulation International.

E-mail address: wolfgang.borutzky@h-brs.de, web-page: http://www2.inf.fh-brs.de/~wborut2m/

A MODEL FOR DYNAMIC FEED-FORWARD CONTROL
OF A SEMI-ACTIVE DAMPER

Andreas Unger[a], Enrico Pellegrini[a], Kay-Uwe Henning[a], Boris Lohmann[a]

[a]Institute of Automatic Control, Fakultät Maschinenewesen,
Technische Universität München, 85748 Garching, Germany

[a]{andreas.unger, lohmann}@tum.de, {enrico.pellegrini, kay-uwe.henning}@mytum.de

ABSTRACT
A new nonlinear model for a semi-active damper which describes a broad variety of hysteresis effects is presented in this paper. The novel part of the model is that the few parameters can be identified easily and cost effective for the automotive industry. Despite the few amounts of parameters, the validation of the model shows, that it describes the behavior of a real semi-active damper with strong hysteresis effects very well. Additionally, the influence of the unknown parameters is discussed in a detailed parameter study. In order to consider the hysteresis in the current control, the model is embedded in a dynamic feed-forward control structure.

Keywords: Damper model, hysteresis, dynamic feed-forward control

1. INTRODUCTION
The fundamental objectives of a car suspension are the isolation from vibrations introduced by the road irregularities and the improvement of road handling by means of a spring and a damper element. The suspension system supports the weight of the vehicle, provides directional control during handling maneuvers and provides effective isolation of passengers from road disturbances. These goals are generally at odds, so that the tuning of parameters in the suspension design involves finding a compromise.

The limitations of passive suspension can be enhanced by mechatronic systems, which can ease the conflict of the objectives ride comfort, ride safety and limited suspension deflection. In the last years, fully active suspension systems have been intensively studied. However, because of high costs and high energy demand their application in production vehicles is limited. Instead, because of the relatively low requested power, semi-active dampers are primarily integrated, which offer performance advantages over passive devices, see e.g. Ahmadian (1999). These suspensions feature "smart" shock absorber, which can vary the damping characteristic depending on the control strategy, mainly skyhook based comfort or handling oriented control laws. The generated force follows the passivity constraints, thus no energy can be introduced into the system. For more information and an overview on the semi-active control de-

sign refer to Guglielmino et al. (2008) and Savaresi et al. (2010) and the references therein.

Because of the complex damper mechanical construction and the switching elements in the valve, the behavior is highly nonlinear (see Duym and Reybrouck (1998) and Savaresi et al. (2010)). To exploit the potential of modern semi-active dampers a detailed damper model based on the static characteristics is desired, in order to be able to incorporate static and dynamical effects, such as hysteresis. In some works an approximation of the damper dynamics has been considered introducing first order lag elements (see e.g. Koch et al. (2010)). In the literature several models have been proposed to capture the hysteresis effects: A survey is given in Visintin (1994) and in Sain et al. (1997). To damper modeling purposes the Bouc-Wen model is frequently adopted. More details can be found in Guglielmino et al. (2008) and in Sain et al. (1997) and the references therein. A review of several idealized mechanical models for electrorheological and magnetorheological dampers based on a Bouc-Wen model is presented in Spencer et al. (1997). The models found in the literature are not able to reproduce the behavior of the twin tube hydraulic semi-active damper with internal switching valve.

The aim of this work is to model the adopted electromechanical device with strong hysteresis effects due to the design specifications and the interaction between fluid and moving mechanical components (valve). The parameter identification of the model is kept easy and only the measurement procedure adopted to obtain the static characteristics are utilized in order to produce no additional costs in the development process of the automotive industry. A model which fulfills these requirements is presented and the parameter influence on the model behavior and on the hysteresis shaping is discussed.

The remainder of the paper is organized as follows: The physical effects due to the fluid dynamics, the electromechanical valve effects and the construction characteristics are introduced in Section 2. In Section 3 the static damper characteristics are presented and compared to the measured damper forces. Based on the resulting insights the hysteresis model is presented in the same Section together with the validation of the damper model. In Section 4 a model based force tracking control in or-

der to reach the desired force considering the nonlinearities is presented. In Section 5 the measurements results are discussed.

2. PHYSICAL EFFECTS CAUSING THE NON-LINEAR BEHAVIOR

In this Section, according to Duym and Reybrouck (1998), Lang and Segel (1981) and White (1986), some hydro-mechanical and thermodynamical aspects, which are able to modify the damper response, are considered. According to models of physical phenomena already described in the literature, a damper model is presented, which considers the major physical aspects in order to explain the measurement data which is shown in the next Section. Comparing the measurement results to the model output, the model performance is considered as basis for decision making of introducing or overcoming a physical effect. Aspects like model costs and complexity, time calculation needed for the determination of parameters, computer memory requirements and numerical problems in the simulation have been considered and thus only relevant effects are included in the modeling.

A common way to describe the behavior of a damper is the static force velocity characteristic which is shown in Figure 1. But as it can be seen in measurements (see e.g. Figure 3), the static relation is not sufficient to completely describe the coherence between the relative velocity and the damping force. In order to reproduce the real damping behavior, the physical characteristics of the fluid are taken into account. The oil flowing through the valve on the piston rod and building up pressure in the compression and rebound chambers is considered compressible and its impact can be modeled by a spring with linear stiffness. The oil compressibility leads to a lag of the pressure build-up in the tube and a phase loss of the damper force especially visible in the high range of velocity measurements.
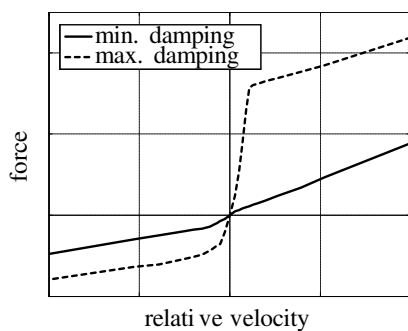

Figure 1: Static damper characteristic

As already analyzed in Duym and Reybrouck (1998) and Lang and Segel (1981), due to the damper design materials, the mechanical components are not perfectly rigid. Above all the rigidity of the cylinder walls has a direct effect on the oil pressure in the chambers. Their effect is similar to the oil compressibility and can be modeled as well as a spring. The two effects are considered together and will appear in the model as a spring component.

In addition to the mechanical friction also a hydro-dynamical friction exits, which appears in elements moving in a viscous medium. This effect is known as Striebeck friction, Beitz and Küttner (1994). This kind of friction probably emerges between the valve mounted on the damper rod and the cylinder. Considering a standard Striebeck friction curve it can be observed, that the friction has a nonlinear dependency on the velocity and on the oil viscosity. Therefore, the amount of required parameters for modeling the friction is high and taking the minor effect on the damping force into account, friction is neglected for the presented model. Additionally, it has to be noted that the viscosity and density depend on the temperature in the chambers, which either has to be measured or estimated in order to consider these dependencies. As this would require high costs and efforts, it is assumed that the damper is heated up at a constant working temperature and thus temperature effects are also omitted.

The complete behavior of the damper depends not only on the characteristics of the fluid and the gas, but also on the mechanical switching element. By adjusting the rod valve, in order to change the damping coefficient of the damper, it can be noticed that the force behavior in this case also depends on the damper state. Defining the switching time, the time lapse in which 90% of the final force is reached it can be noticed that it varies between 10ms and 30ms. According to Heißing and Ersoy (2007), it depends on the absolute value of the damper velocities, on the rebound or compression direction as well as on the switching direction (from soft to hard or vice versa). This effect has to be implemented in the damper model because it describes the relation between the damper velocity (which is generally not the velocity of the oil through the valve assembly), the valve adjustment point (that means the current applied to the valve) and the resulting damping force. This phenomenon is strong nonlinear and very difficult to describe. In the presented model the switching time is approximated with a first order lag element for both directions and both switching directions.

Moreover, the structure of the inner valve assembly, which consists of different check valves, intake valves, port restrictions and blow-off valves (e.g. Duym and Reybrouck (1998)) is strongly nonlinear. The rod and the base valve assemblies are responsible for the static force velocity characteristic of the damper (Figure 1). In order to avoid switching noises of the valves, the pressure gradient is reduced by a damping element, which leads to a delay time while opening and closing. This effect causes nontypical hysteresis effects and peaks in the damper force, especially at high frequencies or high velocities. In order to improve ride and acoustic comfort, a special mount is attached between the damper and the chassis. This element is a flexible compound of metal and rubber and is designed in order to isolate the chassis from high frequency vibration induced by the damper (Heißing and Ersoy (2007)). However, this flexibility causes an additional stiffness, which also affects the damper force. As the measure-

ment data used in this work has been obtained without damper mount, it is not considered in the model but it can easily be approximated by adjusting the upper stiffness ratio. The effect of the damper mount is that a spring-mass mechanical system is generated due to the fact that the damper rod is directly connected to the top mount. By non-optimal design parameters the system can be excited with high frequency vibration, which can deteriorate the damping action. Anyway, these frequencies are not in the range of interest, in which the suspension works. Therefore, the effect is not implemented in the proposed model.

Serveral major physical effects have been presented in this Section. However, there exist many other physical effects affecting the damper force (see e.g. Lang and Segel (1981), Reimpell and Stoll (1989)).

3. THE NONLINEAR DAMPER MODEL
3.1. Model equations

Figure 2 shows a semi-physical model for a semi-active damper, where the major physical effects for hysteresis that have been discussed in the previous Section, are considered. The stiffness of the cylinder walls, the damper mount and the fluid compressibility are substituted by a stiffness k and a damping factor b_l. This spring-damper element is connected in series with the main damper b_{nl}, where the static, nonlinear force-velocity characteristic (Figure 1) is considered.
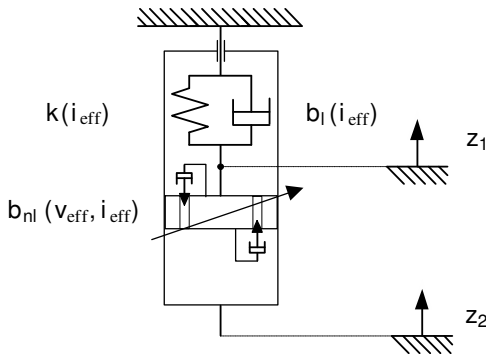

Figure 2: The nonlinear damper model

In order to take the opening time of the check valves into account, two relative velocities have to be distinguished. The velocity between the piston and the rod of the damper is generally used for calculating the damper force and will be denoted as $v_{pr} = \dot{z}_2 - \dot{z}_1$ in this work. As the check valves do not open instantly, the effective relative velocity on the main damper v_{eff} is slightly different. A first order lag element

$$v_{eff} = \frac{1}{T_{cv}s+1} v_{pr} \qquad (1)$$

is suggested in order to approximate the coherence of these two velocities, with T_{cv} being the time constant of the check valves. This effect is symbolized by the dampers acting on the check valves in Figure 2. Also the

dynamics of the current can be considered by a first order lag element

$$i_{eff} = \frac{1}{T_{cur}s+1} i_{des} , \qquad (2)$$

with the time constant T_{cur}.

Balancing the forces and summarizing these equations leads to the differential equations for the nonlinear damper model

$$\dot{z}_1 = -\frac{k}{b_l+b_{nl}} z_1 + \frac{b_{nl}}{b_l+b_{nl}} \dot{z}_2 \qquad (3)$$

$$\dot{i}_{eff} = -\frac{1}{T_{cur}} i_{eff} + \frac{1}{T_{cur}} i_{des} \qquad (4)$$

$$\dot{v}_{eff} = -\frac{1}{T_{cv}} v_{eff} + \frac{k}{T_{cv}(b_l+b_{nl})} z_1 + \frac{b_l}{T_{cv}(b_l+b_{nl})} \dot{z}_2 \qquad (5)$$

where $\dot{z}_2$ and i_{des} are the inputs of the system and the damping force

$$F_{damp} = \frac{kb_{nl}}{b_l+b_{nl}} z_1 + \frac{b_l b_{nl}}{b_l+b_{nl}} \dot{z}_2 \qquad (6)$$

is considered as the output of the system.

3.2. Parameter identification

As it is shown in Section 3.4, the hysteresis behavior can be influenced by the model parameters k, b_l, b_{nl} and T_{cur}. In order to guarantee the highest flexibility of the model for different valve currents and therefore hysteresis types, it is suggested to introduce a dependency of the named parameters on the effective current i_{eff}.

The damping coefficient of the main damper b_{nl} additionally depends on the effective relative velocity v_{eff} and is determined using the static force velocity characteristic of the damper (Figure 1). Therefore, it is considered to be known. The rise time of the current T_{cur} depends on the inductivity, the resistance and the controller parameters of the electrical circuit (Savaresi et al. (2010)). It can be either calculated if the parameters are known or read from a measured step response of the electrical system. The remaining parameters k, b_l and T_{cv} are estimated by optimization such that the difference between the output of the model, i.e. the damping force, O_{model} and the measured output O_{meas}, which can be expressed in the cost function

$$J(k,b_l,T_{cv}) = \frac{1}{n}\sum_{j=1}^{n}(O_{model}(t_j,\dot{z}_2,i_{des},k,b_l,T_{cv}) - O_{meas}(t_j))$$

$$(7)$$

is minimized for a fixed valve current i_{des} and a given velocity input $\dot{z}_2$ for all considered time steps t_j. The model parameters are then interpolated linearly for currents which have not been considered in the parametrization process.

For a cheap parametrization of the model in industrial application, it is desirable to use no additional measurements than those, which are used to obtain the static

characteristic of the damper. The static behavior is determined by exciting the damper by a sinusoidal signal with a fixed stroke amplitude and several predefined velocity amplitudes for fixed valve currents. The damping force is measured for the point, where the velocity reaches its maximum what leads to one point in the force velocity diagram (Figure 1). For further information on the routine the reader is referred to Reimpell and Stoll (1989). In order to excite the hysteresis behavior of the damper, it is suggested to modify the stroke amplitude from 0.05 m to 0.01 m, what can be achieved easily by today's damper test rigs. The values for the amplitudes and the resulting frequencies are listed in Table 1.

Table 1: Input signal for parametrization according to Reimpell and Stoll (1989) and modified stroke amplitude

Stroke[m]	0.01				
Velocity [m/s]	0.052	0.131	0.262	0.393	0.524
Frequency[Hz]	0.83	2.08	4.17	6.25	8.34

By using the whole signals for the minimization of (7) instead of only one point of each data set, the parameters k, b_l and T_{cv} can be estimated. As this measurement data is sufficient, the requirement of producing no additional costs for the parametrization is fulfilled by this procedure.

3.3. Model validation
The model parameters are identified for the semi-active damper using the input signals according to Table 1. In order to validate the parameters, the model output is compared to measurements for different amplitudes and frequencies of the sinusoidal input signal (Table 2).

Table 2: Input signal for validation

Stroke[m]	0.005				
Velocity [m/s]	0.052	0.131	0.262	0.393	0.524
Frequency[Hz]	1.66	4.17	8.34	12.51	16.67

The result of this comparison can be seen in Figure 3 for the maximum damping, in Figure 4 for medium damping and in Figure 5 for the softest damping.
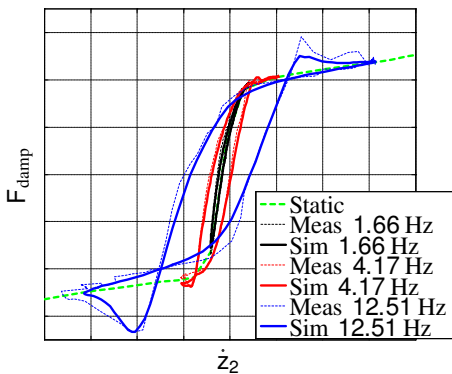

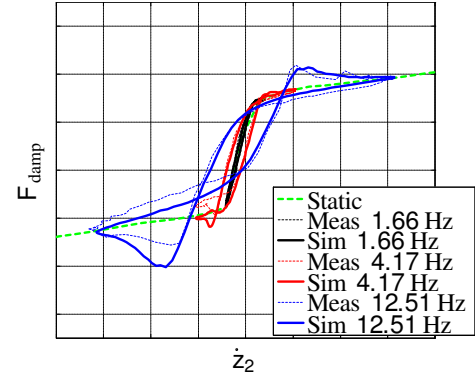Figure 3: Fit to experimental data, $i_{des} = 0A$

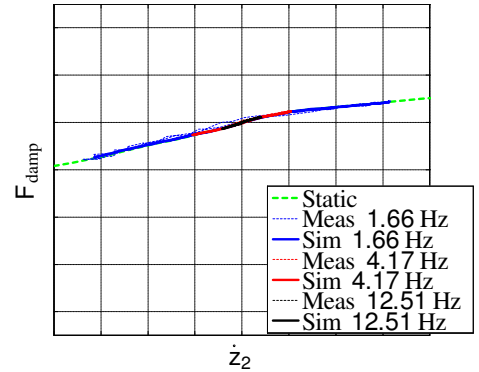Figure 4: Fit to experimental data, $i_{des} = 0.9A$

Figure 5: Fit to experimental data, $i_{des} = 1.8A$

Table 3: Fit values of the model in [%]

Frequency [Hz]	1.66	4.17	8.34	12.51	16.67
0A	84.8	95.4	91.7	91.1	86.4
1A	90.5	86.9	90.2	90.7	87.9
1.8A	85.1	92.6	90.9	90.5	86.9

The model output is compared in regions, where the damping forces nearly match the static characteristic (low velocities and/or soft damping) and also for increased velocities and damping forces, where minor (Figure 3, 4.17 Hz) and major (Figure 3, 12.51 Hz) hysteretic effects appear. It can be seen, that, in all cases, the proposed model matches the measurement data very well. In order to quantify the validity of the model, the fit value

$$fit = 1 - \frac{\left\| F_{meas} - F_{sim} \right\|_{rms}}{\left\| F_{sim} \right\|_{rms}} \qquad (8)$$

for each considered frequency and current is given in Table 3, whereas a mean fit value of 89.4 % can be calculated.

3.4. Parameter variation
Depending on the set of parameters, a broad variety of hysteretic effects can be described using the new damper model. Figure 6 shows simulation results for three different parameter sets which result in minimal hysteresis, common hysteresis and hysteresis with an additional loop.
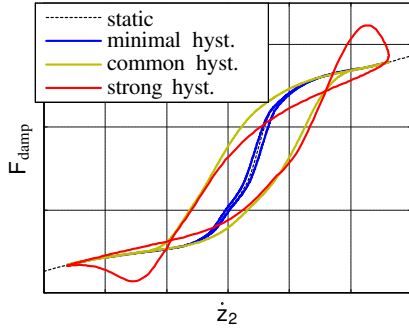
Figure 5: Model possibilities at 0 A and 16.67 Hz

The corresponding parameters are listed in Table 4 and it can be seen, that the hysteretic effects can be increased by increasing the response time of the check valves and reducing the stiffness and damping of the upper mount.

Table 4: Parameters for Figure 6

	k	b_l	T_{cv}
minimal hysteresis	9e7	9e4	2e-5
common hysteresis	5e6	5e3	1e-4
strong hysteresis	4e6	2e3	2.2e-3

In order to understand how the parameters k, b_l and T_{cv} affect the hysteresis behavior in the force velocity diagram, parameter variations have been made and the result is depicted in Figure 7. It can be seen, that the parameters k and b_l have an effect on the general hysteresis shape of the curve. Rising values of k or b_l lead to a smaller area which is enclosed by the damping curve. This seems plausible, because increasing stiffness and damping parameters stiffen the spring-damper element, which is connected in series with the nonlinear static damping element (Figure 2) and therefore, the influence of the static damper becomes higher. Most of the effects are influenced equally by k and b_l, except for the loop in the compression case, where slight differences can be seen. The size of both loops can be manipulated by the time constant of the check valves T_{cv}. Because the influence of the first order lag element (1) is getting smaller with lower time constants, the loop size decreases for faster response times.
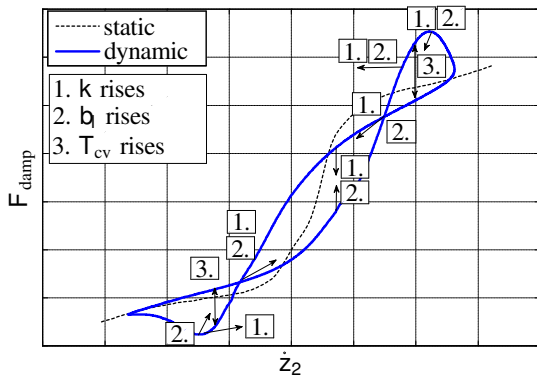

Figure 7: Influence of the model parameters

4. MODEL BASED CURRENT CONTROL

4.1. Dynamic feed-forward control

As seen in the previous Sections, the inversion of the static damping characteristics for a given desired damping force may lead to an improper resulting force due to the hysteresis of the damper. The previously described hysteresis model can be used to consider this nonlinear, dynamic behavior for the calculation of the appropriate valve current by embedding the model in a control structure, based on a dynamic feed-forward control structure, which has been proposed in Franklin et al. (2010) and is shown in Figure 8. It can be seen, that the static inversion of the damping characteristic, which is calculated in the feed-forward block in Figure 8, is extended by a controller, where the current is adjusted whenever an error between the desired and the calculated damping force occurs. The sign of the calculated force is multiplied with the force error because a positive error must lead to an increased damping, regardless of whether the damper is in compression or rebound. The corrected valve current is then applied to the real damper leading to a lower error between the desired force and the resulting force of the real damper. Of course, for proper tracking, it is required that the model matches the behavior of the real damper and that the dynamic feed-forward control is fast enough, i.e. the controller gain is high. A good performance can already be achieved by a proportional gain, which is also used in the present implementation. The gain is determined by optimization using a genetic algorithm such that the error between the desired force and the damping Force F_{damp} is minimized.
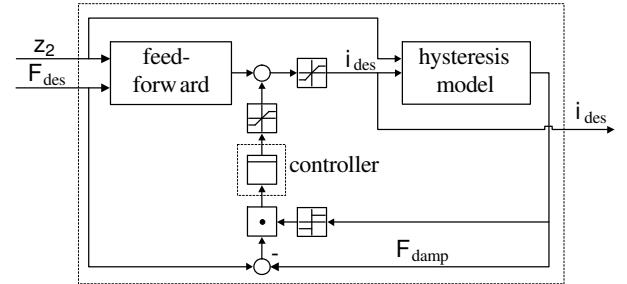

Figure 8: Dynamic feed-forward control structure

4.2. Numerical results

The dynamic feed-forward control approach is compared to the static calculation of the damper current using the hysteresis damper model for the simulation of the real damper. The desired force has been chosen to be proportional to the relative velocity, which has been measured from a real road profile. It is noted, that of course only desired forces, which are between the minimum and maximum characteristic of the damper, can be tracked. The results are shown in Figures 9 and 10 and it can be seen, that the error between the damper force and the desired force can be reduced significantly using dynamic feed-forward control. Especially the force peaks, which lead to high vertical accelerations

and therefore are sensed most by the passenger, are tracked more accurately. The root mean square value for the force error is reduced by 38 % from 125.5 N to 77.2 N by using dynamic feed-forward control instead of the static inversion for this example. In order to estimate the robustness of the controlled system, the parameters of the simulated real damper have been varied by 30 %. The results show, that there is still an improvement of the rms value for the force error of 18.9 % for the variation of k , b_l , b_{nl} and T_{cv} in the positive direction and 5.6 % in the negative direction. A more detailed analysis of the robustness will be part of future work.
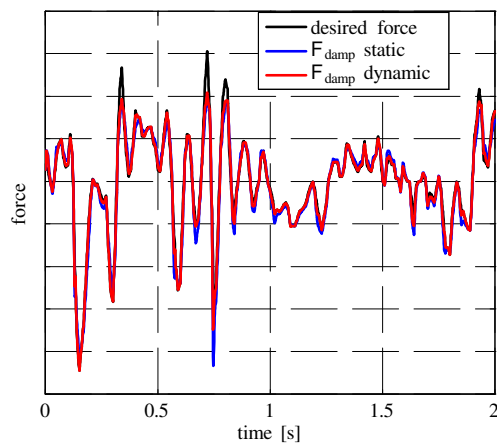
Figure 9: Comparison of absolute forces between static and dynamic current calculation
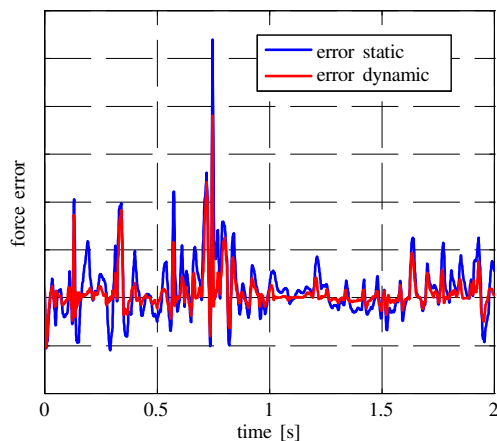

Figure 10: Comparison of force error between static and dynamic current calculation

5. CONCLUSION

In this paper, a novel hysteresis model for a semi-active damper has been presented, which can be used to describe a broad spectrum of dampers. The advantage of the model is that the few unknown parameters can be identified using only measurements, which are already made in the automotive industry in order to obtain the static characteristic of the damper. The identification process has been shown for a real damper with strong hysteresis effects and the validation showed that the model matches the measurement data very well. Fur-

thermore, a study on the parameter variation has been presented and the influence on each parameter on the shape of the hysteresis curves has been discussed. Finally, it has been shown that the model can be used in order to improve the force tracking of the semi-active damper by embedding it into a dynamic feed-forward control structure.

REFERENCES

Ahmadian, M., 1999. On isolation proprieties of semiactive dampers. *Journal of Vibration and Control*, 5, 217-232.

Beitz, W., Küttner, K.H., 1994. *Dubbel Taschenbuch für den Maschinenbau.* Springer-Verlag, Berlin Heidelberg.

Duym, S., Reybrouck, K.,1998. Physical characterization of nonlinear shock absorber dynamics. *European Journal of Mechanical and environmental Engineering*, 43 (4), 181-188.

Franklin, G.F., Powel, J.D., Emami-Naeini, A., 2010. *Feedback Control of Dynamic Systems.* Pearson Prentice Hall, 6th Edition.

Guglielmino, E., Sireteanu, T., Stammers, C., Ghita, G., Giuclea, M., 2008. *Semi-active Suspension Control.* Springer-Verlag, Berlin.

Heißing, B., Ersoy, M., 2007. *Fahrwerkhandbuch.* ATZ/MTZ-Fachbuch.

Koch, G., Spirk, S., 2010, Reference model based adaptive control of a hybrid suspension system, *Proceedings of the 6th IFAC Symposium in Advances in Automotive Control*, 07/2010, Munich Germany.

Lang, H.H., Segel, L., 1981, The mechanics of automotive hydraulic dampers at high stroking frequencies, In: 7th *IAVSD Symp. Dynamics of Vehicles on Roads and Tracks.*

Reimpell, J., Stoll, H., 1989. *Fahrwerkstechnik: Stoß- und Schwingungsdämpfer.* Vogel Verlag Würzburg

Sain, P.M., Sain, M.K., Spencer, B., 1997, Models for hysteresis and application to structural control, *Proceedings of the 1997 American Control Conference*, (1) 16-20, 4-6 Jun,USA.

Savaresi, M., Poussot-Vassalt, C., Spelta, C., Sename, O., Dugard, L., 2010. *Semi-Active Suspension Control Design for Vehicles,* Elsevier Ltd.

Spencer, B.F., Dyke, S.J., Sain, M.K., Carlson, J.D., 1997. Phenomenological model of a magnetorheological damper. *Journal of Engineering Mechanics, ASCE* 123, 230-238.

Visintin, A., 1994. *Differential Models of Hysteresis,* Springer Verlag, Berlin-Heidelberg.

White, F., 1986. Fluid Mechanics*,* McGraw-Hill Book Company

Autofocusing Slow-moving Objects Sensed by an Airborne Synthetic Aperture Radar SAR

José L. Pasciaroni[†] [‡]   Juan E. Cousseau[‡]   Nélida B. Gálvez[†]

‡*DIEC - CONICET - Universidad Nacional del Sur Av. Alem 1253 - (8000) Bahía Blanca*
†*SIAG - Base Naval Puerto Belgrano Av. Colon s/n (8111) Punta Alta*
pasciaro@uns.edu.ar; ngalvez@uns.edu.ar; jcousseau@uns.edu.ar
ARGENTINA

Abstract– During the processing of synthetic aperture radar (SAR) images, unknown-moving objects cause phase modulations in their phase histories. Depending on the motion, such modulations cause undesirable image distortions such as blurring and object-displacement artifacts. This paper discusses the estimation of time-frequency representations of non-stationary signals, by means of autoregressive moving-average (ARMA) model with time-dependent coefficients. This estimate will determine the object's motion law and consequently the possibility of reconstructing and plotting the true trajectory.

Keywords– Syntethic aperture radar, Time-varying auto-regressive moving average models, Kalman filters, Phase estimation.

1. Introduction

In this section we describe some basics of the SAR principle. As shown in Figure 1, a pulsed airborne radar moves in the indicated flight path, transmitting pulses and receiving echoes reflected from the scene. Suppose the cycle starts at point A where the radar transmits a pulse, then others at x_2, x_3,…,x_{n-1} and so on until reaching point B in x_n. What defines the interval size between A and B is the first and last contact of beam with the object (indicated in the figure as range-A and range-B). During that period the object is being iluminated by the radar lobe, receiving x_n pulses. This distance is what is called synthetic aperture. Moreover, the returns of each pulse which contain range information (fast-time), are stored by columns in the SAR data matrix. The cells of each column is called range-increments, where his size is defined by the pulse-width. Thus, in cell C_{11} is stored the farther return in distance and the nearest in cell C_{m1}. The same applies to the transmitted pulse in x_2, but their returns are now stored in column 2 (cells C_{12} to C_{m2}).
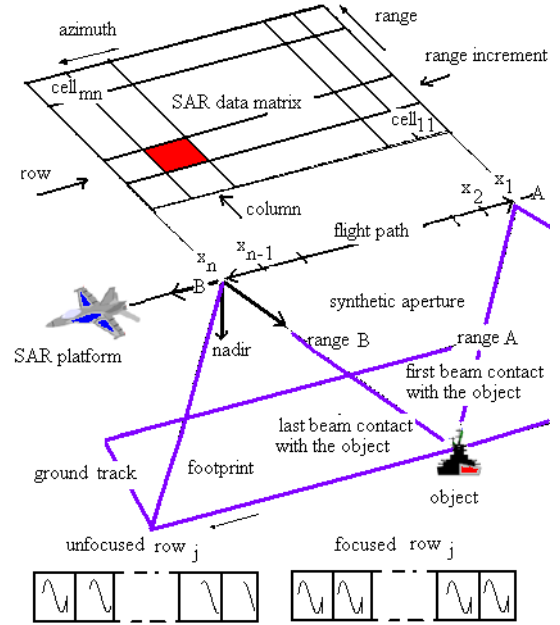

Figure 1: SAR Geometry

This is repeated until the column n (cells C_{1n} to C_{mn}) is reached. Also, the typical assumption is that the platform moves with a constant and known speed and the escene objects is static, which ultimately implies the existence of a Doppler effect. Row cells (slow-time) store "versions" of the returns but shifted in phase (phase history), as shown for the case of row_j. Optimized algorithms focus the distributed energy of the raw data for all scene locations simultaneously. However, if marine or ground objects have an unknown motion they will cause a unknown phase modulation in their phase histories, precisely because of the unknown relative motion between sensor and object (section 2.). In turn, these modulations cause SAR image distortions such as blurring and smearing (Sparr and Krane 2003).

In order to develop this work (Figure 2), we consider: 1) the time serie (TS), comprised of the sequencing radar returns, is modeled by a time-varying auto-

regressive moving average (TV-$ARMA$) model, 2) the $TV-ARMA$ coefficients are updated by a Kalman filter (KF), 3) using the $ARMA$ coefficients updated at each iteration we obtain the instantaneous estimated power spectral density (PSD), 4) the maximum value of PSD provides us with the instantaneous frequency IF, 5) finally, the last objective is to analyze and to compare the performance of a phase $ARMA-KF$ estimator, operating in a non-stationary simulated environment which has a moving object, using as framework the WignerVille distribution WVD.

2. SAR simulated scenario whith a dynamic object

Here we define the scenario that contains an object with rotational movement around a fixed point as shown in Figure 3, which considers the projection of this movement over the line of sight (LOS).
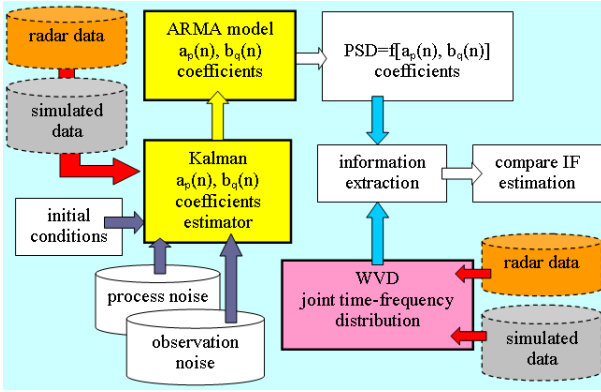

Figure 2: $ARMA-KF$ and WVD processes block diagram

Let the freedom to choose the location of the reference system so that: 1) the radar platform is placed at $[0,0,h]$, 2) the object has a initial coordinates $[x_0, y_0, 0]$ and a absolute position $[x(n), y(n), 0)]$, determined by $\mathbf{R}(0)$ plus the phasor $\mathbf{r}$ (whose initial phase is φ_0), 3) The relative coordinates are the projections of $\mathbf{r}$ on the x and y axis, i.e. with a oscillatory variation proportional to $\sin(n\omega_m)$ and $\cos(n\omega_m)$ respectively. The term ω_m represent the object rotational frequency. The magnitude of range $\mathbf{R}(n)$ measured from the radar to the object is

$$|R(n)| = \sqrt{(x(n) - v_{pl}n)^2 + h^2 + y(n)^2} \qquad (1)$$

The Figure 4 shows the successive positions of the platform. As we have stated, it is necessary to apply a phase correction (applying a rotation to the appropriate phasor) to each return received. The error and therefore the correction phase, is directly proportional to d_j^2, the square of the j-th instantaneous position relative to the center of the aperture. But more important here is that it is inversely proportional to $|\mathbf{R}(n)|^2$. Namely, the correction is (Stimson 1998)
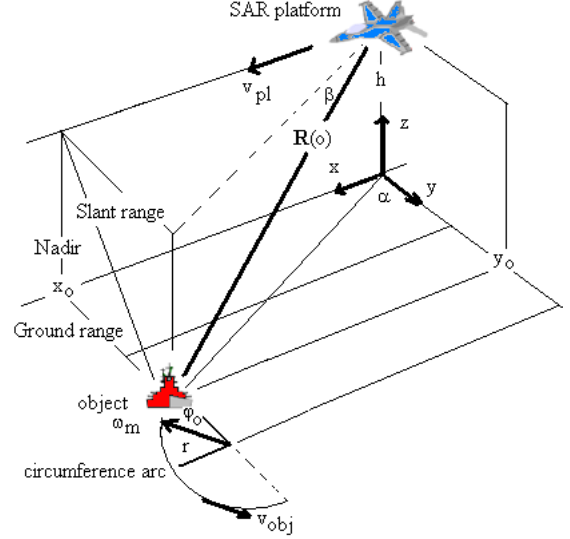

Figure 3: Scenario and coordinate system

$$\varphi_j \approx \frac{2\pi}{\lambda|\mathbf{R}(n)|^2}d_j^2 \qquad (2)$$

It is obvious that lack of knowledge of true coordinates $x(n), y(n)$ of the object, preventing an effectively phase correction. Let's see in the next section how this scenario affects the signal model.

3. Physical signal model

As we will see in the section 4. this scenario allows us to form a TS $\{x(n)\}$ of the trajectory. Then, each transmitted linear frequency modulated LFM pulse is expressed by the complex exponential $s_t(\tau) = g(\tau)e^{j2\pi(\varphi_0 + f_0\tau + \alpha\tau^2)}$, where τ is the pulse round-trip time, $g(\tau)$ is a square pulse of $T_w[\mu sec]$ wide, α the slope of the modulating frequency, φ_0 an arbitrary initial phase, f_0 the carrier frequency, so $(\varphi_0 + f_0\tau + \alpha\tau^2)$ is the phase $\varphi(\tau)$ [rad]. The signal $s_t(\tau)$ is valid in the interval $[-T_w/2, T_w/2]$. Therefore, the received echo from an object at the distance R will be

$$s_t(\tau) = g(\tau)e^{j2\pi(\varphi_0 + f_0(\tau - 2R/c) + \alpha(\tau - 2R/c)^2)} \qquad (3)$$

In this work we are only interested in the resulting signal at the end of the reception chain, i.e. after pulse compression processing. As the SAR platform moves along a hypothetical circumference arc flight track, the radar emits and receives pulses. Considering one object of interest, the difference between pulses will be determined by the attenuation of the signal from the path loss and the antenna pattern as well as by the change in distance to the object. Based on this, the pulse compressed signal before SAR processing is

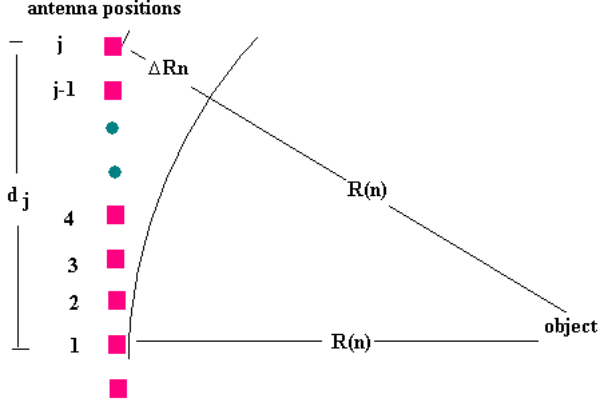$$s_{pc}(n,\tau) = T_w(1 - \frac{|2\mathbf{R}(n)/c - \tau|}{T_w})e^{j2\pi f_0(2\mathbf{R}(n)/c - \tau)}$$

Figure 4: Focused Array

Figure 5: PSD versus time-frequency

$$\left[sinc(2\pi\alpha T_w(2\mathbf{R}(n)/c - \tau)(1 - \frac{|2\mathbf{R}(n)/c - \tau|}{T_w}) \right] \quad (4)$$

function of n, and τ and valid in the interval $[2|\mathbf{R}(n)|/c - T_w, 2|\mathbf{R}(n)|/c + T_w]$. Using the expression above, renamed as $x(n)$ we generate a set $s_{pc}(n, \tau)$ $n = 0, \ldots, N - 1$, where each element is a complex-value signal per pulse. N is the number of pulses emitted during the aperture time. We will use that set to test the proposed $ARMA - KF$ methods and reference WVD time-frequency representation.

4. Building a mathematical model for TS

Now we will develop a mathematical model of the received SAR signal

$$x(n) = -\sum_{k=1}^{p} a_k(n)x(n-k) + \sum_{k=0}^{q} b_k(n)w(n-k) + n(n) \quad (5)$$

where $x(n)$ is considered as one that describes the system and the measurement process, with coefficients $\{a_k, b_k\}$ and $\mathbf{w}(n)$ the clutter noise input driving sequence, independent w.r.t. past values $x(n - k)$. The values of p and q are the order of autoregressive $ARMA$ model. Now, if we consider that the TS samples has varying spectral properties, it is also natural to assume the coefficients to be time dependent. So, it is more realistic: 1) to model the a_k, b_k with a degree of randomness i.e. $a_k(n + 1) = a_k(n) + v_a(n)$ and $b_k(n+1) = b_k(n) + v_b(n)$, where the zero-$th$ order random processes $\mathbf{v}(n) = \{v_a(n), v_b(n)\}$ can be interpreted as the uncertainty of the prediction of the next coefficient value and 2) accept the measurements as being noisy. To denote this new condition, the $x(n)$ expression is modified by the adittion of the observation noise $\mathbf{n}(n)$, as independent of $AWGN$ coefficient noise $\mathbf{v}(n)$ and clutter $\mathbf{w}(n)$ processes.

5. TS expressed as state-space model (SSM)

As we saw, $\{x(n)\}$ is an uni-variate TS represented as an $ARMA$ model with coefficient $a_k(n)$, $b_k(n)$. Now
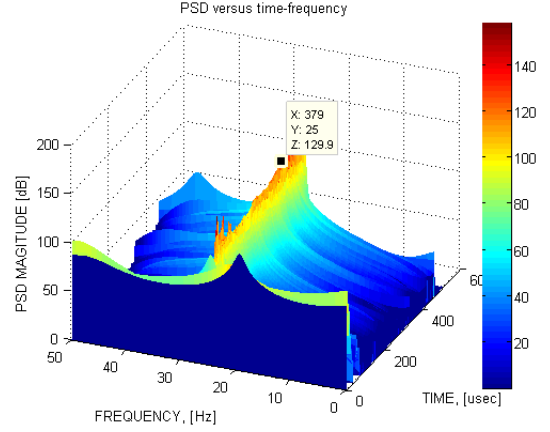
we pose our problem of spectral estimation in SSM term. For this, let the $ARMA$ coefficients represent the state of the system

$$\mathbf{s}(n + 1) = \mathbf{s}(n) + \mathbf{v}(n) \quad (6)$$

where $\mathbf{s}(n) = [a_1(n), \ldots, a_p(n), b_0(n), \ldots, b_q(n)]$ is the (m)-dimension $(m=p+q+1)$, unobservable state vector and $\mathbf{v}^T(n) = [v_1(n), \ldots, v_{p+q}(n)]$ is the $(m - 1)$-dimension noise vector, that as mentioned above, models the unknown statistics of the coefficients. So, get the state vector $\mathbf{s}(n)$ is to get the TV $ARMA$ coefficients. On the other hand, if we define the (m)-dimension vector of returns and noise process as $\mathbf{C}(n) = \{-x(n - 1), \ldots, -x(n - p), w(n), w(n - 1), \ldots, w(n-q)\}$, we can write the measurement equation as

$$x(n) = \mathbf{C}(n)\mathbf{s}(n) + n(n) \quad (7)$$

6. State estimation via the AKF

Comparing equations (6) and (7) with the conventional SSM model equations, we can conclude that, in the problem under study: 1) the transition matrix Φ is the identity matrix I, 2) the measurement matrix H is represented by $\mathbf{C}$ and 3) the estimate of state vector $\mathbf{s}(n)$ (or $ARMA$ coefficients $\{a_k(n), b_k(n)\}$), can be gotten from the observed data $\{x(n)\}$ using a KF (Harashima, Ferrari and Sankar 1996). To illustrate this "recast", let $(p, q) = 2$, $\mathbf{C}(n) = [-x(n-1), -x(n-2), w(n), w(n-1), w(n-2)]$ and $\mathbf{s}(n) = [a_1(n), a_2(n), b_0(n), b_1(n), b_2(n)]^T$, then we can express $x(n)$ in matrix form as

$$x(n) = \begin{vmatrix} -x(n - 1) \\ -x(n - 2) \\ w(n) \\ w(n - 1) \\ w(n - 2) \end{vmatrix}^T \begin{vmatrix} a_1 \\ a_2 \\ b_0 \\ b_1 \\ b_2 \end{vmatrix} + n(n) \quad (8)$$
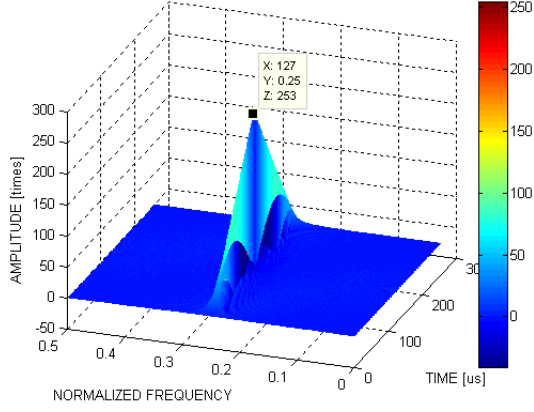
Figure 6: WVD versus time normalized-frequency

Figure 7: PSD of a signal received from a moving object

Moreover, the state $\mathbf{s}(n)$ of the system and the gain matrix $\mathbf{K}(n)$ can be estimated recursively, with the proper interpretation of stochastic characteristics of both, the input driving signal $\{\mathbf{v}(n); \sigma_v^2\}$ and measurement noise $\{\mathbf{n}(n); \sigma_n^2\}$. To apply the AKF, we also assume that $a_k(0)$ and $b_k(0)$ are Gaussian random variables.

7. Initializing the augmented KF

In this section we define the set of parameters required to initialize the AKF algorithms, considering that the only information available is the noisy measure of range represented by the TS $\mathbf{x}(n)$ and covariances of the noises $\mathbf{R}(n)$ and $\mathbf{U}(n)$, as we will see later. Moreover, we will not repeat here the recursive set of equations that make up the AKF, except those related to the initialization process. We only recall that one way of expressing an estimated $\hat{\mathbf{s}}(n)$ of $\mathbf{s}(n)$ that is as close to it as possible in a mean-squared sense is $\hat{\mathbf{s}}(n) = \mathbf{I}\hat{\mathbf{s}}(n-1) + \mathbf{K}(n)[\mathbf{x}(n) - \mathbf{CI}\hat{\mathbf{s}}(n-1)]$. We know that the recursive computation of the variable-gain matrix $\mathbf{K}(n)$ involves three matrix equations (Auger, Flandrin, Gonçalvès and Lemoine 2002)

$$
\begin{aligned}
\mathbf{K}(n) &= \mathbf{J}(n|n-1)\mathbf{C}(n)^T[\mathbf{C}(n)\mathbf{J}(n|n-1)\mathbf{C}(n)^T + \\
&+ \mathbf{R}(n)]^{-1} \\
\mathbf{J}(n|n) &= \mathbf{J}(n|n-1) - \mathbf{K}(n)\mathbf{C}(n)\mathbf{J}(n|n-1) \\
\mathbf{J}_{n+1|n} &= \mathbf{IJ}(n|n)\mathbf{I}^T + \mathbf{U}(n)
\end{aligned}
$$

where $\mathbf{J}(n|n) = \mathbf{E}[[\mathbf{s}(n) - \hat{\mathbf{s}}(n)][\mathbf{s}(n) - \hat{\mathbf{s}}(n)]^T]$ is just the covariance matrix of estimation errors. Note that in our time-varying case the matrices $\mathbf{J}(n|n)$ and $\mathbf{K}(n)$ can not be computed a priori because of the dependence of $\mathbf{C}(n)$ on the actual system measurements. Moreover, as the inspection of the Kalman gain reveals, we must consider that if the measurement noise is small, then $\mathbf{K}(n)$ will be large, i.e. a lot of credibility will be given to the measurement. In contrast,
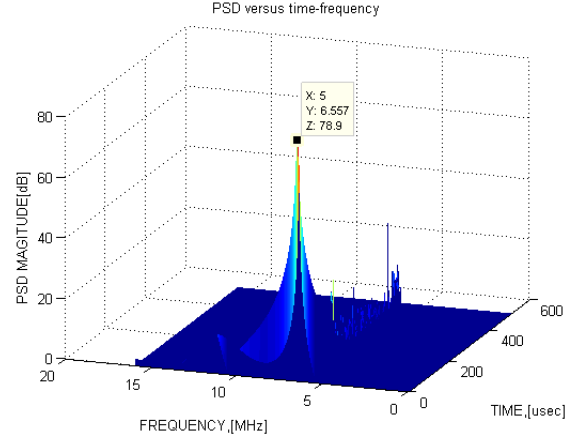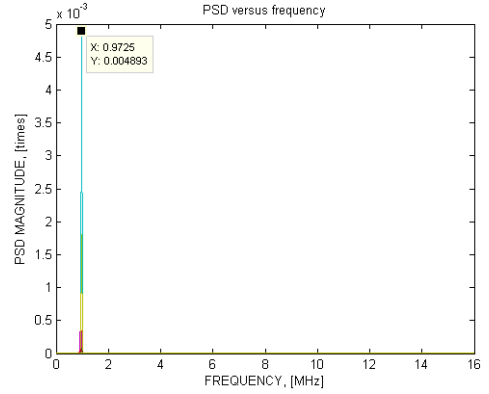

Figure 8: $ARMA - KF$ processes to radar return signal

if the measurement noise is large, then $\mathbf{K}(n)$ will be small, i.e. very little credibility will be given to the measurement, in both cases when computing the next value of $x(n)$.

Now, the key issue is to assume values for the variances σ_n^2, σ_w^2 and σ_v^2. In the case of σ_n^2, it should be assigned on the basis of our knowledge of the noise introduced by measurement sensors.

8. Phase estimation

From the moment that we have the time-varying coefficients of the $ARMA$ model, is possible to obtain the unknown phase from the PSD, expressed as a function of that coefficient a_k, b_k. The ridge-peak of this function is the instantaneous frequency.

$$
\hat{P}_x(e^{j\omega}) = \frac{|\sum_{k=0}^{q}\hat{b}_q(k)e^{-jk\omega}|^2}{|1 + \sum_{k=0}^{p}\hat{a}_p(k)e^{-jk\omega}|^2} \tag{9}
$$

9. Comparison between the $ARMA - KF$ estimator and the WVD

In this section, we compare the performance of the $ARMA - KF$ estimator against WVD. To do this we consider the following design parameters. With respect to the platform: 1) The antenna has a physical inclination (look-angle) of $76[degrees]$ 2) we assume a pulsed linear frequency modulation (LFM) transmitted signal, of $T_w = 10[\mu s]$ wide and slope of the modulating frequency $\alpha = 40[MHz/s]$, 3) a pulse repetition frequency $PRF = 500[Hz]$, frequencying a duty cycle or interval between pulses IP of 1% or $\approx 2.0[msec]$ of maximum processing-time, 4) the transmiter has a carrier frequency $f_0 = 1.275[GHz]$ ($\lambda = 0.0235[m]$), and a intermediate frequency stage FI of $25[MHz]$, where the radar video is sampled at a frequency $f_s = 100[MHz]$, with a sample interval $\Delta t = 1/f_s = 10[ns]$ and time of the sample n defined by $tn = n\Delta t$, 5) the platform operates in stripmap mode, with a spatial and azimuth resolution of $6[m]$ and $1[m]$ respectively (depend on the size of the antenna), 6) besides that exposed, the platform has a speed of $100[m/s]$ and a height of $4000[m]$ results a ground-range value of $1800[m]$ (see x_0 in section 2.), 7) moreover, from the supposed speed of the object and also from an arbitrary value of $\mathbf{d}\sin(\theta) = 50[m]$, we then obtain a simulated signal with frequency $f_m = 0.5[Hz]$, 8) finally, the number of samples N depends on the range-cell resolution (RCR), which is proportional in turn to the compressed pulse wide. If the compression ratio is 20 then the $RCR = 75[m]$ or $500[ns]$. Since $f_s = 100[MHz]$, then an acceptable minimum is $N = 64$. With respect to the $ARMA$ and kalman filter: 1) we use $p = q = 2$, 2) for the variances $\sigma_n{}^2$, $\sigma_w{}^2$ and $\sigma_v{}^2$ we consider a reception chain $SNR \geq 10[dB]$ and measurement errors of 1%, then we can use $\sigma_n{}^2 = 1000$, 3) in the case of $\sigma_w{}^2$, after an assessment of the way the $\mathbf{w}(n)$, representing the clutter, is likely to vary (Anderson and Moore 1979). Again, if the signal to clutter ratio $SCR \geq 10[dB]$, with values of radar video signal on the order of $[mV]$, we consider then $\sigma_w{}^2 = 1$, 4) moreover, for $\sigma_v{}^2$, we consider a trial an error approach method, 5) finally, we need to assume a priori initial values for each $a_k(n)$, $b_k(n)$ of these coefficients, in that case, we assume complete ignorance of them, so $a_k(0)$, $b_k(0)$. We test three case studies, all of them related to the scenario of section 2.

CASE A– Scenary with a constant frequency signal. Is a signal with $f_m = 25[Hz]$, $SNR = 15[dB]$, with amplitude of $\mp15[mV]$. This signal is processed by the $ARMA - KF$ algorithm proposed. Figure 5 shows a three-dimensional representation of the PSD in $[dB]$ versus time-frequency. The crest of the surface manifests a maximum value of PSD for $25[Hz]$. The WVD representation using a Matlab$^{®}$ Time-Frequency Toolbox (Auger, Flandrin, Gonçalvès and Lemoine 2002) is shown in Figure 6. At a sampling
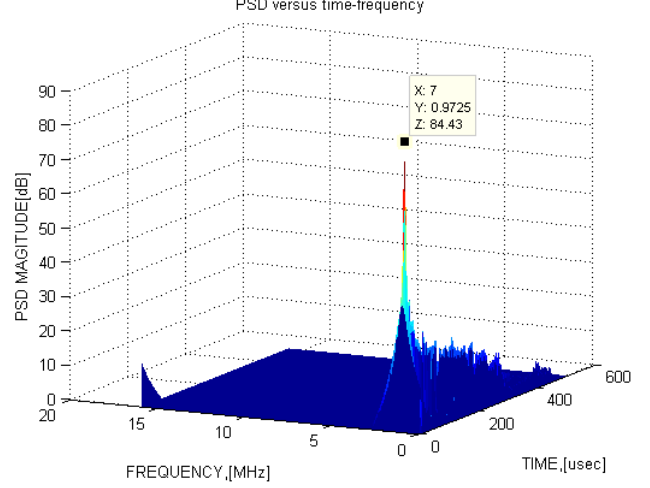

Figure 9: PSD of radar return signal

frequency $f_s = 100[MHz]$, the normalized frequency of 0.25 represents the $25[Hz]$ searched by estimation.

CASE B– Scenary with a object moving at constant speed. The example plots the PSD calculated using equation (9), received from a object moving at the speed $v = 10[m/s]$ ($36[km/h]$, as shown in Figure 7. The rotating frequency is $f_0 = 6.5[Hz]$. The object is moving along a straight line, which gets closer to the observer up to a distance $d = 5000[m]$ and then moves away.

CASE C– Scenary with a radar return signal. The example uses 128 samples of the test radar return signal with doppler frequency $f_m = 0.9578[Hz]$. The estimated instantaneous frequency calculated by the $ARMA - KF$ procedure is represented as the maximum of PSD as shown in Figure 8. The Figure 9 shows a PSD time-frequency representation of the same signal.

10. Results, conclusions and future works

The study aims to estimate the phase caused by the unknown dynamics of an object moving relatively slowly with respect to the platform. To estimate this phase we use an $ARMA - KF$ model and as frame of reference a non-parametric time-frequency representation. The $ARMA - KF$ model is very good in resolution, but totally dependent on the model order, that is, a higher order $ARMA$ would show unnecessary peaks and a lower order may miss the peaks in the clutter spectrum. However, WVD has the problem of cross-spectral components and needs some insight in understanding the distribution, as can be seen clearly in the Figure 10, where we have applied the Hough transform on the WVD image to produce a representation with peaks, whose coordinates give estimates of the linear frequency modulation parameters. This is in contrast to the immediate interpretation of the PSD of the proposed method, Figure 9. The time progression of

radar signal return in complex form, in the presence of additive measurement $AWGN$ and $AWGN$-clutter, was evaluated. This procedure utilizes the adaptive feature of the KF which is carried out recursively for each sample, and offers sufficient resolution in time-frequency domain, in spite of the higher computational severity compared to other parametric or non parametric methods. The first experimental results show that the resolution of the introduced method is equal or higher than other usual time-frequency techniques. The reliability of the procedure was tested using Matlab® programs on simulated data. Compared with standard methods, the $ARMA - KF$-based estimation responded most quickly to parameter changes. The usefulness of our approach in the analysis of Radar returns (SAR) was introduced by an example. Referring to this application, from a general point of view, the use of Kalman filters for the estimation procedure has some advantages compared to other approaches: Kalman filters can be constructed for multivariate systems with stochastic variation of the parameters and the properties of the resulting estimates can be described theoretically. Particularly, this is not necessary to search for a suitable set of base functions to model the temporal evolution of coefficients or to implement procedures for the detection of change points. Furthermore, the Kalman algorithm is appropriate for implementation on microcomputers due to its recursive structure that allows on-line processing, even of huge data sets.

The disadvantage of this procedure is the afore mentioned strong dependence on the order of $ARMA$ model and the establishment of the initial conditions of the variances $\sigma_n{}^2$, $\sigma_w{}^2$ and $\sigma_v{}^2$. It also depends on the initial covariance of the estimation error P and the initial values of the coefficients $a_k(0)$, $b_k(0)$. On the other hand, since the model of system represented by equations (6) and (7) is conditionally Gaussian with covariance $\mathbf{J}(n|n)$, the estimates can be of high variance, thereby a smoothing procedure for the estimates should be involved. Therefore, as future work, to maintain the adaptation speed a nonlinear recursive lowpass filter $\hat{a}_k$ and $\hat{b}_k$ should be used for each component a_k and b_k of the estimated coefficient vector $\mathbf{s}(n)$.

On the other hand, as well as tasks to develop in the future, we list: 1) compare the performace with an AR model in place of the WVD, 2) implement the reconstruction of trajectory as suggested in section 8. 3) improvement the solution of the KF Ricatti equation trought the use of Cholesky factors (known as squared-root filter), 4) use other KF to identify the initial $ARMA - KF$ values, 5) also in the KF, the use of an alternative implementation of state vector, called information filter to improve the numerical stability, specially in cases of very large uncertainties of initial condition estimations, 6) design other types of trajectory dynamics such as the aforementioned rotational trans-
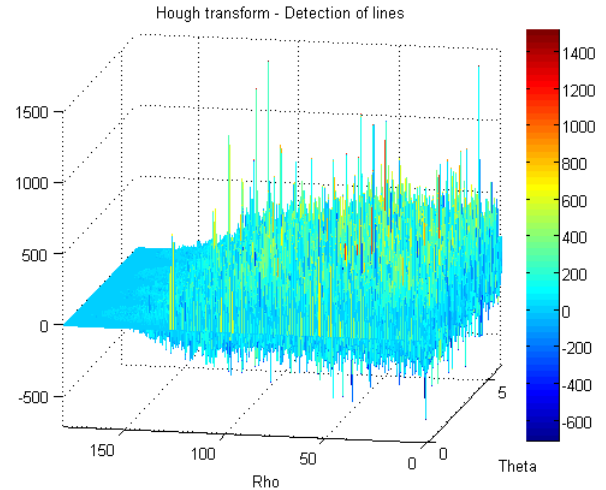

Figure 10: Hough transform to WVD

lation, 7) non-Gaussian clutter with varying SCR, 8) other model for the unknown statistics of the coefficients, 9) analysis and comparison of the computational complexity between $ARMA - KF$ and WVD, 10) extraction of information from a time-frequency image.

References

[1] Brian D.O. Anderson and John B. Moore. *Optimal Filtering.* Prentice-Hall, Inc., 1979.

[2] François Auger, Patrick Flandrin, Paulo Gonçalvès, and Olivier Lemoine. Time-frequency toolbox for use with Matlab. *Free Software Foundation, Inc.,*, Version 1.2,, 2002.

[3] Leonard A. Ferrari Masaharu Harashima and P.V.Sankar. Time-frequency representation for time-varying signals using a Kalman filter. *Electrical and Computer Engineering University of California Irvine, IEEE Proceedings of ASILOMAR-29*, 1996.

[4] Trygve Sparr and Bård Krane. Analysis of phase modulation caused by target motion in SAR images. *Norwegian Defence Research Establishment, Division for Electronics N-2027 Kjeller, Norway*, 2003.

[5] George Stimson. *Introduction to Airborne Radars.* Scitech Publishing, INC, 1998.

A NEW BOND GRAPH MODEL OF PHOTOVOLTAIC CELLS

S. Astier, G. Fontès, X. Roboam, C. Turpin, F. Gailly, L. Ménard

University of Toulouse; INPT, UPS; LAPLACE (Laboratoiry of Plasma and Conversion of Energy);
ENSEEIHT, 2 Rue Camichel, BP7122, 31071 Toulouse Cedex 7, FRANCE

Stephan.astier@laplace.univ-tlse.fr,

ABSTRACT

This paper presents the building of a Bond Graph model of photovoltaic cells based on PN junction by modeling directly the various phenomena involved in the photovoltaic conversion from incident sunlight to final electricity. The basics of photovoltaic conversion and the classical equivalent circuit modeling are presented. Then a new original Bond Graph model of a photovoltaic cell is elaborated. It is built with regards to the physical energy structure of the light and to the energy structure of the PN junction used for the conversion device i.e. the photovoltaic cell. A particular attention is paid to causality problem, as the light appears as a power source of energy which implies very particular constraints explained and leads to a model using bi-causality. Finally, it is shown how this model can be reduced to the classical equivalent circuit model.

Keywords: photovoltaic, modeling, Bond Graph, bi-causality

1. INTRODUCTION

Electricity generated from solar photovoltaics (PV), even if it is so far negligible with regards to the global electric power, with about 40 GW of peak power installed in the world at the end 2010, is characterised by a very significant growth of about 30% by year for the last ten years and more than 100% during the few last years (PV status report 2010, Eurobserv'er 2011). Indeed, in the context of sustainable development, this electricity generation directly from the solar radiation, the most abundant renewable source of energy, is expected to become one of the major sources of electricity in the middle of XXIst century (WWF energy report). Photovoltaic generators began to be used 50 years ago in space and 30 years ago on ground applications after the first oil crisis. Now well known, but not yet as largely developed as other types of electricity sources, the PV field is still strongly evolving with numerous researches and technological developments for new materials and new systems as well for small power remote applications as for grid-connected generators and for large power plants up to tens of MW.

The studies of electric systems are traditionally based on equivalent circuit models or their derivatives. Faithful equivalent electrical circuits of PV cells or arrays are well known. But other types of modeling are exploited, such as Bond Graph (BG), because they are better suited to studies of heterogeneous energy systems within different disciplinary fields (Karnopp 1991). Particularly, the use of Bond Graph representation has proved very fruitful in many domains as for example for the study of systems including PV and/or electrochemical components (Andoulsi 1999a and 1999b, Astier 2004, Saïsset 2006, Ménard 2010). However, this use relied on a direct translation of the electric circuits models of PV generators, not on modeling the PV conversion directly in Bond Graph.

In this paper, considering a solar PV cell classically based on a PN junction, we present the building of a Bond Graph model by modeling directly the various phenomena involved in the photovoltaic conversion from sunlight to electricity. First we recall some basics of photovoltaic conversion and the classical equivalent circuit model. Then we describe the construction of an original Bond Graph model of photovoltaic cell based on PN junction. We pay particular attention to causality problem and analyze it especially in comparison to those of other power converters as electrochemical ones. It is particularly shown that a bi-causal modeling is required. Finally, it is shown how this model can be reduced to the well-known equivalent circuit model.

2. PHOTOVOLTAIC CONVERSION AND CIRCUIT MODEL

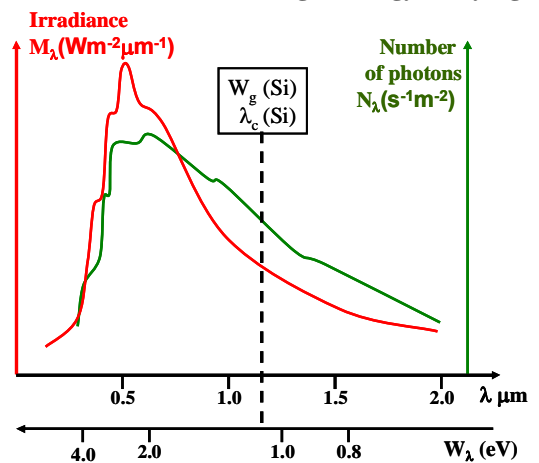### 2.1. Solar radiation and light energy carrying

Figure 1: Solar radiation spectrum in space and associated flux of photons.

Solar light in the vacuum of space has a spectrum of electromagnetic radiation of wavelengths ranging continuously from 0.2 μm to 4 μm as shown on Fig. 1. This spectrum is very near from the one of the radiation emitted by a black body at 5800 K, the irradiance following the theoretical Planck law.

From the energy point of view, this electromagnetic radiation consists in a population of photons (Fig.1), each one carrying the energy w_λ associated with each radiation of wavelength λ and given by (1) where h is the Planck constant and c the speed of light in vacuum.

$$w_\lambda = h\frac{c}{\lambda} \tag{1}$$

The surface density of the power carried by an electromagnetic radiation, called spectral irradiance M_λ [Wm^{-2}μm^{-1}] is therefore as well a surface density of photon flux (Fig. 1) N_λ [s^{-1}m^{-2}] given by (2) where $d\lambda$ represents a slice of spectral wave length of a given width.

$$N_\lambda = M_\lambda \frac{\lambda}{hc} d\lambda \tag{2}$$

At the ends of the visible spectrum:

- for $\lambda = 0.40$ μm (UV); $w_\lambda = 3.10$ eV; N_λ varies between 1 and 3.10^{17} photons.cm^{-2}s^{-1}
- for $\lambda = 0.78$ μm (IR); $w_\lambda = 1.59$ eV; N_λ varies between 3 et 5.10^{17} photons.cm^{-2} s^{-1}

2.2. Photovoltaic conversion and materials

Generating electricity from electromagnetic radiation follows conditions already set by Einstein:

1. the photons must be absorbed by the material (optical absorption) by transmitting their energy to charge carriers of material (electrons);
2. the energy acquired by the excited carriers must be a potential energy recoverable as electrical work (voltaic) but not a thermal energy (kinetic energy);
3. the excited carriers (electrons) must be collected in the external circuit before returning to their initial energy level by recombination.

These conditions lead to define the criteria to realise an ideal photon-electron converter: a semiconductor material and a strong electric field at the location of excited electron-hole pairs. These conditions are particularly satisfied by a PV cell made from a PN junction structure with a semiconductor material. The Fig. 2 shows two classical structures of PV cells using either crystalline silicon or amorphous thin film silicon.

On the one hand the organisation of energy levels of electrons as a band structure with a w_g band gap in the semiconductor allows the photovoltaic conversion:

an electron from the valence band jumps to the conduction band gaining the potential energy w_g.

On the other hand, the strong internal electric field, linked to the space charge around the PN interface, enables to collect the excited free carriers from within the material to the external electrical circuit connected to the front and rear faces of the cell. This function particularly appears on p-i-n structures of thin film amorphous silicon cells (Fig. 2).
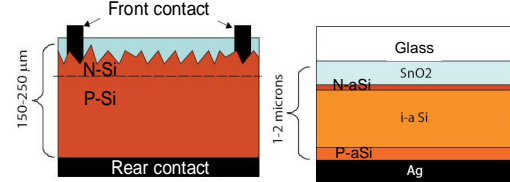

Figure 2: Structures of PV cells based on PN junction: crystalline and amorphous thin film silicon

According to their individual energy w_λ, photons are either reflected or absorbed or transmitted. Only the photons with an energy w_λ greater than the band gap w_g of the material are useful to the generation of photocurrent and to the photovoltaic effect. Thence a first basic necessary condition for the photovoltaic conversion to be realised is given by (3):

$$w_g \le w_\lambda = h\frac{c}{\lambda} \tag{3}$$

The energy of the photon is only partially absorbed in breaking a valence bond, which creates, thus, an electron-hole pair capable of mobility. The energy in excess is rapidly transferred to the lattice as heat (phonons). The electron and hole must be quickly released and collected to then participate in electrical conduction before their recombination. The pairs that reach the area of space charge are separated by the junction field and then collected. An absorption efficiency η_a can thus be defined: for crystalline silicon η_a is about 0.4.

In this complex process, we can consider a simple "two levels model of the photovoltaic conversion" so that:

- a photon with an energy w_λ less than the band gap energy w_g is not absorbed and passes through the material without transmitting any energy;
- a photon with an energy w_λ greater than the band gap energy w_g is completely absorbed and creates an excited electron-hole pair;
- excited electrons only acquire the potential energy equal to w_g. Some of them are involved in recombinations, the other ones participate to the generated photocurrent. However, for the following, the hypothesis of no recombination can be considered in order to simplify.

It results of these phenomena that the band gap energy w_g of the semi-conductor material is a major factor of the photovoltaic conversion.

On the one hand the band gap energy w_g fixes the only part of the spectrum that can be converted and shown on Fig.1, i.e. the radiations of shorter wave length less than λ_c (4):

$$\lambda \leq \lambda_c = \frac{hc}{E_g} \qquad (4)$$

Then, considering the actual solar spectrum on earth, at ground level, the ideal values of bandgap energy to do an efficient conversion with a "two level system" should be comprised between 0.9 and 1.5 eV. This is why, the most used material and technologies are, for example:

- Silicon with w_g (Si) = 1,11 eV
- Gallium Arsenide: w_g (AsGa) =1,35 eV
- Cadmium Telluride: w_g (CdTe) =1,45 eV

On the other hand, the band gap energy w_g also fixes the potential energy gained by the free photo-excited electrons and consequently the maximum output voltage V_g of the theoretical unit cell given by (5) where e ($e=1.6\ 10^{-19}$ C) is the absolute value of electric charge of electron:

$$V_g = \frac{w_g}{e} \qquad (5)$$

As a consequence, any elementary photovoltaic converter, called PV cell, is a low voltage generator. As an example: V_g=1.1 V for silicon cells.

In addition, the energy of radiation being quantified, the number of photons for each wavelength is fixed by the irradiation M_λ and by the spectrum (Planck's law, Fig.1). Consequently, this determines the number of photo-excited electrons and thus the photo-generated current I_{ph}.

In this process, such a theoretical photovoltaic converter appears to behave as a power source P_{ph} (6), imposing simultaneously and separately the voltage V_g AND the current I_{ph} delivered by the voltage source V_g.

$$P_{ph} = V_g I_{ph} \qquad (6)$$

Such a lock is physically incompatible with the connexion to an actual electric circuit of given impedance, except if it can satisfy exactly the relation (6), i.e. with an equivalent resistance R_{eq} such that:

$$V_g = R_{eq} I_{ph} \qquad (7)$$

Fortunately, this lock is broken thanks to the different losses that make the real semi-conductor converters imperfect, particularly with the diode effect linked to the PN junction, different voltage drops linked to conductions and different current leakages.

The well-known simplest equivalent circuit of the ideal PN junction PV cell, given on Fig.3, is deducted from the preceding physical considerations: it includes the current source I_{ph}, which models the photoelectric current, coupled with a diode in parallel which models the diode effect linked to PN junction.
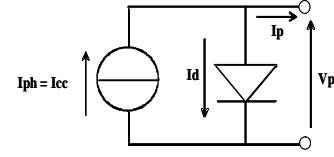

Figure 3: Equivalent circuit of an ideal PV cell based on PN junction.

Let's notice that the current source is well related to the quantified nature of light which behaves as an imposed flow of photons. Many other power converters models involve a voltage source of energy such as for electromechanical or electrochemical ones as recalled in the analysis made further in this paper.

In practice, several factors reduce the efficiency of the PV conversion:

- the reflection of radiations on the cell surface;
- electron-hole recombination which reduces the output current (collecting efficiency);
- the actual output voltage V_p is lower than the theoretical voltage V_g
- the voltage drops at contacts and current leakage at the edges of the junction.

All these additional losses of energy, linked to voltage drops and leakage of current can be globally taken into account by adding to the ideal model of Fig. 3 the two resistors R_s and R_{sh} as indicated on Fig. 4.
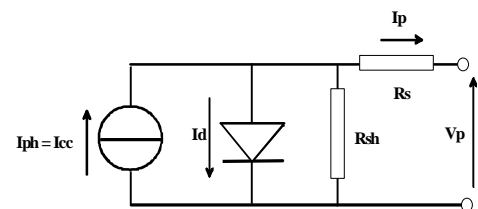

Figure 4: Equivalent circuit of a real PV cell based on PN junction with additional losses.

The equations of this model Fig.4 are (8), (9), (10):

$$I_d = I_s \left(\exp \frac{V_d}{n_D V_T} - 1 \right) \qquad (8)$$

$$I_P = I_{sc} - I_d - \frac{V_d}{R_{sh}} \qquad (9)$$

$$V_P = V_d - R_s I_p \qquad (10)$$

The values of the parameters of this equivalent circuit depend on the actual cell and determine its

performance. For good quality cells, R_{sh} is more than 10 kΩ and R_s is less than 1 Ω.

A more accurate model can use two diodes connected in parallel, which better represents the mechanism of recombination of minority carriers near the middle of the band gap in the area of the space charge. Another simple way to improve the Fig.4 model is to introduce a diode factor n_D in the single diode model as done in (8). The diode factor is slightly greater than 1. Choosing $n_D=1$ neglects the recombination zone in the space charge.

A dynamic model can be obtained by putting a C_D capacitor representing the electric stored space charge associated to the PN junction across the diode on Fig. 4.

This static single diode model is very widely used in studies of PV systems with the well known shape of associated characteristics as indicated on Fig. 5.
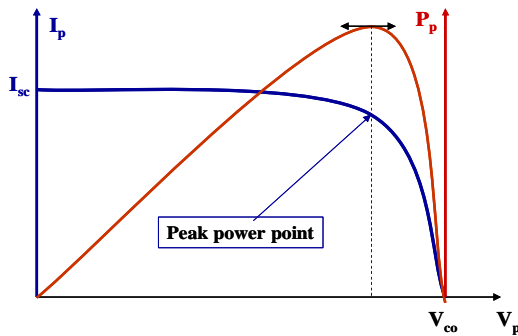

Figure 5: Typical electric characteristic of a PV cell based on PN junction.

The short circuit current I_{sc} is proportional to the power of the light (total irradiance). For a silicon cell, the open circuit voltage V_{oc} is about 0.6 V at 298 K. It varies with the temperature at a rate of about -0.4 % K^{-1}. Thence the basic PV cell is a DC current generator with a very low voltage of about 0.5V at peak power. Higher voltages can be obtained first by connecting a large number of PV cells in series as in commercialized PV modules and PV arrays, and second by means of well chosen static converter as a boost chopper for example.

3. BOND GRAPH MODELS OF PV CELLS

3.1. Bond graph model deducted from equivalent circuit model

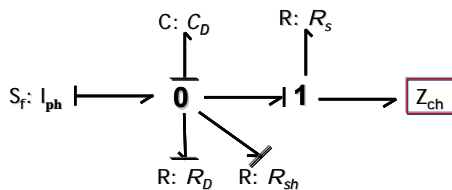A bond graph model can be easily deducted directly from the equivalent circuit given on Fig. 4.

Figure 6: Bond Graph model of a PV cell deducted from the equivalent circuit of Fig. 4 (C_D is added).

The diode is represented by the nonlinear $R: R_D$ element governed by the laws of electrical and thermal behavior of the PN junction. The equivalent capacitor of the junction is added and represented by element C: C_D which determines the causality on the zero junction. This type of model has been used efficiently in many studies on photovoltaic system with Bond Graph (Astier 2004, Andoulsi 1999a and 1999b).

The structure of this dynamic model can be simplified with no fundamental change of the behaviour by involving R_{sh} within R_D and R_s within the load Z_{ch}.

3.2. Bond graph model for monochromatic photovoltaic conversion (without causality)

Now, in order to obtain another model of photovoltaic conversion directly in Bond Graph, let's consider directly the conversion of the radiant energy.

At first let's consider a monochromatic radiation of wavelength λ carrying a power density P_λ. Each photon carries the energy w_λ given by (1). Thus the photons flux density N_λ is given by (11).

$$P_\lambda = N_\lambda \, w_\lambda \tag{11}$$

In order to use a normalized measure of collections of particles, we can consider molar flows densities (by m^2). Then the molar flow density of photons ξ_λ is given by (12) where N_A is the Avogadro number ($N_A = 6.023 \times 10^{23}$).

$$N_\lambda = \xi_\lambda N_A \tag{12}$$

The energy W_λ of one mole of photons is given by (13) and the power density of the radiation P_λ by (14).

$$W_\lambda = N_A w_\lambda \tag{13}$$
$$P_\lambda = N_\lambda w_\lambda = \xi_\lambda W_\lambda \tag{14}$$

In the electric field, the photo-created current I_λ is therefore given by (15) where $F = 96500$ C is the Faraday i.e. the electric charge of one mole of electrons.

$$I_\lambda = k_{pe} N_\lambda e = k_{pe} \xi_\lambda F \tag{15}$$

$k_{pe} < 1$ takes into account the conversion rate and recombination before collecting the excited electrons. But in order to simplify it will be considered $k_{pe} = 1$ for the following which doesn't change the model.

Now, in a first step, if we consider the total transfer of the energy of each photon to each excited electron, the power conservation implies a theoretical voltage V_λ in the electric field depending only on λ by (13).

$$V_\lambda = \frac{P_\lambda}{I_\lambda} = \frac{\xi_\lambda W_\lambda}{I_\lambda} = \frac{W_\lambda}{F} = \frac{w_\lambda}{e} = \frac{hc}{\lambda e} \tag{16}$$

Then, with (15) and (16) it is easy to give the Bond Graph model of this theoretical ideal photovoltaic conversion from light field to electric field as on Fig. 7 with a TF element whose ratio is F.
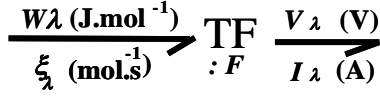

Figure 7: Basic Bond Graph model of an ideal monochromatic photovoltaic conversion.

In the light field the molar flux of photons ξ_λ is a Bond Graph flow while the energy W_λ is the effort. At this step it is independent from the actual semi-conductor material used by the conversion device.

3.3. About causality of PV model

Such a model of Fig.7 looks very similar to the one of the electrochemical conversion of power (Saisset, Ménard) which is recalled on Fig. 8. In this model ξ is the molar flow of the chemical reaction and ΔG the variation of the molar Gibbs free energy. Considering the causality, the molar flow ξ (i.e. the reactant consumption) is a direct consequence of the current I absorbed by the electric circuit. Thence the ξ flow is imposed by the electric side while the effort ΔG is imposed by the chemical side (molar free energy), thus imposing the voltage E in the electric field.
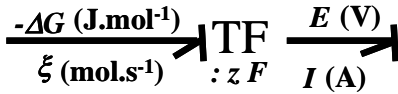

Figure 8: Basic Bond Graph model of an ideal electrochemical conversion.

Differently, the model presented on Fig. 7 does not enable to fix any causality at this step. Indeed the radiation appears as an input power source imposing simultaneously the current I_λ and voltage V_λ. This is the result of the microscopic quantification directly reflected at the macroscopic level by the mechanism of photon-electron conversion. Such a property will require a bi-causal formalism to be rightly represented. But, before, the material behavior also implied in the actual PV conversion has to be introduced.

3.4. Bi-causal BG model of PV conversion

The band energy structure of electrons in the semi conductor imposes to the excited electron the w_g energy instead of w_λ, so that the energy drop ($w_\lambda - w_g$) has to be represented. As all the energy w_λ is first transferred to electric charges, the exeeding power appears as a kinetic power P_k acquired by electric charges and lost by interaction with the lattice becoming heat. Thence, this energy dissipation can be expressed in the electric field as a voltage drop from V_λ down to V_g, as in (17).

$$P_k = W_\lambda - W_g = (V_\lambda - V_g).I_\lambda \qquad (17)$$

We have to represent this phenomenon both with the behavior of the light as a power source. The bi-causal

model on Fig. 9 proposes a solution. In this model both W_λ and ξ_λ are imposed by the light on one side, which is represented by the bicausal source S_eS_f in the light field. On the other side, in the electric field, the potential energy in the semi-conductor is imposed equal to W_g, i.e. an electric potential V_g, which is imposed by the source of effort Se : V_g. The inserted bicausal element D_eD_f has the right constitutive law in order to dissipate precisely the power $P_k = (V_\lambda - V_g)I_\lambda$.
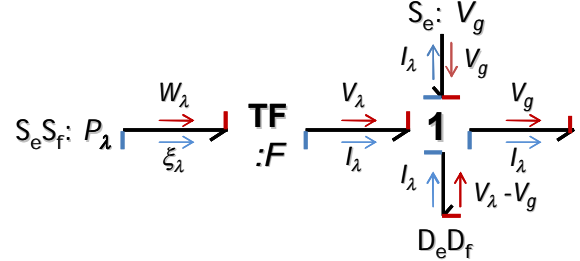

Figure 9: Bond graph model of a monochromatic of PV conversion using a material of V_g bandgap (part1).

At this step, just as before with the circuit modeling approach, the only possible working point in the electric domain is : (V_g, I_λ). The output of this model appears as current source I_λ flowing out from the voltage source V_g. So, all the model of Fig. 9 can be replaced by the equivalent bi-causal source: $S_eS_f : V_g I_\lambda$

Now, just as before, in order to get a more realistic model representing the actual process, it is necessary to introduce the diode effect tightly coupled to the photovoltaic conversion process for collecting excited carriers. Classically, the diode effect involves both a dissipative phenomenon and a capacitive one due to the space charge stored at PN interface. A solution is proposed on Fig. 10. The element R_D has the static constitutive law of a diode. The transition capacitor of the diode fixes the output causality by imposing the voltage V_c. Both with R_D and the electric load Z_{ch}, this constitutes the causal part of the model which has to be connected to the part 1 of Fig. 9, globally replaced on Fig. 10 by the bi-causal source $S_eS_f : V_g I_\lambda$.
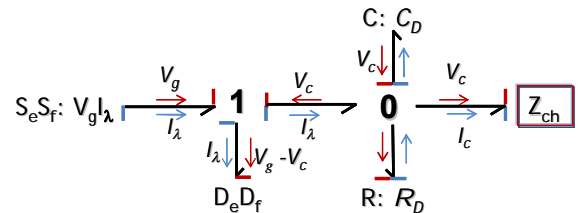

Figure 10: Bond graph model of a monochromatic of PV conversion using a material of V_g bandgap (part2).

On Fig.10, the inserted bi-causal element D_eD_f insures the transition between the bi-causal part and the causal part of the model representing a part of losses causing a voltage drop from V_g to V_c.

Indeed, even in open circuit the open circuit voltage ($V_c = V_{oc}$) is less than V_g. This because a potential barrier equal to the band gap V_g is not realistic in practice, corresponding to the extreme case of doping (p^+n^+) which would imply a too short lifetime of excited

charges to be collected before recombination by tunnel effect. So another doping configuration is chosen in practice for PV cells, such as (pn$^+$), with the consequence of a voltage drop of the potential barrier down to a voltage V_{oc} in open circuit less than the theoretical V_g. More losses resulting in voltage drops and leakage currents can be modeled either with two more dissipative elements, R_s and R_{sh} as on Fig. 6 and not represented here or involved within R_D and Z_{ch}.

Finally, models of Fig. 9 and Fig. 10 can be associated together (part 1 + part 2) so as to constitute the Bond Graph dynamic model of a monochromatic PV conversion. This model satisfies the aim of the study, representing the different phenomena involved in the photovoltaic conversion from the light field to the electric field. But, at this step, it is only valid for monochromatic lights, which is not the case of the sunlight.

3.5. Bond Graph model of photovoltaic conversion with white sunlight

Now, with the white sunlight, a continuous multi-chromatic spectrum (Fig. 1) must be converted with a unique semi-conductor material of w_g band gap energy. In this case, the different values of energy w_λ of every convertible photon are all together downed to w_g when transferred to electrons, the surplus being kinetic energy dissipated as heat in the lattice as already seen.

Therefore as many legs as wavelengths have to be combined to give the global photo-current I_{ph} which is a discrete or a continuous summation of the various I_λ (considering slices $d\lambda$ around λ). And considering the different dissipative phenomena, the common voltage of photo-generated electrons is V_p. Then the model of the electric field is not changed and the Fig. 11 shows a model for a light spectrum with a discrete number n of monochromatic radiations. With a continuous spectrum, the number of legs would be infinite.
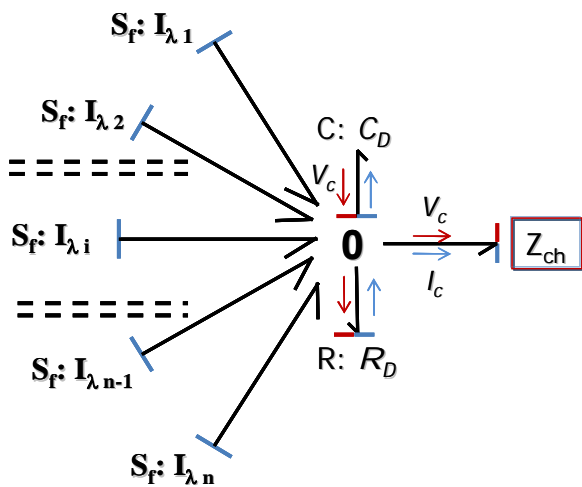

Figure 11: Bond graph model of a poly-chromatic photovoltaic conversion using a PN junction.

Now, the proposed model passing from the light field to the electric field represents the different phenomena involved in the photovoltaic conversion. But, it clearly appears complex, this complexity being related to the continuous spectrum of the white light. In order to avoid this great or infinite number of legs, inddeed not useful in practice, the part of the model at the left side of junction 0 can be replaced by a global flow source S_f: I_{ph} (current source). This operation leads us to the model of Fig. 6 which was directly deduced from the electric circuit model of Fig. 3 ... of course!

4. CONCLUSION

In this paper a new Bond Graph model of a photovoltaic cell based on PN junction, involving both light field and electric field, has been constructed by modeling directly the various phenomena involved in the photovoltaic conversion from sunlight to electricity, doing this instead of just translating the classical equivalent circuit in Bond Graph. If it is truly interesting considering physics, the continuous spectrum of white solar light leads to a model with an infinite number of legs, one for each slice of wavelength! Regrouping all the generated photocurrent in one current source I_{ph} finally leads to the well known model in the electric field. However, this approach showed that the model of PV conversion requires a bi-causal representation in the light field: this is very original among all well known models of power converters. Moreover, such an approach could be interesting for example to model the multi-junction cells (up to six junctions) which are developed for multispectral non linear responses under concentrated light with the aim to reach high efficiencies over 40%.

REFERENCES

Andoulsi, R., Mami, A., Dauphin-Tanguy, G., Annabi, M., 1999. Bond graph modelling of a photovoltaic generator coupled to a dc motor. Proceeding of ICBGM'99, pp 371-378.

Andoulsi, R., Mami, A., Dauphin-Tanguy, G., Annabi, M., 1999. Modelling and simulation by bond graph technique of a DC motor fed from a photovoltaic source via MPPT boost converter. Proceeding of CSSC'99, pp4181-4187.

Astier, S., Saïsset, R., Roboam, X. 2004. Modelling and study of a solar car with embedded photovoltaic array and Li-ion storage Bond. I3M-IMAACA Gènes, October 28-30, 2004.

Saïsset, R., Fontès, G., Turpin, C., Astier, S. 2006, Bond – graph model of PEM fuel cell, Journal of Power Sources 156, pp 100 - 107.

Karnopp, D., Rosenberg, R., 1991. Systems dynamics: a unified approach. John Wiley and Sons.

Ménard, L., Fontès, G., Astier, S., 2010. Dynamic energy model of a lithium-ion battery. Mathematics and computer in simulation, vol 81, issue 2, pp 327-339.

Observ'er 2011 : on website

PV status report 2010, Joint Research Center of European Commission. : on website

WWF, 2010, The energy report : on website.

MODELLING AND CONTROL OF AN EFFORT FEEDBACK ACTUATOR IN HELICOPTER FLIGHT CONTROL USING ENERGETIC MACROSCOPIC REPRESENTATION

Mikaël MARTIN*[(a)], Julien GOMAND*[(b)], François MALBURET*[(c)], Pierre-Jean BARRE*[(d)]

(*) LSIS, Arts et Métier ParisTech – Aix en Provence (Fr)

[(a)]mikael.martin-7@etudiants.ensam.eu, [(b)]julien.gomand@ensam.eu, [(c)]francois.malburet@ensam.fr, [(d)]pierre-jean.barre@ensam.eu,

ABSTRACT

In helicopter field, electromechanical devices controllers are usually designed and tuned from global analysis with transfer functions calculations. This leads to control architectures with a reduced number of controllers. Their regulating loops are usually global PID controllers where parameters are directly set up on dedicated test benches. Energetic representation tools such as Energetic Macroscopic Representation (EMR) aim at simplifying systems analysis and control providing model and control structuring method. In this paper, a simplified helicopter flight axis control is modelled with the intention of controlling the helicopter stick force feedback. Performances of both global PID and energetic model based inversion controllers are discussed through simulation results.

Keywords: causal inversion, model-based control, EMR

1. INTRODUCTION

From the first flight in 1903 of Wright Flyer in Kitty Hawk (US) to the last new technology ones, aircrafts have been equipped with mechanical and hydro-mechanical actuators for flight control. Over time, designers have increased the number of elements of each flight axis control to improve pilots comfort. For example, electrohydraulic actuators controlled by a breakthrough analogical Automatic Pilot Module have been introduced in 1980 on the Dauphin Eurocopter to stand in for overload pilot tasks. From 1983 to 1993, embedded equipments have increased of 50%, from 77 (A310) to 115 (A34) units, in the Airbus industries. Friction effort controls, stability and security improvement systems are at the heart of aeronautical industry and leads to more complex and heavier flight controls. However, mechanical and hydro-mechanical systems are relatively heavy, difficult to adjust and without possibilities of evolution. They also have limited dynamic capabilities compared to electromechanical systems but are able to provide higher forces. Researches are focused on development of new architectures to minimise the number of passive elements. Airbus A320 was the first airliner in 1983 with digital fly-by-wire controls and more recently the

Dassault Falcom 7X became the first business jet with fly-by-wire controls. Evolution towards electronic control architectures sometimes leads up to the mechanical decoupling between pilot control interfaces and final controls on aircraft's steering (Defaÿ, 2010). Pilots generally have mini-sticks and are thus totally disconnected from the force applied on flight controls.

Fly-by-wire brings a lot of advantages in term of weight, cost and ease to connect in axis flight controls. Possibilities to improve or create new flight control functions are enlarged thanks to the flexibility and the high dynamic of the electronic controls which have ability to adapt flight parameter controls in real time. In the helicopter context, interest of fly-by-wires is reduced compared to its high advantages for planes: distance from pilot sticks to hydraulic power control is quite small and mass gain profits against costs are consequently reduced. Mechanical links between pilot sticks and hydraulic power control is maintained except on the fly-by-wire NH90 Eurocopter helicopter (Perrimond 2006). Nevertheless, active force feedback is a mean to dynamically limit the flight envelope. High or sharp inclination of paddles may lead to collisions between paddles and anti-torque rotor (case of nose-up) or may block off turbines air intake (case of nose-down).

Designers are attempted to develop new active flight architectures with more functions, more flexibilities, more securities and more comfort with less elements. Associated control structures are thus becoming more and more complex with difficulties in setting control parameters avoiding over shoots, oscillations and respecting correct response times.

The control of an electromechanical device is usually made separately, locally i.e. independently from other surrounding devices. They are elaborated from global transfer function calculations, using Proportional-Integral-Derivative controllers (PID). Transfer function parameters are estimated from flight structure identifications. PID parameters are consequently directly set up on test benches with typical regulating difficulties. This classical "element by element approach" may be useful for a size limited structure, with specifications on time response, over

68

shoot and stability by customising gain and phase margins. However difficulties come up when the active actuator is surrounded by several physical subsystems. Graphical approaches can be used to help in modelling and controlling multi-physic systems (Hautier 2005). In order to determine a control architecture adapted to the system and correctly set controller parameters, graphic and more specifically energetic methods suggest to properly model the system by respecting physical and natural causality. One of these methods is the Energetic Macroscopic Representation (EMR) which allows deducing a global and a systematic control architecture system representation.

The purpose of this paper is to elaborate a global control strategy of the system using the Energetic Macroscopic Representation tool. A helicopter axis control is described and the model using the EMR is proposed. A systematic model based control is then determined to control the stick force feedback. In the last part, a simulation using Matlab/Simulink shows the interest of that energetic approach compared to a single PID structure.

2. MODELING

This section is dedicated to the presentation and the model of a helicopter flight axis control. The proposed lumped parameters model is then represented using the EMR formalism to facilitate the model inversion exposed in the next section.

2.1. Helicopter axis control

Initially, a helicopter pilot directly had to control the rotor blades angle via direct mechanical chains, without any assistance. With the increase of the aircraft size, and consequently the increase of their mass, the hydraulic powered assistance became necessary to provide the control efforts up to 30kN per axis (depending on the helicopter range). Three hydraulic power-actuators are located under the rotor plate to control the blades angle. The residual frictions of the mechanical chains from stick to the hydraulic actuator control inputs are then one of the most important loads for the pilot.

For the sake of comfort and/or security, additional passive and active devices have been introduced step-by-step, such as (figure 1):

- A damper to reduce the dynamic capabilities of the flight control axes;
- A stabilisation actuator to provide oscillating motion of the control axis;
- A spring based artificial force feedback element;
- An electromechanical actuator controlled by an Auto Pilot Module (APM);
- ...

Nowadays, helicopter flight axis control is thus a very complex system composed of numerous discrete subsystems.

The purpose of this paper is to describe an improvement of one of these subsystems: the spring based artificial force feedback element is replaced by an active force feedback electromechanical system. The objectives are to keep the sensation of moving steers and to give to the pilot a sensation of the flight control states. Force feedback is a feeling indicator to have an indication of the flight envelope and prevent risks for the pilot to overpass high and dangerous inclination of flight controls (to avoid stall for example). The stick is the interface between the model and the pilot. Through the stick, the pilot has the possibility to control paddle angle and thus the helicopter behaviour and trajectory.
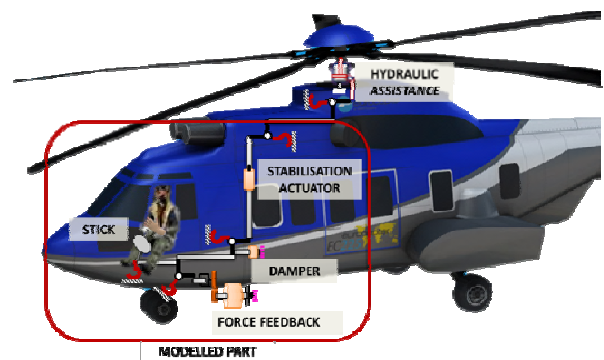

Figure 1: Simplified helicopter axis control (the considered system is delimited by the rectangle)

2.2. Lumped parameters model

A lumped parameters model is further used to establish the EMR of the system. The goal is to locate energy storage and dissipation elements.

Equivalent parameters are listed from the helicopter axis control described in previous section and represented in figure 1. Characteristics (size, materials) of the stick compared to the other elements let us assume that the flexibility of the axis flight control is mainly due to the stick stiffness. This stiffness is thus called K_{stick} and is located in the stick block as shown in the lumped parameters model of figure 2. Each motor (force feedback, damper and stabilisation actuators) and each element with significant mass (load) contributes to the system dynamics (inertia). They are grouped in two parts:

- The first is located in the output shaft of the force feedback element and represents motors and rods inertia ($J_{eqMoteur}$);

- The second represents the load inertia considered as the control axis output (M_{load}).

On the stick side, two configurations are to be considered according to the piloting mode. The first mode, called "hands off", is the case where the pilot does not handle the stick. The stick is free to move and

stick inertia is then considered in the model. The second mode, called "hands on", is the case where the pilot handles the stick. The stick movement is therefore imposed by the pilot. Both cases are represented on the lumped parameters model of figure 2.
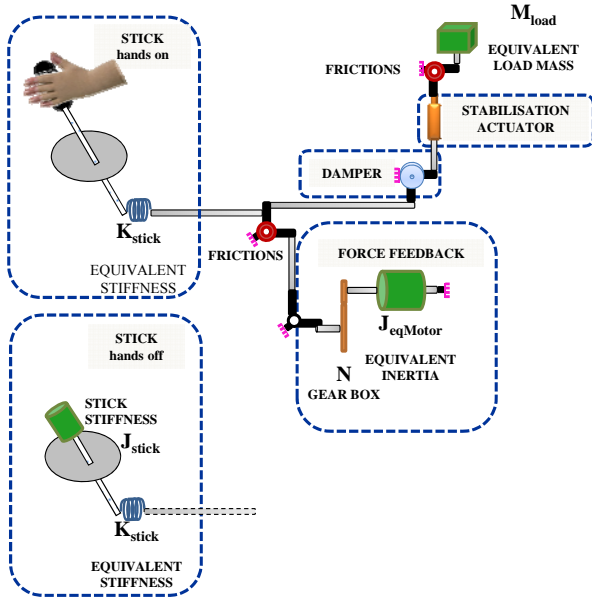

Figure 2: Lumped parameters model in hands on and hands off configurations.

2.3. Energetic Macroscopic Representation formalism

EMR is a graphic and energetic oriented tool defined in 2000 at L2EP research laboratory in Lille (Fr) (Barre 2006). EMR has been invented in order to facilitate analysing and control of the power transfer of a multi-physic complex systems by representing the power fluxes from the energetic sources (electric, mechanical, hydraulic… represented on figure 3.a) to the ending elements. EMR is a general modelling method: it transcends physical fields and has been developed as modelling and analysing tool for complex system. EMR is not limited to either mechanical or hydraulic or electronic systems but every physic field elements may be represented with a unique formalism, on a unique model. It is thus quite close to Bond Graph (BG) (Geitner 2006) which is a more common energetic tool also used for the design, modelling and analysis of complex multi-physic systems. The main particularity of EMR is that it describes physical processes imposing natural causality integration in order to deduce an inversion based control from the energetic model. BG can also be used for system control and analysis of controllability, observability, relative degree and stability of systems. For example Junco S. et al (Junco 2001) propose a BG approach to the trajectory tracking of a series DC motor. However, the BG graphical form doesn't appear like a dedicated control oriented tool and state space representation is often used to determine the control.

Oriented power fluxes on EMR indicate the natural causality which means that each transfer between two elements is either a rigid (timeless relation with no specific direction, figure 3.b) or a causal (integral relation, oriented from cause to consequence relation (figure 3.c). Without natural causality restrictions, unphysical derivative relations may appear on systems which lead to uncontrollable energy storage elements. Application of this essential rule allows a systematic control which is discussed in section 3. Kestelyn (Kestelyn 2009) presents a graphical modelling based on lumped-parameters model of a symmetrical Gantry system. He then deduced from it an inversion based control. In this paper, an equivalent energetic approach is used in application to helicopter field.
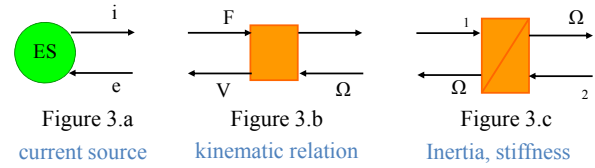

Figure 3: EMR formalism, (a) Energy source, (b) Energy converter (timeless relation) and (c) Energy storage element (Causal relation)

EMR brings a macroscopic vision of systems. Only energy storages elements and power fluxes are represented. If existed, dissipated terms are associated with energy elements they concern. This simplification is particularly helpful to model both multi-physical and complex systems and simplifies the comprehension of the global system.

2.4. EMR of the model

The system is composed of a stick, an active force feedback electro-mechanical converter, a damper and a stabilisation actuator (figures 1). Two configurations corresponding to "hands on" and "hands off" flight mode and described on the lumped parameters of figure 2, are discussed. In "hands on" case, the pilot has hands on the stick and he controls the helicopter motion. Stick inertia J_{stick} is thus negligible compared to pilot arm inertia which is not considered. The pilot imposes a velocity on the stick, i.e. on the stiffness K_{stick}, and feels a force feedback F_{stick} as a reaction of the motion. The pilot is therefore assimilated to a power source imposing the velocity. On a stiffness element, the force is a consequence of the relative velocity between its extremities, as shown in equation 1:

$$F_{stick}(t) = K_{stick} \int_0^t \big(V_{stick}(T) - V_{coupling1}(T)\big) dT \ (1)$$

Where V_{stick} and $V_{coupling1}$ respectively represent the stick and the attached rod velocities.

In the case where pilot does not handle the stick, the helicopter motion is directly controlled by the Auto Pilot Module (APM) via the force feedback actuator. Stick inertia J_{stick} is not anymore negligible and has to be considered. The motion imposed by the APM controller creates a torque Γ_{stick} on stick via the stick

stiffness K_{stick} (equation 2). This torque (cause) implies the stick movement Ω_{stick} (effect).

$$\begin{cases} \Gamma_{stick}(t) = L_{stick} * F_{stick}(t) \\ \Omega_{stick}(t) = \frac{1}{J_{stick}} \int_0^t \Gamma_{stick}(T)dT \end{cases} \quad (2)$$

The active force feedback control consists in controlling motor torque depending on the position of the flight controls. In our simple representation, electric components are assimilated to an electric energy source which delivers a current *i* (figure 3a). As shown on the lumped parameters on figure 2, gear box is supposed to be perfect, without clearances. Motor and part of the equivalent axis flight control inertia are merged and designated by $J_{eqMotor}$ (figure 4). According to equation 2, inertia imposes angular velocity; velocity which is carried towards the first coupling on figure 4. From that coupling, velocity is propagated and modulated by the axis kinematic to the next energy potential storage elements.

Damper effect is added in coupling 2. It is a passive element assimilated to a mechanic energy source (figure 4). Its role is to limit the flight dynamic control by outputting a resistant torque opposed and proportional to the stick velocity, like a viscous force.

The last part of the modelled system is constituted of the stabilisation system. Depending on helicopter measurement and reference attitude, APM calculates and transmits electric orders towards stabilisation actuator. Motors are generally mechanically fixed on non-moving structures. In this paper, the motor is placed in series inside the flight control. Electromotive force (e_{mf}) is then deduced from the velocities subtraction, one imposed by the upper part (V_{load}), the other one from the underneath system ($V_{coupling2}$). Mathematic relation is explained in equation 3, where K_{Φ} is the electromechanical coefficient of the stabilisation motor.

$$e_{mf}(t) = K_{\Phi} * \left(V_{load}(t) - V_{relay_s}(t) \right) \quad (3)$$

The remaining part of the axis helicopter flight control is mainly composed of the hydraulic load assistance. Load is assimilated to an equivalent mass M_{load} and is represented by a mechanical effort source (figure 4). Displacement of that equivalent load is representative of paddles displacement.

The time constant of the motor electrical part is supposed to be negligible compared to the mechanical time constants. The electrical part of the active force feedback actuator is therefore not represented.

3. CONTROL

This section focuses on the determination of the control architecture in order to control the stick force feedback. The energetic model described in previous section allows a systematic model based inversion control.

3.1. Model based inversion control

Control architectures of multi-element systems are often separately elaborated and locally tuned. Their regulating loops are usually PID controllers where gain parameters are directly set up on dedicated bench tests. Difficulties in tuning correctors then appear when the system contains several energetic interactions, in case of complex multi-elements system for example.

Defining control architecture of a system is determining the exact inverse of the model (Barre 2006). As the EMR model has been represented with rigid and causal relations, two different model based inversion control solutions are described in this section.

3.1.1. Direct inversion

The inversion of a bijective rigid conversion is obtained by directly inverse the relation which results in determining regulating value according to a desired value like exposed in equation (4): a torque $\boldsymbol{\Gamma}_{stick}$ is applied to the stick of length L_{stick} which generates a force F_{stick}. The inversion of the model consists in calculating a regulating torque $\boldsymbol{\Gamma}_{stick\_REG}$ to obtain the desired stick force F_{stick_DES}. Rigid inversion using EMR formalism is exposed in figure 5.

$$F_{stick} = \frac{\Gamma_{stick}}{L_{stick}} \xrightarrow{inversion} \Gamma_{stick_{REG}} = L_{stick} * F_{stick_{DES}} \quad (4)$$

3.1.2. Indirect inversion

The inversion of a causal relation consists in determining the input of a storage element depending on
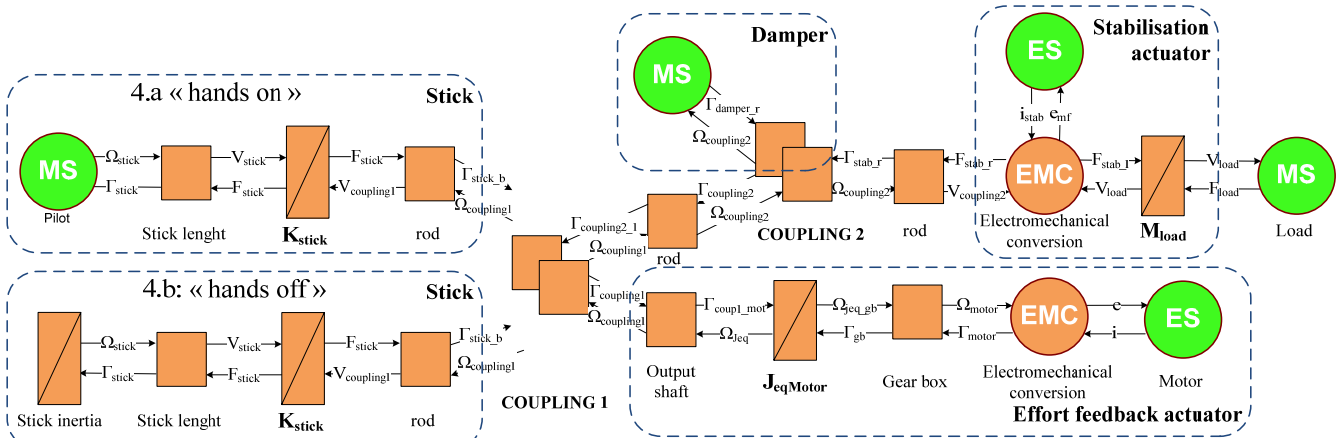

Figure 4: Energetic models of a simplify helicopter control axis in hands on (4.a) and hands off (4.b) configurations

the desired output. The direct inversion would consist in knowing the desired output value and its future evolution to calculate the reference value. This would lead to an unphysical derivate relation. The inversion of the causal relation of the model illustrated by the equation 1 consists in controlling the accuracy of the stick force F_{stick} by comparing and raising the difference ε between the measured and the desired stick force values $(F_{stick_{MES}}, F_{stick_{DES}})$. The rod velocity $V_{coupling1REG}$ which has to be regulated is calculated by modulating the difference ε with a proportional corrector K_p (equation 5). The indirect causal inversion using the EMR formalism is exposed in figure 5.

$$\begin{cases} \varepsilon = F_{stick_{DES}} - F_{stick_{MES}} \\ V_{coupling1_{REG}} = K_p * \varepsilon \end{cases} \quad (5)$$
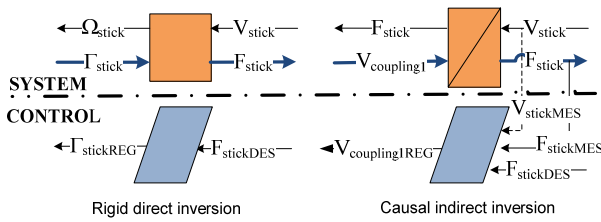

Figure 5: Direct and indirect model based inversion control using EMR formalism.

3.2. Inversion of the model

The control of the simplified axis helicopter flight control modelled in figure 3 has been obtained using the direct and indirect inversion principles and the result is shown in figure 6. At each flight control position is associated a reference force feedback. Stick feel forces is thus regulated thanks to the motor current control. Simple analysis of energy storages on EMR (figure 4) between stick and motor defines number and location of causal storage elements which therefore determine number of correctors. Two energy storage elements are located in the control path: the first one is the kinetic storage element (inertia) which has to be controlled by a primary velocity loop, the other one is the potential energy storage due to the stick stiffness and has to be

managed by a force loop control.

The motor velocity is controlled by the PI controller. EMR is structured to deduce from the model a maximal control structure. It is not only resulting into the direct and indirect inversions control like exposed previously but a maximal control structure is obtained by measuring and compensates the reaction parameters in order to increase the efficiency of the regulating loop. In figure 5, motor torque is a measured factor required to increase the efficiency of the controller. This improvement is only possible when the measure is cost reasonable and conceivable. In the proposed system, there is no torque sensor on output motor shaft. Without compensation the motor torque reaction is seen as a disturbance source. The closest sensor is the force sensor used for the force feedback controller. A more practical control structure is then determined by anticipating the reaction torque control from the stick force reference value and using the kinematic of the model.

Energetic approach lets opportunity to rapidly guess how many signals should be measured, nature of these signals (velocity, effort, current) and how many controllers should be set up. In the ideal case, the proposed methodology determines a maximum control structure: each energetic storage element has its own sensor and its own controller.

4. SIMULATION

For simulation purpose, the simplified axis helicopter control is modelled in a Matlab/Simulink environment. The Simulink model is based on EMR constraints as represented in Figure 6. However, only elements from the stick to the active haptic device have been modelled. Damper and stabilisation actuator are considered as disturbances and are ignored in first approximation. The brushless motor is controlled to create a force feedback to the pilot. The simulation is focused on the force feedback quality and more particularly on the ripple sensibility. The motor technology inevitably generates torque ripples which are uncomfortable and should not be felt by the pilot. An estimate of motor torque ripples is therefore added to the model.
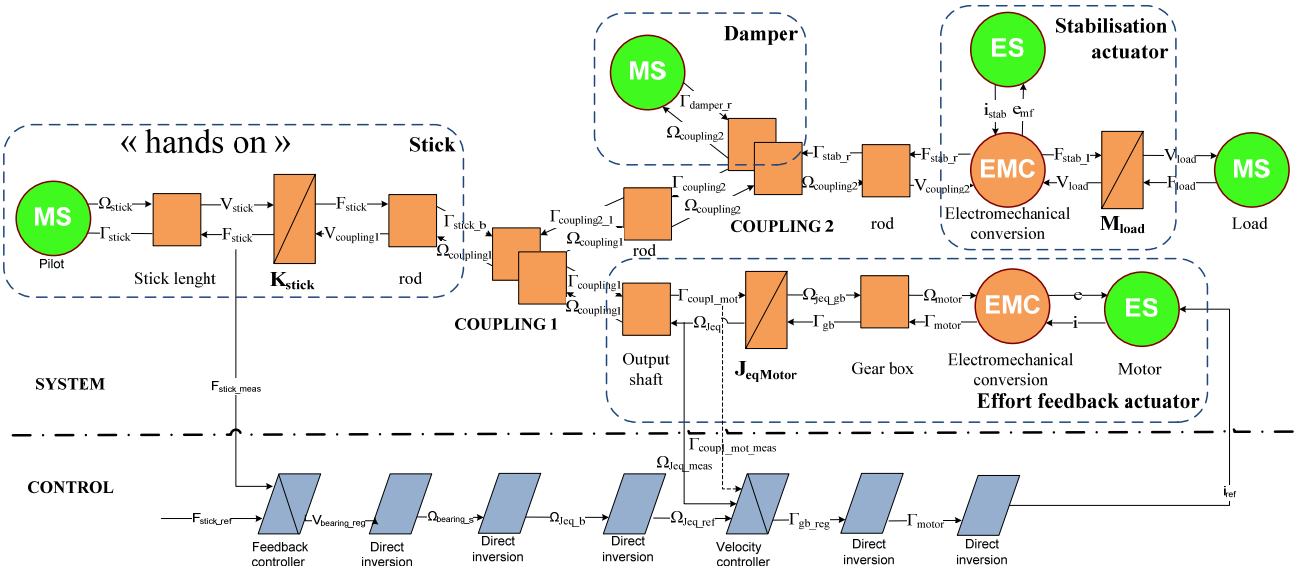
The system is supposed to interact with the pilot ("hands on"). Studies of the relation between muscular activation patterns and movements show that human beings dynamic capabilities are limited and can be modelled with minimum-Jerk trajectories (Harris 2004). Such a smooth motion is generated thanks to a trapezoidal velocity signal filtered by a sliding mean filter. The trapezoidal shape simulates a forward stick displacement whereas the mean filter allows smoothing the velocity signal in order to limit the Jerk. The corresponding position motion is calculated by integrating the velocity and is represented in figure 8.a. The aim of the active actuator is to output a force feedback depending on the stick position. A theoretical effort against position curve is defined to compute the force reference depending on the stick position. This force map is shown on figure 7. The law is composed of two symmetric effort gradients.
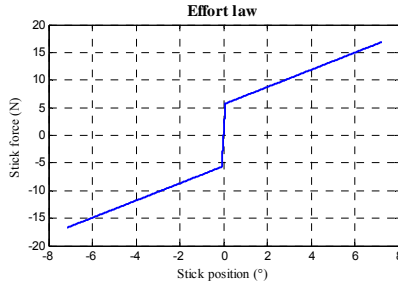

Figure 7: Force feedback map

Two different types of control structures are elaborated:

1. A unique PID control which represents the most common used control loop.
2. An inversion model based control deduced from the energetic methodology presented in this paper.

The model obtained by using the energetic approach points up two energetic storage elements. The common PID control architecture regulates the global system energies with a unique force controller whereas the model based inversion control represented on figure 6 recommends the regulation of both storage elements, as if to regulate a velocity primary loop and a secondary force loop. Control loop characteristics are summarised in table 1. Both regulating loop has been set up regarding the following characteristics:

- maximising phase margin,
- minimising overshooting,
- minimising force error
- 45Hz force bandwidth

Table 1: Controller characteristics

Controller	Phase margin	Bandwidth
Force PID	*66°*	*45Hz*
Speed PI	*66°*	*113 Hz*
Force PI	*35°*	*45Hz*

The human being has an asymmetric sensing control capability. The maximum frequency at which the human fingers can transmit or control a position to their environment is from 0 to 10Hz but human can perceive a force or position signal at up a frequency of 20 to 30Hz (Burdea 1996). The 45Hz force bandwidth is therefore large enough for both control architectures.

Results of simulation are shown on figure 8. The measured force feedback map (figure 8.b) and the stick force feedback error (figure 8.c) for both control structures demonstrate that torque ripples are
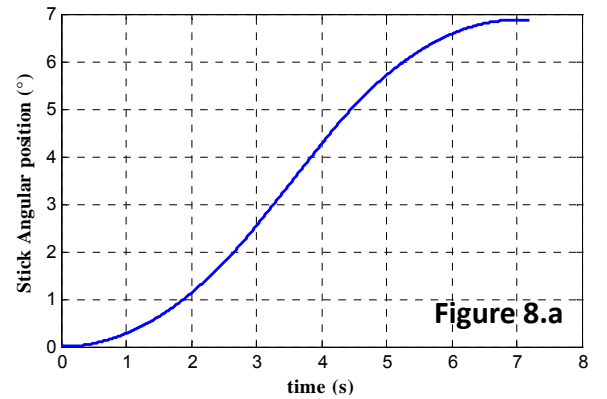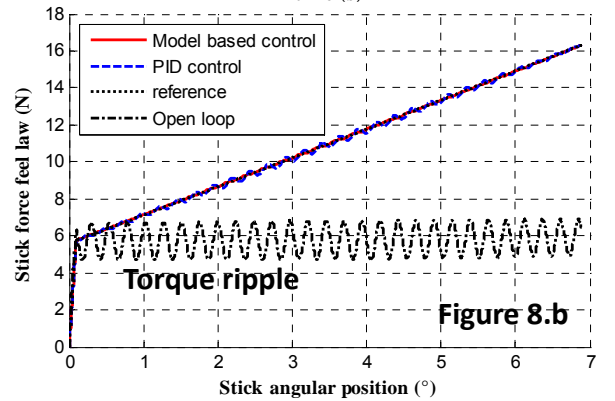

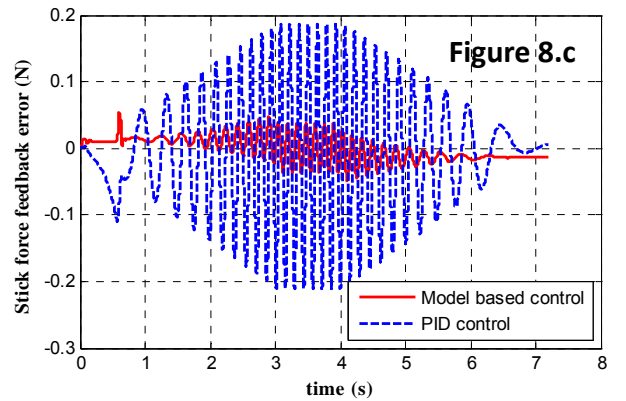Figure 8: Matlab/Simulink simulation results

Figure 6: model inversion: application to an helicopter axis control

significantly reduced thanks to the regulations. Nevertheless, better performances on force feedback error are observed for the model based control with a maximum force error of 100mN peak to peak. The maximum force error on PID control reaches 400mN peak to peak which is largely felt by the human sensibility.

5. CONCLUSION

Energetic tools such as EMR recently became visible to structure models and determine a systematic model based control. An axis helicopter flight control has been modelled and an active force feedback has been generated on stick. The EMR energetic method has been used to define the control architecture respecting natural causality and resulting performances are compared to a more common used PID control loop. A Matlab/Simulink simulation demonstrates that the model based control allows the smoothing of the motor torque ripples whereas the PID controller only reduces the amplitude of the torque ripple. The pilot comfort is therefore improved.

REFERENCES

Barre, P.J., Bouscayrol, A., Delarue, P., Dumetz, E., Giraud, F., Hautier, J.P., 2006. Inversion-based control of electromechanical systems using causal graphical descriptions. *IEEE-IECON'06 - 32nd Annual Conference on Industrial Electronics*, (pp. 5276-5281).

Burdea, G.C., 1996, Force and touch feedback for virtual reality, *John Wiley and Sons*, ISBN 0-471-02141-5

Defaÿ, F., Alazard, D., Dehais, F., 2010. Impedance active control of flight control devices. ISAE, Toulouse University.

Geitner, H.G., 2006, Power flow diagrams using a bond graph library under Simulink, *IEEE-IECON'06 - 32nd Annual Conference on Industrial Electronics*, pp. 5282-5288.

Harris, C.M., 2004, Exploring smoothness and discontinuities in human motor behavior with Fourier analysis, *Mathematical biosciences*, vol. 188, pp. 99-116.

Hautier, J.P. Barre, P.J., 2005, The causal ordering graph - A tool for system modelling and control law synthesis, *Journal of studies in informatics and control*, vol. 13, no. 4, pp. 265-283.

Junco, S., Donaire, A., Garnero, G., 2001, Trajectory Tracking on Bond Graphs. Procedures and applications to DC electrical drives, ESS01, *13th European Simulation Symposium, Simulation in Industry*, , October 18-20, Marseilles, France, SCS Europe Bvba, pp. 799-805.

Kestelyn, X., Gomand, J., Bouscayrol, A., & Barre, P.J., 2009, Control of a symmetrical dual Gantry system using Energetic Macroscopic Representation, *Solid State Phenomena, vol. 144, pp 181-185*.

Perrimond, G., Belan, G., 2006. *NH90, A successful programme, TTU international*, 24 May, Supplement to n°162.

PASSIVE FAULT TOLERANT CONTROL: A BOND GRAPH APPROACH

Nacusse, Matías A.[(a,b)] **and Junco, Sergio J.**[(a)]

[(a)] LAC, Laboratorio de Automatización y Control, Departamento de Control, Facultad de Ciencias Exactas, Ingeniería y Agrimensura, Universidad Nacional de Rosario, Ríobamba 245 Bis – S2000EKE Rosario – Argentina.
[(b)] CONICET – Consejo Nacional de Investigaciones Científicas y Técnicas.

[(a,b)]nacusse@fceia.unr.edu.ar, [(a)]sjunco@fceia.unr.edu.ar

ABSTRACT

The Passive Fault Tolerant Control (PFTC) approach defines a unique, robust control law able to achieve the control objectives even in the presence of a fault. This work addresses the PFTC problem in the Bond Graph (BG) domain. The control law is obtained using an energy and power shaping method in this domain. At its first step, the method proposes a so called Target Bond Graph (TBG) that expresses the desired closed-loop behaviour, i.e., the control system specifications, in terms of a desired closed-loop storage function and a power dissipation function. The control law is obtained in a subsequent step via BG-prototyping. In order to make it fault-tolerant, this control law is further robustified by using a residual signal obtained from a modified version of the diagnostic bond graph (mDBG) which is created from the TBG. This results in a closed-loop behavior under faults that it asymptotically equivalent to the faultless case.

Keywords: Passive Fault Tolerant Control, Bond Graph, Energy and Power Shaping, Diagnostic Bond Graph.

1. INTRODUCTION

Fault tolerant control (FTC) can be classified in two main categories, Passive Fault Tolerant Control (PFTC) and Active Fault Tolerant Control (AFTC). The passive approach defines a unique control law to achieve the control objectives even in the presence of a fault. Generally speaking, the passive approach ensures stability and confers robustness under faults to the control system, but there exists a trade-off between performance and robustness (Isermann 2006). The active approach modifies the control law according to the faults occurred, so that in this approach the faults must be detected and isolated and a decision must be made in order to reconfigurate the control law. Both approaches are usually complemented in the praxis to improve the performance and stability of the fault tolerant system (Blanke, et al. 2006). Refer to (Zhang and Jiang 2008) for a bibliographical and historical review on FTC.

There are many methods for model based FDI defined in the BG domain. Most of these method derive analytical redundant relations (ARR) from the BG model (Ould Bouamama et al. 2003) and other use the BG model for direct numerical evaluation of ARRs (Samantaray et al. 2006), (Borutzky 2009).

This work addresses the problem of Passive Fault Tolerant Control (PFTC) in the BG domain. To obtain the fault tolerant control law, an energy and power shaping method in the BG domain is used (Junco 2004). This method first expresses the control system specifications in terms of desired closed-loop energy and power dissipation functions, capturing them in a so called Target Bond Graph (TBG) matching the desired closed-loop behaviour. The method proceeds further constructing the control inputs to the plant via Bond Graph prototyping in such a way that the coupling of the resulting controller-BG (a BG representing the control law) and the plant-BG is equivalent to the TBG.

The classic approach to solve problems of system parameter dispersion and to simultaneously reject disturbances is adding integral action to the controller (Khalil 1996). The BG solution to this problem presented in (Junco 2004) has been generalized in a Port-Controlled Hamiltonian Theory context in (Donaire and Junco 2009). This previous result is shown in this paper to solve a PTFC problem, but yielding a closed-loop response different than the originally defined in the TBG. In this paper, a new control law based on a modified Diagnostic Bond Graph (mDBG) is proposed. It makes the closed-loop system behave like the original TBG, at least asymptotically.

As presented in (Samantaray et al. 2006) the Diagnostic Bond Graphs (DBG) are originally used to generate residuals for fault detection and isolation (FDI) in Active Fault Tolerant Control (AFTC) problems. The standard version of the DBG uses the plant inputs and the plant measurements to generate a residual signal. This residual signal depends on the model parameters and the real plant parameters. Here, a modified version of the DBG is proposed: instead of feeding the plant nominal model with the measurements, the original TGB, i.e., the nominal control system, is fed with the actual reference signals and measured plant outputs. Thus, the residual signal obtained from the mDBG is a measure of the error between the desired and the actual dynamics of the control system. If the residual signal is

zero, then the control system behaves like the TBG and the control objectives are achieved. So, the control law aims at making the residual signal vanish in time.

The methodology proposed here is developed through a case study and validated numerically in simulation using the software 20sim (Controllab Products B.V).

The rest of the paper is organized as follows. Section 2 presents the system under study, its mathematical model and some background results necessary to introduce the contributions of the paper. Section 3 presents the methodology contributed by the paper illustrated through an example. Section 4 presents some simulation responses that prove the good dynamic response of the control system. Section 5 addresses some issues related to controller reconfiguration and, finally, in Section 6, conclusions and future work are addressed.

2. BACKGROUNDS

This section presents the system under study and some background on the energy and power shaping method in the BG domain.

2.1. System under study

The system under study, depicted in Figure 1a, consists of two tanks located one above the other, where the upper tank discharges into the lower tank.

The tanks are fed with one input flow which is splitted between them through a distribution valve whose parameter $\gamma \in [0,1]$ determinates how the input flow is distributed to the tanks, if $\gamma = 1$ then all the input flow in directed to the upper tank.
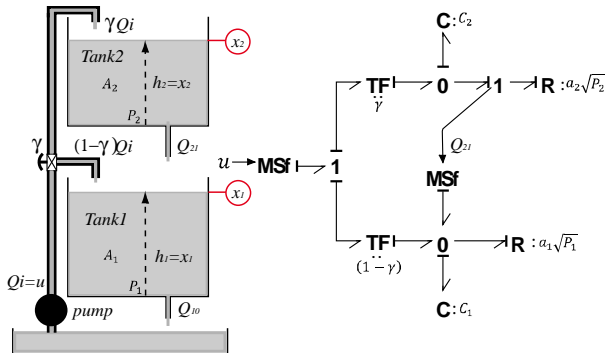

Figure 1. Physical system and its BG model. Measured plant outputs encircled in red.

The state equations can be read from the BG using the stored liquid volumes as state variables, as usual, or the liquid levels, as done here in (1):

$$\dot{x}_1 = -\frac{a_1\sqrt{\rho g x_1}}{A_1} + \frac{a_2\sqrt{\rho g x_2}}{A_1} + \frac{(1-\gamma)}{A_1}u$$
$$\dot{x}_2 = -\frac{a_2\sqrt{\rho g x_2}}{A_2} + \frac{\gamma}{A_2}u \qquad (1)$$

Where x_1 and x_2 represents the liquid level of tank1 and tank2 respectively; A_1 and A_2 are the cross section areas of the tanks, related to the tanks hydraulic

capacities by the relation $C_i = \frac{A_i}{\rho g}$ (with $i=1,2$); ρ is the constant density of the liquid, g is the gravitational acceleration and a_1 and a_2 represents the cross section of outlet hole from the tanks.

Observing the BG model of Figure 1 it is easy to see that the system is structurally state controllable, because all states are input-reachable which means there exist a causal path between the source and every energy store in integral causality, and the energy storage elements can change causality when derivative causality is preferred (Sueur and Dauphing-Tanguy 1991) (this being a result valid for linear systems can be applied to a version of our BG linearized around an equilibrium point). Moreover, the tank level x_1 can be achieved through two different causal paths. The shorter comes directly from the modulated source through the transformer (input-to-state relative degree 1). The other causal path comes from the modulated flow source over tank2 (input-to-state relative degree 2). So, the state variable x_1 is input-reachable by two different paths; this structural redundancy can be used in case of faults in the distribution valve. Indeed, even for a severe fault like $\gamma = 1$, i.e., for a complete obstruction of the lower duct, which directly feeds tank1, the system is still structurally state controllable.

Let x_1 and x_2 be the measures of the system making it observable. In steady state condition the system imposes the following constraint:

$$\overline{x_2} = \left(\frac{\gamma a_1}{a_2}\right)^2 \overline{x_1} \qquad (2)$$

This steady state constraint allows controlling one tank level forcing the other level to a desired value. This dependency can be used in case of sensor faults.

The faults considered in this work can be classified in three different groups:

- Faults in the constitutive relationships of the BG components, for example an obstruction / opening of the discharges orifices.
- Structural faults, for example an obstruction in the distribution valve ($\gamma = 1$ or $\gamma = 0$), Leakage in one or both tanks.
- Sensor faults.

2.2. Power and Energy shaping on bond graphs

The power and energy shaping control technique defines the control problem as a stabilization one, imposing the desired closed-loop energy and power dissipation, and obtaining the control law through matching equations relating the control open-loop energy function (a kind of control Lyapunov function, see Sontag 1998) and the desired closed-loop functions.

In the BG domain the desired stored energy and power dissipation are captured in the TBG. In order to obtain the control law, the controlled sources in the BG model of the plant are prototyped with the aim of obtaining a so called virtual BG that matches the TBG. For further details the reader must refer to (Junco 2004).

This method is exemplarily performed on the discussed two tank model with the following control objectives on tank1:

- Level tracking.
- Disturbance rejections.
- Robustness regarding parametric uncertainties.

The proposed TBG for the closed loop system is shown in Figure 2 with the desired stored energy and power dissipation expressed in terms of the tracking error state variable in (3) and (4). To simplify the notation in the equations the following constants have been introduced: $k_1 = (\rho g)^2 C_1$, $k_2 = \frac{(\rho g)^2}{R_H}$, where R_H represents a hydraulic resistance and C_1 a tank hydraulic capacity.
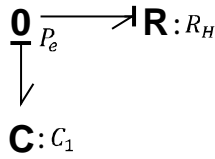

Figure 2. Proposed TBG

$$V(x) = \frac{1}{2} k_1 x_e{}^2 \tag{3}$$
$$\dot{V}(x) = -k_2 x_e{}^2$$
$$\dot{x}_e = -\frac{b}{A_1} x_e \tag{4}$$

In the equations above, the tracking error $x_e = x_1 - x_1^{ref}$ is the state variable of the (incremental) TBG, x_1^{ref} is the tank1 reference level, and $b = \frac{k_2}{\rho g}$.

To enforce the desired closed-loop dynamics specified by the TGB, the virtual BG of Figure 3 is constructed. It shows how to proceed in order to obtain the control law. The left half of the figure is obtained prototyping the controlled power source MS_f in such a way that access is gained to the chosen output, the level x_1, and an overall equivalent behavior to the TBG is achieved (some cancellations can be seen following the causal paths, also note that there are some virtual elements with negative "gains").
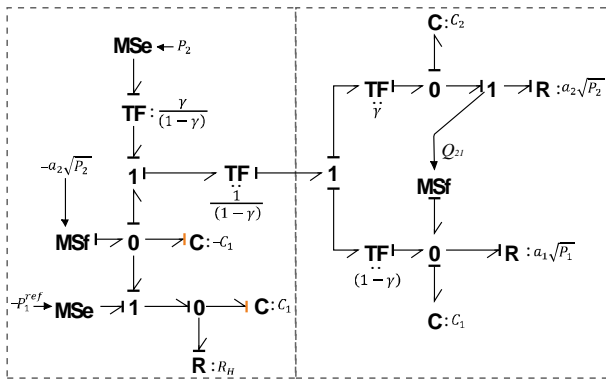

Figure 3. Virtual BG.

Using the standard causality reading procedure, the control law (5) is obtained directly from the virtual BG. This law can be thought of splitted into three components as in (6). The first and second term cancel the perturbation from tank2 and the nonlinearity of the outlet flow of tank1, respectively, and the third imposes the tracking error desired dynamics.

$$u = \left(\frac{1}{1-\gamma}\right)\left[-a_2\sqrt{\rho g x_2} + a_1\sqrt{\rho g x_1} - b\left(x_1 - x_1^{ref}\right) + A_1 \dot{x}_1^{ref}\right] \tag{5}$$

$$u = \left(\frac{1}{1-\gamma}\right)\left[u_1 + u_2 + u_3\right] \tag{6}$$

$$u_1 = -a_2\sqrt{\rho g x_2}, \qquad u_2 = a_1\sqrt{\rho g x_1},$$
$$u_3 = -b\left(x_1 - x_1^{ref}\right) + A_1 \dot{x}_1^{ref}$$

Assuming exact model knowledge and perfect measurements, this control law yields a closed-loop behavior equivalent to the TBG of Figure 2, i.e., the closed-loop dynamics satisfies (7). As no objectives are imposed on tank2 and its dynamics is hidden in closed-loop, its stability must be analyzed after the controller has been designed, property that can be easily verified in this case.

$$\dot{x}_e = -\frac{b}{A_1} x_e \tag{7}$$

Perturbed closed-loop dynamics. Because of parameter dispersion, faults, modeling errors, sensor limited precision, noise, etc., neither the model nor the measurements are exact. To deal with this it is convenient to think the control input as composed by two terms as in (8), where u_r is the "rated" part of u. This means the control input part that performs the power and energy shaping under ideal plant and measurement conditions. In the same expression, δ_u is the unknown controller part due to modeling errors, parametric dispersion, faults, etc. The BG of Figure 4 reflects this situation.
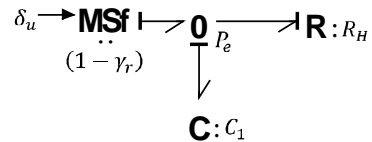
$$u = u_r + \delta_u \tag{8}$$

Figure 4. Perturbed TBG

Under this situation the closed-loop dynamics no longer satisfies (7) but (9), where γ_r is the rated value of the distribution valve parameter:

$$\dot{x}_e = -\frac{b}{A_1} x_e + \frac{(1-\gamma_r)}{A_1} \delta_u \tag{9}$$

It can again be verified that the hidden closed-loop dynamics of tank 2 is stable as it satisfies (10):

$$\dot{x}_2^{CL} = -\frac{a_2\sqrt{\rho g x_2}}{(1-\gamma_r)A_2} + \frac{a_1\gamma_r\sqrt{\rho g x_1}}{(1-\gamma_r)A_2} - \frac{b\gamma_r\left(x_1-x_1^{ref}\right)}{(1-\gamma_r)A_2} +$$
$$\frac{A_1\gamma_r\dot{x}_1^{ref}}{(1-\gamma_r)A_2} + \frac{\gamma}{A_2}\delta_u \qquad (10)$$

It can be seen in (9) that the level error of tank 1 is driven by δ_u. A remedy must be found if this induces inadmissible behavior in closed-loop. The next subsection shows how to do this in the BG domain using existing results. The solution is shown to be good even under the presence of some faults. In the next section a new result is presented which recovers a closed-loop performance closer to that of the original TBG.

2.3. Robustifying the control law adding integral action.

Looking at the perturbed TBG of Figure 4 a solution comes immediately to mind: inject a state-dependent flow into the 0-junction that (asymptotically) cancels the flow injected by the disturbance source. This is attained adding the I-element as shown in Figure 5, which is exactly adding integral action to the control law, a classic approach to solve problems of system parameter dispersion and to simultaneously reject disturbances (Khalil 1996). This kind of BG solution first presented in (Junco 2004) has been theoretically generalized in the context of the Port-Controlled Hamiltonian Systems Theory in (Donaire and Junco, 2009).
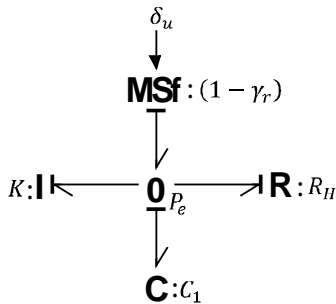

Figure 5. Perturbed TBG with integral action

The control law with integral action has to be calculated on the modified TGB given in Figure 6. Proceeding in the same way as when deriving equation (5) the control law $u = \left(\frac{1}{1-\gamma}\right)[u_1 + u_2 + u_3 + u_4]$ is obtained, with $u_4 = K\int\left(x_1 - x_1^{ref}\right)$.
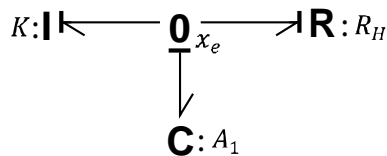

Figure 6. TBG with integral action

The tracking error x_e can be shown to satisfy (11), ie., a second order dynamics is obtained which differs from the first order dynamics error defined by the TBG of Figure 2. Note however that this second order dynamics can be made to arbitrarily close approximate a first order dynamics with the correct choice of $\frac{b}{A_1}$ (i.e., R_H) and K, at least theoretically, then this is mathematically possible but physically limited by the capacity of the actuators.

$$\ddot{x}_e + \frac{b}{A_1}\dot{x}_e + Kx_e = \frac{(1-\gamma_n)}{A_1}\dot{\delta}_u \qquad (11)$$

2.4. Diagnostic Bond Graph
The Diagnostic Bond Graph was first presented by (Samantaray et al. 2006) for numerical evaluation of analytical redundant relationships (ARR). The ARRs are calculated to perform FDI in an AFTC frame.

Basically, the DBG is obtained from a BG model of the plant injecting the plant measurements and inputs through modulated sources. The residual signal is obtained by measuring the power co-variables of the modulated sources, see Figure 7.
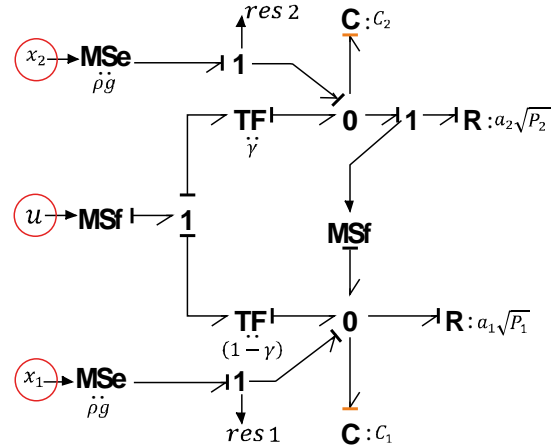

Figure 7. Diagnostic Bond Graph. Plant measurements to be fed into the DBG encircled in red.

Reading directly from the BG the residuals are:

$$res1 = A_1\dot{x}_1 + a_1\sqrt{\rho g x_1} + a_2\sqrt{\rho g x_2} - (1-\gamma)u$$
$$res2 = A_2\dot{x}_2 + a_2\sqrt{\rho g x_2} - \gamma u \qquad (12)$$

As can be noted in (12), the residuals depend on system parameters. If the model represents perfectly the controlled system, then the residual signals are zero. The derivative causality is an advantage in FDI, because no initial states are necessary to evaluate the residuals.

In the sequel only residual 1 is considered, as it is the only one related to the TBG associated to the control problem the paper deals with.

3. MAIN RESULT
In this section, the residual signal obtained from a modified version of the DBG is used to obtain a control law which is robust to faults.

3.1. Modified Diagnostic Bond Graph (mDBG).

The mDBG is defined injecting the tracking error level (as measured on the real control system) into the TBG through modulated sources, see Figure 8.
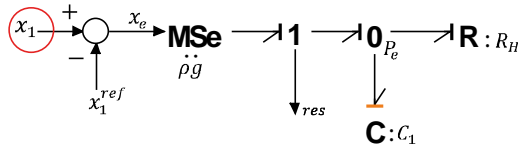

Figure 8. Proposed mDBG. Measurements to be fed into the mDBG encircled in red.

The mDBG yields the new error dynamics in (13), where res is the residual signal read from the mDBG, a measure of the difference between the actual and the ideally expected closed-loop dynamics. This signal is used to make the control law more reliable under faults.

$$\dot{x}_e = -\frac{b}{A_1}x_e + \frac{res}{A_1} \tag{13}$$

As it can be seen in (13), the error dynamics is driven by the residual signal, when $res = 0$, x_e responds as previously defined in the TBG of Figure 2.

3.2. Obtaining the control law.

The control objective is reached when there exists a control law $u = u(x_1, x_2, x_1^{ref}, res)$ such that res tends to zero with growing time.

The residual expression (14) obtained reading the mDBG clearly shows that choosing u as in (5) yields $res = 0$ in absence of faults and modeling errors.

$$res = -a_1\sqrt{\rho g x_1} + a_2\sqrt{\rho g x_2} + b(x_1 - x_1^{ref}) - A_1\dot{x}_1^{ref} + (1-\gamma)u \tag{14}$$

To compensate for faults and improve the control system robustness, the extra term u_4 shown in (15) is added to the expression (5) for u.

$$u = \left(\frac{1}{1-\gamma}\right)\left[-a_2\sqrt{\rho g x_2} + a_1\sqrt{\rho g x_1} - b(x_1 - x_1^{ref}) + A_1\dot{x}_1^{ref} + u_4\right] \tag{15}$$

Choosing $u_4 = -K\int res$ yields the residual dynamics (16):

$$\dot{res} + Kres = (1-\gamma_n)\dot{\delta}_u \tag{16}$$

Thus, with constant δ_u, res goes asymptotically to zero with time constant $1/K$. As already anticipated, this forces x_e to approach asymptotically the desired error dynamics defined in the TBG of Figure 2.

Representing u_4 in terms of $(x_1 - x_1^{ref})$ yields (17), expression showing that, in this case, the residual signal defined in the mDBG has a PI structure. Note however that this does not necessarily generalize, since the resulting structure depends on the TBG.

$$u_4 = -KA_1(x_1 - x_1^{ref}) - Kb\int(x_1 - x_1^{ref}) \tag{17}$$

4. SIMULATIONS RESULTS

The parameters used in the simulations are: $A_i = 28\ cm^2$, $\rho = 2\frac{gr}{cm^3}$, $g = 981\ \frac{cm}{s^2}$ and $a_i = 0.71cm^2$ (Johansson 2000), $b = 1$, $K = 1$, $\gamma = 0.5, \gamma_r = 0.45$. The control law given in (15), (17) is used.

The simulation scenario concerns abrupt faults in the system and the measurements are noise free.

The dynamic response of the control system with different faults occurring at time $T = 75s$ is shown in Figures 9–13. As can be seen from these figures, x_1 recovers its reference level $x_1^{ref} = 30cm$ after the fault occurrence, and the residual signal, which is sensitive to the faults considered, tends to zero while x_2 remains bounded.
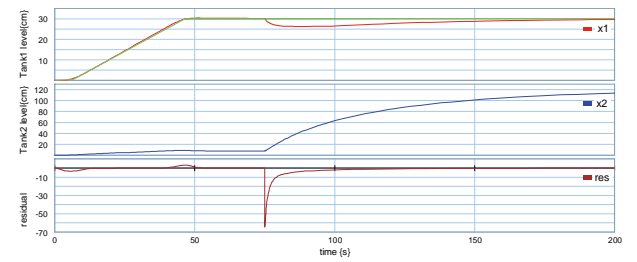

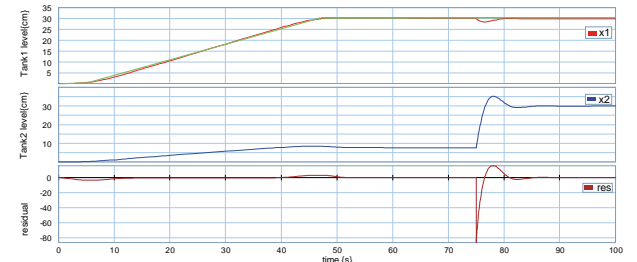Figure 9. 75% obstruction in the outlet hole of tank2

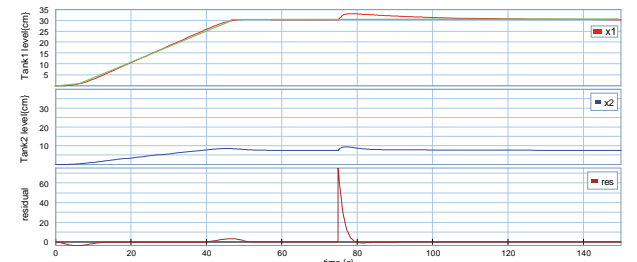Figure 10. Fault in the distribution valve, $\gamma = 1$

Figure 11. Fault in the level sensor of tank2, measured level equal to zero.
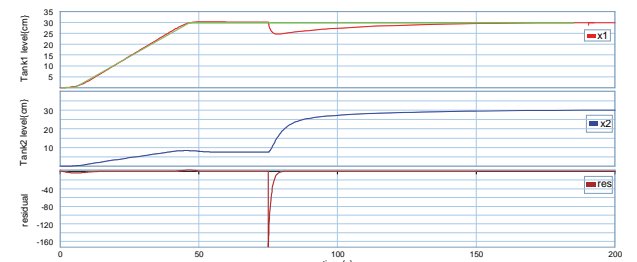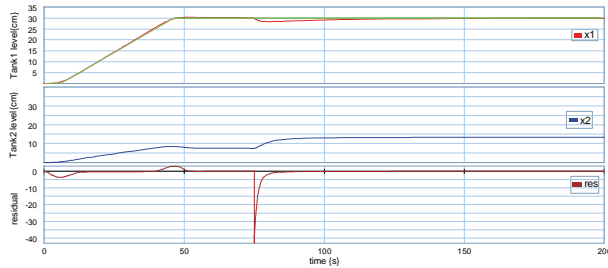

Figure 12. Leakage in tank1.

Figure 13. Fault in the discharge flow from tank2 to tank1, only 50% enters into tank1.

Figure 14 shows the dynamical response of the controlled system under multiple sequential faults. The simulation scenario is as follows, at time $T = 75s$ a 75% obstruction in the outlet hole of tank2 happens, at time $T = 250s$ the distribution valve fails ($\gamma = 1$), at time $T = 400s$ the level sensor of tank2 measures 50% of its actual value, at time $T = 550s$ a leakage in tank1 appears with outlet hole cross section $0.6\ cm^2$, finally, at time $T = 700s$ only the 50% of the outlet flow of tank2 enters into tank1.
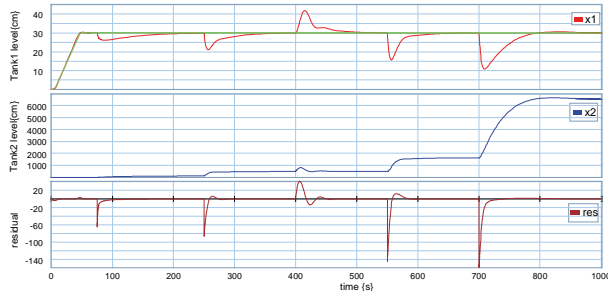

Figure 14. Multiple sequential faults.

As seen in the figures above, the control system is fault tolerant under different faults scenarios. However, this fault tolerance depends on the tank2 capacity, i.e., no overflow is modeled, in a real case tank2 could overflow. In such a case the controller should be reconfigured to manage the faults.

Figure 15 shows the behavior under a simultaneous fault occurrence at time $T = 75s$ in the distribution valve and in tank2 outlet hole ($\gamma = 1$ and 75% obstruction, respectively). Some measurement noise (normal distribution and amplitude $n = 0.1\ cm$) has been considered. In this case the controller again rejects the faults forcing the tank1 level to follow its reference and the residual signal remains close to zero.
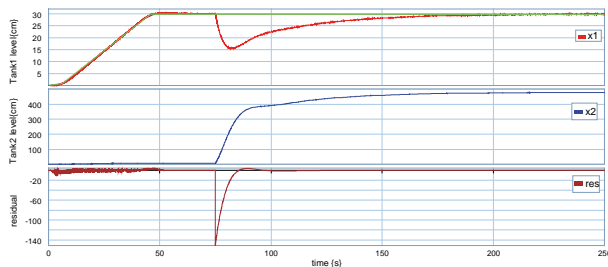

Figure 15. Simultaneous faults, $\gamma = 1$ and 75% obstruction in the tank2 outlet hole.

5. CONTROLLER RECONFIGURATION AND AFTC

This section gives some clues on how to use the method proposed in this paper in the context of AFTC, i.e., when it becomes mandatory, or just convenient, to reconfigure the control law. This is performed on the same case study handled along this paper now considering a structural fault and a sensor fault.

5.1.1. Fault in the distribution valve

Consider again the structural fault $\gamma = 1$ in the distribution valve, i.e., the discharge into tank1 is blocked and all the flow is directed into tank2. This fault has already been simulated with satisfactory results *using $\gamma = \gamma_r = 0.5$ as rated parameter in the controller*. However, if the fault were known it could be of interest to use the real parameter $\gamma = 1$ in the control law. This cannot be done using the former laws (15), (17), as the controller is not defined. This calls for controller reconfiguration. The same method presented before can be used to solve the problem, but in this case the virtual prototyping of the power control source should be made over tank2, the only way to causally access the dynamics of tank1. The procedure of obtaining a closed loop equivalent to the TBG is suggested by the virtual BG of Figure 16.
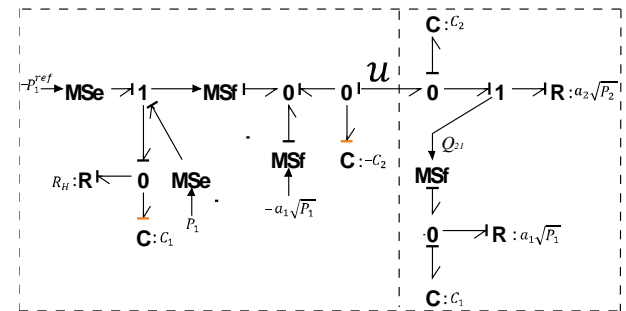
Figure 16. Virtual BG matching the TBG in case of a structural fault in the distribution valve.

The control law obtained reading the BG of figure 16 is:

$$u = -A_2 \dot{x}_2 + u_2 + u_3 \qquad (18)$$

Equation (18) cancels the tank2 dynamics to achieve the TBG. *Mutatis mutandis*, here again an extra integral term can be added to the controller, as shown in the previous section, in order to improve the fault tolerance.

This procedure can be repeated for all structural faults, obtaining a set of control laws which can be switched, with the help of an FDI algorithm, to improve the performance of the control system.

5.1.2. Sensor faults and controller reconfiguration

When a fault occurs in the level sensor of tank2, then the control system becomes unobservable. However, the control objectives can still be achieved just discarding the term in (5) that depends on x_2 because u_4

compensates the flows mass differences. But if the level sensor of tank1 fails, then the control must be reconfigured in order to regulate the level of Tank1, because x_1 cannot be injected into the mDBG.

Using the steady state relationship (2) it is possible to define a new TBG to handle faults in sensor level x_1 as depicted in Figure 17.
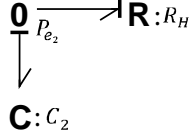

Figure 17. Proposed TBG to handle faults in sensor level $x1$.

$$\dot{x}_{e2} = -\frac{b}{A_2} x_{e2} \qquad (19)$$

In the equation above, the tracking error $x_{e2} = x_2 - x_2^{ref}$ is the state variable of the (incremental) TBG and x_2^{ref} is the tank1 reference level. x_2^{ref} is related to x_1^{ref} through the steady state relationship (2). The same method presented before can be used to solve the problem. Figure 18 shows the associated virtual BG and (20) is the control law obtained.
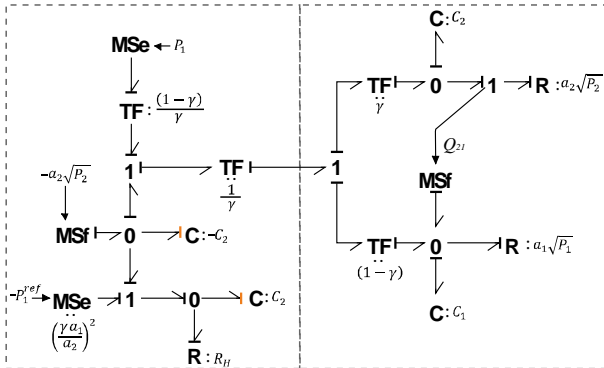

Figure 18. Virtual BG matching the TBG in case of a fault in the sensor level x1.

$$u = \frac{1}{\gamma}\left[-a_2\sqrt{\rho g x_2} - b\left(x_2 - \left(\frac{a_1}{a_2\gamma}\right)^2 x_1^{ref}\right) + A_2\left(\frac{a_1}{a_2\gamma}\right)^2 \dot{x}_1^{ref}\right] \qquad (20)$$

Of course the control systems will follow the reference if and only if only sensor x_1 fails.

Summarizing, controller reconfiguration can be used to get a performance better than that obtained with (15) after a fault occurrence, or in case that the TBG must be modified. As AFTC is beyond the scope of this paper, to get an improved closed-loop behavior with (15) while staying in the context of PFTC, the use of a gain scheduling approach is recommended, performing controller fault accommodation varying the constant K in dependence on the faults.

6. CONCLUSIONS

This work addressed the PFTC approach in the BG domain. The obtained control law is calculated through a energy and power shaping method. An extra term was added to the control law to improve robustness to faults in different components of the control system. Simulations demonstrate the good response and the fault tolerance of the control system.

Further work will be aimed at generalizing the method using the relations between BG and the Port-Controlled Hamiltonian Systems theory.

REFERENCES

Blanke, M., Kinnaert, M., Lunze, J., Staroswiecki, M. *"Diagnosis and Fault-Tolerant Control"*. Second Edition, Springer-Verlag.. 2006.

Borutzky, W. *"Bond graph model-based fault detection using residual sinks"*. Proceedings of the Institution of Mechanical Engineers, Part I: *Journal of Systems and Control Engineering*, Volume 223, 2009, pp. 337-352

Controllab Products B.V., 20-sim, http://www.20sim.com/

Donaire, A. Junco, S. *"Energy shaping, interconnection and damping assignment, and integral control in the bond graph domain"*. Simulation Modelling Practice and Theory 17 (2009) 152?174.

Isermann, R. *"Fault-Diagnosis Systems"* Springer-Verlag. 2006.

Junco, S. *"Virtual Prototyping of Bond Graphs Models for Controller Synthesis through Energy and Power Shaping"*. Conference on Integrated Modeling and Analysis in Applied Control and Automation (IMAACA 2004).

Johansson, K. *"The Quadruple-Tank Process---A multivariable laboratory process with an adjustable zero"*. IEEE Transactions on Control Systems Technology, 2000.

Khalil, H. *"Nonlinear Systems",* Second Edition., Prentice-Hall, New Jersey, 1996.

Ould Bouamama, B. Samantaray, A. Staroswiecki, M. Dauphin-Tanguy, G. *"Derivation of Constraint Relations from Bond Graph Models for Fault Detection and Isolation"* Proc. ICBGM 2003.

Samantaray, A. Medjaherb, K. Ould Bouamama, B. Staroswieckic, M. Dauphin-Tanguy, G. *"Diagnostic bond graphs for online fault detection and isolation"*. Simulation Modelling Practice and Theory Volume 14, Issue 3, April 2006, pages 237-262.

Sontag, E. *"Mathematical Control Theory: Deterministic Finite Dimensional Systems"*. Second Edition, Springer, New York, 1998.

Sueur, C. Dauphin-Tanguy, G. *"Bond graph approach for structural analysis of MIMO linear systems"*, *Journal of the Franklin Institute*. 328 (1), 1991, pages 55-70.

Zhang, Y. Jiang. J. *"Bibliographical review on reconfigurable fault-tolerant control systems"* Annual Reviews in Control, Vol. 32, No. 2. (December 2008), pages 229-252

On Algebraic Approach for MSD Parametric Estimation

Marouene Oueslati[(a,b)], **Stéphane Thiery**[(a,b)], **Olivier Gibaru**[(a,b)], **Richard Bearee**[(a)] and **George Moraru**[(a)]

[(a)] Laboratory of Science of Information and Systems LSIS, Arts et Métiers ParisTech
INSM project-team 8, Bd Louis XIV, 59046 Lille Cedex

[(b)] Non-A project-team, Centre de recherche INRIA Lille-Nord Europe
40, avenue Halley. Bât.A 59650 Villeneuve d'Ascq - France

marouene.oueslati@ieee.org, {stephane.thiery, olivier.gibaru, richard.bearee, george.moraru}@ensam.eu

ABSTRACT

This article address the identification problem of the natural frequency and the damping ratio of a second order continuous system where the input is a sinusoidal signal. An algebra based approach for identifying parameters of a Mass Spring Damper (MSD) system is proposed and compared to the Kalman-Bucy filter. The proposed estimator uses the algebraic parametric method in the frequency domain yielding exact formula, when placed in the time domain to identify the unknown parameters. We focus on finding the optimal sinusoidal exciting trajectory which allow to minimize the variance of the identification algorithms. We show that the variance of the estimators issued from the algebraic identification method introduced by Fliess and Sira-Ramirez is less sensitive to the input frequency than the ones obtained by the classical recursive Kalman-Bucy filter. Unlike conventional estimation approach, where the knowledge of the statistical properties of the noise is required, algebraic method is deterministic and non-asymptotic. We show that we don't need to know the variance of the noise so as to perform these algebraic estimators. Moreover, as they are non-asymptotic, we give numerical results where we show that they can be used directly for online estimations without any special setting.

Keywords: Parameter estimation; Recursive algorithm; Kalman-Bucy algorithm; Forgetting factor; Algebraic approach; Laplace transform; Operational calculus; Leibniz formula; Integral rules; Filtering.

1. INTRODUCTION

Since a wide large of mechanical systems are modeled through coupled or isolated Mass Springer Damper systems, the estimation problem of the MSD parameters is classic in nature. Moberg et al [1] have modeled a 2 link of an ABB industrial robot based on serial MSD system for each axis. This is done aiming to simplify the related elastic dynamic equations. Also, a double mass model of an elastic cam mechanism was described in [2], that gives a more realistic idea of the relationship in mass distribution in the process. The method introduced in this article concerns the parameters estimation problem based on a new algebraic method introduced by Fliess and Ramirez [3] and compare to a conventional algorithm proposed through the Kalman-Bucy filter in parameter estimation. These algebraic parametric estimation techniques for linear systems [4] have been extended for various problems in signal processing for example [5],[6],[7],[8],[9],[10]. Let us emphasize that those methods, which are non-asymptotic, exhibit excellent robustness properties with respect to corrupting noises, without the need of knowing their statistical properties [19]. We propose to apply this algebra based approach for identifying parameters of a Mass Spring Damper (MSD) excited by a sinusoidal input. Similar approach is proposed in [8], however the novelty of this article, is to compare two types of identification algorithms based on finding the optimal input solution in order to well and quickly identify the mechanical system parameters. We perform a numerical study to obtain the optimal solution in case when a wave generator is used as excitation signal. The optimal input signal design depends on two parameters : frequency ω_1 and the amplitude that gives the best training exciting trajectory. We compare the results to the ones obtained via a classical recursive approach [11], [12], [13], [14]. In particular, this method is compared to a weighted Kalman-Bucy filter [13] in order to show the robustness and the efficiency of the proposed technique where measurements are corrupted by a noise. We study the effect of a Gaussian noise added to the output on the estimators variance [15]. This is performed by taking the sampling period into account. We focus on optimal input excitation in order to maximize the convergence rate of estimators based on minimum variance analysis [16]. Hence, we compare the algebraic method variance with the one of the Kalman-Bucy filter. This variance analysis allows us to show that contrary to the recursive approach [17], [18], the algebraic method is less sensitive to the value of the exiting frequency input and more robust to the corrupted noise. Moreover, the Kalman-Bucy approach needs the knowledge of the statistical properties of the noise that is not required for the algebraic method [19]. We show that this method is also robust even for a high frequency sinusoidal disturbance. Online identification of the unknowns undamped angular frequency and the damping ratio is devoted. The identified parameters are obtained in finite time and the noise effect is attenuated by the iterated integrals. Numerical results show the accuracy of the estimation and the best training signal design.

The outline of this paper is as follows. Section 2 describes the problem statement. Section 3 contains a variance analysis of identified parameters thorough the Kalman-Bucy algorithm and the corresponding exciting trajectory design. Mathematical framework for algebraic parameters estimation is presented in Section 4 which also contains the estimation methods following from the rules of operational calculus. Simulations and comparative analysis of estimators are proposed in Section 5, while Section 6 concludes the paper.

2. PROBLEM STATEMENT

Consider a Mass Spring Damper (MSD) system defined by the following continuous-time second-order system :

$$\ddot{x}(t) + 2\zeta\omega_0\dot{x}(t) + \omega_0^2 x(t) = u(t), \tag{1}$$

where $u(t)$ is the input, ω_0 is the natural undamped frequency and $\zeta \in [0,1]$ is the damping ratio. One can note that physically $\omega_0 = \sqrt{\frac{k}{m}}$ is the undamped angular frequency of the mechanical system and $\zeta = \frac{c}{2\sqrt{km}}$ be the damping ratio; Where c is the viscous damping coefficient, k denotes the spring constant and m the mass of the load. We set $\theta_1 = 2\zeta\omega_0$, $\theta_2 = \omega_0^2$ and $u(t) = A_1\sin(\omega_1 t)$. Let $\tilde{x}_i = x_i + \varpi_i$ be a noisy observation of the "true" position $x_i = x(t_i)$ of the system at $t_i = iT_s$ for $i = 0,\ldots,N$. The real value T_s denotes the sampling period. We assume that ϖ is an additive noise corruption which is a second order continuous stochastic process with zero-mean and a known variance σ^2. Consequently, we search the values of ω_1 which allows us to estimate θ_1 and θ_2 with the minimum variance for a given time estimation.

3. KALMAN-BUCY FILTER ESTIMATORS

A. Introduction

This section aims to use the Kalman-Bucy filter [13] so as to estimate the vector $\Theta = (\theta_1, \theta_2)^T$ which is involved in the motion equation (1). In order to quickly identify these parameters through an optimal designed sinusoidal input, a variance analysis of the estimator is described in the following. This will allow us to optimally choose the values of A_1 and ω_1. The input sequence $(u_i)_{i=1,\ldots,N}$ and the output sequence $(\tilde{x}_i)_{i=1,\ldots,N}$ are measured synchronously at the sampling period T_s. Consequently, we obtain the following linear relations from these measurements :

$$Y_k = X_k\Theta + \rho_k, \text{ with } m < k \leq N, \tag{2}$$

where the regression matrix $X_k = \left(\begin{array}{cc} (\dot{x}_i)_e & \tilde{x}_i \end{array} \right)_{i=m+1,\ldots,k}$, the observed signal vector $Y_k = (u_i - (\ddot{x}_i)_e)_{i=m+1,\ldots,k}$ and $(\dot{x}_i)_e$ (resp. $(\ddot{x}_i)_e$) is the velocity estimation (resp. acceleration estimation) at $t_i = iT_s$. We assume that ρ_k is a sequence of independent Gaussian variables with zero mean and known variance σ_ρ^2 issued from the variance estimators due to both of the measurement noises ϖ and the derivative estimation errors. Moreover, the integer m is the minimum value needed so as to calculate $(\dot{x}_i)_e$ and $(\ddot{x}_i)_e$. Usually, these estimators are computed through a filtered finite numerical differentiator [22],[23].

From now on, the problem is to estimate Θ based on the measurements and the observed signal vector. We consider the situation when the observations are obtained one-by-one from the process. We would like to update the parameters estimate whenever new observation to the previous set of observations. In what follows, a recursive formulation is derived. Instead of recomputing the estimates with all available data, the previous parameters estimate are updated with the new data sample. In order to do this, the Kalman-Bucy filter is written in the form of a recursive algorithm. The recursive algorithm is given by the following structure:

$$\begin{cases} K_{k+1} &= P_k X_{k+1}^T (R_{k+1} + X_{k+1} P_k X_{k+1}^T)^{-1}, \\ \alpha_{k+1} &= Y_{k+1} - X_{k+1}\hat{\Theta}_k, \\ \hat{\Theta}_{k+1} &= \hat{\Theta}_k + K_{k+1}\alpha_{k+1}, \\ P_{k+1} &= \lambda^{-1}(P_k - K_{k+1}X_{k+1}P_k), \end{cases} \tag{3}$$

where $\hat{\Theta}_k$ is the parameters estimation vector after the first $k-$samples and $\lambda \in]0,1]$ is a forgetting factor which reduces the influence of old data. In particular, if $\lambda = 1$, then all the data are taken into account in the same manner. In this algorithm (3), one notes that the vector Θ_k and the matrix P_k are involved in the recursions. In order to initialize the algorithm, we must provide initial values for these variables. We choose to apply the Ordinary Least Square solution of this identification problem by using a "small" samples of the first m-measures $(\tilde{x}_i)_{i=1,\ldots,k}$ to compute

$$\hat{\Theta}_m = P_m B_m, \quad \text{where } \begin{cases} P_m = (X_m^T R_m^{-1} X_m)^{-1}, \\ B_m = X_m^T R_m^{-1} Y_m. \end{cases} \tag{4}$$

Let us denote

$$\alpha(i) = k - \max\{i - m, k\} \text{ for } i \in \{m+1,\ldots,k\} \tag{5}$$

After $k \geq m$ stacked samples, by applying recursions (3) initialized with (4), one can recursively obtain the following estimation

$$\hat{\Theta}_k = \frac{\sum_{i=m+1}^k \lambda^{\alpha(i)} X_i Y_i}{\sum_{i=m+1}^k \lambda^{\alpha(i)} X_i^2} \tag{6}$$

with $K_k = \dfrac{X_k}{\sum_{i=m+1}^k \lambda^{\alpha(i)} X_i^2}$ and $P_k = \dfrac{\sigma_\rho^2}{\sum_{i=m+1}^k \lambda^{\alpha(i)} X_i^2}$. $\tag{7}$

B. Variance analysis

In this subsection, we are interested in the variance analysis of the estimation (6), aiming to find the input trajectory $u(t)$ i.e. the values of $(A_1)_{opt}$, and $(\omega_1)_{opt}$, which allow to minimize the variance of (6). The value $(\omega_1)_{opt}$ is investigated in term of the optimal ratio $Z_{opt} = \frac{(\omega_1)_{opt}}{\omega_0}$.

Besides, for small values of ζ, the dynamic equation (1) can be simplified by neglecting the damping effect based on a numerical simulation of the differential equation. For example, in Fig 1, we compare the difference between the exact solutions of (1) with $\zeta = 0.0021$ and $\zeta = 0$.

This will be used so as to simplify the variance analysis. Also, this approximation will take place only in order to
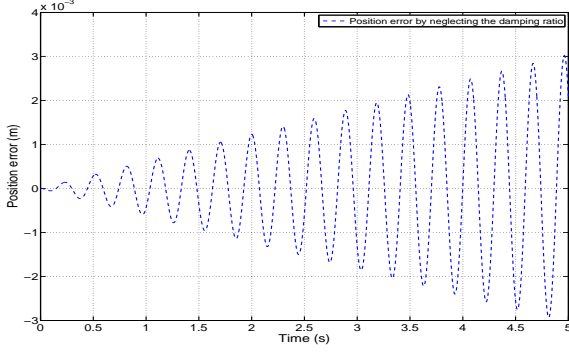
Fig. 1. Error of the exact solution of $x(t)$ with $\zeta = 0.0021$ (real damping ratio) and $\zeta = 0$ and a given sinusoidal exciting trajectory

perform the Kalman-Bucy filter variance of Θ, $Var(\Theta)$, in term of the ratio $Z = \dfrac{\omega_1}{\omega_0}$. This is done in order to find a variance expression of the recursive estimator. However, Kalman-Bucy algorithm in parameter estimation will be rebuilt, by means of (2) and (3), in order to estimate the unknowns parameters θ_1 and θ_2 based on the calculated variance expression. Under this assumption, in order to perform the variance expression, Θ is limited to the scalar variable θ_2. Moreover, the regression matrix X_k can be rewritten $X_k = (\tilde{x}_i)_{i=m+1,\dots,k}$. The explicit solution of this reduced differential equation becomes :

$$x(t) = \frac{A_1 \left[\omega_1 \sin(\omega_0 t) - \omega_0 \sin(\omega_1 t)\right]}{\omega_0(\omega_1^2 - \omega_0^2)}. \tag{8}$$

We denote $P_k = ((X_k R_k^{-1} X_k)^T)^{-1}$, where R_k is a diagonal matrix

$$R_k = diag(\underbrace{r_1, \dots, r_{k-m}}_{k-m \text{ times}}), \tag{9}$$

with the $r_j > 0$ and $e_k = Y_k - X_k \hat{\Theta}_{k-1}$ is the a priori error of estimation. Consequently, the Kalman-Bucy filter consists of two stages. The first part employs an estimate $\hat{\Theta}_k$ using the information already available at time k and the second part provides the main time-update made by the innovation process (a priori errors), denoted α_{k+1} in (3), in order to estimate $\hat{\Theta}_{k+1}$ from measurements Y_{k+1}, regression X_{k+1} and $\hat{\Theta}_k$.

In fact, ρ_k depicts a white noise vector with zero mean and it is defined by the following autocorrelation function

$$\mathbb{E}[\rho(t)\rho^*(t-\tau)] = \begin{cases} \sigma_\rho^2, & \tau = 0, \\ 0, & \tau \neq 0. \end{cases} \tag{10}$$

Concerning the matrix P_k, it represents the variance-covariance matrix of the estimation error.

$$P_k = cov[e_k] = \mathbb{E}[(\hat{\Theta}_k - \Theta)^T(\hat{\Theta}_k - \Theta)].$$

At this stage, the developments below, will be based on the Kalman-Bucy algorithm with a fixed variance, i.e., for any $k \geq m$, $r_{k-m} = \sigma_\rho^2$.

Therefore, by applying the linearity property of the variance, we obtain the above variance expression of (6)

$$Var(\hat{\Theta}_k) = \frac{\sigma_\rho^2 \sum\limits_{i=m+1}^{k} \lambda^{2\alpha(i)} X_i^2}{\left(\sum\limits_{i=m+1}^{k} \lambda^{\alpha(i)} X_i^2\right)^2}. \tag{11}$$

Relation (11) can be expressed by using the explicit solution (8), as follows

$$Var(\hat{\Theta}_k) = \frac{\sigma_\rho^2}{A_1^2} K(Z, \lambda, \omega_0, T_s, m, k)$$

where

$$K(Z, \lambda, \omega_0, T_s, m, k) =$$
$$\frac{(\omega_0^2(Z^2-1))^2 \sum\limits_{i=m+1}^{k} \lambda^{2\alpha(i)}(Z\sin(\omega_0 t_i) - w_0 \sin(Z\omega_0 t_i))^2}{\left(\sum\limits_{i=m+1}^{k} \lambda^{\alpha(i)}(Z\sin(\omega_0 t_i) - \omega_0 \sin(Z\omega_0 t_i))^2\right)^2}. \tag{12}$$

Hence, the minimization of the variance of the Kalman-Bucy estimator may be obtained by increasing the magnitude A_1 of the input force. However, this strategy is naturally restricted by some physical limits. Concerning the variable ω_1 i.e. the ratio $Z = \frac{\omega_1}{\omega_0}$, it will be explained in next subsection.

C. Influence of the forgetting factor λ

In a first series of experiments, we investigate the influence of the forgetting factor λ on the value of $K(Z, \lambda, \omega_0, T_s, m, k)$, Fig 2. In fact, Fig 3 shows the logarithm value of $K(Z, \lambda, \omega_0, T_s, m, k)$ according to a discretized value of Z belonging to $[0.01, 2]$ where the sampling period $T_s = 0.001$ s, $k = 100$ and $m = 3$. A set of different values of the forgetting factor $\lambda = \{0.95, 0.98, 0.99, 1\}$ is choosen. As we can see, $\lambda = 1$ is always the optimal value for our application.

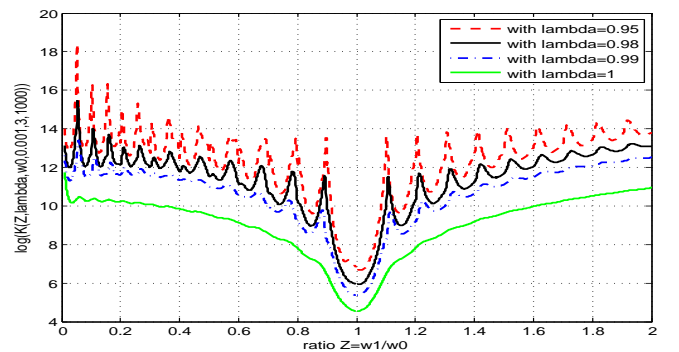Fig. 2. Influence of the forgetting factor λ in variance analysis

D. The optimal input trajectory

Consequently, when $\lambda = 1$, we have

$$K(Z, \omega_0, T_s, m, k) = \frac{\omega_0^4(Z^2-1)^2}{\left(\sum\limits_{i=m+1}^{k} (Z\sin(\omega_0 t_i) - \sin(Z\omega_0 t_i))^2\right)}. \tag{13}$$

A Taylor series at $Z = 1$ allows us to conclude that the minimum value is obtained for $Z = 1$ i.e. $(\omega_1)_{opt} = \omega_0$. Figure 3 depicts the value of $K(Z, \omega_0, T_s, m, k)$ (13), according to Z for different numbers k of samples. The other parameters are the same than those used in the previous subsection.
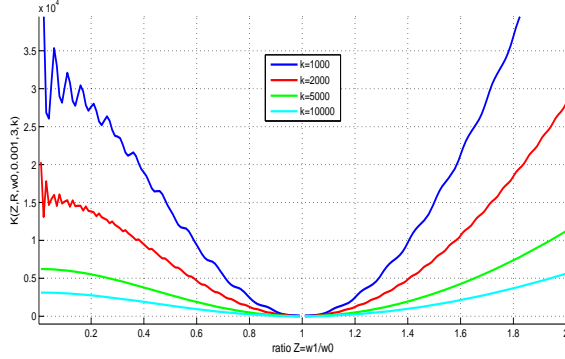

Fig. 3. Influence of the ratio $Z = \frac{\omega_1}{\omega_0}$ for optimal trajectory design

Figure 3 shows that the sensibility of the variance is quite important in the neighborhood of $Z_{opt} = 1$. In conclusion using an input trajectory as closed as possible to the natural frequency of the system, we can consequently minimize the variance of the Kalman-Bucy estimator.

4. ALGEBRAIC PARAMETRIC ESTIMATOR

In this section, we provide the interested reader with rigorous mathematical development in which the algebraic parameter estimation technique, used in this article for the estimation problem, is based. The fundamental developments are based on the module theoretic approach to linear systems [3], [5], [6].

A. Mathematical framework: Generalized expressions of parameters estimation

Set $k = k_0\{\Theta\}$, where k_0 is considered as real, or complex differential field and $\Theta = (\theta_1, ..., \theta_r)$ a finite set of *unknowns parameters* which might not be constant. The unknown parameters $\Theta = (\theta_1, ..., \theta_r)$ are said to be linearly identifiable if, and only if,

$$P(t) \begin{pmatrix} \theta_1 \\ \vdots \\ \theta_r \end{pmatrix} = Q(t) + R(t), \qquad (14)$$

where

- the entries of the matrices $P(t)$ and $Q(t)$, of respective sizes $r \times r$ and $r \times 1$ belong to $span_{k_0(t)[\frac{d}{dt}]}(u, y)$ where (u, y) denotes respectively the vector of inputs and outputs of systems;
- $det(P(t)) \neq 0$;
- R is a $r \times 1$ matrix with entries in $span_{k_0(t)[\frac{d}{dt}]}(\pi)$ which designed the disturbance contribution.

B. Algorithm

We consider the dynamic of a system satisfies the inputs-outputs relation:

$$\sum_{i=0}^{n} a_i \, y^{(i)} = \sum_{i=0}^{m} b_i \, u^{(i)} \qquad (15)$$

where $a_n = 1$, $m < n$. The algebraic parametric estimator is derived using the following steps:

- We apply the Laplace transform of (15):

$$\sum_{i=0}^{n} a_i \, (s^i y(s) - s^{i-1} y_0 - ... - y_0^{(i-1)})$$
$$= \sum_{i=0}^{m} b_i \, (s^i u(s) - s^{i-1} u_0 - ... - u_0^{(i-1)}) \, . \qquad (16)$$

One note the onset of the initial conditions to the $(n-1)^{th}$ order involved in (16).

- By applying n times the derivative operator with respect to s, we can annihilate the initial conditions. This step will be of a great advantage by eliminating these conditions since they are usually unknown. Hence, we obtain algebraic parametric estimator independent from initial conditions. It is aimed to estimate the parameters a_i and b_i in a fast way and on the basis possibly noisy measurements. For this exact expressions of the parameters are derived as a function of the integral of the output and the input, through the following inverse Laplace transform :

Proposition 4.1 *Let Γ be a causal real continuous function and t_0 be a strictly positive real value then for any positive real $T \leq t_0$, $m \in \mathbb{N}^*$ and $n \in \mathbb{N}$ we have*

$$\mathcal{L}^{-1}\left(\frac{1}{s^m} \frac{d^n \hat{\Gamma}(s)}{ds^n}\right)(T) = \frac{(-1)^n T^{m+n}}{(m-1)!} \int_0^1 (1-\tau)^{m-1} \tau^n \gamma(t_0 - \tau T) d\tau, \qquad (17)$$

where $\hat{\Gamma}(s)$ is the Laplace transform of the continuous function $\Gamma(t) = \gamma(t_0 - t)$.

Proof: By applying the Cauchy formula, we obtain for any $T \geq 0$,

$$\mathcal{L}^{-1}\left(\frac{1}{s^m} \frac{d^n \hat{\Gamma}(s)}{ds^n}\right)(T) = \frac{(-1)^n}{(m-1)!} \int_0^T (T-u)^{m-1} u^n \Gamma(u) du.$$

If we assume that for any $u \leq t_0$, $\Gamma(u) = \gamma(t_0 - u)$ then by substituting u by τT we obtain

$$\int_0^T (T-u)^{m-1} u^n \Gamma(u) du = T^{m+n} \int_0^1 (1-\tau)^{m-1} \tau^n \gamma(t_0 - \tau T) d\tau.$$

Consequently (17) holds. ∎

By now on, we set

$$P_{m,n,T}^{(\gamma)}(t_0) = \frac{(-1)^n T^{m+n}}{(m-1)!} \int_0^1 (1-\tau)^{m-1} \tau^n \gamma(t_0 - \tau T) d\tau, \qquad (18)$$

where γ is either the system output $y(t)$ or the input $u(t)$ and $T > 0$ is the time length of the sliding window estimation. Let us denote t_0 is the initial step time for each sliding window i.e. time estimation T along the simulation time vector t. This estimation time T may be small especially in the absence of noise. Meanwhile, T cannot obviously be taken arbitrary small even in a noise-free context. A lower bound for T has been formally characterized in [[19], Prop. 3.2], within the framework of nonstandard analysis.

C. Application

Applying the previous rules to (1), we obtain an explicit formula for the estimates $\hat{\theta}_1$ of θ_1 and $\hat{\theta}_2$ of θ_2, as a function of the estimation time t_0 on a sliding windows of length T. We firstly apply the Laplace transform to the differential equation (1). This gives the following equation:

$$s^2 X(s) - sx(0^+) - \dot{x}(0^+) + \theta_1(sX(s) - x(0^+))$$
$$+ \theta_2 X(s) = u(s).$$

In order to eliminate the initial conditions $x(0^+)$ and $\dot{x}(0^+)$, we apply the derivative operator with respect to s two times. It leads to:

$$2X(s) + 4s\frac{dX(s)}{ds} + s^2\frac{d^2X(s)}{ds^2} + \theta_1\left[s\frac{d^2X(s)}{ds^2} + 2\frac{dX(s)}{ds}\right]$$
$$+ \theta_2\frac{d^2X(s)}{ds^2} = \frac{d^2U(s)}{ds^2}. \tag{19}$$

Multiplying the equation (19) by $s^{-\mu}$, $\mu \geq 3$, and applying the inverse Laplace transform (18), we obtain a set of linear equations in the unknown parameters $\Theta = (\theta_1 \quad \theta_2)^T$ in the time domain. It is expressed in terms of a linear combination of iterated convolution integrals over $x(t)$ and $u(t)$. Consequently

$$\begin{pmatrix}\hat{\theta}_1(t_0) \\ \hat{\theta}_2(t_0)\end{pmatrix} = \begin{pmatrix} 2P^{(x)}_{\mu,1,T}(t_0) + P^{(x)}_{\mu-1,2,T}(t_0) & P^{(x)}_{\mu,2,T}(t_0) \\ 2P^{(x)}_{\mu+1,0,T}(t_0) + P^{(x)}_{\mu,1,T}(t_0) & P^{(x)}_{\mu+1,2,T}(t_0) \end{pmatrix}^{-1}$$
$$\begin{pmatrix} -2P^{(x)}_{\mu,0,T}(t_0) - 4P^{(x)}_{\mu-1,1,T}(t_0) - P^{(x)}_{\mu-2,2,T}(t_0) + P^{(u)}_{\mu,2,T}(t_0) \\ -2P^{(x)}_{\mu+1,0,T}(t_0) - 4P^{(x)}_{\mu,1,T}(t_0) - P^{(x)}_{\mu-1,2,T}(t_0) + P^{(u)}_{\mu+1,2,T}(t_0) \end{pmatrix} \tag{20}$$

As in [6], we could get another estimators by applying a derivation to (19) before applying the $\frac{1}{s}$ operator. For an experimental design of the algebraic estimator, a discretization of the integral in (18) will be held by using the Simpson's rule * [20].

$$P^{(\gamma)}_{m,n,T}(t_0) = \frac{(-1)^n}{(m-1)!}T^{m+n}\int_0^1 (1-\tau)^{m-1}\tau^n\gamma(t_0 - \tau T)d\tau$$
$$\approx \frac{(-1)^n}{(m-1)!}T^{n+m}\frac{Ts}{3}\left[\left((1-\tau)^{m-1}\tau^n\gamma(t_0 - \tau T)\right)_{(0)}\right.$$
$$+ 2\sum_{j=1}^{L/2-1}\left((1-\tau)^{m-1}\tau^n\gamma(t_0 - \tau T)\right)_{(2j)}$$
$$+ 4\sum_{j=1}^{L/2}\left((1-\tau)^{m-1}\tau^n\gamma(t_0 - \tau T)\right)_{(2j-1)}$$
$$\left. + \left((1-\tau)^{m-1}\tau^n\gamma(t_0 - \tau T)\right)_{(n)}\right] \tag{21}$$

where L represents the sampling window T length in samples: $L = \frac{T}{T_s}$. Therefore, let us quote the following remarks:

- The estimation time T may be small, resulting in fast estimation.

*Simpson's rule is employed aiming to reduce the numerical integration error compared to the trapezium rule and not to complicate the numerical implementation.

- The noise effect is attenuated by the iterated integrals (low pass filter).
- The computational complexity is low.

5. SIMULATIONS AND COMPARATIVE ANALYSIS

Computer simulations were carried out with the Matlab-Simulink software. Simulations are achieved on the dynamic equation

$$\ddot{x}(t) + \theta_1\dot{x}(t) + \theta_2 x(t) = A_1 sin(\omega_1 t) \tag{22}$$

where $x(t)$ is corrupted by a noise with zero-mean and a known variance. This stochastic signal is built by means of sequence of random variables by the instruction *awgn* in the Matlab package which adds white Gaussian noise to the vector signal $x(t)$. A step sampling of $T_s = 0.001s$ is used. The noise level is measured by the signal to noise ratio in dB, i.e., $SNR = 10\ log_{10}\ (\frac{\sum_i |x(t_i)|^2}{\sum_i |\rho(t_i)|^2})$. Simulations are achieved for a spring value $k = 400\ N/m$, a damping coefficient $c = 0.05\ N.s/m$ and a load mass $m = 3\ kg$. Concerning the sinusoidal input $u(t)$, the signal amplitude A_1 was chosen so that it will be the maximum allowed with respect to the limited physical properties of the system. In our case, A_1 is set to $333.3333N$. Fig 4 shows the noisy position for t_i ranging from 0 to 5 seconds. Although, the linear time invariant (LTI) MSD system (1) is discretized in order to perform the identification algorithms for each sampling time.

We note that in this section, most of figures depict the natural frequency estimation and are limited to $\omega_0 = \sqrt{\theta_2}$.
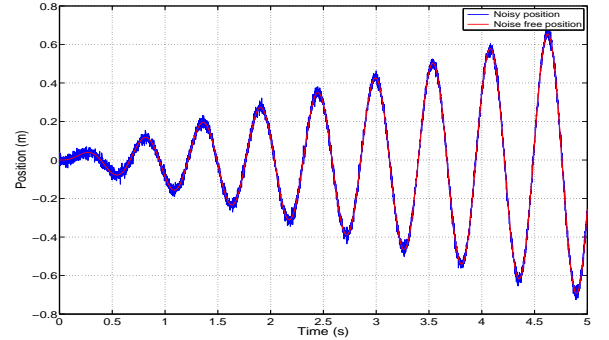

Fig. 4. Noisy Position $x(t)$ with $SNR = 25dB$ for a given sinusoidal input

A. Robustness Analysis

In order to compare the performance of the proposed algebra-based method with the Kalman-Bucy filter, we generate numerical simulations with high level noises which allows us to illustrate the robustness of the parameters estimators involved in (1) with respect to the SNR in dB and the ratio $Z = \frac{\omega_1}{\omega_0}$. Both of estimators algorithms were carried out around two important quantities that reflect the robustness and the performance of the identification methods: signal-to-noise ratio and Z. Fig 5 and 6 depict the weightiness of both SNR and Z in the convergence of each estimation algorithm.
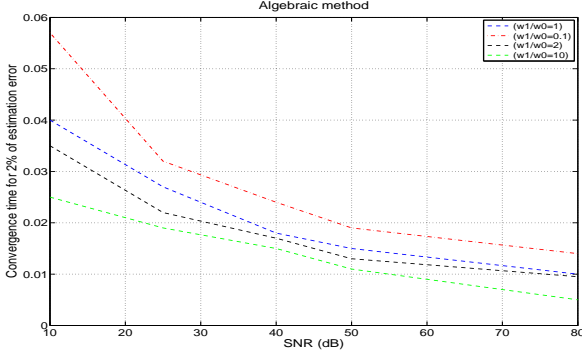
Fig. 5. Convergence time (s) for 2 % of estimation error with Algebraic technique
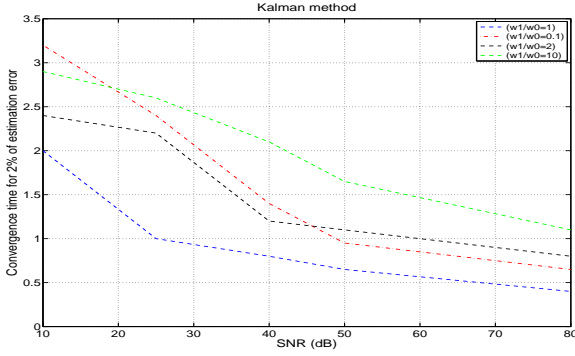

Fig. 7. Algebraic method estimates for ω_0 with SNR = $25dB$

Fig. 6. Convergence time (s) for 2 % of estimation error with Kalman-Bucy algorithm

Fig. 8. Algebraic method estimates for ω_0 with SNR = $50dB$

Fig. 9. Algebraic method estimates for ω_0 with SNR = $80dB$

We assume that the time-estimation is stopped when the absolute value error estimation is less than 2%. Consequently, one can note that the algebraic technique converges as well as possible with respect to the rapidity in time when the period of the signal $u(t)$ is 10 times less than natural period of the MSD system. Moreover, the computation time of ω_0 decreases whenever the SNR and $\frac{\omega_1}{\omega_0}$ is increased. As we can see, for a $SNR = 80$ dB and an angular frequency ratio Z = 10, ω_0 is computed in 0.005s when the sampling time T_s is 0.001s. Fig 7, 8 and 9 depict the algebraic estimation of the natural frequency ω_0 with the presence of a noise effect using the algorithm represented by equations (20). One can note that peaks in Fig 7 are generated from numerical artifacts due to the implementation of (21) and does not exceed 0.0035% of estimation error .We can conclude that this algorithm is non-asymptotic and the noise contribution is attenuated by the presence of iterated integrals after the numerical discretization through the Simpson's rule (21).
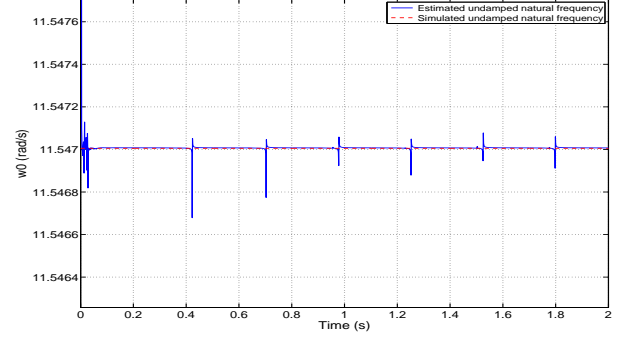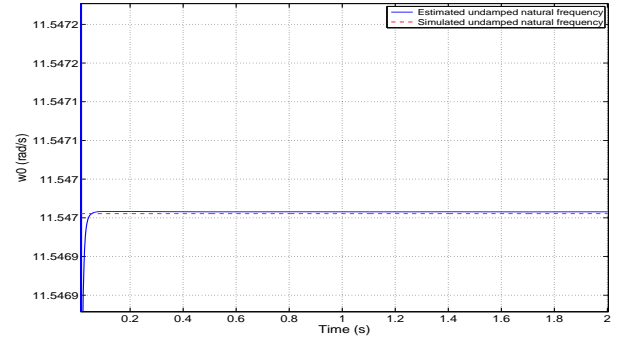
The Kalman-Bucy filter is performed based on the variance analysis as illustrated in section 3.B. This is done through the minimization of the variance expression (11). Fig 3 depicts $Var(\Theta)$ according to $\frac{\omega_1}{\omega_0}$. In fact, as it was shown in Fig 6, $Var(\Theta)$ is minimum when $\omega_1 = \omega_0$. Besides, from Fig 10, 11 and 12 we can conclude that the convergence time decreases when the signal-to-noise ratio in dB increases. However, the convergence time in case of Kalman-Bucy filter is 100 times more as compared to the algebra-based approach for a given SNR. It should be emphasized that for recursive algorithm, the convergence is made asymptotically. From this, it was noted that algebraic technique presents an online-estimator due to the quickness of the parameter computation.

Fig. 10. Kalman-Bucy algorithm estimates for ω_0 with SNR = 25dB

Fig. 11. Kalman-Bucy algorithm estimates for ω_0 with SNR = 50dB

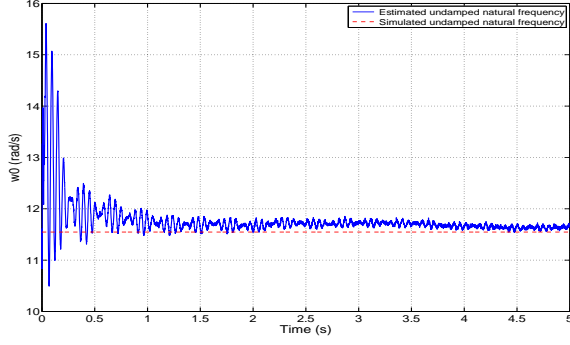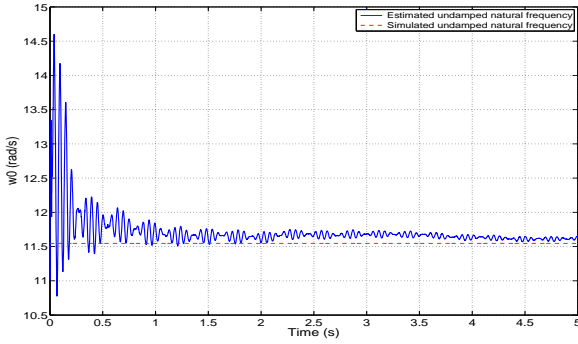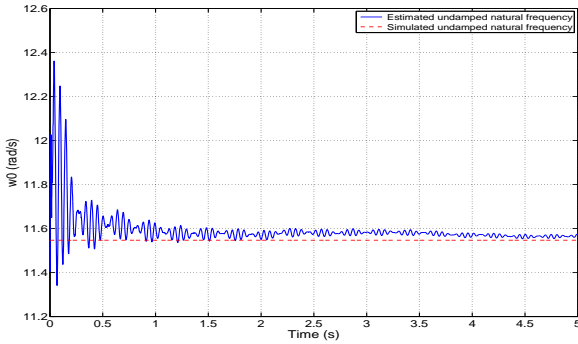Fig. 12. Kalman-Bucy algorithm estimates for ω_0 with SNR = 80dB

B. High frequency sinusoidal perturbation

This section devoted a numerical simulation to evaluate the performance of the proposed algebraic approach compared to the Kalman-Bucy algorithm on a current perturbed position. Unfortunately, this type of experiment involves a severe tempering of the compared estimation algorithms. We consider that the measured position $x(t)$ is corrupted by another sinusoidal perturbation with a higher frequency generator (which is not satisfy the sampling limit) and a white noise process $\rho(0, 0.001)$ with a high signal-to-noise ratio. This can be expressed as $\widetilde{x}(t) = x(t) + \rho(t) + A_2 sin(\omega_2 t)$ where, $A_2 = 0.1$ and $\omega_2 = 500\omega_0$ (Fig 15). The experiments are performed with the optimal conditions for the Kalman-Bucy ($\omega_1 = \omega_0$) and with $\omega_1 = 10 \omega_0$ for the algebra-based algorithms. Fig (14) and (15) depict the estimation of the angular frequency with presence of

a higher frequency sinusoidal. We note that, even the "true" position is highly corrupted Kalman-Bucy filter and the proposed algorithm converge. For the algebra-based technique, the estimations are achieved in about $5 \times T_s$ and $50 \times T_s$ for the recursive algorithm for a 2% of estimation error. It should be noted that the convergence time is faster than for highly Gaussian noisy measurement. Indeed, the robustness of the obtained estimations with respect to the unknown measurement noise is quite high.
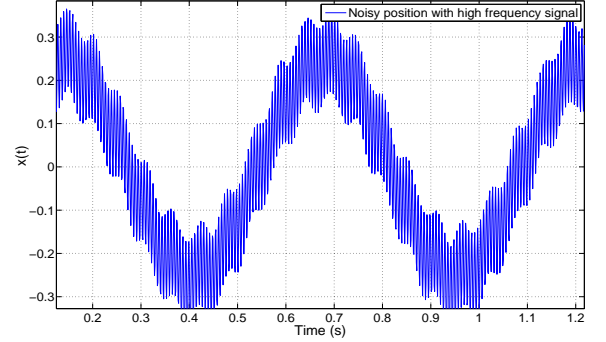
Fig. 13. Position measurement with a high frequency sinusoidal noise
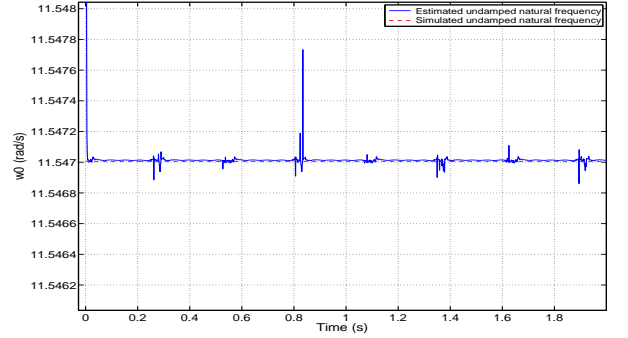
Fig. 14. Algebraic method estimates for ω_0 with SNR = 80dB and high frequency sinusoidal noise
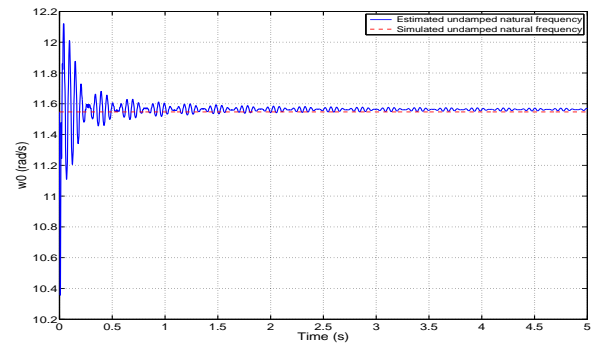
Fig. 15. Kalman filter estimates for ω_0 with SNR = 80dB and high frequency sinusoidal noise

C. Variable parameter estimation

There are many applications where the involved parameters vary in time due to behavior of the system or some physical change [21]. For example, due to the thermal effect, the angular frequency of the MSD system may change with respect to time. This is explained through the fluctuation of spring constant k or the viscous damping coefficient c. To evaluate the performance of the algebraic

89

algorithm, we made an interesting experience where we have simulated the variation of ω_0 with a discontinuity change-point. However, the system still LTI in that range where ω_0 is constant. Therefore, we can apply directly our algebraic algorithm so as to estimate different abrupt changes of the values of $\omega_0 = \sqrt{\frac{k}{m}}$.

Fig 16 shows the performance of the estimation approach where the first jump is carried for $t = 0.1$ s. That result proves the accuracy of the proposed algorithm even if the unknown parameter is time varying with a specific behavior.
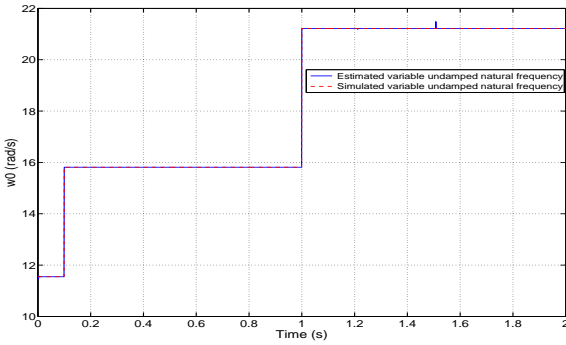
Fig. 16.   Algebraic methods estimates for a variable ω_0

6. CONCLUSION

The objective of this paper has been to study optimal sinusoidal input design so as to minimize the variance parameters estimation of a range of mechanical system. We have presented an algebraic approach to the fast and reliable identification of dynamical parameters of a Mass Spring Damper system, compared to a conventional algorithm introduced by the Kalman-Bucy filter in parameter estimation. The calculation of the characteristic parameters of these mechanical systems was interesting for many reasons touching the engineering theme. As it known, a lot of mechanical structures are modeled via coupled or isolated MSD systems aiming to simplify both of their static and dynamic behaviors . It results an important problem in the control theory such as feedback and feed-forward control where the involved parameters are unknown and should be identified for each time step.

By analyzing the variance estimators, contrary to the Kalman-Bucy filter, we show that the proposed algebraic approach is less sensitive to the choice of the input frequency and is more robust to additive noise on the output. In this study, the numerical differentiation of the output signal employed for the recursive algorithm, was a simple finite difference technique with a low pass filtering. This latter has an influence in the robustness and fastness of the identification for small SNR. Such problems can be minimized by using numerical algebraic differentiators (see [7], [22], [23]). Moreover, we can also directly address the real-time identification of the parameters, where the computational complexity is low as it shown is in the algebra-based approach.

REFERENCES

[1] S. Moberg, J. Ohr, and S. Gunnarsson, (2008), A benchmark problem for robust control of a multivatiable nonlinear flexible manipulator, 17th IFAC World Congress, Seoul, Korea.

[2] B. Horeni, (1992), Double-mass model of an elastic cam mechanism , Mech. Mach Theory, Vol 27 No 4, pp 443-449.

[3] M. Fliess, H. Sira-Ramírez, (2003), An algebraic framework for linear identification, ESAIM Control Optim. Calc. Variat., 9, pp. 151-168.

[4] M Fliess and H Sira-Ramírez, (2008) Closed-loop parametric identification for continuous-time linear systems via new algebraic technique. *Identification of Continuous-time Models from sampled Data*. H. Garnier & L. Wang (Ed.) pp. 362-391

[5] M Fliess, C Join, and H Sira-Ramírez, (2008), Non-linear estimation is easy. *Int. J. Modelling Identification and Control*, 4, 1:12–27.

[6] M. Mboup, (2009), Parameter estimation for signals described by differential equations, Applicable Analysis, Vol. 88 (1), pp. 29-52.

[7] M Mboup, C Join, and M Fliess, (2009), Numerical differentiation with annihilators in noisy environment. *Numerical Algorithms*, 50, pp.439–467.

[8] J.R. Trapero, H. Sira-Ramírez, and V.F. Batlle, (2007), An algebraic frequency estimator for a biased and noisy sinusoidal signal, Signal Processing 87(6), pp. 1188–1201.

[9] J.R. Trapero, H. Sira-Ramírez, V.F. Battle, (2007), A fast on-line frequency estimator of lightly damped vibrations in flexible structures, J. Sound Vibration, 307, pp. 365-378.

[10] J.R. Trapero, H. Sira-Ramírez, V.F. Battle, (2008), On the algebraic identification of the frequencies, amplitudes and phases of two sinusoidal signals from their noisy sums, Int. J. Control, 81, pp. 505-516.

[11] R.E Kalman, P.L Falb, and M.A Arbib (1969), *Topics in mathematical system theory*. McGraw-Hill.

[12] R.E Kalman (1960),   A new approach to linear filtering and prediction. *Journal of Basic Engineering*, 82, D:35–45.

[13] R.E Kalman and R.S Bucy, (1961), New results in linear prediction and filtering theory. *Journal of Basic Engineering*, 83, D:95–100.

[14] L Ljung and V Söderstmö (1983), *Theory and Practice of Recursive Identification*. The MIT Press.

[15] J Schoukens, Y Rolain, and R Pintelon, (2010), On the use of parametric and non-parametric noise-models in time- and frequency domain system identification. , pp. 316–321.

[16] Bo Wahlberg, Hakan Hjalmarsson, and Petre Stoica, (2010), On optimal input signal design for frequency response estimation. , pp. 302–307.

[17] J Jiang and Y Zhang, (2004), A revisit to block and recursive least squares for parameter estimation. *Computers and Electrical Engineering*, 30, Issue 5, pp. 403–416.

[18] I.D Landau (1993), *Identification et commande des systèmes*. 2e édition revue et augmentée. Hermès.

[19] M. Fliess, (2006), Analyse non standard du bruit, CRAS, Série 1, Mathématiques, 342, pp. 797–802.

[20] M Abramowitz and I.A Stegun (1965), *Handbook of mathematical functions*. Dover.

[21] W Yin and A. Saadat Mehr (2010). Least square identification of alias components of linear periodically time-varying systems and optimal training signal design. *IET Signal Processing*, Vol 4, n° 2, pp. 149-157

[22] D. Liu, O. Gibaru, W. Perruquetti (2011), Differentiation by integration with Jacobi polynomials, Journal of Computational and Applied Mathematics, Vol. 235, pp. 3015-3032.

[23] D. Liu, O. Gibaru, W. Perruquetti (2011), Error analysis of Jacobi derivative estimators for noisy signals, Numer. Algo., available online.

[24] D. Liu, O. Gibaru, W. Perruquetti, M. Fliess, M. Mboup, (2009), An error analysis in the algebraic estimation of a noisy sinusoidal signal, IEEE Conference on Control and Automation, pp. 1296 - 1301.

[25] A.Olabi, R. Béarée, O. Gibaru, M. Damak, (2010), Feedrate planning for machining with industrial six-axis robot, Control Engineering in Practice, Vol 18, pp. 471-482.

[26] Kian Jafari, Jerome Juillard, and Eric Colinet (2010), A recursive system identification method based on binary measurements. , 49th IEEE Conference on Decision and Control (CDC'10), pp. 1154–1158.

[27] E Pereira, J.R Trapero, I.M Diaz, and Feliu.V. (2009), Adaptive input shaping for maneuvering flexible structures, *Automatica*, 40 (4), pp.1046–1051.

[28] I Peñarocha and R Sanchis (2010), Adaptive extended kalman filter for recursive identification under missing data. , 49th IEEE Conference on Decision and Control (CDC'10), pp. 1165–1170.

[29] M Verhaehen and V Verdult, (2007), *Filtering and System Identification. A Least Squares Approach.* Cambridge University Press.

[30] K Yosida, (1984), *Operational Calculus.* Springer, New York.

IDENTIFICATION BY HYBRID ALGORITHM
APPLICATION TO THE IDENTIFICATION OF DYNAMICS PARAMATERS OF SYNCHRONOUS MACHINE

Jean Claude RAKOTOARISOA[a], Jean Nirinarison RAZAFINJAKA[b]

[a]Laboratoire d'Automatique, Ecole Supérieure Polytechnique, Diego Suarez, Madagascar
[b]Laboratoire d'Automatique, Ecole Supérieure Polytechnique, Diego Suarez, Madagascar

[a]rajeanclaud1976@yahoo.fr, [b]razafinjaka@yahoo.fr

ABSTRACT
This work treats the identification of the dynamic parameters of a synchronous machine using the test by symmetrical three-phased short circuit. The fitting of these parameters is carried out through the minimization of the quadratic errors between the short circuit currents measured and calculated.
. The approach of hybridization used in this work is the combination of the genetic algorithm (AG) and the algorithm of Levenberg-Marquardt (LM). The results of the optimization are validated first by comparing the parameters obtained by the approach adopted with those estimated with the graphic method of the IEEE standard then by the test of "voltage of reappearance".

Keywords: Identification, synchronous machine, Genetic Algorithm, Levenberg-Marquardt Algorithm

1. INTRODUCTION

The synchronous machine is used mainly for the generation of electrical energy. The dynamic behaviour of this machine is characterized by parameters which take into account the type of pole (smooth pole or projecting pole), the type of windings (with or without shock absorber) and the initial condition. These parameters are identified in the majority of cases by tests.

The reliable and precise determination of these parameters presents better latitude of optimization.
Various measurement techniques and procedures of estimation are used for this purpose. The graphic procedure to estimate the parameters using the short circuit test is well known like dynamic method. It requires the technical graphic for the construction of the envelope and the tangents of the currents of short circuit (IEEE Guide 1995). However, the manual construction of the tangents is not precise.

With the appearance of reliable algorithms of optimization, the idea of this article is to exploit them to optimize the quadratic errors between the measured values and those calculated of the short circuit currents of armature of the synchronous machine in order to identify its dynamic parameters. Various methods were used for this kind of problem. They can be subdivided

in two groups: the first is the least square of which the algorithm of Levenberg-Marquardt is the most used. This algorithm converges quickly but its major disadvantage is the risk of convergence towards the local optima.

The second group is consisted by the method based on the evolution of an individual for example the set of the parameters. The genetic algorithm belongs to this group. The advantage of this group of optimization is the convergence towards the global minimum but it should be noted that it is slow because it requires several evaluations.

In this work, an approach of hybridization is used in order to exploit only the advantages of these two groups of algorithm. The principle of the approach consists in launching a global research with the genetic algorithm then to pass to local research with the algorithm of Levenberg-Marquardt in order to refine the result.

This article is organized as follows: first, the description of the test and the expressions of the currents induced at the moment of an abrupt three phased short circuit are detailed in the section 2. Then, a description of the algorithms, genetic algorithm and algorithm of Levenberg-Marquardt used for optimization are made in the section 3. The last section is devoted to the presentation of the results of identification and the comparison of these results between those obtained by using the method of estimate graphic. The article will be finished by a conclusion and prospect.

2. TEST IN THREE-PHASE SHORT-CIRCUIT SYMETRIQUE OF A SYNCHRONOUS MACHINE

2.1. Description of the experimental test
This test consists in applying a short circuit to the terminals of the stator phases from the no-load march of the machine and recording the evolution of the currents of the machine. It is often used to determine the dynamic parameters of the machine which are the time-constants and the transitory and sub transitory reactance. The test is applied on a standard synchronous

machine LORENZO DL 1026 with projecting poles without shock absorber winding whose characteristics given by the manufacturer are shown by the following table:

Table 1: Characteristic of the machine

Power	1.1kVA
Nominal voltage	380V
Nominal current	1.67A
Factor of power	0.8AR
Frequency	50Hz
Number of poles	4

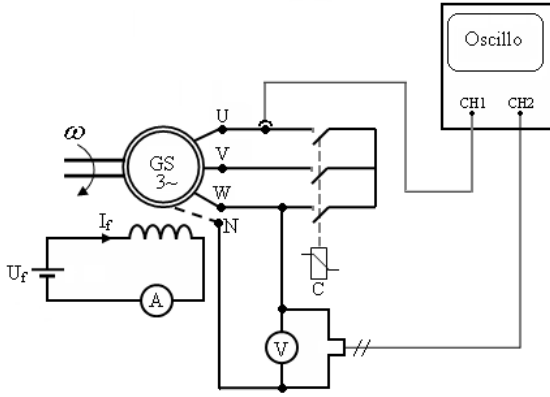The following figure shows the assembly diagram:

Figure 1: Assembly diagram of the test in abrupt three-phased short circuit

2.2. Expression of the current of short-circuit

The expression of the current in a phase at an abrupt three-phased short circuit on the stator without winding shock absorbers according to a no-load march is given (Amal 1979; Chatelain 1983):

$$i_a(t) = -V_m \left(\frac{1}{x_d} \right) \cos(\omega_0 t + \varphi_0)$$

$$-V_m \left(\frac{1}{x_d'} - \frac{1}{x_d} \right) e^{\frac{-t}{T_d'}} \cos(\omega_0 t + \varphi_0)$$

$$-V_m \left(\frac{1}{x_d''} - \frac{1}{x_d'} \right) e^{\frac{-t}{T_d''}} \cos(\omega_0 t + \varphi_0) \quad (1)$$

$$+\frac{1}{2} V_m \left(\frac{1}{x_d''} + \frac{1}{x_q''} \right) e^{\frac{-t}{T_a}} \cos(\varphi_0)$$

$$+\frac{1}{2} V_m \left(\frac{1}{x_d''} - \frac{1}{x_q''} \right) e^{\frac{-t}{T_a}} \cos(2\omega_0 t + \varphi_0)$$

With,

x_d, x_d' and x_d'' : Reactances synchronous, transitory and sub transitory of the direct axis D.

x_q'' : The sub transitory reactance of the transverse axis Q.

T_d' , T_d'' : Transitory and sub transitory time-constants corresponding to the direct axis D.

T_a : Time constant of the armature.

V_m : Maximum value of the voltage of armature before short circuit.

The current in the phases B and C is obtained while replacing respectively φ_0 by ($\varphi_0 - 2\pi/3$) and ($\varphi_0 - 4\pi/3$).

The examination of the expression of the current shows that it is the sum of five terms which are:

- A permanent sinusoidal term with pulsation ω_0 whose amplitude is lowest than transient and sub transitory: $x_d'' < x_d' < x_d$
- A deadened sinusoidal term (transitory) with pulsation ω_0 whose damping coefficient is $\frac{1}{T_d'}$, it is the transitory component of the current : $T_a < T_d'$.
- A deadened sinusoidal term (sub transitory), with pulsation ω_0 whose damping coefficient is $\frac{1}{T_d''}$; this term decreases relatively quickly.
- A deadened aperiodic term whose damping coefficient is $\frac{1}{T_a}$, this damping is relatively fast $T_a < T_d'$.
- A deadened sinusoidal term with pulsation $2\omega_0$; its amplitude is the lowest $x_d'' < ou \approx x_q''$.

3. ALGORITHM OF OPTIMIZATION

As already mentioned in the introduction, the problem of identification becomes a problem of optimization through the minimization of the quadratic errors between the measured current and the current calculated by the expression (1). One leads then to the identification of the vector of parameters:

$$\theta^T = [T_a \ T_d' \ T_d'' \ x_d \ x_d' \ x_d'' \ x_q'' \ \omega_0 \ \varphi_0] \quad (2)$$

The value of V_m is determined by the initial conditions.

3.1. The Genetic algorithm

The structure of this algorithm (Ariba 2008) is given by the flow chart shown on Figure 2.
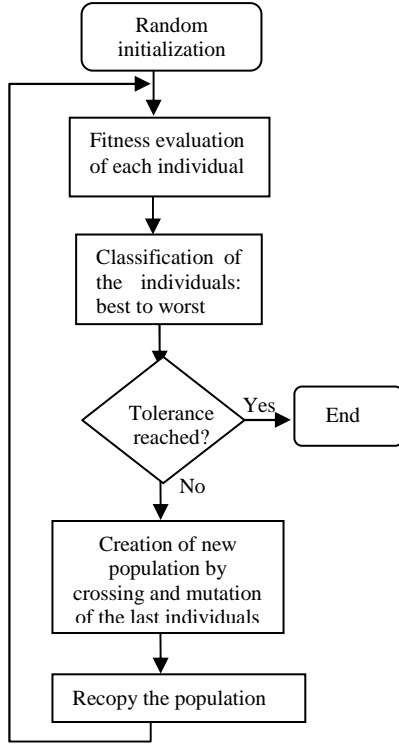
Figure 2: Structure of the genetic algorithm

This algorithm begins with the creation of the initial population individuals generated by random way and ends by converging to the best individual of the population corresponding to the solution of optimization. The passage of a generation through another is made by applying the mechanisms of evaluation, selection and modification until obtaining a stop criterion.

Each individual, called here chromosome, of the population is defined by a gene chain which corresponds to the various parameters to identify (see Figure 3).

Figure 3: Representation of current chromosome.

To avoid the difficulties which can be raised on the binary coding and decoding of chromosome, the AG with real coding is used (Bontemps 1995). The values of the parameters are limited in an interval $[\theta_{min}, \theta_{max}]$.

The principal reason to establish such limits is to make the process of research more effective by reducing its space.

The initial population is selected in a random way. That makes possible to begin research starting from various solutions of the space of research. The function of fitness defined by the equation (3) allotted to the algorithm is the sum of the quadratic errors between the digitized values of the measured current of phase and those calculated by the equation (1) by using the parameters of the evaluated chromosome.

$$Fit = \sum_{k=1}^{N} \left(I_m(k) - I_c(k) \right)^2 \qquad (3)$$

The stochastic operators adopted by the algorithm are:
1. The selection by tournament which consists in taking randomly a sample of N individuals (2 at least) to each tournament. Then, the best of this sample is selected to be relative.
2. The crossing of the type $BLX - \alpha$. This type of crossing is applied parameter by parameter by using the following mechanism: x_i' is selected randomly in the interval

$$[\theta_{min} + I\alpha, \theta_{max} - I\alpha]$$

With:

$$\theta_{min} = \min(x_i, y_i)$$
$$\theta_{max} = \max(x_i, y_i)$$,
$$I = \theta_{max} - \theta_{min}$$

and α is a random number with $\alpha \in [0,1]$.

x_i and y_i are of the same parameters row of two parents, x_i' the parameter of corresponding row of a child obtained.

This type of crossing makes it possible to create diversity in the population and to move away from the risk of uniformity of the produced chromosomes (Bahloul and Ouali 2009). The rate of crossing is 90 %.
3. A uniform change applied with a probability of 1 %. It acts to modify a parameter by choosing a value in a way random defined by the constraints of the field.

3.2. Algorithm of Levenberg-Marquardt

The function cost defined by the equation (4) used is in the algorithm is the sum of the quadratic errors between the digitized values of the measured current of phase and those calculated by the equation (1) by using the parameters defined in the equation (2).

$$J(\theta) = \sum_{k=1}^{N} \left(I_m^k - I_c^k(\theta) \right)^2 \qquad (4)$$

This algorithm works by starting from the best parameters found by the genetic algorithm and ends in convergence towards the best parameters corresponding to the solution of the problem optimization. The formula updating the parameters is given by:

$$\theta^{k+1} = \theta^k - \left(H(\theta^k) + \mu_{k+1} I \right)^{-1} \nabla J(\theta^k) \qquad (5)$$

Where $H(\theta^k)$ is the Hessian matrix of the function cost J, I is the matrix identity, and where μ_{k+1} is a scalar called step. For small step values μ_{k+1}, this

method approaches the Newton one (Dennis and More 1977) while for great step values, the method tends towards the gradient simple one (More 1977). By judiciously choosing the step value during the algorithm, it is then possible to avoid preliminary implementation like method of simple gradient to approach the minimum.

The calculation of the reverse of the matrix $[H(\theta^k) + \mu_{k+1}I]$ can be done by methods of direct inversion. Nevertheless, taking into account the function cost considered in (4), it is preferable to implement a method of iterative inversion, founded by the following property: being given four matrixes A, B, C and D.

$$(A + BCD)^{-1} = A^{-1} - A^{-1}B(C^{-1} + DA^{-1}B)^{-1}DA^{-1} \quad (6)$$

The expression of the Hessian matrix is as followed:

$$H(\theta^k) = \sum_{n=1}^{N} \left(\frac{\partial e_n}{\partial \theta^k}\right)\left(\frac{\partial e_n}{\partial \theta^k}\right)^T + \sum_{n=1}^{N} \frac{\partial^2 e_n}{\partial \theta^k (\partial \theta^k)^T} e_n \quad (7)$$

Where $e_n = y_n^p - y_n$ is the prediction error.

Neglecting the second term in the relation (7), which is proportional to the error, the following approximation of the Hessian matrix is obtained:

$$H(\theta^k) = \tilde{H}(\theta^k) = \sum_{n=1}^{N} \left(\frac{\partial e_n}{\partial \theta^k}\right)\left(\frac{\partial e_n}{\partial \theta^k}\right)^T \quad (8)$$

This approximate Hessian matrix obeys to the formula of following recurrence:

$$\tilde{H}^n = \tilde{H}^{n-1} + X^n (X^n)^T$$
$$\text{with} \quad X^n = \frac{\partial e_n}{\partial \theta^k} = \frac{\partial y_n}{\partial \theta^k}, \quad n = 1,...,N \quad (9)$$

By fixing like initial value $\tilde{H}^0 = \mu_k I$, the following expression is obtained: $\tilde{H}^N = \tilde{H} + \mu_k I$

Using the lemma of inversion stated previously with $A = \tilde{H}^{n-1}$, $B = X^n$, $C = I$, $D = (X^n)^T$, it is possible to write:

$$(\tilde{H}^n)^{-1} = (\tilde{H}^{n-1})^{-1} - \frac{(\tilde{H}^{n-1})^{-1} X^n (X^n)^T (\tilde{H}^{n-1})^{-1}}{I + (X^n)^T (\tilde{H}^{n-1})^{-1} X^n} \quad (10)$$

It is then possible to calculate the reverse of the matrix repeatedly $\tilde{H}^N = \tilde{H} + \mu_k I$

Let us note that this method of approximate calculation of the reverse of the matrix rises from the approximation (9), which is valid only for low error values of prediction e_n, and thus for the values of θ near the optimal value. The field of validity of the Newtonian approximation, extended a priori by the addition of the term $\mu_k I$ in the formula (7) is finally restricted to be able to effectively calculate the reverse of this increased Hessian matrix.

4. RESULTS AND DISCUSSIONS

4.1. Experimentation

The alternator is actuated at its nominal speed by an engine with D.C. current of 3kW, under a tension of excitation reduced not to run up against property damages due to electromechanical constraints.

This also makes it possible to avoid erroneous results bus in the zone of saturation the parameters change values.

The maximum tension at the boundaries of the winding before the short circuit is of $E_m = 312\sqrt{2}$ [V]. The following curve shows the recorded current:
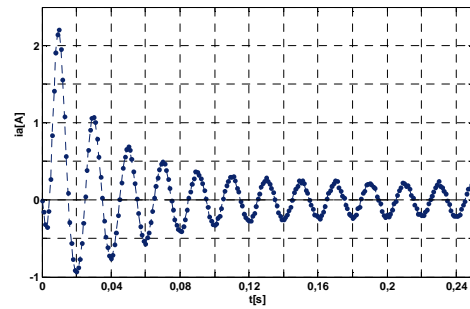
Figure 4: Current of three-phase short-circuit

The oscilloscope used has an interface making possible to record in a PC the discrete values of the current. These discrete values are then used to feed the program of optimization to identify the parameters of the machine. Figure 5 (resp. Figure 6) shows the evolution of the average quadratic error between the estimated current and the current measured during cycle of optimization by AG (resp. Hybrid Algorithm).
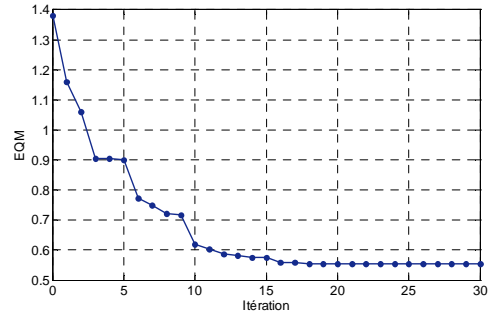
Figure 5: Evolution of the global error (AG)
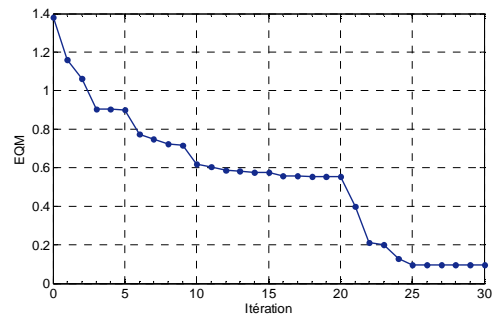
Figure 6: Evolution of global error (Hybrid)

It is noted that convergence is ensured in both cases. The use of AG alone gives a global error approximately 0.5 at the end of 20 iterations (Figure 5). This error is still reduced by continuing optimization with the algorithm of LM. One obtains a final error of 0.1 in 30 iterations (Figure 6). These results justify the adopted approach.

Figure 7: Estimation of the current phase $i_a(t)$ by AG

Figure 8: Estimation of the current phase $i_a(t)$ by hybrid algorithm.

Figure 7 (resp. Figure 8) shows the estimation between the measured current and the current estimated by AG (resp. Hybrid Algorithm). It is noted that these two methods offer a good estimate as well in transient state as in steady operation. The superiority of the hybrid algorithm is shown in comparison of the global errors presented in Figure 5 and Figure 6.

4.3. Comparison
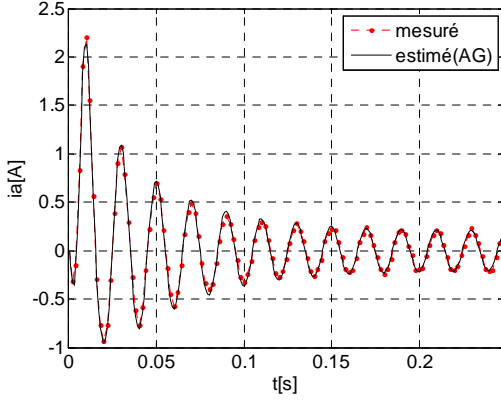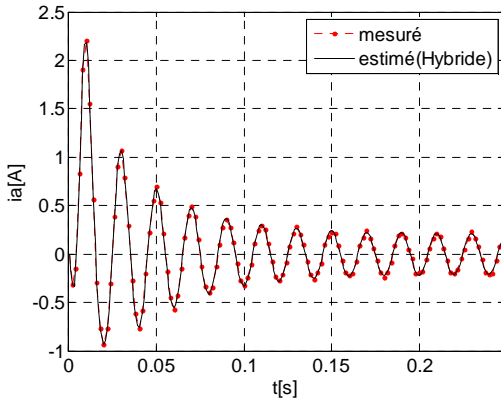To evaluate the parameters identified by the method suggested a comparison with those estimated by AG and detailed graphic method (IEEE Guide 1995; Rakotoarinoro 1999) is presented in the following table:

Table 2: Comparison of the parameters

	AG	Hybrid	Graphic
$T_a[ms]$	10.15	10.15	10
$T_d^{'}[ms]$	44.21	38.72	38
$T_d^{''}[ms]$	38.92	35.16	38
$x_d[\Omega]$	218.45	209.87	210
$x_d^{'}[\Omega]$	24.88	26.31	23.3
$x_d^{''}[\Omega]$	24.59	23.40	23.3
$x_q^{''}[\Omega]$	120.9	102.03	98
$\omega_0[rad.s^{-1}]$	314.46	314.18	314.16
$\varphi_0[rad]$	0.0001	0.0001	0

For a synchronous machine without shock absorbers, the sub transitory and transient parameters (reactances and time-constants) are supposed to be identical during a graphic estimate. However, in reality, there is a light difference. This difference is shown well by the adopted approach. It is important to notice that the vector of parameters which is obtained with the suggested approach is only one average vector of parameters i.e. it is the vector which corresponds to the best compromise to represent the machine in its operation.

4.3 Experimental validation
To check the parameters identified by the hybrid algorithm proposed, it is useful to proceed a phase of experimental validation. For that, one carried out under test known as "of tension of reappearance" (Lessenne, Notelet and Seguier 1981). The alternator is initially in three-phase short-circuit, one opens the short-circuit by the intermediary of the three poles switch C (Figure 1) and one records the curve of re-establishment of current.

The simplified expression of this tension, for a synchronous machine without shock absorbers is (Rakotoarinoro 1999) :

$$V_a = V_m \left(1 - \left(1 - \frac{x_d^{'}}{x_d} \right) e^{\frac{-t}{T_{d0}^{'}}} \right) \sin(\omega t + \varphi) \qquad (11)$$

With $T_{d0}^{'} = T_d^{'} \dfrac{x_d}{x_d^{'}}$

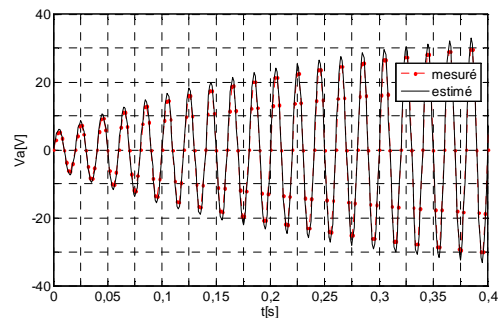The following curve shows the evolution of this tension of reappearance:

Figure 9: Voltage waveform of reappearance

According to Figure 9, the tension of reappearance curve simulated by using the parameters identified with hybrid algorithm follows well the experimental one. That increases the certainty with the suggested method.

5. CONCLUSION

A method of optimization hybrid combining the genetic algorithm (AG) and the algorithm of Levenberg-Marquardt (LM) was implemented in this article.

The objective was to minimize the quadratic errors between measured and estimated of short circuited currents.

In their work, Bahloul et al. (Bahloul 2009) proposed the identification of the dynamic parameters of a synchronous machine by using only the genetic algorithm. They arrived at notable results. However, the combination of this algorithm with the Levenberg-Marquardt one led to finer results on by the way of the global errors.

The approach of hybridization adopted in this work makes possible to find the parameters dynamic of a machine synchronous using the test in short-circuit, independently of the initial parameters. The space of research of the parameters is given to help the algorithm to find quickly the optimum.

The approach suggested is tempting and offers a very interesting alternative compared to the traditional methods, which require "to start" near the optimal vector and whose behaviour is dubious in front of the local minima.

The comparison between the measured current and the current estimated by simulation makes it possible to validate the method. Indeed, there is agreement of the results as well in transient state as in steady operation showing the precision of the given parameters. This proves the performance and the robustness of the adopted method of optimization.

This work gives multiple possibilities of research that it is interesting to dig. First of all with regard to the genetic algorithm, it would be interesting to compare our approach with that where the change of individuals is fixed not by probability but by sizes of under population. Then concerning the choice of the moment when the algorithm of Levenberg-Marquardt takes over in its turn, it would be interesting to exploit the evolution of each parameter.

Finally, it was said that the suggested method combines the genetic algorithm and the Levenberg-Marquardt one. Here, it is noted that the number of iterations Ni is approximately equal to two for three times the number of parameters Np. The identification is started with AG and after 18 iterations; the commutation to the LM algorithm is made. It would be useful to find a relation between Ni and Np.

REFERENCES

IEEE Guide, 1995. *Test Procedures for Synchronous Machines*, IEEE Std 115

Amail, P., 1979, Régime transitoire et stabilité des machines à courant alternatif, *EDF Division technique générale.*

Chatelain, J., 1983, Machines Electriques. *Traité d'Electricité, Vol. X.* ed. Georgi

Ben Ariba, H., 2008, *Apports à la résolution de problèmes pour le dispatching des réseaux d'énergie électrique*, Thesis (Doctorat). ENIS University.

Bontemps, C., 1995, *Principes Mathématiques et Utilisations des Algorithmes Génétiques.* Extract of course.

Bahloul W., Ouali M., 2009, Identification des paramètres dynamiques de la machine synchrone par algorithme génétique. *IREC'09* (Sousse, Tunisia)

Dennis, J.E., More, J.J. Holland, M., 1977. Quasi-Newton methods, motivation and theory. *SIAM Review, Vol. 19, No. 1,* pp. 46-89.

More, J.J., 1977. The Levenberg-Marquardt Algorithm: Implementation and Theory. *Lecture Notes in Mathematics,* ed. Springer Verlag. pp.105-116.

Rakotoarinoro, R.C., 1999. *Contribution à l'étude du comportement des machines synchrones de réseau en régimes transitoires.* Thesis (DEA) ESP-University of Diego Suarez.

Lessenne, J., Notelet, F., Seguier, G., 1981. Introduction à l'électrotechnique approfondie. *Technique de documentation.* Paris.

AUTHORS BIOGRAPHY

Razafinjaka N. Jean. Born in 1956. Doctor Engineer. He works now in the Laboratory of Automatic, Polytechnic Higher School of Diego Suarez, Madagascar. His field of research relate to the advanced commands applied to electric and electromechanical systems.

Rakotoarisoa A. Jean Claude. Born in 1976. Holder of DEA. He works at the laboratory of Automatic and currently prepares his doctorate in collaboration with the University of Paris 8. His theme consists of research on the identification of a mechanical micro system using different algorithms and their combination.

STUDY OF COUPLED DYNAMICS BETWEEN BODY AND LEGS OF A FOUR LEGGED WALKING ROBOT

V. L. Krishnan[a], P. M. Pathak[b], S. C. Jain[c]

Robotics and Control Laboratory
Mechanical and Industrial Engineering Department
Indian Institute of Technology, Roorkee, 247667, India

(a) vlk08dme@iitr.ernet.in, [b] pushpfme@iitr.ernet.in, [c] sjainfme@iitr.ernet.in

ABSTRACT

The paper presents a study of coupled dynamics between body and legs of a four legged robot with two articulate joints per leg. This study brings out the influence produced by the ground reaction forces through robot legs on the posture of robot body i.e. body rotations about X, Y and Z axis. An object oriented approach has been used for the bond graph modeling of locomotion dynamics of the four legged robot while taking into consideration the robot-ground interaction forces. Detailed kinematics analysis of a single leg has been carried out. Sub-model created for a single leg is repeatedly used for developing the bond graph model of the four legged walking robot. A multi bond graph is used to represent the system. A dynamical gait is proposed and implemented through joint control. Joint control has been achieved using a proportional derivative control law for each joint of the four legged robot. The robot locomotion has been successfully demonstrated through simulation and experiments on a robot prototype.

Keywords: Four legged walking robot, coupled dynamics between body and legs, bond graph modeling

1. INTRODUCTION

Legged robots have evolved as a better alternative compared to their wheeled counterparts for field applications such as military combat or transport operations, material handling and rescue operations, hazardous site inspection; and for extraterrestrial applications viz. Mars or space exploration etc. Legged robots are more suitable for such applications as they have greater agility and also they can move well on all kind of terrain whereas wheeled robots require only paved paths for better performance. Among the legged robots, six or more legged robots are suitable from the locomotion stability perspective. However, four legged robots offer a good compromise between locomotion stability and speed.

Research in rigid legged robots started almost in early 1980's with an attempt to realize rigid legged locomotion mainly on flat terrain. Waldron and McGhee (1986) presented the design of Adaptive Suspension Vehicle (ASV), which had morphology analogous to a six legged insect. Hartikainen et al. (1992) developed MECANT-I, a hexapod walking machine for forests applications. These robots mainly used statically stable gaits for their locomotion. Zhang and Song (1993) presented a study of the stability of generalized wave gaits. Gonzalez de Santos and Jimenez (1995) introduced discontinuous gaits and Estremera and Gonzalez de Santos (2002) proposed free gaits for locomotion of quadruped robots on irregular terrain. Later on researchers got motivated to explore dynamically stable gaits for realizing faster locomotion. Furusho et al. (1995) realized bounce gaits on SCAMPER, a rigid legged quadruped with its design similar to mammals. Estremera and Waldron (2006) proposed a leg thrust control method for the stabilization of dynamic gaits in rigid legged quadruped robot KOLT. Garcia et al. (2003) suggested that in addition to developing suitable gaits for legged locomotion, consideration of actuator dynamics and friction is essential for getting the real legged locomotion behaviour. Bowling (2005) examined the robot's ability to use ground contact to accelerate its body. Yoneda and Hirose (1992) employed biologically inspired approach to realize smooth transition from static to dynamic gait in the quadruped walking robot 'TITAN-IV'. Biological inspired approach refers to an extensive use of sensory feedback and reflex mechanisms (similar to that found in animals) for locomotion control. Further Kurazume et al. (2001) accomplished dynamic trot gait control for 'TITAN-VIII'. Inagaki et al. (2006) proposed a method for the gait generation and walking speed control of an autonomous decentralized multi-legged robot by using a wave Central Pattern Generator (CPG) model. Wyfells et al. (2010) has presented a design and realization of quadruped robot locomotion using Central pattern generators.

Legged robot is a multi-body dynamic system. Realization of various locomotion behaviors viz. walking, running etc. requires a precise understanding of the coupled dynamics between the body and legs of a legged robot. Also influence of ground reaction forces on the robot body is an important matter of

investigation because it is quite intuitive that instantaneous leg tip velocity of a legged robot depends upon the body state variables. Hence with the objective to investigate the coupled dynamics between the body and legs of a four legged walking robot, the present work has been carried out. This paper presents a three
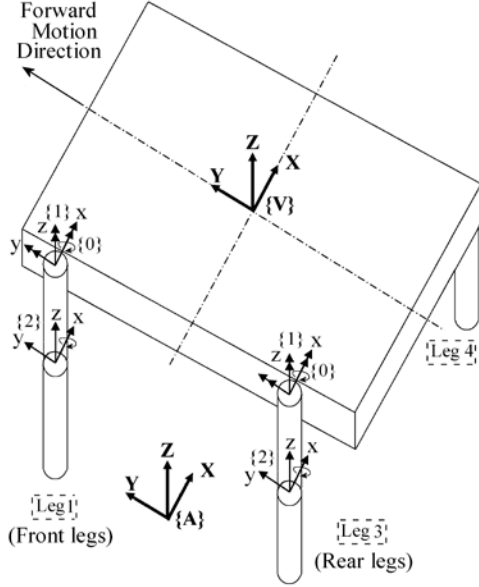

Figure 1: Schematic diagram of four legged robot

dimensional study of coupled dynamics between body and legs and generation of stable walking in a four legged robot. The coupled dynamics and robot locomotion has been demonstrated through simulation results and experiments on a robot prototype.

Bond graph technique (2006) has been used for object oriented modelling of the four legged robot.

Bond graph is an explicit graphical tool for capturing the common energy structure of systems. It gives power exchange portray of a system. It provides a tool not only for the formulation of system equations, but also for intuition based discussion of system behavior viz. controllability, observability, fault diagnosis, etc. The language of bond graphs aspires to express general class physical systems through power interactions. The factors of power i.e., effort and flow, have different interpretations in different physical domains. Yet power can always be used as a generalized element to model coupled systems residing in different energy domains. In order to avoid repetitive modeling of same type of structure and to express a very large system in modular form, objects are created and then joined together to create an integrated system model. In the present work bond graph model of the four legged robot is created and simulated in SYMBOLS Shakti (2006), a bond graph modelling software.

2. BOND GRAPH MODELLING OF FOUR LEGGED ROBOT

Modelling of four legged robot consists of modelling of translational and angular dynamics of robot legs and body. Figure 1 shows the schematic diagram of a four legged robot model. In Fig. 1, {A} is inertial frame of reference and {V} is body frame. Each leg of the robot has two links. The joint between links i and $i+1$ is numbered as $i+1$. A coordinate frame $\{i+1\}$ is attached to $(i+1)$ joint.

The linear dynamics of a body is governed by Euler's first law and angular dynamics by Euler's second law. Linear dynamics of a body can be given by

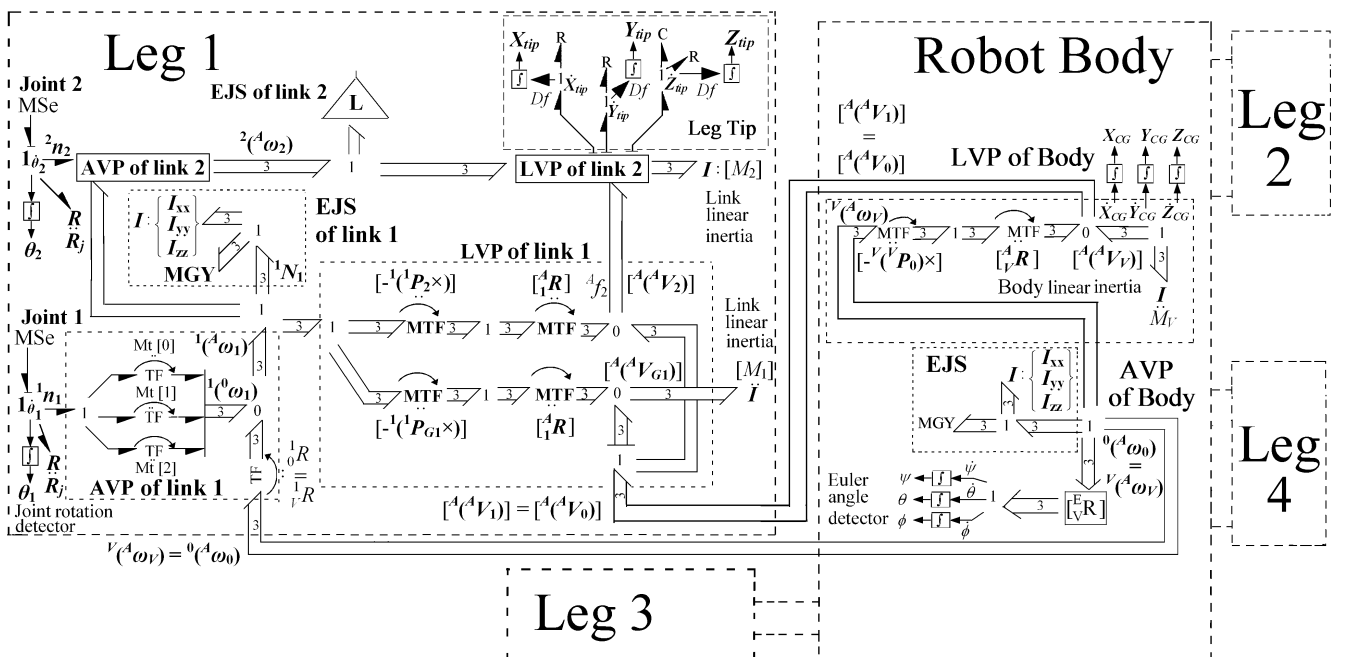$$^{A}F_{V} = M_{V} \; ^{A}(\; ^{A}\dot{V}_{G_{V}}) \qquad (1)$$

Figure 2: Multi bond graph of four legged robot

Where, $^{A}F_{V}$ is resultant of forces at joints ends and external forces acting at center of gravity (CG), expressed with respect to frame $\{A\}$; $^{A}(\overset{\bullet}{\overset{A}{V}}_{G_{V}})$ is acceleration of CG of body with respect to frame $\{A\}$; M_{V} is the mass of the body.

Angular dynamics of the body can be given by

$$^{V}N_{V} = {}^{V}(\overset{A}{\overset{\bullet}{h}}_{V}) + {}^{V}(^{A}\omega_{V}) \times {}^{V}(^{A}h_{V}) \qquad (2)$$

Where, $^{V}N_{V}$ is the moment acting on the body; $^{V}(^{A}h_{V})$ and $^{V}(^{A}\omega_{V})$ are respectively the angular momentum and angular velocity of body with reference to inertial reference frame $\{A\}$.

These fundamental equations of motion (1) and (2) along with linear velocity and angular velocity propagation relations for leg links (presented later in this section) guides the bond graph modeling of a four legged robot. Figure 2 shows the multi bond graph model of a four legged walking robot. The multi bond graph consists of sub model for the robot body and the four legs. Figure 3(a) and (b) respectively presents the bond graph of the 'LEG' and the 'Joint actuator' sub-models.
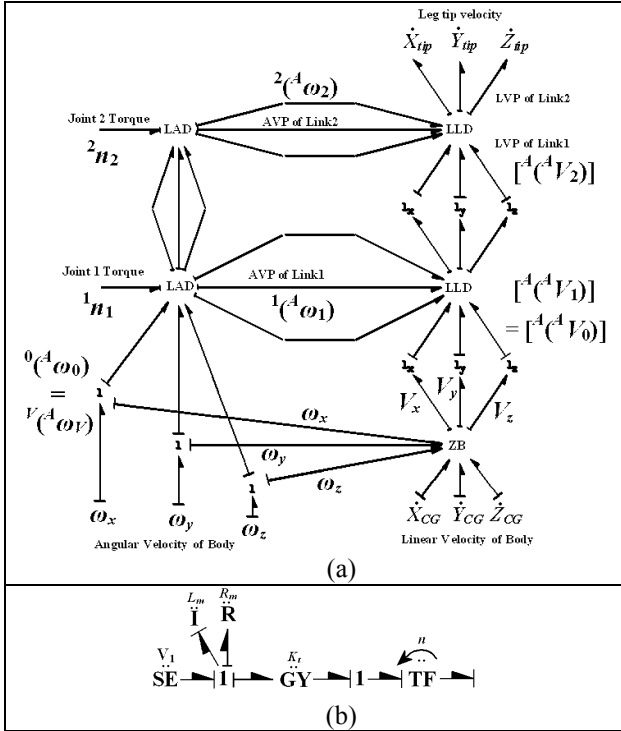

(a)

(b)

Figure 3(a) LEG sub-model (b) Joint actuator sub-model

Robot body sub model represents its translation and angular dynamics. The velocity of the body CG frame $\{V\}$ is obtained from the linear inertia of the body. Euler junction structure (EJS) can be used to represent the angular dynamics of a body (Mukherjee et al. 2006).

Hence, it has been used to represent the rotational dynamics of the robot body as well as that of the links of robot legs in the present work.

The Euler equations used in the creation of the EJS sub-model of body are deduced from Eq. (2) and are given by Eq. (3) as

$$N_{x} = I_{x}\dot{\omega}_{x} + (I_{z} - I_{y})\omega_{y}\omega_{z}, \qquad (3a)$$

$$N_{y} = I_{y}\dot{\omega}_{y} + (I_{x} - I_{z})\omega_{z}\omega_{x}, \qquad (3b)$$

$$N_{z} = I_{z}\dot{\omega}_{z} + (I_{y} - I_{x})\omega_{x}\omega_{y}. \qquad (3c)$$

Where N_{x}, N_{y}, N_{z} are the torques and $\dot{\omega}_{x}$, $\dot{\omega}_{y}$ and $\dot{\omega}_{z}$ are angular velocities acting about the principal axes of the corresponding body fixed frame. Linear velocity of the $\{0\}$ frame of each leg is given as

$$^{A}(^{A}V_{0}) = {}^{A}(^{A}V_{V}) + {}_{V}^{A}R[-{}^{V}(^{V}P_{0}) \times {}^{V}(^{A}\omega_{V})] \qquad (4)$$
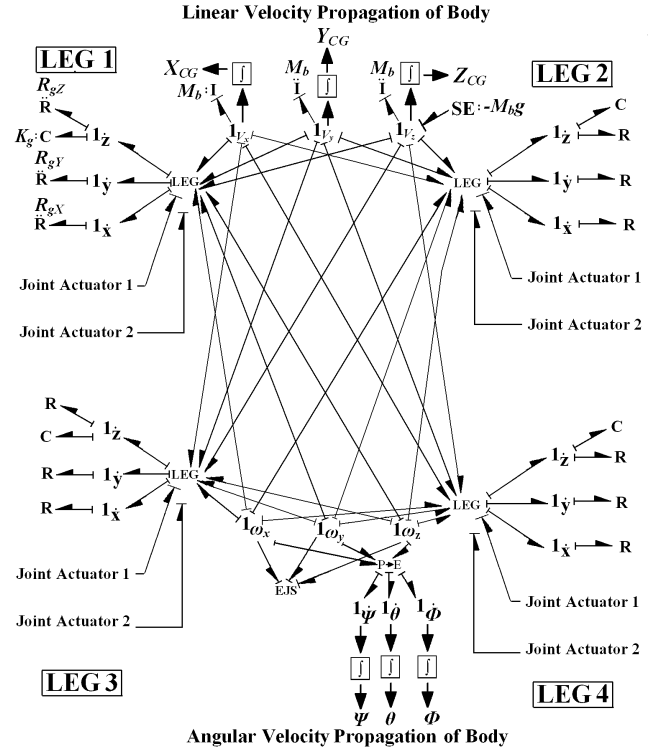
Figure 4: Bond graph of four legged robot locomotion dynamics

In Eq. 4, $_{V}^{A}R$ represent the transformation from body frame $\{V\}$ to inertial frame $\{A\}$ and can be expressed as,

$$_{V}^{A}R = \begin{pmatrix} c\theta\,c\phi & s\psi\,s\theta\,c\phi - c\psi\,s\phi & c\psi\,s\theta\,c\phi + s\psi\,s\phi \\ c\theta\,s\phi & s\psi\,s\theta\,s\phi + c\psi\,c\phi & c\psi\,s\theta\,s\phi - s\psi\,c\phi \\ -s\theta & s\psi\,c\theta & c\psi\,c\theta \end{pmatrix}$$

ψ, θ and ϕ are the Euler angles representing robot body rotation about X, Y, Z axis of the body fixed frame $\{V\}$. $^{V}(^{V}P_{0})_{i}$ represent the position vector of frame $\{0\}$ of ith

leg with respect to body CG frame $\{V\}$. It can be expressed as $^V(^VP_0)_i =[R_{ix} \quad R_{iy} \quad R_{iz}]^T$, where '$i$' denotes leg 1 to 4. Value of R_{ix}, R_{iy} and R_{iz} corresponding to leg 1 to 4 is listed in Table 1 in appendix.

Linear Velocity Propagation (LVP) sub model shown in the 'Body' part of the multi bond graph in Fig. 2 takes the angular velocity from body $^V(^A\omega_V)$ (obtained from EJS) and linear velocity $^A(^AV_V)$ (decided by body mass) as input and gives out the velocity of $\{0\}$ frame to the link 1 of each leg. Frame $\{0\}$ and $\{1\}$ are coincident for each leg. Hence, the velocity of frame $\{1\}$ is same as frame $\{0\}$ i.e. $^A(^AV_1) = {}^A(^AV_0)$.

'Leg' sub model, shown in detail in Fig. 3(a), represents a two DOF leg. 'Leg' takes angular and linear velocity of body and joint torques about X-axis as input. 'LEG' sub model uses Angular Velocity Propagation (AVP) and LVP sub models of links 1 and 2 and gives out leg tip velocity as output. Link lengths l_1 and l_2 are taken along the principal Y-axis of the links and hence represented in vector form as,
$^0P_1 = [0 \quad 0 \quad 0]^T$, $^1P_2 = [0 \quad l_1 \quad 0]^T$, $^2P_3 = [0 \quad l_2 \quad 0]^T$.
Thus 'LEG' furnishes complete dynamics of a two link leg. The various sub models shown in Fig. 2 for leg '1' can also be used to model leg 2, 3 and 4.

LVP and AVP sub models for leg links can be used to find the velocity of tip of link1 and link 2 of a leg i.e. frame $\{2\}$ and frame $\{3\}$ respectively. Governing equation for AVP of links of a leg can be given as per theory (Craig 2006),

$$^{i+1}(^A\omega_{i+1}) = {}^{i+1}_iR \; ^i(^A\omega_i) + {}^{i+1}(^i\omega_{i+1}) \qquad (5)$$

Where, $^{i+1}(^i\omega_{i+1})$ is the angular velocity of $(i+1)$ link as observed from i^{th} link and expressed in $(i+1)^{th}$ frame. The term can be expressed for link 1 and 2 respectively as,
$^1(^0\omega_1) = [\dot{\theta}_1 \quad 0 \quad 0]^T$, $^2(^1\omega_2) = [\dot{\theta}_2 \quad 0 \quad 0]^T$.
$^i(^A\omega_i)$ is the angular velocity of the ith link with respect to inertial frame $\{A\}$ and expressed in i^{th} frame. $^{i+1}(^A\omega_{i+1})$ is the angular velocity of $(i+1)$ link with respect to inertial frame and expressed in $(i+1)^{th}$ frame. The above equation (5) is represented by the sub model 'AVP of Link' in each leg of the robot.

Governing equation for the link tip velocity and link CG velocity are given as,

$$^A(^AV_{i+1}) = {}^A(^AV_i) + {}^A_iR[\,^i(^A\omega_i) \times {}^i(^iP_{i+1})] \qquad (6)$$

This can be simplified as,

$$[^A(^AV_{i+1})] = [^A(^AV_i)] + [^A_iR][-(^iP_{i+1}\times)][^i(^A\omega_i)] \qquad (7)$$

For position of a link CG, $^i(^iP_{G_i}) = [0 \; l_{Gi} \; 0]^T$

$$[^A(^AV_{G_i})] = [^A(^AV_i)] + [^A_iR][-(^iP_{G_i}\times)][^i(^A\omega_i)] \qquad (8)$$

Equations (6), (7), (8) represent the LVP of Link in each leg of the robot. CG velocity of links depends on link inertia. In bond graph model 'I' elements (representing mass of a link), are attached at flow junctions. They yield the CG velocities of links. The starting point of the current link is same as the previous link tip. Hence, the tip velocity of the previous link and the angular velocity of the current link are used to find the tip velocity and CG velocity of the current link. $^i(^A\omega_i)$ in above equations can be obtained from the AVP for the current link.

The leg tip sub-model in Fig. 2 represents the modelling of leg tip-ground interaction. The robot is assumed to be walking on a hard surface with no slipping of legs. An 'R' element is appended to '1' junction of each leg in the X and Y direction, to model the frictional resistance offered by ground. Similarly, 'C' and 'R' elements are attached in Z-direction to model the normal reaction force from the ground. Leg tip position detectors in each direction yields the leg tip position coordinates.

The integrated bond graph model representing the four legged robot locomotion dynamics is presented in Fig. 4. Parametric values assumed for the purpose of simulation of the four legged robot model are shown in Table 2 in appendix.

3. GAIT PATTERN

Gait pattern represents the sequence of leg movements required for realizing locomotion of a robot while maintaining body stability. In the present work, a bounding walk gait pattern has been implemented for achieving robot locomotion. The gait has been used by Lasa and Buehler (2000) for their single-link legged quadruped robot SCOUT-II. In this gait pattern, either the front or rear legs of the quadruped robot are simultaneously lifted up or brought down to the ground in a particular phase of the gait. Figure 5(i-viii) represents schematically the eight phases of the gait pattern of a locomotion cycle.

To implement the gait pattern, position of the joints of each leg must be controllable. The voltage supplied for controlling the position of joints can be given as

$$V_i = K_P\left(\theta_{di} - \theta_i\right) + K_V\left(\dot{\theta}_{di} - \dot{\theta}_i\right) \qquad (9)$$

Where, V_i is the input voltage supplied at the ith joint of a leg. The voltage supply to a joint actuator is implemented in bond graph, through an 'SE' element of joint actuator sub-model shown in Fig. 3(b). K_P and K_V are respectively the proportional and derivative gains; θ_{di} is the desired value of rotation, θ_i is the actual value of rotation, $\dot{\theta}_{di}$ is the desired joint velocity and $\dot{\theta}_i$ is the actual joint velocity.
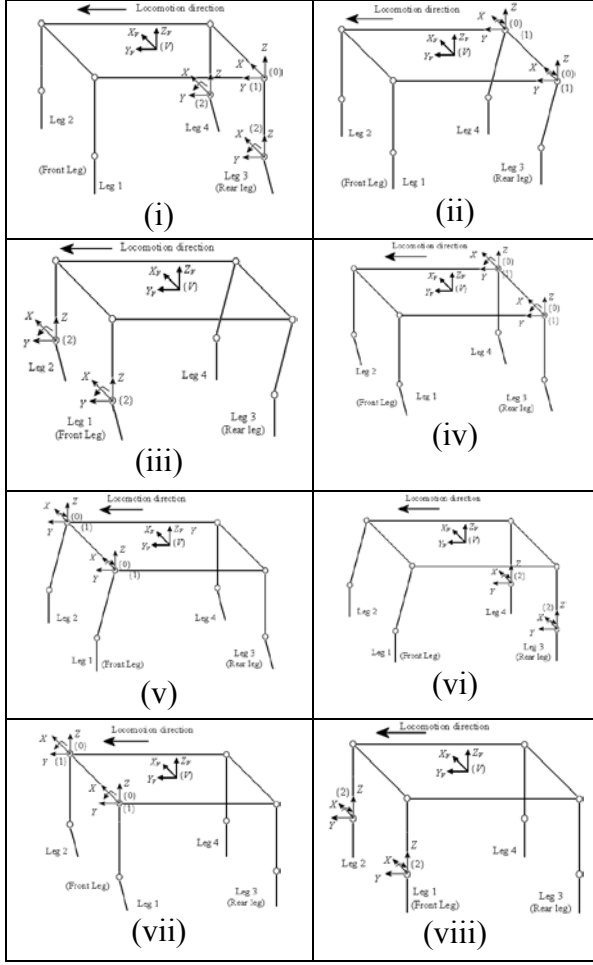
Figure 5: Phases of gait pattern

The joint reference command used for moving a particular joint to the desired position 'θ_{i+1}' at the end of a certain time interval Δt_{i+1} i.e. (t_{i+1}-t_i), can be expressed by the following equation,

$$\theta_{i+1} = \theta_i \pm k(1 - e^{-\lambda(t_{i+1}-t_i)}) \qquad (10)$$

In Eq. (14), θ_i and θ_{i+1} respectively represents the joint angular displacement values at the beginning and end of a time interval Δt_{i+1}; k is a factor by which the joint angle is to be increased or decreased, λ is an integer. It can be noted from Eq. (14) if $t = t_i$, then $\theta = \theta_i$ and when $t = t_{i+1}$ then $\theta = \theta_i \pm k$, for very large values of λ which leads the exponential term to a zero value.

4. SIMULATION RESULTS AND DISCUSSION

For the selected gait pattern and robot parameters, simulation has been carried out for 3 cycles. A locomotion cycle takes 3.2 seconds. Simulation results are presented in Fig. 6, 7, 8 and 9. Fig. 6(a) and (b) respectively shows front leg joint 1 displacement (θ_{1F}) and front leg joint 2 displacements (θ_{2F}) versus time. Similarly, rear leg joint 1 (θ_{1R}) and rear leg joint 2 (θ_{2R}) displacement versus time is presented in Fig. 7(a) and

(b). It can be noted that the leg joint angular displacement plots corresponds with the specified gait pattern.

Figure 8(a), (b) and (c) respectively presents the variation of robot body Euler angles ψ, θ and ϕ versus time. ψ, θ and ϕ represents the robot body rotation about X, Y and Z axis. As a consequence of the selected gait

(a)

(b)

Figure 6: Simulation results of rigid legged quadruped robot locomotion: Front leg joint angular displacements (a) θ_{1F} (radians) versus time (s) (b) θ_{2F} (radians) versus time (s)

pattern for locomotion (in which the two front or rear legs are simultaneously lifted or brought down simultaneously), the body angular displacement 'ψ' is oscillatory, as shown in Fig. 8(a). Variation of 'ψ' with respect to time signifies the coupled dynamics between the legs and the body. In the figure, second cycle of locomotion has been further split into eight phases (of the specified gait pattern) to explain the coupled dynamics. It can be noted that a correspondence between the various phases of the gait pattern and the body pitching motion 'ψ' exists. For instance, in the first two phases of the gait pattern (i.e. Fig. 5(i) and (ii)) the rear leg links rotation should result in body pitching about +X-axis. The reason being the rear leg hip joint will be at lesser height as compared to its front leg
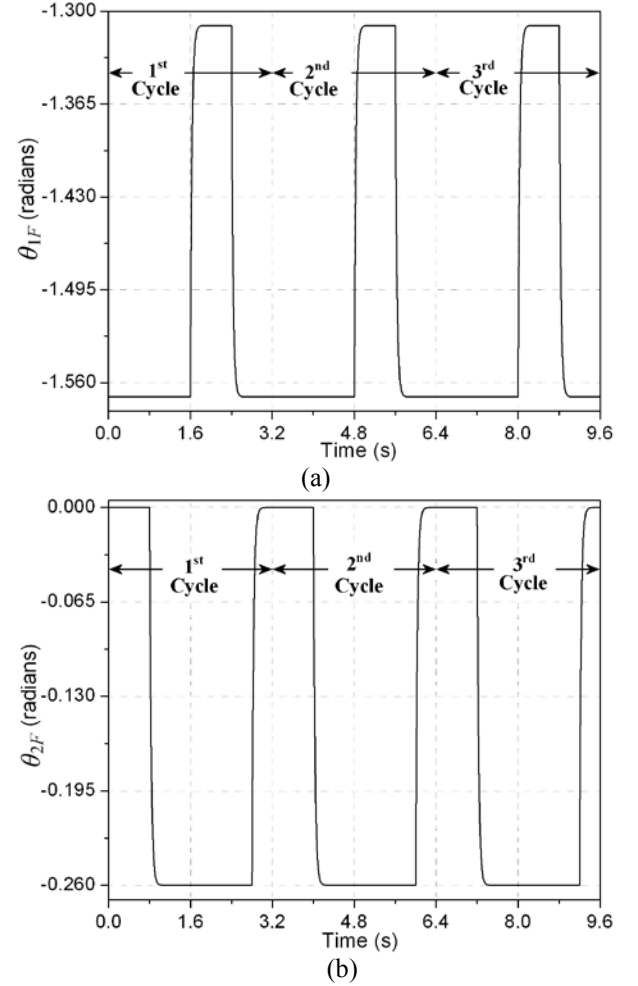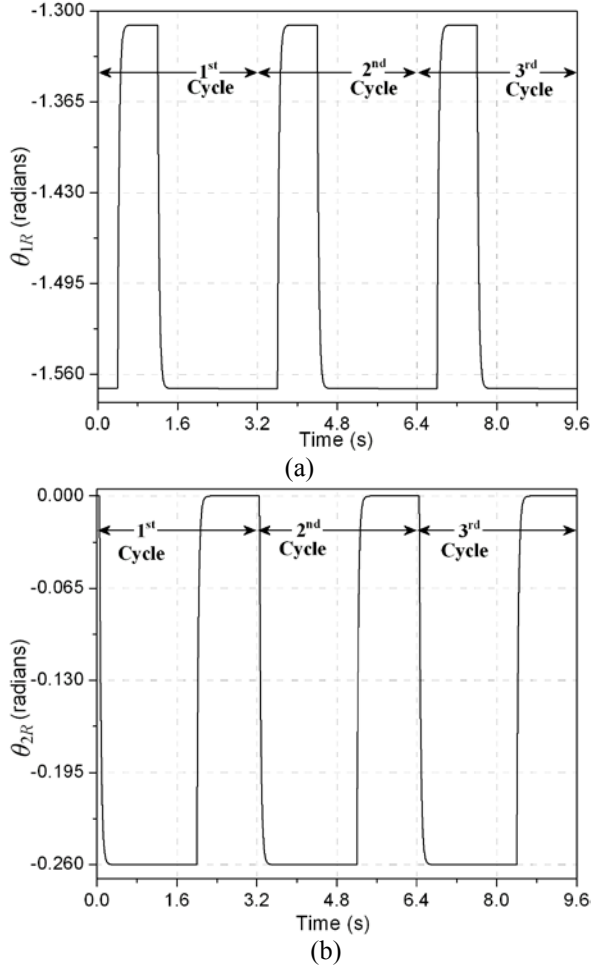
(a)

(b)

Figure 7: Simulation results of rigid legged quadruped robot locomotion: Rear leg joint angular displacements (a) θ_{1R} (radians) versus time (s) (b) θ_{2R} (radians) versus time (s)

counterpart due to the specified rotation. Simulation result in Fig. 8(a) validates this observation. Similarly corresponding to the gait pattern phase shown in Fig. 5(iv), pitching about X axis varies from positive to negative values due to the variation of the difference in the vertical height of front and rear leg hip joints. Similarly it can be observed that 'ψ' is almost equal to zero radians i.e. robot body is absolutely horizontal, at the beginning of a locomotion cycle, at the end of the fourth and eighth phase of a locomotion cycle. The reason for the fact is that the front and rear leg hip joints in the respective phases are at almost equal elevation.

Figure 8(b) indicates that there is no significant angular displacement 'θ' about Y-axis (as expected), because the joint torque is supplied only about X-axis. Figure 8(c) shows that the body is turning about +Z-axis. Figure 8(d) indicates quadruped robot progression, as since the robot body CG displacement occurs along positive Y-axis i.e. the direction of locomotion.

Figure 9(a), (b), (c) and (d) respectively shows the displacement of tip of legs 1, 2, 3 and 4, along positive Y-axis. It can be noted from the simulation results that the leg tip displacement values for leg 2 and 4 are greater than that of leg 1 and leg 3. Thus the simulation
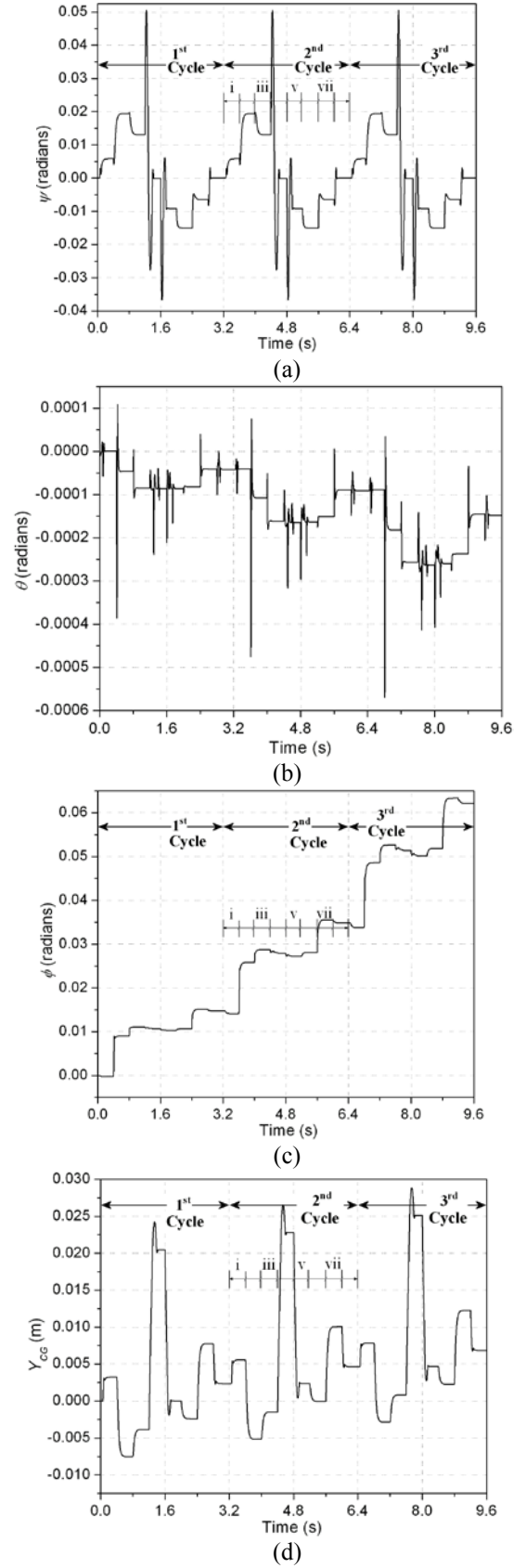

(a)

(b)

(c)

(d)

Figure 8: Simulation results of rigid legged quadruped robot locomotion (a) ψ (radians) versus time (s) (b) θ (radians) versus time (s) (c) ϕ (radians) versus time (s) (d) Y_{CG} (m) versus time (s)
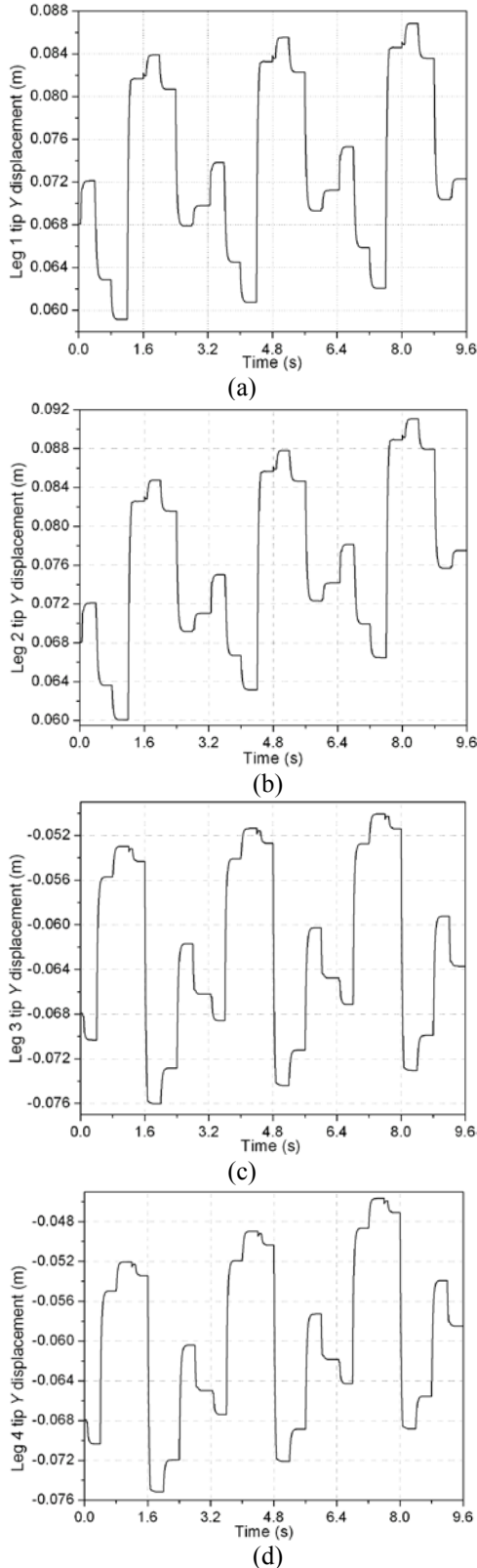
(a)

(b)

(c)

(d)

Figure 9: Simulation results of rigid legged quadruped robot locomotion: (a) Leg 1 tip Y displacement (m) versus time (s) (b) Leg 2 tip Y displacement (m) versus time (s) (c) Leg 3 tip Y displacement (m) versus time (s) (d) Leg 4 tip Y displacement (m) versus time (s)

results indicate that the robot is turning about positive

Z-axis. The turning about Z-axis can be attributed to the modelling of leg tip-ground interaction.

The frictional resistance offered by the ground to legged robot locomotion has been modeled, by appending a dissipative R-element (bond graph element) at the 1-junction representing leg tip velocity. The value of the parameter 'R' representing frictional resistance has been assigned greater value in X-direction compared to that in the locomotion direction i.e. $R_{gX} > R_{gY}$.

In the next section, the experimental realization of locomotion of quadruped robot prototype is presented.

5. EXPERIMENTAL RESULTS

Experimental set-up designed for realizing robot locomotion is presented in Fig. 10. The quadruped robot comprises of four legs and a body over which the controller CM5+ is mounted. The front and rear legs have been designed identical in all respects. Each leg of the quadruped robot has two rigid links connected through revolute joints, one at the hip and second one at the knee of a leg. The links are rotated through 'Dynamixel' series AX-12+ actuators, deployed at the joints.

Bioloid control behavior interface of the Robotis Inc. is used for the purpose of controlling robot locomotion. Control algorithm is fed to the CM5+ controller through a personal computer. A serial to USB data cable is used for the communication between the PC and the CM5+ controller. SMPS is used for supplying required power to the actuators and electronic circuitry. CM5+ controller uses ATMega 128 (128Kbyte flash memory) as the main processor in it. It operates in the voltage range of 7V-12V.

A program corresponding to the specified gait pattern is communicated to the controller CM5+. Consequently the robot locomotion is accomplished. Figure 11 presents the snapshots of locomotion of the quadruped robot. Snapshots indicate a clear progression of the robot. Figure 12 presents the leg tip and body CG
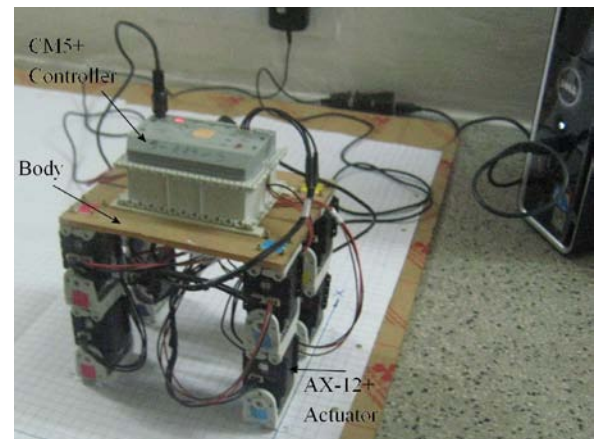

Figure 10: Experimental prototype of four legged robot

trajectory in XY plane, plotted using the experimental data. The body CG trajectory and the leg tip
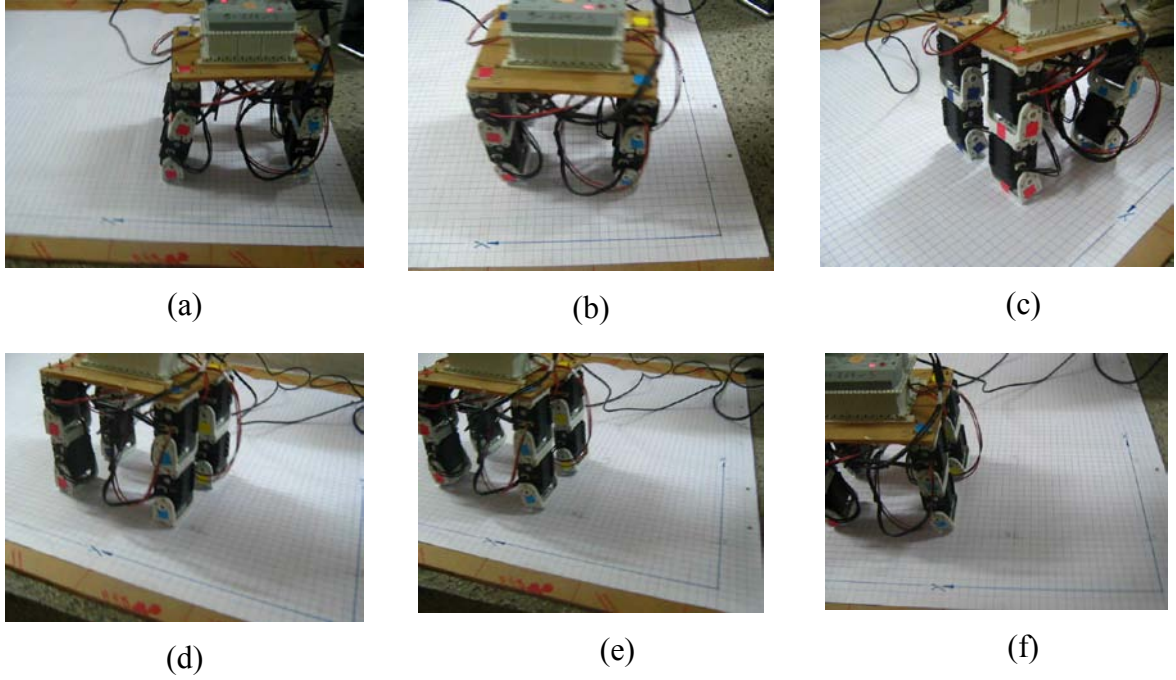
(a)

(b)

(c)

(d)

(e)

(f)

Figure 11: Snapshots of rigid legged quadruped robot locomotion

displacement plots indicate that the robot is almost progressing as desired, along a straight line in the locomotion direction. Slight deviation of the body CG and leg tip trajectories from the straight line path can be attributed to the unpredictable and arbitrary lag in command signals to the front or rear leg joint actuators.

Figure 13(a), (b), (c) and (d) respectively presents the actual and commanded joint angle displacement trajectories of joint 1 and 2 of the front and rear legs. There is a slight deviation in the commanded and actual joint trajectory along with some arbitrary lag in the front and rear joint trajectory.

6. CONCLUSIONS

The paper presents the coupled dynamics between the

Figure 12: Experimental results of rigid legged quadruped robot locomotion: Leg tip and body CG displacement along X-axis (m) versus Y-axis (m)

legs and body during locomotion. This study has been carried out through simulation as well as experiment on a four legged robot design with two articulate joints per leg. Modeling and simulation of the four legged robot has been carried out using bond graph technique. For the specified gait pattern, the body angular displacement trajectory for ψ and θ demonstrates the influence of ground reaction forces, transmitted through legs, on the robot body. The gait pattern has been tested on a quadruped robot experimental model and locomotion has been successfully realized.

This dynamic model of the four legged robot locomotion can be further extended for simulating running behavior, obstacle avoidance etc. Flexible leg concept may be incorporated to enable the robot to walk on uneven terrain. In fact the motivation for this work originated with a desire to build a model of four legged walking robot with flexible legs. Flexible legged robot model can be used to study the effect of flexible leg dynamics on the stability and impact tolerance of the robot.

APPENDIX

Table 1:Position of Frame {0} with respect to Body CG

Leg 'i'	R_{ix}	R_{iy}	R_{iz}
Leg 1	-0.042	0.068	-0.045
Leg 2	0.042	0.068	-0.045
Leg 3	-0.042	-0.068	-0.045
Leg 4	0.042	-0.068	-0.045

Table 2: Four Legged Robot Parameters

Parameters	Value
Robot body mass	M_b = 0.43Kg;
Moment of Inertia (M. I.) of body	I_{BXX} = 0.007Kg-m^2, I_{BYY} = 0.004Kg-m^2, I_{BZZ} = 0.002Kg-m^2
Leg link lengths	l_1 = 0.068m, l_2 = 0.025m
Leg link mass	M_{l1}= 0.075Kg; M_{l2} = 0.015Kg
M. I. of link '1'	I_{xx1} = 0.0002Kg-m^2, I_{yy1} = 0.00025Kg-m^2, I_{zz1} = 0.00001Kg-m^2
M. I. of link '2'	I_{xx2} = 0.00001Kg-m^2, I_{yy2} = 0.000003Kg-m^2, I_{zz2}= 0.000004Kg-m^2
Joint actuator parameters	Inductance: L_m =0.001H; Resistance: R_m= 0.1Ohms; Motor constant: K_t = 0.2 N-m/A; Gear ratio: n = 254
Leg tip-ground interaction	Stiffness: K_g = 100000N/m; Damping: R_{gZ} = 1000N-s/m; Frictional resistance: R_{gX} = 800N-s/m, R_{gY} = 400N-s/m
Gain Values	Proportional gain: K_P = 2500V/rad; Derivative gain: K_V =25V/rad/s

REFERENCES

Bowling A., 2005. Mobility and dynamic performance of legged robots. *Proceedings of IEEE International Conference on Robotics and Automation*, pp. 4100-4107, 18-22 April, Barcelona, Spain.

Craig J. J., (2006), *Introduction to Robotics– Mechanics and Control*, Pearson Education, Inc.

Estremera J. and Gonzalez de Santos P., 2002. Free gaits for quadruped robots over irregular terrain. *The International Journal of Robotics Research*, 21, 115-130.

Estremera J. and Waldron K. J., 2006. Leg thrust control for stabilization of dynamic gaits in a quadruped robot. *Proceedings of ROMANSY*, pp. 213-220, June 20-22, Warsaw, Poland.

Furusho J., Sano A., Sakaguchi M., and Koizumi E., 1995. Realization of bounce gait in a quadruped robot with articulate joint type legs. *Proceedings of IEEE International Conference on Robotics and Automation*, pp. 697-702.

Garcia E., Galvez J. A. and Gonzalez de Santos P., 2003. On finding the relevant dynamics for model based controlling walking robots. *Journal of Intelligent and Robotic Systems*, 37(4), 375-398.

Gonzalez De Santos P., Jimenez M. A., 1995. Path tracking with quadruped walking machines using discontinuous gaits. *Computers and Electrical Engineering*, 21(6), 383-396.

Hartikainen K. K., Halme A. J., Lehtinen H. and Koskinen K.O., 1992. MECANT I: A six legged
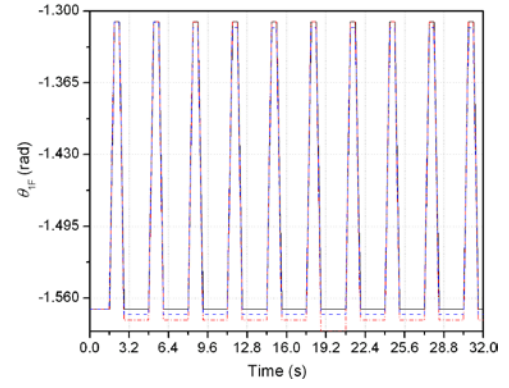
(a)
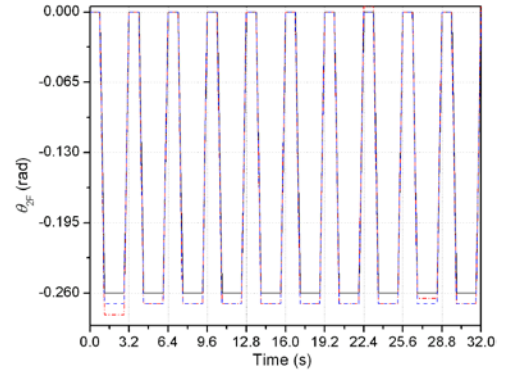
(b)
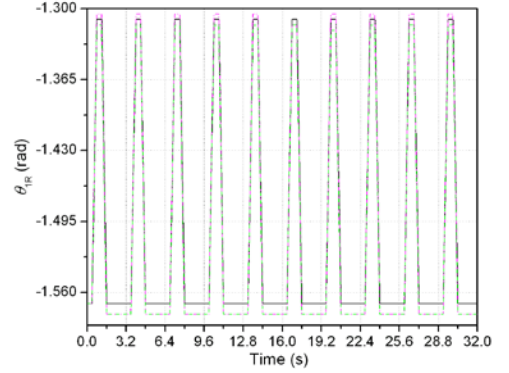
(c)
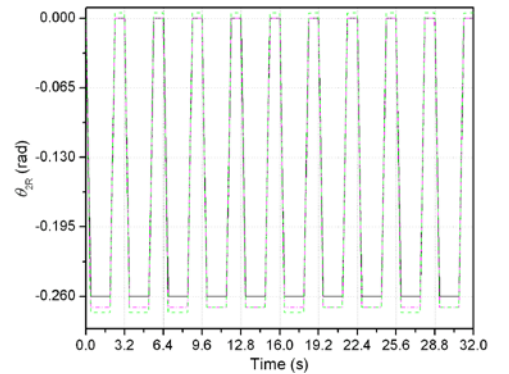
(d)

Figure 13: Experimental results: Joint angular displacements (a) θ_{1F} (radians) versus time (s) (b) θ_{2F} (radians) versus time (s) (c) θ_{1R} (radians) versus time (s) (d) θ_{2R} (radians) versus time (s)

walking machine for research purposes in outdoor environment. *Proceedings of IEEE International Conference on Robotics and Automation*, pp. 157-163. 12-14 May, Nice, France.

Inagaki S., Yuasa H., Suzuki T. and Arai T., 2006. Wave CPG model for autonomous decentralized multi-legged robot: gait generation and walking speed control", *Robotics and Autonomous Systems*, 54(2), 118-126.

Kurazume R., Yoneda K., and Hirose S., 2001. Feedforward and feedback dynamic trot gait control for a quadruped walking vehicle. *Proceedings of IEEE International Conference on Robotics and Automation*, Vol. 3, pp. 3172-3180, 21-26 May, Seoul, Korea.

Lasa Martin de and Buehler M., 2000. Dynamic compliant walking of a quadruped robot: preliminary experiments. *Proceedings of 3rd Int. Conf. on Climbing and Walking Robots*, 2-4 October, Madrid, Spain.

Mukherjee A., Karmarkar R. and Samantray, (2006). *Bondgraph in modeling simulation and fault identification*, New Delhi, I. K. International Publishing House Pvt. Ltd.

Mukherjee A., 2006. Users Manual of SYMBOLS Shakti, http://www.htcinfo.com/, High-Tech Consultants, S.T.E.P., Indian Institute of Technology, Kharagpur.

Waldron K. and McGhee R., 1986. The adaptive suspension vehicle. *Control Systems Magazine, IEEE*, 6(6), 7-12.

Wyffels, F., D'Haene, M., Waegeman, T., Caluwaerts, K., Nunes, C., Schrauwen, B., 2010. Realization of a passive compliant robot dog. *3rd IEEE International Conference RAS & EMBS on Biomedical Robotics and Biomechatronics (BioRob)*, pp. 882-886, 26-29 September, Tokyo, Japan.

Yoneda K. and Hirose S., 1992. Dynamic and static fusion gait of quadruped walking vehicle on winding path. *Proceedings of IEEE International Conference on Robotics and Automation*, pp. 143-148, Nice, France.

Zhang C. D. and Song S. M., 1993. A study of the stability of generalized wave gaits. *Mathematical Biosciences*, 115(1), 1-32.

A Final Marking Planning Method for Join Free Timed Continuous Petri nets

Hanife Apaydın Özkan and Aydın Aybar

Abstract— In this paper, an online control method is developed and corresponding algorithm is proposed for driving Join Free continuous Petri net from its initial marking, to target marking through a linear trajectory by minimizing the time. Then, the control problem in which some components of the target marking are not specified is considered and developed control method is used for that case.

I. INTRODUCTION

Discrete Petri Nets (PNs) are powerful graphical and mathematical tools for modeling, analysis and synthesis of *Discrete Event System*s (DESs) [1], [2]. The distributed state or marking of a PN is given by a vector of natural numbers which represent the number of tokens in each place. This is a significant advantage with respect to other formalisms such as automata, where the state space is a symbolic unstructured set.

Like in most modeling formalisms for DESs, PNs suffer from the so called *state explosion* which leads to an exponential growth of the size of state space with respect to the size of the system and population of initial state. Some relaxation techniques are studied to overcome this difficulty and to reduce the computational complexity of the analysis and synthesis of PNs (i.e. decomposition techniques, Lagrangian relaxations, fluidification and the others [3], [4], [5], [6]). *Fluidification* may be very useful relaxation technique when applied to highly populated systems.

For PNs, fluidification was introduced in [3], [7] aiming at giving fluid (continuous) approximation of original PN in the sense of behaviours and properties, and these models are called *continuous Petri nets*. The idea is to try to overcome, at least partially, the potentially very high computational complexity arising in many practical situations.

Different techniques have been proposed for control of continuous Petri nets in the literature [8], [9], [10], [11], [12]. Steady state optimal control of continuous Petri nets was studied in [13] where it is shown that, the optimal steady state control problem of continuous Petri net system can be solved by means of *Linear Programming Problem* (LPP) in the case that all transitions are controllable and the objective function is linear. For the problem of reaching a given steady state from an initial marking, implicit and explicit *Model Predictive Control* (MPC) methods minimizing a certain performance index are proposed in [8]. The step tracking problem, i.e. design of control laws to drive the system states to target references was considered in [10] and

Hanife Apaydın Özkan and Aydın Aybar are with Department of Electrical and Electronics Engineering, Anadolu University, Eskişehir, Turkey hapaydin1,aaybar@anadolu.edu.tr

a Lyapunov-function-based dynamic control algorithm was proposed for the problem. That method requires solving a *BiLinear Programming Problem* (BLP) for the computation of intermediate states. In [12] an efficient heuristics for minimum time control of continuous Petri nets, which aims at driving the system from an initial state to a target one by minimizing the time of a piecewise linear trajectory is developed.

In some control problems, final states of some places may not be specified, while that others are specified. Because reaching desired final states of other places in minimum time is more important then the unspecified final states. Accordingly, corresponding components of the target state are not specified in the control problem. In this paper, this problem is considered for continuous Petri nets for the first time in the literature.

For calculating the value of unspecified components under the objective of time minimization, we developed an online control strategy. This strategy focuses on timed continuous Petri nets (contPN) without synchronizations called Join Free contPN and drives JF contPN to a specified target state through a linear trajectory by means of LPP. This method proposes an online algorithm and requests solving BLP for calculating unspecified components.

The remainder of the paper is organized as follows. Section 2 briefly introduces the required concepts of contPN systems and introduces the formulation of applied control. A new control scheme for JF contPN is given in Section 3. In Section 4, a method for calculating unspecified components of target marking under the objective of time minimization is adressed. Finally, some conclusions and future directions are drawn in Section 5.

II. BASIC CONCEPTS AND NOTATION

We assume that the reader is familiar with Petri nets. A continuous Petri net system is a pair $\langle \mathcal{N}, \boldsymbol{m}_0 \rangle$ where $\mathcal{N} = \langle P, T, \boldsymbol{Pre}, \boldsymbol{Post} \rangle$ is a net structure where $P = \{p_1, p_2, ... p_{|P|}\}$ and $T = \{t_1, t_2, ... t_{|T|}\}$ are the sets of places and transitions, respectively; $\boldsymbol{Pre}, \boldsymbol{Post} \in \mathbb{N}^{|P| \times |T|}$ are pre and post matrices connecting places and transitions; $\boldsymbol{m}_0 \in \mathbb{R}_{\geq 0}^{|P|}$ is initial marking (state).

For a place $p_i \in P$ and a transition $t_j \in T$, Pre_{ij} and $Post_{ij}$ represent the weights of the arcs from p_i to t_j and from t_j to p_i, respectively. Each place p_i has a marking denoted by $m_i \in \mathbb{R}_{\geq 0}$. The vector of all token loads is called *state* or *marking*, and is denoted by $\boldsymbol{m} \in \mathbb{R}_{\geq 0}^{|P|}$. For every node $v \in P \cup T$, the sets of its input and output nodes are denoted as $^\bullet v$ and $v^\bullet$, respectively.

A transition $t_j \in T$ is enabled at $\boldsymbol{m}$ iff $\forall p_i \in^\bullet t_j$, $m_i > 0$ and its enabling degree is given by

$$enab(t_j, \boldsymbol{m}) = \min_{p_i \in^\bullet t_j} \left\{ \frac{m_i}{Pre_{ij}} \right\} \qquad (1)$$

which represents the maximum amount in which t_j can fire. An enabled transition t_j can fire in any real amount α, with $0 < \alpha \leq enab(t_j, \boldsymbol{m})$ leading to a new state $\boldsymbol{m'} = \boldsymbol{m} + \alpha \cdot \boldsymbol{C}_{\cdot j}$ where $\boldsymbol{C} = \boldsymbol{Post} - \boldsymbol{Pre}$ is the token flow matrix and $C_{\cdot j}$ is its jth column. If $\boldsymbol{m}$ is reachable from $\boldsymbol{m}_0$ through a finite sequence σ, the state (or fundamental) equation is satisfied: $\boldsymbol{m} = \boldsymbol{m}_0 + \boldsymbol{C} \cdot \boldsymbol{\sigma}$, where $\boldsymbol{\sigma} \in \mathbb{R}_{\geq 0}^{|T|}$ is the firing count vector, i.e., σ_j is the cumulative amount of firings of t_j in the sequence σ. The set of reachable markings from $\boldsymbol{m}_0$ is denoted by $RS(\mathcal{N}, \boldsymbol{m}_0)$.

Left and right natural annullers of the token flow matrix $\boldsymbol{C}$ are called *P-semiflows* (denoted by $\boldsymbol{y}$) and *T-semiflows* (denoted by $\boldsymbol{x}$), respectively. If $\exists \ \boldsymbol{y} > 0$, $\boldsymbol{y} \cdot \boldsymbol{C} = 0$, then the net is said to be *conservative*. If $\exists \ \boldsymbol{x} > 0$, $\boldsymbol{C} \cdot \boldsymbol{x} = 0$ it is said to be *consistent*.

A timed continuous Petri net (contPN) is a continuous Petri net together with a vector $\boldsymbol{\lambda} \in \mathbb{R}_{>0}^{|T|}$ where λ_j is the firing rate of t_j. As in untimed continuous Petri nets state equation summarizes the way the marking evolves along time. The state equation of contPN has an explicit dependence on time $\boldsymbol{m}(\tau) = \boldsymbol{m}_0 + \boldsymbol{C} \cdot \boldsymbol{\sigma}(\tau)$ where τ is global time. But, in continuous systems, the marking is continuously changing, so we may consider the derivative of $\boldsymbol{m}$ with respect to time. This way, $\dot{\boldsymbol{m}}(\tau) = \boldsymbol{C} \cdot \dot{\boldsymbol{\sigma}}(\tau)$ is obtained. Here, $\dot{\boldsymbol{\sigma}}(\tau)$ is flow throgh transitions and it is denoted by $\boldsymbol{f}(\tau) = \dot{\boldsymbol{\sigma}}(\tau)$. Hence, the state equation is

$$\dot{\boldsymbol{m}}(\tau) = \boldsymbol{C} \cdot \boldsymbol{f}(\tau) \qquad (2)$$

Different semantics have been defined for continuous timed transitions [14], [3]. Infinite server semantics is considered in this paper. Under this semantics, the flow of transition t_j is the product of firing rate, λ_j, and enabling of transition $enab(t_j, \boldsymbol{m}(\tau))$:

$$f_j(\tau) = \lambda_j \cdot enab(t_j, \boldsymbol{m}(\tau)) = \lambda_j \cdot \min_{p_i \in^\bullet t_j} \left\{ \frac{m_i(\tau)}{Pre_{ij}} \right\} \qquad (3)$$

For the sake of simplicity τ is omitted in the rest of the paper.

We consider Join Free contPNs (JF contPNs) which satisfy $|^\bullet t_j| \ \forall j \in \{0...|T|\}$. Let us define matrix $\boldsymbol{\Pi} \in \mathbb{R}_{\geq 0}^{|T| \times |P|}$ as:

$$\Pi_{ji} = \begin{cases} \frac{1}{Pre_{ij}}, & if \ \ Pre_{ij} \neq 0 \\ 0, & \text{otherwise} \end{cases} \qquad (4)$$

The state equation of uncontrolled JF contPN is as follow:

$$\dot{\boldsymbol{m}} = \boldsymbol{C} \cdot \boldsymbol{f} = \boldsymbol{C} \cdot \boldsymbol{\Lambda} \cdot \boldsymbol{\Pi} \cdot \boldsymbol{m} \qquad (5)$$

where $\boldsymbol{\Lambda} = \text{diag}\{\lambda_1, ... \lambda_{|T|}\}$. Reachability set of JF contPN is denoted by $RS(\mathcal{N}, \ \boldsymbol{\lambda}, \ \boldsymbol{m}_0)$. Given that the continuous Petri nets that we are considering are Join Free and every transition is fireable, the set of reachable markings is equal to the solutions of the state equation $RS(\mathcal{N}, \ \boldsymbol{\lambda}, \ \boldsymbol{m}_0) = \{\boldsymbol{m} \mid \boldsymbol{m} = \boldsymbol{m}_0 + \boldsymbol{C} \cdot \boldsymbol{\sigma}\}, \boldsymbol{\sigma} \in \mathbb{R}_{\geq 0}^{|T|}$.

A. Control Scheme

Now let us introduce control concept that we consider in this paper. In contPN, a transition is associated in general to a machine and this machine can not work faster then its maximum firing rate, the only control action we consider is to brake it down. In other words, we assume that the only action that can be applied to contPN is to reduce the flow of transitions [15]. If a transition can be controlled (its flow can be reduced or even stopped), we will say that it is a *controllable* transition [13]. In this paper, it is assumed that all transitions are controllable.

The controlled flow, $\boldsymbol{w}$, of a contPN is defined as $\boldsymbol{w}(\tau) = \boldsymbol{f}(\tau) - \boldsymbol{u}(\tau)$, with $\boldsymbol{0} \leq \boldsymbol{u}(\tau) \leq \boldsymbol{f}(\tau)$, where $\boldsymbol{f}$ is the flow of the uncontrolled system, i.e., defined as in (5), and $\boldsymbol{u}$ is the control action.

Therefore, the control input $\boldsymbol{u}$ is dynamically upper bounded by the flow $\boldsymbol{f}$ of the corresponding unforced system. Under these conditions, the overall behaviour of a JF contPN system in which all transitions are controllable is ruled by the following system:

$$\begin{aligned} \dot{\boldsymbol{m}} &= \boldsymbol{C} \cdot [\boldsymbol{f} - \boldsymbol{u}] = \boldsymbol{C} \cdot \boldsymbol{w} \\ \boldsymbol{0} &\leq \boldsymbol{u} \leq \boldsymbol{f} \end{aligned} \qquad (6)$$

The constraint $\boldsymbol{0} \leq \boldsymbol{u} \leq \boldsymbol{f}$ can be rewritten as $\boldsymbol{0} \leq \boldsymbol{f} - \boldsymbol{u} \leq \boldsymbol{f}$. From the definition, $\boldsymbol{w} = \boldsymbol{f} - \boldsymbol{u}$, the constraint can be expressed as:

$$\boldsymbol{0} \ \leq \ \boldsymbol{w} \ \leq \ \boldsymbol{\Lambda} \cdot \boldsymbol{\Pi} \cdot \boldsymbol{m} \qquad (7)$$

The following sections focus on the control problem for JF subclass in the case that some components of target markings are not specified. In Section 3, an online control method is developed and corresponding algorithm is proposed for driving JF contPN from its initial marking, $\boldsymbol{m}_0$, to target marking, $\boldsymbol{m}_f$, through a linear trajectory by minimizing the time. Section 4 makes use of developed method for the control problem in which some components of target marking are not specified. We assume that $\boldsymbol{m}_0$ and $\boldsymbol{m}_f$ are strictly positive. The assumption that $\boldsymbol{m}_0$ is positive ensures that the system can move at $\tau = 0$ in the direction of $\boldsymbol{m}_f$ [16]; the assumption that $\boldsymbol{m}_f$ is positive ensures that $\boldsymbol{m}_f$ can be reached in finite time [13].

III. A CONTROL METHOD FOR JF CONTPNS

In this section, an online control method that drives the system from the initial marking $\boldsymbol{m}_0$ to a desired target marking $\boldsymbol{m}_f$ through a linear trajectory will be introduced.

Our procedure consists of using discrete time representation of the system, and calculating control input at each sampling instant by using the maximum flows of transitions at $\boldsymbol{m}_0$ in the direction to $\boldsymbol{m}_f$ and maximum flows of transitions at $\boldsymbol{m}_f$ in the direction from $\boldsymbol{m}_0$ to $\boldsymbol{m}_f$.

Maximum flow of transitions at $\boldsymbol{m}_0$ in the direction to $\boldsymbol{m}_f$ is denoted by $\boldsymbol{w}_0$ and it is calculated by the following

LPP, where $\boldsymbol{s}_0 = \boldsymbol{w}_0 \cdot \tau_0$:

$$
\begin{aligned}
\min_{\boldsymbol{s}_0} \quad & \tau_0 \\
\text{s.t.} \quad & \boldsymbol{m}_f = \boldsymbol{m}_0 + \boldsymbol{C} \cdot \boldsymbol{s}_0 && (a) \\
& 0 \le s_{0_j} \le \lambda_j \cdot \Pi_{ji} \cdot m_{0i} \cdot \tau_0 \\
& \forall\, j \in \{1,..,|T|\} \text{ where } i \text{ satisfies } \Pi_{ji} \ne 0 && (b)
\end{aligned} \qquad (8)
$$

The equations correspond to: (a) the straight line connecting $\boldsymbol{m}_0$ to $\boldsymbol{m}_f$, (b) flow constraints at $\boldsymbol{m}_0$. Notice that (b) is a linear constraint because m_{0i} and m_{f_i} are known $\forall i \in \{1,2,..,|P|\}$.

Maximum flow of transitions at $\boldsymbol{m}_f$ in the direction from $\boldsymbol{m}_0$ to $\boldsymbol{m}_f$ is denoted by $\boldsymbol{w}_f$ and it is calculated by the following LPP, where $\boldsymbol{s}_f = \boldsymbol{w}_f \cdot \tau_f$:

$$
\begin{aligned}
\min_{\boldsymbol{s}_f} \quad & \tau_f \\
\text{s.t.} \quad & \boldsymbol{m}_f = \boldsymbol{m}_0 + \boldsymbol{C} \cdot \boldsymbol{s}_f && (a) \\
& 0 \le s_{f_j} \le \lambda_j \cdot \Pi_{ji} \cdot m_{f_i} \cdot \tau_f \\
& \forall\, j \in \{1,..,|T|\} \text{ where } i \text{ satisfies } \Pi_{ji} \ne 0 && (b)
\end{aligned} \qquad (9)
$$

The equations correspond to: (a) the straight line connecting $\boldsymbol{m}_0$ to $\boldsymbol{m}_f$, (b) flow constraints at $\boldsymbol{m}_f$. Notice that (b) is a linear constraint because m_{0i} and m_{f_i} are known $\forall i \in \{1,2,..,|P|\}$.

Proposition: Let $\langle \mathcal{N}, \boldsymbol{\lambda}, \boldsymbol{m}_0 \rangle$ be a contPN system with $\boldsymbol{m}_0 > \mathbf{0}$. If $\boldsymbol{m}_f$ belongs to $RS(\mathcal{N}, \boldsymbol{\lambda}, \boldsymbol{m}_0)$ and $\boldsymbol{m}_f > 0$, then LPPs in (8) and (9) are feasible.

Proof: Since $\boldsymbol{m}_f$ is a reachable marking, then there exists $\boldsymbol{s}$ such that the state equations (8)(a) and (9)(a) are satisfied. By taking τ_0 and τ_f sufficiently large (8)(b) and (9)(b) can be satisfied since $\lambda_j \cdot \Pi_{ji} \cdot m_{0i} > 0$ and $\lambda_j \cdot \Pi_{ji} \cdot m_{f_i} > 0$. $\square$

Since it is assumed that, $\boldsymbol{m}_0 > \mathbf{0}$ and $\boldsymbol{m}_f > \mathbf{0}$, then the linear trajectory from $\boldsymbol{m}_0$ to $\boldsymbol{m}_f$ can be followed by the system [16]. At each marking $\boldsymbol{m}$ on the line connecting $\boldsymbol{m}_0$ to $\boldsymbol{m}_f$ (i.e. $\boldsymbol{m} = \alpha \cdot \boldsymbol{m}_0 + (1-\alpha) \cdot \boldsymbol{m}_f$, $\alpha \in [0\ 1]$), the maximum flow in the direction to $\boldsymbol{m}_f$ is calculated by

$$
\boldsymbol{w} = \alpha \cdot \boldsymbol{w}_0 + (1-\alpha) \cdot \boldsymbol{w}_f \qquad (10)
$$

Corresponding control action is calculated easily by $\boldsymbol{u} = \boldsymbol{f} - \boldsymbol{w}$ where $\boldsymbol{f}$ and $\boldsymbol{w}$ are controlled and uncontrolled flow vectors, respectively.

In order to calculate control inputs to drive the system from $\boldsymbol{m}_0$ to $\boldsymbol{m}_f$ through a linear trajectory and drive the system to $\boldsymbol{m}_f$ by using the calculated control we propose to use discrete-time representation of contPN. The discrete-time representation of the continuous-time system (6) is given by:

$$
\begin{aligned}
& \boldsymbol{m}[k+1] = \boldsymbol{m}[k] + \Theta \cdot \boldsymbol{C} \cdot \boldsymbol{w}[k] \\
& 0 \le \boldsymbol{w}[k] \le \boldsymbol{\Lambda} \cdot \boldsymbol{\Pi} \cdot \boldsymbol{m}[k]
\end{aligned} \qquad (11)
$$

Here Θ is the sampling period ($\tau = k \cdot \Theta$) and $\boldsymbol{m}[k]$ is the marking at step k, i.e., at time $k \cdot \Theta$. The sampling period should be small enough to avoid to reach negative markings. Let us consider a place p_i with $p_i^\bullet = \{t_1, t_2, ..., t_j\}$ and $m[k]_i > 0$. Then state equation can be written as $m[k+1]_i = m[k]_i + \Theta \cdot \boldsymbol{C}(i,:) \cdot \boldsymbol{w}[k] \ge m[k]_i - \Theta \cdot (\lambda_1 + \lambda_2 + ... + \lambda_j) \cdot m[k]_i = m[k]_i \cdot (1 - \sum_{t_j \in p^\bullet} \lambda_j \cdot \Theta) \ge 0$. Hence, if Θ is chosen

Algorithm 1

Input: $\langle \mathcal{N}, \boldsymbol{m}_0 \rangle$, $\boldsymbol{m}_f$, Θ
Step 1) Solve LPP in (8) and LPP in (9)
Step 2) k=0
Step 3) If $\boldsymbol{m}[k] \ne \boldsymbol{m}_f$
Step 4) Measure

$$
\boldsymbol{m}[k+1] = \boldsymbol{m}[k] + \Theta \cdot \boldsymbol{C} \cdot \boldsymbol{w}[k] \qquad (13)
$$

Step 5) Calculate corresponding α

$$
\alpha = \frac{\boldsymbol{m}[k+1]_1 - \boldsymbol{m}_{f_1}}{\boldsymbol{m}_{0_1} - \boldsymbol{m}_{f_1}} \qquad (14)
$$

Step 6) If $\alpha > 1$

$$
\Theta = \frac{\boldsymbol{m}_{f_1} - \boldsymbol{m}[k]_1}{\boldsymbol{C}(1,:) \cdot \boldsymbol{w}[k]} \qquad (15)
$$

Step 7) Advance one step

$$
\boldsymbol{m}[k+1] = \boldsymbol{m}[k] + \Theta \cdot \boldsymbol{C} \cdot \boldsymbol{w}[k] \qquad (16)
$$

Step 8) Calculate corresponding $\boldsymbol{w}[k+1]$

$$
\boldsymbol{w}[k+1] = \alpha \cdot \boldsymbol{w}_0 + (1-\alpha) \cdot \boldsymbol{w}_f \qquad (17)
$$

Step 9) $k = k + 1$

such that

$$
\sum_{t_j \in p^\bullet} \lambda_j \cdot \Theta < 1 \qquad (12)
$$

then any marking reachable from $\boldsymbol{m}_0 = \boldsymbol{m}[0] \ge 0$ is nonnegative [8].

In order to calculate controlled flows and drive the system to $\boldsymbol{m}_f$, Algorithm 1 is developed. In this algorithm, controlled flow is calculated at each sampling instant by using the fact given in (10). Here, the controlled flows $\boldsymbol{w}[0] = \boldsymbol{w}_0$ and $\boldsymbol{w}_f$ are obtained by solving LPPs in (8) and (9), respectively. Then, at the first iteration, $\boldsymbol{m}[1]$, α and $\boldsymbol{w}[1]$ (by using obtained α) are calculated. And obtained controlled flow is realized. At the next iteration, $\boldsymbol{m}[2]$, α and $\boldsymbol{w}[2]$ are calculated by similar way. This procedure is repeated until $\boldsymbol{m}_f$ is reached. During the execution of the algorithm, if α is obtained as bigger than 1, that is $\boldsymbol{m}_f$ is passed at the current iteration, Θ is recalculated to reach $\boldsymbol{m}_f$ accurately. We developed MATLAB program for Algorithm 1. This program is implemented on a PC with Intel(R) Core(TM) 2CPU T5600 @ 1.83GHz, 2.00 GB of RAM.

Example 1: Let us consider JF contPN in Fig. 1 with $\boldsymbol{\lambda} = [1\ 1\ 1\ 1\ 1\ 1]^T$. The only minimal P-semiflow is $\boldsymbol{y} = [2\ 2\ 1\ 1\ 1]$ and there are two minimal T-semiflows $\boldsymbol{x}^1 = [1\ 1\ 0\ 0\ 0\ 0]^T$ and $\boldsymbol{x}^2 = [0\ 0\ 1\ 1\ 1\ 1]^T$. Our aim is to drive the system from $\boldsymbol{m}_0 = [13\ 3\ 4\ 4\ 5]^T$ to a final state $\boldsymbol{m}_f = [10\ 6\ 6\ 3\ 2]^T$ by using the proposed control method. The system dynamics can be described as follows:

$$
\begin{aligned}
\dot{m}_1 &= m_2 + \tfrac{1}{2} \cdot m_5 - m_1 - m_1 \\
\dot{m}_2 &= m_1 - m_2 \\
\dot{m}_3 &= m_1 - m_3 \\
\dot{m}_4 &= m_1 - m_4 \\
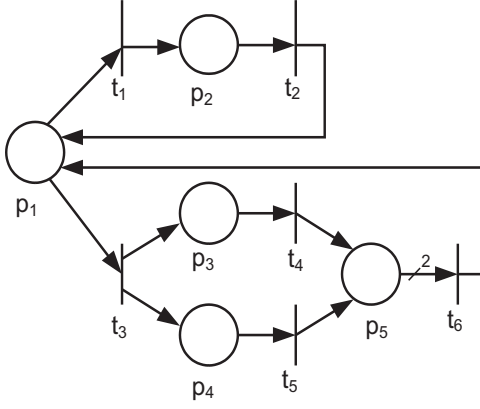\dot{m}_5 &= m_3 + m_4 - m_5
\end{aligned} \qquad (18)
$$

Fig. 1: A JF Petri net [9]

Fig. 3: Evolution of control actions and controlled flows of t_1 and t_2 for Example 1

At $\boldsymbol{m}_0$, solving LPP (8) yields $\boldsymbol{w}_0 = [5.8 \ 2.05 \ 2.5 \ 0 \ 1.25 \ 2.5]^T$ $\tau_0 = 0.8$ t.u. At $\boldsymbol{m}_f$, solving LPP (9) yields $\boldsymbol{w}_f = [4.04 \ 2.54 \ 1 \ 0 \ 0.5 \ 1]^T$ $\tau_f = 2$ t.u.

By executing Algorithm 1 ($\Theta = 0.01$), $\boldsymbol{m}_f$ is reached by 122 discrete steps, which corresponds to 1.22 time unit (t.u.) Evolution of markings m_1, m_2 and m_3 and, control actions-controlled flows of transitions t_1 and t_2 are shown in Fig. 2 and 3, respectively.
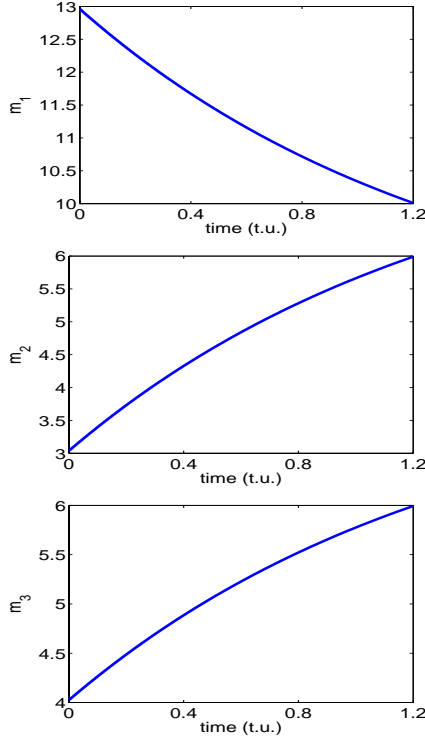

Fig. 2: Evolution of m_1 and m_2 for Example 1

IV. FINAL MARKING PLANNING FOR JOIN FREE CONTPNS

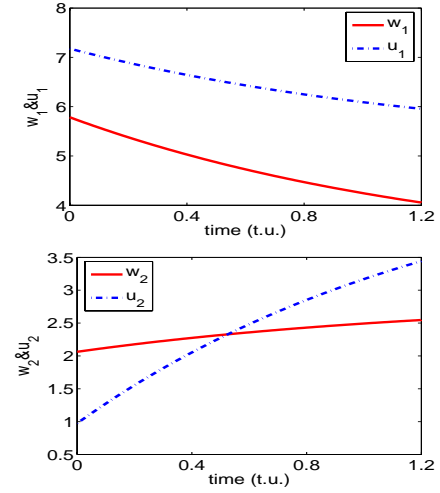So far, we introduced an online control method for driving the system from a given initial state $\boldsymbol{m}_0$ to the specified target marking. In this section, we consider the control problems that final markings of some places are specified, while the others are not. The set of places whose final markings are specified precisely is denoted by $P_{sp} = \{p_i \mid m_{f_i} \text{ is specified}\}$ and the set of places whose final markings are not specified is denoted by $P_{un} = \{p_i \mid m_{f_i} \text{ is un-specified}\}$. In order to calculate the m_{f_i} $\forall p_i \in P_{un}$ by minimizing the time (or maximizing controlled flows $\boldsymbol{w}_0$ and $\boldsymbol{w}_f$), Algorithm 1 is used again with only one modification. Thus, as differ from the first case we propose to solve BLP in (19) instead of LPPs (8) and (9) with variables τ_0, τ_f, $\boldsymbol{s}_0$, $\boldsymbol{s}_f$, m_{f_i} $\forall p_i \in P_{un}$ where $\boldsymbol{s}_0 = \boldsymbol{w}_0 \cdot \tau_0$, $\boldsymbol{s}_f = \boldsymbol{w}_f \cdot \tau_f$:

$$
\begin{aligned}
\min \quad & \tau_0 + \tau_f \\
\text{s.t.} \quad & \boldsymbol{m}_f = \boldsymbol{m}_0 + \boldsymbol{C} \cdot \boldsymbol{s}_0 && (a1) \\
& 0 \le s_{0j} \le \lambda_j \cdot \Pi_{ji} \cdot m_{0i} \cdot \tau_0 \\
& \quad \forall \, j \in \{1,..,|T|\} \text{ where } i \text{ satisfies } \Pi_{ji} \neq 0 && (a_2) \\
\\
& \boldsymbol{m}_f = \boldsymbol{m}_0 + \boldsymbol{C} \cdot \boldsymbol{s}_f && (b1) \\
& 0 \le s_{fj} \le \lambda_j \cdot \Pi_{ji} \cdot m_{f_i} \cdot \tau_f \\
& \quad \forall \, j \in \{1,..,|T|\} \text{ where } i \text{ satisfies } \Pi_{ji} \neq 0 && (b_2) \\
\\
& \boldsymbol{m}_f = \boldsymbol{m}_0 + \boldsymbol{C} \cdot \boldsymbol{\sigma}, \; \boldsymbol{\sigma} > \boldsymbol{0} && (c)
\end{aligned}
\tag{19}
$$

The equations correspond to: (a1)&(b1) the equation of the straight line connecting $\boldsymbol{m}_0$ to $\boldsymbol{m}_f$; (a2)&(b2) flow constraints at $\boldsymbol{m}_0$ and $\boldsymbol{m}_f$, respectively; (c) reachability condition of $\boldsymbol{m}_f$. Note that, since the net we consider is consistent (19)(c) is equivalent to $\boldsymbol{B}_y^T \cdot \boldsymbol{m}_f = \boldsymbol{B}_y^T \cdot \boldsymbol{m}_0$, $\boldsymbol{m}_f > \boldsymbol{0}$ where $\boldsymbol{B}_y^T$ is basis of *P-semiflows* [13].

Example 2: Let us go back to Example 1 with the same initial marking $\boldsymbol{m}_0 = [13 \ 3 \ 4 \ 4 \ 5]^T$ and Θ. But in this case, final marking of some places (not all) are specified: $m_{f_1} = 10$, $m_{f_4} = 3$, $m_{f_5} = 2$, that is $P_{sp} = \{p_1, p_4, \ p_5\}$ and $P_{un} = \{p_2, p_3\}$. Our objective is to find the final markings of m_{f_2} and m_{f_3} by minimizing the time and drive the system

to fulfilled final marking. For this example, BLP in (19) leads to:

$$
\begin{aligned}
\min \quad & \tau_0 + \tau_f \\
\text{s.t.} \quad & 10 = 13 + s_{02} + s_{06} - s_{01} - s_{03} \\
& m_{f_2} = 3 + s_{01} + s_{02} \\
& m_{f_3} = 4 + s_{03} - s_{05} \qquad\qquad (a1) \\
& 6 = 4 + s_{03} - s_{05} \\
& 6 = 5 + s_{04} + s_{05} - 2 \cdot s_{06} \\[4pt]
& 0 \le s_{01} \le 13 \cdot \tau_0 \\
& 0 \le s_{02} \le 3 \cdot \tau_0 \\
& 0 \le s_{03} \le 13 \cdot \tau_0 \\
& 0 \le s_{04} \le 4 \cdot \tau_0 \qquad\qquad (a2) \\
& 0 \le s_{05} \le 4 \cdot \tau_0 \\
& 0 \le s_{06} \le 2.5 \cdot \tau_0 \\[6pt]
& 10 = 13 + s_{f2} + s_{f6} - s_{f1} - s_{f3} \\
& m_{f_2} = 3 + s_{f1} + s_{f2} \\
& m_{f_3} = 4 + s_{f3} - s_{f5} \qquad\qquad (b1) \\
& 6 = 4 + s_{f3} - s_{f5} \\
& 6 = 5 + s_{f4} + s_{f5} - 2 \cdot s_{f6} \\[4pt]
& 0 \le s_{f1} \le 10 \cdot \tau_f \\
& 0 \le s_{f2} \le m_{f_2} \cdot \tau_f \\
& 0 \le s_{f3} \le 10 \cdot \tau_f \qquad\qquad (b2) \\
& 0 \le s_{f4} \le m_{f_3} \cdot \tau_f \\
& 0 \le s_{f5} \le 6 \cdot \tau_f \\
& 0 \le s_{f6} \le 3 \cdot \tau_f \\[4pt]
& 2 \cdot m_{f_2} + m_{f_3} = 18 \qquad\qquad (c)
\end{aligned}
\tag{20}
$$

By solving the BLP in (20), markings of unspecified places are obtained as $m_{f_2} = 6.5$ and $m_{f_3} = 5$, that is $\boldsymbol{m}_f = [10\ 6.5\ 5\ 3\ 2]^T$. By executing Algorithm 1, $\boldsymbol{m}_f$ is reached by 90 discrete steps (0.90 t.u.). Evolution of markings m_1, m_2 and m_3 and control actions-controlled flows of transitions t_1 and t_2 are shown in Fig. 4 and 5, respectively.

V. CONCLUSION

An online control method is developed for JF contPN. The method takes discrete time representation. In this method, in order to drive the system from its initial marking to target marking, corresponding control action is calculated and applied at each time step. Algorithm 1 which uses LPP is developed for this method.

In some control problems, target markings of some places are given while that of others are not specified. In that case, we propose to calculate unspecified target markings of places by solving a BLP with time minimization objective. Then Algorithm 1 is executed again with a simple modification.

An interesting point is to extend this work to more general structures, that is for other types of contPNs.

REFERENCES

[1] D. Mandrioli, A. Morzenti, M. Pezze, P. Pietro S., and S. Silva. A Petri net and logic approach to the specification and verification of real time systems. In *Formal Methods for Real Time Computing*. John Wiley & Sons Ltd., 1996.
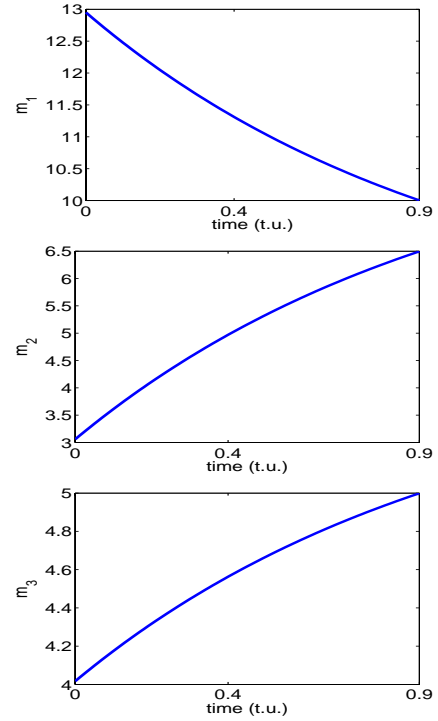
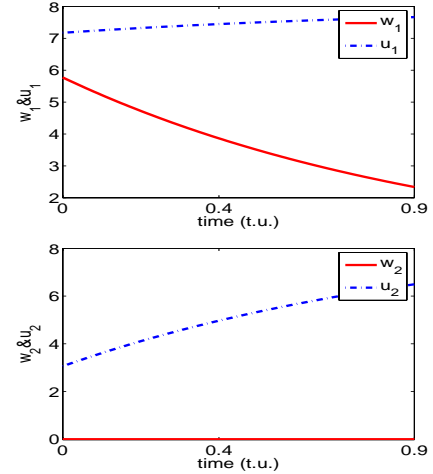Fig. 4: Evolution of m_1 and m_2 for Example 2

Fig. 5: Evolution of control actions and controlled flows of t_1 and t_2 for Example 2

[2] T. Murata. Petri nets: properties, analysis and applications. *Proceedings of the IEEE*, 77:541–580, 1989.
[3] M. Silva and L. Recalde. Petri nets and integrality relaxations a view of continuous Petri nets. *IEEE Trans on Systems Man and Cybernetics*, 32(4):314–327, 2002.
[4] A. Aybar, A. İftar, and H. Apaydin-Özkan. Centralized and decentralized supervisory controller design to enforce boundedness, liveness, and reversibillity in petri nets. *International Journal of Control*, pages 537–553, 2005.
[5] A. Aybar and A. İftar. Decentralized structural controller design for large-scale discrete event systems modelled by petri nets. *Kybernetika*, 1(45):3–14, 2009.
[6] T. Nishi, K. Shimatani, and M. Inuiguchi. Lagrangian relaxation technique for solving scheduling problems by decomposition of timed petri nets. In *Proceedings of the 17th IFAC World Congress*, volume 17,

COEX, Korea, South, 2008.

[7] R. David and H.Alla. Continuous Petri nets. In *8th European Workshop on Application and Theory of Petri Nets*, Zaragoza, Spain, 1987.

[8] C. Mahulea, A. Giua, L. Recalde, C. Seatzu, and M. Silva. Optimal model predictive control of timed continuous Petri nets. *IEEE Transactions on Automatic Control*, 53(7):1731 – 1735, August 2008.

[9] J. Xu, L. Recalde, and M. Silva. Tracking control of join-free timed continuous Petri net systems under infinite servers semantics. *Discrete Event Dynamic Systems*, 18(2):263–288, 2008.

[10] J. Xu, L. Recalde, and M. Silva. Tracking control of timed continuous Petri net systems under infinite servers semantics. In *17th World Congress of IFAC*, pages 3192–3197, Seoul, Korea, 2008.

[11] C.R. Vázquez and M. Silva. Piecewise-linear constrained control for timed continuous petri nets. In *IEEE Conference on Decision and Control*, Shanghai, China, 12/2009 2009.

[12] Hanife Apaydin-Özkan, Jorge Júlvez, Cristian Mahulea, and Manuel Silva. Approaching minimum time control of timed continuous petri nets. *Nonlinear Analysis: Hybrid Systems*, 5(2):136 – 148, 2011. Special Issue related to IFAC Conference on Analysis and Design of Hybrid Systems (ADHS'09) - IFAC ADHS'09.

[13] C. Mahulea, A. Ramirez, L. Recalde, and M.Silva. Steady state control reference and token conservation laws in continuous Petri net systems. *IEEE Transactions on Automation Science and Engineering*, 5(2):307–320, April 2008.

[14] H.Alla and R.David. Continuous and hybrid Petri nets. *Journal of Circuits, Systems, and Computers*, 8:159–188, 1998.

[15] M. Silva and L. Recalde. On fluidification of Petri net models: from discrete to hybrid and continuous models. *Annual Reviews in Control*, 28(2):253–266, 2004.

[16] J. Júlvez, L. Recalde, and M. Silva. On reachability in autonomous continuous Petri net systems. In *24th International Conference on Application and Theory of Petri Nets*, pages 221–240, Eindhoven, Netherlands, June 2003. Springer.

MULTI-SCALE EXTENSION OF DISCRIMINANT PLS FOR FAULT DETECTION AND DIAGNOSIS

Mohammad Sadegh Emami Roodbali [a], Mehdi Shahbazian[b]

[a]Petroleum University of Technology, Ahwaz, Iran

[b]Petroleum University of Technology, Ahwaz, Iran

[a]emamimohamads@yahoo.com

ABSTRACT

A new approach based on the Partial Least Squares (PLS) and Wavelet Transform is presented for the industrial process monitoring. A different scheme for applying PLS for multiple faults diagnosis is used in this approach. Because of multi-scale nature of the variable measurements in the most of industrial processes the Discrete Wavelet Transform (DWT) is applied to extract the multi-scale features of these measurements. Comparison of the ability of this Multi-Scale PLS (MSPLS) algorithm with the PLS to diagnosis the multiple faults in the Tennessee Eastman process (TEP) benchmark, demonstrates the efficiency of the proposed approach and indicates that this MSPLS algorithm can be useful for process monitoring and detection and diagnosis multiple faults.

Keywords: fault, detection, diagnosis, discriminant PLS, MSPLS

1. INTRODUCTION

P Partial Least Squares (PLS) structure is one of the Statistical Process Monitoring (SPM) methods that widely used for monitoring the abnormal situations that happen in the processes. PLS projects the input-output data down into a latent space, extracting a number of main factors with an orthogonal structure, while capturing most of the variance in the original data (Geladi and Kowalski 1986; Wold et al. 1984). A popular application of PLS is to select the predictor block X, containing the variables measurements and the predicted block Y, containing the product quality data (Raich and Cinar 1995). This model can be used for detecting, identifying and diagnosing the faults (Piovoso and Kosanovich 1994). Another application of PLS mainly focusing on fault diagnosis is to define Y as class membership (Chaing Russell, and Braatz 2000). This PLS model is known as discriminant Partial Least Squares. To diagnosing the multiple faults in the process, discriminant Partial Least Squares is applied in this study.

Similar to the other statistical process monitoring methods, there are some limitations for applying PLS on process monitoring. Most processes in modern industrial plants are typically complex, and such a complexity seems reflected in collected data, which contain the cumulative effect of many underlying phenomena and disturbances, with different form in the time and frequency domain (Reis Saraiva and Bakshi 2008).Therefore, the overall systems are composed of processing units that have different time scales and frequency bands (Reis Saraiva and Bakshi 2008).For detecting, identifying and diagnosing events in these systems using statistical (data-driven) methods, the collected data blocks, containing the measured variables should be assayed and treated in several scales. In discriminant Partial Least Squares, all of the process variables data and quality variables will be gathered into one data block. Therefore the other limitation is the autocorrelation of variables.

Wavelet Transform is able to decompose the variables into different scales representation. Also, the online wavelet decomposition (includes downsampling) is useful to decorrelate the autocorrelation between the measurements (Ganesan Das and Venkataraman 2004).

In this study, an online Wavelet Transform is applied to the discriminant Partial Least Squares to build a MSPLS model for process monitoring. The Tennessee Eastman Process (TEP) data with multiple faults is used to examine the ability of the proposed MSPLS algorithm to diagnosis these multiple faults.

2. TENNESSEE EASTMAN PROCESS

The Tennessee Eastman Process (TEP) was created by the Eastman Chemical Company to provide a realistic industrial process for evaluating process control and monitoring methods (Downs and Vogel 1993).

The test process is based on a simulation of an actual industrial process where the components, kinetics, and operating conditions have been adjusted for specific aims. The process consists of five major units: a reactor, condenser, compressor, separator, and stripper; and, it contains eight components: A, B, C, D, E, F, G, and H (Chaing Russel and Braatz 2001).
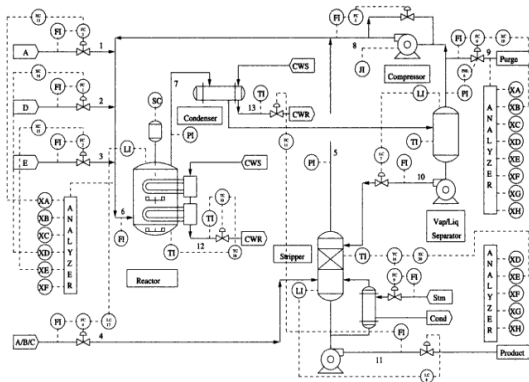
Figure1: Tennessee Eastman Process (Lyman 1995)

As shown in Figure 1, the gaseous reactants A, C, D, and E and the inert B are fed to the reactor where the liquid products G and H are formed and the species F is a by-product of the reactions. The labels in Figure 1 represent flow meters (FI), thermometers (TI), pressure gauges (PI), level detectors (LI), agitator speed control (SC), steam supply (Stm), and cooling water supply/recycle (CWS/CWR) (Wilson and Irwin 2000).

The process contains 53 variables containing 41 measured and 12 manipulated variables. The measurements of these 53 variables, is generated from the open-loop and the closed-loop simulations for the Tennessee Eastman process (TEP) as well as the training and testing data files used for evaluating the statistical methods (PCA, PLS, FDA, and CVA).

The training set used in this study consists of 500 observations for each variable which was generated with no fault and 1440 observation generated under three programmed faults. Fault 1 is connected to the step change in the cooling water. Fault 2 is a low drift in the reaction kinetics, and Fault 3 is associated with one of the sticking valves. The testing set contains of 3840 observations which starts with normal operation. Then, each of the faults mentioned above occurs to the system at determinate times.

3. MODEL DESCRIPTION

3.1. Discriminant PLS modeling

Discriminant PLS selects the matrix X, containing all process variables and selects the matrix Y, to focus PLS on the task of fault diagnosis (Chaing, Russell and Braatz 2000). To determine the predicted class in the prediction step, discriminant analysis is used (Nouwen et al. 1997).

To apply discriminant PLS for multiple fault diagnosis, the model is trained with observations of normal operation and also with faulty observations. The output (Y) of the training data is no longer the quality variables, the predicted variables are dummy variables (0 or 1), where 0 is corresponds to the faultless observation and the 1 is faulty observation. In the case that there is only one possible fault in the process, the

predicted block is one column vector. In this study, however there are 3 possible faults and discriminant PLS model needs to be built for each of those faults as shown in Figure 2.
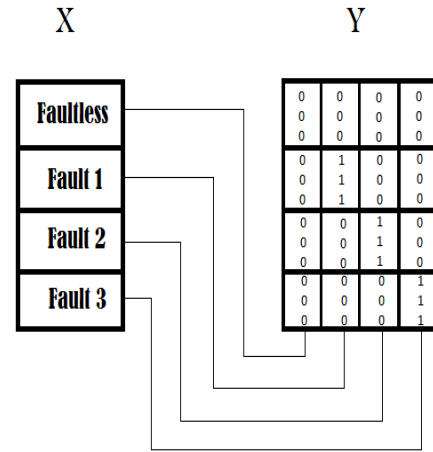

Figure2: Discriminant PLS structure. Each sub block of x is corresponds to one sub mode in predicted block.

One problem about this model is that the predicted outputs are not exactly 0 and 1, and need to be assigned to 0 or 1. One way to do that is to assign the nearest value to the predicted value.

3.2. Wavelet Transform and multi-scale modeling

Wavelet Transform analyses the signal containing multi-frequency content at different resolutions. The family of wavelet basis functions may be represented as:

$$\psi_{su}(t) = \frac{1}{\sqrt{s}} \psi\left(\frac{t-u}{s}\right) \qquad (1)$$

Where, s and u represent the dilation and translation parameters, respectively. ψ (t) is the mother wavelet.

Any signal may be decomposed to its contribution at multiple scales by convolution with the corresponding filters. Using online Wavelet compels the translation parameters to be discretized dyadically as $u = \left(2^m k\right)$ and so the wavelet decomposition downsamples the coefficients at each scale. This approach permits the use of orthonormal wavelets, which approximately decorrelate autocorrelated measurements (Ganesan Das and Venkataraman 2004).

To construct the MSPLS structure, in this study, the Wavelet Transform is used to decompose the measurements to its contribution at multiple scales, and the discriminant PLS is applied to each scale to diagnosis the multiple faults mentioned earlier. Scales at which the current coefficient detects the faults are selected as being relevant at the current time. The signal and covariance at the selected scales are reconstructed by the inverse wavelet transform. The schematic diagram of MSPLS algorithm is shown in Figure 3.
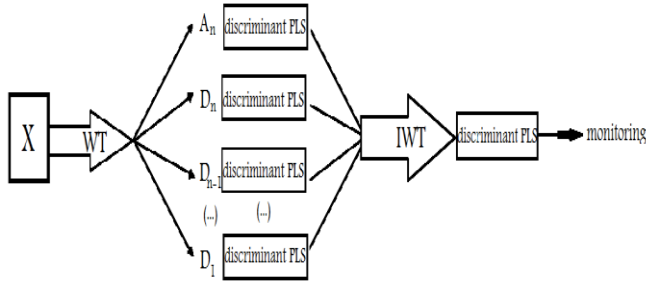
Figure3: Multi-scale PLS for process monitoring

The type of mother wavelet and the optimum decomposition level are two important selective items that should be first determined for the implementation of the online DWT (Lee Lee and Park 2009). Choosing a proper mother Wavelet usually depends on the purpose of its application. As mentioned above, an orthonormal Wavelet can approximately decorrelate the autocorrelated measurements. Based on a systematic approach proposed by Maulud (Maulud Wang, and Romagnoli 2006) for selecting the optimum decomposition level, the 'Haar' wavelet with level three is used in this study. The 'Haar' Wavelet has a simple mother function and is a common Wavelet which is applicable in discrete signal processing and also doesn't make the non-causality problems (Aradhye et al. 2003).

4. MONITORING AND RESULTS

The training data set used in this process contains data which are generated under different conditions. After the modeling with MSPLS, the model is applied to monitoring the TE process. To show the ability of MSPLS to diagnosis multiple faults in process, the testing data used in this study contains three different faults which occur at determinate times. Fault1 start at the sample time 1100 and is connected to the step change in the cooling water. Fault2 start at the sample time 2100 and is a low drift in the reaction kinetics. Fault3 start at the sample time 3000 and is associated with one of the sticking valves.

Table 1: Table1: The percent variance captured by MSPLS and discriminant PLS for the simulated faults.

| | PLS | MSPLS | Scales | | | |
		Global level	A3	D3	D2	D1
Fault1	0.50	0.89	0.81	0.12	0.04	0.08
Fault2	0.53	0.92	0.88	0.09	0.18	0.22
Fault3	0.44	0.74	0.55	0.23	0.28	0.71

To compare the monitoring ability of the MSPLS algorithm and PLS algorithm, the monitoring diagrams of the PLS and MSPLS are shown in the Figure 4 and Figure 5 respectively. In these figures the predicted variable for each fault is drawn. This value is between '0' and '1' where '0' indicates the faultless observation and '1' a faulty observation. Different colors for different faults have been used. The blue sketch is for Fault1, red is for Fault2 and green is for Fault3.
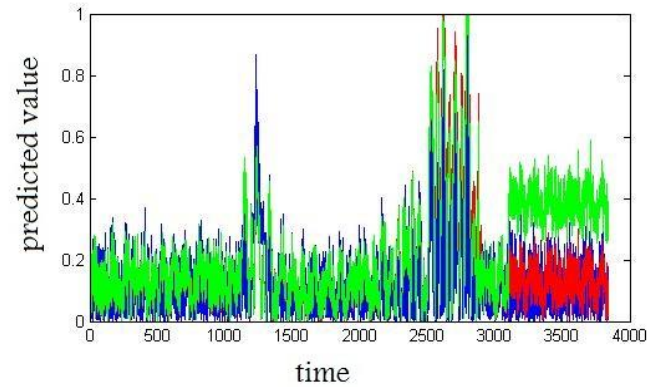

Figure4: Monitoring results of discriminant PLS. The blue graph is relevant to fault1, red is for fault2, and green is corresponds to fault3.

As shown in Figure 5 the MSPLS can detect all of three faults at acceptable time delay, while PLS could not do this adequately. This issue also is emphatic in table 1 where the percent of variance captured by PLS and MSPLS at each scale is shown for each fault, and the final MSPLS algorithm. It is obvious that this value for MSPLS is more than PLS.
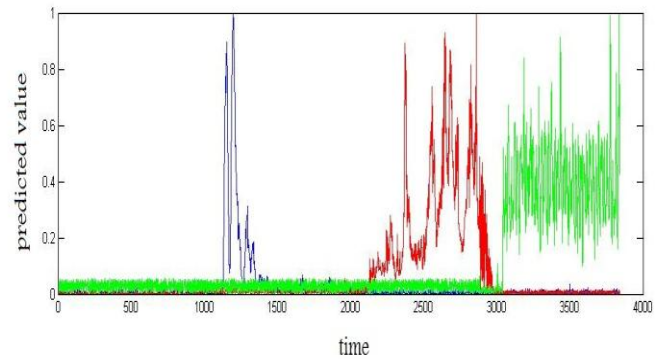

Figure5: Monitoring results of MSPLS. The blue graph is relevant to fault1, red is for fault2, and green is correspond to fault3.

5. CONCLUSIONS

In this paper, a discriminant PLS is used for constructing a new MSPLS algorithm for monitoring on the processes with multiple possible faults. This methodology has the potential of detection and diagnosing each abnormal event, denoting the time and frequency location of each event. In addition to exploiting the useful properties of Wavelet Transform, a MSPLS model which could separately construct a sub model for each possible fault and can detect each of these faults separately. The presented MSPLS algorithm can do this job adequately.

Future works can bring into focus on use of other types of mother wavelet, with the aim of improving detection, diagnosis and identification capabilities. The application of this methodology to other processes is also visualized as an interesting field for future research activities to be carried out.

REFERENCES

Aradhye H. B., Bakshi, B. R., Strauss, R. A. and Davis,J.F. 2003. AIChE J. 49:939–958.

Chaing, L. H., Russel, E. L. and Braatz, R.D. 2001. Fault Detection and Diagnosis in Industrial Systems, Springer, London.

Chaing, L. H., Russell, E. L. and Braatz, L. D. 2000. Fault diagnosis in chemical processes using Fisher discriminant analysis, discriminant partial least squares and principal component analysis. Chemometrics and Intelligent Laboratory Systems, 50:243-352

Downs, J.J. and Vogel, E.F. 1993. A plant-wide industrial process control problem. Computer and chemical engineering 17: 245-255.

Ganesan, R., Das, T. K. and Venkataraman, V. 2004. Wavelet-based multiscale statistical process monitoring, IIE Transactions (2004), 36: 787–806

Geladi P. and Kowalski B.R. 1986. Partial least-squares regression: a tutorial. Anal Chim Acta, 185:1–17.

Lee D.S. and Vanrolleghem P. A. 2003. Monitoring of sequencing batch reactor using adaptive multiblock pricipal component analysis. Biotechnol Bioeng 82(4):489–97.

Lee, D.S., Lee, M.W., Woo, S.H., Kim, Y.J. and Park, J.M. 2006. Nonlinear dynamic partial least squares modeling of a full-scale biological wastewater treatment plant. Process Biochemistry, 41:2050-2057.

Lee, H.W., Lee, M.W. and Park, J.M. 2009. Multi scale extension of PLS for advanced on-line process monitoring, Chemometrics and Intelligent Laboratory Systems, 98:201-212.

Lyman, P. R. 1995. Plant-wide control structure for Teenesse Eastman process. Computers and chemical engineering, 19:321-331.

Maulud A., Wang, D. and Romagnoli, J.A. 2006. Batch/Semi-Batch Process Fault Detection and Diagnosis using Orthogonal Nonlinear Multi-Way PCA Process Control. ADCHEM - International Symposium on Advanced Control of Chemical Processes, Vol. 1: 273-279

Nouwen J., Lindgren F., Hanasen W. K. B., Verharr, H. J. M., et.al. 1997. Classification of environmentally occurring chemicals using structural fragments and PLS discriminant analysis, Environ Sci. Technol, 31:2313-2318.

Piovoso, M. J. and Kosanovich, K. A. 1994. Application of multivariate statistical methods to process monitoring and control design. International Journal of Control, 59:743-765

Raich, A.C. and Cinar, A. 1995. Multivariate statistical methods for monitoring continue processes. Chemometrics and Intelligent Laboratory Systems, 30:37-48

Reis, M. S., Saraiva, P. M. and Bakshi, B. R. 2008. Multiscale statistical process control using wavelet packets (Citations: 1)Aiche Journal - AICHE J , vol. 54, no. 9:2366-2378

Wilson, D. J. H. and Irwin, G. W. 2000. PLS modelling and fault detection on the Tennessee Eastman benchmark, International Journal of Systems Science, 31(11): 1449–1457.

Wold S., Ruhe A., Wold H. and Dunn W.J. 1984. The Collinearity problem in linear regression. The partial least squares approach to generalized inverse. SIAM J Sci Stat Comput, 3:735–43.

MODELING OF AERODYNAMIC FLUTTER ON A NACA 4412 AIRFOIL WIND BLADE

Drishtysingh Ramdenee[a], H. Ibrahim[b], N.Barka[a], A.Ilinca[a]

[a]Wind Energy Research Laboratory, Université du Québec à Rimouski, Canada.G5L3A1
[b]Wind EnergyTechnocentre, Murdochville, Canada. G0E1W0

[a]dreutch@hotmail.com, [b]hibrahim@eolien.qc,ca

ABSTRACT

Study of aeroelastic phenomena on wind turbines (WT) has become a very important issue when it comes to safety and economical considerations as WT tend towards gigantism and flexibility. At the Wind Energy Research Laboratory (WERL), several studies and papers have been produced, all focusing on computational fluid dynamics (CFD) approaches to model and simulate different aeroleastic phenomena. Despite very interesting obtained results; CFD is very costly and difficult to be directly used for control purposes due to consequent computational time. This paper, hence, describes a complementary lumped system approach to CFD to model flutter phenomenon. This model is based on a described Matlab-Simulink model that integrates turbulence characteristics as well as characteristics aerodynamic physics. From this model, we elaborate on flutter Eigen modes and Eigen values in an aim to apply control strategies and relates ANSYS based CFD modeling to the lumped system.

Keywords: flutter, Computational fluid dynamics, lumped system, Matlab-Simulink, ANSYS

1. INTRODUCTION

As wind turbines become increasingly larger and more flexible, concerns are increasing about their ability to sustain both static and dynamic charges. When it comes to static loads, the calculation is fairly easy and IEC norms adequately set the standards for the manufacturing industry. However, when it comes to dynamic loads, the modeling is far more complex as we need to include the rotational movement, the bending, the wind speed, turbulence and other complex fluid-structure interactions that can generate divergence, dynamic stall or flutter. The main aim of modeling these phenomena is to be able to apply mitigation actions to avoid them as they are extremely damageable for wind turbines. In this article, we will model one of the most destructive aeroelastic phenomena - flutter via Matlab/Simulink and compare our results with ANSYS – CFX based CFD generated results. The aim of the Simulink based modeling is to set up an integrated model that can more easily be incorporated in a control strategy to limit operation in critical vibration conditions. Aerodynamic flutter is a dynamic aeroelastic phenomenon characterised by blade response with respect to changes of the fluid flow such as external atmospheric disturbances and gusts. Flutter is a very dangerous phenomenon resulting from an interaction between elastic, inertial and aerodynamic forces. This takes place when the structural damping is not sufficient to damp the vibration movements introduced by the aerodynamic effects. Flutter can take place for any object in an intense fluid flow and condition of positive retroaction. In other words, the vibratory movement of the object increases an aerodynamic solicitation, which, in turn, amplifies the structural vibration. When the energy developed during the excitation period is larger than the normal system dumping, the vibration level will increase leading to flutter. The latter is characterized by the superposition of two structural modes – the pitch and plunge movement. When wind speed increases, the frequency of these vibration modes coalesce to create the resonance of flutter.

2. NOMENCLATURE

α   Angle of attack   θ   Plunge angle
w_g Centre of gravity   M   Aerodynamic moment
ψ_{lo} Longitudinal Speed Turbulence Spectrum
ψ_{la} Lateral Speed Turbulence Spectrum
ψ_v Vertical Speed Turbulence Spectrum

3. FLUTTER PHENOMENON

As previously mentioned, flutter is caused by the superposition of two structural modes – pitch and plunge. The pitch mode is described by a rotational movement about the elastic centre of the airfoil whereas the plunge mode is a vertical up and down motion at the blade tip. Theodorsen [1-3] developed a method to analyze aeroelastic stability. The technique is described by equations (1) and (2). α is the angle of attack (AoA), α_0 is the static AoA, C(k) is the Theodorsen complex valued function, h the plunge height, L is the lift vector positioned at 0.25 of the chord length, M is the pitching moment about the elastic axis, U is the free velocity, ω is the angular velocity and a, b, d1 and d2 are geometrical quantities as shown in figure 1.
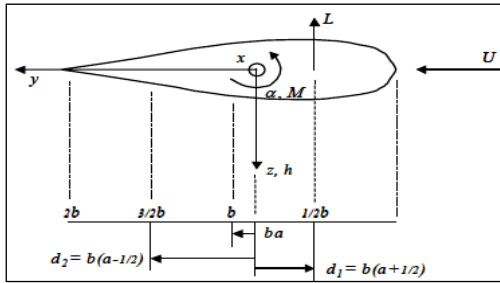

Figure 1: Model defining parameters

$$L = 2\pi\rho U^2 b \left\{ \frac{i\omega C(k)h_0}{U} + C(k)\alpha_0 \right.$$
$$+ [1 + C(k)(1-2a)]\frac{i\omega b\alpha_0}{2U}$$
$$- \frac{\omega^2 b h_0}{2U^2}$$
$$\left. + \frac{\omega^2 b^2 a\alpha_0}{2U^2} \right\} \qquad (1)$$

$$M$$
$$= 2\pi\rho U^2 b \left\{ d_1 \left[\frac{i\omega C(k)h_0}{U} + C(k)\alpha_0 \right. \right.$$
$$\left. + [1 + C(k)(1-2a)]\frac{i\omega b\alpha_0}{2U} \right] + d_2 \frac{i\omega b\alpha_0}{2U} - \frac{\omega^2 b^2 a}{2U^2} h_0$$
$$+ \left(\frac{1}{8} \right.$$
$$\left. \left. + a^2 \right) \frac{\omega^2 b^3 \alpha_0}{2U^2} \right\} \qquad (2)$$

The Theodorsen equation can be rewritten in a form that can be entered and analyzed in Matlab Simulink as follows:

$$L = 2\pi\rho U^2 b \left\{ \frac{C(k)}{U}\dot{h} + C(k) \right.$$
$$\alpha +[1 + C(k)(1-2a)]]\frac{b}{2U}\dot{\alpha}$$
$$\left. + \frac{b}{2U^2}\ddot{h} - \frac{b^2 a}{2U^2}\ddot{\alpha} \right\} \qquad (3)$$

$$M = 2\pi\rho U^2 b \left\{ d_1 \left[\frac{C(k)\dot{h}}{U} + C(k) \right. \right.$$
$$\left. \alpha + \left[1 + C(k)(1-2a)\frac{b}{2U}\dot{\alpha} \right] \right]$$
$$+ d_2 \frac{b}{2U}\dot{\alpha} + \frac{ab^2}{2U^2}\ddot{h} - (\frac{1}{8}$$
$$\left. + a^2)\frac{b^3 \ddot{\alpha}}{2U^2} \right\} \qquad (4)$$

4. FLUTTER MOUVEMENT

Flutter can be triggered by a rotation of the profile (t=0 seconds in figure 2). The increase in the force adds to the lift such that the profile tend to undertake a vertical upward movement. Simultaneously, the torsion rigidity of the structure returns the profile to the zero pitch position (t=T/4 in figure 2). The flexion rigidity of the structure tries to return the profile to its neutral position but the profile now adopts a negative angle of attack (t=T/2 in figure 2). Once again, the increase in the aerodynamic force imposes a vertical downwards movement and the torsion rigidity returns the profile to zero angle of attack position. The cycle ends when the profile returns to a neutral position with a positive angle of attack. With time, the vertical movement tends to get damped whereas the rotational movement diverges. If the movement is left to repeat, the rotation induced forces will lead to failure of the structure.
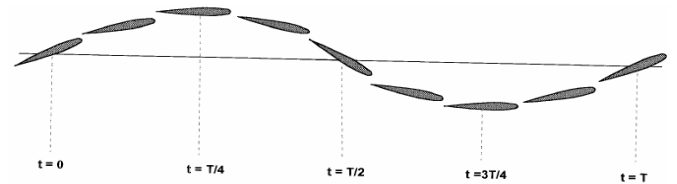

Figure 2: Illustration of the flutter movement

In order to understand this complex phenomenon, we describe flutter as follows: Aerodynamic forces excite the mass – spring system illustrated in figure 3. The plunge spring represents the flexion rigidity of the structure whereas the rotation spring represents the rotation rigidity.
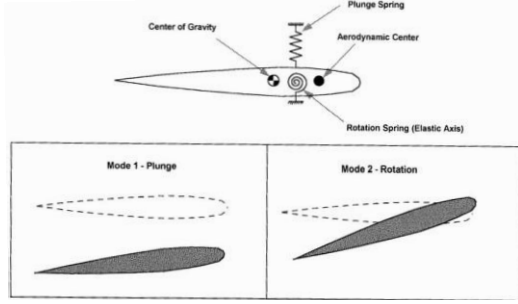
Figure 3: Illustration of both pitch and plunge

5. FLUTTER EQUATIONS

Initially, it is important to find a relationship between the generalized coordinates and the angle of attack of the model. This will be essential in the computation of the aerodynamic forces. From [4], the relationship between the angle of attack and the coordinates can be written as:

$$\alpha(x,y,t)$$
$$= \theta_T + \theta(t) + \frac{\dot{h}(t)}{U_0} + \frac{l(x)\dot{\theta}(t)}{U_0}$$
$$- \frac{w_g(x,y,t)}{U_0} \tag{5}$$

From these energy equations, the Lagrangian equations are constructed for the mechanical system. The first one corresponds to the vertical displacement z and the other is subject to the angle of attack α.
Hence:
$$J_0\ddot{\alpha} + md\cos(\alpha)\ddot{z} + c(\alpha - \alpha_0)$$
$$= M_0 \tag{6}$$
and
$$m\ddot{z} + md\cos(\alpha)\ddot{\alpha} - m\sin(\alpha)\dot{\alpha}^2$$
$$+ kz \tag{7}$$

In order to enable numerical solving of these equations, we need to express F_z and M_o as polynomials of α. Moreover; $F_z(\alpha) = \frac{1}{2}\rho SV^2 C_z(\alpha)$ and $M_o(\alpha) = \frac{1}{2}\rho LSV^2 C_{m0}(\alpha)$ for S being the surface of the blade, C_z, the lift coefficient, C_{m0} being the pitch coefficient, F_z being the lift, M_o, the pitch moment. C_z and C_m values are extracted from NACA 4412. Degree 3 interpolations for C_z and C_m with respect to the AoA are given below:

$$C_z$$
$$= -0.0000983\,\alpha^3 - 0.0003562\alpha^2 + 0.1312\alpha$$
$$+ 0.4162 \tag{8}$$
$$C_{m0}$$
$$= -0.00006375\alpha^3 + 0.00149\alpha^2 - 0.001185\,\alpha$$
$$- 0.9312 \tag{9}$$

6. MATHLAB-SIMULINK AND ANSYS-CFX TOOLS

Reference [5] describes the Matlab included tool Simulink as an environment for multi-domain simulation and Model-Based Design for dynamic and embedded systems. It provides an interactive graphical environment and a customizable set of block libraries that let you design, simulate, implement, and test a variety of time-varying systems. For the flutter modeling project the aerospace blockset of Simulink has been used. The Aerospace Toolbox product provides tools like reference standards, environment models, and aerospace analysis pre-programmed tools as well as aerodynamic coefficient importing options. Among others, the wind library has been used to calculate wind shears and Dryden and Von Karman turbulence. The Von Karman Wind Turbulence model uses the Von Karman spectral representation to add turbulence to the aerospace model through pre-established filters. Turbulence is represented in this blockset as a stochastic process defined by velocity spectra. For a blade in an airspeed V, through a frozen turbulence field, with a spatial frequency of Ω radians per meter, the circular frequency ω is calculated by multiplying V by Ω. For the longitudinal speed, the turbulence spectrum is defined as follows:

$$\psi_{lo}$$
$$= \frac{\sigma^2_\omega}{VL_\omega} \cdot \frac{0.8(\frac{\pi L_\omega}{4b})^{0.3}}{1 + \left(\frac{4b\omega}{\pi V}\right)^2} \tag{10}$$

where L_ω represents the turbulence scale length and σ is the turbulence intensity. The corresponding transfer function used in Simulink is expressed as:

$$\psi_{lo}$$
$$= \frac{\sigma_u\sqrt{\frac{2}{\pi}\frac{L_v}{V}}(1 + 0.25\frac{L_v}{V}s)}{1 + 1.357\frac{L_v}{V}s + 0.1987\left(\frac{L_v}{V}s\right)^2 s^2} \tag{11}$$

For the lateral speed, the turbulence spectrum is defined as:

$$\psi_{la} = \frac{\mp\left(\frac{\omega}{V}\right)^2}{1 + \left(\frac{3b\omega}{\pi V}\right)^2} \cdot \varphi_v(\omega) \tag{12}$$

and the corresponding transfer function can be expressed as :

$$\psi_{la} = \frac{\mp \left(\frac{s}{V}\right)^1}{1 + \left(\frac{3b}{\pi V}s\right)^1} \cdot H_v(s) \qquad (13)$$

Finally, the vertical turbulence spectrum is expressed as follows:

$$\psi_v = \frac{\mp \left(\frac{\omega}{V}\right)^2}{1 + \left(\frac{4b\omega}{\pi V}\right)^2} \cdot \varphi_\omega(\omega) \qquad (14)$$

and the corresponding transfer function is expressed as follows:

$$\psi_v = \frac{\mp \left(\frac{s}{V}\right)^1}{1 + \left(\frac{4b}{\pi V}s\right)^1} \cdot H_\omega(s) \qquad (15)$$

The Aerodynamic Forces and Moments block computes the aerodynamic forces and moments about the center of gravity. The net rotation from body to wind axes is expressed as:

$$C_{\omega \leftarrow b} = \begin{bmatrix} \cos(\alpha)\cos(\beta) & \sin(\beta) & \sin(\alpha)\cos(\beta) \\ -\cos(\alpha)\sin(\beta) & \cos(\beta) & -\sin(\alpha)\sin(\beta) \\ -\sin(\alpha) & 0 & \cos(\alpha) \end{bmatrix} \qquad (16)$$

On the other hand, the fluid structure interaction to model aerodynamic flutter was made using ANSYS multi domain (MFX). As we mentioned in the abstract of this paper, the drawback of the ANSYS model is that it is very time and memory consuming. However, it provides a very good option to compare and validate simplified model results and understand the intrinsic theories of flutter modelling. On one hand, the aerodynamics of the application is modelled using the fluid module CFX and on the other side, the dynamic structural part is modelled using ANSYS structural module. An iterative exchange of data between the two modules to simulate the flutter phenomenon is done using the Workbench interface. Details of this modelling are available in [6].

7. EXPERIMENT FOR VALIDATION

Reference [7] makes a literature review of work performed on divergence and flutter. It is clear from there that most work has been performed on the control and mitigation of such phenomena without emphasizing on the modelling. This is mainly because the latter is very complex and the aim is primarily to avoid these phenomena. The aims of the studies conducted in by Heeg [8] were to: 1) to find the divergence or flutter dynamic pressure; 2) to examine the modal characteristics of non-critical modes, both in subcritical and at the divergence condition; 3) to examine the eigenvector behaviour. The test was conducted by setting as close as possible to zero the rigid angle of attack, $\square_0$, for a zero airspeed. The divergence/flutter dynamic pressure was determined by gradually increasing the velocity and measuring the system response until it became unstable. The results of [8] will be compared with our aerospace blockset-based obtained model.

8. RESULTS

We will first present the results obtained by modeling AoA for configuration # 2 in [8] for an initial AoA of 0°. As soon as divergence is triggered, within 1 second the blade oscillates in a very spectacular and dangerous manner. This happens at a dynamic pressure of 5,59 lb/pi^2 (268 N/m^2). Configuration #2 uses, in the airfoil: 20 elements, unity as the normalized element size and unity as the normalized airfoil length. Similarly, the number of elements in the wake is 360 and the corresponding normalized element size is unity and the normalized wake length is equal to 2. The result obtained in [8] is illustrated in figure 2:
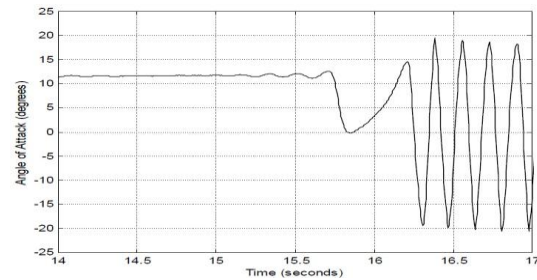

Figure 4: Flutter response- an excerpt from [8]

We can notice that at the beginning there is a non-established instability followed by a recurrent oscillation. The peak to peak distance corresponds to around 2.5 seconds, that is, a frequency of 0.4 Hz. The oscillation can be defined approximately by amplitude of $0^0 \pm 17^0$. The same modelling was performed using the Simulink model and the result for the AoA variation and the plunge displacement is shown below:
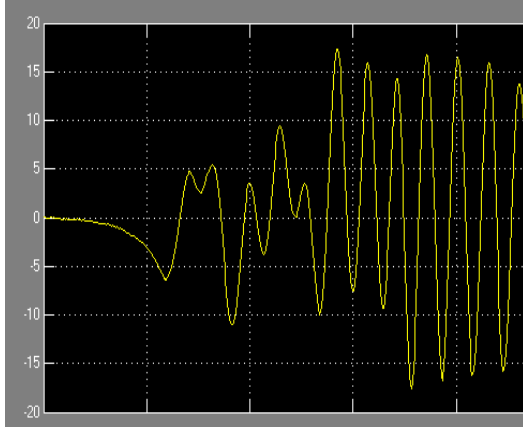
Figure 5: Flutter response obtained from Matlab Aerospace blockset

We can note that for the AoA variation, the aerospace blockset based model provides very similar results with J. Heeg results. The amplitude is, also, around $0^0 \pm 17^0$ and the frequency is 0.45 Hz. Furthermore, we notice that the profile of the variation is very similar. We can conclude that the aerospace model does represent the flutter in a proper manner. It is important to note that this is a special type of flutter. The frequency of the beat is zero and, hence, represents divergence of "zero frequency flutter". Using Simulink, we will vary the angular velocity of the blade until the eigenmode tends to a negative damping coefficient. The damping coefficient, ξ is obtained as: $\xi = \frac{c}{2m\omega}$, ω is measured as the Laplace integral in Simulink, c is the viscous damping and $\omega = \sqrt{\frac{k}{m}}$.

Table 1 below gives a summary of the obtained results of damping coefficient against rotor speed which are plotted in figure 4.

Table 1: Damping coefficient and frequency mode

Rotor Speed (Hz)	Damping Coefficient	Frequency of flutter mode (Hz)
0.1	0.0082	9.4
0.3	0.0731	8.721
0.45	0.1023	8.2532
0.6	0.2013	7.5324
0.65	0.15343	7.01325
0.7	0.08931	6.4351
0.75	-0.09321	6.33
0.8	-0.099315	5.5835

We can note that as the rotation speed increases, the damping becomes negative such that the aerodynamic instability which contributes to an oscillation of the profile is amplified. We also notice that the frequency reduces and becomes nearer to the natural frequency of the system. This explains the reason for which flutter is usually very similar to resonance as it occurs due to a coalescing of dynamic modes close to the natural vibrating mode of the system.
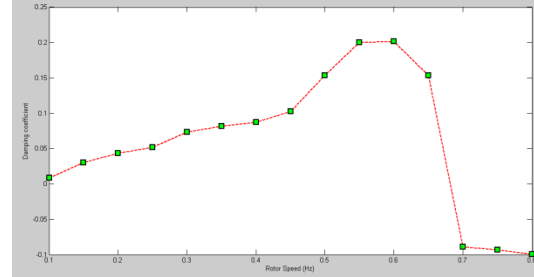

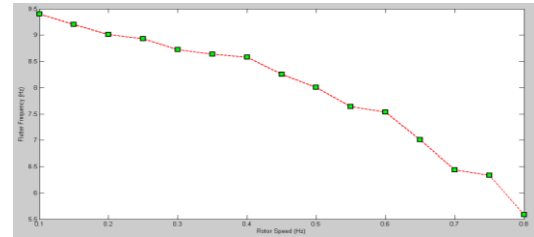Figure 6: Damping coefficient against rotational speed

Figure 7: Flutter frequency against rotor speed

We, now present the results obtained for the same case study using ANSYS – CFX. We notice that the frequency of the movement using Matlab is 6.5 Hz that using the ANSYS-CFX model, 6.325 Hz and that obtained from Jennifer Heeg experiments 7.1Hz. Furthermore, the amplitudes of vibration are very close as well as the trend of the oscillations. For points noted 1, 2 and 3 on the flutter illustration, we exemplify the relevant flow over the profile. The maximum air speed at moment noted 1 is 26.95 m/s. we note such a velocity difference over the airfoil that an anticlockwise moment will be created which will cause an increase in the angle of attack. Since the velocity, hence, pressure difference, is very large, we note from the flutter curve, that we have an overshoot. The velocity profile at moment 2, i.e., at 1.88822 s shows a similar velocity disparity, but of lower intensity. This is visible as a reduction in the gradient of the flutter curve as the moment on the airfoil is reduced. Finally at moment 3, we note that the velocity profile is, more or less, symmetric over the

airfoil such that the moment is momentarily zero. This corresponds to a maximum stationary point on the flutter curve. After this point, the velocity disparity will change position such that angle of attack will again increase and the flutter oscillation trend maintained, but in oposite direction. This cyclic condition repeats and intensifies as we have previously proved that the damping coefficient tends to a negative value.
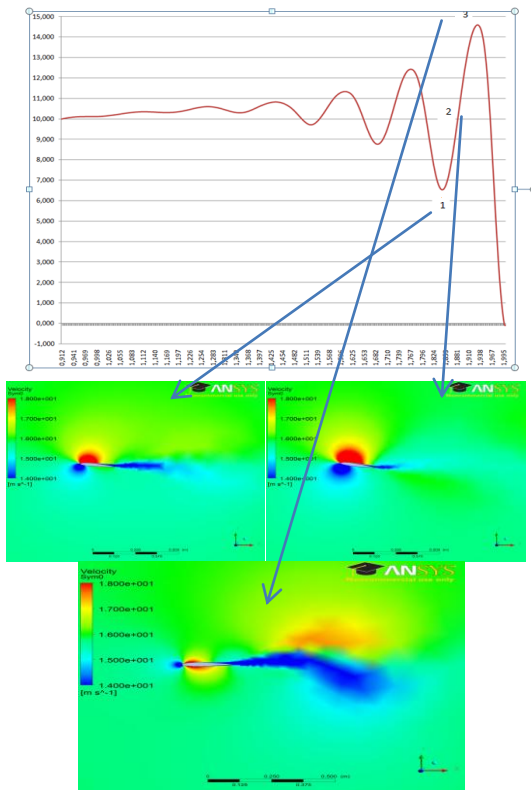

Figure 7: Flutter simulation with ANSYS-CFX at 1) 1.8449 s, 2) 1.88822 s and 1.93154s

9. DISCUSSION AND FUTURE OF THE PROJECT

In this article, we have detailed the aims and steps of modeling flutter using Simulink. The obtained results are very close to those obtained by Heeg [8]. The model furthermore enables monitoring of the damping with respect to rotational speed. Coupled with the eigenvalues and eigen frequencies analysis, the model enables a satisfactory representation of the phenomenon and a very conducive form for incorporation in a control strategy. However, this model needs to be further tried and refined to include other aspects such as rapid change in the wind speed (gusts), the flexibility of the model to adapt to different airfoils, etc. In future studies, the model will be used on different airfoils and

for various wind regimes. Furthermore, additional variables will be entered in the model such as thermal variability. Once, the model is enough optimized to approach experimental results, control strategies will be applied to damp the vibrations. In an initial phase, classic control strategies such as the "Proportional Integration Differentiator Filter" models, cascade models, internal models, and Smith predictive models will be used. In a second phase, if required, a neural network control will be tried on the model.

10. CONCLUSION

In this article we modeled the very complex and dangerous flutter phenomenon. In an initial phase, we described the phenomenon and the equations characterizing it analytically. This was done by emphasizing on the required fluid-structure interaction. The article then ponders on the Matlab and ANSYS models used to simulate the phenomena as well as the experimental work used to validate our results. Both ANSYS and Matlab have given very interesting results. However, it can be noted that Matlab can only propose the aerodynamics coefficient curves while ANSYS can provide both the aerodynamic characteristics of the response and visualisation of the different flow fields along the airf oil at all time. It must be emphasized that the use of one model or the other must be based on several criteria. ANSYS requires very large computational capacity whereas the Matlab model is very less demanding. For academic and research needs, the ANSYS model proves to be very interesting as the generated flow fields help to understand the intrinsic phenomena that causes flutter. On the other hand, the Matlab model is better suited for industrial applications, as the model can be directly integrated in a control strategy and the flutter phenomenon avoided.

References

[1] Theodorsen, T., General theory of aerodynamic instability and the mechanism of flutter, NACA Report 496, 1935.

[2].Fung, Y. C., An Introduction to the Theory of Aeroelasticity. Dover Publications Inc.: New York, 1969; 210-216.

[3].Dowell, E. E. (Editor), A Modern Course in Aeroelasticity. Kluwer Academic Publishers: Dordrecht, 1995; 217-227.

[4] A.G Chervonenko et al, '' Effect of attack angle on the nonstationary aerodynamic characteristics and flutter resistance of a grid of bent vibrating compressor blades'' UDC 621.515/62-752

[5] mathworks.com/products/simulink

[6] D.Ramdenee et al. "Numerical Simulation of the Divergence Phenomenon on a NACA 4412 Airfoil-Part 2" Canadian Society of Mechanical Engineering Conference, University of Victoria, British Columbia, June 2010.

[7] D.Ramdenee et al. "Numerical Simulation of the Divergence Phenomenon on a NACA 4412 Airfoil-Part 1" Canadian Society of Mechanical Engineering Conference, University of Victoria, British Columbia, June 2010.

[8] J. Heeg "Dynamic Investigation of Static Divergence: Analysis and Testing" Langley Research Center, Hampton, Virginia

[9] D.Ramdenee et al. "Numerical Simulation of the StallPhenomenon on an S 809 Airfoil" CFD Society , University of West Ontario,London, Ontario. 2010

[10] D.Ramdenee et al. "Numerical Simulation of the Divergence phenomenon" CFD Society , Montreal . 2011

CONTROL ALLOCATION STRATEGIES
FOR AN OVERACTUATED ELECTRIC VEHICLE

J.-Y. Dieulot

LAGIS, Polytech-Lille/IAAL, Cité Scientifique, 59650 Villeneuve d'Ascq, France phone: 33-3-28-76-74-95
jean-yves.dieulot@polytech-lille.fr

ABSTRACT

Overactuated Intelligent Autonomous Electric Vehicles may possess up to 6 degrees of freedom, such as two steering devices and four independent traction wheels, which serve to control yaw, lateral and longitudinal velocities. It is shown that the coupled nonlinear control problem can be set as an optimal control strategy, which consists of a distribution of contact forces or steering angles according to yaw and lateral speed control. Then, the motor torques of the DC drives are computed in order to ensure dominant wheel rolling operation for a 4 X 4 motion or to ensure that all wheels are in the rolling state when using steering angles

Keywords: Hybrid Vehicle, Nonlinear Control, Control Allocation, Contract forces, Vehicle Steering

1. INTRODUCTION

Intelligent autonomous vehicles (IAVs) is a class of intelligent transportation systems which are operated without a human driver. As an example, these vehicles can be used in harbor environments for goods low-speed transportation, where they ensure safe, reconfigurable and, while electric, low emission traffic (Djeziri et al., 2009). Often, such vehicles embed multi-actuated traction and steering systems, which allows to consider redundancy in control, design of different scenarios to run the vehicle on a segment of the road, and control/operating modes reconfigurable solutions. A good knowledge of the kinematics and dynamics is important to design a robust MIMO controller of such intelligent autonomous vehicles, in order to compensate for external perturbations and local nonlinearities (Merzouki et al., 2007, Merzouki et al., 2009).

Since the 1980s, various active chassis vehicle control approaches have been investigated, some of which can be transposed to IAVs. In particular, research into vehicle dynamics control (VDC) or vehicle stability control systems has become very active (e.g. Furukawa and Abe, 1997). Variables to be controlled are typically longitudinal velocity, lateral velocity, and yaw rate, while the actuation generally includes individual wheel drive and steering devices. Recently, control allocation approaches have been introduced into vehicle control systems to take advantage of actuator redundancy for improving system performance and achieving reconfigurable control solutions e.g. Wang and Logoria (2009), Tjonnas and Johansen (2010). A control allocation approach is generally used when different combinations of effector commands can produce the same result and when the number of effectors available exceeds the number of states being controlled. In these suggested control allocation (CA) schemes, the generalized forces are allocated to longitudinal and/or lateral tire forces. Different CA schemes have been proposed, either by optimal control, nonlinear control or fuzzy logic e.g. Raffo et al (2009), Feiqiang et al. (2009), Partouche et al., (2007), Tjønnås and Johansen (2010) and very few on IAVs.

Vehicle motion is governed by forces induced by each tire interacting with the road, and these forces mainly depend on the slip velocity/slip angle and tire-friction coefficient Canudas de Wit et al.(2003), Merzouki et al. (2007), Bakker et al. (1989). While allocating the control effort to tires, it is important to take these factors into account, to ensure that the tire can actually yield the desired forces.

This study focuses mainly on the control of an over-actuated autonomous electric vehicle, using simplified dynamic and kinematic models (Djeziri et al., 2009). The developed dynamics concerns the longitudinal, lateral, vertical and yaw of the chassis, including the dynamics of the electromechanical systems and the wheel-ground interactions. Two kinds of control strategies are proposed, whether the automated steering systems are used or not.
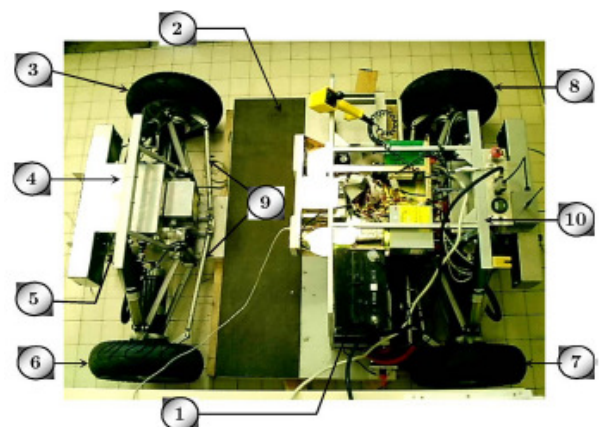
2. VEHICLE MODELLING

Figure 1: Robucar vehicle and actuators

The autonomous overactuated electric vehicle in Fig. 1, named Robucar, owns four actuated traction wheels and two actuated steering systems with a total of 6 degrees of freedom, thus allowing to handle actuator or sensor

defaults. Figure 1 shows the following composition : 1) 12-V 60-Ah sealed batteries; 2) a honeycomb chassis; 3) a front right wheel; 4) a front control cabinet; 5) a front steering electrical jack; 6) a front left wheel; 7) a rear left wheel; 8) a rear right wheel; 9) a rear steering electrical jack; and 10) a rear control cabinet. All technical details are supplied in previous papers (see e.g. Djeziri et al., 2009).

Basically, the overall model can be split up into 4 parts: the kinematic model, the behaviour of the chassis along the trajectory, the electromechanical model which links the motor voltage to the wheel speed and contact forces, and the modelling of the wheel-tire-road contact itself.
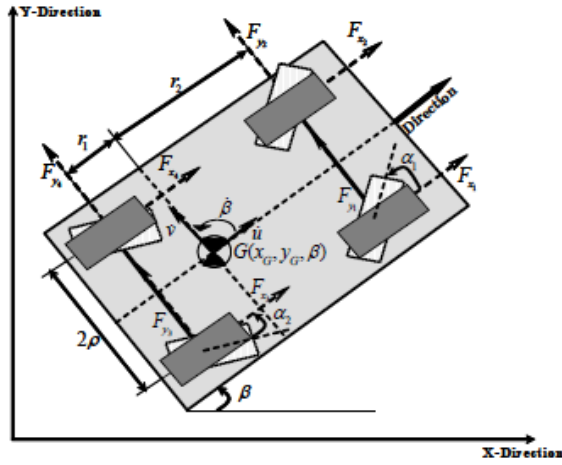
2.1. Chassis dynamics

Figure 2 Contact forces and kinematic parameters

Owing to the low vehicle velocity (< 20 km/h), all centrifugal forces are neglected in the sequel and only the longitudinal, lateral and yaw dynamics are considered which yields:

Longitudinal dynamics :
$$m\ddot{u} = (F_{x1} + F_{x2})\cos\alpha_1 + (F_{x3} + F_{x4})\cos(\alpha_2)$$
$$- (F_{y1} + F_{y2})\sin(\alpha_1) + (F_{y3} + F_{y4})\sin(\alpha_2) \quad (1)$$

Lateral dynamics :
$$m\ddot{v} = (F_{x1} + F_{x2})\sin\alpha_1 - (F_{x3} + F_{x4})\sin(\alpha_2)$$
$$+ (F_{y1} + F_{y2})\cos(\alpha_1) + (F_{y3} + F_{y4})\cos(\alpha_2) \quad (2)$$

Yaw dynamics :
$$I\ddot{\beta} = \left[(F_{x1} + F_{x2})\sin\alpha_1 + (F_{y1} + F_{y2})\cos(\alpha_1)\right]x_2$$
$$+ \left[(F_{x3} + F_{x4})\sin(\alpha_2) - (F_{y3} + F_{y4})\cos(\alpha_2)\right]x_1 \quad (3)$$
$$+ \frac{\rho}{2}\left[\begin{array}{l}(F_{x1} - F_{x2})\cos\alpha_1 + (F_{x3} - F_{x4})\cos(\alpha_2) \\ -(F_{y1} - F_{y2})\sin(\alpha_1) + (F_{y3} - F_{y4})\sin(\alpha_2)\end{array}\right]$$

where $u, v, \beta, \alpha_1, \alpha_2$ are respectively the center of mass longitudinal, lateral speeds, yaw angle, back and rear

steering angles, F_{xi}, F_{yi} the longitudinal and lateral contact (tire) forces, I is the moment of inertia of the C.O.G with respect to the vertical axis, ρ is the axle track and m is the vehicle mass (Figure 2).

2.2. Kinematics

The heading, velocities and position of the vehicle in the absolute frame can be obtained using :

$$\dot{x}_G = \dot{u}\cos\beta - \dot{v}\sin\beta$$
$$\dot{y}_G = \dot{u}\sin\beta + \dot{v}\cos\beta$$

and the positions and yaw angle can be obtained by integration.

2.3. Electromechanical Model

The electromechanical model of a quarter-Robucar can be represented as a DC drive monitoring a two-mass-spring damper system which figures out the mechanical flexibilities of the transmission system, i.e., neglecting the current loop (with fast dynamics):

$$J_{ej}\ddot{\theta}_{ej} = -f_{ej}\dot{\theta}_{ej} - K_j(\theta_{ej} - N_j\theta_{sj}) + \Gamma_j$$
$$J_{sj}\ddot{\theta}_{sj} = -f_{sj}\dot{\theta}_{sj} + K_j N_j(\theta_{ej} - N_j\theta_{sj}) - \upsilon F_{xj} \quad (4)$$

where θ_{ej}, θ_{sj} are the motor and wheel angle, Γ_j is the electrical motor torque, $J_{ej}, J_{sj}, f_{ej}, f_{sj}, K_j$ are the corresponding modal inertias, frictions and elasticity constants of the motor-wheel system, υ is the wheel radius, N_j is the speed ratio.

2.4. Modelling road-tire contact

As pointed out before, there are many ways to describe the tire-road modelling. Basically, one can introduce the slip velocity $\dot{x}_{sj} = \dot{u} - \upsilon\dot{\theta}_{sj}$, where υ is the wheel radius. The longitudinal effort can be estimated using such different models as the LuGre model (Canudas de Wit et al., 2003), the well-known Pacejka model (Bakkeer et al., 1989) or other models embedding a combination of different phenomena (Merzouki et al., 2007). Let the longitudinal forces be described by the Pacejka model for which:

$$G = (\upsilon\dot{\theta}_{sj} - \dot{x}_{sj})/\max(\upsilon\dot{\theta}_{sj}, \dot{x}_{sj}),$$
$$F_{xj} = f(G)$$
$$= D\sin\{C\arctan[Bx - E(Bx - Arc\tan(Bx))]\} + S_v \quad (5)$$

where F_{xj} is a so-called « canonical » curve, and the parameters A, B, C, D, E, S_v depend on external conditions. (Fig. 3)

2.5. Available Measurements & state reconstruction

To improve the efficiency of intelligent and autonomous vehicle, the following information should be determined and available in real time (Djeziri et al., 2009):

1. Position localization of the vehicle;
2. Kinematic and dynamic states of the vehicle;
3. Evolutive state of the environment surrounding the vehicle;
4. State of the traction and steering controls in presence of obstacles or referred targets;
5. Communication between vehicle to vehicle or vehicle to infrastructure;
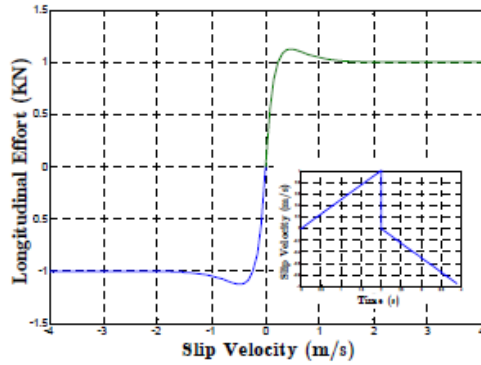6. Access to the coordinates of the trajectory.

Figure 3. Canonical contact forces curve

On the *Robucar* system, a GPS system ensures the localization of the vehicle whereas optical encoders and accelerometers allow to yield wheel and motors position, speed and accelerations. Some variables as the contact forces and the real transmitted torque have to be estimated using a state observer based on second-order sliding mode (Merzouki et al., 2009). Nevertheless, every variable in the previous and remaining equations can be either measured, or estimated (Djeziri et al., 2009).

3. CONTROL STRATEGIES

Generally speaking, monitoring a fully automated vehicle which purposes are goods transportation should meet the following requirements :

- track a geometric trajectory $g(x_c, y_c) = 0$ with a required tolerance
- use the vehicle safely (e.g. beware of tire-road contact, obstacle such as steps); this includes yaw and lateral velocity control
- embed all electromechanical constraints
- given all constraints, operate at maximum longitudinal velocity.

Details about the trajectory and contact forces are given in Fig. 4. Monitoring an autonomous electric vehicle can thus be viewed as an optimal control under constraints. This paper presents preliminary results where driving along a straight road $y_c = 0$ is only considered. The optimal control along the road can thus be formulated as follows:

$$\max x(t)$$
$$s.t. \ \lim_{t \to \infty} \beta = 0, \lim_{t \to \infty} y = 0$$
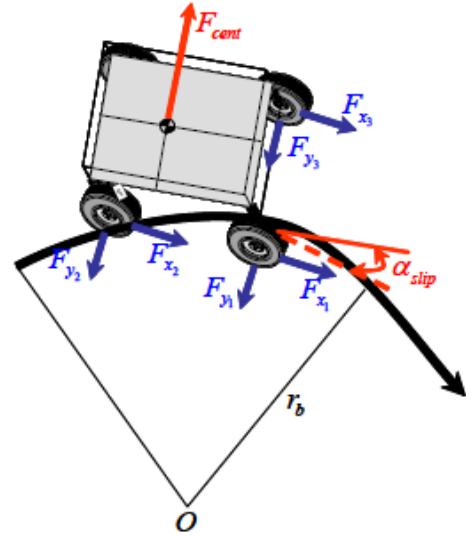$$|\dot{x}(t)| \le V_{max}, |u(t)| \le U_{max}$$

Figure 4. Motion and contact forces

3.1. 4 X 4 actuation with dominant wheel

The basic idea of the control scheme is to use the traction wheels control, by letting a combination of contact forces ensure an appropriate control of the yaw and lateral speed (see Figures 6-7) . The remaining degrees of freedom are available to ensure that at least one wheel operates in the rolling stage.

3.1.1. Hierarchical control

From (1-2), one has

$$m\ddot{u} \cos\left(\frac{\alpha_2 - \alpha_1}{2}\right) + m\ddot{v} \sin\left(\frac{\alpha_2 - \alpha_1}{2}\right) =$$
$$(F_{x1} + F_{x2}) \cos\alpha_1 \cos\left(\frac{\alpha_2 - \alpha_1}{2}\right) +$$
$$(F_{x3} + F_{x4}) \cos\left(\frac{\alpha_2 - \alpha_1}{2}\right) \cos(\alpha_2)$$
$$+ (F_{x1} + F_{x2}) \sin\alpha_1 \sin\left(\frac{\alpha_2 - \alpha_1}{2}\right)$$
$$- (F_{x3} + F_{x4}) \sin\left(\frac{\alpha_2 - \alpha_1}{2}\right) \sin(\alpha_2)$$

and

$$m\ddot{u}\cos\left(\frac{\alpha_2-\alpha_1}{2}\right)+m\ddot{v}\sin\left(\frac{\alpha_1-\alpha_2}{2}\right)=$$
$$\left(F_{x1}+F_{x2}\right)\cos\left(\frac{\alpha_2+\alpha_1}{2}\right)+\left(F_{x3}+F_{x4}\right)\cos\left(\frac{\alpha_2+\alpha_1}{2}\right) \quad (6)$$

One can draw the following chassis control strategy:

- yaw control using $\Delta F_x = F_{x1} - F_{x2} = F_{x3} - F_{x4}$ as input
- lateral speed control using
  $\sum F_x = F_{x1} + F_{x2} + F_{x3} + F_{x4}$ as input
- maximum speed control $Vx = \dot{u}$
  using $F_x^+ = \sup\{F_{xi}\}$ as input

When $F_x^+ = \sup\{F_{xi}\}$ is known, the other forces can be allocated and, then, every motor torque can be derived.

- Lateral speed control

Using lateral and longitudinal acceleration feedbacks, one yields $\sum F_x$

$$\left(\sum F_x\right)\cos\left(\frac{\alpha_2+\alpha_1}{2}\right) = m\ddot{u}\cos\left(\frac{\alpha_2-\alpha_1}{2}\right)$$
$$+k_v^v m\dot{v}\sin\left(\frac{\alpha_1-\alpha_2}{2}\right)+k_v^x m v\sin\left(\frac{\alpha_1-\alpha_2}{2}\right) \quad (7)$$

which ensures the closed-loop lateral speed dynamics:
$\ddot{v}+k_v^v\dot{v}+k_v^x v=0$ and $\lim_{t\to\infty} v \to 0$, where k_v^v, k_v^x are adequately chosen.

- Yaw control

In equation (3), the first group of terms is related to the lateral behavior, which is enforced to zero by equation (7). Hence, *only the slip phenomenon is accounted for,* which simplifies equation (3) to:

$$I\ddot{\beta}=\frac{\rho}{2}\Delta F_x\left[\cos\alpha_1\cos\left(\alpha_2\right)\right] \quad (8)$$

and ΔF_x can be deduced easily

$$\frac{\rho}{2}\Delta F_x\left[\cos\alpha_1\cos\left(\alpha_2\right)\right]=I\dot{\beta}k_\beta^v+I\beta k_\beta^x, \quad (9)$$

Hence, the controlled yaw dynamics is
$I\ddot{\beta}+I\dot{\beta}k_\beta^v+I\beta k_\beta^x=0$,

which, when coefficients k_β^x, k_β^v are adequately chosen, ensures the yaw closed-loop dynamics to converge to zero, $\lim_{t\to\infty}\beta\to 0$

3.1.2. Estimation of contact forces

Let us suppose, for example, that the dominant side is the left one, that the same difference $\Delta F_x = F_{x1} - F_{x2} = F_{x3} - F_{x4}$ applies. If one can find the

adequate dominant contact force, say F_{x1}, the other desired forces can be deduced in real time from (8-9), as summed up in Figure 6. The next goal is then to be able to monitor the dominant force and the other forces, which depend on the tire-road contact, for a quarter-car. This aspect is handled in paragraph 3.3.

3.2. Non-slip condition using steering actuators

As the vehicle is overactuated, it could be possible to use the steering actuators to monitor both yaw and lateral speed. In this case, the used method consists of lettings all wheels converge to the rolling (non-slip) stage (Figure 5).
In this case, one can simplify equations (2-3) to:

$$m\ddot{v}\approx\left(F_{x1}+F_{x2}\right)\sin\alpha_1-\left(F_{x3}+F_{x4}\right)\sin\left(\alpha_2\right)$$
$$I\ddot{\beta}=\frac{\rho}{2}\left[\left(F_{x1}-F_{x2}\right)\cos\alpha_1+\left(F_{x3}-F_{x4}\right)\cos\left(\alpha_2\right)\right]$$

Adequate control can be obtained using lateral and yaw acceleration feedback:

$$\left(\hat{F}_{x1}+\hat{F}_{x2}\right)\sin\alpha_1-\left(\hat{F}_{x3}+\hat{F}_{x4}\right)\sin\left(\alpha_2\right)=$$
$$m\left(k_v^v\dot{v}+k_v^x v\right)$$
$$\frac{\rho}{2}\left[\left(F_{x1}-F_{x2}\right)\cos\alpha_1+\left(F_{x3}-F_{x4}\right)\cos\left(\alpha_2\right)\right]= \quad (10)$$
$$I\left(k_\beta^v\dot{\beta}+k_\beta^x\beta\right)$$

which yields :

$$\ddot{v}+k_v^v\dot{v}+k_v^x v=0$$
$$\ddot{\beta}+\dot{\beta}k_\beta^v+\beta k_\beta^x=0$$

where parameters $k_{(.)}^{(.)}$ are adequately chosen to ensure the convergence of v, β to zero
Of course, one has to solve equation (10), which will consists of finding the roots of a 4th order polynomial at each time (one can transform this set of equations into two polynomials using the transformation in $\tan\left(\frac{\alpha_i}{2}\right)$).

On a straight line, if $\left(F_{x1}-F_{x2}\right)=\left(F_{x3}-F_{x4}\right)=0$, equation (10) may have no solution, and, in this case, it is necessary that at least one of the wheels be in slip stage.

3.3. Control of a wheel with rolling condition

3.3.1. Dominant wheel control

In order to accelerate properly, one would like to obtain a slip value $G \to 0$ so that $\dot{x}-v\dot{\theta}_{sj} \to 0$

Let us not forget that, in practice, $|\dot{x}| = |\dot{u}| \leq V_{max}$ where V_{max} is the maximum speed of the vehicle.

Combining the two equations in (4) yields, neglecting viscous friction::

$$N_j J_{ej} \ddot{\theta}_{ej} + J_{sj} \ddot{\theta}_{sj} = N_j \Gamma_j - \upsilon F_{xj} \qquad (11)$$
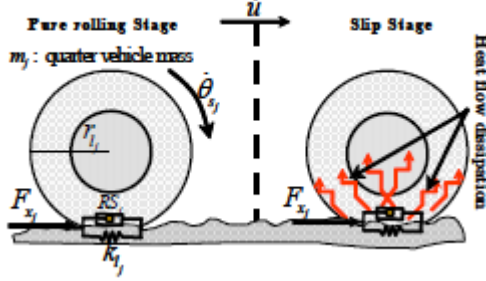

Figure 5. Rolling and slip stages

Acceleration feedback is a powerful control structure which allows to eliminate the vibration mode of such linear flexible systems at the expense of mounting acceleration sensors on the wheels and motor (see (Dieulot and Colas, 2009; Dumetz et al., 2006) for practical demonstrations on 2-mass-spring systems):

$$N_j \Gamma_j = N_j J_{ej} \ddot{\theta}_{ej} + J_{sj} \left(-\ddot{x}/\upsilon \right)$$
$$+ k_{\upsilon}^{v} J_{sj} \left(\dot{\theta}_{sj} - \dot{x}/\upsilon \right) + k_{\upsilon}^{x} J_{sj} \left(\theta_{sj} - x/\upsilon \right) + \upsilon \hat{F}_{xj} \qquad (12)$$

where $\hat{F}_{xj}$ is the estimate of F_{xj}.

Controller (12) ensures that:

$$\left(\ddot{\theta}_{sj} - \ddot{x}/\upsilon \right) + k_{\upsilon}^{v} \left(\dot{\theta}_{sj} - \dot{x}/\upsilon \right) + k_{\upsilon}^{x} \left(\theta_{sj} - x/\upsilon \right) = 0 \qquad (13)$$

and thus that $G \to 0$.

This methodology can be applied for the dominant wheel when not using steering commands or for each wheel, independently, when using steering commands.

3.3.2. Slave wheels control

In the case where steering actuators are not used, slave forces are given by (8-9), and the goal is to ensure convergence of the corresponding slip angle $x - \upsilon \theta_{sj}$ of the jth wheel to the appropriate one, using the model given in (5).

In the same way as before, since

$$N_j J_{ej} \ddot{\theta}_{ej} + J_{sj} \ddot{\theta}_{sj} = N_j \Gamma_j - \upsilon F_{xj} . \qquad (14)$$

One wants that $F_{xj} = \tilde{F}_{xj}$.

Equation (5) or the canonical curve deriving from it provides us with a slip reference $\left(\dot{x} - \upsilon \dot{\theta}_s \right)_{ref}$ which corresponds to the computed reference contact force $\hat{F}_{xj}$ and then one should have:

$$N_j \Gamma_j = N_j J_{ej} \ddot{\theta}_{ej} + J_{sj} \ddot{\theta}_{sj} + J_{sj} \overline{\left(\ddot{\theta}_{sj} - \ddot{x}/\upsilon \right)}$$
$$+ k_{\upsilon}^{v} J_{sj} \overline{\left(\dot{\theta}_{sj} - \dot{x}/\upsilon \right)} + k_{\upsilon}^{x} J_{sj} \overline{\left(\theta_{sj} - x/\upsilon \right)} + \upsilon \hat{F}_{xj} \qquad (15)$$

where $\overline{\left(\theta_{sj} - x/\upsilon \right)} = \left(\theta_{sj} - x/\upsilon \right)_{ref} - \left(\theta_{sj} - x/\upsilon \right)$.

Controller (15) ensures the convergence of $\left(\dot{x} - \upsilon \dot{\theta}_{sj} \right)$ towards $\left(\dot{x} - \upsilon \dot{\theta}_{sj} \right)_{ref}$.

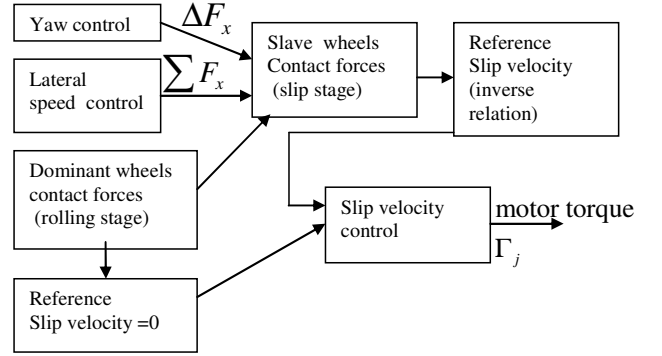The two schemes are summarized in figure 6 and 7.
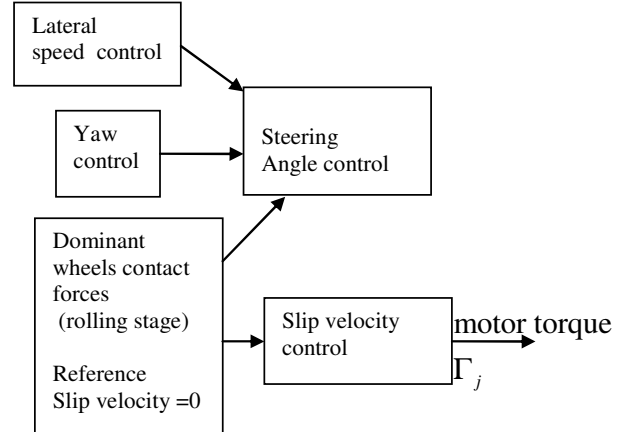

Figure 6. 4X4 control algorithm

Figure 7 Steering actuators plus rolling conditions control algorithm

4. SIMULATIONS

Simulations are carried out with the following parameters: mass $m = 350$ kg , wheel radius $\upsilon = 0.2$ m , axle track $\rho = 1.4$ m , maximum speed $V_{max} = 5$ m.s^{-1} , ellipsoidal inertia $Iz = 46.65$ kg.m^2 . Simulations are carried out for car misalignement (initial nonzero yaw).

The two control schemes show both their performance and limitations; tuning was achieved by imposing as fast closed-loop poles as possible until equations (8-9)

or (10), depending on the control scheme, could find no solution; a less heuristic procedure has to be found and will be the topic of a future paper. The all 4 X 4 control scheme is actually adequate but performances would deteriorate in case of fault – one sees that the overall speed is reduced in order to recover a proper direction. Control using only steering angles results in a strategy where at first the right angle is tracked at low speed and then a higher speed is reached.
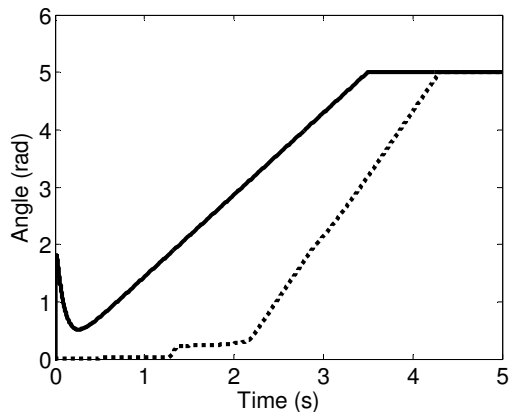

Figure 8 Longitudinal speed
full line: 4x4 monitoring
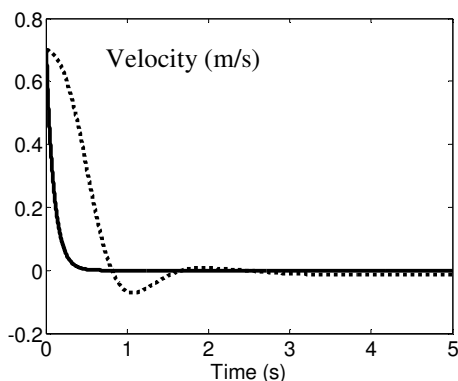dotted line: full steering angle monitoring

Figure 9 Yaw angle
full line: 4x4 monitoring
dotted line: full steering angle monitoring

5. CONCLUSION

Different control algorithms have been proposed for an autonomous overactuated electric vehicle, which owns four actuated traction wheels and two actuated steering systems, based on full 4X4 control or with steering systems control scheme. Preliminary results are given, which show the feasibility of these control algorithms, and the need for a mixed control strategy. Next work will include a generalization of previous results to specific situations (corner crossing, emergency situations, driving on slippery road …) and rapid

control reconfiguration (fault in actuators…), allowing to skip from one control scheme to another.

REFERENCES

Merzouki , R., B. Ould-Bouamama , M.A. Djeziri , M. Bouteldja , Modelling and estimation of tire–road longitudinal impact efforts using bond graph approach, *Mechatronics*, 17, 93–108, 2007

Canudas de Wit, C., P. Tsiotras, E. Veenis , M. Basset , G. Gissinger, Dynamic friction models for road/tire longitudinal interaction. *Vehicle Syst Dyn* , 39, 189–226, 2003.

Merzouki, R., M. A. Djeziri, B. Ould-Bouamama, Intelligent Monitoring of Electric Vehicle, 2009 IEEE/ASME International Conference on Advanced Intelligent Mechatronics, Singapore, July 14-17, 2009

Djeziri, M.A., R. Merzouki, B. Ould Bouamama, Robust Monitoring of an Electric Vehicle With Structured and Unstructured Uncertainties, *IEEE Transactions on Vehicular Technology,* 58, 2009

Wang, J., R. G. Longoria, Coordinated and Reconfigurable Vehicle Dynamics Control, *IEEE Transactions on Control Systems Technology*, 17, 2009

Raffo, G.V., G. K. Gomes, J. E. Normey-Rico, C. R. Kelber, L. B. Becker,, A Predictive Controller for Autonomous Vehicle Path Tracking, *IEEE Transactions on Intelligent Transportation Systems*, 10, 2009

Feiqiang, L., W. Jun, L. Zhaodu, Intelligent Speed Adaptation Using a Self-Organizing Neuro-Fuzzy Controller, IEEE Conference on Vehicle Power and Propulsion, Dearborn, MI, 2009.

Partouche D., Pasquier M., Spalanzani A., IEEE Intelligent Vehicles Symposium, Istanbul, 2007

Furukawa, Y., M. Abe, Advanced Chassis Control Systems for the Vehicle Handling and Active Safety. *Vehicle System Dynamics,* 28, 59-86, 1997

Tjønnås, J., T. A. Johansen, Stabilization of Automotive Vehicles Using Active Steering and Adaptive Brake Control Allocation, *IEEE Transactions on Control Systems Technology*, 18, 2010.

Bakker, E., L. Nyborg, H. B. Pacejka A new tyre model with an application in vehicle dynamics studies. *SAE Transactions: Journal of Passenger Cars*, 98(890087), 1989.

Dieulot, J.-Y., F. Colas, Robust PID Control of a linear mechanical axis: a case study, *Mechatronics.* 19 , pp 269-273, 3-2009.

Dumetz E., J.-Y. Dieulot, P.-J. Barre, F. Colas, T. Delplace, Control of an industrial robot using acceleration feedback , *Journal of Intelligent and Robotic Systems*, 46, pp 111–128, 2006.

MODELING OF WHEEL/TRACK INTERACTION WITH WHEEL DEFECTS AND DIAGNOSIS METHODS: AN OVERVIEW

R. Fesharakifard[(a)], A. Dequidt[(b)], F. Brunel[(c)], Y. Desplanques[(d)], O. Coste[(e)], P. Joyez[(f)]

(a,b) LAMIHUniversity of Valenciennes, Le Mont Houy, 59313 Valenciennes, France
(c,d) LML, Ecole Centrale de Lille, Cité Scientifique, 59651 Villeneuve d'Ascq, France
(e) Signal Développement, 12 boulevard Chasseigne, 86000 Poitiers, France
(f) Eurotunnel, 62231Coquelles, France

[(a)]fesharaki@univ-valenciennes.fr, [(b)]antoine.dequidt@univ-valenciennes.fr, [(c)]florent.brunel@ec-lille.fr, [(d)]yannick.desplanques@ec-lille.fr, [(e)]o.coste@signal-developpement.com, [(f)]patrick.joyez@eurotunnel.com

ABSTRACT

The contact defects between wheel and rail which are originated from wheel profile irregularities and train overloading or unbalancing could deteriorate railway tracks. The interaction defects influence more drastically the maintenance of railways with high traffic like the ones of Eurotunnel tracks. Several works have been performed on different issues of dynamic interaction model between rail and wheel. Such a model can be used for analysis and numerical simulation for wheel profile geometry and train load characteristics. The aim of this paper is to review the major methods and models and classify them, so that different parts of the model are demonstrated: wheel profile geometry, wheel/rail contact model, vehicle and track structure and substructure models, calculation methods and transient simulation models. To achieve the accurate prediction by the model, different methods of exploiting the experimental data to identify the load and diagnosis the interaction defects are also presented.

Keywords: Track/train interaction model, wheel profile irregularity, track response, identification and diagnosis

1. INTRODUCTION

The reduction of railway operating cost is a key issue for infrastructure managers and railway operators. In the scope, maintenance cost reduction is a major target. Highest speed, increased traffic density and load result in accelerated degradation (fatigue). In the same time, interoperability, variety of different vehicles, induces a wider variety of potential degradation sources. One of the most important problems facing the railway maintenance is the monitoring of dynamic behavior of tracks subjected to moving loads (freight and passenger trains) and the defect diagnosis. The structures are therefore subjected to severe vibrations and dynamic stresses, which in turn are much more than the corresponding static stresses.

The dynamic force of railway interaction is influenced by geometrical characteristics of wheel and rail and dynamic characteristics of the load. The goal of diagnosis is to identify these characteristic and the defects related to them. The main geometrical defects are out-of-roundness of wheel profile, rail corrugation, rail joints discontinuity and wheel/rail roughness.

The out-of-roundness (OOR) defects concern the deviation of the wheel tread geometry from its circular shape. Different types of out-of-roundness are catalogued by the International Union of Railways (ETF 04). Two major types of OOR are wheel polygonal and wheelflat. When the brakes are applied to a railway wheel, it can sometimes happen that the wheel locks and slides along the rail. The reason for this may be poorly adjusted, defective or frozen brakes or lack of adhesion at the wheel/rail interface, for example, due to leaves on the rail head. This sliding causes severe wear of the part of the wheel in contact with the rail, leading to the formation of a wheelflat. Such flats on the wheel may be typically 50 mm long but can extend to over 100 mm long.

A gathering of several flats leads to the creation of a polygon. When the wheels rotate, wheelflats generate large impact forces between the wheel and track. Polygonal wheels with a few dominating harmonics (1 to 5 wavelengths around the wheel circumference) have previously been detected especially on high-speed trains. Simulation results show that the most important wavelength-fixing mechanisms of the wheel OOR are the vertical resonance of the coupled train-track system and the frequency region including the lowest vertical track antiresonance (Johansson and Andersson 05).

Wheelflats and rail roughness are very important in the context of dynamic wheel-rail interaction and track deterioration. Dynamic characteristics of the load as variable moving speed and unbalance of wagons load are other important issues to study subsequently. Hereby the parameters like train charge and its vertical position are considered to be fixed.

In order to investigate the track/train dynamic interaction with various load balancing and wheel profile defects, train speeds and static loads, a proper mathematical model is essential. Such a model can be used for analysis and numerical simulation for different

train characteristics and parameters (speed, load …) and method of diagnosis. The aim of this paper is to review the major methods and models and classify them, so that different parts of the model are demonstrated: wheel profile geometry, wheel/rail contact model, vehicle and track structure and substructure models, calculation methods and transient simulation models. Such a model is then usable to identify the moving train characteristic and diagnosis the defects on profile of railway wheel. This paper reports on an overview within the framework of the wheel defect topic of the "Track Train System Availability" (TTSA) project of i-Trans competitiveness cluster.

2. PREVIOUS WORKS

The interaction defects between wheel and rail causes noise and safety problem. When wheel locks and slides along the rail, it produces wear and flattening the wheel. Consequently large vibration amplitudes are created which lead to the damage of track and propagation of fatigue cracks. Several works have been performed on different issues of dynamic interaction between rail and wheel. In this part the related more recent studies are addressed contingent on the following topics.

2.1. Geometry of Wheelflat

The force of contact is expressed as a function of the relative displacement between wheel and rail at the contact point, and it depends on the un-deformed wheel–rail geometry and the elastic characteristics of the wheel–rail contact. Two main kinds of model have been used to study wheel/rail interactions, a moving irregularity between a stationary wheel and rail, and a wheel rolling on the track.

For a discretely supported rail the moving irregularity model cannot deal with the parametric excitation at the sleeper-passing frequency caused by the varying dynamic stiffness in a sleeper bay, and may underestimate the interaction force level at high speeds (Wu and Thompson 06). The moving wheel model is therefore essential to investigate the effects on wheel/rail interaction due to the parametric excitation (Sheng and Thompson 04). This model incorporates vehicles, a track and a layered ground, and uses the moving axle loads and the vertical rail irregularities such as wheelflat in its inputs (Baeza, Roda, and Nielsen 06b).

In a rough approximation, the relation between d, depth of a wheelflat, l_0, its length, and R, the radius of the wheel could be $d \cong \frac{l_0^2}{8R}$. The geometric equations of a wheelflat is detailed in (Baeza, Roda, Carballeira, and Giner 06a; Pieringer and Kropp 08b).

Two kinds of wheelflat geometry are generally considered: the fresh wheelflat with sharp edges as occurring right after formation and the rounded wheelflat, which rapidly develops from the fresh wheelflat as a result of wheel tread wear and plastic deformation. In the literature, a rounded flat is given the same depth as the fresh flat but with a greater length.

Above a certain critical train speed, the wheel separates from the rail when the interface encounters certain types of discontinuities. The strength of the impact and the frequency of repetition are proportional to the train speed (Vér, Ventres, and Myles 76). For a fresh wheelflat, of depth 2 mm and length 86 mm, loss of contact is found to occur for speeds above 30 km/h. For a rounded flat of the same depth but overall length 121 mm the speed at which loss of contact first occurs increases to about 50 km/h (Wu and Thompson 02).

Pieringer and Kropp (08b) derived the equations for the geometric parameters of a wheelflat related to the centre angle, Φ_0. Because the contact algorithm requires the wheel profile expressed in Cartesian coordinates in the wheel-following coordinate system (x',z'), the orientation of the wheel in this coordinate system was described by the angle φ, $0 \leq \varphi < 2\pi$, see figure 1.
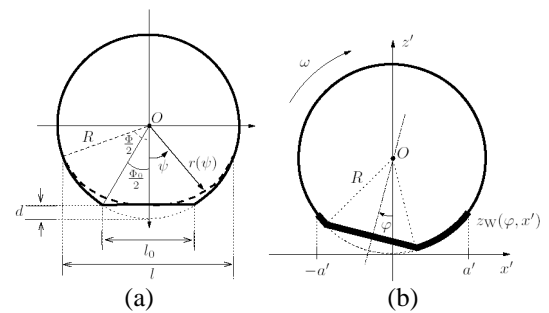

(a)   (b)

Figure 1. Geometry of wheelflats in φ=0 (a); and in angular position φ>0 (b)

In order to distinguish the fresh and rounded flats with simpler equations, Baeza et al. defined the irregularity function as the vertical displacement of the wheel form a reference point for bringing wheel and rail in contact (Baeza, Roda, Carballeira, and Giner 06a).

2.2. Model of contact

This still-relevant classical solution provides a foundation for modern problems in contact mechanics. In linear elastic context, where the area of contact is much smaller than the characteristic radius of the body, each body can be considered an elastic half-space. An elastic sphere of radius R indents an elastic half-space to depth d, and thus creates a contact area of radius a. The applied force F is related to the displacement d by $F = \frac{4}{3}E^* R^{1/2} d^{3/2}$ where $\frac{1}{E^*} = \frac{1-v_1^2}{E_1} + \frac{1-v_2^2}{E_2}$ and E_1, E_2 are the elastic modules and v_1, v_2 the Poisson's ratios associated with each body.

This equation is usable for interaction of a round wheel with a rail surface. To consider a wheelflat, it is considered that a rigid cylinder is pressed into an elastic half-space. So it creates a pressure distribution described by $p(r) = p_0 (1 - \frac{r^2}{a^2})^{-1/2}$ where r is a the radius of the cylinder $p_0 = \frac{1}{\pi} E^* \frac{d}{a}$ and the relationship between the indentation depth and the normal force is given by $F = 2aE^* d$.

For the first time, Hertz solved the problem involving contact between two elastic bodies with curved surfaces. Based on geometrical effects on local elastic deformation, the Hertz formulation relates the normal transmitted force between the bodies by $F_c = K_H \delta^{1.5}$. The application of 2D Hertzian contact had been dominating related to the other methods just as these cases (Vér, Ventres, and Myles 76; Nielsen and Igeland 95; Zhai and Cai 97; Nordborg 02; Sun and Dhanasekar 02; Sheng, Jones, and Thompson 04; Ford and Tompson 06; Wu and Thompson 06).

Despite its popularity, limits of this approach are that geometric requirement in Hertzian model by which the non-deformed surface collide, should be elliptic paraboloids. Other assumptions which are made in determining the solutions of Hertzian contact problems are:

• the strains are small and within the elastic limit,
• each body can be considered an elastic half-space, i.e., the area of contact is much smaller than the characteristic radius of the body,
• the surfaces are continuous and non-conforming,
• the surfaces are frictionless.

This requirement is not satisfied when rail-wheel contact coincides with flat. Additional complications arise when some or all these assumptions are violated and such contact problems are usually called non-Hertzian. The theoretical results of a Hertzian and a non-Hertzian contact model are compared and it is found out that Hertzian model tends to overestimate the peak impact forces (Baeza, Roda, Carballeira, and Giner 06a; Pieringer and Kropp 08b). Non-linearities in the wheel/rail interaction cannot be neglected in the case of excitation by wheelflats because of the resulting large contact forces and the occurrence of loss of contact for train speeds above the critical speed.

A two dimensional model consisting of a Winkler bedding of independent springs between wheel and rail is introduced (Pieringer and Kropp 10b). Figure 2 shows that for the calculation of the normal contact force, this model takes into account one line of combined wheel/rail roughness, in the rolling direction. The springs in the bedding are independent and allow for loss of contact.
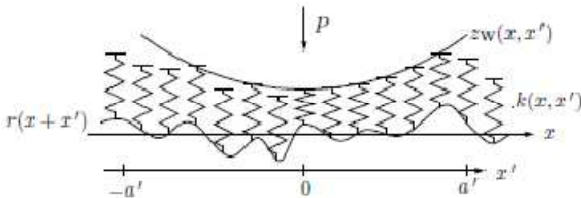

Figure 2. Bedding model of contact

The most general non-Hertzian model used in railway diagnostic is proposed by Kalker (Kalker 95). As mentioned, loss of contact which is occurred on wheelflats, passing over rail joints, and non-linearity could not be considered by Hertzian approach. Thereby the pretabulated Kalker model has taken much attention lastly (Baeza, Roda, Carballeira, and Giner 06a; Baeza, Roda, and Nielsen 06b; Mazilu 07; Mazilu 10). To summarize the process, Baeza, Roda, Carballeira, and Giner (06a) defined the potential contact area (PCA) in such way that it contains every point of the contact area and is rectangular. A discretization of the PCA is established in equal rectangular elements within which the magnitudes to be defined in each element are considered to be constant, as seen in the figure 3.
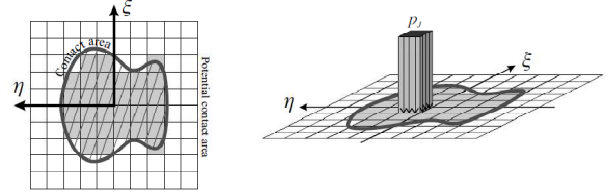

Figure 3. Definition of the potential contact area

2.3. Vehicle Model

In most of cases, a simple model for vehicle is used to achieve the necessary frequency band (Nordborg 02; Wu and Thompson 02; Mazilu 07; Pieringer and Kropp 08b; Steenbergen 08; Pieringer and Kropp 10a). The vehicle motion is governed by the wheel-rail contact. The simplest 1D vehicle model is an unsprung mass. In this model, as shown in figure 4(a), an unsprung mass which represents the wheel is connected to a sprung mass assimilated by a static load for the rest of vehicle. Because the primary suspension filters the high frequency vibration of contact, only the vertical dynamics are considered (Baeza, Roda, Carballeira, and Giner 06a). This model has 2 degrees of vertical movement and the masses are connected through a suspension.
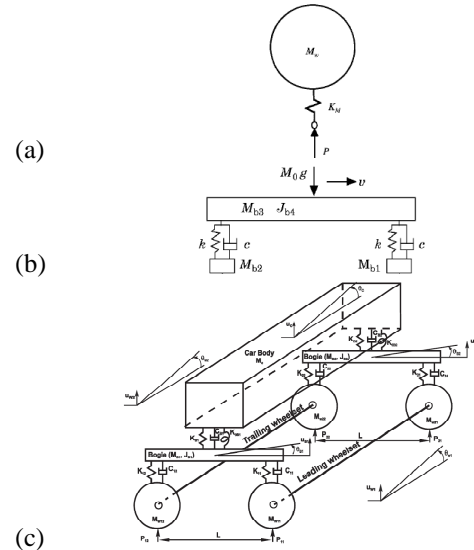

Figure 4. Vehicle model in 1D (a), 2D (b) and 3D (c)

Using just an unsprung mass to show whole vehicle is not the case in practice, where multiple wheels roll on the rail. It is shown that the high-frequency excitation from each wheel can be treated independently by using the superposition principle, provided that the rail vibration is considered as a frequency band average (Wu and Thompson 01).

Although a single unsprung mass on a nonlinear Hertzian contact was the most common vehicle model,

it is shown that this model may largely underestimate the dynamic response (Nielsen and Igeland 95). Therefore a 2D vehicle model is proposed which contains 2 unsprung masses. Figure 4(b) presents the complete configuration of this model. The other parts of the model are 1 sprung mass and 2 suspensions and it has 6 dof. By comparing the results calculated for vehicle models including one and two unsprung masses, it was found that the interaction between the two wheels of the bogie model was primarily due to the vibration of the track structure, whereas the bogie frame was not much a effected by the imperfections studied. This is not surprising since bogie suspensions are designed to isolate the unsprung masses from the rest of the vehicle (Nielsen and Igeland 95).

For vibration analysis in dissymmetric load and stability analysis, 3D vehicle model is proposed (Zhai and Cai 97; Szolc 01; Sun and Dhanasekar 02; Hou, Kalousek, and Dong 03; Johnsson and Andersson 05). Figure 4(c) (Hou, Kalousek, and Dong 03) shows a complete bogie model which is consisted of two wheelsets and a bogie frame. The primary suspension between the wheelsets and the bogie frame consists of linear springs and viscous dampers (Johnsson and Andersson 05). The considered vehicle model is supported on two double-axle bogies at each end and is described as a 10 dof lumped mass system comprising the vehicle body mass and its moment of inertia, the two bogie masses and their moments of inertia, and four wheelset unsprung masses (Zhai and Cai 97; Sun and Dhanasekar 02).

2.4. Track Model

The model of rail and its supports is the most important part which affects the accuracy and speed of the simulation and has attracted much attentions. The simplest model which is widely used to represent an infinite rail is Euler-Bernoulli beam (Zhai and Cai 97; Szolc 01; Nordborg 02; Steenbergen 08; Hammoud, Duhamel, and Sab 10). The dynamic response of an Euler-Bernoulli beam under a moving load is analyzed (Lee 94). This model is limited for high frequencies. Timoshenko beam model was developed to consider the shear deformation and large deflections, and has been used to study vehicle/track dynamics to examine the effect of wheelflats since 1926 (Sun and Dhanasekar 02; Wu and Thompson 02; Baeza, Roda, and Nielsen 06b; Mazilu 07). Timoshenko beam model is known to provide a good representation of the vertical vibration of the rail up to about 2 kHz, above which the rail cross-section deformation should be taken into account. The shear deformations can be included not only for mechanical reasons, but also in order to optimize the discretization in the space-domain. So the physically more realistic Timoshenko beam model offers additional numerical advantages when dealing with transient dynamic problems in unbounded domains.

UIC60 is the most frequent rail type, so its characteristic is frequently used. To have a better representation of the rail, it is divided into two parts: the upper part representing the head and the lower part representing the foot. Both the head and the foot are represented by infinite Timoshenko beams in the rail axis direction. These two beams are connected by continuously distributed springs to allow relative motion between them (Wu and Thompson 99).

Although Euler-Bernoulli model is satisfactory at low frequencies, to consider the shear deformation and rotary inertia at higher frequencies a Rayleigh-Timoshenko finite beam elements is used (Pieringer, Kropp, and Thompson 10b). As an example, curve squealing of railway wheels occurs erratically in narrow curves with a frequency of about 4 kHz (Pieringer and Kropp 10a).

A simple model to fix the rail to the ground is the continuous foundation like an elastic half-space and a Winkler Bedding. The main difference between the elastic half-space and the Winkler bedding lies in the fact that in the elastic half-space the points are coupled with each other, while in the Winkler bedding a set of non-coupled springs is used. Because the points in the elastic half-space are coupled an iterative procedure is necessary to determine the displacements due to the roughness profile at each time step. Therefore this model is computationally more expensive than the Winkler bedding. However, the Winkler system cannot represent a real foundation. It's because the coupling between a beam and an elastic mass under a mobile charge is a problem. Complete comparison between maximum moment and maximum displacement obtained by Winkler bedding and halfspace showed that for a given maximum moment, the Winkler bedding yields 1.5 to 1.8 times higher spring constant than the halfspace (Fischer and Gamsjäger 08).

In figure 5, the vertical geometry of the rail, which is continuously supported, is described by $z(x)$. Continuously supported beam is a significant simplification, which will especially affect the model results when irregularities having a predominant wavelength that is comparable to the sleeper spacing distance (Steenbergen 08).
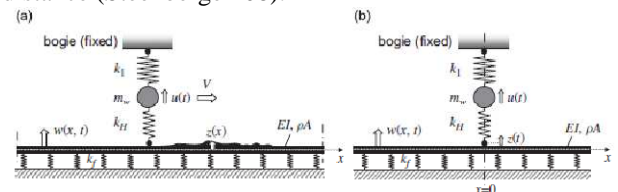

Figure 5. Moving wheel on irregularity z(x) (a) and equivalent excitation z(t) (b)

Figure 6 shows a slab track is considered as an infinite structure consisting of elastically supported double Euler-Bernoulli beams (Mazilu 10).

Some more examples of using the continuous support are presented here:

A simple model is developed based on essential cross-sectional deformation of a double Timoshenko beam in vertical vibration at high frequencies (Wu and Thompson 99).

The foundation consists of distributed non-interacting springs and dampers (Winkler foundation). A track model which has been discretized by use of standard polynomial finite elements will therefore be sought (Nielsen and Igeland 95).
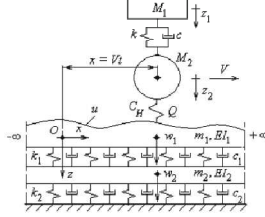

Figure 6. Mechanical model of a moving mass and an infinite homogenous structure

The dynamic behavior of the isolated substructures (rails and sleepers) is described by linear beam theory. The sleepers are modeled as Euler beams on a Winkler foundation, where the ballast is the elastic foundation. Therefore, initially, ballast is considered as having linear characteristics. However, it is possible to include nonlinearities by incorporating them as external forces as a function of sleeper displacements (Baeza, Roda, and Nielsen 06b).

Despite these numerous applications, continuously supported beam is a significant simplification, which will especially affect the model results when irregularities having a predominant wavelength that is comparable to the sleeper spacing distance, and so is not preferred related to discretely supported beams. In cases when loss of contact occurs the effect of the discreteness of the supports becomes important.

While conventional Timoshenko beam model may be used only up to about 2000 Hz, the discretely supported rail model could consider the vertical vibration from about 1000 to 6500 Hz. For high frequencies (at least 5 kHz), the rail cross-sectional deformation is significant. The equation of Timoshenko beam with discrete supports is given in (Sun and Dhanasekar 02).

A discretely supported rail could consider higher frequency vibration modes. Properties associated with the discrete supports are sleeper-passing frequency $f_s = v/l$, and pinned-pinned frequency f_{pp} and the wavelength $\lambda_{pp} = 2l$, where v is the train's speed and l is the distance between 2 consecutive sleepers. Numerical simulations show that it is necessary to include discrete supports in rail modeling to describe the response at low frequencies, determined by the sleeper-passing frequency f_s, and around the pinned–pinned frequency f_{pp}, usually around 1 kHz-in particular if the rail is very smooth or has a corrugation with a wavelength corresponding to the pinned–pinned frequency. If the rail has a corrugation it may also be necessary to include the nonlinear contact spring, since loss of contact occurs for great corrugation amplitudes, e.g., if the corrugation amplitude r_0 is greater than 15 μm when the preload P is 65 kN.

This model is a flexible approach which could be developed in different degree of freedom for the

support. The rail is normally described as an infinitely long beam discretely supported at rail/sleeper junctions by a series of springs, dampers and masses. In (Zhai and Cai 97), the three layers of discrete springs and dampers represent the elasticity and damping effects of the rail pads, the ballast, and the subgrade, respectively. The two layers of discrete masses below the rail represent the sleepers and the ballast, respectively. In order to account for the shear continuity of the interlocking ballast particles, shear springs and dampers are introduced between the ballast masses to model the shear coupling effects in the ballast. Also the transient differential equations for 3 layer of supports and a 10 dof car model is developed.

Railpad and ballast/subgrade have a large influence on the track dynamics at low frequencies. In figure 7, two different visco-elastic models of the rail pad is shown (Nielsen 08).
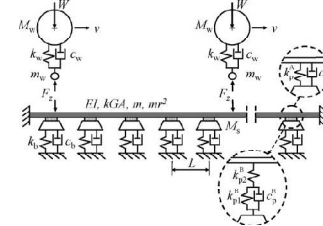

Figure 7. Vehicle model (2 wheelsets) and2 different visco-elastic models of the rail pad

Two track models that include different models of rail pads and ballast/subgrade are compared (Nielsen 08). In track model A, each rail pad is modeled as a discrete linear elastic spring and a viscous damper in parallel (Kelvin model). In track model B, each pad is modeled by a three-parameter visco-elastic model (standard solid model), see figure 7. In track model A, the support under each sleeper is modeled by a Kelvin model, whereas in track model B, a four-parameter visco-elastic model is adopted. The four-parameter model means two spring–damper sets coupled in series, with each set containing one elastic spring and one viscous damper coupled in parallel.

A model is proposed in (Baeza, Roda, and Nielsen 06b) for wheel-rail contact, railpad and ballast with the interaction forces between them and the vehicle. It provides a model with a reduced number of coordinates, so a low computational cost. The developed track model is based on a sub-structuring approach, where a modal description of each isolated rail and sleeper is adopted. Ballast and railpads are considered as connection elements, where the ballast connects sleepers and ground and the railpads connect sleepers and rails.

2.5. Calculation in Time and Frequency Domain

The issue of wheel/rail interaction may be studied by applying the frequency-domain analysis or the time-domain analysis. The most wide spread prediction model for railway noise is a frequency-domain model. It establishes a relationship between receptance and external force at different frequencies using a mathematical transformation under set assumptions,

thereby avoiding the solution of complicated differential equations. It is in the nature of frequency-domain models that they can only include a linear contact model. So the frequency-domain analysis requires the assumption that the track is a linear (steady state) structure. Steady-state interaction for an asymmetrical vehicle/track model is calculated by (Hou, Kalousek, and Dong 03). For steady state, each wheel/rail force is a periodic function of time and can be expressed as a Fourier series. The noise is usually presented in frequency bands such as one-third octaves.

Frequency-domain method mainly proposed by Remington is premature, requiring evidence from more advanced description of the generation mechanism. To study the non-linear effect of roughness on interaction force and provide an excitation input to exciting linear model of noise, 2 procedures are examined (Remington and Webb 96). Alternative excitation mechanisms include nonlinearities and parametric excitation due to charging rail and contact receptance with position. For continuously supported track, nonlinear wheel/rail contact effects are mostly unimportant, unless the roughness level is very high or the static load is very small.

Because of the fact that the model is simplified and its linearity, it is not suitable for quantitative predictions with a high level of accuracy. So the nonlinear characteristics of the track structure and the non-linear wheel/rail contact should be modeled in time-domain. However, the time-domain models are more complex than frequency-domain models.

Based on a derivation of the translational and rotational dynamic stiffness of both infinite Timoshenko and Euler-Bernoulli beams on Winkler foundation in the frequency-domain and time-domain beam models, the use of Timoshenko's beam model leads to an asymptotic behavior in the frequency-domain which is linear with respect to $i\omega$. Thus, the corresponding expression in the time-domain is a first-order time derivative. Contrary to Timoshenko's model, Euler-Bernoulli's beam theory generates rational powers of $i\omega$ in the frequency-domain and consequently fractional derivatives in the time-domain with memory integrals to be solved. Their evaluation asks for nonlocal time-solvers with much higher computational effort than local solvers (Ruge and Birk 07).

Two methods of time & frequency-domain are presented to diagnosis the noise generation in wheel/rail contact including discrete supports, parametric excitation, and nonlinear contact spring (Nordborg 02).

A wavenumber-based frequency-domain calculation method is used for the response of a periodically supported rail to a moving harmonic load, allowing dynamic wheel/rail forces to be calculated for single or multiple wheels moving over an initially smooth or rough track. It is shown that this approach is more efficient than a full time-domain approach, but extends earlier work by including the effect of forward motion explicitly.

Another application of time-domain method is that the roughness spectrum is not derived from the wheelflat geometry directly, but from the results of the time-domain calculation. Thus, it is used as a means of converting the wheel/rail interaction force into an equivalent roughness input. It is shown that a time-domain wheel/rail interaction analysis model gives similar results to quasi-static roughness filtering with a constant load for moderate roughness, but dynamic effects became significant when the roughness amplitudes were large, particularly with dipped rail joints (Ford and Thompson 06).

In order to predict the consequent noise radiation, the wheel/rail interaction force is transformed from time-domain into the frequency-domain and then converted back to an equivalent roughness spectrum. This spectrum is used as the excitation to a linear, frequency-domain model of wheel/rail interaction to predict the noise. This hybrid approach has been shown to be adequate by comparing direct and hybrid calculations for a wheel with a single, lightly damped resonance (Wu and Thompson 02).

3. TRANSIENT SIMULATION MODELS

Displacements in physical coordinates, required to calculate the forces transmitted through railpads and contacts, and the values of the force terms which appear in the differential equations (Baeza, Roda, and Nielsen 06b). From the geometry, the kinematics, and the dynamics of the wheel/rail system, analytical models could be developed to identify the major variables controlling the generation of impact noise. The coupling between normal and lateral directions was introduced through the track dynamics due to an offset of the wheel/rail contact point from the rail centre line which is assumed as an input to the model (Pieringer, Kropp, and Nielsen 08a).

The beam model could be considered in time-domain via its receptance. Vertical vibration receptances of the rail have been calculated using both continuously and discretely supported rail models (Wu and Thompson 99). Nielsen studied the vertical dynamic behavior of a bogie moving on a rail discretely supported via railpads by sleepers resting on an elastic foundation. The transient interaction problem is numerically solved by use of an extended state-space vector approach in conjunction with a complex modal superposition for the track (Nielsen and Igeland 95).

Hammoud showed that the fully continuous solutions could not give a satisfactory response for the frequency response model. A formulation for coupling discrete and continuum models for both dynamic and static analyses was given, which offered the better simulations of material properties than the discrete calculations (Hammoud, Duhamel, and Sab 10).

Consideration of wheels moving along a discretely supported rail is normally achieved in the time-domain by solving differential equations as an initial-value problem. Pretension in the rail affects the dynamic response of the rail. The wave propagation velocity in

the track is greater than 1000 km/h. For train speeds not exceeding 350-400 km/h the proper estimations performed have shown that the wave effects can be neglected and the most significant influence of vehicle-track relative motion is expressed by periodic fluctuation of track properties during run over successive sleepers.

In analytic solution for a continuously supported rail equation, the derivations related to time and positioned are replaces with its Laplace and Fourier transformed equivalent. But the analytical method will never give a simple and precise solution for a discretely supported rail. To obtain the satisfactory results, one method is using the Green's function; where wheel and rail are represented by impulse response functions. Otherwise the time-domain approaches require the track to be truncated into a finite length. To minimize wave reflections from the truncations and to be able to account for high-frequency vibration, the track section must include at least 100 sleepers and the rail must be modeled using either the finite element method or the modal superposition method employing more than 100 modes (Sheng, Li, Jones, and Thompson 07). The mentioned solutions are explained in this part.

3.1. Finite Element Method
It is generally found that the infinite beams representing the rail resting on an elastic foundation provides only a limited insight into the dynamic response of various track components. An improvement to such models is achieved by accounting the discrete spacing of the sleepers. The discrete support models and the finite element model allow improved prediction of the rail response and offer the potential for refinement by including all conceivable track components as layers.

The advantages of finite element modeling is that the dynamic analysis solved by use of numerical time-stepping routine, non-linear components and contacts may be included. Using real valued modal analysis and U-transformation, to find the dynamic response of an infinite uniform beam, by exact or approached methods. Disadvantage of finite element is that non-physical discontinuity in slope over element boundaries may occur.

A finite element time-domain model for is proposed to determine the dynamic responses to wheel/rail interaction (Hou, Kalousek, and Dong 03). The solution of the wheelflat case reveals that the wheel/rail impact on one rail significantly affects the wheel/rail interaction on the other side of the track.

3.2. Assumed Mode Simulation for Discrete Support
The modal properties of rail and sleepers can be calculated from an Euler-Bernoulli or Timoshenko beam model. The equation of motion in matrix form is formulated by using Hamilton's principle and the assumed mode method (Lee 94). The intermediate point constrains are located arbitrarily along the beam and are modeled as linear springs of very large stiffness.

From a modal basis viewpoint, the vertical displacement of a point located through the longitudinal coordinate x at the instant of t is $v^i(x,t) = \sum_{m=1}^{N_i} \phi_m^i(x) q_m^i(t)$, where $\phi_m^i(x)$ is the mth vibration mode and $q_m^i(t)$ is a set of modal coordinates.

Problem size depends on the number of rails and sleepers, the number of vibration modes considered in the modal descriptions of these elements, and the number of coordinates considered in the vehicle model (Baeza, Roda, Carballeira, and Giner 06b).

3.3. Green's Function
As described, one approach for track model is a finite/boundary method which should be used for a truncated model. To achieve a high precision in this method, the track model is sufficiently long, which is time consuming. Other approach based on analytical models and Green's function method, consider the track as an infinite structure.

Because the time-dependent stiffness of the track is obtained by inversing the receptance due to a unit stationary harmonic load, some authors have tried to study the transient simulation model by Green's function (Mazilu 07; Pieringer, Kropp, and Nielsen 08a; Mazilu 10). Mazilu investigated the interaction between a moving vehicle and a slab track by using new forms of the time-domain Green's functions for both slab track and vehicle, that are suitable to account for the nonlinear wheel/rail contact and the Doppler effect, due to substructure technique. From the receptance of the rail and slab in the stationary coordinate system, and applying the inverse Fourier transform, the time-domain Green's functions of the rail and slab are calculated step-by-step taking into account the moving impulse force and assembled in the so-called the Green's matrix of the track, invoking the damping feature of the track structure. From this equation, the contact force results and then, the wheel and bogie displacements and the rail and slab deflection may be calculated separately. More accurate results for the high-frequency range might be obtained by replacing the Euler–Bernoulli beam model with the Timoshenko beam model.

On the other hand, the Green's functions of the two-mass oscillator (the wheel and the suspended mass of the bogie) for the time-domain analysis are expressed via the Laplace transform. Starting from the equations of the wheel and rail displacement, the Green's matrix of the track (only the vector corresponding to the rail) and the Green's function at wheel meet in the nonlinear equation of the wheel/rail contact. From this equation, the contact force results and then, the wheel and bogie displacements and the rail and slab deflection at the section of the contact point may be calculated separately. No limitative condition regarding the wheel/rail contact or the irregularities of the rolling surfaces is required. In the quasi-static conditions, all Green's functions are set to zero. The quasi-static approach proposed by Wu is adequate to account for the nonlinearity of the wheel/rail contact, but they are not able to simulate the presence of the Doppler Effect.

Green function could be used for both frequency-domain and time-domain models. The time-model proposed by Heckel's determines the vertical rail deflection by time integrating Green function. The frequency-domain model solves a descretized integral equation by coupling the frequency components which are the Fourier coefficients of expansion of the varing receptance along the track, described by the track's Green function. Transforming of Green's function in frequency-domain into time-domain is down by a discrete Fourier transform.

Pieringer also presented the wheel Green's functions and the track by moving Green's functions (Pieringer, Kropp, and Nielsen 08a). One major advantage of the presented interaction model is its high computational efficiency. Even if combined with a complex finite element model of the track, the calculation time for a simulation is typically less than 20 seconds. Because the Green's matrix of the track has to be calculated once for a particular speed and then it can be used to simulate the interaction between the track and the vehicle for any set of vehicle parameters. One drawback of the moving Green's functions is that it should be calculated for each train speed.

4. VEHICLE PARAMETERS IDENTIFICATION AND WHEEL DIAGNOSIS

One method to evaluate the life assessment of the rail is direct measurement of the forces exerted by the train. A transducer called MPQY is developed for measuring the interaction train/track forces to forecast life of the rail (Delprete and Rosso 09). If the interaction forces could be measured, a damage model can be used to evaluate the track life. So in a continuous procedure, every vehicle should be equipped by a force measuring instrument. Since this solution is not suitable for infrastructure managers and is expensive for railway operators, a stationary instrumentation on the rail is often preferred.

4.1. Vehicle parameters identification

In recent years, the technique of moving load identification has been developing very rapidly. Although different systems have been developed for measuring static weigh of the moving vehicles, the important problem is to identify the dynamic load railways. This problem has been discussed with different approaches for bridge deck. Here some cases are presented and they illustrate the possibilities for identifying a vehicle moving on ground-jointed rail. Other parameters like the friction coefficient could be characterized by the load and contact conditions.

The implementation of bridge weigh-in-motion technology is an inexpensive method to measure vehicle characteristics and true dynamic response (Liljencrantz, Karoumi, and Olofsson 07). An existing method of moving load identification on a single-span bridge deck was generalized for a continuous bridge deck with general boundary conditions (Zhu and Law 06). Based on modal superposition and regularization technique,

this method was used to identify the moving loads on the elastically supported bridge deck. Numerical simulations showed that the method has been effective to identify accurately the moving loads on the bridge with elastic bearings using different types of measured responses. Measured acceleration gives better results than those from strains. Therefore, data acquisition is based primarily on the use of accelerometers.

Au, Jiang, and Cheung (04) simulated the acceleration measurements from the solution to the forward problem of a continuous beam under moving vehicles, together with the addition of artificially generated measurement noise. The identification was carried out through a robust multi-stage optimization scheme based on genetic algorithms, which searches for the best estimates of parameters by minimizing the errors between the measured accelerations and the reconstructed accelerations from the moving vehicles.

While most of moving force identifications are based on modal decomposition and modal truncation error, a finite element based method was developed (Law, Bu, Zhu, and Chan 04). The measured displacements were formulated as the shape functions of the finite elements of the structure are modeled as straight beam. In this method, the identified results are relatively not sensitive to the sampling frequency, velocity of vehicle, measurement noise level and road surface roughness.

4.2. Railway wheel diagnosis

Wheel defects like OOR and wheelflat cause high impact forces to train and track components. The accelerometer based methods for determining the wheel impact load resulting from wheel defect could prevent catastrophic failure of these components. This approach relies on the dynamic response of the track in determining the magnitude of the impact load imposed on the track by a defective wheel (Wasiwitono, Zheng, and Chiu 07).

The wavelet transform could be used to develop a diagnostic tool for quantifying the wheelflat defect in different train speeds. It is claimed that this diagnostic method is very effective to detect all the damaged wheel and to measure the train speed with a single rough sensor set-up (Belotti, Crenna, Michelini, and Rossi 06). To detect the performance robustness, they varied the threshold value of 10% and verified how results modify.

Skarlatos, Kleomenis, and Trochidis (04) proposed a fuzzy-logic method for wheel defects diagnosis. First, the vibration signatures caused by the rail-wheel interactions were recorded on the rail both in case of healthy wheels and wheels with defects known a priori. The measurements were made on new rails without any defect. Consequently, any expected change in the vibration signatures would reflect the condition of the wheels. Next, the measured data were statistically analyzed and confidence intervals for the wheel condition depending on train speed and frequency of analysis were established.

Another method of detecting the wheelflats by accelerometer is energy and cepstrum analysis of rail acceleration. Energy analysis is useful to estimate the global stress which the undergoes, while ceptrum analysis, a signal processing technique capable of detecting echoes even in strongly noisy signals, allows the detection of the independently from the presence of other defects, even when their effects are hidden (Bracciali and Cascini 97).

Recently one flat detector method using Doppler Effects is presented (Brizuela, Ibanez, Nevado, and Fritsch 10). This system analyses the rail/wheel contact by frequency and phase shifts. When a wheel moves at constant speed, the receiving signal presents a regular shift related to the movement speed. The difference between the emitted and the received frequencies changes if any defect on the wheel tread is detected.

5. CONCLUSION

In this review, some different models which provide an efficient prediction of wheel/rail dynamic behavior relating to the interaction defects were presented. Despite the large volume of researches which have been done on this subject, few work analyze comprehensively the problem for an experimental case.

In order to specify a global dynamic model for railway vehicle/track interaction, some precise and quick methods will be developed. Following the overview, the models comprise 3 different sub-models for track, contact and vehicle, where each part could be refined with increasing complexity. The rail is modeled by Euler-Bernoulli, Timoshenko and double Timoshenko beam methods, considering the necessary frequency bandwidth. For the same reason, type of beam support could be continuous or discretized. For the model of contact, after using the nonlinear Hertzian force, a Winkler bedding contact will be implemented. Two central defects of interaction stated in our study are different types of wheel contour irregularity and asymmetrically loaded vehicle. To analyze the wheel profile irregularities, a 1D vehicle model with 2 dof will be used to reduce the calculation time of the model. But the asymmetrical loading problem could only be analyzed by implementing a 3D vehicle model.

Firstly, each model should face to theoretical results reported in the literature. These models will be corrected and their parameters will be identified later by experimental results that are given via different experimental campaign on Eurotunnel platform. Based on the experimental data, the proposed methods of diagnosis will predict the defects of profile of railway wheel and identify the train load.

REFERENCES

Au, F.T.K., Jiang, R.J. and Chenug, Y.K., 2004. Parameter identification of vehicles moving on continuous bridges. *Journal of Sound and Vibration*, 269, 91-111.

Baeza, L., Roda, A., Carballeira, J. and Giner, E., 2006. Railway train-track dynamics for wheelflats with improved contact models. *Nonlinear Dynamics*, 45 (3-4), 385-397.

Baeza, L., Roda, A. and Nielsen, J.C.O., 2006. Railway vehicle/track interaction analysis using a modal substructuring approach. *Journal of Sound and Vibration*, 293 (1-2), 112-124.

Belotti, V., Crenna, F., Michelini, R.C. and Rossi, G.B., 2006. Wheel-flat diagnostic tool via wavelet transform. *Mechanical Systems and Signal Processing*, 20, 1953-1966.

Bracciali, A. and Cascini, G., 1997. Detection of corrugation and wheelflats of railway wheels using energy and cepstrum analysis of rail acceleration. *IMechE*, 211 (F), 109-116.

Brizuela, J., Ibanez, A., Nevado, P. and Fritsch, C., 2010. Railway Wheels Flat Detector Using Doppler Effect. *Physics Procedia*, 3, 811-817.

Delprete, C. and Rosso, C., 2009. An easy instrument and a methodology for the monitoring and the diagnosis of a rail. *Mechanical Systems and Signal Processing*, 23, 940-956.

Fischer, F.D. and Gamsjäger, E., 2008. Beams on Foundation, Winkler Bedding or Halfspace - a Comparison. *Technische Mechanik*, 28 (2), 152-155.

Editions Techniques Ferroviaires (ETF), 2004. Atlas of wheel and rail defects. *A report commissioned by the Steering Group of UIC/WEC*, ISBN 2-7461-0818-6.

Ford, R.A.J. and Tompson, D.J., 2006. Simplified contact filters in wheel/rail noise prediction. *Journal of Sound and Vibration*, 293 (3-5), 807-818.

Hammoud, M., Duhamel, D. and Sab, K., 2010. Static and dynamic studies for coupling discrete and continuum media; Application to a simple railway track model. *International Journal of Solids and Structures*, 47 (2), 276-290.

Hou, K., Kalousek, J. and Dong, R., 2003. A dynamic model for an asymmetrical vehicle/track system. *Journal of Sound and Vibration*, 267 (3), 591-604.

Johansson, A. and Andersson, C., 2005. Out-of-round railways - a study of wheel polygonalization throught simulation of three-dimensional wheel-

rail interaction and wear. *Vehicle System Dynamics*, 43 (8), 539-559.

Kalker, J.J., 1990. Three-Dimensional Elastic Bodies in Rolling Contact. *Kluwer Academic Publishers*, Dordrecht, ISBN 0-7923-0712-7.

Law, S.S., Bu, J.Q., Zhu, X.Q. and Chan, S.L., 2004. Vehicle axle loads identification using finite element method. *Engineering Structures*, 26, 1143-1153.

Lee, H.P., 1994. Dynamic response of a beam with intermediate point constraints subject to a moving load. *Journal of Sound and Vibration*, 171 (3), 361-368.

Liljencrantz, A., Karoumi, R. and Olofsson, P., 2007. Implementing bridge weigh-in-motion for railway traffic. *Computers and Structures*, 85, 80-88.

Mazilu, T., 2007. Green's functions for analysis of dynamic response of wheel/rail to vertical excitation. *Journal of Sound and Vibration*, 306 (1-2), 31-58.

Mazilu, T., 2010. Interaction between a moving two-mass oscillator and an infinite homogeneous structure: Green's functions method. *Archive of applied mechanics*, 80 (8), 909-927.

Nielsen, J.C.O and Igeland, A., 1995. Vertical dynamic interaction between train and track - influence of wheel and track imperfections. *Journal of Sound and Vibration*, 187 (5), 825-839.

Nielsen, J.C.O., 2008. High-frequency vertical wheel-rail contact forces - Validation of a prediction model by field testing. *Wear*, 265 (9-10), 1465-1471.

Nordborg, A., 2002. Wheel/rail noise generation due to nonlinear effects and parametric excitation, *Journal of the Acoustical Society of America*, 111 (4), 1772-1781.

Pieringer, A., Kropp, W. and Nielsen, J.C.O., 2008. A Time-Domain Model for Wheel/Rail Interaction Aiming to Include Non-linear Contact Stiffness and Tangential Friction. *Notes on Numerical Fluid Mechanics and Multidisciplinary Design*, 99, 285-291.

Pieringer, A. and Kropp, W., 2008. A fast time-domain model for wheel/rail interaction demonstrated for the case of impact forces caused by wheel flats. *Proceeding of Acoustics'08*, Paris.

Pieringer, A. and Kropp, W., 2010. A Time-Domain Model for High-Frequency Wheel/Rail Interaction Including Tangential Friction. *Proceedings of 10th French Congress of Acoustics (CFA10)*.

Pieringer, A., Kropp, W. and Thompson, D.J., 2010. Investigation of the dynamic contact filter effect in vertical wheel/rail interaction using a 2D and a 3D non-Hertzian contact model. *Wear*, 11 pages.

Remington, P. and Webb, J., 1996. Estimation of wheel/rail interaction forces in the contact area due to roughness. *Journal of Sound and Vibration*, 193 (1), 83-102.

Ruge, P. and Birk, C., 2007. A Comparison of Infinite Timoshenko and Euler–Bernoulli Beam Models on

Winkler Foundation in the Frequency- and Time-Domain. *Journal of Sound and Vibration*, 304, 932–947.

Sheng, X., Jones, C.J.C. and Thompson, D.J., 2004. A theoretical model for ground vibration from trains generated by vertical track irregularities. *Journal of Sound and Vibration*, 272 (3-5), 937-965.

Sheng, X., Li, M., Jones, C.J.C. and Thompson, D.J., 2007. Using the Fourier-series approach to study interactions between moving wheels and a periodically supported rail. *Journal of Sound and Vibration*, 303 (3-5), 873-894.

Skarlatos, D., Karakasis, K. and Trochidis, A., 2004. Railway wheel fault diagnosis using a fuzzy-logic method. *Applied Acoustics*, 65, 951-966.

Steenbergen, M.J.M.M., 2008. Quantification of dynamic wheel–rail contact forces at short rail irregularities and application to measured rail welds. *Journal of Sound and Vibration*, 312 (4-5), 606-629.

Sun, Y.Q. and Dhanasekar, M., 2002. A dynamic model for the vertical interaction of the rail track and wagon system. *International Journal of Solids and Structures*, 39 (5), 1337-1359.

Szolc, T., 2001. Simulation of Dynamic Interaction between the Railway Bogie and the Track in the Medium Frequency Range. *Multibody System Dynamics*, 6 (2), 99-122.

Vér, I.L., Ventres, C.S. and Myles, M.M., 1976. Wheel/rail noise - Part III: Impact noise generation by wheel and rail discontinuities. *Journal of Sound and Vibration*, 46 (3), 395-417.

Wasiwitono, U., Zheng, D. and Chiu, W.K., 2007. How useful is track acceleration for monitoring impact loads generated by wheel defects? *Proceeding of 5th Australasian Congress on Applied Mechanics*, 10-12 December 2007, Brisbane, Australia.

Wu, T.X. and Thompson, D.J., 1999. A double Timoshenko beam model for vertical vibration analysis of railway track at high frequencies. *Journal of Sound and Vibration*, 224 (2), 329-348.

Wu, T.X. and Thompson, D.J., 2001. Vibration analysis of railway track with multiple wheels on the rail. *Journal of Sound and Vibration*, 239 (1), 69-97.

Wu, T.X. and Thompson, D.J., 2002. A hybrid model for the noise generation due to railway wheel flats. *Journal of Sound and Vibration*, 251 (1), 115-139.

Wu, T.X. and Thompson, D.J., 2006. On the rolling noise generation due to wheel/track parametric excitation. *Journal of Sound and Vibration*, 293 (3-5), 566-574.

Zhai, W. and Cai, Z., 1997. Dynamic interaction between a lumped mass vehicle and a discretely supported continuous rail track. *Computers & Structures*, 63 (5), 987-997.

Zhu, X.Q. and Law, S.S., 2006. Moving load identification on multi-span continuous bridges with elastic bearings. *Mechanical Systems and Signal Processing*, 20, 1759–1782.

Diagnosing photovoltaic systems: an iterative process

Abed Alrahim Yassine, Stéphane Ploix, Hussein Joumaa

G-SCOP lab
46, avenue Félix Viallet F-38031 Grenoble cedex 01

Abed-Alrahim.Yassine@g-scop.inpg.fr, Stephane.Ploix@inpg.fr, Hussein.Joumaa@imag.fr

Abstract— This paper presents the principles of a fault diagnosis system for photovoltaic plants that makes it possible iterative diagnosis processes. It proposes a solution for auto generating tests from models and consecutive measurements. Moreover, it is shown that overall possible diagnoses may be compute from the symptoms coming from the generated tests. The operation principles are detailed and technological aspects for implementation are discussed. Thanks to these diagnosis systems, maintenance operators just have to depict the photovoltaic plant and to introduce iteratively the measurements they do. The system to be diagnosed is transformed into a MILP Problem (Mixed Integer Linear Programming) and a GLPK solver is used to find out all the possible diagnoses.

Index Terms— Fault diagnosis, detection test design, Symptom generation, constraint satisfaction problem, MILP solver.

I. INTRODUCTION

Today's competitive global economy requires continued advances in fault diagnostic technology to achieve and maintain cost advantages. This challenge can be envisaged by using, as much as possible, efficient diagnostic systems that makes it possible to detect an isolate the faults the wich occurs in systems.

In this paper, the model-based diagnosis method has been adopted because of its analysis capabilities and its generality. This method provides an alternative to the traditional techniques based on experience, such as rule-based or case-based reasoning systems. In the model-based method, tree main kinds of approaches might be found: The FDI approach, which focuses on fault detection in dynamic systems, has been summarized in [2]. The community **FDI** (Fault Detection and Isolation ([19], [12], [4], [10], [14]) bases the foundations of its solving approaches on engineering disciplines such as control theory and static decision making. The DX approach, summarized in [13], focuses on the diagnosis reasoning. The main works of the **DX** approach ([22], [7]) base the foundations of solution approaches on the field of computer science and artificial intelligence. Recently, a bridge approach between FDI and DX has been proposed ([17], [20]).

The bridge approach has been adopted for the design of diagnosis system. It can be decomposed into two stages: detection and isolation. In the first stage, the detection tests are computed for the system to be diagnosed from the system model composed of a set of constraints. Then, symptoms are deduced from these detection tests. One of the proposed methods for the symptom generation is to transforme the system to be diagnosed into a MILP Problem (Mixed Integer Linear Programming) which is a regular constraint satisfaction problem and to use GLPK solver which makes it possible to compte symptomes.

In the second stage, the fault diagnosis is performed from the generated symptoms.

In order to find out the detection tests, several sensors have to be installed. The performance of a diagnostic system highly depends on the number and on the location of actuators and sensors. Several sensor placement methods are presented in this paper.

[23] has proposed a method based on consecutive retraction of sensors, which takes into account diagnosability criteria. This approach is based on Analytical Redundancy Relations (**ARR**) [2]. However, this method requires an a priori design of all the **ARR** for a given set of sensors.

Recently, the sensor placement problem satisfying diagnosability objectives becomes possible without designing **ARR**s. [11] has proposed a efficient method based on a Dulmage-Mendelsohn decomposition [9]. Nevertheless, this method only applies to just-determined sets of constraints while most practical systems are under-determined when sensors are not taken into account, and over-determined afterwards.

Another sensor placement method without designing **ARR**s has be presented in ([27], [28]). This method improves the possibility of detecting and localizing faults in systems for which only the structure is known. It considers the complete range of specifications with respect to the constraints, i.e. the set of constraints that must be diagnosable, the set of constraints that must be non discriminable but detectable and the set of constraints that must be non detectable.

The paper is organized as follows. Section II describes the necessary steps for fault diagnosis. Section III presents the concepts used to model systems in a standardized form for diagnosis. Section IV details the methods to find out all the testable subsystems. Section V presents a new method of symptom generation from the testable subsystems by transforming the system to be diagnosed into a MILP Problem and by using a GLPK solver. In section VI, the diagnosis reasoning method is presented. In section VIII, the steps of the fault diagnosis system is applied on a photovoltaic
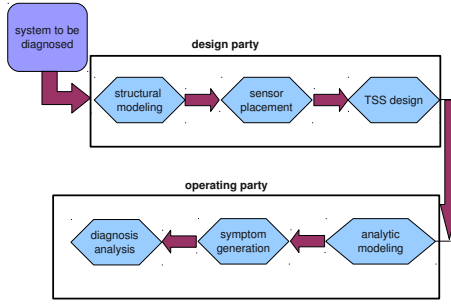
Fig. 1. Fault diagnosis system

system. Finally, the conclusions are drawn in Section IX.

II. FAULT DIAGNOSIS SYSTEM SCHEMA

The goal of a fault diagnosis system is to detect and isolate the possible faults which can occur in systems. The necessary steps of the diagnostic system are given in figure 1.

1) the first step is the structural modeling of the system to be diagnosed. This step consists in colleting a structural model of the systsem behavior
2) the second step is the sensor placement. This step makes it possible to find the optimal sensor placement satisfying objectives like observability, monitorability and diagnosability
3) the third step is the design of testable subsystems (**TSS**s) i.e a constraint set which leads to a test. This step makes it possible to compute the necessary **TSS**s for diagnosis
4) the fourth step is the analytical modeling of the system to be diagnosed. This step consists in colleting a model the system behavior according to section III
5) The fifth step is the symptom generation. This step makes it possible to generate the symptoms from the **TSS**s found.
6) The last step is the diagnostic analysis. This step computes out all the diagnoses from the symptoms found.

III. SYSTEM MODELING FOR DIAGNOSIS

This section deals with concepts used to model systems in a standardized form for diagnosis. Let us firstly introduce some basic definitions.

A variable models an information, material or energy flow. It is named *shared variable* when it is shared by several components. Shared variables play the role of *port* ([6]) for connection between components. Variables are therefore potentially observable elements of information about the actual state of system. It is important to distinguish between a variable which is related to a physical phenomena from a parameter which is model-dependent. Generally speaking, even if the variable is observable, it is not possible to merge it with data because in fault diagnosis, data are only known providing that some actuators or sensors behave properly.

This fact leads to the concept of *data flow*. It is used to model data provided by a source of information concerning a variable. This source may be a sensor or an actuator. The set of all possible data flows, that can be collected from a system when spanning all the possible values for controlled variables, is named OBS. A set of observed values is written obs.

In fault diagnosis, a system is not supposed to remain in a given mode. Indeed, diagnostic analysis aims at retrieving the actual behavior modes of the components of a system. At least, two modes have to be defined: the *ok* mode, which corresponds to the normal behavior and *cfm* mode, which is the complementary fault mode: it refers to all the behaviors that do not fit to the normal behavior. Sometimes, specific fault modes may be modeled but in this paper, for the sake of clarity, only the modes ok and cfm are considered.

Except for the complementary fault mode, behavioral modes are modeled by cause-effect relationships between variables. These relationships are represented by constraints. Each constraint refers to a set of mappings containing unknown variables and known data flows represented by $obs \in OBS$. This fact can be formalized as follows:

$$ok(c_i) \leftrightarrow \forall obs \in OBS, k_i(obs, V_i) = 0 \qquad (1)$$

where $k_i(obs, V_i) = 0$ is the constraint set to be satisfied by component c_i in mode $ok(c_i)$ and V_i is the set of variables involved in k_i.

However, it is obvious that in most of practical fault diagnosis problems, it is not possible to check all the context. Equation (1) leads to:

$$ok(c_i) \rightarrow \forall obs \in \mathcal{O}, k_i(obs, V_i) = 0 \qquad (2)$$

where $\mathcal{O} \subset OBS$.

For the sake of simplicity, elementary models are formalized by:

$$(ok(c_i), k_i(obs, V_i) = 0) \qquad (3)$$

Consequently, the model of a system Σ is composed of a set of components and a set of behavioral modes related to these components.

Let's now present the way of constructing the system model. A complex system can be broken into a set of connected components. In other words, a system can also be constructed by joining together its components. The model of a component is assumed to be embedded into it thanks to a communication mechanism: component could be named *smart* or *intelligent components*. This model can represent one of the possible modes (*ok* or *cfm* here). As mentioned before, a component can connect others via its variables. Several components can be joined together via a connection point, which is a shared variable. The connection points between the different components of a system have to be specified by the installer. A variable can also be connected to sensors which provide data flows.

In order to illustrate the fault diagnosis system, consider for instance the photovoltaic system shown in figure 2.
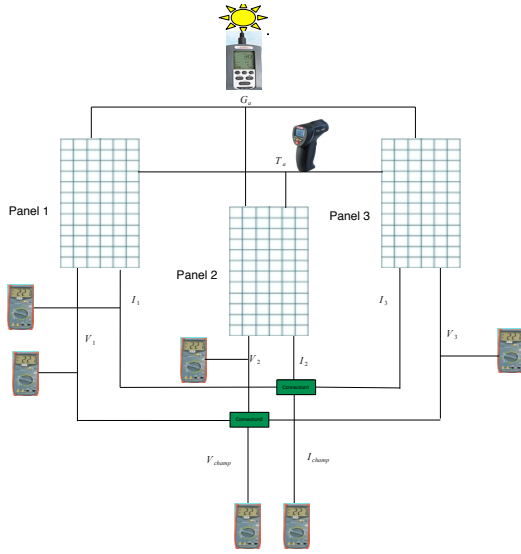

Fig. 2. A photovoltaic system composed of 3 panels.

This system is basically composed of tree main panels Photowatt PW1650 each: $panel_1$, $panel_2$ and $panel_3$. The panels are connected in series. The variables of $panel_1$ are ambient irradiation, denoted $panel_1.G_a$, the ambient temperature denoted $panel_1.T_a$, the cell temperature denoted $panel_1.T_c$, the panel current denoted $panel_1.I$, the panel voltage denoted $panel_1.V$, the short circuit current $panel_1.I_{sc}$ and the open circuit voltage denoted $panel_1.V_{oc}$. The model corresponding to the normal behavior (the ok mode) of $panel_1$ is given by the constraints k_{1-1}, k_{1-2}, k_{1-3}, k_{1-4} and k_{1-5}. In these constraints, the mentionned variables of $panel_1$ are indicated, for short, by: G_a, T_a, T_c, I_1, V_1, I_{sc_1} and V_{oc_1}.

$$
\begin{aligned}
&(ok(panel_1), k_{1-1} : I_{sc_1} = 5.3 \times G_a) \\
&(ok(panel_1), k_{1-2} : T_c = T_a + 0.031 \times G_a) \\
&(ok(panel_1), k_{1-3} : V_{oc_1} = 0.6 - 0.0023(T_c - 25)) \\
&(ok(panel_1), k_{1-4} : V_{t_1} = 0.000345(273 + T_c)) \\
&(ok(panel_1), k_{1-5} : I_1 = \ldots \\
&I_{sc_1}(1 - exp((V_1 - 72 \times V_{oc_1} + 0.45 \times I_1)/72V_{t_1})))
\end{aligned}
\tag{4}
$$

By replacing the constraints k_{1-1}, k_{1-2}, k_{1-3}, k_{1-4} in the constraint k_{1-5}, the following constraint is obtained:

$$
\begin{aligned}
&k_1 : I_1 = 5.3 \times G_a(1 - exp((V_1 - 72(0.6 - \ldots \\
&0.0023(T_a + 0.031 \times G_a - 25)) + 0.45 \times I_1) \ldots \\
&/72(0.000345(273 + T_a + 0.031 \times G_a)))).
\end{aligned}
\tag{5}
$$

For a given measure of G_a, the photocurrent can be calculated by the following relationship:
$I_{ph} = \frac{G}{G_{STC}}[I_{ph,STC} + \alpha(T_c + T_{STC})]$ where:
G: irradiation received by the photovoltaic cell
T_c: the cell temperature
α: coefficient of temperature of the short circuit current
G_{STC}: the irradiation at standard condition $[1000W/m^2]$

T_{STC}: the temperature at standard condition $[25^C]$
$I_{ph,STC}$: the photocurrent at standard condition

In the same, The model corresponding to the normal behavior (the ok mode) of $panel_2$ and $panel_3$ is given by the constraints k_2 and k_3.

$$
\begin{aligned}
&k_2 : I_2 = 5.3 \times G_a(1 - exp((V_2 - 72(0.6 - \ldots \\
&0.0023(T_a + 0.031 \times G_a - 25)) + 0.45 \times I_2) \ldots \\
&/72(0.000345(273 + T_a + 0.031 \times G_a)))). \\
&k_3 : I_3 = 5.3 \times G_a(1 - exp((V_3 - 72(0.6 - \ldots \\
&0.0023(T_a + 0.031 \times G_a - 25)) + 0.45 \times I_3) \ldots \\
&/72(0.000345(273 + T_a + 0.031 \times G_a)))).
\end{aligned}
\tag{6}
$$

They do not give any information about the relation between the panels. The connections between panels have to be taken into account. The same current passes through $panel_1$, $panel_2$ and $panel_3$. Also, the total voltage V is the sum of V_1, V_2 and V_3. These connections are expressed by the constraints k_4, k_5, k_6 and k_7.

$$
\begin{aligned}
&(ok(connection1), k_4 : I_1 = I_2) \\
&(-(connection_1), k_5 : I_2 = I_3) \\
&(-(connection_1), k_6 : I_3 = I_{champ}) \\
&(ok(connection_2), k_7 : V_{champ} = \ldots \\
&V_1 + V_2 + V_3)
\end{aligned}
\tag{7}
$$

(7) represents the connection model for the system. This model has to be deduced from data provided by the installer. The sensors are modeled by the following constraints:

$$
\begin{aligned}
&(ok(sensor_1), k_8 : G_a = \tilde{G}_a \\
&(ok(sensor_2), k_9 : T_a = \tilde{T}_a \\
&(ok(sensor_3), k_{10} : I_1 = \tilde{I}_1 \\
&(ok(sensor_4), k_{11} : V_1 = \tilde{V}_1 \\
&(ok(sensor_5), k_{12} : V_2 = \tilde{V}_2 \\
&(ok(sensor_6), k_{13} : V_3 = \tilde{V}_3 \\
&(ok(sensor_7), k_{14} : I_{champ} = \tilde{I_{champ}} \\
&(ok(sensor_8), k_{15} : V_{champ} = \tilde{V_{champ}}
\end{aligned}
\tag{8}
$$

Requesting models from component database of the photovoltaic system, permits to get component models (5) and (6) together with the observation models (8). The photovoltaic system installer provides data that make it possible to establish (7).

The set of all the elementary models can be decomposed into several sets:

- E^{comp} that contains the elementary models with modes and without dataflows (for example, the constraint sets 5, 6 and 7).
- E^{obs} that contains the elementary models with modes and with dataflows, which contain terminal constraints (for example, the constraint set 8).

IV. DESIGN OF TESTABLE SUBSYSTEMS (**TSS**S) FOR FAULT DIAGNOSIS

The goal of a diagnosis system is to detect abnormal behaviors to avoid failures, breakdowns and damages and to propose hypotheses about the possible faulty components.

143

The faults in a physical system can be diagnosed by checking the consistency between knowledge represented by a behavior model and observations. In this research, the diagnosis system is a two-stage process: detection and isolation stages. In the first stage, the detection tests are computed from the model of the system to be diagnosed. This model is composed of a set of behavioral constraints, a set of terminal constraints containing dataflows. The principle used for testing is to form testable subsets of behavioral and terminal constraints. In the following, the testable subsets of constraints is called as testable subsystems (or TSSs). One can notice the equivalence between a detection test and a TSS.

In the second stage, a fault diagnosis analysis is performed using the symptoms generated during the previous stage.

Generally speaking, each detection test is related to a subset of behavioral constraints.

According to expressions (1), (2) and (3), a testable subsystem can be written as:

$$\bigcup_{i \in T} (ok(c_i), k_i(obs, V_i) = 0) \tag{9}$$
$$\leftrightarrow \quad \left(\bigwedge_{i \in T} ok(c_i), \bigcup_{i \in T} k_i(obs, V_i) = 0 \right)$$

By definition, if (9) is testable subsystems, it means that it exists a constraint k_T, depending only on data flows, such as $k_T(obs) = 0 \leftrightarrow \bigcup_{i \in T} k_i(obs, V_i) = 0$. Therefore, the test is given by:

$$\left(\bigwedge_{i \in T} ok(c_i), k_T(obs) = 0 \right) \tag{10}$$

In the scientific literature, there are several methods for finding all the testable subsystems: the more general ones rely on a structuro-behavioral modeling. A method based on a bipartite graph approach has been used in ([3]), ([8]), ([25]). It makes it possible to compute the testable subsystems. Another method has been proposed in ([26]). This method is based on the retraction of sensors measuring the variables of a system. A method based on the Dulmage-Mendelsohn decomposition ([9]) has been proposed in ([15]). A general method for finding all the possible testable subsystems has been proposed in ([21]). This method is based on a structural analysis. It provides the constraints that have to be used for the design of each detection test and manages situations where constraints contain non deductible variables.

V. New method for symptom generation by using MILP problem et GLPK solver

The symptoms generation from the testable subsystems is a very important step for the fault diagnosis. In literature, there are several methods for symptom generation. The value propagation ([1]) and the constraints propagation ([24] are efficient methods for symptom generation. But, most approaches require the definition of propagation heuristics that depend on the problem to be solved whereas for fault diagnosis, general approaches are searched

In this section, a new method for symptom generation is presented. The main idea is to transform the constraints
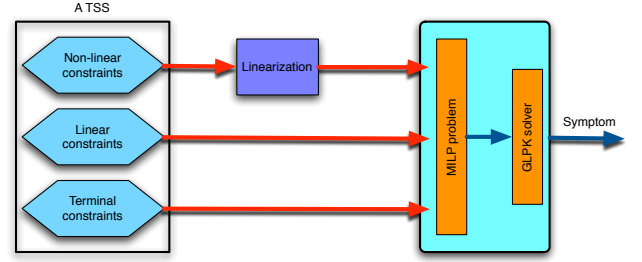

Fig. 3. Symptom generation by using the Milp problem and GLPK

of each TSS into a MILP Problem i.e a regular constraint satisfaction problem and then to use an MILP solver that check if each TSS has a solution or not. If the solver gives a solution then the test resulting from a TSS is positive and consequently, the behavior is consistent with reference model. If the solver does not find any solution then the test resulting from the TSS is negative and consequently, there is a inconsistency between observation and reference model.

MILP problems (Mixed Integer Linear Programming) make it possible to solve the linear or linearized problems. In order to formalize a MILP problem, a linear model describing the behavior of the system has to be formulated. This mathematical model has been implemented automatically in Java.

In order to solve the formalized problem, a MILP solver is needed. Several solvers are available:

- the GLPK solver (open-source software). This solver is suitable for medium problems.
- the CPLEX solver. This solver is suitable for great size problems

In this paper, the GLPK solver has been used. Figure 3 shows the necessary steps for symptom generation.

A. Linearization of a decreasing curve

The model of the photovoltaic system presented in figure 2 contains linear constraints (see relationships 7) and non-linear constraints (see relationships 5 and 6). Because the MILP can only manages linear or linearized problems, the constraints 5 and 6 have to be linearized (see figure 3). Because the curve of each panel is a decreasing curve, this section details the linearisation method of a decreasing curve (see figure 4).

Let consider a point (x, y) on the curve. The following constraints can be written:

$$\begin{array}{ll} x_1 \leq x \leq x_2 \leftrightarrow & y = y_1 - \frac{y_1 - y_2}{x_2 - x_1}(x - x_1) \\ x_2 \leq x \leq x_3 \leftrightarrow & y = y_2 - \frac{y_2 - y_3}{x_3 - x_2}(x - x_2) \end{array} \tag{11}$$

with $dom(x) = [x_1, x_3]$ and $dom(y) = [y_3, y_1]$.

By taking into account the variables δ_1 and δ_2, the constraints 11 can be rewritten by the following equations:

$$\begin{array}{ll} (\delta_1 = 1) \wedge (\delta_2 = 0) \leftrightarrow & y = y_1 - \frac{y_1 - y_2}{x_2 - x_1}(x - x_1) \\ (\delta_1 = 1) \wedge (\delta_2 = 1) \leftrightarrow & y = y_2 - \frac{y_2 - y_3}{x_3 - x_2}(x - x_2) \end{array}$$
$$\tag{12}$$

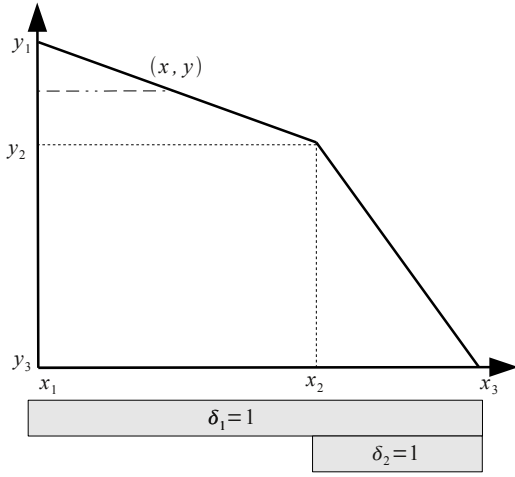Fig. 4.   A decreasing curve

Fig. 5.   Metamodel structure of the optimisation problem

or:

$$\begin{array}{ll} \delta_1 - \delta_2 = 1 \leftrightarrow & y = y_1 - \frac{y_1 - y_2}{x_2 - x_1}(x - x_1) \\ \delta_1 + \delta_2 = 2 \leftrightarrow & y = y_2 - \frac{y_2 - y_3}{x_3 - x_2}(x - x_2) \end{array} \quad (13)$$

Constraints (13) can be rewritten in a MILP format:

$$\begin{array}{l} (y_1 - y) - \frac{y_1 - y_2}{x_2 - x_1}(x - x_1) \le \quad (y_1 - y_3)(1 - \delta_1 + \delta_2) \\ (y_1 - y) - \frac{y_1 - y_2}{x_2 - x_1}(x - x_1) \quad \ge \frac{(y_2 - y_1)(x_3 - x_1)}{x_2 - x_1}(1 - \cdots \\ \delta_1 + \delta_2) \\ (y_2 - y) - \frac{y_2 - y_3}{x_3 - x_2}(x - x_2) \le \quad (y_1 - y_3)(2 - \delta_1 - \delta_2) \\ (y_2 - y) - \frac{y_2 - y_3}{x_3 - x_2}(x - x_2) \ge \quad \frac{(y_3 - y_2)(x_3 - x_2)}{x_3 - x_2}(2 - \cdots \\ \delta_1 - \delta_2) \end{array} \quad (14)$$

In order to describe the variables δ_i, the following constraints are written:

$$\begin{array}{ll} \delta_1 = 1 \leftrightarrow & x_1 \le x \\ \delta_2 = 1 \leftrightarrow & x_2 \le x \end{array} \quad (15)$$

These constraints can be rewritten in a MILP format:

$$\begin{array}{l} x_1 - x \le 0 \\ x_1 - x \ge (x_1 - x_3)\delta_1 \\ x_2 - x \le (x_2 - x_1)(1 - \delta_2) \\ x_2 - x \ge (x_2 - x_3)\delta_2 \end{array} \quad (16)$$

Because all the values of δ_i are not possible, the following constraints have to be added:

$$\begin{array}{l} \delta_2 \le \delta_1 \\ \delta_1 + \delta_2 \ge 1 \end{array} \quad (17)$$

Relationships (14), (16) and (17) represent the linearized equations of the decreasing curve.
This linearization approach has applied for equations (5) and (6).

B. MILP metamodel

A metamodel can be defined as a language from which a model can be generated. The figure 5 shows all the necessary classes for the model construction.
The computer implementation of the metamodel is done without user intervention, but obviously, the user should complete the metamodel with the necessary data. This is made easy by using external files of type xml.

C. GNU Linear Programming Kit (GLPK)

GLPK is free, open source software. It makes possible to solve mathematical programs. Specifically, it solves linear programs (**LP**) via revised simplex method and primal-dual interior point method. It also solves linear mixed-integer programs (MILP) via branch-and-bound algorithm, together with advanced cut routines.
GLPK is intended to solve large-scale problems. It provides an optional presolver, which transforms the problem into one that has better numerical properties for the simplex algorithm; this is particularly useful for large-scale problems. GLPK was developed, and is maintained, by Andrew Makhorin, Department for Applied Informatics, Moscow Aviation Institute.

VI. DIAGNOSIS ANALYSIS

The diagnosis analysis aims at computing all the possible diagnoses from the symptoms provided by the detection tests. Different kinds of approaches may be used to analyze the symptoms. The decision tree approach is a general approach to analyze symptoms but not relevant for complex systems ([5]). Case-based reasoning compares the current symptoms with patterns coming from a knowledge database in order to retrieve a similar situation with the current one ([16]). This approach is not exhaustive and does not benefit of the knowledge about the behavioral modes of the system components. The signature based approach relies implicitly on a no exoneration assumption ([20]), which may lead to erroneous diagnoses. The bridge approach shows that the consistency-based reasoning can be used to analyze the symptoms coming from the detection tests depicted by the involved modes ([18]), ([20]). The bridge approach presented in ([20]) shows that a detection test based on a TSS, can be depicted as a set of modes of related components.

Depending on the way of testing, test (10) yields to:

$$\bigwedge_{i \in T} ok(c_i) \leftrightarrow \forall obs \in OBS, k_T(obs) = 0 \qquad (18)$$

or

$$\bigwedge_{i \in T} ok(c_i) \rightarrow \forall obs \in \mathcal{O}, k_T(obs) = 0 \qquad (19)$$

where $\mathcal{O} \subset OBS$.

Practical diagnosis relies on expressions like (19). For each, unsatisfied test $k_T(obs) \neq 0$ for a given $obs \in OBS$ yields:

$$k_T(obs) \neq 0 \rightarrow \bigvee_{i \in T} cfm(c_i) \qquad (20)$$

where the set $\{cfm(c_i); i \in T\}$ is named explanation for T: $Expl(T)$.

If a test is satisfied, $k_T(obs) = 0$, it does not lead to any formal conclusion.

If a test has been performed for all the possible observations OBS, equation (20) becomes:

$$\bigwedge_{i \in T} ok(c_i) \leftrightarrow \forall obs \in OBS, k_T(obs) = 0 \qquad (21)$$

But it is almost never possible to test all the possible situations OBS because:

- even if the variables are observable, the system has to be controllable to be able to check all the possible OBS
- sensors and actuators may be faulty, so it is not possible to prove that all the situations OBS have been checked.

Therefore, to compute diagnoses, expressions like (20) are used. A diagnosis is thus a conjunction of faulty modes such that it may explain all the current known test explanations. Then, in each diagnosis, the fault mode of every faulty component is taken into account. In other words, consider a set of inconsistent detection tests $\{T_1, \ldots, T_n\}$. Finding a diagnosis consists in searching a conjunction of modes which verifies the following expression:

$$\bigwedge_{T_i} \bigvee_{mode_j \in Expl(T_i)} mode_j \qquad (22)$$

The second method which can be used to sort the diagnosis exploits the priori reliability of components: $p(mode(c_i) = ok(c_i))$. Then, because there are only two considered modes, $p(mode(c_i) = cfm(c_i)) = 1 - p(mode(c_i) = ok(c_i))$. Then, if $d = \bigwedge_i cfm(c_i)$ is a diagnosis, its a priori probability is given by:

$$\mu_T^a(d) = \prod_i (1 - p(mode(c_i) = ok(c_i))) \qquad (23)$$

VII. DIAGNOSIS STRATEGY OF PHOTOVOLTAIC SYSTEM

In this section, the diagnosis strategy of photovoltaic systems is presented.
The list of the necessary sensors for diagnosis analysis are:

1) current sensor
2) voltage sensor
3) temperature sensor
4) pyranometer
5) laptop and diagnosis software

Suppose that the variables to be measured are determined, the following steps are necessary for fault detection and isolation:

1) the sensors measuring the variables G_a and T_a are installed in order to plot the curve of each photovoltaic panel in function of I and V. This step is very important to linearize the non-linear constraints modelling these panels.
2) the other sensors are consecutively installed (one by one, two by two....)
3) measurements of installed sensors are collected
4) these measurements are entered into the diagnosis software
5) the diagnosis software finds the testable subsystems
6) this software generates the symptoms by using the testable subsystems
7) this software finds all the possible diagnoses
8) if the goal is satisfied then the analysis diagnosis is finish, else, the steps 2, 3, 4, 5, 6 et 7 are repeated

Consequently, the diagnosis software provides all the possible diagnoses.

VIII. APPLICATION

In this section, the iterative diagnosis process presented in this paper is applied on the photovoltaic system (Fig. 2) in order to satisfy the following specifications: the faults on panels have to be isolated and the faults on connections have to be at least detected. The necessary sensors are already determined. The necessary sensors are already determined. The photovoltaic system with sensors is modelled by the constraint sets (5, 6, 7 and 8).
By applying the diagnosis strategy presented in section VII, all the sensors already determined have to be installed. Suppose that the measured values are: $T_a = 330$, $I_1 = 3.40$, $V_1 = 21.6851$, $V_2 = 28.9135$, $V_3 = 28.9135$, $V_{champ} = 79.5121$, $I_{champ} = 3.40$. Suppose also that the value of photocurrent related to the measured irradiation G_a is $I_{ph} = 3.77$.
By applying one of methods of testable subsystem design, 8 minimal testable subsystems (**TSS**s) are obtained. This **TSS**s are:

- $TSS_1 : k11, k3, k5, k8, k4, k15, k14$
- $TSS_2 : k12, k6, k5, k8, k4$
- $TSS_3 : k12, k11, k3, k6, k15, k14$
- $TSS_4 : k11, k10, k9, k7, k13$
- $TSS_5 : k11, k3, k10, k2, k5, k15, k14$
- $TSS_6 : k12, k6, k10, k2, k5, k15, k14$
- $TSS_7 : k9, k1, k8, k15, k14$
- $TSS_8 : k10, k2, k8, k4, k15, k14$

By applying the method of symptom generation, the following results are obtained: the tests resulting from the testable subsystems $SST1$, $SST2$, $SST3$, $SST4$, $SST5$, $SST6$, $SST8$ are positive (observations are consistent with models) and the test resulting from $SST7$ is negative (observations are inconsistent with models).
The symptoms found are used for diagnosis analysis. A

diagnosis software DXLAB (developed in the laboratory G-SCOP) has been used to find out all the possible diagnoses. These symptoms are declared in DXLab as in figure 6.

Once symptoms are declared, the diagnoses shown in figure

```
abed@ubuntu-pioneer:~$ dxlab
Welcome into DXLAB version 0.2.2
[1] test k3 k4 k5 k8 k11 k14 k15 positive
test t001 [enabled], symptom:0.0 % (positive)
   k3, k4, k5, k8, k11, k14, k15
[2] test k4 k5 k6 k8 k12 positive
test t002 [enabled], symptom:0.0 % (positive)
   k4, k5, k6, k8, k12
[3] test k3 k6 k11 k12 k14 k15 positive
test t003 [enabled], symptom:0.0 % (positive)
   k3, k6, k11, k12, k14, k15
[4] test k7 k9 k10 k11 k13 positive
test t004 [enabled], symptom:0.0 % (positive)
   k7, k9, k10, k11, k13
[5] test k2 k3 k5 k10 k11 k14 k15 positive
test t005 [enabled], symptom:0.0 % (positive)
   k2, k3, k5, k10, k11, k14, k15
[6] test k2 k5 k6 k10 k12 k15 k16 positive
test t006 [enabled], symptom:0.0 % (positive)
   k2, k5, k6, k10, k12, k15, k16
[7] test k1 k8 k9 k14 k15 negative
test t007 [enabled], symptom:100.0 % (negative)
   k1, k8, k9, k14, k15
[8] test k2 k4 k8 k10 k14 k15 positive
test t008 [enabled], symptom:0.0 % (positive)
   k2, k4, k8, k10, k14, k15
```

Fig. 6.   The sysmptom declaration

7 were obtained.

Suppose that the sensors are reliable (faults on sensors are

```
[9] show diagnoses
#0 (score:100.0 %, apriori:10.0%, vote:100.0%)
   k1 is not ok

#1 (score:100.0 %, apriori:10.0%, vote:87.5%)
   k9 is not ok

#2 (score:100.0 %, apriori:10.0%, vote:62.5%)
   k8 is not ok

#3 (score:100.0 %, apriori:10.0%, vote:50.0%)
   k14 is not ok

#4 (score:100.0 %, apriori:10.0%, vote:37.5%)
   k15 is not ok
```

Fig. 7.   The possible diagnoses

not considered), the diagnosis software shows that a fault has occured in the panel 1. Consequently, the problem has been isolated and the specifications are satisfied.

IX. CONCLUSIONS

This paper points out how iterative fault diagnosis system can be achieved. The necessary steps are detailed. The modelling aspect is firstly discussed. The information required to depict a system to be diagnosed has been presented. Then, all the possible testable subsystems are generated and the way of testing them has been presented. Testable subsystems are transformed into a MILP problem and a GLPK solver is used to illustrate a possible implementation. Finally, a photovoltaic system that makes it possible to clarify the iterative diagnosis process is presented.

REFERENCES

[1] Krzysztof Apt. *Principles of Constraint Programming*. Cambridge University Press, 2003.

[2] M. Blanke, M. Kinnaert, J. Lunze, and M. Staroswiecki. *Diagnosis and Fault-tolerant Control*. Springer-Verlag, 2006.

[3] M. Blanke, M. Kinnaert, and M. Staroswiecki. *Diagnosis and fault tolerant control*. Springer, 2003.

[4] J.P. Cassar, M. Staroswiecki, and V. Cocquempot. Optimal residual design for model-based fault detection and isolation. In *3rd European Control Conference ECC95*, Roma, Italy, 1995.

[5] M. Chen, Alice X. Zheng, Jim Lloyd, Michael I. Jordan, and Eric Brewer. Failure diagnosis using decision trees. volume 0, pages 36–43, Los Alamitos, CA, USA, 2004. IEEE Computer Society.

[6] L. Chittaro and R. Ranon. Hierarchical model-based diagnosis based on structural abstraction. *Artificial Intelligence*, 47(1-2):147–182, 2004.

[7] J. De Kleer and B. C. Williams. Diagnosing multiple faults. *Artificial Intelligence*, 32:97–130, 1987.

[8] P. Declerck and M. Staroswiecki. Characterization of the canonical components of a structural graph for fault detection in large scale industrial plants. In *European Control Conference*, pages 298–303, Grenoble, France, 1991.

[9] A. L. Dulmage and N. S. Mendelsohn. A structure theory of bi-partite graphs of finite exterior extension. *Transactions of the Royal Society of Canada*, 53(III):1–13, 1959.

[10] P. Frank. Analytical and quantitative model-based fault diagnosis- a survey and some new results. *European Journal of Control*, 2:6–28, 1996.

[11] E. Frisk and M. Krysander. Sensor placement for maximum fault isolability. In *The 18th International Workshop on Principles of Diagnosis (DX-07)*, 2007.

[12] J. Gertler. Analytical redundancy methods in fault detection and isolation. In *International Conference on fault diagnosis Tooldiag'93*, 1993.

[13] W. Hamscher, L. Console, and J. De Kleer. *Readings in Model-Based Diagnosis*. Morgan Kaufmann, 1992.

[14] R. Isermann. Model-based fault detection and diagnosis-status and applications. In *16th symposium on Automatic Control in Aerospace (ACA'2004)*, 2004.

[15] M. Krysander, J. Åslund, and M. . Nyberg. An efficient algorithm for finding minimal overconstrained subsystems for model-based-diagnosis. *IEEE Transactions on systems, Man, and Cybernetics-Part A: Systems and Humans*, 38(1):197–206, 2008.

[16] D. McSherry. Interactive case-based reasoning in sequential diagnosis. *Applied Intelligence*, 14(1):65–76, 2001.

[17] M. Nyberg and M. Krysander. Combining ai, fdi, and statistical hypothesis-testing in a framework for diagnosis. In *IFAC Safeprocess'03*, Washington, U.S.A., 2003.

[18] M. Nyberg and Mattias Krysander. Combining AI, FDI, and statistical hypothesis-testing in a framework for diagnosis. 2003.

[19] R.J. Patton and J. Chen. A review of parity space approaches to fault diagnosis. In *IFAC SAFEPROCESS Symposium*, Baden-Baden, 1991.

[20] S. Ploix, S. Touaf, and J. M. Flaus. A logical framework for isolation in fault diagnosis. In *SAFEPROCESS'2003*, Washington D.C., U.S.A., 2003.

[21] S. Ploix, A.A. Yassine, and J.M. Flaus. An improved algorithm for the design of testable subsystems. In *The 17th IFAC WORLD CONGRESS*, Seoul, Korea, 2008.

[22] R. Reiter. A theory of diagnosis from first principles. *Artificial Intelligence*, 32:57–95, 1987.

[23] H. Ressencourt, L. Travé-Massuyès, and J. Thomas. Hierarchical modelling and diagnosis for embedded systems. In *SAFEPROCESS'2006*, aug. 30th-sep. 1st 2006.

[24] S. Russell and P. Norvig. *Artificial Intelligence, a modern approach, 2nd ed.* Prentice Hall, 2003.

[25] M. Staroswiecki and P. Declerck. Analytical redundancy in nonlinear interconnected systems by means of structural analysis. In *IFAC AIPAC'89 Symposium*, Nantes, France, 1989.

[26] L. Travé-Massuyès, T. Escobet, and X. Olive. Diagnosability analysis based on component supported analytical redundancy relations. *IEEE transactions on Systems, Man, And Cybernetics - Part A: Systems and Humans*, 36(6):1146–1160, November 2006.

[27] A. Yassine, S. Ploix, and J.-M. Flaus. A method for sensor placements taking into account diagnosability criteria. *Applied Mathematics and Computer Science*, 18(4), 2008.

[28] A. Yassine, S. Ploix, and J.-M. Flaus. Structural approach for sensor placement with diagnosability purpose. In *The 17th IFAC World Congress*, Seoul, Corea, 2008.

APPLICATION OF NONLINEAR MODEL PREDICTIVE CONTROL BASED ON DIFFERENT MODELS TO BATCH POLYMERIZATION REACTOR

Sevil Cetinkaya[a,*], **Duygu Anaklı[a]**, Zehra Zeybek[b], Hale Hapoğlu[b], Mustafa Alpbaz[b]

[a] *Cumhuriyet University, Department of Chemical Engineering, 58140, Sivas, Turkey*

[b] *Ankara University, Department of Chemical Engineering, 06100 Tandogan, Ankara,*

Turkey

* *e-mail:* cetinkaya.sevil@gmail.com

Phone: +905055671933, Fax: +903462191179

ABSTRACT

In the present work, at the previously determined optimal conditions, to control batch polymerization reactor, Linear Generalized Predictive Control (LGPC) and Nonlinear Generalized Predictive Control (NLGPC) algorithm were utilized. Several system models were applied to the control algorithms. The effect of different optimal conditions has been examined on monomer conversion, average viscosity molecular weight and chain length. At the same operating conditions of LGPC and NLGPC temperature control was used for comparison. According to the experimental results, the performance of NLGPC was obtained well than LGPC control method. In addition, the results denoted that the NLGPC control performances depend on different models and the optimum conditions.

Keywords: Styrene polymerization, Generalized Predictive Control, Nonlinear Model Predictive Control

1. INTRODUCTION

Control of polymerization reactors is often difficult and sensors to provide on-line measurement of polymer properties are generally not available (Altınten, Erdogan, Hapoglu, and Alpbaz 2003; Altınten, Erdogan, Hapoglu, Alıev, and Alpbaz 2006; Cetinkaya 1996). The most significant task for a polymerization reactor control strategy is to maintain the major design because of having complex and nonlinear reaction and operational variables like product quality is also important to preserve smooth and stable operation.

Physical, chemical and mechanical properties of polymers are generally closely related with their molecular weights. But the weight of all polymer molecules within a polymer sample is not equal to each other. For this reason, the molecular weight of polymers that were determined in any way shows average number not the exact values. The full molecular weight distribution (MWD) of a specific polymer and ratio of moments of this distribution, such as the number average ($\overline{M}_n$) or the weight average ($\overline{M}_w$) molecular weight, indicate the mechanical properties of the polymer. Initial initiator concentration and temperature are the primary control ways to influence the molecular weight of a polymer produced in free radical polymerization (Barner-Kowollik and Davis 2001). The temperature change has been observed that has great influence on the kinetics of polymerization process, and physical properties and quality of produced polymer. The main objective of the temperature control of polymerization reactor is to remove a great amount of heat from the exothermic reaction to achieve the desired number average chain and a desired conversion in a minimum time (Yuce, Hasaltun, Erdogan, and Alpbaz 1999). Then, reactor temperature should be controlled effectively to satisfy the desired polymer quality.

Various control methods have been applied both theoretically and experimentally to the systems at constant and changing set points (Zeybek, Cetinkaya, Hapoglu, and Alpbaz 2006; Seki, Ogawa, Ooyama, Akamatu, Ohshima, and Yang 2001). Lewis, Nguyen, and Cohen (2007), highlighted to the effect of initiator amount on the radical polymerization. In free-radical polymerization, the reaction temperature and both the initiator and the chain-transfer agent's concentrations are usually chosen as controlled variables. These variables can also affect the rate of polymerization and the molecular weight of the polymer (Ponnuswamy, Shah, and Kiparissides 1987).

Generalized Predictive Control (GPC) algorithm is commonly used in polymerization reactors (Ozkan, Hapoglu, and Alpbaz 1998; Yuce (Çetinkaya) 2001). But, NLGPC algorithm was rarely practiced with different models and at constant temperature in a batch polymerization reactor. Zeybek, Cetinkaya, Hapoglu, and Alpbaz (2006) developed the generalized delta rule

(GDR) algorithm with generalized predictive control (GPC) for two different changing temperature path, and used experimentally in a batch polymerization reactor. The application of a GDR for system identification is a feasible alternative when model equations are not known or only historical input-output data are available.

In this study, Linear Generalized Predictive Control (LGPC) and Nonlinear Generalized Predictive Control (NLGPC) algorithm were used. Two models of NLGPC were applied to the control algorithms and compared with LGPC and both of constant and changing temperatures. According to the experimental results obtained at the constant and changing set points, the performance of NLGPC in terms of efficiency were also obtained better than LGPC control method. Desired monomer conversion, average viscosity molecular weight and chain length were affected by the controller performance. In addition, the results denoted that the NLGPC control performances depend on different models and the optimum conditions.

2. MATHEMATICAL MODEL OF THE REACTOR

In polymerization, physical, chemical and mechanical properties are based on polymer quality of the final product, which means molecular and structural characteristics of a polymer. The ability of a mathematical model is important to predict exact molecular properties of a polymer manufactured by means of polymerization reactor for optimal production cost in polymer industry.

A basic free radical polymerization mechanism has three fundamental reaction steps: initiation, propagation and termination.

Considering the standard free radical polymerization and assuming constant density, no chain transfer and no gel effects, and using quasi steady-state and long chain approaches for live radials, the equations for monomer conversion, initiator conversion and the dimensionless zeroth moment of the molecular weight distribution are used as follows (Chen and Jeng 1978):

$$\frac{dX}{dt} = k_1 (1-c)^{0.5} \frac{(1-x)}{g} \tag{1}$$

$$\frac{dc}{dt} = k_d (1-c) \tag{2}$$

$$\frac{dq_0}{dt} = (2-c)\alpha k_d (1-c) \tag{3}$$

With the initial conditions, $X(0) = c(0) = q_0 = 0$

$$k_1 = k_p \left(\frac{2 f k_d I_0}{k_t} \right)^{\frac{1}{2}} = I_0^{\frac{1}{2}} A_1 \exp(-E_1 y) \tag{4}$$

$$X = 1 - \frac{M}{M_0}; \quad c = 1 - \frac{I}{I_0}; \quad g = 1;$$

$$\alpha = \frac{fI_0}{M_0}; \quad v = \frac{k_{tc}}{k_{td}} = 1 \tag{5}$$

$$E_1 = E_p + \frac{E_d}{2} - \frac{E_t}{2}$$

$$A_1 = (2f)^{\frac{1}{2}} A_p A_d^{\frac{1}{2}} A_t^{-\frac{1}{2}} \tag{6}$$

The number average chain length can be given as:

$$L_n = \frac{X}{q_0} \tag{7}$$

The control variables in the isothermal batch jacketed reactor are taken as reaction temperature and initial initiator concentrations. In order to acquire optimal operating conditions, the method of Lagrange's Multiplier and Hamiltonian Maximum Principle are used and the optimum conditions are given in the Table 1.

3. DESIGN OF NONLINEAR GPC

Consider the control of a linear state space process with the output corrupted by noise. The system has the ARMAX representation:

$$\sum_{i=0}^{n} a_i y(t-i) = \sum_{i=0}^{m} b_i u(t-\tau-i) + \sum_{i=0}^{n} c_i v(t-i) \tag{8}$$

where $\tau \geq 1$ is a delay.

All zeros of $B(q^{-1}) = (b_0 + b_1 q^{-1} + ... + b_m q^{-m})$ and $C(q^{-1}) = (1 + c_1 q^{-1} + ... + c_n q^{-n})$ are strictly inside the unit circle, and $A(q^{-1}) = (1 + a_1 q^{-1} + ... + a_n q^{-n})$ is monic. The objective is to cancel the noise by manipulating the control $u(t)$. Only $y(t)$ is measured. Such problems arise in signal processing and numerous control applications. The control objective is achieved by the feed-forward law:

$$u(t) = \frac{1}{\beta_0} \left[\sum_{i=1}^{1} \alpha_i y(t+1-i) - \sum_{i=1}^{m} \beta_i u(t-i) - \sum_{i=0}^{d-1} f_i v(t+1+i) \right] \tag{9}$$

Where the coefficients f_i, α_i and β_i are found by solving a Diophantine identity:

$$C(q^{-1}) = F(q^{-1})A(q^{-1}) + q^{-d}G(q^{-1})$$

For F and G and setting $\alpha(q^{-1}) = G(q^{-1})$ and $\beta(q^{-1}) = F(q^{-1})B(q^{-1})$. The control u(t) (Eq. (9)) cannot be applied since noise signals are required "d" steps into the future. Stochastic control overcomes the problem by assuming that $v(t)$ is white noise, which

may be generated from the sampling of a Brownian motion. We may then exploit the fact that $E\{v(t+1)F(t)\}=0$ for $i \geq 1$. The optimal control, in the sense of minimizing the output variance, is then implemented using measured signals Ydstie (1990). The control law that sets the predicted output to zero satisfies:

$$u(t)=\frac{1}{\beta_0}\left[\sum_{i=1}^{1}\alpha_i y(t+1-i)-\sum_{i=1}^{m}\beta_i u(t-i)-\sum_{i=0}^{d-n+1}g_i v_f(t+1+i)\right] \quad (10)$$

Where $\alpha(q^{-1})=G(q^{-1})$ and $\beta(q^{-1})=F(q^{-1})B(q^{-1})$,

$$H(q^{-1})=(1+h_1 q^{-1}+...+h_m q^{-m})=F(q^{-1})C(q^{-1})$$

To establish the GPC algorithm, it is supposed that a model of the linearized plant is expressed in terms of the NARMAX . The most general NARMAX model structure would take the form below, where some function of the previous output $y(k-T)$, $y(k-2T)$ and inputs $u(k-T), u(k-2T)$, would give the current output.

$$y(k)=F[(y(k-T), y(k-2T), y(k-n_y T), u(k-T), u(k-2T),...,u(k-n_u T))] \quad (11)$$

This is a natural model for any sampled dynamic system with inputs at discrete times. It is extremely difficult to determine the function F which fits the input/output data well over a range of operating condition.

Consider the vector error (ε) composed of predicted future system errors $W(t+j)-\hat{y}(t+j)$. W is the references signal. The suggested future control sequence $\{u(t+j)\}$ is chosen by GPC at time t to minimize a cost-function such as

$$J(N_1, N_2, NU, \lambda)=\sum_{J=N_1}^{N_2}\varepsilon^2(t+j)+\lambda\sum_{j=1}^{NU}\Delta U^2(t+j-1) \quad (12)$$

N_1 is the minimum costing horizon,
N_2 is the maximum costing horizon,
N_U is the control horizon, and
λ is the (optional) control weighting
According to GPC strategy, equation (8) can be written as
$$\hat{y}=\overline{G}\Delta\overline{u}+f \quad (13)$$

$$\overline{G}=\begin{bmatrix} g_0 & 0 & \cdots & 0 \\ g_1 & g_0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ g_{N_U-1} & g_{N_U-2} & \cdots & g_0 \\ \vdots & \vdots & & \vdots \\ g_{N_2-1} & g_{N_2-2} & \cdots & g_{N_2-N_U} \end{bmatrix}_{N_2 x N_U}$$

$\overline{G}$ matrix is $N_2 x N_U$ dimensional lower triangular matrix.

4. EXPERIMENTAL SYSTEM

Experimental system is shown in Fig. 1. Water was used as coolant in the jacket and the temperature inside the reactor was measured at each sampling period using thermocouples. A computer with A/D and D/A converters was employed for data acquisition and the control of experimental reactor. In this study, the monomer of styrene was used and benzoyl peroxide was used as initiator. Toluene was chosen as a solvent. The reactor contents were 70% styrene and 30% toluene. The experimental studies were carried out in a cylindrical glass jacketed reactor which has the internal volume of 1600 ml and jacket volume of 980 ml. The temperature of the reactor and the temperatures of the jacket inlet and outlet were measured by thermocouples. The converter modules were connected to adjust the flow rate of the pump and the heat input which is given into the reactor. In this work, the converter module has two outputs (pump and heater) and three inputs (reactor temperature, cooling water inlet temperature and outlet temperature). All A/D connections to the reactor were realized on the converter module using VisiDAQ packet program. Reactor was heated by an immersed heater inside the reactor. In all the experiments, the manipulated variable was accepted as the heat input which was transferred by diver heater to the reactor. It was accepted that enough mixing was provided by the mixer in the reactor.
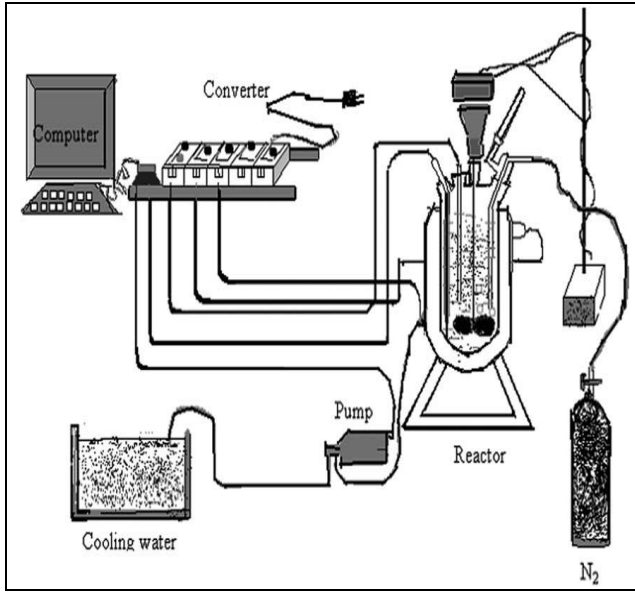
Figure 1: Experimental System

5. RESULTS AND DISCUSSION

The proposed control strategy was deeply tested on a batch polymerization reactor. The optimum reactor temperature set point for minimum polymerization time was obtained at different benylzoyl peroxide initial initiator concentrations. Two of these optimum reactor set points were selected for control purposes. GPC was applied to control the optimal temperature of a styrene polymerization reactor.

FORTRAN program has been used in the simulation of the system. For NLGPC control, models of the system have been taken in the following from (Yuce 2001).

Model1:

$$\left(y(t)-y(t-1)\right)=\frac{b_0 z^{-1}}{1+a_1 z^{-1}+a_2 z^{-2}}\left(U(t)-U(t-1)\right)$$

(14)

$a_1 = -0.5227 \quad a_2 = 0.0001 \quad b_1 = -0.4773$

Model2:

$$\left(y(t)-y(t-1)\right)=\frac{b_0 z^{-1}}{1+a_1 z^{-1}}\left(U(t)-U(t-1)\right) \quad (15)$$

$a_1 = -0.9998 \quad b_1 = 0.0001$

Where $U(t)=u^4(t-1)$

These coefficients (b_0, a_1 and a_2) were found by using a pseudo random binary sequence as the input function. A second order polynomial is sufficient to represent the denominator plant dynamics. The system is defined and the model parameters are calculated using the least squares regression method given as follows. Regression coefficient (R^2) is found as 0.96.

Some experimental studies have been carried out to acquire the quality of polymer product. The effect of different optimal conditions has been examined on monomer conversion, average viscosity molecular weight and chain length. The polymerization reactor temperature was controlled by manipulating the powder to the heater and monitored to see the control performance.

To see the performance of NLGPC control and to compare with LGPC and different models, some experiments have been carried out on the experimental system. To obtain the reaction curve, a step change in the heat input was applied after the system reached to the steady state conditions. The temperature responses to this effect obtained from experimental work are given in Fig 2.
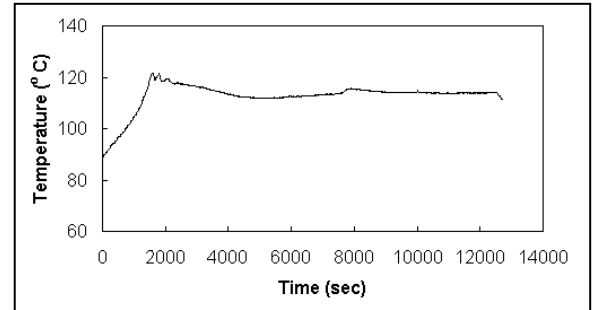
Figure 2: Temperature response of the reactor inlet (I=0.0185 mol l^{-1}, T=89.04°C)

As described in the previous section, the optimum reactor temperature for a minimum polymerization time was obtained at different benylzoil peroxide initial initiator concentrations. For control purposes, two of these optimum reactor constant temperature and paths were selected and shown in Table 1. Experimental control results are shown in Figs. 2-7 for LGPC and NLGPC.

The experimental results show some oscillations due to the fact that this is a batch exothermic reaction with constantly changing conditions and heat generation is not constant during the reaction (Zeybek, Cetinkaya, Hapoglu, and Alpbaz 2006).

As it is seen from Fig 3, 6 and 7, for variable temperature control the fluctuations occurrence are more significant and the frequency is higher than constant temperature control.

At the same operating conditions of LGPC and NLGPC temperature control was used for comparison. An overshoot is observed at the beginning of the reaction for this two control method and then the NLGPC and LGPC controller bring the reactor temperature back to set point but after 6000 seconds LGPC controller continuous as increasing oscillation until the end of the experiment (Fig. 3a).

Figs. 4-5 show the results of experimental temperature response of NLGPC for different models at constant set point. Both of model 1 and model 2 for

NLGPC is very suitable, fast and robust in these applications. Also, desired values chain length and conversion were closely achieved (Table 2).

In Figs. 6-7, the experiments used for two models of NLGPC were carried out using two different temperature pathways which have 89.04 and 92.7 °C of initial temperature respectively shown in Table 1. The chain length of 500, conversion 50% were target in these pathways. As seen in Figs.6-7, at the temperatures best result was reached for model 2 and both chain length and conversion, target was nearly reached (Table 2). The experimental results demonstrated that NLGPC at the constant temperature has been good control performance, because some oscillations in the temperature around the set point are seen due to continuously changing conditions, gel effect and strong nonlinearities (Yuce, Hasaltun, Erdogan, and Alpbaz 1999).

The progression of the manipulated variable (Q) executed by the computer during the experiments is also monitored in Figures 4(b)-7(b).

Table1: Optimal operating conditions used in experimental studies

Run	T_R (°C)	I_0 (mole l^{-1})	M_0 (mole l^{-1})	t_f (s)	X_d (%)	L_{nd} (gmol^{-1})	T_c °C
1	103.8	0.0126	6.092	9036	50	500	21
2	105.5	0.0038	6.092	7440	30	1000	21
3	89.04	0.0185	6.092	12720	50	500	21
4	92.7	0.0150	6.092	10200	50	500	21
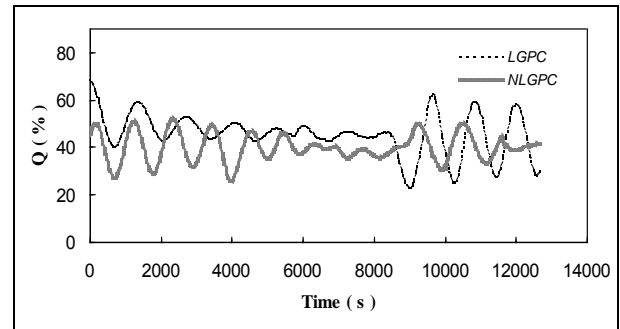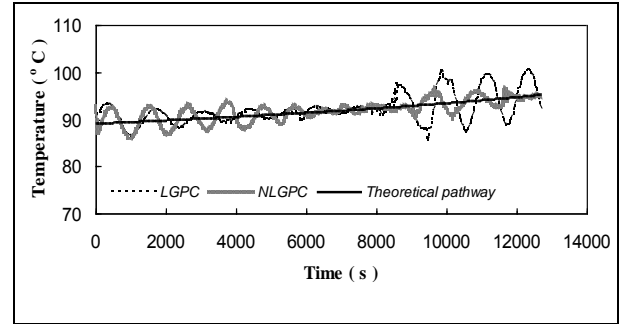
Fig. 3: (a) Temperature response, (b) manipulated variable with time under LGPC and NLGPC respectively ($I_0 = 0.0185$ mole l^{-1}, Model 1)
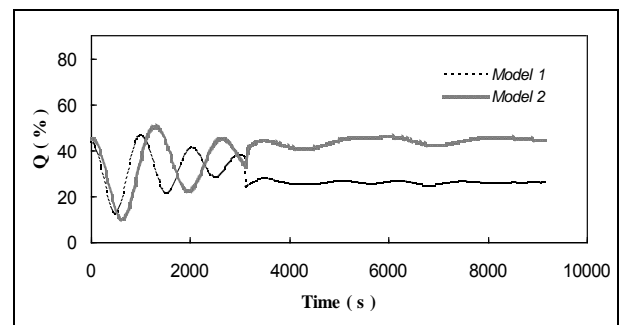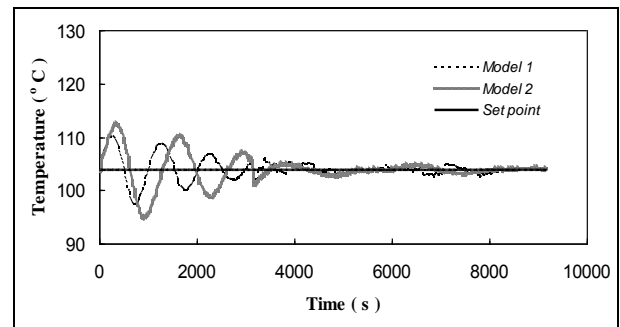
Fig. 4: (a) Temperature response, (b) manipulated variable with time under NLGPC respectively for model 1 and model 2 ($I_0 = 0.0126$ mole l^{-1})
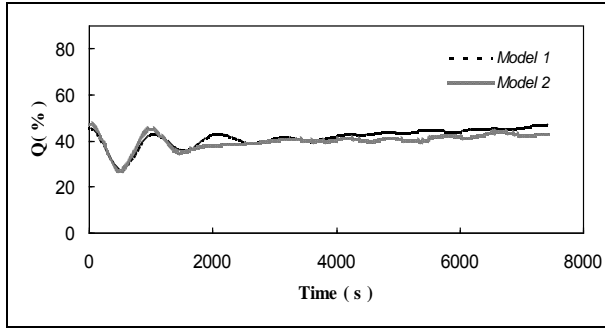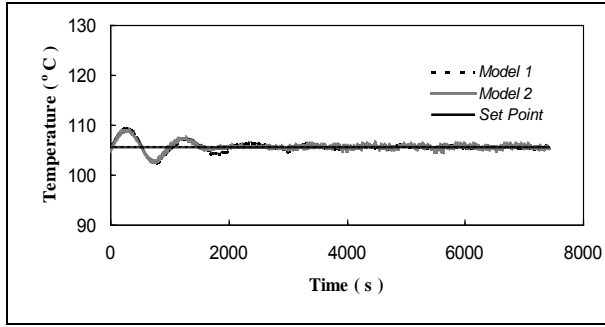
Fig. 5: (a) Temperature response, (b) manipulated variable with time under NLGPC respectively for model 1 and model 2 ($I_0 = 0.0038$ mole l^{-1})

Fig. 6: (a) Temperature response, (b) manipulated variable with time under NLGPC respectively for model 1 and model 2 ($I_0 = 0.0185$ mole l^{-1})
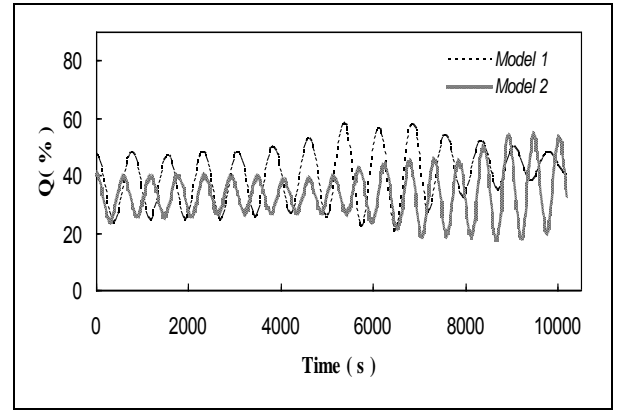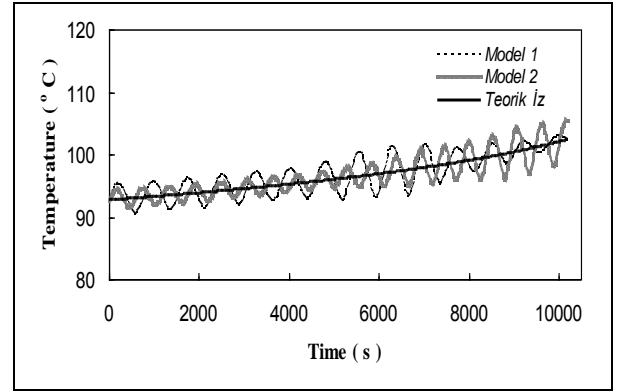
Fig.7: (a) Temperature response, (b) manipulated variable with time under NLGPC respectively for model 1 and model 2 ($I_0 = 0.0150$ mole l^{-1})

153

Table 2: Comparison of experimental results with target values

T (^{o}C)	103.8	105.5	89.04	92.7
I_0 (mole l^{-1})	0.0126	0.0038	0.0185	0.0150
Desired conversion, m*(%)	50	30	50	50
Experimental conversion, m(%) (NLGPC,Model 1)	44	33	71.5	65
Experimental conversion, m(%) (LGPC, Model 1)	-	-	57.8	-
Experimental conversion, m(%) (NLGPC,Model 2)	45.9	29.88	69.95	54.84
Desired chain length, X_n^*	500	1000	500	500
Experimental chain length, X_n (NLGPC,Model 1)	425	940	321	269
Experimental chain length, X_n (LGPC, Model 1)	-	-	204	-
Experimental chain length, X_n (NLGPC,Model 2)	412	955	352	452
Desired average viscosity molecular weight	52000	104000	52000	52000
Experimental average viscosity molecular weight (NLGPC,Model 1)	44253.2	97869	32983.82	27681.35
Experimental average viscosity molecular weight (LGPC, Model 1)	-	-	21249	-
Experimental average viscosity molecular weight (NLGPC,Model 2)	42851.5	99445.3	36324.5	47155.6

6. CONCLUSION

Generalized Predictive Control has been presented in this paper. Linear Generalized Predictive Control (LGPC) and Nonlinear Generalized Predictive Control (NLGPC) algorithm were used to keep the temperature of a jacketed batch polymerization reactor at constant and variable optimal conditions. Good performance is achieved using NLGPC. The experimental results obtained have also confirmed that this control performs well particularly at constant optimal conditions. It is also observed that by the use of NLGPC, the desired values of molecular weight are achieved at the end of the batch. In addition, the performance of NLGPC is better than the LGPC.

APPENDIX A. NOMENCLATURE

A_d, A_p, A_t :Frequency factor for initiator decomposition, propagation and termination respectively, s^{-1}, L mol s^{-1}

a_I : Parameters of A polynomial

b_o : Parameters of A polynomial

c : initiator conversion

E : a polynomial

E_d, E_p, E_t :Activation energies for initiator decomposition, propagation and termination, respectively, kJ kmole^{-1}K^{-1}

f : Initiator efficiency

g : gel effect

G : a polynomial

$\overline{G}$: a matrix

I, I_0 :Initiator concentration, initial initiator concentration, respectively, mole L^{-1}

k_d : Initiator decomposition rate constant

k_p : Propagation rate constant

k_t : Termination rate constant

L_n, L_{nd} : Number average chain length, desired number average chain length

M, M_0 : Monomer concentration, initial monomer concentration, mole L^{-1}

N_1 : The minimum costing horizon

N_2 : The maximum costing horizon.

N_U : The control horizon

q_0 : Dimensionless zeroth moment

T_c, T_{ci}, T_{co} :Average, inlet and outlet coolant temperatures (^{0}C)

t, t_f : Time, polymerization time, s

u(t) : Input variable at time t

v : k_{tc}/k_t , constant

X : Monomer conversion,

y(t) : Output variable at time t

$y(t-1)$: Output variable at time t-1.

Greek symbols

α : Coefficient of momentum

REFERENCES

Altınten, A., Erdoğan, E., Hapoğlu, H., & Alpbaz, M., 2003. Control of a polymerization reactor by fuzzy control method with genetic algorithm. *Comp. Chem. Eng.*, 27(7), 1031-1040.

Altınten, A., Erdoğan, E., Hapoğlu, H., Alıev, F., & Alpbaz, M., 2006. Application of fuzzy control metod with genetic algorithm to a polymerization reactor at constant set point. *Chem. Eng. Res. Des. (IChemE)*, 84(A11), 1012-1018.

Barner-Kowollik C., & Davis, T. P., 2001. Using kinetics and thermodynamics in the controlled synthesis of low molecular weight polymers in

free-radical polymerization. *Macromol. Theory Simul.*, 10(4), 255-261.

Cetinkaya, S., 1996. *Dynamic matrix control of a batch polymerization reactor under the optimal conditions.* Thesis (M.Sc). Department of Chemical Engineering, Ankara University, Turkey.

Chen, S.A., & Jeng, W.F., 1978. Minimum end time policies for batch-wise radical polymerization. *Chem Eng Sci.*, 33(6), 735-743.

Lewis, G.T., Nguyen, V., & Cohen, Y., 2007. Synthesis of poly(4-vinylpyridine) by reverse atom transfer radical polymerization. *J. Polym. Sci.*, 45(24) (Part A), 5748-5758.

Ozkan, G., Hapoglu, H., & Alpbaz, M., 1998. Generalized predictive control of optimal temperature profiles in a polystyrene polymerization reactor. *Chemical Engineering and Processing*, 37(2), 125-139.

Ponnuswamy, S.R., Shah S.L., Kiparissides C.A., 1987. Computer optimal control of batch polymerization reactors. *Ind. Eng.Chem.Res.*, 26(11), 2229-2236.

Seki, H., Ogawa, M., Ooyama, S., Akamatsu, K., Ohshima, M., & Yang, W., 2001. Industrial application of a nonlinear model predictive control to polymerization reactors. *Cont. Eng. Pract.*, 9(8), 819-828.

Ydstie, B.E., 1990. Forecasting and Control Using Adaptive Connectionist Networks, *Computers Chem. Eng.,* 14(4/5), 583-599.

Yuce (Cetinkaya), S., 2001. *Nonlinear Model Predictive Control of Reactor Temperature in Agitated Batch Polymerization Reactor Operating Optimal Condition.* Thesis (Ph.D). Department of Chemical Engineering, University of Ankara, Turkey.

Yüce, S., Hasaltun, A., Erdoğan, S., & Alpbaz, M., 1999. Temperature control of a batch polymerization reactor. *Chem. Eng. Res. Des.*, 77(5), 413-420.

Zeybek, Z., Cetinkaya, S., Hapoglu, H., & Alpbaz, M., 2006. Generalized delta rule (GDR) algorithm with generalized predictive control (GPC) for optimum temperature tracking of batch polymerization. *Chem. Eng. Sci.*, 61(20), 6691-6700.

AUTHORS BIOGRAPHY

Associate Professor Sevil CETINKAYA has both M.sc and Ph.D. degree in Chemical Engineering Department of Ankara University. She is a lecturer in the Department of Chemical Engineering of Cumhuriyet University(Sivas), Turkey.
Tel: +905055671933
e-mail: cetinkaya.sevil@gmail.com

Professor Mustafa ALPBAZ is a senior lecturer in the Department of Chemical Engineering Department of Ankara, Turkey, who holds a Ph.D. in Chem. Eng. Dept., Aston University, Birmingham/U.K. (1976). His research interests are, Modelling, Process Control and Simulation.
Tel:+90-312-2033445-3433
e-mail:alpbaz@eng.ankara.edu.tr

Professor Hale HAPOGLU has a PhD in Chemical Engineering Dept., Wales University, U.K.(1993). Since 2004, she is Professor at the Chemical Engineering Dept., University of Ankara. She has published several papers in conference proceedings. Her research interests are also on the Process Control, Modelling and Simulation.
Tel: +90-312-203 3514 - 3433
Fax:+90-312-2121546
e-mail: hapoglu@eng.ankara.edu.tr

Associate Professor Zehra ZEYBEK is a lecturer in the Department of Chemical Engineering of Ankara University(Turkey). She has a Ph.D. degree from the same university. Her research interest areas are Artificial Intelligence, Neural Network, Optimization, Fuzzy Logic and Modelling and Control of Chemical Plant with Al techniques.
Tel: +90-312-203 3522 – 3433
e-mail: zeybek@eng.ankara.edu.tr

Research Assistant Duygu ANAKLI has a M.Sc. degree in Chemical Engineering Department of Cumhuriyet University(Sivas), and also her Ph.D. program continues at the same university.
Tel: +905054996765
e-mail: danakli80@gmail.com

Active Vibration Reducing of a One-link Manipulator Using the State Feedback Decoupling and the First Order Sliding Mode Control

Mohammed BAKHTI [(a)] , Badr BOUOULID IDRISSI [(b)]

Ecole Nationale Supérieure d'Arts et Métiers
The University of Moulay Ismail _ Meknes
Morocco
[(a)]mdbakhti@yahoo.fr, [(b)]bbououlid@yahoo.fr,

ABSTRACT

This article presents an application of the first order sliding mode control in order to damp the mechanical vibrations of a flexible one-link manipulator using piezoceramic actuators.
The flexible manipulator, seen as a multiple inputs multiple outputs system is decoupled using state feedback scheme.
The model of the system is deduced using the finite element method, and its response to a control torque is calculated using the Lagrange equations. The state space equations are expressed in order to simplify the simulation and the controller implementation.

Keywords: flexible structure, first order sliding mode control, state feedback decoupling, piezoceramic actuators.

1. INTRODUCTION

The active control of flexible structures has been covered by a large number of research papers due to its high potential for industrial applications. For small amplitude vibration on very flexible structures, active approaches lead to lightweight and high performance control systems [1].
Used as sensors or actuators, piezoelectric materials have been well-studied [1], with [2] the first to suggest this idea.
Bailey and Hubbard [3] used distributed-parameter control theory and a piezoceramic actuator to actively control vibration on a cantilever beam actively.
Other researchers [4,5] studied the effect of the actuators on the host structures for vibration control through modal shape analysis. A variable structure adaptive controller developed by [6] to control contact forces on a cantilever beam used only the output force as feedback, resulting in undesirable chattering. Artificial neural networks (ANN) for identification and state feedback control of flexible structures have been implemented with good preliminary results [7].
Robust control focuses on the ability to have good control performance and stability in the presence of uncertainty in the system model as well as its exogenous inputs, including disturbances and noise. The H1 controller compensates for some of these uncertainties in active vibration control [8]. Recently, [9] developed a robust rejection method using a Kalman filter to estimate the system states under persistent excitation.
A large amount of research has focused on the optimization of sensors/actuators numbers and location, an example being [10]. Wang et al. [11] have very recently introduced PC for vibration suppression on a motor driven flexible beam, with very good simulation and practical results. These initial results were used also to diminish tip vibration on flexible beams [12].

The problem of decoupling linear multi-inputs multi-outputs systems has been subject of a great number of researchers works. Morgan [13] gave a sufficient condition for decoupling systems and defined a rather restrictive class of control laws which decouple. Additional results were obtained by Rekasius [14]. More recently, Falb and Wolovich [15] gave necessary and sufficient conditions for decoupling. They also restricted the class of control laws which decouple, which subsumes the classes introduced in [14] and [15]. Still more recently they obtained necessary and sufficient conditions on control laws for decoupling.

The control design methods to dominate the uncertainties and disturbances of systems have also attracted great research interest [16_18].

Among existing control strategies, the sliding mode control (SMC), first proposed in the early fifties, is one of the most suitable control design methods to dominate the uncertainties and disturbances acting on a system [19]. In the early sixties, the method has gained significant research attention, and has been widely applied in a variety of applications [18]. This technique has been applied to control a one flexible robot arm by [20].It also has been developed in association to piezoceramic actuators[21, 22].

Advantages of the SMC are robustness to parameter uncertainty, insensitivity to bounded disturbances, fast dynamic response, remarkable computational simplicity with respect to other robust control approaches and easy implementation [18].

The problem of the traditional (first-order) SMC method is mainly the problem of chattering. The chattering is due to the inclusion of the sign function in the switching term and it can cause the control input to start highly oscillating around the sliding surface, resulting in undesired actuators effects. This problem has been processed using two approaches: the first is to smoothen the switching term as the sliding surface gets closer to zero (soft switching) by using the continuous approximations of the discontinuous sign function, and the second is to generate ``higher-order sliding modes'', first introduced by Levant in 1987 [23]. In first order sliding mode controller design, the sliding surface is selected such that it has relative degree one with respect to the control input. That means the control input acts on the first derivative of the sliding surface. Higher-order sliding mode is the generalization of the

first-order sliding mode and the control input is performed such that it acts on higher derivatives of the sliding surface.

This paper will present a contribution based on the association of the state feedback decoupling and the first order sliding mode control to actively damp a flexible one-link manipulator.

The one-link manipulator considered as a Bernoulli beam is modeled in section 2 using the Euler - Lagrange formulation. The proper modes of vibration are calculated based on a finite element analysis and truncated to the first two more significant modes. Piezoceramic actuators effect on the model is introduced via the stiffness and force matrices. In section 3, the formulation of the decoupling problem is presented, and the control laws to decouple are calculated. Section 4 will deal with the first order sliding mode controller implementation. Thus, the sliding surfaces will be designed, and the control laws will be expressed. Finally, results of simulation will be illustrated in section 5, and conclusions given in section 6.

2. MODELING OF THE FLEXIBLE ONE-LINK MANIPULATOR

This section deals with the modeling of the flexible one-link manipulator. A model is developed for the case of a manipulator with one flexible link constrained to acting on a horizontal plane, and which is rigidly attached at one end to the shaft of an electric servomotor.

Two reference systems are defined:
1. X-Y-Z system: An inertial system with its origin at the centre of the shaft of the actuator, its z-axis aligned with the shaft, and the x-axis aligned with the home position of the manipulator.
2. r-w-Z system: A rotating system, attached to the actuator's shaft, with its z-axis coincident with the inertial system's z-axis and its r-axis tangent to the link at the shaft.
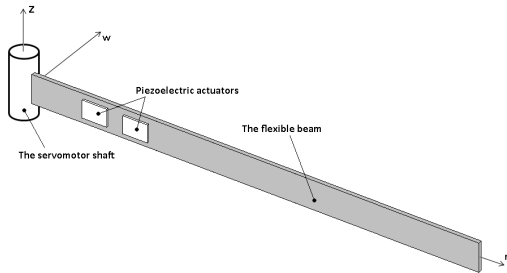
Figure 3.1: Configuration of the flexible one-link manipulator

The total displacement of the manipulator is considered to be the sum of rigid body rotation plus the flexible motion, as follows:

$$y(r,t) = \theta(t).r + w(r,t) = \theta(t).r + \sum_{i=1}^{N} a_i(t).\varphi_i(r)$$
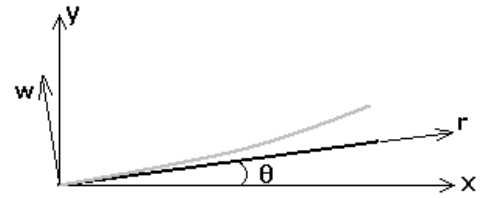
Figure 3.1: The flexible one- link manipulator displacement

The relative displacement is truncated to the first two flexible modes that have the most significant contribution to the global behavior of the system.

$$w(r,t) = a_1(t).\varphi_1(r) + a_2(t).\varphi_2(r)$$

Two layer of piezoelectric film are bonded to the manipulator in order to control the flexible modes.

2.1 Effect of the servomotor

The behavior of the link relative to the r- w -Z system can be analyzed using beam theory. Assuming that the link is a homogeneous slender beam of constant cross-section, the equations of motion can be found using the Bernoulli-Euler theory.

The flexible modes are calculated using the finite element method.

The elementary matrices of mass end stiffness are given by:

$$M = \frac{\rho s L}{420}\begin{bmatrix} 156 & 22l & 54 & -13l \\ 22l & 4l^2 & 13l & -3l^2 \\ 54 & 13l & 156 & -22l \\ -13l & -3l^2 & -22l & 4l^2 \end{bmatrix}$$

$$K = EI/l^3\begin{bmatrix} 12 & 6l & -12 & 6l \\ 6l & 4l^2 & -6l & 2l^2 \\ -12 & -6l & 12 & -6l \\ 6l & 2l^2 & -6l & 4l^2 \end{bmatrix}$$

And the analysis leads to an equation like:

$$M\ddot{q} + Kq = 0$$

The proper modes are given by the eigen vectors of the matrix: $M^{-1}K$, and the pulsations of vibration are given by the eigen values of the same matrix.

The magnitude of the velocity of any point of the link can be obtained by:

$$\dot{y}(r,t) = \dot{\theta}(t).r + \dot{w}(r,t) = \dot{\theta}(t).r + \sum_{i=1}^{2} \dot{a}_i(t).\varphi_i(r)$$

The kinetic energy of the manipulator can be written as the sum of rotational and translational components as:

$$T = \frac{J_m.\dot{\theta}(t)^2}{2} + \frac{1}{2}\int_0^L \dot{y}(r,t)^2\, dm$$

$$= \frac{J_m.\dot{\theta}(t)^2}{2} + \frac{1}{2}\int_0^L \left(\dot{\theta}(t).r + \sum_{i=1}^{2} \dot{a}_i(t).\varphi_i(r)\right)^2 \rho.S.dr$$

Where ρ is the density of the beam's material, S and I respectively its cross-sectional area and moment of inertia, and J_m is the motor fixture inertia.

The potential energy is stored in the manipulator as strain energy in the flexible link, and it can be expressed in terms of the mode shapes and modal coordinates as follows:

$$V = \frac{1}{2}EI \int_0^L \left(\frac{\partial^2 y(r,t)}{\partial r^2}\right)^2 dr$$

$$= \frac{1}{2}EI \int_0^L \left(\sum_{i=1}^{N} a_i(t).\frac{\partial^2 \varphi(r)}{\partial r^2}\right)^2 dr$$

E is the Young's modulus.

The joint angle and the modal coordinates can be grouped to form a vector of generalized coordinates, defined as:

$$q_{3\times1} = [\theta\ a_1\ a_2]^T$$

Similarly, the torque T applied at the joint and the modal forces can be grouped to form a vector of generalized forces.

Since there are no modal forces being applied to the system, the vector of generalized forces can be written as:

$$Q_{3\times1} = [T\ 0\ 0]^T$$

The work done by non-conservative external forces can then be written in terms of the generalized coordinates and forces as:

$$W_{nc} = Q^T q = T\theta$$

Since the kinetic and potential energy are already expressed in terms of the vector of generalized coordinates, the Lagrange's equations can be expressed:

$$\frac{d}{dt}\left(\frac{\partial(T-V)}{\partial \dot{q}_i}\right) - \frac{\partial(T-V)}{\partial q_i} = Q_i \qquad i = 1\dots N$$

Substituting the expressions for kinetic and potential energy and performing the required operations, one obtains the following matrix equation, which is a set of 3 ordinary differential equations, that model the dynamic behavior of the system:

$$M\ddot{q} + Kq = \begin{bmatrix} k_m \\ 0 \\ 0 \\ . \\ : \\ 0 \end{bmatrix} u(t) = f.u(t)$$

Where u is the control voltage sent to the servomotor amplifier. The system matrices are given by [12]:

$$M = \begin{bmatrix} Jm+\frac{\rho SL^3}{3} & \rho S \int_0^L \varphi_1(r).r.dr & \rho S \int_0^L \varphi_2(r).r.dr \\ \rho S \int_0^L \varphi_1(r).r.dr & \rho S \int_0^L \varphi_1{}^2(r).r.dr & 0 \\ \rho S \int_0^L \varphi_2(r).r.dr & 0 & \rho S \int_0^L \varphi_2{}^2(r).r.dr \end{bmatrix}$$

$$K = \begin{bmatrix} 0 & 0 & 0 \\ 0 & EI \int_0^L \left(\frac{\partial^2 \varphi_1(r,t)}{\partial r^2}\right)^2 dr & 0 \\ 0 & 0 & EI \int_0^L \left(\frac{\partial^2 \varphi_2(r,t)}{\partial r^2}\right)^2 dr \end{bmatrix}$$

The motor and hub parameters needed for the numeric calculation of the mass and stiffness matrices are shown in Table 2.2. The link parameters are listed in Table 2.1.

Table 2.1: Link Properties

Density	$\rho = 2700$ Kg/m^3
Length	L = 1.1 m
Young's modulus	E = 6.9 10^{10} Pa
Cross-section area	S = 8 10^{-8} m²
The quadratic moment	I = 2.67 10^{-8} m^4

Table 2.2: Motor-Hub Properties

Motor-Fixture Inertia	$J_m = 4\ 10^{-3}$ Kg m^2
Friction Coefficient	$B_m = 6.79\ 10^{-2}$ Nm/rad/s
Motor Constant	$k_m = 1$ Nm/V

The effects of actuator friction and the link structural damping can be included in the model via a viscous damping matrix given by:

$$H = \begin{bmatrix} B_m & 0 & 0 \\ 0 & 2\xi_1 m_{22}\omega_1 & 0 \\ 0 & 0 & 2\xi_2 m_{33}\omega_2 \end{bmatrix}$$

Where ω_i $i = 1..2$ are the clamped-free frequencies of the link, m_{ii} are the corresponding elements of the mass matrix, and ξ_i are the modal damping coefficients.

This yields to the following modified dynamic model:

$$M\ddot{q} + H\dot{q} + Kq = k_m u$$

2.2 Effect of the piezoceramic actuators

When the piezoceramic film (PZT) is used as an actuator, its effect on the dynamic model is through the passive stiffness and the force produced by the actuator.
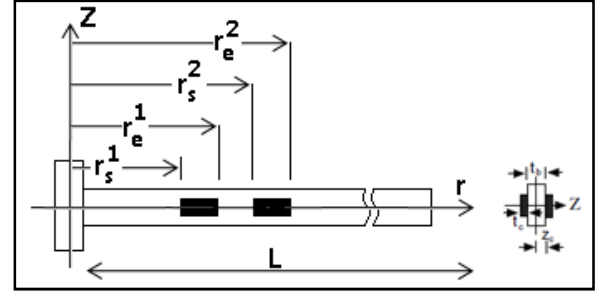

Figure 3.3: The PZT actuators setup

The contributions of the nth PZT to the overall stiffness matrix and global force matrix are given by [24]:

$$K_{pij}^n = \frac{E_c W_c t_c t_b^2}{4} \int_{r_s^n}^{r_e^n} \left(\frac{d^2\phi_i(x)}{dx^2}\frac{d^2\phi_j(x)}{dx^2}\right) dx$$

$$F_{pi}^n = f_{pi}^n u_p^n(t)$$

Clearly, $K_{pij}^n = 0$ when i does not equal j due to the orthogonality characteristics of mode shapes. F_{pi}^n is the generalized force associated with the ith mode. The coefficient of the generalized force is defined as:

$$f_{pi}^n = -sgn(z_c) \frac{d_{31}E_c W_c t_b}{2} \int_{r_s^n}^{r_e^n} \Xi_c^n \left(\frac{d^2\phi_i(x)}{dx^2}\right) dx$$

Where $u_p^n(t)$ is the voltage applied to the nth PZT actuator.

Introducing the effect of piezoelectric films, the new system dynamic model can be described as

$$M_{ps}\ddot{q} + H_{ps}\dot{q} + K_{ps}q = f_{ps}u$$

The new modal stiffness matrix K_{ps} is a combination of matrix K and K_{pij}

$$K_{ps} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & E_b I_{cs} \int_0^L \frac{d^2\phi_1(r)}{dr^2} dr + K_{p11} & 0 \\ 0 & 0 & E_b I_{cs} \int_0^L \frac{d^2\phi_2(r)}{dr^2} dr + K_{p22} \end{bmatrix}$$

Where $K_{pij} = \sum_{n=1}^2 K_{pij}^n$ and K_{pij}^n is obtained by evaluating Eq.(12) with the parameters on nth PZT film.

The modal force coefficient matrix f_{ps} is

$$f_{ps} = \begin{bmatrix} ku & 0 & 0 \\ 0 & f_{p1}^1 & f_{p1}^2 \\ 0 & f_{p2}^1 & f_{p2}^2 \end{bmatrix}$$

Where f_{pi}^n is obtained by evaluating Eq. (14) with the parameters of nth PZT film associated with mode i.
The input voltage vector u to the motor that rotate the beam and the voltage to the PZT actuators are:

$$u = [u_m(t) \quad u_p^1(t) \quad u_p^2(t)]^T_{3\times1}$$

Therefore, the system's generalized force matrix now includes two components: the force produced by the voltage applied to the motor $u_m(t)$ and the force produced by the voltage applied to the piezoelectric actuators $u_p^n(t)$ (n = 1, 2).

The PZT parameters are shown in Table 2.3.
Table 2.3: Piezoceramic actuators properties and placement

Material	PZT
Application	Actuator
Charge constant d_{31} (C/N)	$175 \cdot 10^{-12}$
Capacitance C_c (F)	$2 \cdot 10^{-8}$
Young's modulus E_c (N/m2)	$6.5 \cdot 10^{10}$
Width W_c (m)	0.025
Thickness t_c (m)	$5 \cdot 10^{-4}$
Length (m)	0.025
Shape function Ξ_c	1
r_s^1 (m)	0.1
r_e^1 (m)	0.125
r_s^2 (m)	0.150
r_e^2 (m)	0.175

2.3 State Space equations formulation

For modeling and control purposes it is convenient to write the model of the system in state-space form, as follows:

$$\dot{x}(t) = Ax(t) + BQ(t)$$
$$\text{Where } x = [\theta \; a_1 \; a_2 \; \dot\theta \; \dot a_1 \; \dot a_2]^T$$

Matrices A and B are defined in terms of the stiffness, mass and damping matrices K_{ps}, M_{ps} and H_{ps}, respectively, and the force vector f_{ps}:

$$A = \begin{bmatrix} 0_{3\times3} & I_{3\times3} \\ -[M_{ps}^{-1}K_{ps}]_{3\times3} & -[M_{ps}^{-1}H_{ps}]_{3\times3} \end{bmatrix}_{6\times6}$$

And

$$B = \begin{bmatrix} 0_3 \\ [M_{ps}^{-1}f_{ps}]_3 \end{bmatrix}_{6\times3}$$

Since the tip displacement is considered, the output vector is given by:

$$y = C.x = \begin{bmatrix} y_1(t) \\ y_2(t) \\ y_3(t) \end{bmatrix}$$

With :

$$\begin{cases} y_1(t) = L.\theta \\ y_2(t) = \varphi_1(L)a_1(t) \\ y_3(t) = \varphi_2(L)a_2(t) \end{cases}$$

So the total tip displacement is expressed as the sum of the rigid body motion plus the flexible cmodal contributions.

$$d(L,t) = L.\theta + a_1(t).\varphi_1(L) + a_2(t).\varphi_2(L)$$

L is the beam length.

Thus the matrix C is defined as follows:

$$C = \begin{bmatrix} L & 0 & 0 & 0 & 0 & 0 \\ 0 & \varphi_1(L) & 0 & 0 & 0 & 0 \\ 0 & 0 & \varphi_2(L) & 0 & 0 & 0 \end{bmatrix}$$

3. STATE FEEDBACK DECOUPLING

This section deals with the application of a feedback control in order to decouple the closed-loop system.

3.1 The decoupling method

Consider a linear dynamical system S:

$$\dot{x} = A.x + B.u$$
$$y = C.x + D.u$$

With inputs:
$$u(t) = [u_1 \quad u_2 \quad ... \quad u_m]^T_{1\times m}$$
State:
$$x(t) = [x_1 \quad x_2 \quad x_3 \quad ... \quad x_n]^T$$
And outputs:
$$y(t) = [y_1 \quad y_2 \quad ... \quad y_m]^T_{1\times m}$$

The control law used, as originally proposed by Morgan is given by [13]:

$$u = F.x(t) + G.v(t)$$
$$= F.x(t) + G[v_1 \quad v_2 \quad v_3 \quad ... \quad v_m]^T_{1\times m}$$

Where F and G are real, constant matrices of appropriate size, and v(t) is the new input to the closed-loop system.
The key to the solution of the decoupling problem is a canonical representation of integrator decoupled systems.
The system (S) can be decoupled if and only if the matrix given by :

$$[CB \quad CAB \quad CA^2B \quad ... \quad CA^{n-1}B]$$

is nonsingular.
The transfer function of the decoupled system is given by:

159

$$H(s, F, G) = C.(I_n.s - A - B.F)^{-1}.B.G$$

Where I_n is the n×n identity matrix, and s is the Laplace transform variable.

Decoupling the system consist on computing matrices F end G in order to have H diagonal and nonsingular.

Let C_i be the ith row of the output matrix C, and $D_i = C_i.A^{d_i}.B$, with:

$$d_i = \begin{cases} 0 \; if \; C_i.B \neq 0 \\ j \; if \; C_i.B = 0 \end{cases}$$

Where j is the largest integer from $\{1, 2, \ldots n - 1\}$ such that $C_i.A^{d_i}.B = 0$ for $k = 0, 1, \ldots j - 1$.

Once d_i and D_i are calculated, the system is decoupled by the following matrices F and G [26]:

$$\begin{cases} F = -D^{-1}.\tilde{A} \\ \quad G = D^{-1} \end{cases}$$

Where:

$$D = \begin{bmatrix} D_1 \\ D_2 \\ D_3 \end{bmatrix} and \tilde{A} = \begin{bmatrix} C_1 A^{d_1+1} \\ C_2 A^{d_2+1} \\ C_3 A^{d_3+1} \end{bmatrix}$$

And the ith output is the $(d_i + 1)$-fold integral of the ith input.

3.2 Application to the flexible manipulator

For the flexible manipulator, new inputs to the system are going to be defined in order have a decoupled system.
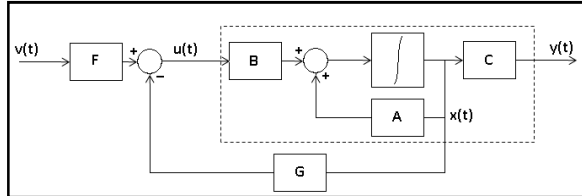
Figure3.1. State feedback decoupling scheme

The initial input vector is :

$$u = [u_m(t) \quad u_p^1(t) \quad u_p^2(t)]_{1\times3}^T$$

And the state vector is :

$$x = [\theta \; a_1 \; a_2 \; \dot{\theta} \; \dot{a}_1 \; \dot{a}_2]^T$$

The control law used is defined as:

$$\begin{aligned} u &= F.x(t) + G.v(t) \\ &= F.x(t) + G[v_m(t) \quad v_p^1(t) \quad v_p^2(t)]_{1\times3}^T \end{aligned}$$

Calculation results on:
$$d_i = 1 \; for \; i = 1,2 \; and \; 3$$

So, matrices $D$and $\tilde{A}$ are given by:

$$D = C.A.B \; and \; \tilde{A} = C.A^2$$

And each output of the system is going to be the double integral of its appropriate control input.
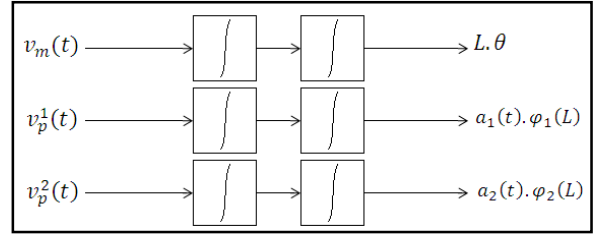
Figure3.2. Decoupled system

4. FIRST ORDER SLIDING MODE CONTROL

Since the flexible manipulator has been decoupled into three, single-input single-output second order, slightly independent systems, the first order sliding mode controller can be easily formulated.

1.1. Controller formulation

Considering the decoupled model of the flexible one-link manipulator, the system model is:

$$\begin{cases} \ddot{\theta} = v_m(t) \\ \ddot{a}_1 = v_p^1(t) \\ \ddot{a}_2 = v_p^2(t) \end{cases}$$

The tracking errors for state variables are defined as:

$$\begin{cases} e_\theta = \theta - \theta_d \\ e_{a_1} = a_1 - a_{1d} \\ e_{a_2} = a_2 - a_{2d} \end{cases}$$

The crucial and the most important step of SMC design is the construction of the sliding surface $s(t)$, also called sliding function.

A sliding mode is said "first order sliding mode" if and only if $s(t) = 0$ and $s(t)\dot{s}(t) < 0$. The inequality $s(t)\dot{s}(t) < 0$ is the fundamental condition for sliding mode. The aim of the first-order sliding mode control is to force the state (error) to move on the switching surface $s(t) = 0$.

The sliding surface, in the traditional SMC depends on the tracking error $e(t)$ and its derivative(s) as [19]:

$$s(t) = \left(\lambda_c + \frac{d}{dt}\right)^{n-1} e(t)$$

where n denotes order of uncontrolled system and λ_c is a positive constant.

If the system concerned is assumed to be of second-order, the first time derivative of the sliding surface is:

$$\dot{s}(t) = \lambda_c \dot{e}(t) + \ddot{e}(t)$$

The second time derivative of the error, ë(t) relatively to each state, can be written in terms of the plant new inputs as:

$$\begin{cases} \ddot{e}_\theta = \ddot{\theta} - \ddot{\theta}_d = v_m(t) \\ \ddot{e}_{a_1} = \ddot{a}_1 - \ddot{a}_{1d} = v_p^1(t) \\ \ddot{e}_{a_2} = \ddot{a}_2 - \ddot{a}_{2d} = v_p^2(t) \end{cases}$$

And, the switching surfaces may be defined as:

$$\begin{cases} s_\theta(t) = \lambda_\theta e_\theta(t) + \dot{e}_\theta(t) \\ s_{a_1}(t) = \lambda_{a_1} e_{a_1}(t) + \dot{e}_{a_1}(t) \\ s_{a_2}(t) = \lambda_{a_2} e_{a_2}(t) + \dot{e}_{a_2}(t) \end{cases}$$

The control input can be given as[19]:

$$u(t) = v_{eq}(t) + v_{sw}(t)$$

where $v_{eq}(t)$ and $v_{sw}(t)$ are the equivalent control and the switching control, respectively.

The equivalent control, $v_{eq}(t)$, proposed by Utkin[25], is based on the nominal (estimated) plant parameters and provides the main control action. The switching control, $v_{sw}(t)$, ensures the discontinuity of the control law across sliding surface, supplying additional control to account for the presence of matched disturbances and unmodeled dynamics.

If the initial error is not on the sliding surface $s(t)$ due to parameter variations and disturbances, the controller must be designed such that it can drive the error to the sliding surface. The error under the condition that will move toward and reach the sliding surface is said to be on the reaching phase.

4.1 Equivalent control

The equivalent control approach is defined as the smooth feedback control law that locally sustains the evolution of the error ideally restricted to the smooth sliding surface $s(t)$ when the initial error of the system is located precisely on it. This control allows an asymptotic convergence of the sliding function to zero according to a desired dynamic and the control forces the system to evolve on the sliding surface.

Thus, in the design of sliding mode controllers, an equivalent control is first given so that the states can stay on sliding surface. Consequently, the system dynamics in sliding motion is independent of the original system and a stable equivalent control system is achieved.

The equivalent control is obtained when $\dot{s}(t) = 0$, which is the necessary condition for the tracking error to remain on the sliding surface.

Now, an equivalent control for the three states can be defined:

$$\begin{cases} \dot{s}_\theta(t) = \lambda_\theta \dot{e}_\theta(t) + \ddot{e}_\theta(t) = 0 \Leftrightarrow v_{eq}^\theta(t) = -\lambda_\theta \dot{e}_\theta(t) \\ \dot{s}_{a_1}(t) = \lambda_{a_1} \dot{e}_{a_1}(t) + \ddot{e}_{a_1}(t) = 0 \Leftrightarrow v_{eq}^{a_1}(t) = -\lambda_{a_1} \dot{e}_{a_1}(t) \\ \dot{s}_{a_2}(t) = \lambda_{a_2} \dot{e}_{a_2}(t) + \ddot{e}_{a_2}(t) = 0 \Leftrightarrow v_{eq}^{a_2}(t) = -\lambda_{a_2} \dot{e}_{a_2}(t) \end{cases}$$

4.2 Switching control

The switching control is introduced as[19]:
$$v_{sw}(t) = -k.sign(s(t))$$
where k is a positive constant, chosen to dominate the matching uncertainties, and sign(.) denotes sign function defined by:

$$sign(s) = \begin{cases} 1 & if\ s > 0 \\ 0 & is\ \ s = 0 \\ -1 & if\ s < 0 \end{cases}$$

Consequently, the control laws for the three states are given as:

$$v_m(t) = v_{eq}^\theta(t) + v_{sw}^\theta(t) = -k_\theta sign(s_\theta) - \lambda_\theta \dot{e}_\theta(t)$$
$$v_p^1(t) = v_{eq}^{a_1}(t) + v_{sw}^{a_1}(t) = -k_{a_1} sign(s_{a_1}) - \lambda_{a_1} \dot{e}_{a_1}(t)$$
$$v_p^2(t) = v_{eq}^{a_2}(t) + v_{sw}^{a_2}(t) = -k_{a_2} sign(s_{a_2}) - \lambda_{a_2} \dot{e}_{a_2}(t)$$

5. SIMULATION AND RESULTS

This section illustrates the results of the simulations conducted on the flexible beam that is solidly coupled to the shaft of a servomotor.

The motor is rotated and controlled to an angular set-point using a sliding mode controller.

During rotation, the beam's vibrations are suppressed using piezoceramic actuators until set-point of the joint angle has been completed.

First, results are illustrated when only joint angle is controlled. The single-input single-output (SISO) controller uses $v_m(t)$ as the control variable while the controlled variable is θ.

Then, the piezoceramic actuators are used to diminish the beam vibrations during the rotation. The effect of each mode of vibration will be controlled separately. The single-input single-output (SISO) controllers uses respectively $v_p^1(t)$ and $v_p^2(t)$ as the control variables for the manipulated variables a_1 and a_2.

5.1 Joint angle control

The Joint angle set-point is 45°, and the control law is given by:

$$v_m(t) = v_{eq}^\theta(t) + v_{sw}^\theta(t) = -k_\theta sign(s_\theta) - \lambda_\theta \dot{e}_\theta(t)$$

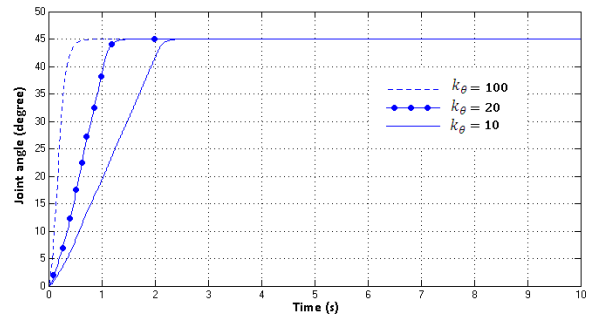The simulation illustrate the effect of the parameters: k_θ and λ_θ.

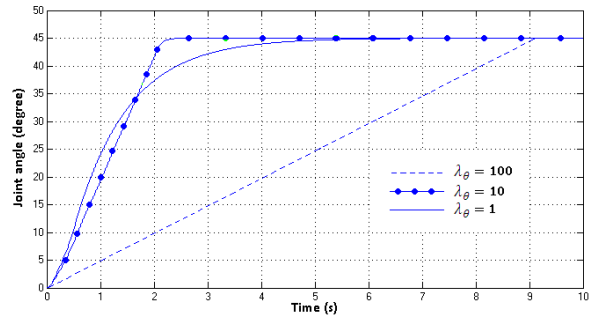Figure 5.1: Joint angle sliding mode control evaluation of the parameter k_θ (λ_θ=10)

Figure 5.2: Joint angle sliding mode control evaluation of the parameter λ_θ (k_θ=10)

The response of the system comes to be faster as the parameter k_θ increases.

The parameter λ_θ doesn't have an important effect on the system acceleration because its influence start once the response reach the switching surface and not during the reaching phase.

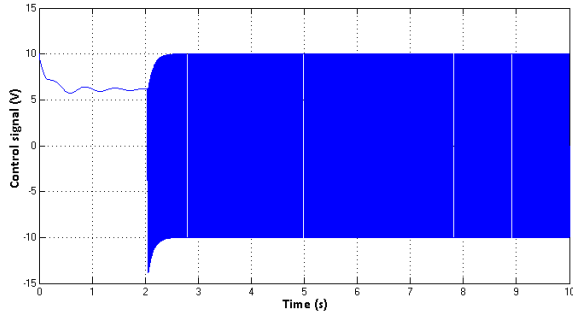The equivalent control input to the servomotor is very oscillatory as it's illustrated in Figure 5.3.

Figure 5.3: Joint angle sliding mode control evaluation of the control signal sent to the servomotor (k_θ=10 and λ_θ=10)

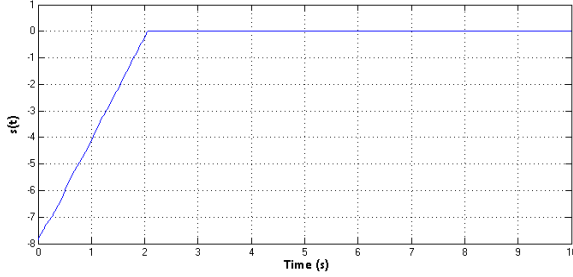The figures 5.4 and 5.5 illustrate the sliding surface and the phase plane respectively.
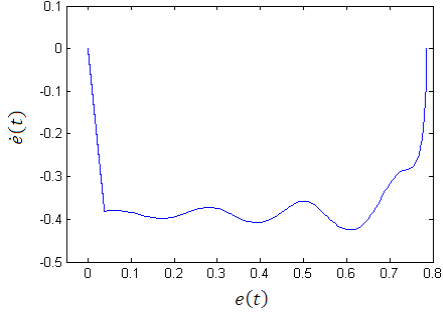

Figure 5.4: Sliding surface s(t)

Figure 5.5: The phase plane $e(t)$ and $\dot{e}(t)$

The sliding mode is in the reaching phase up to 2s and then arrives to the sliding phase. The error and its first derivative comes to zero at steady-state conditions.

5.2 Piezoelectric actuators control

The tip displacement set-points are chosen to be equal to 0.1mm and 0.01mm respectively for the first and second mode contributions, and the control inputs send to the piezoelectric actuators are :

$$v_p^1(t) = v_{eq}^{a_1}(t) + v_{sw}^{a_1}(t) = -k_{a_1}sign\left(s_{a_1}\right) - \lambda_{a_1}\dot{e}_{a_1}(t)$$
$$v_p^2(t) = v_{eq}^{a_2}(t) + v_{sw}^{a_2}(t) = -k_{a_2}sign\left(s_{a_2}\right) - \lambda_{a_2}\dot{e}_{a_2}(t)$$

Similarly to the joint angle control, the consequences of manipulating $v_p^1(t)$ and $v_p^2(t)$ on the first and second mode contributions to the tip displacement. are illustrated.
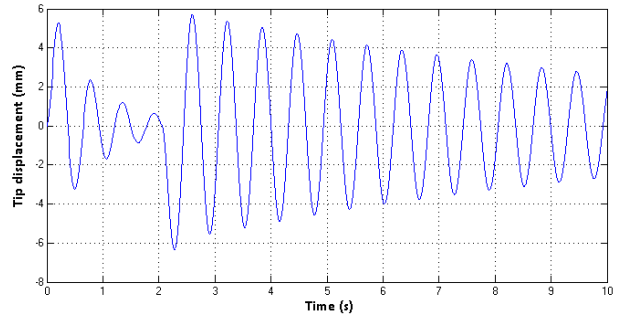
Figure 5.6: First mode contribution to the tip displacement (only joint angle is controlled with k_θ=10 and λ_θ=10)
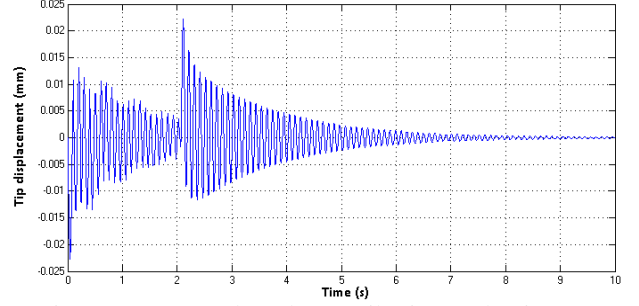
Figure 5.7: Second mode contribution to the tip displacement (only joint angle is controlled with k_θ=10 and λ_θ=10)
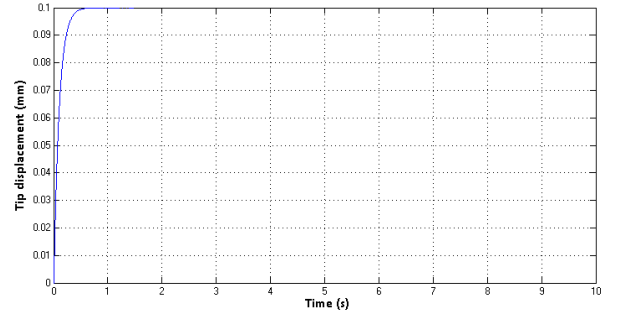
Figure 5.8: First mode contribution to the tip displacement (PZT is controlled with k_{a_1}= 0.1 and λ_{a_1}=10)
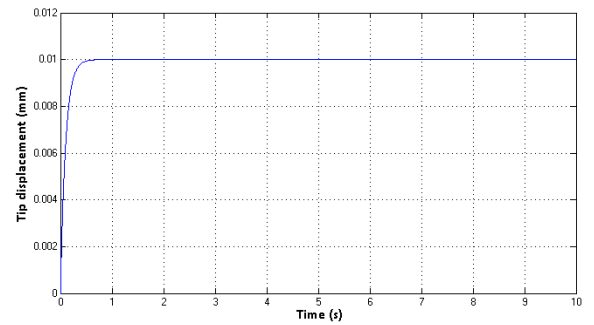
Figure 5.9: Second mode contribution to the tip displacement (PZT is controlled with k_{a_2}= 0.01 and λ_{a_2}=10)

It's clear that the objective is reached. The tip displacements due to vibrations of 1st and 2nd mode are reduced to 0.1mm and 0.01mm respectively.

The problem is on the control signals sent to the piezoelectric actuators, and it's illustrated in Figures 5.10 and 5.11.
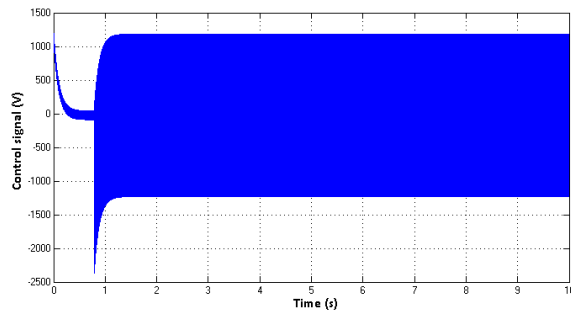
Figure 5.10: Control signal sent to the 1st PZT
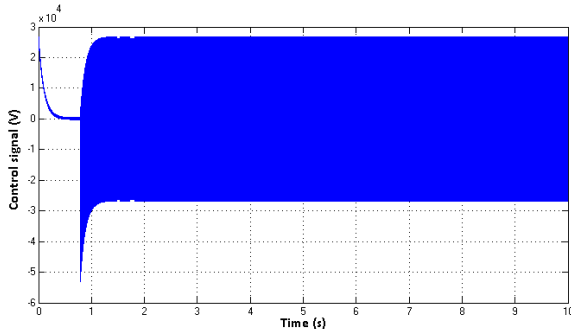

Figure 5.11: Control signal sent to the 2nd PZT

In addition to the fact that they are very oscillatory, control signal reaches very high values that can't be supported by piezoelectric actuators.

6. CONCLUSIONS

State space equations of the flexible manipulator, that describe the rigid and flexible motions, were expressed using the Euler–Lagrange formulation. The flexible displacement was calculated using modal analysis and beam theory.

The piezoelectric actuators effect, on the manipulator model, was properly introduced via the stiffness and force matrices.

The system was decoupled into three independent single-input single-output systems using a state feedback decoupling scheme. The advantage of this strategy was to provide direct control on proper modes of vibration contributions to the flexible manipulator tip displacement.

A sliding mode controller was, then, implemented for each one of the three independent systems. Although the control signal was very oscillatory and of higher magnitude, and no comparisons to other controllers were conducted, this method of control has successfully demonstrated its ability to perfectly control the rigid body rotation. It has also proved a good vibration decreasing within a very satisfying time.

Both, for decoupling and for controlling the system, state variables should be measured. Those measures were not discussed in this paper. However; in practice, the modal coordinates are well estimated using piezoelectric films as sensors, and the joint angle is simply measured using a potentiometer.

REFERENCES

[1] Recent advances in sensing and control of flexible structures via piezoelectric materials.
Sunar M, Rao SS. Appl Mech Rev 999;52(10):1–16.

[2] Modeling of an active constrained layer damper proc.
Plump JM, Hubbard Jr JE. In: 12th intl congress on acoustics, Toronto, Canada, 1986; #D4-1.

[3] Distributed piezoelectric-polymer active vibration control of a cantilever beam.
Bailey T, Hubbard Jr JE. J Guidance Control Dynam 1985;8(5):605–11.

[4] Performance of an active control system with piezoelectric actuators.
Baz A, Poh S. J Vib Control 1988;126(2):327–43.

[5] Control of a miniature gripper driven by piezoceramic bimorph cells.
Chonan S, Jiang ZW, Sakuma S. Force J Adv Automat Technol, 1994;6:247–54.

[6] Variable structure adaptive control of a cantilever beam using piezoelectric actuator.
Yim W, Singh SN. J Vib Control, 2000; 6:1029–43.

[7] Identification and control of a benchmark flexible structure using piezoelectric actuators and sensors.
Lin CL, Lee GP, Liu VT. J Vib Control, 2003;9(12):1401–20.

[8] Robust control for structural systems with parametric and unstructured uncertainties.
Wang S, Yeh H, Roschke PN. J Vib Control, 2001;7:753–72.

[9] A robust disturbance rejection method for uncertain flexible mechanical vibrating systems under persistent excitation.
Zheng LA. J Vib Control, 2004;10(3):342–57.

[10] Optimal design of number and locations of actuators in active vibration control of a space truss.
Yam LH, Yan YJ. Smart Mater Struct, 2002;11:496–503.

[11] Predictive control of flexible structures.
Wang R, Hassan M, Dubay R. In: Flexible automation & intelligent manufacturing conference, Toronto, Ontario, 2004. p. 98–104.

[12] Vibration control of flexible structure using smart materials.
Bravo R. PhD dissertation. McMaster University, Canada, 2000.

[13] The synthesis of linear multivariable systems by state variable feedback
B. S. MORGAN, JR. Proc. 1964 JACC, Stanford, California, pp. 468-472.

[14] Decoupling of multivariable systems by means of statefeedback
Z.V. REKASIUS. Proc. Third Allerton Conference on Circuit and System Theory, Monticello, Illinois, 1965, pp. 439-448.

[15] On the decoupling of multivariable systems.
P. L. F,LB ,ND W. A. WOLOVICn, Proc. 1967 JACC, Philadelphia, Pennsylvania, pp. 791-796.

[16] Second order sliding mode control of vehicles with distributed collision avoidance capabilities.
Ferrara A, Vecchico C. Mechatronics, 2009; 19:471_7.

[17] Sliding mode control of quadruple tank process.
Biswas PP, Srivastava R, Ray S, Samanta AN. Mechatronics 2009;19:548_61.

[18] High-order sliding control of mechanical systems: Theory and experiments.
Cavallo A, Natale C. Control Eng Pract, 2004; 12:1139_49.

[19] Applied nonlinear control.
Slotine JJ, Li W. Englewood Cliffs (New Jersey): Prentice Hall Inc.; 1991.

[20] Regulation of a one-link flexible robot arm using sliding-mode technique
K.S. Yeung; Y.P. Chen, International Journal of Control, 1989.

[21] Active vibration control of a flexible beam using a non-collocated acceleration sensor and piezoelectric patch actuator
Z Qiu, J Han, X Zhang, Y Wang – Journal of sound and vibration, 2009.

[22] A PZT actuator control of a single-link flexible manipulator based on linear velocity feedback and actuator placement
DSun, JK Mills, J Shan, Mechatronics, 2004.

[23] Sliding order and sliding accuracy in sliding mode control.
Levant A. Internat J Control, 1993; 58:1247_63.

[24] Composite modeling of flexible structures with bonded piezoelectric film actuators and sensors.
Vaz A. IEEE Trans Intrum Measur1998; 24(2):513–20.

[25] Sliding modes in optimization and control problems.
Utkin VI. New York: Springer-Verlag; 1992.

[26] The decoupling of multivariable systems by state feedback .
Elmer, G. Gilbert, SIAM J. CONTROL Vol. 7, No. 1, February 1969

DEVELOPMENT OF A TRANFER-CASE CONTROL SYSTEM
FOR REAR WHEEL BASED 4WD VEHICLES

Woosung Park[(a)], Hyeongjin Ham[(b)],Kanghee Won[(c)], Hyeongcheol Lee[(d)*]

[(a),(b),(c)] Department of Electrical Engineering, Hanyang University, 17 Haengdang-dong, Sungdong-gu, Seoul 133-791, Korea
[(d)] Division of Electrical and Instrumentation Biomedical Engineering, Hanyang University, 17 Haengdang-dong, Seongdong-gu, Seoul 133-791, Korea

[(a)]Jacklos@hanyang.ac.kr, [(b)]postechj@hanyang.ac.kr, [(c)]wkh@hanyang.ac.kr, [(d)]hclee@hanyang.ac.kr

ABSTRACT

This paper presents a control method of torque distribution for Four Wheel Drive (4WD) vehicles. The 4WD system is a device which distributes optimal torque between front and rear propeller shafts to improve traction ability and stability of a vehicle. While the object of the traditional 4WD system is to improve only traction ability by distribution of the transmission torque, the recent 4WD system is used for to meliorate the traction ability, stability and fuel consumption efficiency. This paper proposes the control algorithm for the vehicle which is equipped the electronic controlled 4WD system to improve traction ability. In this paper, the mathematical equation-based control method by vehicle dynamics analysis is presented. The proposed control algorithm of the 4WD system is validated by real vehicle test using the Rapid Control Prototyping (RCP) and the mid-size Sport Utility Vehicle (SUV) equipped 4WD system of multi plate clutch type actuated by solenoid coil.

Keywords: Vehicle Dynamics, Power train & Drive train Control, Traction and Yaw control

1. INTRODUCTION

The 4WD system distributes the engine torque to the front or rear wheels according to the driving condition and the vehicle statement to improve traction ability and stability of the vehicle. Nowadays, there is a requirement of traction ability such as not only Sports Utility Vehicles (SUV) and Multi Purpose Vehicles (MPV) but also luxury sedan and sports car. Especially, the person who lives in the snowy region or country side needs the traction performance of the 4WD vehicles.

The 4WD system is classified by coupling of the Front Wheel Drive (FWD) vehicle and transfer-case of the Rear Wheel Drive (RWD) vehicle. Moreover according to the existence of controller, it is categorized by mechanical type and electronic type. The electronic type 4WD system is assorted again to hydraulic motor type and clutch type controlled by solenoid. Compared to the traditional mechanical 4WD, the electronically controlled 4WD system provides better controllability and fuel economy.

Controlling vehicle dynamics by the 4WD control system is a quite challenging area for the following reasons: [1]

a. The 4WD control is meaningful only when the engine generates driving torque or meaningful engine brake torque.

b. The 4WD control system can control only the maximum biasing torque that the coupling can support without slipping. The actual biasing torque can be any value under this maximum value, depending on the amount of engine torque, the tire–road surface condition, and the weight shift between the front and rear wheels.

c. The arbitration between the 4WD control system and the brake control system (e.g. ABS, TCS, and ESP) is not clear. Most of the time, the 4WD control system deactivates the coupling when the brake control system is activated.

Owing to these difficulties, the 4WD control algorithm has been frequently developed using the look-up table method. These types of controller proposed by [2], [3] have some problems, such as rough control (bad control accuracy with the large duty cycle increment) or slow response (with the small duty cycle increment).

In this paper, the Equation control strategy for the 4WD vehicles is proposed. For this, the 2 DOF vehicle model and the driveline model including torque biasing devices for RWD based 4WD are developed in section 2. A proposed 4WD control algorithm is given in section 3 and the simulation results are shown in section 4.

2. SYSTEM MODELING

The system modeling is provided in this section to figure out 4WD device and to design controller.
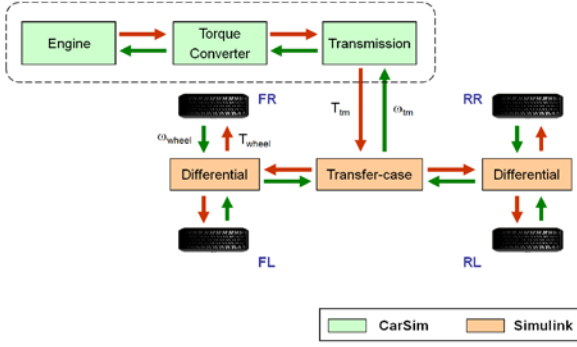
Figure 1: Vehicle Model

Fig. 1 shows a graphical description of the overall system configuration.

From this diagram, the modeling of transfer-case for the system is formulated in the following section 2.1.

2.1. Transfer-case Model

As explained previously in introduction, two different types of drive-line systems that one is coupling and the other is transfer-case, are typically used for the 4WD device model depending on the placement of a vehicle engine.

From Fig. 2, it implies that T_{bias} is the bias torque which is passed through the coupling device.
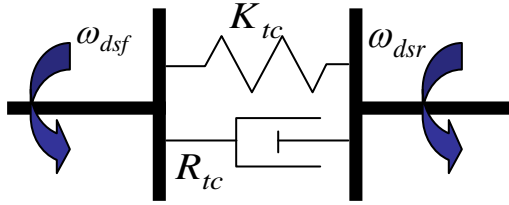
Figure 2: Torque biasing device

By assuming that the inertia is 0, the bias torque equation is derived as follow.

$$T_{bias} = K_{tc}(\phi_{dsf} - \phi_{dsr}) + R_{tc}(\omega_{dsf} - \omega_{dsr}) \quad (1)$$

where K_{tc}, R_{tc}, ϕ_{dsf}, ϕ_{dsr}, ω_{dsf} and ω_{dsr} are a spring coefficient, a damping coefficient, the displacements of the front and the rear drive shaft, and the angular velocities at the front and the rear drive shaft, respectively.

In this case, the bias torque defined by both the spring and the damper has an upper limit so that the bias torque can be always smaller than the input torque.

As shown in Fig. 3, if the speed difference between the front propeller shaft and the rear propeller shaft is detected, the electric signal is sent to solenoid by the control module so that the pressure is given to the

ball/lamp and it give rise to the friction forces between clutch plates. As a result, the driving force is delivered to the front propeller shaft.
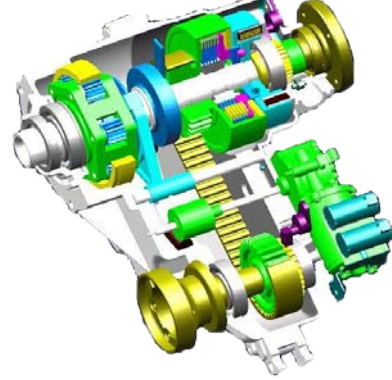
Figure 3: Solenoid type transfer-case device

By using the transfer-case device model as shown in Fig. 2, the model of the clutch plates can be expressed as equation 1, previously obtained.

However, since the torque for controlling clutch, provided by the control module, is not equal to the bias torque of equation 1, the lock and the slip should be analyzed in the different states.

If the clutch control torque provided by control module is smaller than the bias torque ($T_{control} < T_{bias}$), the clutch control torque transfers to the driven wheels. In this case, because the control torque exceeds the tolerance range of the transfer-case, the clutch slips. On the other hand, if the clutch control torque is larger than the bias torque ($T_{control} > T_{bias}$), the bias torque transfers to the driven wheels in the locking state of the transfer-case.

In other words, the bias torque is a criterion of conversion, whether lock state or slip state.

The torque relationship of each state is represented as follows:

$$* \left| T_{dsf\_lock} + T_{bias} \right| < T_{control\_c} : \text{Lock state}$$

$$T_{dsr} = \frac{G_{taf}}{G_{tar}} T_{tm} - T_{dsf} \qquad (2\text{-}1)$$

$$T_{dsf} = T_{dsf\_lock} + T_{bias}$$

$$* \left| T_{dsf\_lock} + T_{bias} \right| \geq T_{control\_c} : \text{Slip state}$$

$$T_{dsr} = G_{tar} \cdot T_{tm} - \frac{G_{tar}}{G_{taf}} T_{dsf} \qquad (2\text{-}2)$$

$$T_{dsf} = T_{control\_c}$$

where G_{taf}, G_{tar}, T_{dsf_lock} and T_{dsr_lock} are a gear ratio between transfer-case and front drive shaft, a gear ratio between transfer-case and rear drive shaft, lock torque of front drive shaft, lock torque of rear drive shaft, respectively.

2.2. Consideration Of The Road Condition

In general, a flow of the torque is represented from the engine to the wheels. The torque distribution neglected the road condition affect to the accuracy of vehicle drive-train model. In this paper, therefore, a limitation of transferring torque to the driven wheels is decided according to the road condition as follows equations.

A transmission torque equals summation of the front drive shaft torque and rear drive shaft torque.

$$T_{tm} = (T_{dsf} / G_{taf}) + (T_{dsr} / G_{tar}) \qquad (3)$$

Since the front and rear drive shaft torque equals product of longitudinal force of wheels and effective wheel radius, the difference between rear drive shaft and front drive shaft is derived as follow.

$$T_{dsr} - T_{dsf} = \{(F_{rl} + F_{rr}) \times R_{wr} / G_{fdr}\} \\ - \{(F_{fl} + F_{fr}) \times R_{wf} / G_{fdf}\} \qquad (4)$$

where G_{fdf}, G_{fdr}, R_{wf} and R_{wr} are a final reduction gear ration of front and rear differential, a effective wheel radius of front and rear wheels, respectively.

From the equation (3) and (4), the limitation of front and rear drive shaft torque is derived using appropriate math operation.

$$T_{dsr\_Lock} = \begin{bmatrix} G_{taf} \times G_{tar} \times T_{tm} + G_{tar} \\ \times \begin{cases} (F_{rl} + F_{rr}) \times R_{wr} / G_{fdr} \\ -(F_{fl} + F_{fr}) \times R_{wf} / G_{fdf} \end{cases} \end{bmatrix} \\ / (G_{taf} + G_{tar}) \qquad (5)$$
$$T_{dsf\_Lock} = (G_{taf} \times G_{tar} \times T_{tm} \\ - G_{taf} T_{dsf}) / G_{tar}$$

3. CONTROL ALGORITHM

The controller of look-up table methods was used until from the now many Conventional vehicles. The method which it proposes from the present paper collects the strong point of Look-up table methods and Equation methods and it makes and they are Look-up table methods which formulate.

The purposes of the controller are to improve on traction ability, stability ability and robustness against model uncertainties. The traction performance and stability performance of the proposed control strategy are validated by vehicle simulation using MSC's CarSim®.

3.1. Traction Control

The 1st Traction Control strategy is the front wheel torque distribution which it follows in vehicle speeds and APS(Acceleration Pedal Signal). The 2nd Traction Control strategy is the front wheels torque distribution in order to reduce the difference between front and rear wheel speed.

The engine torque graph which it relationship with vehicle speed shown as follows:
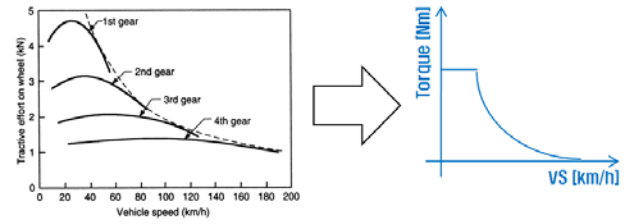

Figure 4: The torque limit with a engine performance

Because of this, transfer torque also has to decrease as the velocity of vehicle increases.

Next consideration such as APS and DeltaN(Wheel speed difference between Front wheel and Rear wheel) is the show of will and the condition of the car. Because of this, the output torque has to be above such level from the time when the input value is small. This could lead to the satisfaction of the driver and at the same time, can induce the car to a safe condition.

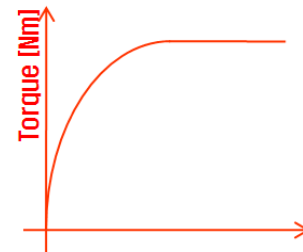The relationship is defined as follows.

Figure 5: Torque of APS or DeltaN

The equation has been derived by considering all the factors above. The equation can be fine-tuned by varying the internal parameters.

First, the feed forward controller that could be determined by APS and vehicle speed are as follows.
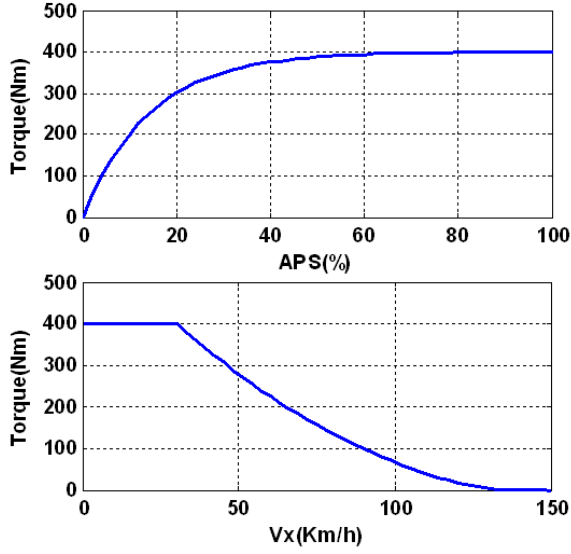
Figure 6: Feed-Forward equation graph

$$T_{TR\_FF} = C_{FF\_slope} \times T_{FF\_max}$$
$$\times \left(1 - e^{(-APS \times p_{slope\_e})}\right) \quad (6)$$

$$C_{FF\_slope} = \frac{\left(p_{slope\_q} \times V_x - p_{sat}\right)^2}{p_{sat}^2} \quad (7)$$

Seconds, feedback controller has been derived by using DeltaN and Vehicle speed. To achieve fast responses, instead of using feed forward control, we have decreased the vehicle speed and increased control amount.
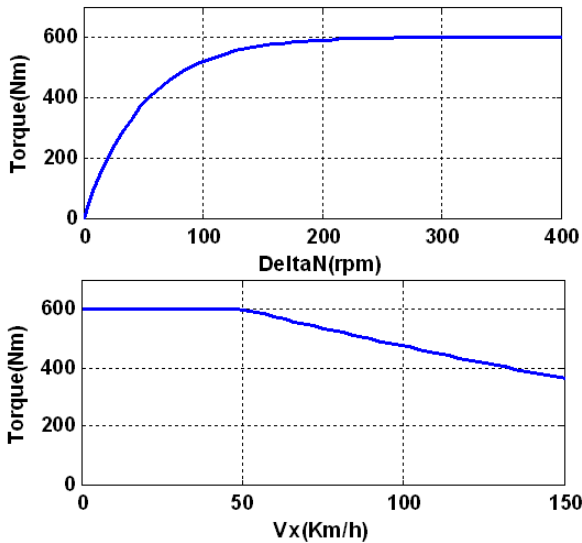

Figure 7: Feed-back equation graph

$$T_{TR\_FB} = C_{FB\_slope} \times T_{FB\_max}$$
$$\times \left(1 - e^{(-DeltaN \times p_{slope\_e})}\right) \quad (8)$$

$$C_{FB\_slope} = \frac{\left(p_{slope\_q} \times V_x - p_{sat}\right)^2}{p_{sat}^2} \quad (9)$$

We have developed equation controller based on the characteristic of the car and the will of the driver.

Compared to look-up table, control at more elaborate and dynamic range can be possible.
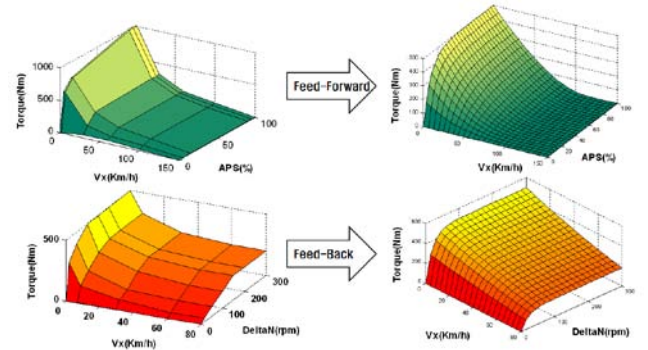
Figure 8: Comparison of Look-up table and Equation Control

3.2. Yaw Control

If the torque amount for Front and Rear wheel has been varied, the cornering force affecting the wheel can be changed and this could potentially lead to yaw moment in the car.

Yaw controller has been designed to decrease the yaw error using this method.

To find out yaw error which acts as input data for yaw control, desired yaw rate has been estimated by utilizing various methods shown below.
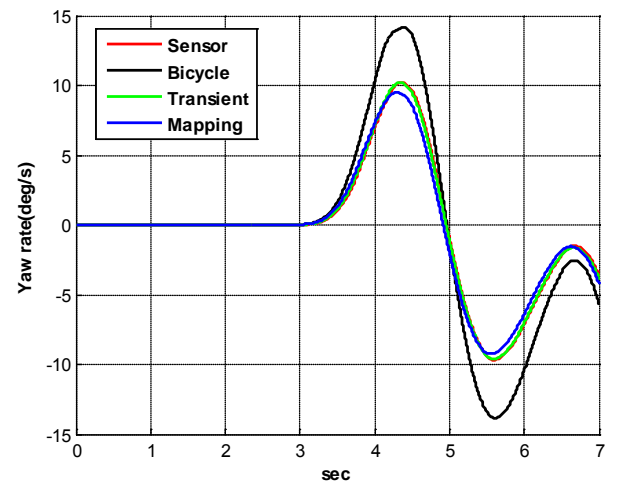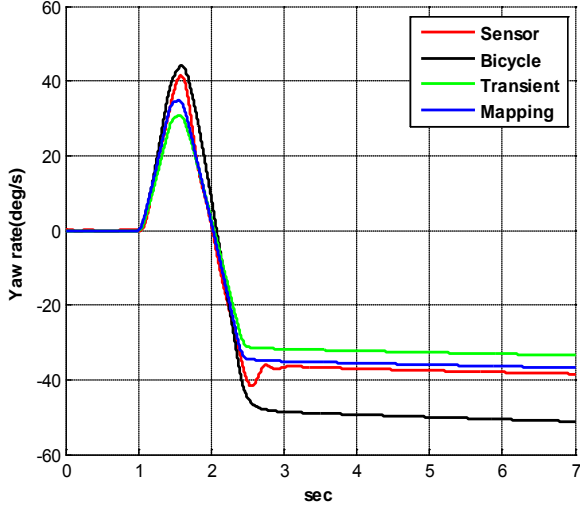

Figure 9: Double lane change desired yaw rate

Figure 10: Fishhook desired yaw rate

Desired yaw rate can be estimated by using bicycle mode, transient model and simulation-based mapping look-up table. The following data is the comparison between the desired yaw rate and actual yaw rate.

Transient model which acts the most similar has been used from the simulation result.

Transient model has been accomplished by deriving space state eqution of Equation (10) from the bicycle model.

After Laplace transform has been applied, transfer function for equation (11) has been constructed.

$$\begin{pmatrix} \dot{v}_y \\ \ddot{\psi} \end{pmatrix} = \begin{bmatrix} -\dfrac{2(C_f + C_r)}{mv_x} & -v_x - \dfrac{2(C_f l_f - C_r l_r)}{mv_x} \\ -\dfrac{2(C_f l_f - C_r l_r)}{I_z v_x} & -\dfrac{2(C_f l_f^2 + C_r l_r^2)}{I_z v_x} \end{bmatrix} \begin{pmatrix} v_y \\ \dot{\psi} \end{pmatrix} \qquad (10)$$

$$+ \begin{pmatrix} \dfrac{2C_f}{m} \\ \dfrac{2C_f l_f}{I_z} \end{pmatrix} \delta_f$$

$$\Phi_{\delta_f \to \psi} = \frac{b_2 s + (a_1 b_2 - a_3 b_1)}{s^2 + (a_1 + a_4)s + a_1 a_4 - a_2 a_3} \qquad (11)$$

$$a_3 = \frac{2(C_f l_f - C_r l_r)}{I_z v_x}, a_4 = \frac{2(C_f l_f^2 + C_r l_r^2)}{I_z v_x},$$

Where,

$$b_1 = \frac{2C_f}{m}, b_2 = \frac{2C_f l_f}{I_z}$$

By using yaw error derived, equation controller has been designed in a similar way for traction controller.
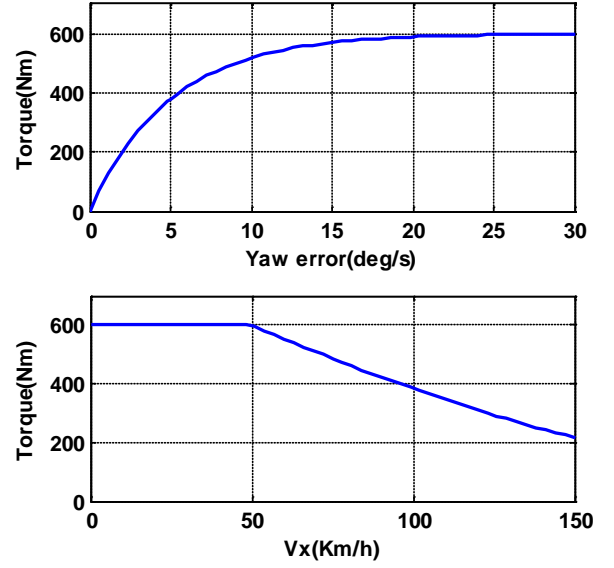

Figure 11: yaw control equation graph

$$T_{Yaw} = C_{Yaw\_slope} \times T_{Yaw\_max} \times \left(1 - e^{(-Yaw\_err \times p_{slope\_e})}\right) \qquad (12)$$

$$C_{Yaw\_slope} = \frac{\left(p_{slope\_q} \times V_x - p_{sat}\right)^2}{p_{sat}^2} \qquad (13)$$

If above two equations are put together, yaw control equation can be derived below.
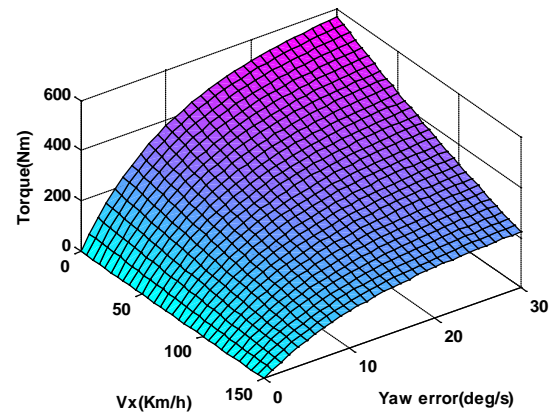

Figure 12: 3D Plot of Yaw control

3.3. Blending Control
Conventional controllers independently utilize traction control and yaw control, but in this paper, the two methods have been mixed together.

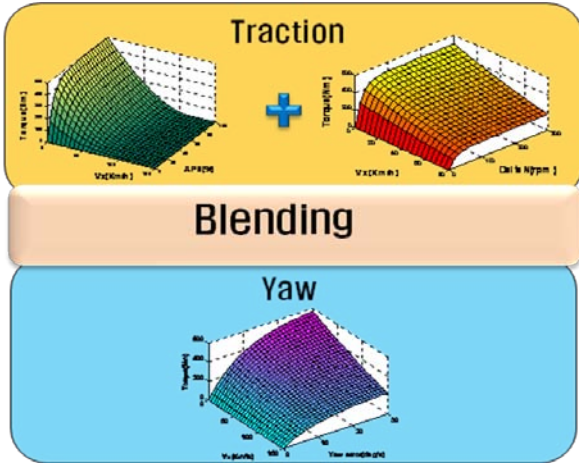The concept of blending the traction and yaw controller is as follows.

169

Figure 13: Blending Concept

Weighting functions varying along the error of traction control and yaw control can be multiplied. Each

Distributed output from the control algorithm can be used. The equation is shown below.
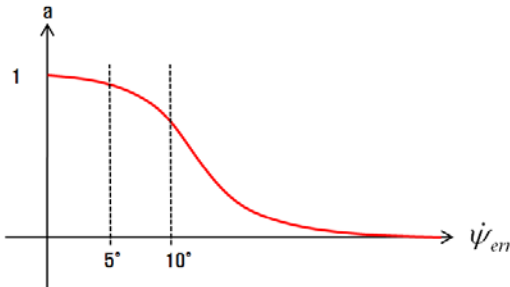

Figure 14: Weighting Function for Yaw control

$$T_{total} = a \cdot (T_{FF} + T_{FB}) + (1-a)T_{yaw} \qquad (14)$$

$$a = f(\dot{\psi}_{err}) \qquad (15)$$

4. SIMULATION RESULT

This research has been preceded on top of method of "Model-based control design". To prove the performance of proposed control algorithm, computer simulation for driver performance and handling performance has been executed.

4.1. Simulation Configuration

To construct control logic for model base control design Matlab /simulink has been used. Carsim, which is fast in calculation but inexpensive, one of the powerful software in vehicle dynamics has been used. The following figure shows the interlinked simulation environment.
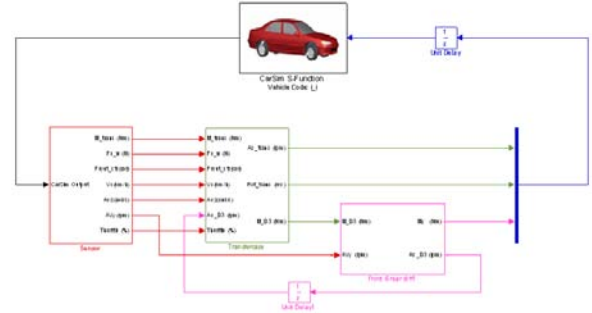

Figure 15: Overall system configuration

4.2. Traction control simulation

Since traction performance is one of the fundamental performances for 4WD, verifications were undergone at three difference conditions. To prove this, conventional look-up table method, PID control method and the proposed equation control methods were compared.
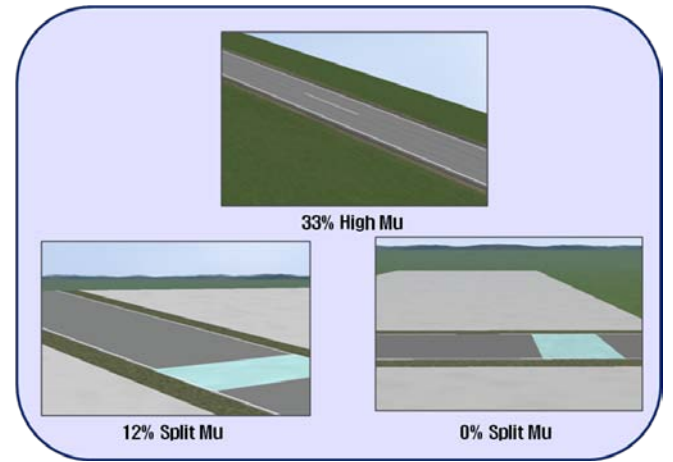

Figure 16: Traction Simulation Condition

The first condition is climbing experiment at 33% high mu. Without controlling the steering wheel, results shown below at Wide Open Throttle situation.

When at PID control in climbing situation, Front wheel torque had to be distributed as widely as possible,

But because of force loss from tire slip, controlling performance was very bad. Opposed, look up table control showed fairly bad performance compared to equation control due to the lack of control amount.
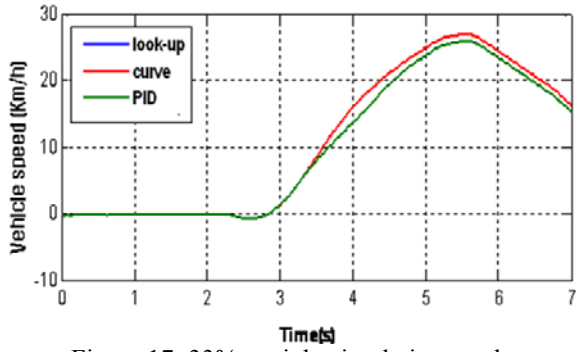

170
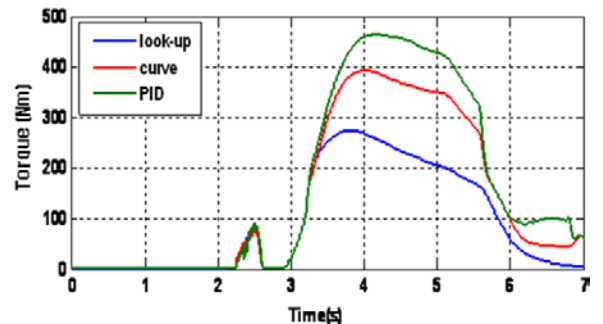
Figure 17: 33% straight simulation result

Second and third conditions are both split mu and the slope is 12% and 0% respectively.

Split mu condition at rear wheel has developed low mu (mu=0.05). So we have tested escape ability for Rear wheel based 4WD vehicle.
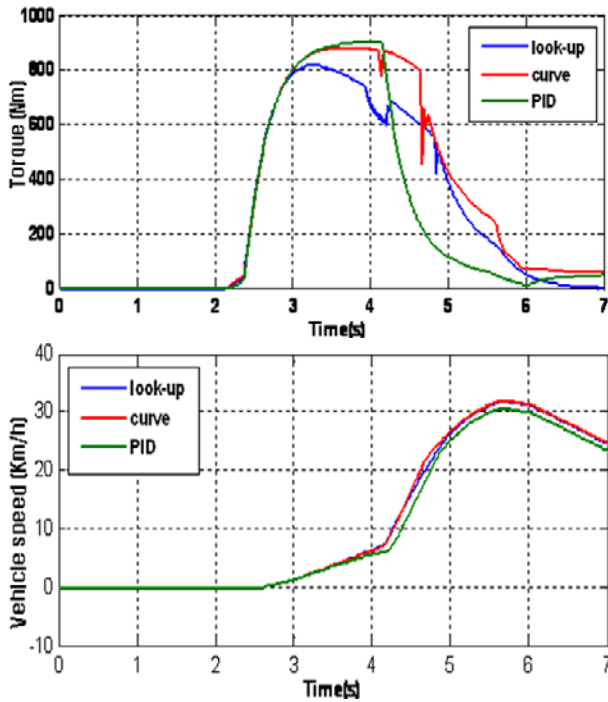

Figure 18: 12% Roller simulation result

As we can see from above, control amount for equation control is almost the same compared to PID and look up table.

The reason is because transfer torque limit is approximately 900Nm. To sum up, compared to vehicle speed, this shows better performance compared to different control method.
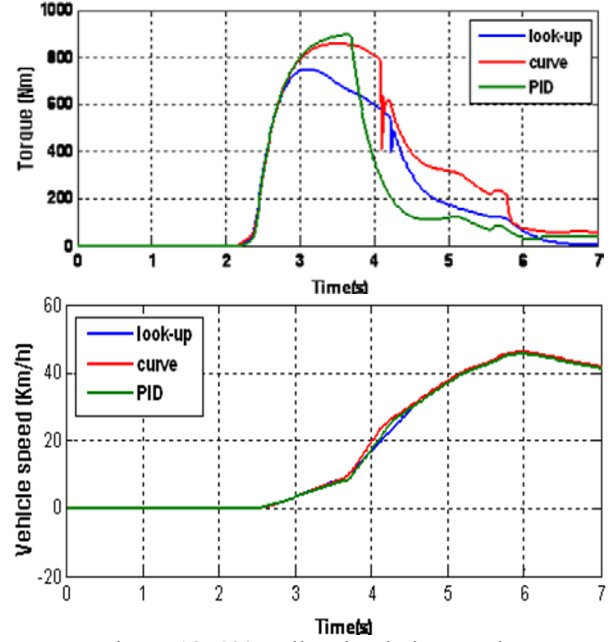

Figure 19: 0% Roller simulation result

We have come up with simulation result similar to that of 12% split mu. This shows the dominance of equation control.

4.3. Yaw control simulation

As mentioned earlier at yaw controller, yaw moment changes due to torque distribution for forward wheel.

This is why we have proved validity by comparing yaw rate and path under fishhook simulation condition.

Table 1: Yaw Simulation Condition

	Simulation 'a'	Simulation 'b'
Vehicle speed	40km/h ➡ 70km/h	40km/h ➡ 60km/h
Steering wheel angle	0° ➡ 60°	0° ➡ 100°
Road Condition	Mu = 0.5 , Slope = 0%	

Comparison results among back-wheel based 2WD model, only traction control model and traction+yaw control model is as follows.

If front-wheel torque increases for traction performance, the path for the vehicle goes under-steer.

Yaw rate also chatters as well. Opposed to this, model which considers both traction and yaw, the driver can follow the path by proper torque distribution.

Ahead of the proposed simulation 'a', 'b' conditions, both the Traction + Yaw model shows a more stable performance.
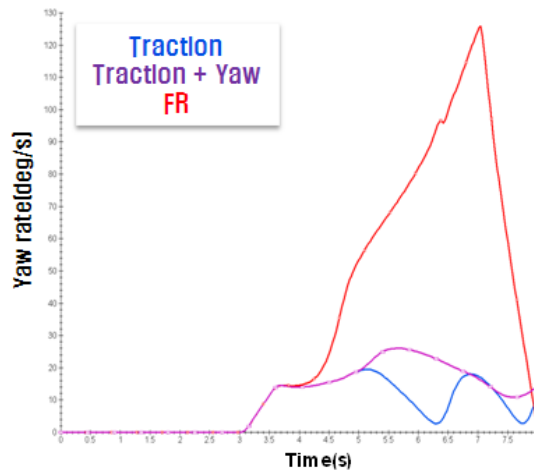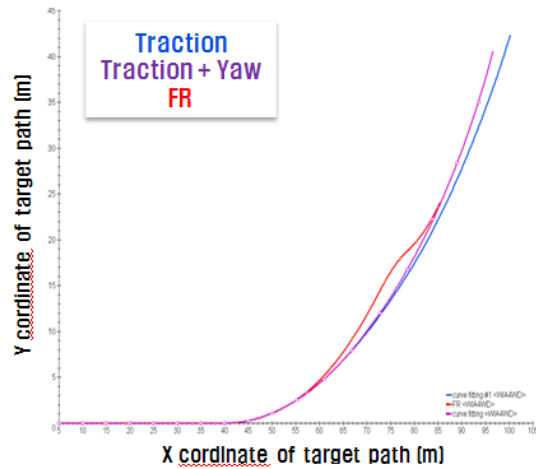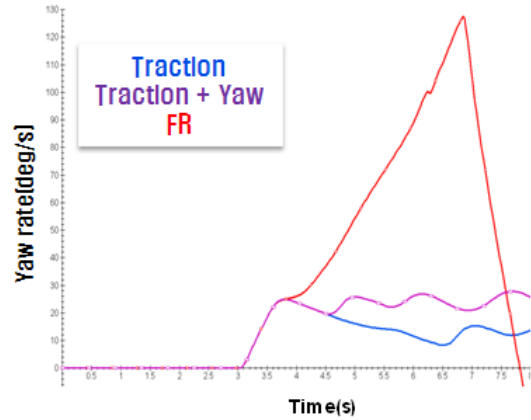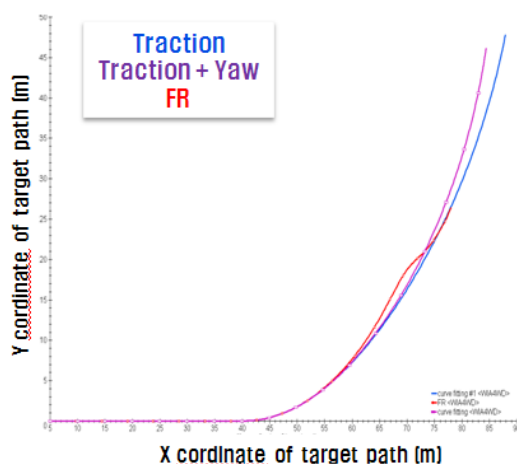

Figure 13: Fishhook simulation 'a' result

Figure 20: Fishhook simulation 'b' result

5. CONCLUSION

This paper presented a new traction and wheel slip control algorithm for the 4WD vehicle. A power-train model was derived for analyzing the torque flow of the 4WD vehicle. The proposed control algorithm was developed by using the Equation Method.

The control algorithm was also implemented and verified through the simulation. The simulation results showed that the proposed control algorithm can transfer the driving torque to the driven wheels and reduce the wheel slip. The proposed control algorithm can improve vehicle performance in terms of the ride comfort and power distribution.

REFERENCES

Hyeongcheol L., "Four-wheel drive control system using a clutchless centre limited slip differential", Proc. IMechE Vol.220 Part D: J. Automobile Engineering, pp.665-681.

McKenzie, C. A. et al., "Ford Explorer control trac 4X4 system", SAE paper 952645, 1995.

Hiroyuki, A. et al., "AWD Vehicle Simulation with the Intelligent Torque Controlled Coupling as a Fully Controllable AWD System", 2005 SAE World Congress.

Mitsuru, O. et al., "Development of a New Electronically Controlled 4WD System: Toyota Active Torque Control 4WD", International Congress and Exposition Detroit, 1999.

Woojong Bong., "AWD Simulation with the Intelligent Torque Split Control Strategy for Improving Traction and Handling performance", 2007 KSAE

E. Galvagno, "Differentials Modeling for Four Wheels Drive", 2006 SAE World Congress

PROCESS MONITORING BASED ON MEASUREMENT SPACE DECOMPOSITION

José L. Godoy[a], Jorge R. Vega[a][b] and Jacinto L. Marchetti[a]

[a] INTEC (CONICET-Universidad Nacional del Litoral), Güemes 3450, 3000, Santa Fe, Argentina
[b] Facultad Regional Santa Fe, UTN, Lavaisse 610, 3000, Santa Fe, Argentina

[a]jlgodoy@santafe-conicet.gov.ar, jvega@santafe-conicet.gov.ar, jlmarch@santafe-conicet.gov.ar.

ABSTRACT

A useful decomposition of the input and output variable spaces is obtained by using the relationships supporting Partial Least Squares Regression (PLSR). The resulting technique is capable to classify faults or anomalies according to four types: those associated to measurements of input variables, those related to measurements of output variables, those linked to the inner latent structure of the complex processes data and those related to upsets that follow its correlation structure. This classification is suggested by the different subspaces in which the whole measurement space can be decomposed by using PLSR modeling and specific statistics actuating on each subspace. Hence, using an in-control PLSR model, the tool is able to detect anomalies and to diagnose its kind. The approach can be used for monitoring closed loop systems and to detect abnormal controller functioning. Several features of the proposed classification technique are analyzed through static and dynamic simulation examples.

Keywords: PLSR, space decomposition, multivariate process monitoring, fault detection indexes.

1. INTRODUCTION

The need for associating input and output data sets obtained by online data login of several variables originated in complex processes constitutes a problem that requires increasing attention. Monitoring strategies are used in industry to detect and diagnose abnormal processes behavior. Fault detection indices are lately being proposed to indicate abnormalities in process operations: a fault signal appears basically when one single or combined index goes beyond its control limit. After a fault is detected, the need of having a diagnostic about its cause becomes almost mandatory. Typically, this diagnostic is performed by analyzing the contributions made by each measured variable to the index has given the alarm.

Lately, PLSR has become a powerful approach to find multivariable linear structures in the data, mainly because significant co-linearity, which is frequently implicit in there, can be adequately overcome.

Nowadays, the chemical process industries appeal to multivariable control technologies for controlling complex multi-input multi-output (MIMO) processes systems. However, in general, it is clear that not only multivariable control systems but also traditional multi-

loop control systems must be monitored for improving quality and performance standards and for stabilizing production throughputs. Hence, monitoring medium or large processes to detect abnormalities or undesirable low performances has become an important research field.

As mentioned before, the multivariate nature of most processes frequently shows highly correlated variables. However, they usually move in an effective subspace of lower dimension, this is because latent variable (LV) methods can transform noisy and correlated data into a smaller informative set free of ill-conditionings. Since PLSR lays on a reduced set of input-output LV's, it allows fitting a MIMO linear model using an identification data set that might have been originally ill-conditioned.

Lately, some applications based on PCA/PLS are being done through data acquisition systems collecting measurements directly from the plant. An important contribution in this area was provided by AlGazzawi and Lennox (2009), who investigated the ability of multivariate statistical process control (MSPC) techniques to monitor industrial model predictive control (MPC) systems. This has also encouraged the development of a data-driven LV-MPC method, where PLSR is used for fitting a multi-step ahead prediction model and the MPC is then implemented in the space of the LV's (Laurí et al., 2010). This type of works confirms the actual efforts to improving the way of managing multivariable data sets.

This work attempts a contribution in this area by providing an online data processing technique capable of orientating process engineers towards the main causes of alarming faults signals and undesirables performances. Hence, the presentation is as follow: in Section 2 we recall the main relationships arising when developing a PLSR model; in particular the Sub-section 2.1 formalizes the PLSR implicit decomposition and describes several geometric properties useful for better understanding of the contribution. In Section 3 the main statistical tools are introduced and their roll in the measurement decomposition is described. Sub-section 3.2 and the numerical application example in Sub-section 4.1 have a core importance in understanding the proposed technique. Sub-section 4.2 presents a more realistic case through the application to a non-isothermal continuous stirred tank reactor (CSTR). Finally, the conclusions are given in Section 5.

2. PLSR MODELING

The PLSR model developed here is calculated by simultaneously deflating the data matrices, using the classical PLS-NIPALS algorithm (Geladi and Kowalski, 1986). This procedure gives better results for multivariate prediction and for process monitoring that others alternative PLS algorithms (Gang et al., 2010). Besides, the simultaneous deflation on both data matrices allows the detection of predictor variables that can play an interfering effect (Godoy et al., 2011).

Given a predictor matrix $\mathbf{X}=[\mathbf{x}_1 \ ... \ \mathbf{x}_N]'$ ($N \times m$), consisting of N samples with m variables per sample, and a response matrix $\mathbf{Y}=[\mathbf{y}_1 \ ... \ \mathbf{y}_N]'$ ($N \times p$), with p variables per sample, PLSR can be used to find a regression model between the measurement vectors $\mathbf{x}=[x_1...x_m]'$ and $\mathbf{y}=[y_1...y_p]'$, even when their correlation matrixes ($\mathbf{R}_x$ and $\mathbf{R}_v$) are both positive semi-definite (i.e. $\mathbf{X}$ and $\mathbf{Y}$ have collinear variables). The method projects $\mathbf{X}$ and $\mathbf{Y}$ onto correlated low-dimension spaces defined by a common (small) number of A latent variables. The implicit objective of the PLS-NIPALS algorithm in each a-run is to find the solution of the following problem:

$$\max_{\mathbf{w}_a, \mathbf{q}_a}\left(\mathbf{w}'_a \mathbf{X}'_a \mathbf{Y}_a \mathbf{q}_a\right) \quad s.t. \quad \|\mathbf{w}_a\|=1, \|\mathbf{q}_a\|=1 \quad (1)$$

where $\mathbf{X}_a$ and $\mathbf{Y}_a$ are the deflated a-versions of $\mathbf{X}_1=\mathbf{X}$ and $\mathbf{Y}_1=\mathbf{Y}$.

This provides an external, an internal, and a regression model. The external model decomposes $\mathbf{X}$ and $\mathbf{Y}$ in score vectors ($\mathbf{t}_a$ and $\mathbf{u}_a$), weight vectors ($\mathbf{p}_a$ and $\mathbf{q}_a$), and residual error matrices ($\widetilde{\mathbf{X}}$ and $\widetilde{\mathbf{Y}}_2$), as follows:

$$\mathbf{X} = \mathbf{TP}' + \widetilde{\mathbf{X}} = \sum_{a=1}^{A}\mathbf{t}_a\mathbf{p}'_a + \widetilde{\mathbf{X}} = \widehat{\mathbf{X}} + \widetilde{\mathbf{X}}, \quad (2)$$

$$\mathbf{Y} = \mathbf{UQ}' + \widetilde{\mathbf{Y}}_2 = \sum_{a=1}^{A}\mathbf{u}_a\mathbf{q}'_a + \widetilde{\mathbf{Y}}_2 = \widehat{\mathbf{Y}}^* + \widetilde{\mathbf{Y}}_2, \quad (3)$$

where $\mathbf{P}=\left[\mathbf{p}_1...\mathbf{p}_A\right]$, $\mathbf{Q}=\left[\mathbf{q}_1...\mathbf{q}_A\right]$, $\mathbf{T}=\left[\mathbf{t}_1...\mathbf{t}_A\right]$, and $\mathbf{U}=\left[\mathbf{u}_1...\mathbf{u}_A\right]$ are orthogonal by columns. Call $\mathbf{R}$ to the pseudo-inverse of $\mathbf{P}'$ ($\mathbf{P}'\mathbf{R} = \mathbf{R}'\mathbf{P} = \mathbf{I}$); then, the prediction of $\mathbf{T}$ is directly obtained from Eq. (2), as: $\mathbf{T} = \mathbf{XR}$, because the row space of $\widetilde{\mathbf{X}}$ belongs to the null-space of the linear transformation $\mathbf{R}'$, i.e. $\widetilde{\mathbf{X}}\mathbf{R}=0$ (Meyer, 2000). If $\mathbf{P}' = \mathbf{W}_A\mathbf{\Sigma}_A\mathbf{V}'_A$ is the compact singular value decomposition (SVD) of $\mathbf{P}'$, then $\mathbf{R}=(\mathbf{P}')^- = \mathbf{V}_A\mathbf{\Sigma}_A^{-1}\mathbf{W}'_A$ where $^-$ denote a generalized inverse (Meyer 2000). Equivalently, from Eq. (3), $\mathbf{S}$ is the pseudo-inverse of $\mathbf{Q}'$ ($\mathbf{Q}'\mathbf{S}=\mathbf{S}'\mathbf{Q}=\mathbf{I}$), and since $\widetilde{\mathbf{Y}}_2\mathbf{S}=0$, then: $\mathbf{U}=\mathbf{YS}$. For the internal model, $\mathbf{t}_a$ is linearly regressed against the y-score vector $\mathbf{u}_a$, i.e.:

$$\mathbf{U} = \mathbf{TB} + \mathbf{H} = \widehat{\mathbf{U}} + \mathbf{H}, \quad \mathbf{B} = diag(b_1...b_A) \quad (4)$$

where $b_1...b_A$ are the regression coefficients determined by minimization of the residual matrix $\mathbf{H}$. Then, the following $\mathbf{X}$-$\mathbf{Y}$ regression model is obtained:

$$\mathbf{Y} = \mathbf{XRBQ}' + \mathbf{HQ}' + \widetilde{\mathbf{Y}}_2 = \widehat{\mathbf{Y}} + \widetilde{\mathbf{Y}}_1 + \widetilde{\mathbf{Y}}_2 \quad (5)$$

where $\widetilde{\mathbf{Y}}_2 = \mathbf{Y} - \mathbf{YSQ}'$ and $\widetilde{\mathbf{Y}}_1 = \mathbf{YSQ}' - \widehat{\mathbf{Y}}$. Also, $\widehat{\mathbf{X}} = \mathbf{XRP}'$ verifies:

$$\widehat{\mathbf{X}}\mathbf{RBQ}' = \mathbf{XRBQ}' = \widehat{\mathbf{Y}} \quad (6)$$

The selection of A is determined by supervising the simultaneous deflation of $\mathbf{X}$ and $\mathbf{Y}$.

2.1. PLS decomposition of the input and output spaces

After synthesizing an *in-control* PLSR model, the measurement vectors $\mathbf{x} \in \mathbb{R}^m$ and $\mathbf{y} \in \mathbb{R}^p$ can be decomposed as described in the following Lemma (see Proof 1 in Appendix).

Lemma 1. *Call* $\mathbf{\Pi}_{\mathbf{P}|\mathbf{R}^\perp}$ *(*$\mathbf{\Pi}_{\mathbf{Q}|\mathbf{S}^\perp}$*) the projector on the model subspace* $Span\{\mathbf{P}\}\subseteq\mathbb{R}^m$ *(*$Span\{\mathbf{Q}\}\subseteq\mathbb{R}^p$*), along the residual subspace* $Span\{\mathbf{R}\}^\perp$ *(*$Span\{\mathbf{S}\}^\perp$*). Then:*

$$\mathbf{\Pi}_{\mathbf{P}|\mathbf{R}^\perp} = \mathbf{PR}', \quad \mathbf{\Pi}_{\mathbf{R}^\perp|\mathbf{P}} = \mathbf{I} - \mathbf{PR}', \quad (7)$$

$$\mathbf{\Pi}_{\mathbf{Q}|\mathbf{S}^\perp} = \mathbf{QS}', \quad \mathbf{\Pi}_{\mathbf{S}^\perp|\mathbf{Q}} = \mathbf{I} - \mathbf{QS}', \quad (8)$$

where $^\perp$ *denotes the orthogonal complement of the subspace.*

From Lemma 1, we propose the following theorem on the PLS decomposition (see Proof 2 in Appendix).

Theorem 1. *Input and output variable spaces can be decomposed (by PLSR) in complementary oblique subspaces, with both modeled subspaces interrelated according to:*

$$\mathbf{x} = \widehat{\mathbf{x}} + \tilde{\mathbf{x}} \in \mathbb{R}^m,$$

$$\widehat{\mathbf{x}} = \mathbf{PR}'\mathbf{x} \in S_{MX} \equiv Span\{\mathbf{P}\},$$

$$\tilde{\mathbf{x}} = \mathbf{x} - \widehat{\mathbf{x}} = \left(\mathbf{I} - \mathbf{PR}'\right)\mathbf{x} \in S_{RX} \equiv Span\{\mathbf{R}\}^\perp \quad (9)$$

$$\mathbf{y} = \widehat{\mathbf{y}}^* + \tilde{\mathbf{y}}_2 \in \mathbb{R}^p,$$

$$\widehat{\mathbf{y}}^* = \mathbf{QS}'\mathbf{y} \in S_{MY} \equiv Span\{\mathbf{Q}\},$$

$$\tilde{\mathbf{y}}_2 = \mathbf{y} - \widehat{\mathbf{y}}^* = \left(\mathbf{I} - \mathbf{QS}'\right)\mathbf{y} \in S_{RY} \equiv Span\{\mathbf{S}\}^\perp \quad (10)$$

$$\widehat{\mathbf{y}}^* = \widehat{\mathbf{y}} + \tilde{\mathbf{y}}_1,$$

$$\widehat{\mathbf{y}} = \mathbf{QBR}'\widehat{\mathbf{x}} \in S_{MY},$$

$$\tilde{\mathbf{y}}_1 = \widehat{\mathbf{y}}^* - \widehat{\mathbf{y}} = \mathbf{QS}'\mathbf{y} - \mathbf{QBR}'\mathbf{x} \in S_{MY} \quad (11)$$

The **x** and **y** modeled projections are related to the latent spaces and to each other, as follows:

$$\mathbf{t} = \mathbf{R}'\mathbf{x}, \quad \hat{\mathbf{u}} = \mathbf{B}\mathbf{t} = \mathbf{S}'\hat{\mathbf{y}}, \quad \hat{\mathbf{x}} = \mathbf{P}\mathbf{t}, \quad \hat{\mathbf{y}} = \mathbf{Q}\hat{\mathbf{u}}, \quad (12)$$

where $\mathbf{t} = [t_1 \cdots t_A]'$ and $\hat{\mathbf{u}} = [\hat{u}_1 \cdots \hat{u}_A]'$ are the latent coordinate vectors on the model hyper planes. Hence, their correlation matrixes are related through:

$$\boldsymbol{\Lambda} = (N-1)^{-1}\mathbf{T}'\mathbf{T} = diag(\lambda_1 ... \lambda_A), \quad (13a)$$

$$\boldsymbol{\Delta} = (N-1)^{-1}\hat{\mathbf{U}}'\hat{\mathbf{U}} = \mathbf{B}\boldsymbol{\Lambda}\mathbf{B} = diag(\delta_1 ... \delta_A), \quad (13b)$$

$$\mathbf{R}_{\hat{\mathbf{x}}} = (N-1)^{-1}\hat{\mathbf{X}}'\hat{\mathbf{X}} = \mathbf{P}\boldsymbol{\Lambda}\mathbf{P}' \quad (14a)$$

$$\mathbf{R}_{\hat{\mathbf{y}}} = (N-1)^{-1}\hat{\mathbf{Y}}'\hat{\mathbf{Y}} = \mathbf{Q}\boldsymbol{\Delta}\mathbf{Q}' \quad (14b)$$

where λ_i and δ_i are the estimated variances of t_i and $\hat{u}_i$, respectively.

Figure 1 illustrates all the geometric properties mentioned in this work. Each measurement vector is decomposed and their projections are compared with their limits.
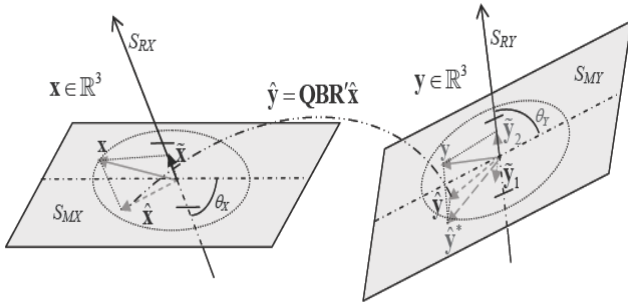

Figure 1: Induced PLS-decomposition with their relations and control limits.

3. PROCESS MONITORING BASED ON PLSR
3.1. Fault detection indexes
The multivariate process monitoring strategy uses statistical indexes associated to each subspace for fault detection. Based on the *in-control* PLSR model, we can analyze any future process behavior by projecting the new **x** and **y** measurements onto each subspace. Thus, for detecting a significant change in S_{MX}, the following Hotelling's T^2 statistic for **t** is defined:

$$T_t^2 = \left\| \boldsymbol{\Lambda}^{-1/2}\mathbf{t} \right\|^2 = \left\| \left(\boldsymbol{\Lambda}^{-1/2}\mathbf{R}' \right)\mathbf{x} \right\|^2 = \left\| \left(\boldsymbol{\Lambda}^{-1/2}\mathbf{P}' \right)\hat{\mathbf{x}} \right\|^2 \quad (15)$$

When a new special event (originally not considered by the in-control PLS model) occurs, the new observation **x** will move out from S_{MX} into S_{RX}. The squared prediction error of **x** (SPE_X), or distance to the **x**-model, is defined as:

$$SPE_X = \left\| \tilde{\mathbf{x}} \right\|^2 = \left\| (\mathbf{I} - \mathbf{P}\mathbf{R}')\mathbf{x} \right\|^2 \quad (16)$$

Then, SPE_X can be used for detecting a change in S_{RX}. Similarly, the T^2 statistic for $\hat{\mathbf{u}}$, is given by:

$$T_u^2 = \left\| \boldsymbol{\Delta}^{-1/2}\hat{\mathbf{u}} \right\|^2 = \left\| \left(\boldsymbol{\Delta}^{-1/2}\mathbf{S}' \right)\hat{\mathbf{y}} \right\|^2. \quad (17)$$

In the same way, this statistic can be used for detecting changes in S_{MY}, and the distance to the regression model in S_{MY} is defined by:

$$SPE_{Y1} = \left\| \tilde{\mathbf{y}}_1 \right\|^2 = \left\| \begin{bmatrix} \mathbf{QS}' & -\mathbf{QBR}' \end{bmatrix} \begin{bmatrix} \mathbf{y} \\ \mathbf{x} \end{bmatrix} \right\|^2 = \left\| \mathbf{u} - \hat{\mathbf{u}} \right\|^2 \quad (18)$$

and the distance to the **y**-model in S_{RY} is defined as:

$$SPE_{Y2} = \left\| \tilde{\mathbf{y}}_2 \right\|^2 = \left\| (\mathbf{I} - \mathbf{QS}')\mathbf{y} \right\|^2. \quad (19)$$

Frequently, $\mathbf{R}_{\hat{\mathbf{x}}}$ and $\mathbf{R}_{\hat{\mathbf{y}}}$ are singular. Then, the generalized Mahalanobis' distance for $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ are:

$$D_{\hat{\mathbf{x}}} = \hat{\mathbf{x}}'\mathbf{R}_{\hat{\mathbf{x}}}^- \hat{\mathbf{x}}, \quad (20a)$$

$$D_{\hat{\mathbf{y}}} = \hat{\mathbf{y}}'\mathbf{R}_{\hat{\mathbf{y}}}^- \hat{\mathbf{y}}. \quad (20b)$$

Besides, Proof 3 in Appendix demonstrate the following

Theorem 2. *The metrics on* $\hat{\mathbf{x}}$, $\mathbf{t}$, $\hat{\mathbf{u}}$ *or* $\hat{\mathbf{y}}$ *are equivalents, i.e.:* $D_{\hat{\mathbf{x}}} = T_t^2 = T_u^2 = D_{\hat{\mathbf{y}}}$.

The above theorem 2 tells that output variables can be monitored through a PLSR-based metric of the input variables. Therefore, we propose monitoring using four non-overlapped metrics (SPE_X, T_t^2, SPE_{Y1}, and SPE_{Y2}) which completely cover both measurement spaces, each one on a different subspace.

3.2. Fault diagnosis by means of alarmed subspaces
An anomaly is a change in the measurements following or not the correlation structure captured by the PLSR model. If the change produces an out-of-control point, the anomaly source can be classified according to a) an excessively large operation change of the normal operation; b) a significant increase of variability; c) the alteration of cross-correlations, and d) sensor faults. Cases a) and b) involve changes in the measurement vector following the modeled correlation structure; while cases c) and d) involve changes in some variables altering the correlation pattern with the others. In fact, an abnormal process behavior involves a deviation of the modeled correlations, thus increasing the value of the proper metric being used. In order to classify the abnormalities, we analyze the effect on each subspace (see Table 1). Rows 1), 2), 3) and 6) feature complex process changes; while rows 4) and 5) represent localized sensor faults. By analyzing the contributions to an alarmed index (Alcala and Qin, 2011), such as SPE_X (or SPE_Y), it could be possible to

discriminate changes in the **X**- / **Y**-outer part against sensor fault in **x** / **y** (see this ambiguity in Table 1).

In summary, the proposed monitoring strategy is based on an input and output space PLS decomposition, which classifies the type of process fault or anomaly according to the statistic that triggers the alarm condition.

Table 1. Fault diagnosis based on alarmed index.

Fault/Anomaly in	T_t^2	SPE_X	SPE_{Y1}	SPE_{Y2}
1- Inner part, $d\mathbf{B}$	–	–	×	–
2- **X** outer part, $d\mathbf{P}$	–	×	o	–
3- **Y** outer part, $d\mathbf{Q}$	–	–	o	×
4- **x** sensor	–	×	o	–
5- **y** sensor	–	–	o	×
6- latent space upset	×	–	–	–

×: high value. –: negligible value. o: high/low value.

4. SIMULATION EXAMPLES
4.1. A numerical evaluation example
To understand the proposed decompositions and statistics as monitoring tools, we simulated different faults in a synthetic system. The system includes an internal and an external part and an input-output relation given by:

$$\mathbf{t} = \mathbf{t}^0 + \xi, \quad \xi \sim N\left(\mathbf{0}, 0.1^2 \mathbf{I}_2\right),$$

$$\mathbf{u} = \mathbf{Bt} + \gamma, \quad \mathbf{B} = diag\left(2, 0.5\right), \quad \gamma \sim N\left(\mathbf{0}, 0.05^2 \mathbf{I}_2\right),$$

$$\mathbf{x} = \mathbf{Pt} + \tilde{\mathbf{x}}, \quad \mathbf{P} = \left[\mathbf{p}_1 \quad \mathbf{p}_2\right], \quad \tilde{\mathbf{x}} \sim N\left(\mathbf{0}, 0.05^2 \mathbf{I}_7\right),$$

$$\mathbf{y} = \mathbf{Qu} + \tilde{\mathbf{y}}_2, \quad \mathbf{Q} = \left[\mathbf{q}_1 \quad \mathbf{q}_2\right], \quad \tilde{\mathbf{y}}_2 \sim N\left(\mathbf{0}, 0.05^2 \mathbf{I}_3\right),$$

$$\hat{\mathbf{y}}^* = \mathbf{QBR'x} + \mathbf{Q}\gamma, \quad \mathbf{p}_i = \mathbf{p}_i^* / \left\|\mathbf{p}_i^*\right\|, \quad \mathbf{q}_j = \mathbf{q}_j^* / \left\|\mathbf{q}_j^*\right\|,$$

$$\mathbf{p}_1^* = \left[1.5, 0, 2, 1, 0.5, 0, 2.5\right]',$$

$$\mathbf{p}_2^* = \left[0, 2.5, 0.5, -0.5, -1.0, 1.5, 0\right]',$$

$$\mathbf{q}_1^* = \left[1.5, 1, 0.5\right]', \quad \mathbf{q}_2^* = \left[0, 0.5, -1\right]',$$

where $\mathbf{R'}$ is the pseudo-inverse of $\mathbf{P}$. The normal process operation follows a sequence of 4 states: $\left\{\left(t_1^0, t_2^0\right)\right\}_{1...4} = \left\{(1,1), (1,3), (3,3), (3,1)\right\}$. Matrixes **X** and **Y** are obtained on the basis of 32 realizations of **x** and **y**. The PLSR model was adjusted with centered data, without scaling, to identify a centered version of the latent process. The differences between the identified matrixes and the true ones are negligible, although the columns of $\hat{\mathbf{R}}$ and $\hat{\mathbf{Q}}$ may include opposite signs to those of **R** and **Q**. Since such sign alteration is present in both matrixes, then: $\mathbf{QBR'} \cong \hat{\mathbf{Q}}\hat{\mathbf{B}}\hat{\mathbf{R}}'$. Since $\hat{\mathbf{y}} = \sum_{a=1}^{A} b_a \mathbf{q}_a \mathbf{r}_a' \mathbf{x}$, when $\hat{u}_2$ exhibits an opposite sign to that of the real u_2 (e.g., because $\hat{\mathbf{r}}_2 = -\mathbf{r}_2$), then $\hat{\mathbf{q}}_2 = -\mathbf{q}_2$. Six faults were simulated (see Table 1): 1) a change of the internal model at $k=11$: $db_2 = 0.7 \ b_2$; 2)

change of the process coefficient (plant) at $k=19$, given by $d\mathbf{p}_2 = [0, 0.32, 0, 0, -0.08, 0.16, -0.16]'$ (high SPE_X if $d\mathbf{p}_2 \in S_{RX}$), and at $k=27$; 3) $d\mathbf{q}_1 = [-0.08, 0.06, 0]'$ (high SPE_{Y2} if $d\mathbf{q}_1 \in S_{RY}$); 4) a multiple sensor fault in the measurement **x** at $k = 35$, given by $\mathbf{x} = \mathbf{x} + d\mathbf{x}$, with $d\mathbf{x} = [0.25 \ 0 \ 0 \ 0 \ 0 \ 0.25 \ 0]'$; 5) a multiple sensor fault in the measurement **y** at $k=43$, with $d\mathbf{y} = [-0.23, 0.27, 0.14]' \in S_{RY}$, and 6) a change (following the modeled correlation structure) at $k = 51$, given by $d\hat{\mathbf{x}} = [0, 3.1623, 0.6325, -0.6325, -1.2649, 1.8974, 0] \equiv \mathbf{P} [0 \ 4]'$. Table 1 shows the diagnosis expected in each simulated fault/anomaly. Figure 2 shows the evolution of each statistic with time, where it is possible to detect and classify the type of simulated fault on the basis of the information given in Table 1.
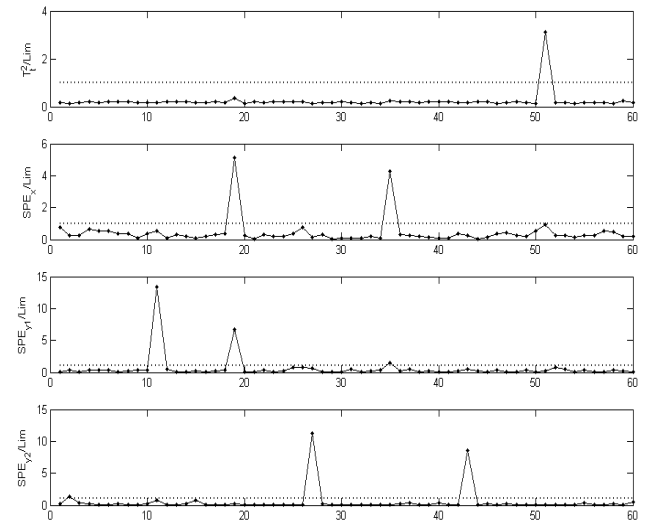

Figure 2: Time evolution of each normalized PLSR-statistic.

4.2. Monitoring of a controlled reactor
A simulated case study is used to evaluate the performance of the proposed monitoring technique for a variety of fault scenarios. The process is a continuous stirred tank reactor (CSTR) with a cooling jacket. The chemical kinetics is characterized by a classical first-order irreversible reaction A→B. The instrumentation includes level, flow rate and temperature transmitters and two PI controllers, with set-points SP_T and SP_h. The controlled variables are the reactor temperature (T) and the tank level (h). The manipulated variables are the coolant flow-rate (Q_J) and the output flow-rate from the reactor (Q). Input conditions T_F, T_{JF}, and C_{AF} are assumed constant to the values in Table 2. Figure 3 shows a schematic diagram of the reactor.

The mathematical model of the system is represented by (Singhal and Seborg, 2002):

$$\frac{dC_A}{dt} = -k_0 e^{-E/RT} C_A + \frac{Q_F C_{AF} - Q C_A}{Ah},$$

$$\frac{dT}{dt} = \frac{(-\Delta H)k_0 e^{-E/RT} C_A}{\rho C_p} + \frac{Q_F T_F - QT}{Ah} + \frac{UA_J\left(T_J - T\right)}{\rho C_p Ah},$$

$$\frac{dT_J}{dt} = \frac{Q_J \left(T_{JF} - T_J \right)}{V_C} + \frac{UA_J \left(T - T_J \right)}{\rho_J C_{pJ} V_C},$$

$$\frac{dh}{dt} = \frac{Q_F - Q}{A},$$

$$\frac{dQ_J}{dt} = K_{C1} \frac{d(SP_T - T)}{dt} + \frac{K_{C1}}{T_{i1}} (SP_T - T),$$

$$\frac{dQ}{dt} = K_{C2} \frac{d(SP_h - h)}{dt} + \frac{K_{C2}}{T_{i2}} (SP_h - h),$$

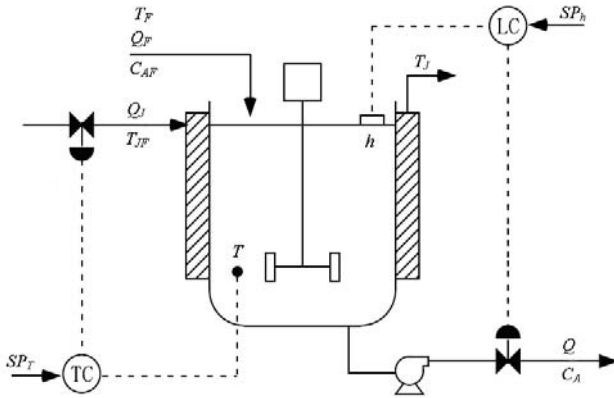$$x = \frac{C_{AF} - C_A}{C_{AF}}, \qquad \dot{H} = \Delta H\, k_0 e^{-E/RT} C_A A h.$$

Figure 3: Schematic of the CSTR system with a multi-loop control system.

Table 2: Nominal Operating Conditions and Model Parameters for the CSTR in Fig. 3

$Q = 100$ L/min	$A = 0.1666$ m^2
$Q_J = 15$ L/min	$k_0 = 7.2 \times 10^{10}$ min^{-1}
$T_F = 320$ °K	$\Delta H = -5 \times 10^4$ J/mol
$T_{JF} = 300$ °K	$\rho C_p = 239$ J/(L °K)
$T = 402.35$ °K	$\rho_J C_{pJ} = 4175$ J/(L °K)
$T_J = 345.44$ °K	$E/R = 8750$ °K
$C_{AF} = 1.0$ mol/L	$UA_J = 5 \times 10^4$ J/(min °K)
$C_A = 0.037$ mol/L	$V_J = 10$ L
$h = 0.6$ m	$K_{C1} = -1,\ T_{i1} = 1,\ K_{C2} = -15,\ T_{i2}=10,$

Since data are auto-correlated, **X** and **Y** include time-lagged measurements to a given order. Notice that PLSR modeling works in this case as an alternative identification method for discrete dynamics like ARX. In this application, the adopted ARX order is 3, so the representation takes the form

$$\hat{\mathbf{y}}(k) = [\mathbf{QBR'}]\mathbf{x}(k) = [\mathbf{B}_1 \mathbf{A}_1 \mathbf{A}_2 \mathbf{A}_3] \begin{bmatrix} \mathbf{z}(k-1) \\ \mathbf{y}(k-1) \\ \mathbf{y}(k-2) \\ \mathbf{y}(k-3) \end{bmatrix} =$$

$$= \mathbf{B}_1 \mathbf{z}(k-1) + \mathbf{A}_1 \mathbf{y}(k-1) + \mathbf{A}_2 \mathbf{y}(k-2) + \mathbf{A}_3 \mathbf{y}(k-3) \quad (21)$$

where $\mathbf{y} = \begin{bmatrix} Q_J & Q & C_A & T & T_J & h & x & \dot{H} \end{bmatrix}'$ and $\mathbf{z} = \begin{bmatrix} SP_T & SP_h & Q_F \end{bmatrix}'$ are centered and scaled. The

sampling time is adjusted to T_s=5 min, a quite reasonable time interval considering an approximated settling time of T_{st}=120 min. To build the identification data set, 50 random set-point changes, each one every 120 min., around the nominal values are implemented, and the step responses of the system were collected along 120 min after each step change. The input variables in **z** followed uniform distributions around the nominal value, with ranges SP_h: 0.6 ± 0.05 m, SP_T: 402.35 ± 3 °K, and Q_F: 100 ± 5 L/min. The dynamic in Eq. (21) is adjusted by PLSR modeling and gave a standard error of = 4.2%. The determined model order is A=5 and, the post modeling verification with the Variable Importance in Projection Index (VIP) - for each x_i (i=1…27) - gives greater than 0.8 for all of them, indicating that all the used predictors are significant (Godoy et al., 2011).

Figure 4 shows the time evolution of each statistic normalized by its control limit, which clearly exhibit the four simulated abnormalities (see Table 3). In the samples k = 121, 122 and 123, the discrete ARX model fails to follow the fast change experienced by the continuous process, resulting in a high SPE_{Y1}, SPE_{Y2}, SPE_X. During the first transition (starting at k = 121) K_{C2} is away from the proper tuning. On the contrary, during the second transition (starting at k = 141) K_{C2} is close to the right tuning. Thus, the control system takes a longer time for stabilizing the controlled variables during the first transition than when the process system goes back to the previous condition (see Fig. 4a).

Table 3: Simulated fault scenarios.

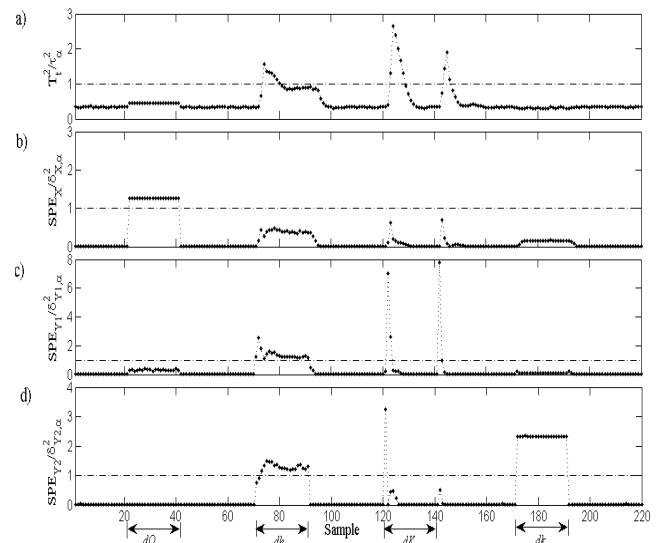| Type	Location	Magnitude	Diagnosis
sensor fault	k=21..41	dQ_F=20 L/min	2/4
sensor fault propagated	k=71..91	$dh = 0.15$ m	6 and 3/5
system fault	k=121..141	$dK_{C2} = -0.75$	3/5→6
process fault -disturbance	k=171..191	dk_0= 0.36 10^{10} min^{-1}	3/5

Figure 4: Simulation of statistical monitoring during 4 different faults on the closed-loop operation of CSTR.

The simulation results show that the developed prototypes are able to identify abnormalities attributed to poor control performance, process upsets and disturbances, as well as sensor faults. The general approach proposed for closed-loop monitoring is based on designing a PLSR model that should be then used online to identify abnormalities. The model supporting this monitoring approach is based on operating data taken along periods where the controller is assumed to be working satisfactorily. Process data recollection constitutes a critical step when developing empirical models for monitoring control performances; it is highly desirable to be operating as close as possible to an optimal condition during the period in which the controlled process data is collected.

5. CONCLUSIONS AND FUTURE WORKS

The multivariate process monitoring approach proposed in this work is based on a PLSR model that represents 'in-control' conditions. Thus, a meaningful deviation of the variables from their expected trajectories serves for the detection and diagnosis of abnormal process behaviors. The results obtained using two simulation examples illustrate that the proposed strategy is efficient and accurate enough as to deserve additional future work. In particular, a controlled CSTR is successfully monitored during different fault conditions or process abnormalities using the new detection indices. The combination of alarm/no-alarm conditions occurring during the simultaneous use of the proposed indices can be used for locating and classifying the disturbing event and contributes to a preliminary diagnosis.

APPENDIX A

Proof 1. The oblique projector onto $Span\{\mathbf{A}\}$ along $Span\{\mathbf{B}\}$ can be obtained by the following equation (Meyer, 2000):

$$\mathbf{\Pi}_{\mathbf{A}|\mathbf{B}} = \mathbf{A}\left(\mathbf{A}'\mathbf{\Pi}_{\mathbf{B}}^{\perp}\mathbf{A}\right)^{-1}\mathbf{A}'\mathbf{\Pi}_{\mathbf{B}}^{\perp} \qquad (A1)$$

where $\mathbf{\Pi}_{\mathbf{B}}^{\perp}$ is the orthogonal projector onto $Span\{\mathbf{B}\}^{\perp}$. Since $\mathbf{R}$ and $\mathbf{S}$ are full-column-ranked, then:

$$\mathbf{\Pi}_{\mathbf{R}^{\perp}}^{\perp} = \mathbf{\Pi}_{\mathbf{R}} = \mathbf{R}\left(\mathbf{R}'\mathbf{R}\right)^{-1}\mathbf{R}' = \mathbf{R}\mathbf{R}', \qquad (A2)$$

$$\mathbf{\Pi}_{\mathbf{S}^{\perp}}^{\perp} = \mathbf{\Pi}_{\mathbf{S}} = \mathbf{S}\left(\mathbf{S}'\mathbf{S}\right)^{-1}\mathbf{S}' = \mathbf{S}\mathbf{S}'. \qquad (A3)$$

Since $\mathbf{P}'\mathbf{R} = \mathbf{R}'\mathbf{P} = \mathbf{I}$ (or $\mathbf{Q}'\mathbf{S} = \mathbf{S}'\mathbf{Q} = \mathbf{I}$), then Eq. (A1) and Eq. (A2) [or Eq. (A3)] yield: $\mathbf{\Pi}_{\mathbf{P}|\mathbf{R}^{\perp}} = \mathbf{P}\mathbf{R}'$ (or $\mathbf{\Pi}_{\mathbf{Q}|\mathbf{S}^{\perp}} = \mathbf{Q}\mathbf{S}'$). Similarly, we have $\mathbf{\Pi}_{\mathbf{R}^{\perp}|\mathbf{P}} = \mathbf{I} - \mathbf{P}\mathbf{R}'$ (or $\mathbf{\Pi}_{\mathbf{S}^{\perp}|\mathbf{Q}} = \mathbf{I} - \mathbf{Q}\mathbf{S}'$). □

Partially, Lemma 1 has already been proved by Gang et al., 2010). Each oblique projector in Eq. (7, 8) works as an identity matrix applied onto every vector belonging to its range.

Proof 2. Eqs. (9, 10) can be proved by taking into account that: (i) $Span\{\mathbf{I} - \mathbf{P}\mathbf{R}'\} = Span\{\mathbf{R}\}^{\perp}$ and $Span\{\mathbf{I} - \mathbf{Q}\mathbf{S}'\} = Span\{\mathbf{S}\}^{\perp}$ (see Lemma 1); and (ii) the projections belong to complementary subspaces, because $rank\left(\mathbf{P}\mathbf{R}'\left(\mathbf{I} - \mathbf{P}\mathbf{R}'\right)\right) = \dim(S_{MX}) + \dim(S_{RX})$ $= m$ and $rank\left(\mathbf{Q}\mathbf{S}'\left(\mathbf{I} - \mathbf{Q}\mathbf{S}'\right)\right) = \dim\left(S_{MY}\right) + \dim\left(S_{RY}\right) = p$. Then, Eq. (11) is directly derived from Eq. (6). □

Proof 3. Since $\mathbf{Q}$ is orthogonal by columns, the property of the generalized inverse of a SVD yields: $\mathbf{R}_{\hat{\mathbf{y}}}^{-} = \left(\mathbf{Q}\mathbf{\Delta}\mathbf{Q}'\right)^{-} = \mathbf{Q}\mathbf{\Delta}^{-1}\mathbf{Q}'$. Then, from Eq. (20b): $D_{\hat{\mathbf{y}}} = \hat{\mathbf{y}}'\mathbf{R}_{\hat{\mathbf{y}}}^{-}\hat{\mathbf{y}} = \hat{\mathbf{y}}'\mathbf{Q}\mathbf{\Delta}^{-1}\mathbf{Q}'\hat{\mathbf{y}} = \hat{\mathbf{u}}'\mathbf{\Delta}^{-1}\hat{\mathbf{u}} = T_u^2$. Equivalently, by replacing $\mathbf{R}_{\hat{\mathbf{x}}}^{-} = \left(\mathbf{P}\mathbf{\Lambda}\mathbf{P}'\right)^{-} = \mathbf{P}\mathbf{\Lambda}^{-1}\mathbf{P}'$ into Eq. (20a), the distance results: $D_{\hat{\mathbf{x}}} = \hat{\mathbf{x}}'\mathbf{R}_{\hat{\mathbf{x}}}^{-}\hat{\mathbf{x}} = \hat{\mathbf{x}}'\mathbf{P}\mathbf{\Lambda}^{-1}\mathbf{P}'\hat{\mathbf{x}} = T_t^2$. Furthermore, by combining Eqs. (12, 13, 14, 17), $T_u^2 = \mathbf{t}'\mathbf{B}\mathbf{B}^{-1}\mathbf{\Lambda}^{-1}\mathbf{B}^{-1}\mathbf{B}\mathbf{t} = T_t^2$ is obtained. From all these equalities, Proposition 1 is proven. □

REFERENCES

Alcala C.F. and Qin S.J., 2011. Analysis and generalization of fault diagnosis methods for process monitoring. *J. Process Control* 21, 322-330.

AlGazzawi A. and Lennox B., 2009. Model predictive control monitoring using statistics. *J. Process Control* 19, 314-327.

Gang L.G., Qin S.J. and Zhou D., 2010. Geometric properties of partial least squares for process monitoring. *Automatica* 46, 204-210.

Geladi P. and Kowalski B., 1986. Partial least-squares regression: A tutorial. *Anal. Chim. Acta*, 185, 1-17.

Godoy J.L., Minari R.J., Vega J.R., Marchetti J.L., 2011. Multivariate Statistical Monitoring of an Industrial SBR Process. Soft-sensor for Production and Rubber Quality. *Chem. Intell. Lab. Syst.* DOI: 10.1016/j.chemolab.2011.04.008.

Laurí D., Rossiter J.A., Sanchis J., Martínez M., 2010. Data-driven latent-variable model-based predictive control for continuous processes. *J. Process Control* 20, 1207–1219.

Meyer D., 2000. *Matrix analysis and applied linear algebra*. SIAM, USA.

Singhal A. and Seborg D.E., 2002. Pattern Matching in Multivariate Time Series Databases Using a Moving-Window Approach. *Ind. Eng. Chem. Res.* 41, 3822-3838.

Task space control of grasped object in the case of multi-robots cooperating system

A. Khadraoui [(a)] , **C. Mahfoudi** [(b)] , **A. Zaatri** [(c)], **K. Djouani** [(d)]

[(a)] *U. Larbi Ben M'Hidi , Institute of Mechanics, Ain Beida, Algeria*
Email: omar.khadraoui@gmail.com
[(b)] *U. Larbi Ben M'Hidi , Institute of Mechanics, Ain Beida, Algeria*
Email: c-mahfoudi- dz@yahoo.fr
[(c)] *U. Mentouri, Institute of Mechanics, Constantine,Algeria*
Email: azaatri@yahoo.com
[(d)] *LISSI Lab/University Paris 12, France*
Email: djouani@univ-paris12.fr

Abstract - This paper deals with multi-robots grasping problem. We present an optimal force distribution strategy for holding and manipulating objects by multiples manipulators. The force distribution issue is formulated in terms of a nonlinear programming problem under equality and inequality constraints. In particular, the friction constrains are transformed from non linear inequalities into a combination of linear equalities and linear inequalities. As a result, the original nonlinear constrained programming problem is then transformed into a quadratic optimization problem. The dynamical model of multi-robot cooperation has been used for determining force control distribution through proportional derivative (PD) controller. Simulation has been performed and some results are presented and discussed.

Keywords : dynamic modeling and control , friction constraints , grasping , Multi-robots cooperation , optimal force distribution , quadratic programming.

1 Introduction

Cooperative systems are generally understood as several coordinated robots simultaneously performing a common given task such as changing the space position of an object, grasping an object, gripping, lifting, lowering, releasing, withdrawing. The purpose of controlling a cooperative system consists of controlling contact forces between the environment and the object under consideration. Several approaches in the literature have been proposed to address the robot coordination problem. A robust control method is developed in [1] for a planar dual-arm manipulator system. In this approach, Contact and friction constraints for grasp conditions are considered and an optimization algorithm is developed which minimizes the total energy E consumed by the actuators. In [2], a robust force-motion control strategies are presented for mobile manipulators and force control constraints are developed using the passivity of hybrid joint rather than force feedback control. A new approach for computing force-closure grasps of two-dimensional and three-dimensional objects was proposed in [3],[4],[5], where a new necessary and sufficient conditions for n-finger grasps to achieve force-closure property are developed. In [6] a detailed analysis of grasping of deformable objects by a three finger hand was carried out, and it has been proved that the internal forces required for grasping deformable objects vary with size of object and finger contact angle. Xydas and al [7] have studied soft finger contact mechanics using finite element analysis and experiments. In [8], Hirai and al proposed a robust control law for grasping and manipulation of deformable objects. They developed a

control law to grasp and manipulate a deformable object using a real time vision system. Cutkosky [9] showed that grasp stability is a function of the fingertip contact models and small changes in the grasp geometry. Mason and Salisbury [10] gave conditions for complete restraint of an object in terms of internal forces. Kerr and Roth [11] proposed a method to determine the optimal internal forces based on approximated frictional constraint at the object and fingertip. Meer and al [12],Patton and al [13] designed a system, which focuses on force control versus deformation control. In these systems the robot manipulator is designed to control the deformation of the object. In [14] Howard and Bekey developed a generalized learning algorithm for handling of 3-D deformable objects in which prior knowledge of object attributes is not required. They used neural network with mass, spring and damping constant as input and the force needed to grasp the object as output of the network.

The object force control needs of real-time calculation of force distribution on the robot's effectors. Due to the existence of more than three actuated joints in each robot, the manipulator system has redundant actuation leading to more active joints than the object degree-of-freedom (6 dof). Thus, when formulating the force distribution problem, we find fewer force moment balancing equations than unknown variables. So, the solution of these equations is not unique. Moreover, some physical constraints, that concern the contact nature, friction, ...etc , must be taken into account in the calculation of force distribution. In addition, joints torque saturation must also be considered. Thus the Force Distribution Problem (FDP) can be formulated as a nonlinear constrained programming problem under nonlinear equality and inequality constraints. In this context, several approaches have been proposed for solving such a problem,using mainly the following four principal methods :

1. Linear-Programming (LP) Method [15],[16]

2. Compact-Dual LP (CDLP) Method [17],[18]

3. Quadratic Programming (QP) Method [19], [20]

4. Analytical Method [21],[22],[23]

A comparative study for the four cited methods can be found in [24]. Some researchers proposed the optimal force distribution scheme of multiple cooperating robots by combining the Dual with the QP Method [25].

In this paper we propose an approach to solve the problem of real-time force distribution for multiple manipulators system grasping an object based on the approach proposed in [26],[27] for an hexapod robot. This approach consists of the combination of the QP method with the reduction technique of problem size. The main idea concerns the transformation of the original nonlinear constrained problem into a linear one, by reducing the problem size and transforming the nonlinear constraints into a linear ones, respecting some physical considerations. The rest of the paper is organized as follows: The direct and inverse geometrical models of the robot manipulator are presented in section 2. Section 3 concerns the force distribution problem. Problem reduction and optimal solution are presented in section 4. Before presenting some remarks and perspectives, a Matlab simulation results of two cooperating manipulators is presented in section 5 to show the efficiency of this approach.

2 Geometrical Modeling

Before presenting the direct and inverse geometrical model, let us consider the robot architecture. Since the robots are similar , only one robot modeling is considered. The dual-arm manipulator system holding a common object and its architecture is shown in figure(1) and figure(2). Every "j" robot j=1,...,n is grasping the object, which is located at a_j distance from the center of gravity of the object. The angle ϕ_j represents the orientation of the coordinate frame $(x_{1,j}, y_{1,j}, z_{1,j})$ fixed at the base of the robot and the world ground coordinate frame (X, Y, Z). Multiple manipulators system is considered as an arborescent robot comporting some closed loops. So to study this kind of robots we use the method defined by Khalil and Kleinfinger [28].

The transformation matrix from ith joint's attached coordinate frame to the (i-1)th joint's attached coordinate frame is given by figure (3):

$$^{i-1}\boldsymbol{T}_i = R(Z,\gamma)T(Z,b)R(X,\alpha)T(X,d)R(Z,\theta)T(Z,r) \tag{1}$$

The table (1) describes the transformation from the world ground coordinate frame (X, Y, Z) to the coordinate frame at the contact point "6" of each robot"j", j=1,...,n

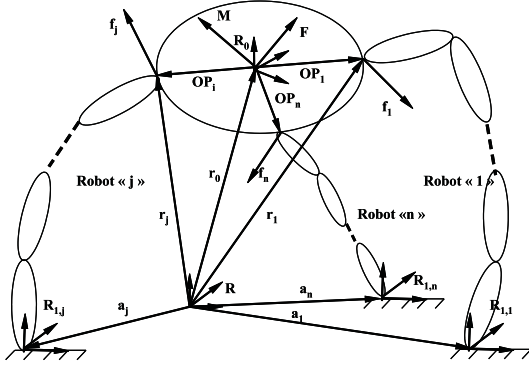The transformation provides the exact position of the contact point"6" of each robot in the absolute

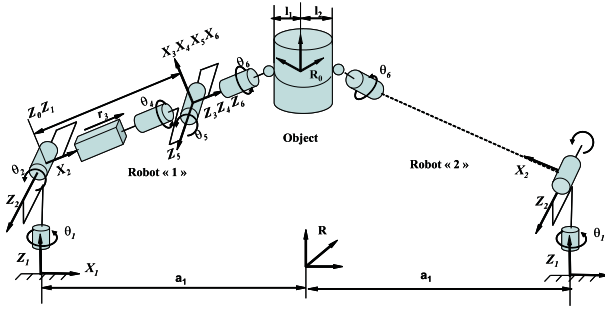Figure 1: **Geometrical parameters of multiple manipulators system**
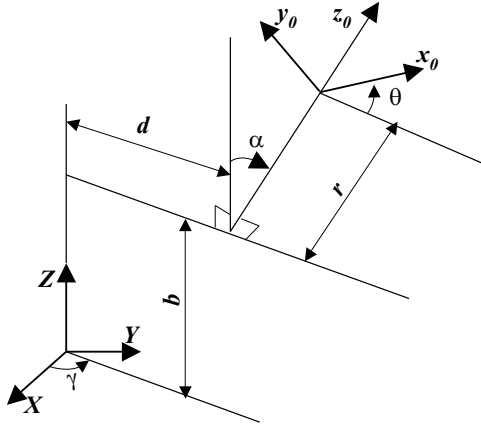

Figure 2: **Exemple of multiple manipulator system**

Figure 3: **Geometrical model**

coordinate frame fixed at the ground which is given by :

$$^{R}\boldsymbol{T}_6 =^{R} \boldsymbol{T}_0^0 \boldsymbol{T}_1^1 \boldsymbol{T}_2^2 \boldsymbol{T}_3^3 \boldsymbol{T}_4^4 \boldsymbol{T}_5^5 \boldsymbol{T}_6 \qquad (2)$$

When the position and the orientation of the last coordinate frame fixed to the end of each robot"j"

frame	α	d	θ	r	b	γ
object	α	d	θ	r	h	β
robot(1)	0	OP_1	$\theta_{1,1}$	0	0	0
robot"j"	0	OP_j	$\theta_{1,j}$	0	0	0
robot"n"	0	OP_n	$\theta_{1,n}$	0	0	0

Table 1: **Geometrical parameters of the system**

frame	α	d	θ	r	b	γ
joint"1"	0	a_j	θ_1	0	0	ϕ_j
joint"2"	0	0	$\theta_2 + \pi/2$	0	0	0
joint"3"	$\pi/2$	0	0	r_3	0	0
joint"4"	0	0	θ_4	r_4	0	0
joint"5"	$-\pi/2$	0	θ_5	0	0	0
joint"6"	$\pi/2$	0	θ_6	0	0	0

Table 2: **Geometrical parameters of the jth robot**

U_0 are known

$$U_0 = \begin{bmatrix} sx & nx & ax & Px \\ sy & ny & ay & Py \\ sz & nz & az & Pz \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad (3)$$

The values of the joints coordinates $\theta_{i,j}$ where : $\theta_{i,j}\,(i=1,2,3,4,5,6)\,(j=1,..,n)$ are given as follow:

$$\begin{cases} \theta_{1,j} = \arctan(Py/(-Px+d_1)) \\ \theta_{2,j} = \arctan(Pz-r_1)/ \\ \quad (C1(Px-d_1)+S1Py)) \\ r_{3,j} = C2(C1(Px-d_1)+S1Py)+ \\ \quad S2(Pz-r_1)-R4-l_2 \\ \theta_{4,j} = \arctan((S1ax-C1ay)/ \\ \quad (S2(C1ax+S1ay)-C2az)) \\ \theta_{5,j} = \arctan((C4S2(C1ax-S1ay)- \\ \quad C4C2az+S4(C1ay-C1sy))/ \\ \quad (C2(C1ax+S1ay)+S2az)) \\ \theta_{6,j} = \arctan((-S4(-S2(C1ax-S1ay)+ \\ \quad C2sz+C4S1sx-C1sy))/ \\ \quad (C5(-C4S2C1sx-S2S1sy+C2sz \\ \quad +S4(S1sx-C1sy)- \\ \quad S5(C1C2sx+S1C2sy+S2sz)) \end{cases} \qquad (4)$$

Remark : S*=sin(*); C*=cos(*); $(Px=P_{x,j}, Py=P_{y,j}, Pz=P_{z,j})$, are the coordinates of the point "6" of the jth robot expressed in (X,Y,Z). by using the Roulis, Tangage and Lacet (RTL) angles ,the desired

position can be found as follow:

$$
\left[\begin{array}{ccc} sx & nx & ax \\ sy & ny & ay \\ sz & nz & az \end{array} \right] = \left[\begin{array}{cc} C\phi C\theta & C\phi S\theta S\psi - S\phi C\psi \\ S\phi C\theta & S\phi S\theta S\psi + C\phi C\psi \\ -S\theta & C\theta S\psi \end{array} \right.
$$

$$
\begin{array}{l} C\phi S\theta C\psi + S\phi C\psi \\ S\phi S\theta C\psi - C\phi S\psi \\ C\theta C\psi \end{array} \ \left. \phantom{\begin{array}{l}a\\a\\a\end{array}}\right] .(5)
$$

The Jacobian matrix developed above is sometimes called the geometric Jacobian to distinguish it from the analytic Jacobian, denoted by $Ja(q)$, which is based on a minimum representation for the orientation of the effector's coordinate frame $X = [x_p \, x_r]^T$. A representation of the situation of coordinate frame R_n in R_0, where x_p represents the three coordinates operational position and x_r is the coordinate operational orientation, the operational velocities are $\dot{X} = [\dot{x}_p \, \dot{x}_r]^T$ however we must find the relationship between these velocities and velocity vectors $^0V_n \, ^0\omega_n$ as follow :

$$
\left[\begin{array}{c} \dot{x}_p \\ \dot{x}_r \end{array} \right] = \left[\begin{array}{cc} \Omega_p & 0_3 \\ 0_3 & \Omega_p \end{array} \right] \left[\begin{array}{c} ^0V_n \\ ^0\omega_n \end{array} \right] = \Omega \left[\begin{array}{c} ^0V_n \\ ^0\omega_n \end{array} \right] \quad (6)
$$

$$
\left[\begin{array}{c} \dot{x}_p \\ \dot{x}_r \end{array} \right] = \Omega^0 J_n \dot{q} = J_a \dot{q} \quad (7)
$$

In general, the sub matrix Ω_p is equal to the unitary matrix I_3 because the operational coordinates of position are simply the Cartesian coordinates of the position of the tool. In contrast, the matrix Ω_r depends on the choice of the operational coordinates of rotation, as it is defined, the orientation vector we chose is $x_r = [\phi \, \theta \, \psi]^T$: representing the rotation called : roll, pith, and yaw (PRY). in this case [KHA 99]:

$$
\Omega_r = \left[\begin{array}{ccc} C\phi tg\theta & S\phi tg\theta & 1 \\ -S\phi & C\phi & 0 \\ C\phi/C\theta & C\phi/C\theta & 0 \end{array} \right]
$$

we can show that,we have singularity when $\theta = +-\pi/2$.

3 Force Distribution Problem

3.1 Problem Formulation

The force system acting on the object is shown in figure (4). For simplicity, only the force components on the contact point are presented.
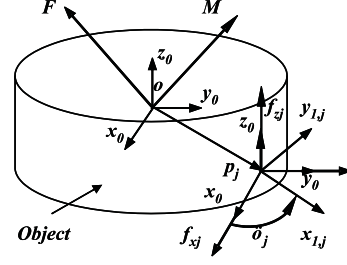

Figure 4: Orientation of coordinate frame

In the general case, rotational torques at the contact are neglected. Let (x_0, y_0, z_0) be the coordinate frame in which the object is located and $(x_{6,j}, y_{6,j}, z_{6,j})$ denote the coordinate frame fixed at the contact point of the jth robot. The $(x_{6,j}, y_{6,j})$ plane which is assumed to be parallel to the (x_0, y_0) plane and its z axis is normal to the surface of the object. $\boldsymbol{F} = [F_X F_Y F_Z]^T$ and $\boldsymbol{M} = [M_X M_Y M_Z]^T$ denote respectively the object force vector and moment vector, which results from the gravity and the external force acting on the object [29],[30],[31] and [32]. Define $f_{x,j}, f_{y,j}$, and $f_{z,j}$ as the components of the force acting on the supporting robot "j" in the directions of x_0, $y_0 \, and \, z_0$, respectively. The number of cooperative robots, can vary between 2 and 3 for this studies. The object's quasi-static force/moment equation can be written as

$$
\left\{ \begin{array}{l} \sum_{j=1}^n \boldsymbol{f}_j = \boldsymbol{F} \\ \sum_{j=1}^n \boldsymbol{OP}_j \wedge \boldsymbol{f}_j = \boldsymbol{M} \end{array} \right. \quad (8)
$$

where $\boldsymbol{OP}_j$ is the position vector joining contact point of the robot "j" and the gravity center of the object. The general matrix form of this equation can be written as:

$$
\boldsymbol{A} \, \boldsymbol{G} = \boldsymbol{W} \quad (9)
$$

with:

$$
\left\{ \begin{array}{ll} \boldsymbol{G} = [\boldsymbol{f}_1^T \boldsymbol{f}_2^T \cdots \boldsymbol{f}_n^T]^T & \in \Re^{3n} \\ \boldsymbol{f}_j^T = [f_{x,j} \, f_{y,j} \, f_{z,j}]^T & \in \Re^3 \\ \boldsymbol{W} = [\boldsymbol{F}^T \boldsymbol{M}^T]^T & \in \Re^6 \end{array} \right.
$$

$$
\boldsymbol{A} = \left(\begin{array}{cccc} \boldsymbol{I}_3 & \cdots & \cdots & \boldsymbol{I}_3 \\ \boldsymbol{B}_1 & \cdots & \cdots & \boldsymbol{B}_n \end{array} \right) \quad \in \Re^{6 \times 3n}
$$

$$
\boldsymbol{B}_j \equiv \widehat{\boldsymbol{OP}}_j \equiv \left(\begin{array}{ccc} 0 & -P_{z,j} & P_{y,j} \\ P_{z,j} & 0 & -P_{x,j} \\ -P_{y,j} & P_{x,j} & 0 \end{array} \right) \quad \in \Re^{3 \times 3}
$$

where $\boldsymbol{I}_3$ is the identity matrix and $\mathbf{G}$ is the robots force vector, corresponding to three $(\boldsymbol{G} \quad \in \Re^9)$.$\boldsymbol{A}$ is a coefficient matrix which is a function of the

positions of the cooperative robots, and $\boldsymbol{B}_j$ is a skew symmetric matrix consisting of $(P_{x,j}, P_{y,j}, P_{z,j})$, which is the position coordinate of contact point of the multi-robot systems robot "j" in (x_0, y_0, z_0). $\boldsymbol{W}$ is a total body force/moment vector. It is clear that (9) is an undetermined system and its solution is not unique. In other words, the robots forces have many solutions according to the equilibrium equation. however, the robot forces must meet the needs for the following physical constraints, otherwise they become invalid :

1. Supported object should not slip when the robots move. It results in the following constraint:

$$\sqrt{f_{x,j}^2 + f_{y,j}^2} \leq \mu\, f_{z,j} \qquad (10)$$

where μ is the static coefficient of friction of the surface of the object

2. Since the robots forces are generated from the corresponding actuators of joints, the physical limits of the joint torques must be taken into account. It follows that :

$$-\tau_{jmax} \leq \boldsymbol{J}_j^{T\,j}\boldsymbol{A}_{0j} \begin{pmatrix} f_{x,j} \\ f_{y,j} \\ f_{z,j} \end{pmatrix} \leq \tau_{jmax} \qquad (11)$$

for $(j = 1, ..., n)$, where $\boldsymbol{J}_j \in \Re^{3\times 6}$ is the Jacobian of the robot "j", $\tau_{jmax} \in \Re^{6\times 1}$ is the maximum joint torque vector of the robot "j", and $\boldsymbol{A}_{0j} \in \Re^{3\times 3}$ is the orientation matrix of $(x_{6,j}, y_{6,j}, z_{6,j})$ with respect to (x_0, y_0, z_0).

3. In order to have definite contact with the object, there must exist a $f_{z,j}$ such that :

$$f_{z,j} \geq 0 \qquad (12)$$

In the following, we propose an approach for problem size reduction, linearization and solving for the three manipulators case. Clearly, it is difficult to solve such a nonlinear programming problem for real-time multi-robots force distribution with complex constraints.

3.2 Problem Size Reduction

The equation (10) is a formulation of the friction cone figure(5). In order to overcome the non lin-
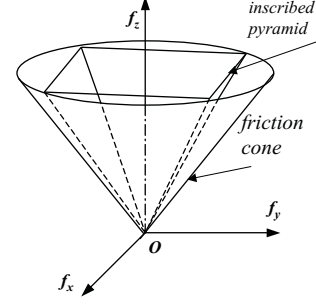

Figure 5: conservative inscribed pyramid

earities induced by the following equations, most researches substitute this friction cone by the inscribed pyramid [16],[33],[19],[34]. Thus, the nonlinear friction constraints are approximately expressed by the following linear inequalities :

$$f_{x,j} \geq \acute{\mu}\, f_{z,j}, \quad f_{y,j} \geq \acute{\mu}\, f_{z,j}, \quad j = 1, ..., n \qquad (13)$$

where $\acute{\mu} = \frac{\sqrt{2}\mu}{2}$ is for the inscribed pyramid.

Thus, the initial non linear constrained programming problem, substituting the non linear constraint Eq(10) by the linear one of Eq(13), becomes a linear programming problem [16],[17], and [19]. The possibility of slipping can be minimized, by optimizing the ratio of tangential to normal forces at the robot. This leads Liu and Wen [22] to find the relationship between the robot forces and transform the initial friction constraints from the nonlinear inequalities into a set of linear equalities. Let us define the global ratio by the ratio of the tangential to normal forces at the object. The advantage of the existing methods lies in the fact that part of component of the robots forces satisfy the global ratio relationship and lets the other components satisfy the linear inequality constraints as Eq (13). For example, defining $f_{x,j}$ $(j = 1, ..., n)$ and $f_{y,j}$ $(j = 1, ..., n)$, for a robot j. Chen et al [1], show that :

$$f_{x,j} = k_{xz} f_{z,j} \quad , (i = 1, ..., n) \qquad (14)$$

$$f_{y,j} \leq \mu^{\star} f_{z,j} \quad , (i = 1, ..., n) \qquad (15)$$

where $k_{xz} = \frac{F_X}{F_Z}$ is the global ratio of forces at the object in direction of x_0 and z_0. $\mu^{\star}$ is the given coefficient for friction constraints. According to Eq(10), we have $\mu^{\star} = \sqrt{\mu^2 - k_{xz}^2}$. Finally, the force distribution problem is transformed into a linear one by replacing Eq (6) with Eqs (10) and (11).

3.3 Problem transformation and Continuous solution

In modelling these systems, we consider that three robots grasp the object at a time. so G and A become a vector of 9×1 and a matrix of 6×9, respectively. Expression (9) contains nine unknown variables with six equations. By adding the Eq (14) to the Eq (9) we obtain nine equations.

$$\overline{A}\, G = \overline{W} \qquad (16)$$

with :

$$\overline{A} =$$

$$\begin{pmatrix} I_3 & & I_3 & & I_3 \\ B_1 & & B_2 & & B_3 \\ 1 & 0 & -k_{xz} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & -k_{xz} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -k_{xz} \end{pmatrix}$$

$$G = \begin{pmatrix} f_1 \\ f_2 \\ f_3 \end{pmatrix} \qquad \overline{W} = \begin{pmatrix} F \\ M \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Using some rows combination of the matrix $\overline{A}$, Eq (16) can be written as :

$$\widehat{A}\, G = \widehat{W} \qquad (17)$$

With: $\widehat{A} =$

$$\begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & -k_{xz} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & -k_{xz} \\ -P_{y,1} & P_{x,1} & 0 & -P_{y,2} & P_{x,2} & 0 \\ P_{z,1} & 0 & -P_{x,1} & P_{z,2} & 0 & -P_{x,2} \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -P_{z,1} & P_{y,1} & 0 & -P_{z,2} & P_{y,2} \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & -k_{xz} \\ -P_{y,3} & P_{x,3} & 0 \\ P_{z,3} & 0 & -P_{x,3} \\ 0 & -P_{z,3} & P_{y,3} \end{pmatrix} \qquad \widehat{W} = \begin{pmatrix} F_X \\ F_Y \\ 0 \\ 0 \\ M_Z \\ M_Y \\ 0 \\ M_X \end{pmatrix}$$

where $\widehat{A} \in \Re^{8 \times 9}$ is the resulting matrix of $\overline{A}$ after combination. $G \in \Re^8$ is the robots force vector. $\widehat{W} \in \Re^8$ is the resulting vector of W after combination. Thus, the force distribution problem is subjected to the inequality constraints expressed by (11), (12) (14).

4 Quadratic Problem Formulation and Solution

The solution to the inverse dynamic equations of this system is not unique, but it can be chosen in an optimal manner by minimizing some objective functions. The approach taken here consists of minimizing the sum of the weighted torque of the robot, which results in the following objective function [20],[25],[35] :

$$f_G = p^T G + \frac{G^T Q G}{2} \qquad (18)$$

with:

$$p^T = [\widehat{\tau}_1^T J_1^T \dots \dots, \widehat{\tau}_n^T J_n^T] \quad \in \Re^{3n}$$

$$Q = \begin{pmatrix} J_1 q_1 J_1^T & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & J_n q_n J_n^T \end{pmatrix} \quad \in \Re^{3n \times 3n}$$

where $\widehat{\tau}_j$ is the joint torque vector due to the weight and inertia of the robot "j", J_j is the Jacobian of the robot "j", and q_j is a positive definite diagonal weighting matrix of the robot"j". This objective function is strictly convex. Because the time for obtaining a solution does not depend on an initial guess, a quadratic programming is superior to linear programming in both speed and quality of the obtained solution [20]. The general linear-quadratic programming problem of the force distribution on robot is stated by :

$$p^T G + \frac{G^T Q G}{2} \qquad (19)$$

$$\widehat{A}\, \widehat{G} = \widehat{W} \qquad (20)$$

$$B\widehat{G} \le C \qquad (21)$$

where $G \in \Re^9$ is a vector of the design variables. It should be pointed out that the expression (21) is the resulting inequality constraints for the combination of equations (11),(12) and (15) where :

$$B = [B_1^T \, B_2^T \, B_3^T \, B_4^T]^T \quad \in \Re^{9 \times 24}$$
$$C = [\tau_{1max} \, .. \, \tau_{6max} \, - \, \tau_{1max} \, .. \, -$$
$$\tau_{6max} \, \boldsymbol{0} \, \boldsymbol{0} \, \boldsymbol{0} \, \boldsymbol{0} \, \boldsymbol{0}]^T \quad \in \Re^{24}$$

with

$$B_1 = \begin{bmatrix} \boldsymbol{J}_1^T \boldsymbol{R}_1 & 0 & 0 \\ 0 & \boldsymbol{J}_2^T \boldsymbol{R}_2 & 0 \\ 0 & 0 & \boldsymbol{J}_3^T \boldsymbol{R}_3 \end{bmatrix} \quad \in \Re^{9 \times 9}$$

$$\boldsymbol{B}_2 = \begin{bmatrix} -\boldsymbol{J}_1^T \boldsymbol{R}_1 & 0 & 0 \\ 0 & -\boldsymbol{J}_2^T \boldsymbol{R}_2 & 0 \\ 0 & 0 & -\boldsymbol{J}_3^T \boldsymbol{R}_3 \end{bmatrix}$$

$$B_3 = \begin{bmatrix} 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \end{bmatrix}$$

$$B_4 = \begin{bmatrix} 0 & 1 & -\mu^\star & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -\mu^\star & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -\mu^\star \end{bmatrix}$$

In Eq (20), we have eight linear independent equations with nine unknown variables. By using Gauss algorithm, this equation is transformed as follow :

$$[\boldsymbol{I}_8 \quad \widehat{\boldsymbol{A}}_r] \begin{bmatrix} \widehat{\boldsymbol{G}}_b \\ \widehat{\boldsymbol{G}}_r \end{bmatrix} = \widehat{\boldsymbol{W}}_r \qquad (22)$$

where $\boldsymbol{I}_8 \in \Re^{8 \times 8}$ identity matrix, $\widehat{\boldsymbol{A}}_r \in \Re^8$ is the remaining column of the matrix $\widehat{\boldsymbol{A}}$ after transformation. $\widehat{\boldsymbol{G}}_b \in \Re^8$ is the partial vector of $\boldsymbol{G}$. $\widehat{\boldsymbol{G}}_r \in \Re$ is the unknown element of $\widehat{\boldsymbol{G}}$ which denotes the design variable. $\widehat{\boldsymbol{W}}_r \in \Re^8$ is the resulting vector of $\widehat{\boldsymbol{W}}$ after transformation. Equation (22) may be rewritten by the following form

$$\boldsymbol{I}_8 \widehat{\boldsymbol{G}}_b + \widehat{\boldsymbol{A}}_r \widehat{\boldsymbol{G}}_r - \widehat{\boldsymbol{W}}_r = 0, \qquad (23)$$

Which yields to

$$\widehat{\boldsymbol{G}}_b = \widehat{\boldsymbol{W}}_r - \widehat{\boldsymbol{A}}_r \widehat{\boldsymbol{G}}_r. \qquad (24)$$

Finally, it results in

$$\boldsymbol{G} = \begin{bmatrix} \widehat{\boldsymbol{G}}_b \\ \widehat{\boldsymbol{G}}_r \end{bmatrix} = \begin{bmatrix} \widehat{\boldsymbol{W}}_r \\ 0 \end{bmatrix} + \begin{bmatrix} -\widehat{\boldsymbol{A}}_r \\ 1 \end{bmatrix} \widehat{\boldsymbol{G}}_r. \qquad (25)$$

Now let $\widehat{\boldsymbol{G}}_0 = [\widehat{\boldsymbol{W}}_r^T \ 0]^T \in \Re^8$ and $N = [-\widehat{\boldsymbol{A}}_r^T \ 1]^T \in \Re^9$, then Eq (25) becomes

$$\boldsymbol{G} = \widehat{\boldsymbol{G}}_0 + \boldsymbol{N}\widehat{\boldsymbol{G}}_r. \qquad (26)$$

Substituting Eq (26) into Eqs (19) and (21), the linear quadratic programming problem can be expressed by :

$$minimize \quad f(\widehat{\boldsymbol{G}}_r), \qquad (27)$$

$$subject \, to \quad \boldsymbol{B}N\widehat{\boldsymbol{G}}_r \leq \boldsymbol{C} - \boldsymbol{B}\widehat{\boldsymbol{G}}_0. \qquad (28)$$

where

$$f(\widehat{\boldsymbol{G}}_r) = \boldsymbol{p}^T \widehat{\boldsymbol{G}}_0 + \tfrac{1}{2} \widehat{\boldsymbol{G}}_0^T Q \widehat{\boldsymbol{G}}_0 + \boldsymbol{p}^T \boldsymbol{N}\widehat{\boldsymbol{G}}_r$$
$$+ \tfrac{1}{2} \widehat{\boldsymbol{G}}_0^T \boldsymbol{Q}\boldsymbol{N}\widehat{\boldsymbol{G}}_r + \tfrac{1}{2} \widehat{\boldsymbol{G}}_r^T \boldsymbol{N}^T \boldsymbol{Q}\widehat{\boldsymbol{G}}_0$$
$$+ \tfrac{1}{2} \widehat{\boldsymbol{G}}_r^T \boldsymbol{N}^T \boldsymbol{Q}\boldsymbol{N}\widehat{\boldsymbol{G}}_r$$

Since $\widehat{\boldsymbol{G}}_r$ is a single variable denoted by x, the optimal force distribution can be written as :

$$minimize \quad a_0 x^2 + a_1 x + a_2 \qquad subject \, to \quad x \in [b_1 \, b_2] \qquad (29)$$

With

$$a_0 = \tfrac{1}{2} \boldsymbol{N}^T \boldsymbol{Q}\boldsymbol{N}$$
$$a_1 = \boldsymbol{p}^T \boldsymbol{N} + \tfrac{1}{2} \widehat{\boldsymbol{G}}_0^T \boldsymbol{Q}\boldsymbol{N} + \tfrac{1}{2} \boldsymbol{N}^T \boldsymbol{Q}\widehat{\boldsymbol{G}}_0$$
$$a_2 = \boldsymbol{p}^T \widehat{\boldsymbol{G}}_0 + \tfrac{1}{2} \widehat{\boldsymbol{G}}_0^T \boldsymbol{Q}\widehat{\boldsymbol{G}}_0$$

Where $[b_1 \, b_2]$ denotes the bound resulted from Eq(28). Since it is clear that $a_0 \geq 0$ because of the positive-definite matrix Q, There must be an optimal solution for the force distribution problem.

5 Robot dynamic modeling

For each robot, the serial robot dynamic model [28], [36], is given by the following equations :

$$\Gamma = f(\boldsymbol{\theta}, \dot{\boldsymbol{\theta}}, \ddot{\boldsymbol{\theta}}) = \boldsymbol{M}(\boldsymbol{\theta})\ddot{\boldsymbol{\theta}} + \boldsymbol{C}(\boldsymbol{\theta}, \dot{\boldsymbol{\theta}})\dot{\boldsymbol{\theta}} + \boldsymbol{Q}(\boldsymbol{\theta}) + \boldsymbol{J}^T \boldsymbol{f} \qquad (30)$$

$\boldsymbol{C} \in \Re^{6 \times 1}, \boldsymbol{Q} \in \Re^{6 \times 1}$ and $\boldsymbol{J}^T \boldsymbol{f} \in \Re^{6 \times 1}$. Where,

- $\boldsymbol{\theta}$, $\dot{\boldsymbol{\theta}}$ and $\ddot{\boldsymbol{\theta}}$ are, respectively, the generalized coordinate, speed and acceleration vectors,

- $\boldsymbol{M}(\boldsymbol{\theta})$, $n \times n$ matrix representing the inertia of the robot,

- $\boldsymbol{C}(\boldsymbol{\theta}, \dot{\boldsymbol{\theta}})$, $n \times 1$ vector representing coriolis and centrifuge terms,

- $\boldsymbol{Q}(\boldsymbol{\theta})$, $n \times 1$ vector of gravity terms

- $\boldsymbol{J}^T$, jacobian matrix of the robot

- $\boldsymbol{f}$, force acting on the supporting robot "j"

Equation 30 can be written as follows :

$$\Gamma = \boldsymbol{M}(\boldsymbol{\theta})\ddot{\boldsymbol{\theta}} + \boldsymbol{H}(\boldsymbol{\theta},\dot{\boldsymbol{\theta}}) + \boldsymbol{J}^T \boldsymbol{f} \qquad (30)$$

with :

$$\boldsymbol{H}(\boldsymbol{\theta},\dot{\boldsymbol{\theta}}) = \boldsymbol{C}(\boldsymbol{\theta},\dot{\boldsymbol{\theta}})\dot{\boldsymbol{\theta}} + \boldsymbol{Q}(\boldsymbol{\theta})$$

By setting $\ddot{\boldsymbol{\theta}}$ and $\boldsymbol{f}$ to $\boldsymbol{0}$, we can compute $\boldsymbol{H}(\boldsymbol{\theta},\dot{\boldsymbol{\theta}})$ using Newton-Euler algorithm.

5.1 Object Dynamic

Consider an object of mass m_0 being carried by the two manipulators, so when the load is grasped by the end-effectors of robots , the force exerted from the tips of the manipulators act to translate or rotate the object in any direction, then the dynamic model of the object is given by :

$$\begin{pmatrix} m_0\boldsymbol{I}_3 & \boldsymbol{0} \\ \boldsymbol{0} & \underline{\boldsymbol{\varphi}}^0 \end{pmatrix} \begin{pmatrix} \boldsymbol{\gamma}_{0,0} \\ \dot{\boldsymbol{\omega}}_{0,0} \end{pmatrix} + \begin{pmatrix} -m_0\boldsymbol{g}_0 \\ \boldsymbol{\omega}_{0,0} \wedge (\underline{\boldsymbol{\varphi}}^0\boldsymbol{\omega}_{0,0}) \end{pmatrix} =$$

$$\begin{pmatrix} \boldsymbol{F} \\ \boldsymbol{M} \end{pmatrix} (31)$$

$\boldsymbol{F}$ and $\boldsymbol{M}$ are given by the following equations.

$$\begin{cases} \boldsymbol{F} = \sum_{j=1}^{6} \boldsymbol{F}_{1,j}^0 \\ \boldsymbol{M} = \sum_{j=1}^{6} (\boldsymbol{C}_{1,j}^0 + \boldsymbol{P}_{01,j}^0 \wedge \boldsymbol{F}_{1,j}^0) \end{cases} \qquad (32)$$

where :

- $\boldsymbol{\gamma}_{0,0}$ and $\boldsymbol{\omega}_{0,0}$ be, respectively, the object linear and angular accelerations in the coordinate frame (x_0, y_0, z_0).

- $\boldsymbol{F}_{1,j}^0$, the force applied by the robot "j" at the articulation "6" on the object "0".

- $\boldsymbol{C}_{1,j}^0$, the moment applied by the jth robot in the articulation "6" on the object.

- $\boldsymbol{P}_{01,j}^0$, the distance between the articulation "6" of the jth robot and the origin of the coordinate frame (x_0, y_0, z_0) expressed in the same coordinate frame.

- m_0 and $\underline{\boldsymbol{\varphi}}^0$ are, respectively, the mass and the inertia matrix of the grasped object.

- $\boldsymbol{I}_3$, the identity matrix (3×3).

- $\boldsymbol{g}_0$: gravity vector

6 Systems Dynamic Joint Control

Let us define $X_d(t)$, the desired trajectory for the arm motion. To ensure trajectory tracking by the joint variable errors and the optimal force distribution algorithm, computed torque control is used for robots joint level control (figures 6 and 7).

with;

$$X_0^d = [P_{x0}, P_{y0}, P_{z0}, \theta_0, \phi_0, \psi_0]$$

$$X_1^d = [P_{x1}, P_{y1}, P_{z1}, \theta_1, \phi_1, \psi_1]$$

$$X_2^d = [P_{x2}, P_{y2}, P_{z2}, \theta_2, \phi_2, \psi_2]$$

where:

X_0^d, X_1^d and X_2^d denote respectively the desired position and orientation of object, robot1 and robot2

$$\boldsymbol{e}(t) = X_d - X(t) \qquad (33)$$

Then, the overall robot arm input becomes:

$$\Gamma = \boldsymbol{A}\boldsymbol{J}^{-1}(\ddot{\boldsymbol{X}} - \dot{\boldsymbol{J}}\dot{\theta}) + \boldsymbol{H} \qquad (34)$$

$$\boldsymbol{X}(t) = \ddot{\boldsymbol{X}} + \boldsymbol{k}_v(\dot{\boldsymbol{X}}_d - \dot{\boldsymbol{X}}) + \boldsymbol{k}_p(\boldsymbol{X}_d - \boldsymbol{X}) \quad (35)$$

This controller is shown in figures (6) and (7)

6.1 Choice of PD Gains

It is usual to take the $n \times n$ matrices diagonal so that:

$$\boldsymbol{k}_v = diag[k_{vi}], \quad \boldsymbol{k}_p = diag[k_{pi}]$$

and $k_{pi} = \omega_n^2$, $k_{pi} = 2\xi\omega_n$ with ξ the damping ratio and ω_n the naturel frequency.
The PD gains are usually selected for critical damping ξ=1. Then, to avoid exciting the resonant mode, we should select natural frequency to half the resonant frequency $\omega_n < \omega_r/2$.
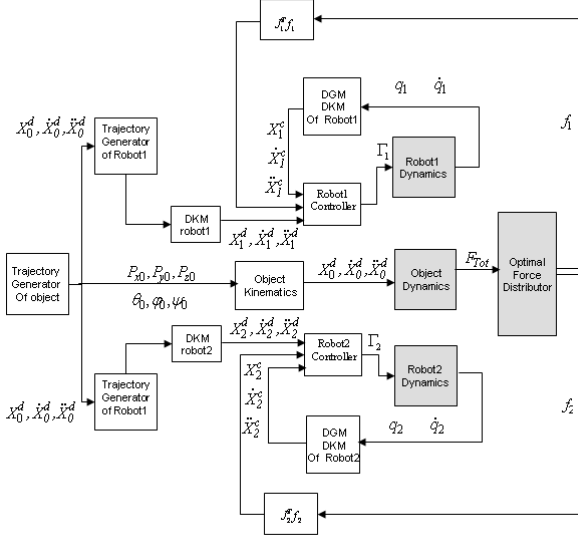
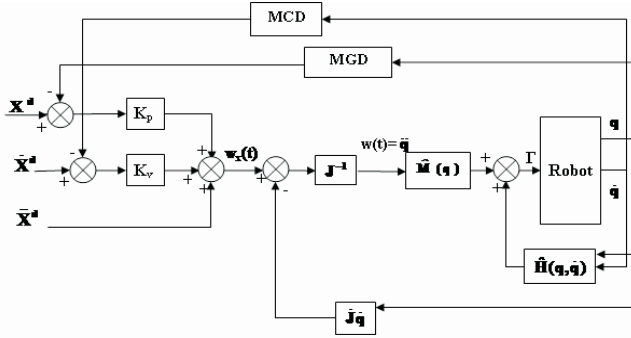Figure 6: **System overall control**

Figure 7: **Torque control of the robot** $"j"$

7 Simulations

In order to show the effectiveness of proposed approach, we consider two identical 6-DOF manipulators handling a rigid object in a task space. The base of each robot is located at a distance a_1 see figure(2), We denote $X_i \in \Re^6$ where: $X_i = [P_{xi}, P_{yi}, P_{zi}, \theta_i, \phi_i, \psi_i]^T$ The end-effector position of robot i is designed by: $[P_{xi}, P_{yi}, P_{zi}]^T$, while the orientation is represented by $[\theta_i, \phi_i, \psi_i]^T$. Furthermore the load vector coordinates (generalized forces) at the object is known. The values of masse $m(i)$ in Kg, the moment of inertia $Ix(i)$, $Iy(i)$,$Iz(i)$ in $Kg.m^2$ and the lenghts in meter of different links of each robot used in simulation , are given in table (3)

and $\mu = 0.05$ is the static coefficient of friction.

Link"i"	$Ix(i)$	Iy(i)	$Iz(i)$	m(i) Kg	l(i)
Link"1"	0.01	0.05	0.06	25	1
Link"2"	0.1	0.2	0.3	30	0.75
Link"3"	0.4	0.4	0.2	15	0.50
Link"4"	0.3	0.5	0.01	4	0.25
Link"5"	0.01	0.6	0.03	4	0.10
Link"6"	0.2	0.2	0.6	2	0.1

Table 3: **Physical parameters of the robot**

Some simulations were conducted with Matlab. We consider that the desired trajectory of the object in the operational space is given as follows : $x = 1 + 2.5.sin(0.3)t$,
$y = -3 + 2.t$, $z = r_{obj} = r_1 + sin(2.t)$, $\theta(t)_z = \pi/2 + 1.5.sin(t)$ is the rotation about the ox axis and $\alpha_x(t) = \pi/2 + 0.5.sin(0.3t)$ the rotation about the oy axis. $\tau_j = [10\,10\,10\,10\,10\,10]^T$ is the maximum joint torque vector and For the objective function Eq (18), the weighting matrix are choosen as follow : $\boldsymbol{p = 0}$ and $\boldsymbol{Q = I}$ (the identity matrix).
The figure 8 shows the trajectory of the object in the X-Y-Z supported by the two robots.
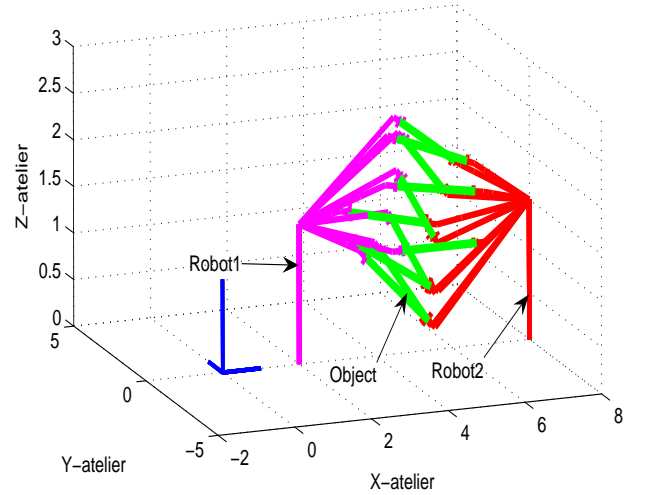
Figure 8: View of the object suported by tow robots

The associated joints coordinates are obtained by using the direct and inverse geometric model (Eq (2)and Eq (3)). The Figure (9.a) and (9.b) show respectively the profiles of the position P_x,

P_y, P_z and orientation θ, ϕ, ψ of the grasped object in the operational space.
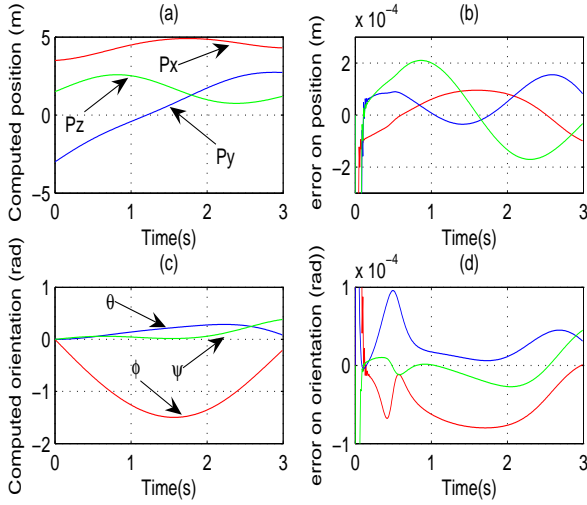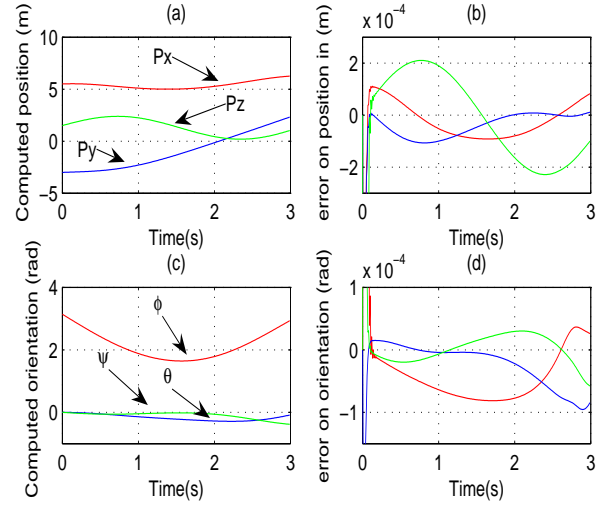

Figure 9: Computed position and orientation of robot1

Figure 10: Computed position and orientation of robot2
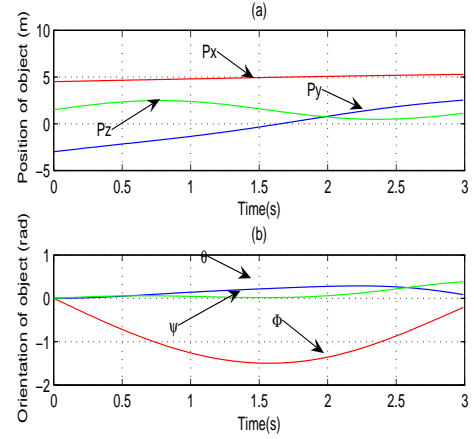
Figure 11: Position and orientation of the grasped object

However the position P_x, P_y, P_z and orientation θ, ϕ, ψ respectively of robot1 and robot2 are presented in Figures (10.a),(10.c) and (11.a),(11.c). The figures (10.b),(10.d), (11.b),(11.d),(12.c),(12.d) and (13.c),(13.d) demonstrates convergence of corresponding signals illustrated in figures cited above. In the figures (14.a) and (14.b) the force distribution acting respectively by robot(1) and robot(2) on the grasped object are given. We can show that, this distribution validate the following force equilibrium equation :

$$\Sigma f x_j = Fx, \quad ,\Sigma f y_j = Fy, \quad \Sigma f z_j = Fz$$

Elsewhere, the z-force components $f z_j$ are never negative, respecting the contact constraint . We can also show that the constraint Eq (10) are always satisfied as shown in figure (14c). and (14d) for the robot(j) ,j=1,2 where the curve $f x_j^2 + f y_j^2$ is always under the curve $\mu^2 f z_j^2$

8 Conclusion

In this paper, the force distribution problem has been presented in the case of multiple manipulators system grasping a same object. First, the robot inverse and direct geometric models are presented. Then, the real time force distribution problem are formulated in terms of non linear programming problem. After problem size reduction and transformation, the initial problem is solved in terms of quadratic programming problem. After then, we present the development for real-time operational space control of the system.Finally, Simulations results are presented in order to show the effectiveness of the
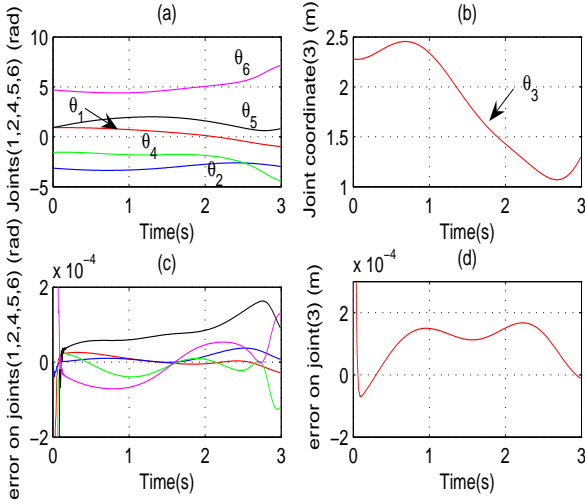
Figure 12: Joints coordinates of the robot2

Figure 14: Forces on the two robots

Figure 13: Joints coordinates of the robot1

proposed approach. Actually, we are working on the generalization of this approach for $n > 3$ robots case.

References

[1] Jin-Fan Liu Karim Abdel-Malek1. Robust control of planar dual-arm cooperative manipulators. *Robotics and Computer-Integrated Manufacturing.*, pages 109–120, (2000).

[2] Zhijun Li1*; Weidong Chen1 and Hong Liu2. Robust control of wheeled mobile manipulators using hybird joints. *International Journal of Advanced Robotic Systems*, 5:83–90, (2008).

[3] B.Bounab D.Sidobre and A.Zaatri. On computing multi-finger force closure grasps of 2d objects. *ICINCO Int. Conf. Informatics in Control, Automat. Robot.*, (2007).

[4] B.Bounab D.Sidobre and A.Zaatri. Central axis approach for computing n-finger force-closure grasp. *in Proc. IEEE Int. Conf. Robot. Automat.*, pages 1169–1174, (2008).

[5] N.Saito T.Satoh and H.Okano. Grasping force control in consideration of translational and rotational slippage of an object by a flexible contact sensor. *Electrical Engineering in Japan*, 165:51–57, 2008.

[6] Sohil Garg and Ashish Dutta. Grasping and manipulation of deformable objects based on internal force requirements. *International Journal of Advanced Robotic Systems*, 3:107–114, (2006).

[7] Xydas. Study of soft finger contact mechanics using finite element analysis and experiments. *Proc. of the IEEE Intl. cong on robotics and Automation*, 3:2179 –2184, 2000.

[8] Hirai.S Tsuboi.T and Wada .T. Robust grasping manipulation of deformable objects. *IEEE Symposium on Assembly and*

Task Planning.Japan, pages 2179 –2184, 2001.

[9] M.R. Cutkosky. On grasp choice, grasp models, and the design of hands for manufacturing tasks. *Trans Robot. Automat*, 5:269–279, 1989.

[10] Mason and J.K. Salisbury. Robot hands and the mechanics of manipulation. *Cambridge, MA: MIT Press*, 5, 1985.

[11] J. Kerr and B. Roth. Analysis of multifingered hands. *International Journal of Robotics Research*, 4:3–17.

[12] D.W. Meer and S.M. Rock. Experiments in object impedance control for flexible objects. *Proceedings International Conference on Robotics and Automation*, 4:1222–1227.

[13] S. Tricamo R. Patton, F. Swern and A. van der Veen. Automated cloth handling using adaptive force feedback. *Journal of Dynamic Systems and Measurements and Control*, 114:731–733.

[14] A. M. Howard and G.A. Bekey. Intelligent learning for deformable object manipulation.

[15] D. E. Orin and S. Y. Oh. Control of force distribution in robotic mechanisms containing closed kinematic chains. *" Trans. of the ASME, J. of Dynamic Systems, Measurement, and Control*, 102:134–141, 1981.

[16] C. A. Klein and S. Kittivatcharapong. Optimal force distribution for the legs of a walking machine with friction cone constraints. *IEEE Trans. on Robotics and Automation*, 6:73–85, 1990.

[17] F. T.Cheng and D. E. Orin. Efficient formulation of the force distribution equations for simple closed-chain. *IEEE Trans. on Systems, Man, and Cybernetics*, 21:25–32, 1991.

[18] J.Cornella R.Suarez and R.Carloni. Dual programming based approach for optimal grasping force distribution. *" Mechatronics (Oxford)*, 165:348–356, 2008.

[19] Nahon.MA.and Angeles.J. Optimization of dynamic forces in mechanical hands. *J. Mech. Trans. Autom. Design*, pages 167–173, 1991.

[20] M. A. Nahon and J. Angeles. Real-time force optimization in parallel kinematic chains under inequality constraints. *IEEE Trans. on Robotics and Automa.*, 12:87–94, 1999.

[21] KOO Patrick. An analytical expression for the generalized forces in multibody lagrange equations. *IEEE transactions on robotics and automation*, 20:340–343, 2004.

[22] H. Liu and B. Wen. Force distribution for the legs of quadruped robot. *Journal of Robotique Systemes*, 14:1–8, 1997.

[23] G.Lin D.Wang and LI. The analytical dynamic model of six-dof industrial robotic manipulators of containing closed chain. *Mechanism and machine theory*, 40:385–393, 2005.

[24] X. Chen and K.Watanabe. Optimal force distribution for the legs of quadruped robot. *Machine inteligence and robotique control*, 1:87–94, 1999.

[25] W. Kwon and B.C. Wen. A new optimal force distribution scheme of multiple cooperating robots using dual method. *J. of Intelligent and Robotic Systems*, 21:301–326, 1998.

[26] C. Mahfoudi K. Djouani S. Rechak and M.Bouaziz. Optimal force distribution for the legs of an hexapod robot. *IEEE Conference on Control Application CCA Istambul*, 1:651670, 2003.

[27] C. Mahfoudi K. Djouani M.Bouaziz and S. Rechak. General method for dynamic modeling and control of an hexapod robot including optimal force distribution. *WSEAS, Transaction on Signal Processing*, 2:323330, 2006.

[28] W. Khalil and E. Dombre. *Modélisation identification et commande des robots*. Hermes sciences, Paris, 1999.

[29] T.watanabe and T.yochikawa. Grasping optimization using a required external force set. *IEEE transactions on automation science and engineering*, 4:52–66, 2007.

[30] S.Daniel ; EDIN Benoni B. Prediction of object contact during grasping. *Experimental brain research*, 3:265–277, 2008.

[31] X.Zhu and H.Ding. Computation of force-closure grasps. *IEEE Trans. Robot. Autom.*, 22(1):172–179, (2006).

[32] J.Gerardo and S.Mendez. *Sensors, Focus on Tactile, Force and Stress Sensors*. I-Tech, Vienna, Austria, Austria, December 2008.

[33] F. T. Cheng and D. E. Orin. Efficient algorithm for optimal force distribution-the compact-dual lp method. *IEEE Trans. on Robotics and Automation*, 6:178–187, 1990.

[34] Xiaoping Yun Eric Paljug and Vijay Kumar. Control of rolling contacts in multi-arm manipulation. *IEEE TRANSACTIONS ON ROBOTICS AND AUTOMATION.*, (1994).

[35] Nilanjan Sarkar anf al. Dynamic control of 3d roling contacts in two arm manipulation. *" IEEE Trans. on Robotics and Automation*, page 364376., 1997.

[36] Q. Xiding and G. Yimin. Analysis of the dynamics of a six-legged vehicle. *" The International Journal of Robotics Research*, 14:1–8, 1995.

Design of a trajectory generator for a hexapod walking robot

Chawki Mahfoudi [(*)]__ **Mohamed Bouaziz** [(**)]**Allaoua Boulahia**[(*)]__ **Karim Djouani** [(***)]__ **Said Rechak** [(**)]

(*) Mechanical Engineering Department
Faculty of sciences and Technology
Larbi ben M'Hidi University , Oum El Bouaghi 04 000 , Algeria
c_mahfoudi_dz@yahoo.fr ,   *all_boulahia@yahoo.fr*

(*) Mechanical engineering and development laboratory
National Polytechnic Superior School,
Mechanical Engineering Department
Hassene Badi avenue , Nb 10 , El Harrach (Algiers) 16200 , Algeria
mbouazizdz@yahoo.f

(***) Paris 12 University, LISII Laboratory , France
djouani@univ-paris12.fr

ABSTRACT
We present in this paper an original investigation on a walker hexapod robot for which we define a locomotion strategy considering the walk appearance characterized by the succession in time of the synchronized events of legs putting down and raising. The legs movements' synchronization is determined by the temporal phase displacement between the instants of legs putting down for which we define, for each one, the ratio between the duration passed spent on the ground and on the duration devoted to the replacement overhead phase. For the appearance implementation, we take into consideration the space distribution of legs ends trajectories, the one with regard to the other, and with regard to the gravity centre of the platform which plays a determining role on the locomotion stability. A simulation is made to illustrate the proposed approach results.

Keywords: *hexapod robot, walking robot, trajectories generation, robot cinematic modelization, robot locomotion, robot walking simulation*

1. INTRODUCTION

Papers When you want to pull robot's legs, we must provide high level orders to indicate in which direction and how fast it should move. This is called task-oriented instructions. These high level instructions come from a supervisory system, which can be represented by a module loaded on the machine or by an operator who directs the robot remotely. As appropriate, the robot can be autonomous or tele - operated. Once the machine is aware of the orders that have been assigned, it must generate instructions for walking working gait. To this end, it has a locomotion module, responsible for selecting the speed to travel on, set the parameters and calculate the corresponding instructions to be sent to the legs actuators engines.

The walking gaits generation can be open or closed loop. It depends on the sensory capabilities of the robot. In open-loop locomotion, it is necessary to know at least the status of the actuators (typically, position and speed). These are associated proprioceptive sensors that provide such information. On the contrary, during a closed -loop locomotion, the command may consider certain information reflecting the robot's interactions with its environment. (Mahfoudi and al. 2003), (Martin, 1998).

The strategy of locomotion defines the nature of the walking working gait and its implementation. A walking working gait is characterized by the time sequence of events to putting down and raising legs. The timing of these events differs according to walking pattern. To define the synchronization mode, we define the temporal phase shifts between times to put down different legs. Furthermore, it must be determined for each leg the ratio of time spent on the ground and the time devoted to the replacement of air phase. Thus, the definition of a walking working gait requires control of these parameters. However, during the implementation of the gait, other parameters are taken into account. These depend on the structure of the robot. They concern the spatial distribution of trajectories of the ends of the legs relative to each other and from the gravity center of the platform. The position of gravity center with respect to the trajectories of legs plays a decisive role for locomotion stability. The definition of a walking motion reference is important for designing a walking robot and the study of control laws. These laws are often written to follow the imposed trajectories.

For one leg of a walking robot, the generation of a walking motion is to calculate the instructions reference: position, velocity and acceleration, which are a function of time and ensure the passage of the foot of the leg by an imposed trajectory, defined by mode of travel of the robot. (Hugel, 1999), (Randal, 1999).

As the robot moves in the operational space, the generation of movement of the leg is described in this space, this implies the use of reverse geometric method (Khalil et al., 1999), (Paul, 1981) to transform each point of this trajectory in corresponding articular coordinates.

The principle of the generation of motion in operational space is shown in Figure 1 (the exponents i, f and d designate respectively the initial, final and desired position).
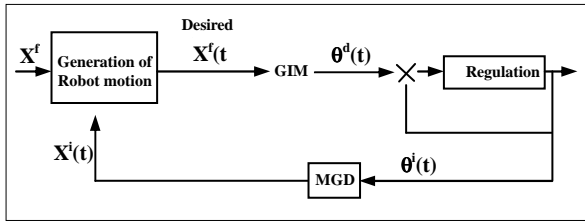
Figure 1: Generation of motion in the operational space

2. THE WALKING OF HEXAPOD

The type of walking seen in some insects is the alternated tripod where we observe at every moment, legs alternate in transfer and support phase, threes by threes (Ridderström, 2003) (Pfeiffer et al., 1995)

The hexapod walks using three legs at once, so a complete cycle consists of two phases (Figure 2): In the first phase, the robot is supported by legs (1,2 and 3); and the platform moves over a distance λ_0 The legs (4,5 and 6) move into the air on a cycloidal path between times t_1 and t_2.

During the second phase, the cycle of phase "1" is repeated except that it is the legs (4, 5, and 6) that are carrier and legs (1, 2 and 3) are in the air. Phase "2" occurs between times t_2 and t_3 (Figure 2).
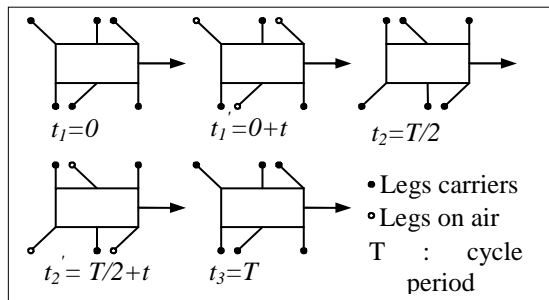
Figure 2: Composition of the hexapod motion during a cycle

2.1. Study of leg motion

During the walking, a step motion of the leg is done in two phases:

> - Support Phase: the leg is in contact with the ground, it must bear   the
> body and   propel the robot forward;

> - Transfer Phase: the leg is lifted and moved in the air passing from one
> step to the next.

2.1.1. Trajectory of lifted legs

The motion of a lifted leg affects the platforms in unstable manner. The choice of a trajectory of leg in the air must take into account the desired characteristics for the lift and pose moments, i.e. flexibility of motion and minimizing collisions with the ground, carrying out circuits through imposed time by the executed gait, adapting to changes in ground elevation.

A trajectory that respects these conditions is that of a cycloid (Villard, 1993), (Quinn et al., 2003). A cycloid is the trajectory in a plane of a point of a circle which rolls without slipping on a line on this plane as shown in Figure 3. Thus, for our case and as defined earlier a cycloid is the trajectory of the lifted leg in the absolute reference. Equations of x and z coordinates of the cycloid as a function of the rolling angle δ are:

$$\overrightarrow{OM}\begin{cases} x = a(\delta - \sin\delta) \\ z = a(1 - \cos\delta) \end{cases} \quad (1)$$
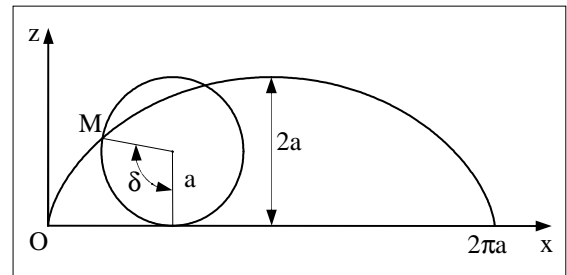

Figure 3: Leg Trajectory in a cycloid form

3. Procedure for generating hexapod motion

To apply the changes to relative motions between the center of gravity of the robot body and feet legs, we define the following parameters (Mahfoudi et al., March and May 2006) as shown in Figure 4:
- L_1, l_2, l_3, a, b: geometric parameters of respectively the leg and the platform
- R_0: Landmark linked to the center of gravity. Its location is calculated with respect to the benchmark base R using the homogeneous transformation

- λ_0 : displacement Value of center of gravity for a sample of walking
- $P = [P_1, P_2, P_3]^T$: Vector coordinates of the leg foot with respect to the reference R
- $PP = [PP_1, PP_2, PP_3]^T$: Vector coordinates of the foot of the leg with respect to the benchmark R_0
- PP_0: starting point of the cycloid with respect to the reference R.
- PP_n: end point of the cycloid, with respect to the reference R.
- λ_{00}: stride of the cycloid measured in the same plane of the cycloid
- N: a period of gait cycle
- S: a simple index that splits the interval n
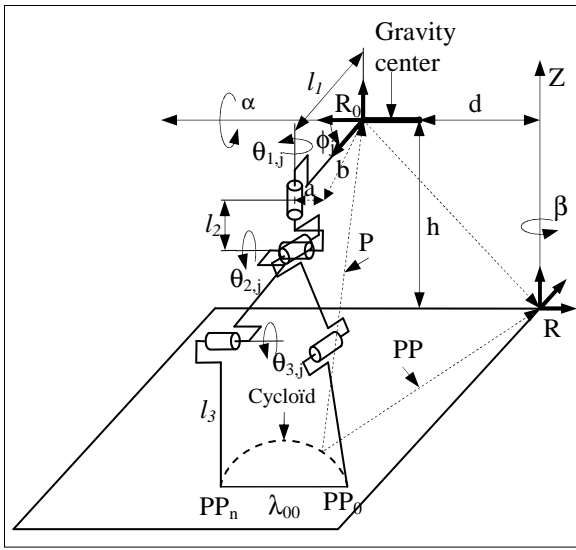
Figure 4: Generation Scheme of a walking leg

3.1. Approach of the platform motion generation

To generate a walking step of the hexapod, we must ensure coordination of the motion of the legs feet with the one of the center of gravity of the robot body. The procedure is as follows:
- Define the geometric parameters of the leg and the platform: l_1, l_2 and l_3, (length of leg components), a, b (half-width and half length of the platform, respectively)
Initialize a location of the legs with respect to the center of gravity of the platform

$$P_{01} = [0, l_2 + a, -h, 1]^T$$

$$P_{02} = [-b, -l_2 - a, -h, 1]^T$$

$$P_{03} = [b, -l_2 - a, -h, 1]^T \qquad (1)$$

$$P_{04} = [0, -l_2 - a, -h, 1]^T$$

$$P_{05} = [b, l_2 + a, -h, 1]^T$$

$$P_{06} = [-b, l_2 + a, -h, 1]^T$$

3.1.1 Equation of motion of the platform center of gravity.

We determine the equation of motion of the center of gravity of the robot (of origin R_0) in the reference R by the following parameters X, Y, Z, β, h, θ_0, d and α. This is actually the homogeneous transformation $^R T_0$ that permits to locate R_0 with respect to R. we define:

- $X = f(t)$: where f (t) is a function of time
- $Y = g(x)$: The vertical axis of center of gravity may be a function of time and X
- $Z = h$: height of center of gravity. It can be fixed or variable with time, X and Y
- $\theta_0 = \mathrm{arctg}(\dot{y})$: Rotation around the axis z0
- $d = \sqrt{x^2 + y^2}$: Distance covered by the center of gravity
- $\beta = \mathrm{arctg}(y/x)$: Angle of overall rotation around the z axis of the reference
- α: angle of rotation around the axis x_0 of the platform. It can be fixed or variable with respect to time

3.1.2. Motion of carrying legs

During the t_i and t_{i-1} moments the platform receives a combined motion of a rotation $\theta_{0i} - \theta_{0i-1}$ around the z_0 axis and a translation of the center of gravity λ_0. Assuming that the carrying legs are (1, 2 and 3) then the points of support $P_j = [P_{1,j}, P_{2,j}, P_{3,j}]^T$ will push back with respect to the center of gravity (Figures 5, 6).

$$\mathrm{Rot}(\theta_{0i} - \theta_{0i-1}) = \begin{pmatrix} \cos(\theta_{0i} - \theta_{0i-1}) & -\sin(\theta_{0i} - \theta_{0i-1}) & 0 & 0 \\ \sin(\theta_{0i} - \theta_{0i-1}) & \cos(\theta_{0i} - \theta_{0i-1}) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$
$$(2)$$

$$\lambda_0 = \sqrt{(X_i - X_{i-1})^2 + (Y_i - Y_{i-1})^2} \qquad (3)$$

The corresponding homogeneous translation matrix is:

$$\mathrm{Trans}(\lambda_0) = \begin{pmatrix} 1 & 0 & 0 & -\lambda_0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \qquad (4)$$

The new coordinates of carrying legs (1, 2, and 3) with respect to the center of gravity are:

$$P_j = Rot(\theta_{0i} - \theta_{0i-1})Trans(\lambda_0)PO_j \quad (j = 1, 2, 3) \quad (5)$$

3.1.3. The motion of legs in the air phase

The legs (4, 5, and 6) are in the air phase. They each carry a cycloid that has parameters $PP_{0,j}$, $PP_{n,j}$ and $\lambda_{00}(j)$ defining respectively the starting, the ending and the stride of each leg. These parameters are calculated in advance.

For legs that carry motion on the air phase along a cycloid, we must calculate the intermediate points belonging to the cycloid. This calculation is made with respect to the R reference as follows (Figures 5 and 6):

- **Calculation of the takeoff point of the leg :**

$$Rot(\beta_0) = \begin{pmatrix} \cos\beta_0 & -\sin\beta_0 & 0 & 0 \\ \sin\beta_0 & \cos\beta_0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (6)$$

$$Rot(\beta_0 - \theta_{00}) = \begin{pmatrix} \cos(\beta_0 - \theta_{00}) & -\sin(\beta_0 - \theta_{00}) & 0 & 0 \\ \sin(\beta_0 - \theta_{00}) & \cos(\beta_0 - \theta_{00}) & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

$$(7)$$

$$Trans(d_0) = \begin{pmatrix} 1 & 0 & 0 & d_0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (8)$$

$$Trans(\lambda_0) = \begin{pmatrix} 1 & 0 & 0 & h_0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (9)$$

The initial positions of the feet coordinates of legs (4, 5, and 6) with respect to R are:

$$PP_{0j} = Rot(\beta_0)Trans(h_0)Trans(d_0)Rot(\beta_0 - \theta_{00})P_{0j}$$
$$(10)$$

- Calculation of the landing point of the leg:

Although we are at the starting point, one expects to predict the landing point of the leg to determine the equation of the cycloid that must follow the foot of the

leg. For this, we assume that we are at the point $n/2$; we then obtain PP_{nj} the same way as PP_{0j}.

$$PP_{nj} = Rot(\beta_{n/2})Trans(h_{n/2})Trans(d_{n/2})Rot(\beta_{n/2} - \theta_{0n/2})P_{0j}$$
$$j = (4, 5, 6) \quad (11)$$
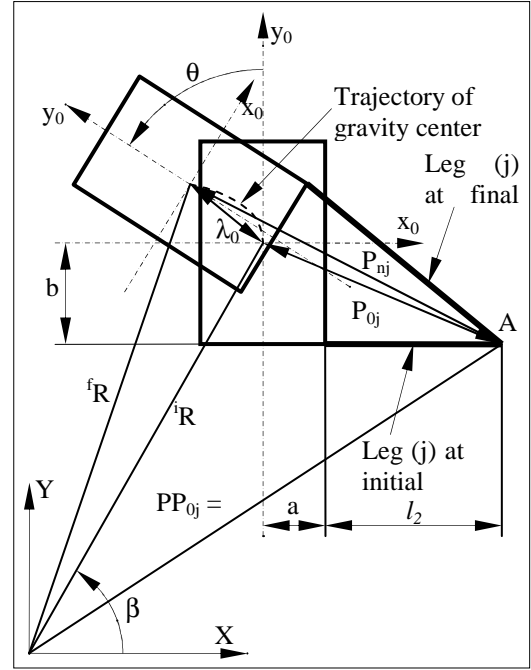
Figure 5 : Carrying leg

- Calculation of travelled distances λ_{00} (j) by the feet of legs:

$$\lambda_{00}(4) = \sqrt{(PP_{n4x} - PP_{04x})^2 + (PP_{n4y} - PP_{04y})^2}$$
$$(12)$$

$$\lambda_{00}(5) = \sqrt{(PP_{n5x} - PP_{05x})^2 + (PP_{n5y} - PP_{05y})^2}$$
$$(13)$$
$$(8)$$

$$\lambda_{00}(6) = \sqrt{(PP_{n6x} - PP_{06x})^2 + (PP_{n6y} - PP_{06y})^2}$$
$$(14)$$

- Calculation of intermediate points (i) constituting the cycloid: (9)

To find the equation of the cycloid corresponding to our case, we introduce the following parameters:

$$\begin{cases} a_{cy} = \dfrac{\lambda_{00}(j)}{2\pi} \\ \zeta = \dfrac{1}{2}, \quad t = s - \dfrac{n}{2}, \quad T = n \\ \delta = \dfrac{2\pi}{(1-\zeta)T}t = \dfrac{4\pi}{n}t \end{cases} \quad (15)$$

With

a_{cy}: radius of the circle

t : time variable of transfer phase
T : period of a walking step

A limitation of this choice is the trajectory height $2a_{cy}$ which is related to the travelled distance. To obtain a usable path, we propose a cycloid corrected time. Using a factor k, we can choose the height of the trajectory. The equation of the cycloid in the travel plane for each leg "j" is:

$$\begin{cases} X_\lambda(i) = \dfrac{\lambda_{00}(j)}{2\pi}\left\{\dfrac{4\pi}{n}t - \sin\left(\dfrac{4\pi}{n}t\right)\right\} \\ Z_\lambda(i) = k\dfrac{\lambda_{00}(j)}{2\pi}\left\{1 - \cos\left(\dfrac{4\pi}{n}t\right)\right\} \\ Y_\lambda(i) = 0 \end{cases} \quad (16)$$

$k \leq 1$: correction factor of the height of the cycloid, i varies from 0 to n/2

The foot velocity at the ideal moment for both leg lifting and setting down is equal to zero. The events of setting down and lifting can then occur without collision or serious interruption of motion of the leg.

The speed components expressions $\dot{X}$ and $\dot{Z}$ are:

$$\begin{cases} \dot{X}_\lambda(i) = \dfrac{2\lambda_{00}(j)}{n}\left\{1 - \cos\left(\dfrac{4\pi}{n}t\right)\right\} \\ \dot{Z}_\lambda(i) = k\dfrac{2\lambda_{00}(j)}{n}\sin\left(\dfrac{4\pi}{n}t\right) \\ \dot{Y}_\lambda(t) = 0 \end{cases} \quad (17)$$
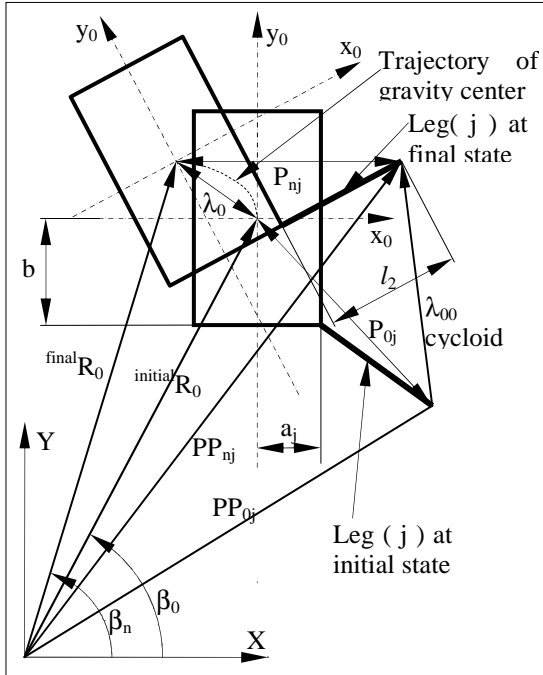
Figure 6 : Leg in air phase

Like any motion must be referenced to the reference base R, then the equation of the cycloid in this benchmark is given by:

$$\begin{cases} X_{PP}(i) = \left(X_{PP_n} - X_{PP_0}\right)\dfrac{X_\lambda(i)}{\lambda_{00}(j)} + X_{PP_0} \\ Y_{PP}(i) = \left(Y_{PP_n} - Y_{PP_0}\right)\dfrac{X_\lambda(i)}{\lambda_{00}(j)} + Y_{PP_0} \quad (18) \\ Z_{PP}(i) = Z_\lambda(i) \end{cases}$$

Its speed in R is given by:

$$\begin{cases} \dot{X}_{PP}(i) = \left(X_{PP_n} - X_{PP_0}\right)\dfrac{\dot{X}_\lambda(i)}{\lambda_{00}(j)} + X_{PP_0} \\ \dot{Y}_{PP}(i) = \left(Y_{PP_n} - Y_{PP_0}\right)\dfrac{\dot{X}_\lambda(i)}{\lambda_{00}(j)} + Y_{PP_0} \quad (19) \\ \dot{Z}_{PP}(i) = \dot{Z}_\lambda(i) \end{cases}$$

Finally, we must rewrite all vectors PP_j, j = (4, 5, and 6) in the benchmark R_0:

$$P_j = \text{Trans}(-h_i)\text{Trans}(-d_i)\text{Rot}\left(-\beta_i - \theta_{0i}\right)\text{Rot}(-\beta_i)PP_j$$
$$(20)$$

- Calculation of new initial positions:

At the end of each calculation step we must reset the new positions of the feet of the legs with respect to the benchmark R_0 as follows:

$$P_{0j} = P_j \quad j = (1, 2....6)$$

Transition from phase "one" to phase "two" corresponding to i = n / 2 in the gait cycle

We separate the two phases of a walking step so that the transfer phase begins when ends one of the support. This separation is achieved through the division of the interval n into two halves and using a test variable s which is compared with the value n / 2

We then define the motion performed during each phase. When s exceeds the value n/2, starts the second phase which consists to swap the legs (1,2,3)in air phase while (4,5,6) will be carriers. Finally the walking cycle ends when s reaches the value of n.
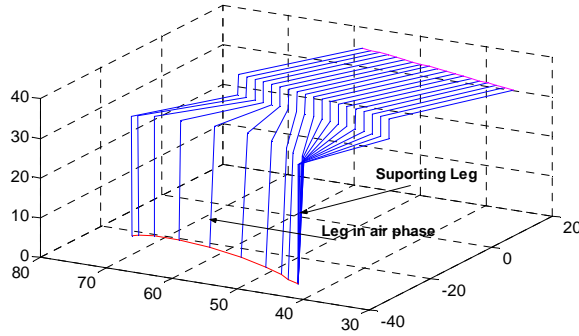
– **Calculation of articular coordinates**

we use the geometric inverse method (GIM) to calculate the articular coordinates $\theta_{1,j}$, $\theta_{2,j}$ and $\theta_{3,j}$ and for j = (1,2, .. 6) as a function of operational coordinates PP_j of legs feet with respect to the reference R. Finally, we draw at each step calculation the corresponding configuration of the hexapod.

3.2. Simulation of hexapod motions

To implement this procedure a set of subroutines linked by a main program was programmed with Matlab software to generate walking gaits for hexapod. A graphical representation of the hexapod is thus possible.

The following figures show the movement of the hexapod during a cycle.

Figure 7 shows a leg tip simulation for a walking cycle

Figure 7: Simulation of a walking step

Figure 8 shows in the operational space and in degree unit the profile of the three suitable articular coordinates and also the desired coordinates of the end of the leg (in cm). Figure 9 shows the hexapod moving in a straight line on a horizontal plane
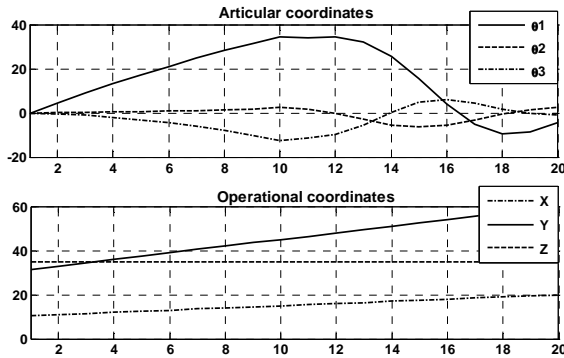
Figure 8: Pattern Evolution of articular and operational coordinates

Figure 9 shows the hexapod walking in accordance with a trajectory of equation Y = 3X and making 80 cycles (for not to clutter up the diagram, we have represented 1 cycle of 15. Figures 10 and 11 represent the detailed movement of the hexapod for a walk cycle.We have imposed to the platform of the hexapod a sinusoidal trajectory in a vertical plane that intersects with the horizontal plane according to the equation Y = 3X.
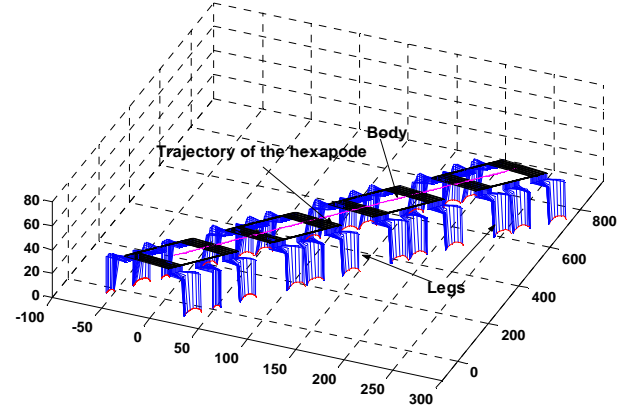
Figure 9: Hexapod motion performing several cycles
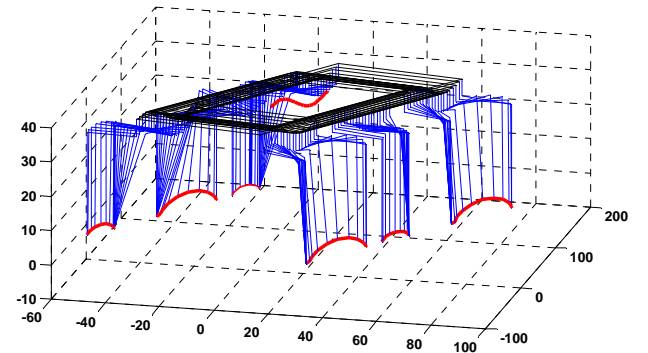
Figure 10: The hexapod walking following a sinusoidal vertical trajectory during a cycle
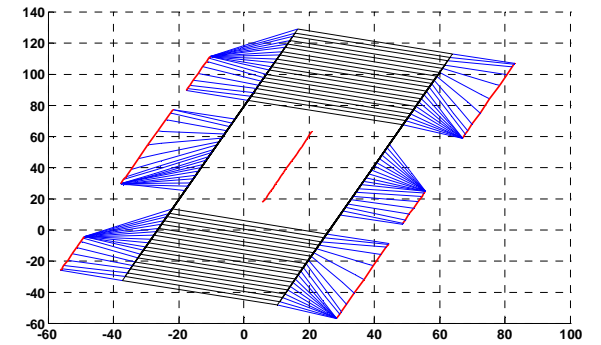
Figure 11: Top view of the hexapod motion

Figures 12 and 13 show respectively the articular coordinates and articular velocities. The hexapod is walking on a straight line of equation Y = 3X, on a horizontal plane for two cycles.

Figures 14 show the operational coordinates of the tips of legs expressed in the benchmark associated to the platform. The hexapod is walking on a straight line of equation Y = 3X on a horizontal plane for two cycles.
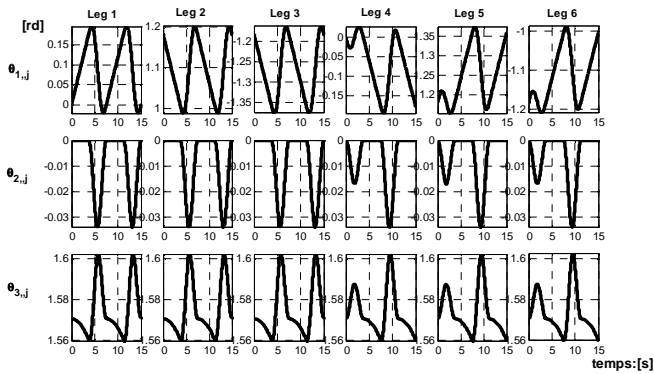
Figure 13: Articular velocities of the six legs of the hexapod
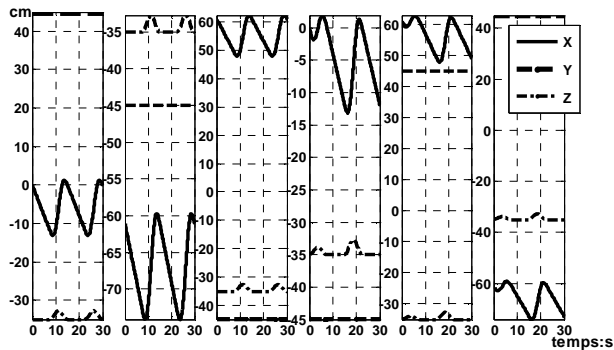
Figure 14: Operational tips Coordinates of the hexapod six legs

4. Experimentation

Therefore, a First experimentation was conducted on our experimental walking machine figure (15). Our experimental hexapod uses two hardware cards. The First one, based on a Microchip 2 micro-controller is dedicated for low level motors control. While the second one, which is a PC104 card under real time OS, is dedicated for computed torque control and real time force distribution computing. Finally, a PC is used for trajectory planning and real time monitoring.
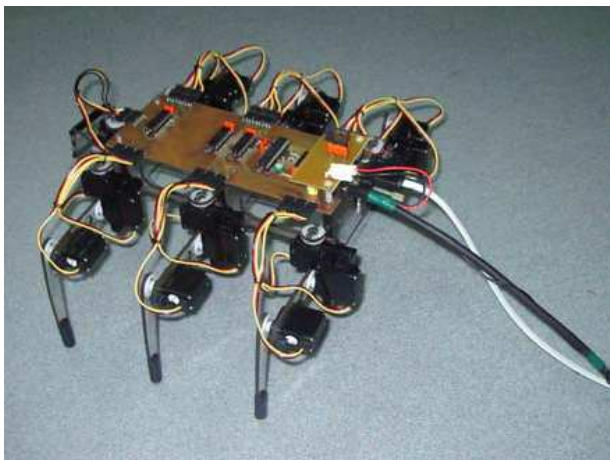
Figure 15: Experimental hexapod

5. Conclusion

The main objective of this work is to implement a quite general walking generator. The trajectory generator uses the alternating tripod gait. The platform can move following any time of law, but according to the constraint of attainable field legs carriers. Thus the robot simulates walking on flat field following trajectories with a quite large radius of curvature. The trajectory generator conceived in such manner permit to read back at any time values of the articular coordinates, velocities and accelerations, all the center of gravity kinematic of the platform, as well as coordinates of the tips of the legs with their corresponding velocities and accelerations in the workshop benchmark.

2.1. Length of the Paper
The International Program Committee will accept two types of camera ready papers: extended paper (10 pages, 2 columns); regular paper (6 pages, 2 columns) both for regular and special sessions.

2.2. Margins
The left and the right margins must be 1,9 cm on each page, and the gutter between the two columns of text must be 1,25 cm in width. The top and the bottom margins must be 2,54 cm.

6. LIST OF REFERENCES

[1] Mahfoudi C., Djouani K., S. Rechak S., Bouaziz M. « Optimal force distribution for the legs of an hexapod robot ». *IEEE 2003, Conference on Control Application CCA 2003*, June 23-25, Instambul Turkey.

[2] Marttin L., Une approche de supervision de marche pour robot quadrupèdes intégrant commande et décision, Thèse de doctorat, Université Paul Sabatier de Toulouse, 1998.

[3] Randal M.J., Stable adaptive neural control of systems with closed kinematic chains applied to biologically-inspired walking robots. PhD Thesis, Université de Paris VI, 1999.

[4] Khalil W., Dombre E., *Modélisation, identification et commande des robots*. Hermès Sciences, Paris, 1999.

[5] R.C. Paul R.C., Robots manipulators, mathematics, programing and control. MIT press, 1981.

[6] Villard C., Etude du comportement dynamique d'un robot quadrupède, PhD Thesis. Université Pierre et Marie Curie, 1993

[7] Ridderström C., Legged locomotion : Balance, control and tools-from equation to action. Thèse de doctorat, Institut Royal de la technologie, Stocklom, Suède, 2003.

[8] Pfeiffer P., Eltze J., Weidemann H.J., « Six-legged technical walking considering », *Bbiological principal Robotics and Autonomous Systems*, pages 223–232, 1995.

[9] Quinn R.D., Laksanacharoen S. and Roy Ritzamann E., Modeling of insects legs by inverse kinematics analysis, *Proc. of the 2ⁿᵈ International Symposium on Adaptive Motion of Animals and Machines*, Kyoto, 2003.

[10] Mahfoudi C, Rechak S., A. Aissaoui A., Djouani K, Bouaziz M., Réalisation et étude de la commande d'une patte d'un robot marcheur. *5ᵉᵐᵉ Journée de Mécanique JM'05*, Ecole Militaire Polytechnique, Bordj El Bahri (Alger), 28-29 mars 2006.

[11] M. Sufierdem and K. Leblebiciogla. Free gait generation with reinforcement learning for a six-legged robot, Robotics and Autonomous systems, August 2007

[12] C. Mahfoudi K Djouani, M. Bouaziz S. and Rechak. General methode for dynamic modeling and control for an hexapod robot. WSEAS, transaction on signal processing, P323-330*2003, Issue 2,vol 2, February 2006*

[13] T.D. Barfoot and al Experiments in learning distribution control for a hexapod robot, Robotics and Autonomous systems, Avril 2006.

USING LINEAR PROGRAMMING FOR THE OPTIMAL CONTROL OF A CART-PENDULUM SYSTEM

Luiz Vasco Puglia[(a)], Fabrizio Leonardi[(b)], Marko Ackermann[(c)]

[(a)]UNIP Universidade Paulista, São Paulo, Brazil, and Faculdades Integradas Torricelli, Guarulhos, Brazil
[(b) (c)] Centro Universitário da FEI, São Bernardo do Campo, Brazil

[(a)] lvpuglia@gmail.com, [(b)]fabrizio@fei.edu.br, [(c)]mackermann@fei.edu.br

ABSTRACT

This paper discusses the use of linear programming for the optimal control of a cart pendulum system. The objective function and the constraints are designed to minimize the control effort and the time duration of the operation. Simulations and experimental tests were performed. Restrictions of null angle and angular velocity at the extremes were incorporated in the design specification as well as other physical constraints. In order to compensate for the modeling errors and disturbances, the optimal trajectory was kept within a prescribed precision by means of a closed loop system. The obtained results illustrate that the technique is simple, powerful and always conclusive.

Keywords: Linear Programming, Optimal Control, Anti-oscillatory Control.

1. INTRODUCTION

The problem of optimal control of cranes has been receiving attention from the scientific community because of undeniable practical relevance. As pointed out by Sorensen, Singhose, and Dickerson (2007), the control schemes developed in the literature may be grouped into three categories: time-optimal control, command shaping, and feedback control. Their paper addresses the perturbations by feedback control and, since the trajectories are not known a priory, they used a feed-forward control in the form of an input shaping in order to reduce the motion-induced oscillation.

Several studies (Cheng and Chen, 1996; Auernig and Troger, 1987; Cruz, Leonardi, and Moraes, 2008; Chen, Hein, and Wörn, 2007; Nassif, Domingos, and Gomes, 2010; Garrido *et al.* 2008; Lee, 2004) address the problem of minimum time and differ, for example, with respect to the model utilized, the constraints imposed and the performance index optimized.

In load transfer operations by a crane, a major problem is optimizing the movement from origin to destination, satisfying constraints related to the equipment and to the kinematics of the movement. The carrier may be considered primarily as a cart-pendulum system, where the length of the pendulum is usually variable, representing the lifting.

One difficulty in solving optimal control problems such as the optimal load transfer by a crane is the necessity of solving a two-point boundary value problem, i.e., with constraints on the initial and final states. For instance, a linear quadratic regulator generates an optimal control law but the final state cannot be pre-determined. This limitation is discussed e.g. by Bemporad, Borelli, and Morari (2002) and by Blanchini (1994).

This work addresses the problem for known trajectory boundaries, which is typical for ship unloading operations. For this kind of problem it is highly desirable to have a motion planning scheme that ensures swing reduction and minimum time operation. Feedback control is used to reduce external perturbations and the optimal control trajectory is obtained by solving a simple Linear Programming problem. Thus, the physical constraints can be included explicitly in the design. In a manner similar to the proposed by Sorensen, Singhose, and Dickerson (2007), the cart kinematics is determined by means of an independent feedback control.

This paper discusses the use of Linear Programming as an alternative for solving this type of optimal control problem, assuming that the system dynamics are linear in the state space in the discrete time domain. In this scenario the discrete values of the control vector are the free design variables and the state vector at any sampling time may be written as a linear combination of the control vector and the initial condition. This results in the standard structure of a Linear Programming (LP) problem.

A cart-pendulum lab system was considered to illustrate the proposed approach. The movement cycle begins and ends at given positions and the load is at rest in both, the beginning and end of cycle. Moreover, in the application considered here, the lifting takes place at the beginning and end of the cycle with the cart stationary, i.e., there is no lift during cart movement.

A more efficient strategy is obviously to perform lifting and cart translation simultaneously, but this work intended to show the potential of the methodology, applying it to a lab-scale system that has no motorized lifting.

2. METHOD

The optimal control problem of a dynamical linear system in the discrete time state space can be written in the form of a standard LP problem.

2.1. Linear Dynamics as LP Constraints

Consider the dynamical system in discrete time with a constant sampling time T and described in the state space

$$x_{(k+1)} = A x_{(k)} + B u_{(k)} \quad (1)$$

For any sampling time nT we can write

$$x_{(n)} = A^n x_{(0)} + A^{n-1} B u_{(0)} + A^{n-2} B u_{(1)} + \cdots$$
$$\cdots + A^1 B u_{(n-2)} + A^0 B u_{(n-1)}$$

$$x_{(n)} = F x_{(0)} + G U \quad (2)$$

where

$$F = A^n$$
$$G = \left[A^{n-1} A^{n-2} \cdots A^1 A^0 \right] diag[BB \cdots BB]$$
$$U = \left[u_{(0)} u_{(1)} \cdots u_{(n-2)} u_{(n-1)} \right]^T$$

Note that it is possible to represent the dynamic model as constraints in the form of $AX = B$, which may include the initial conditions $x_{(0)}$ and the final conditions $x_{(n)}$ at the nT instant

$$GU = x_{(n)} - F x_{(0)}$$
$$\mathbf{A}X = \mathbf{B} \quad (3)$$

where $\mathbf{A} = G$, $X = U$ and $\mathbf{B} = x_{(n)} - F x_{(0)}$.

Note that the system dynamics was represented by linear constraints on the control vector. For state constraints in the form of inequalities of type $x_{(m)} \geq \eta$, we can write $x_{(m)} = F_1 x_{(0)} + G_1 U_1$. Thus,

$$F_1 x_{(0)} + G_1 U_1 \geq \eta$$
$$G_1 U_1 \geq \eta - F_1 x_{(0)}$$
$$A_1 X_1 \geq B_1 \quad (4)$$

To completely define the LP problem it remains to define a cost function which is linear on states and controls. The choice of cost function depends on the optimization problem to be solved. For example, one can maximize the average speed or minimize the fuel consumption to travel a given distance.

The objective function can be adapted to the particular optimization problem to be solved. A possible objective in optimal control problems is minimizing the sum of the absolute control values at each sampling time. Another possibility is maximizing the average

speed to indirectly solve the minimum time problem. These and other objective functions are easily written in the standard form of a Linear Programming problem, i.e., as a linear combination of the control vector.

2.2. Mechanical Model

A scheme of the cart-pendulum system used is shown in Fig. (1), where m_T is the cart mass, m_L the load mass, x_T the cart position and ϕ the load angle.
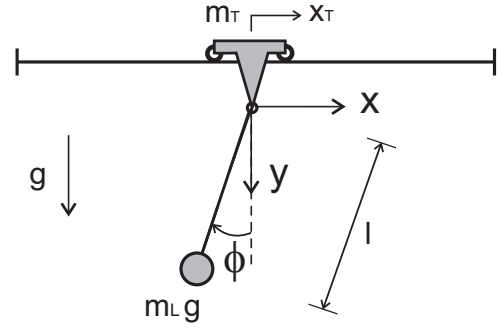

Figure 1: Cart-pendulum scheme

The equations of motion describing the dynamics of the cart-pendulum model were derived using the Newton-Euler formalism as described in Schiehlen (1997) yielding

$$-\ddot{x}_T \cos\phi + \ddot{\phi} l = -g \, sen\,\phi - \frac{\overline{c}\dot{\phi}}{m_L}, \quad (5)$$

where g is the gravity acceleration and $\overline{c}$ a damping constant. In handling anti-oscillatory problems, it is expected that the maximum oscillation angle be small (<10°). This condition leads to the approximations $sen\phi \approx \phi$ and $\cos\phi \approx 1$. These approximations simplify the equations of motion to

$$-\ddot{x}_T + \ddot{\phi} l = -g\phi - \frac{\overline{c}\dot{\phi}}{m_L}. \quad (6)$$

2.3. Optimal Control

The objective function chosen here is the control effort in the form of the sum of the absolute control values $|u_1| + |u_2| + \cdots + |u_n|$.

The standard LP formulation admits a single objective function but we propose an approach to minimize both the control effort and the time of operation. The minimum time is obtained by solving a series of minimal-effort LP problems with decreasing final time until constraints can no longer be satisfied. From (6) and defining $x_1 = \phi$, $x_2 = \dot{\phi}$, $x_3 = x_T$, $x_4 = \dot{x}_T$, $c = \dfrac{\overline{c}}{m_L}$ and $u = \ddot{x}_T$ as the control variable, we get

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -\dfrac{g}{l} & -\dfrac{c}{l} & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} + \begin{bmatrix} 0 \\ \dfrac{1}{l} \\ 0 \\ 1 \end{bmatrix} u \tag{7}$$

with the initial condition $x(t_0) = \begin{bmatrix} 0 & 0 & 0 & 0 \end{bmatrix}^T$ and final condition $x(t_f) = \begin{bmatrix} 0 & 0 & 0.25 & 0 \end{bmatrix}^T$. Based on the limitations of the real plant we have limited $u_{max} = 0.9\,m/s^2$.

Defining the minimum time $t = t_f$ and assuming that t_f is given, consider a problem of minimizing the following functional subject to the two boundaries constraints and denote by $S(t_f)$ this minimum-effort optimal control problem.

$$\min_{u} |u_1| + |u_2| + \cdots |u_n| \tag{8}$$

The following algorithm solves a series of $S(t)$ problems with a tolerance $\epsilon > 0$, were x is an arbitrary real value, such that $x > t_f$.

```
01.  t1 = 0
02.  t2 = x
03.  while |t2-t1| > ϵ
04.      t = (t1+t2)/2
05.      if S(t) exists
06.          t2 = t
07.      else
08.          t1 = t
09.      end
10.  end
11.  tf = t2
12.  u* = S(tf)
```

Thus $u*$ solves $S(t_f)$ with a minimum time t_f. In other words, it is a solution for simultaneous minimum time and minimum effort problem.

2.4. Testing Apparatus

In order to validate the numerical results and implement the proposed control law, it was used a Bytronic lab equipment that allows inverted pendulum or simple pendulum experiments. The schematic diagram of the equipment is shown in Fig. 2.

The pendulum consists of a *0.215 kg* mass connected to the cart by a rod. The mass can be fixed on the rod at different distances from the cart. The cart driver has a position control system with speed compensation. In this loop there is access to the reference signal and the cart position and speed signals.

There is also access to the pendulum angular position signal (not shown in the figure). Since the

implementation details of this internal control system are not well documented we chose to consider this as part of the cart sub-system and its transfer function was experimentally identified via step excitation. A second order transfer function was selected as

$$P(s) = \frac{K_n \omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2} \tag{9}$$

and the parameters $K_n = 0.025$, $\omega_n = 31\,rd/s$, $\zeta = 0.35$ were determined. The gain values for the speed and cart position sensors are $Ktaco = 0.25V/m/s$ and $Kpot = 0.06V/degrees$, respectively.
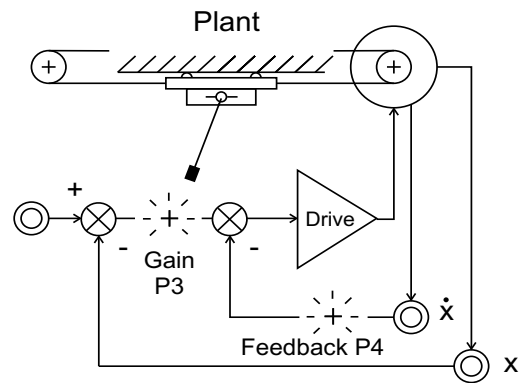

Figure 2: Schematic diagram of the equipment

3. RESULTS

3.1. Simulation

The state equation was discretized with a sampling time $T = 15\,ms$ and the optimal control vector was obtained for *N* sampling periods and including the two boundary constraints. The rod length was set to *l = 0.124m* and after simple binary search, we found *N=74* as the minimum number of sampling periods that still leads to a feasible solution. The results are presented in Fig. 3.

3.2. Experimental Results

The optimal control problem was formulated considering the cart acceleration as the manipulated variable. So a way to impose the cart kinematics in the presence of modeling errors and disturbances is necessary. This was achieved through a state feedback control system with a feed-forward action for the acceleration as illustrated in the block diagram of Fig. 4. Note that, although there is a control loop, the optimal control itself is open loop, since there is no feedback for the angle trajectory.

The control system shown in Fig. 4 has three references that are consistent with each other - position, velocity and acceleration of the cart. The position and velocity are states and, therefore, their references apply to the loop, while the desired acceleration enters as a

feed-forward action through a block that contains the inverse plant model. The gains $\overline{K}_1$ and $\overline{K}_2$ are the gains from state feedback and were tuned interactively in order to obtain good tracking of the reference signals and for disturbances rejection.
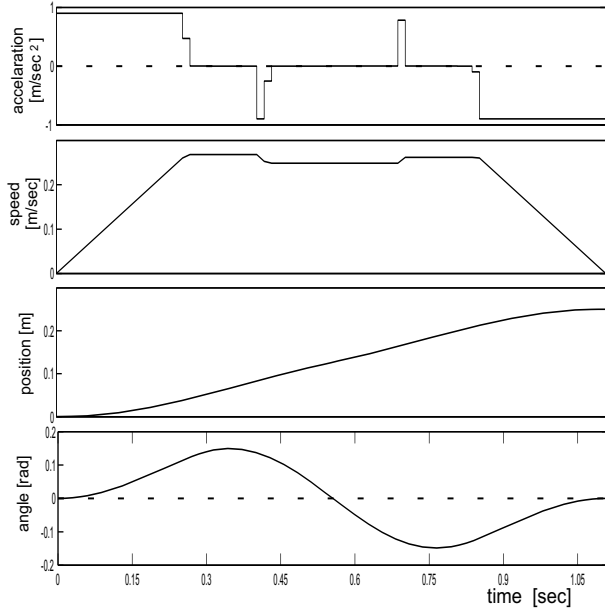

Figure 3: Simulation with *74* sampling periods

Because this control loop cannot be perfect, the optimal control problem will contain errors because the cart acceleration will never be imposed with an infinite precision. The control system along with the trajectory generation and data acquisition, were implemented in Simulink in real time through a data acquisition board and Matlab Windows Target.
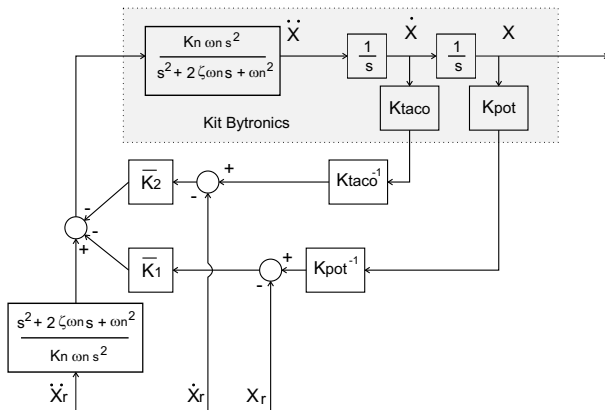
Figure 4: Cart control system

Fig. 5, 6 and 7, show a comparison between the computed and measured position, speed and pendulum angle, respectively. Thin lines were used for generated signals and thick lines for the real signals. Since the experimental apparatus does not have an acceleration sensor for the cart, the real acceleration is not compared with the optimum computed acceleration in none of the following cases.
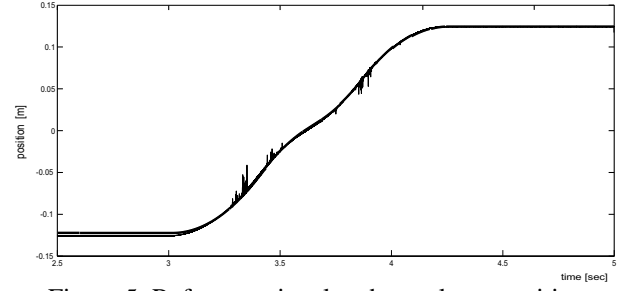
Figure 5: Reference signal and actual cart position
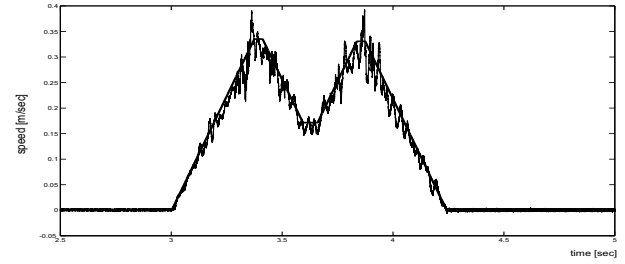
Figure 6: Reference signal and actual cart speed
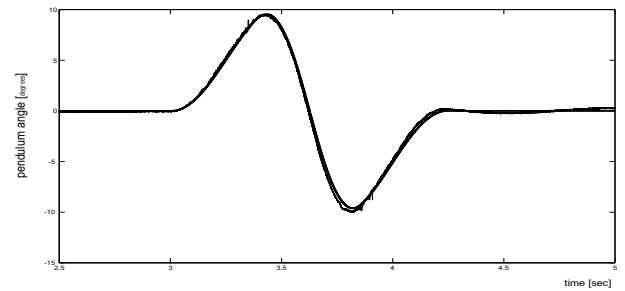
Figure 7: Reference signal and actual pendulum angle

3.3. Sensitivity

A change imposed on the plant model, such as a new mass position, modifies its response. Using the same optimal control vector obtained for *l=0.24m* we evaluate the sensitivity of the system response to variations in rod length *l*. Starting from *l = 0.24 m* changes of length of $\pm 15\,mm$ and $\pm 30\,mm$ in the mass position were performed and the results compared to the optimal trajectory, while keeping the optimal control vector calculated for $l = 0.24\,m$. Figs. 8 to 11 show the effect of changing the rod length on pendulum angle trajectory.
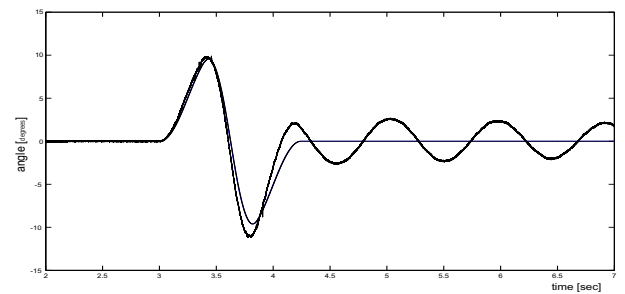

Figure 8: Optimal response for $l\ =\ 0.24m$ and measured response $l\ =\ 0.21m$

Figure 9: Optimal response for $l = 0.24m$ and measured response $l = 0.225m$

Figure 10: Optimal response for $l = 0.24m$ and measured response $l = 0.255m$

Figure 11: Optimal response for $l = 0.24m$ and measured response $l = 0.27m$

The results suggest that this optimal control problem is sensitive to modeling errors. That is, if the plant model is not well known the optimal trajectory will not be assured. This was expected as the optimal control is running with no feedback.

3.4. Closed Loop Control

To make the system less sensitive to the modeling errors and disturbances, a closed loop optimal control strategy was used. To do so, the angular optimal trajectory is used as a reference for a feedback control system applied to the pendulum angle. Thus, the optimal control signal acts as a feed-forward action and this control loop does just the corrections of deviations from the optimal trajectory. Note that modeling errors such as those arising from considering $sin\phi \approx \phi$ and $cos\phi \approx 1$ are also reduced by the feedback control.

Since the positioning of the cart is made by means of three references, it is preferable to work directly with the manipulated variable. The diagram of Fig. 12 shows the complete control system, i.e., the state feedback loop for the cart positioning and the closed loop control

for the pendulum angle. The transfer function used as the angle controller was $G_C(s) = 25/s + 1$ and its parameters were tuned interactively to produce the best insensitivity to the variations of length l.

Figure 12: Complete control system

For evaluation of differences between the two methods, data was collected in both closed loop pendulum angle and open loop. The responses are shown in the Fig. 13 and 14 against the optimal angular trajectory generated by the LP. Fig. 13 shows the effect of keeping $l = 0.24$ m. Fig. 14 shows the effect of changing from $l = 0.24$ m to $l = 0.15$ m.

Figure 13: Cart positions, cart speed and pendulum angle for $l=0.24$ m, closed loop and open loop

In the first case, since the actual length adjusted was exactly the same used for the design, the closed-loop control does not have a noticeable influence. In the second case, shifting the mass position to approximately half of the rod course, it is noted that the angular control loop can practically restore the original behavior of the reference. This is accomplished with the expenditure of

an additional control effort. That is, even though the trajectory is close to the original, there is no guarantee that minimum possible cost is achieved.

Figure 14: Cart positions, cart speed and pendulum angle for *l=0.15 m*, closed loop and open loop

4. CONCLUSIONS

This paper discusses the use of Linear Programming for the solution of a minimum time optimal control problem, including simultaneously the minimization of control effort.

The approach was applied to a cart-pendulum system and the minimal effort problem was solved explicitly by means of an objective function. Since the minimum time problem cannot be written in the form of a linear combination of the control variables, it was solved indirectly by means of a search over the sampling instants. To compensate for modeling errors and disturbances, the trajectory was maintained by means of a closed loop control system with a feed-forward action.

It was found that, by operating in open loop, the trajectory is very sensitive to modeling errors. In these tests, the modeling errors were deliberate and achieved by changing the plant dynamics by shifting the mass position attached to the rod.

The results of this study suggest that the procedure used is very suitable to cases where the real plant can be well approximated by a linear plant model. As a proposal for extending this work the following investigations are suggested.

- Use different objective functions.

- Since the problem is treated in discrete time, it seems to be reasonably simple to extend the results to time-varying systems.

- Use a model matching control structure so that the optimal control is always applied to the same plant.

Thus, if the transfer function of the plant can be maintained within a certain precision, the optimal trajectory will be kept within a pre-established precision as well.

REFERENCES

Auernig, J. W., Troger, H., 1987. Time Optimal Control of Overhead Crane with Hoisting of the Load, *Automatica*, 23(4), 437–447.

Bemporad, A., Borelli, F., Morari, M., 2002. Model Predictive Control Based on Linear Programming – The Explicit Solution, *IEEE Transaction on Automatic Control,* 47(12), 1974–1985.

Blanchini, F. 1994. Ultimate Boundedness Control for Uncertain Discrete-Time Systems Via Set-Induced Lyapunov Functions. *IEEE Trans. Automatic Control*, 39(2), 428-433.

Chen, S. J., Hein, B., Wörn, H. 2007. Swing Attenuation of Suspended Objects Transported by Robot Manipulator Using Acceleration Compensation. *Proceedings of the International Conference on Intelligent Robots and Systems*, pp. 2919–2924 (San Diego, USA).

Cheng, C. C., Chen, Y. C., 1996, Controller Design for an Overhead Crane with Uncertainty, *Control Eng. Practice*, 4(5), 645–653.

Cruz, J. J., Leonardi, F., Moraes, C. C., 2008. Controle Anti-Balanço de tempo Mínimo usando Programação linear Aplicado a um Descarregador de Navios, *Anais do XVII Congresso Brasileiro de Automática*. (Juiz de Fora, Brazil).

Garrido, S., Adberrahim, M., Giménez, A., Diez, R., Balaguer, C. 2008, Anti-Swinging Input Shaping Control of an Automatic Construction Crane. *IEEE Transactions on Automation Science and Engineering*, 5(3), 549–557.

Lee, H.-H. 2004. A New Motion-Planning Scheme for Overhead Cranes With High-Speed Hoisting. *Journal of Dynamic Systems, Measurement, and Control,* 126(2), 359-365

Nassif, L. C. J., Domingos, D. F. B., Gomes, A. C. D. N. 2010. Estratégias de Controle para Reduzir Oscilações em Cargas Pendulares, *Anais do XVIII Congresso Brasileiro de Automática*. pp. 916–922 (Bonito, Brazil).

Olabi, A., Béarée, R., Gibaru, O., Damak, M. 2010. Feedrate Planning for Machining with Industrial Six-axis Robots. *Control Engineering Practice*, 471–482.

Schiehlen, W. 1997. Multibody System Dynamics: Roots and Perspectives". *Kluwer Academic Publishers*, 566–594.

Sorensen, K., Singhose, W., Dickerson, S. 2007. A Controller Enabling Precise Positioning and Sway Reduction in Bridge and Gantry Cranes. *Control Engineering Practice*, 15(7), 825–837.

Algebraic characterization of the invariant zeros structure of LTV bond graph models

D. Yang, C. Sueur, G. Dauphin-Tanguy,

Ecole Centrale de Lille, LAGIS FRE 3303 CNRS, BP 48, 59651, Villeneuve d'Ascq Cedex France
E-mail: christophe.sueur@ec-lille.fr

Abstract— In this paper, invariant zeros structure of linear time varying systems modeled by bond graph is derived by using module theory. Infinite structure of the bond graph model is used to get the number of invariant zeros. In the linear time invariant case, null invariant zeros can be directly pointed out. It is no more true for linear time varying models, the combination of graphic and algebraic methods must be considered. A new procedure based on the finite structure of the bond graph model is given to determine the null value of invariant zeros. Algebraic calculations of torsion modules clarify this difference. Based on a simple RLC circuit, different comparative approaches are proposed. A theoretical form based on Jacobson forms of system matrices is proposed and developed with a Maple programm. Some simulations with 20-sim illustrate the results.

Keywords: invariant zeros, bond graph, LTV system, module theory

1. INTRODUCTION

Linear systems have been intensively studied since fifty years. Invariant zeros are important for the stability analysis of the controlled systems for several well-known control problems such as the disturbance rejection problem, the input-output decoupling problem and some other problems such as the conception of full order or reduced order observers. Different approaches are proposed, according to the choice of a model, such as state space models, transfer models or graphical representations in case of linear time invariant (LTI) models.

The problems have been tackled under various resolution techniques which are often similar, even if formulations are different. Among these techniques, the structural approach, the algebraic approach and the geometric approach which are popular in control theory, appeared to be very effective. Different steps are often proposed. The first step is mainly at an analysis level (study of the internal structure) and the last step deals with synthesis methods. For linear time varying (LTV) systems, algebraic approach has been proposed and developed by several authors [9], [3]. Nevertheless, the extension to the LTV case is not so easy, even if the problem formulations are similar to the LTI case. From the point view of algebraic approach, a linear system is a finitely generated module over a non commutative polynomial ring of differential operator δ. The issue about system poles/zeros is related to solve differential polynomial equations. In [14] it is shown that a skew polynomial can be written as a product of elementary factors $(\delta - \gamma_i)^{d_i}$. This allows one to give intrinsic definition of the poles and zeros of LTV systems by algebraic approach initiated by Malgrange [13] and Fliess [9].

Because of non commutative properties and derivations of time varying coefficients, the bond graph rules proposed in the LTI case for determining invariant zeros are not valid for the LTV case. Some complementary rules must be added. The invariant zeros structure can be studied with an algebraic approach, but the algebraic calculations are often complex. By combining the two methods, a simple procedure to determine the invariant zeros structure is pointed out. In this paper, null invariant zeros are considered. From the point of view of module theory, a null zero corresponds to the factorization of the term $\delta^n (n \geq 0)$ related to *right roots* of torsion module polynomial representation.

This paper is first concerned with some tools for analysis of LTI bond graph models, first with controllability/observability properties and then with the infinite structure related to input-output causal paths. Then, the procedure to get invariant zeros is explained, from bicausal bond graph models. In section 3, algebraic approach with notions of module and polynomial rings are recalled. From this intrinsic point view, the controllability is related to a certain submodule of system. The procedure to determine invariant zeros structure by using algebraic and bond graph methods is then introduced. A LTV system with several cases is studied. The Jacobson form of system matrices and simulation curves are shown.

2. ANALYSIS OF LTI BOND GRAPH MODELS

In a bond graph model, causality and causal paths are useful for the study of properties, such as controllability, observability and systems poles/zeros. Bond graph models with integral causality assignment (BGI) can be used to determine reachability conditions and the number of invariant zeros by studying the infinite structure. In the LTI case, the rank of controllability matrix is derived from bond graph models with derivative causality (BGD). Systems invariant zeros are poles of inverse systems. Inverse systems can be constructed by bond graph models with bicausality (BGB) which are thus useful for the determination of invariant zeros. The number of null invariant zeros can be calculated by studying the infinite structure of BGD models. All these concepts are recalled in this part for LTI bond graph models.

2.1. Controllability/Observability

Usually, when studying the solvability conditions and stability conditions for various problems, the controllability and observability properties of the model must be first studied.

The controllability/observability properties have been first derived from a graphical approach using the causality concept in [16]. Controllability conditions on the bond graph representation are recalled.

Property 1: [16] A LTI bond graph model is controllable if and only if the two following conditions are verified:

- there is a causal path between each dynamical element and one of the input sources
- Each dynamical element can have a derivative causality assignment in the bond graph model with a preferential derivative causality assignment (with a possible duality of input sources)

2.2. Infinite structure

The infinite structure of multivariable linear models is characterized by different integer sets: two sets are recalled here. $\{n'_i\}$ is the set of infinite zero orders of the global model $\Sigma(C, A, B)$ and $\{n_i\}$ is the set of row infinite zero orders of the row sub-systems $\Sigma(c_i, A, B)$. The infinite structure is well defined in case of LTI models [6] with a transfer matrix representation or with a graphical representation (structured approach), [7], and can be easily extended to LTV models with the graphical approach.

The row infinite zero order for the row sub-system $\Sigma(c_i, A, B)$ is the integer n_i, which verifies condition $n_i = \min \left\{ k \mid c_i A^{(k-1)} B \neq 0 \right\}$. n_i is equal to the number of derivations of the output variable $y_i(t)$ necessary for at least one of the input variables to appear explicitly. The global infinite zero orders [8] are equal to the minimal number of derivations of each output variable necessary so that the input variables appear explicitly and independently in the equations.

The graphical procedure for the determination of the row and global infinite structures of a bond graph model is recalled.

Definition 1: The causal path length between an input source and an output detector in the bond graph model is equal to the number of dynamical elements met in the path.

Two paths are different if they have no dynamical element in common.

The order of the infinite zero for the row sub-system $\Sigma(c_i, A, B)$ is equal to the length of the shortest causal path between the ith output detector y_i and the set of input sources. The global infinite structure is defined with the concepts of different causal paths. The orders of the infinite zeros of a global invertible linear bond graph model are calculated according to equation (1), where L_k is the smallest sum of the lengths of the k different input-output causal paths.

$$\begin{cases} n'_1 = L_1 \\ n'_k = L_k - L_{k-1} \end{cases} \quad (1)$$

The study of the infinite structure of LTV bond graph models is quite similar to the LTI case.

2.3. Invariant zeros with graphic approach

The number of invariant zeros is determined by the infinite structure of BGI model.

Assumption 1 In section 2.3, it is supposed that LTI bond graph models are controllable, observable, invertible and square. The state matrix is invertible and the order of the model is n.

Proposition 1: The number of invariant zeros associated to the bond graph model is equal to $n - \sum n'_i$, where n is the number of state variables and $\{n'_i\}$ is the set of infinite zero orders of the global model.

For bond graph models, invariant zeros equal to zero can be directly deduced from the infinite structure of the BGD model. Some definitions are recalled, [2]. The row infinite structure and the global infinite structure in the BGD are defined with the two sets $\{n_{id}\}$ and $\{n'_{id}\}$

Property 2: The infinite zero order n_{id} associated to the ith output variable $y_i(t)$ is equal to the shortest causal path length between the output detector associated to the output variable $y_i(t)$ and the set of input sources in the BGD models.

Property 3: The set of infinite zero orders $\{n'_{id}\}$ is obtained with:

$$\begin{cases} n'_{1d} = L_{1d} \\ n'_{id} = L_{id} - L_{(i-1)d} \end{cases} \quad (2)$$

where L_{id} is equal to the smallest sum of i causal path lengths between i output detectors and i input sources (these paths must be different) in the BGD.

Property 4: The infinite zero order n_{id} of a LTI bond graph model is equal to the number of null invariant zeros associated to the output variable $y_i(t)$ and the number of null invariant zeros associated to the bond graph model is equal to $\Sigma n'_{id}$.

The previous properties are not always valid in the LTV case. The new procedure related to null invariant zeros is proposed in the sequel.

3. ALGEBRAIC APPROACH

For non linear models, variational models can be written with Kahler derivation. These new models are linear time varying (LTV) models. Non linear bond graph models are transformed in LTV bond graph models with some graphical procedures proposed in [1]. In case of the stability analysis, the finite structure must be studied, and in that case LTI extension to the LTV case is not so easy. One approach proposed by [9] for the study of linear systems and then extended to non linear systems for flatness analysis in [10] is the algebraic approach. Different classical problems such as the controllability/observability analysis [5], [12] or input-output decoupling problem [11] are proposed as a direct extension of the LTI case due to some properties of the bond graph representation. This approach is a good solution for studying classical control problems with the stability property on bond graph models. A simple example is first proposed and then the algebraic approach is recalled.

3.1. LTV example 1

Consider the bond graph model with integral causality of a LTI system, figure (1). According to Property 1, the first condition is satisfied.
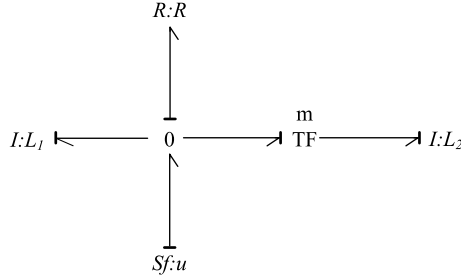

Fig. 1.   Bond graph model with integral causality

As shown in figure (2), the second condition for the bond graph model with derivative causality is not verified. So in the LTI case, this system is not controllable. However, one may get a different result if the system becomes a linear time varying model with for example parameter $m(t)$. In that case, the controllability matrix is of full rank. Property 1 gives in that case only sufficient conditions. From a structural point of view, the model is not structurally controllable if some linear relations can be written between the rows of matrix $[A\ B]$. These relations can be directly written from a graphical analysis on the bond graph model BGD.

In figure (1), the two state variables are x_{L_1} and x_{L_2}. In figure (2), one dynamical element has an integral causality assignment, thus a mathematical relation can be written between state variables, $\dot{x}_{L_1} - m(t)\dot{x}_{L_2} = 0$ (same relation obtained between the rows of matrix $[A(t) \vdots B(t)]$ when applying a structural approach). According to the properties of this equation, the controllability property can be pointed out. The algebraic approach must be applied. If parameter $m = m(t)$, equation $\dot{x}_{L_1} - m(t)\dot{x}_{L_2} = 0$ is not associated to a torsion element, thus this model is controllable. If $m(t) = m$, $\dot{x}_{L_1} - m\dot{x}_{L_2} = 0$ is equivalent to $\delta(x_{L_1} - mx_{L_2}) = 0$, which is the equation of a torsion submodule and in that case the model is not controllable.
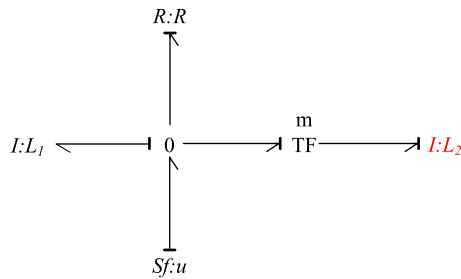
Fig. 2.   Bond graph model with derivative causality

The graphical procedure for the study of the observability property is very closed to the one defined in case of the controllability property for LTI models [16]. From a structural point of view, the model is not structurally observable

if some linear relations can be written between the columns of matrix $[C^t(t)\ A^t(t)]^t$. These relations cannot be written directly from a graphical analysis on the bond graph model. In that case, the concept of duality [15] can be applied.

3.2. Module and linear systems

The classical state representation, for a linear time varying system is given by the Kalman form (3), with $x \in \Re^n$, $u \in \Re^m$ and $y \in \Re^p$. Models described in equation (3) are denoted as $\Sigma(C(t), A(t), B(t))$.

$$\begin{cases} \dot{x}(t) = A(t)x(t) + B(t)u(t) \\ y(t) = C(t)x(t) \end{cases} \tag{3}$$

In Fliess's theoretic approach the systems are the modules. The definitions in this section are the same with the one introduced in [9]. A linear system $\sum$ is a finitely generated left **R**-module.

Definition 2: A(linear) dynamics $\mathscr{D}$ is a system in which a finite set $u = \{u_1, u_2, ..., u_m\}$ of input variables is such that the quotient module $\mathscr{D}/[u]_{\mathbf{R}}$ is torsion. It means that any element in $\mathscr{D}$ can be calculated from u by a linear differential equation.

Equation (3) is equivalent to (4) in a module framework representation, if the output variable is not considered.

$$\begin{pmatrix} I\delta - A(t) & -B(t) \end{pmatrix} \begin{pmatrix} x \\ u \end{pmatrix} = R(\delta, t) \begin{pmatrix} x \\ u \end{pmatrix} = 0 \tag{4}$$

The entries of the matrix $R(\delta, t)$ belong to the non commutative ring **R** and $\begin{bmatrix} x & u \end{bmatrix}_{\mathbf{R}} \in \Omega$ (here x and u are considered as row vectors, Ω is the **R**-module and $\delta = d/dt$).

The structural properties of (4) are then translated to a module framework. The controllability property is thus directly deduced from the Jacobson form of matrix $R(\delta, t)$, or from the torsion module associated to the state equation, if this torsion element exits.

3.3. Controllability of LTV systems

From an algebraic point of view, controllability is related to the torsion submodule $\mathscr{T}(\Sigma)$ which is the module of input decoupling zeros, where $\mathscr{T}(\Sigma)$ is the torsion submodule of system module Σ.

For the controllability analysis [12], in the LTV case, the second condition in property (1) is not a necessary condition. Some differential equations can be written between state variables: if these equations define a torsion element then the model is not controllable, otherwise it is controllable.

Property 5: [9] A linear system is controllable iff it is a free **R**-module, i.e. the torsion submodule is trivial $\mathscr{T}(\Sigma)=0$.

Definition 3: [9] The controllability matrix of LTV systems can be written as the form: $\mathscr{C}(\delta, t) = \begin{bmatrix} B(t), (A(t) - \delta I_n)B(t), ..., (A(t) - \delta I_n)^{n-1}B(t) \end{bmatrix}$. LTV systems are controllable iff $rk(\mathscr{C}(\delta, t)) = n$.

Example 1: (continued) For the system with a time varying transformer $m(t)$ shown in figure (1), the state matrix

$A(t)$ and input matrix $B(t)$ are equal to:

$$A(t) = \begin{bmatrix} -\frac{R}{I_1} & -\frac{R}{m(t)I_2} \\ -\frac{m(t)R}{I_1} & -\frac{R}{I_2} \end{bmatrix} \quad B(t) = \begin{bmatrix} 1 \\ m(t) \end{bmatrix}$$

The controllability matrix is

$$\mathscr{C}(t,\delta) = [B(t),(A(t)-\delta I)B(t)] = \begin{bmatrix} B(t), A(t)B(t) - \frac{dB(t)}{dt} \end{bmatrix}$$

$$= \begin{bmatrix} 1 & -\frac{R}{I_1} - \frac{R}{I_2} \\ m(t) & -\frac{m(t)R}{I_1} - \frac{m(t)R}{I_2} - \frac{dm(t)}{dt} \end{bmatrix}$$

Compared with the LTI case, the controllability matrix has an additional term $B'(t)$. In the LTI case, the rank of matrix $\mathscr{C}$ is smaller than system order. The LTI system is noncontrollable. With the term $\frac{dB(t)}{dt}$, $rk(\mathscr{C}(t,\delta)) = 2$ means that the LTV system is controllable.

3.4. Invariant zeros with algebraic approach

Considering a LTV system $\Sigma(C(t),A(t),B(t))$ which is a finitely generated module M over the ring $\mathbf{R} = K[\delta]$. The module $\mathscr{T}(M/[y]_\mathbf{R})$ is torsion and is called the module of invariant zeros of Σ. Let z be a generator of $\mathbf{R} = K[\delta]$. There exists $Z(\delta) \in \mathbf{R}$ such that $Z(\delta)z = 0$ and let $\bar{K}$ an extension of K over which a set of zeros of $Z(\delta)$ can be derived.

Definition 4: [4] $M_{iz} = \mathscr{T}(M/[y]_\mathbf{R})$ is the module of the invariant zeros of the LTV system. The invariant zeros of the LTV system are the conjugacy classes of the elements of a full set of Smith zeros of M_{iz}.

For the LTV system represented by equation (3), the module of the invariant zeros M_{iz} is defined by $P(\delta,t)\begin{bmatrix} \bar{x} \\ \bar{u} \end{bmatrix}$, where

$P(\delta,t) = \begin{bmatrix} \delta I - A(t) & -B(t) \\ C(t) & 0 \end{bmatrix}$ is the system's matrix and $\bar{x},\bar{u}$ are the images of x,u in module M_{iz}. $P(\delta)$ is singular with certain $\delta = \alpha_i$. With these values of δ, for an input $u(t) = u_0 e^{\alpha t}, t \geq 0$, there exist initial state variables x_0 such that the output is null:$y \equiv 0, t \geq 0$.

4. MAIN CONTRIBUTIONS

In case of LTV models, the controllability/observability matrices are quite difficult to derive. From a structural point of view, the bond graph approach is simple if the algebraic and structural approaches are combined (see previous sections). For the study of the finite structure and particularly for the invariant zeros which play an essential role in the stability property of the controlled model, solutions are proposed in [17], with the algebraic approach combined with the bicausality assignment on the bond graph model. For classical control problems, invariant zeros must be studied for global models (all input and output variables) and also for row submodels (only one output variable). In [17], it is proved that some uncontrollable parts of the BGB models must be compared from an algebraic point of view (torsion submodules). In this paper the focus is on the invariant zeros with a zero value. For LTI models, the infinite structure of the BGD models is directly related to these particular invariant zeros. For LTV models, conditions are only sufficient and

a quite similar extension to the study of the controllability/observability is proposed for the study of null invariant zeros.

4.1. Invariant zero: bond graph procedure

Assumption 2 In section 4, it is supposed that LTV bond graph models are controllable, observable, invertible and square. The state matrices are invertible and the models order is n.

1) Bond graph procedure: A bond graph interpretation of the procedure for the determination of invariant zeros of LTV systems is implemented in figure (3). Combined with the algebraic method, one can finally get the torsion module concerning the invariant zeros structure. In the first step, BGI model can verify the existence of invariant zeros by proposition (1). Because invariant zeros are the poles of inverse system in the general case, BGB model is utilized to get the inverse model. In the third step, the procedure introduced in [17] is required to get equations of four kinds of elements in BGB model:

- Output detectors (variables are set to a zero value)
- Dynamical elements with a derivative causality
- Input sources
- Dynamical elements with an integral causality

Mathematical relations are written for the unknown variables associated to elements of the BGB model (effort or flow variable depending on the causality assignment). From this set of mathematical relations, the torsion module can be highlighted and some polynomials can be written from dynamical elements with an integral causality. For the last step, some state variables dependent on u are substituted by u to get equations of torsion module $\mathscr{T}(\Sigma/[y]_\mathbf{R})$. And the torsion module has the form:

$$P(\delta)\xi = 0, \quad P(\delta) = \delta^n + \sum_{i=1}^{n} a_i \delta^{n-i} \tag{5}$$

where $P(\delta)$ is a differential polynomial and equation (5) is the generate equation of torsion module $\mathscr{T}(\Sigma/[y]_\mathbf{R})$. ξ is a generator of torsion module, in this case, it is equal to u.

A new proposition of null invariant zeros of LTV systems is given.

Proposition 2: The null invariant zeros for a row LTV bond graph model with Assumption 2 can be derived from BGD model. Considering a LTV bond graph model BGD with an input-output causal path length greater or equal to 1, there exist null invariant zeros if there is no time varying dynamic element in the causal path. Otherwise, in the case of existence of time varying elements in causal path, there still exist null invariant zeros if the differential polynomial equation of torsion module $\mathscr{T}(\Sigma/[y]_\mathbf{R})$ has the form: $G(\delta)u = 0 \rightarrow G'(\delta)\delta^{\tilde{n}_{id}}u = 0$, where $G^{(')}(\delta)$ are polynomials of δ and $\delta^{\tilde{n}_{id}}$ is the right factor of $G(\delta)$. $G(\delta)$ is the input-output causal path gain and $\tilde{n}_{id}$ is the number of null invariant zeros. $G'(\delta) \in \mathbf{R}$ has the form $\sum_{i=0}^{n} a_i \delta^i, a_i \in K$, where $a_0 \neq 0$.
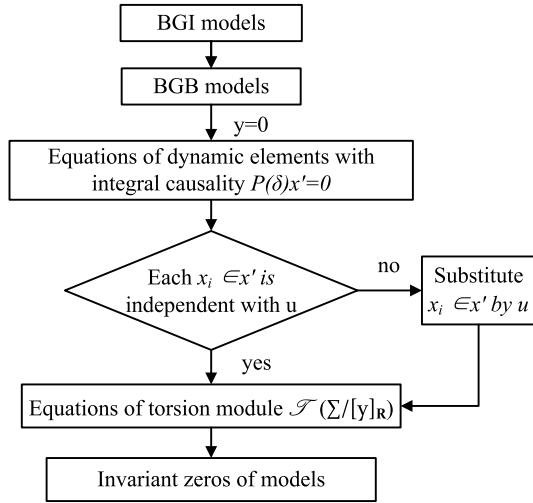
Fig. 3. Procedure for invariant zeros of linear systems

4.2. Examples

In this section, the invariant zero structure of a SISO LTV RLC circuit is studied by algebraic and bond graph approaches. The bond graph model of circuit is shown by figure (4). Table (I) gives numerical values of LTI system components.

TABLE I
NUMERICAL VALUES OF RLC CIRCUIT COMPONENTS

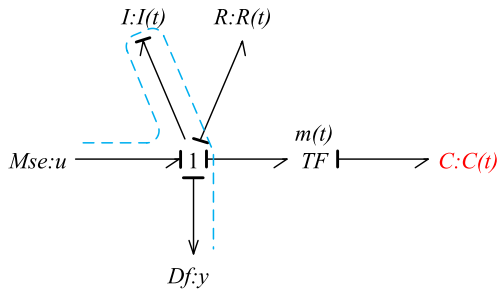| Input u	element I	element R	element TF	element C
1 V	1 H	1 Ω	2	1 F

Fig. 4. Bond graph model with integral causality of RLC circuit

The infinite structure of the BGI model, figure (4), is defined as $n = 1$ (causal path $Df \rightarrow I(t) \rightarrow MSe$). By proposition (1), there exist one invariant zero.

Figure (5) gives the bond graph model with derivative causality. The model is controllable. There is an input-output causal path $(Df \rightarrow C(t) \rightarrow MSe)$. If the system is time invariant, there is a null invariant zero. In case of a LTV system, the causal path gain must be studied. By proposition (3), the input-output causal path gain is equal to $\frac{1}{m(t)} \delta \frac{C(t)}{m(t)}$. If the coefficient $\frac{C(t)}{m(t)}$ is a constant, i.e. $C(t)$ and $m(t)$ are proportional, there exist a right root of differential polynomial which is a null invariant zero, otherwise, the value of the invariant zero is not equal to zero. This result

will be confirmed by calculation and with a simulation in the sequel.
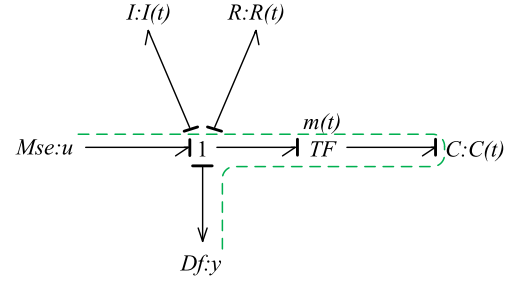
Fig. 5. Bond graph model with derivative causality of RLC circuit

Now the procedure for the determination of the torsion module with the bond graph model with bicausality (BGB) defined in [17] is used. In this simple example, it is not useful, but it will prove the first conclusion obtained from the study of the input output causal path and above all, it is a simple way for the study of torsion submodules associated to any kind of invariant zeros.

In figure (6), the bicausal path is drawn between the input source $Mse : u$ and the output detector $Df : y$. In this simple example, the element $C(t)$ is associated to the torsion module $\mathscr{T}(\Sigma/[y]_{\mathbf{R}})$. The torsion module is the non controllable part of inverse system. The element $C(t)$ (more precisely the state variable) is not controllable, because it is not reachable when the bicausal path is eliminated. Now the procedure to derive the torsion module is given.
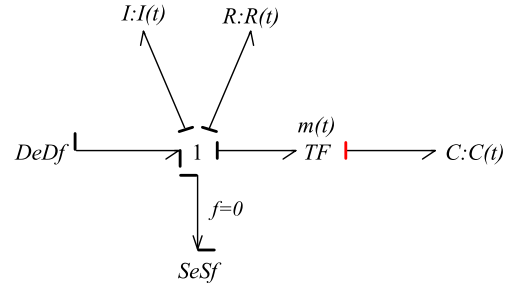
Fig. 6. Bond graph model with bicausality of RLC circuit

Step 1: output variable

For the output detector, the flow at the 1 junction is equal to zero. The equation of output variable is $y = \frac{1}{I(t)} x_1$. One relation is thus rewritten: $y = 0$, thus $x_1 = \dot{x}_1 = 0$.

Step 2: element with derivative causality

Element $I : I(t) \longrightarrow f_{I_1} = \frac{p_{I_1}}{I_1} = \frac{x_1}{I_1} = 0$

Step 3: input source

Source $MSe : u \longrightarrow u = \dot{x}_1 + e_R + \frac{m(t)}{C(t)} x_2 = \frac{m(t)}{C(t)} x_2$

Step 4: element with an integral causality

Element $C : C(t) \longrightarrow \dot{x}_2 = f_{C_5} = 0$

Step 5: expression of torsion module

In this step, previous equations must be used. According to the length of the input-output causal path, there is only one invariant zero and the degree of the polynomial equation is equal to 1. Equations obtained in step 4 must be redefined according to the previous one. In step 4, one has $\delta x_2 = 0$. Substituting x_2 by u, it comes

$$\delta \frac{C(t)}{m(t)} u = \left(\frac{C(t)}{m(t)} \delta + \left(\frac{C(t)}{m(t)} \right)' \right) u = 0 \qquad (6)$$

which is the equation of torsion module $\mathscr{T}(\Sigma/[y]_{\mathbf{R}})$. According to Section 3.4, the invariant zero of system is

$$\Delta_{\mathbb{C}(t)} \left(\frac{C(t)m'(t) - C'(t)m(t)}{C(t)m(t)} \right) \qquad (7)$$

which is the conjugacy class of a full set of Smith zeros of M_{iz}. So if parameters $C(t)$ and $m(t)$ are constant or proportional, there exist a null invariant zero because of $\delta u = 0$, otherwise the value of the invariant zero is different of zero, which has already be proved.

4.3. Several cases

In this section, several cases related to different time varying parameters of the studied system are considered. With different expressions of input source, one can verify if the invariant zero is null. First, the case without time varying element in causal path of BGD model is studied. Simulation results show that there is no influence for the existence of null invariant zero. Then, with time varying elements in the causal path, three situations are proposed. Firstly, there exists only one time varying element $C(t)$ or $m(t)$, then parameters $C(t)$ and $m(t)$ are not proportional. Finally, the case with one time varying element $I(t)$ outside the causal path and one time varying element $C(t)$ in the path will be studied. In each case, the Jacobson form of the system matrix $P(\delta)$ is given to illustrate the simulation result by algebraic method.

1. Without time varying elements in the causal path

The existence of null invariant zeros can be verified if there is no time varying elements in the causal path. Elements I and R with time varying parameters are considered respectively.

1). Inertial element I is time varying

Let $I(t) = t^2 + 1$, the invariant zero of system is $\Delta_{\mathbb{C}(t)}(0)$ according to equation (7). So the system has a null invariant zero. The curve of output Df is shown in figure (7) and the output variable is equal to 0 in the steady state part of the curve.

2). Resistive element R is time varying

Let $R(t) = t + 1$, the invariant zero of system is $\Delta_{\mathbb{C}(t)}(0)$ according to equation (7). So the system has a null invariant zero. The curve of output Df is shown in figure (8).

3). Two element I and R are simultaneously time varying

Let $I(t) = t^2 + 1$ and $R(t) = t + 1$, the invariant zero of system is $\Delta_{\mathbb{C}(t)}(0)$ according to equation (7). So the system has a null invariant zero. The curve of output Df is shown in figure (9).
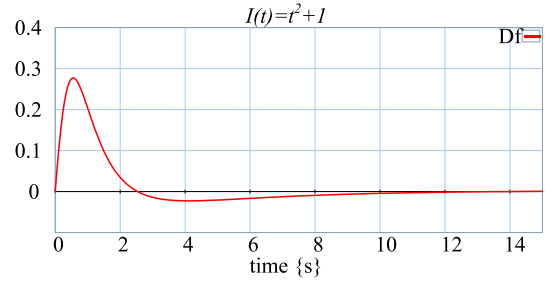
Fig. 7. The output Df curve with $I = t^2 + 1$
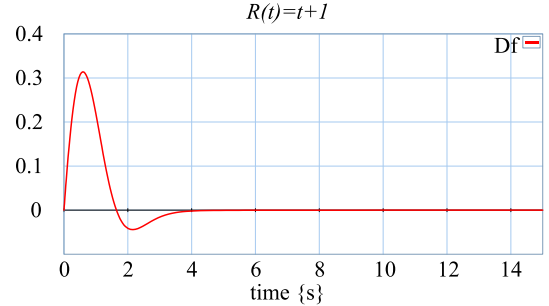
Fig. 8. The output Df curve with $R = t + 1$
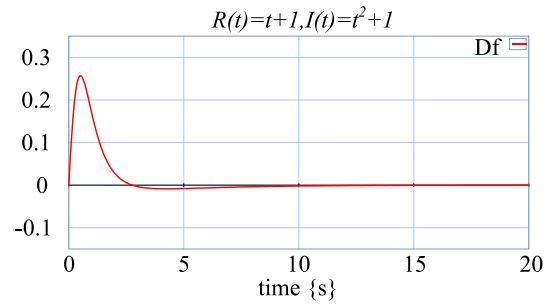
Fig. 9. The output Df curve with $R = t + 1, I = t^2 + 1$

From these three situations, one can verify the proposition (3): if there is no element with time varying parameter in causal path of BGD model, the LTV system has null invariant zero if the causal path length is equal to 1.

2. Time varying elements in the causal path

There are two elements C and TF in the causal path of BGD model. Four situations are considered here.

1). The transformer element TF is time varying

Let $m(t) = \frac{1}{t+1}$, the invariant zero of system is $\Delta_{\mathbb{C}(t)} \left(-\frac{1}{t+1} \right)$ according to equation (7). So there is no null invariant zero for the system. The curve of output Df is shown in figure (10).

If $m(t) = sint(t) + 2$, the curve of output Df which is not stable is shown in figure (10). As for the instability of the curve, one should calculate if there is(are) unstable pole(s) of system. From algebraic approach, poles are related to the module $\Sigma/[u]_{\mathbf{R}}$. Poles of system can be derived from the Jacobson form of matrix $D(\delta) = (\delta I - A)$. For poles and stability property of LTV systems, [14] is recommended.

2). The element C is time varying
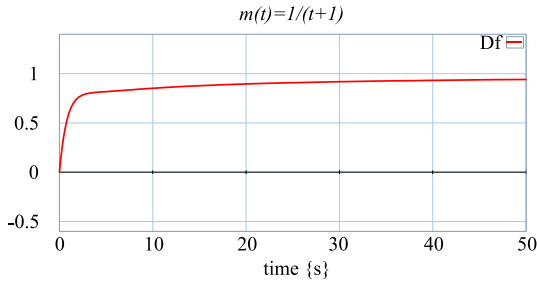
Let $C(t) = t + 1$, the invariant zero of system is

211
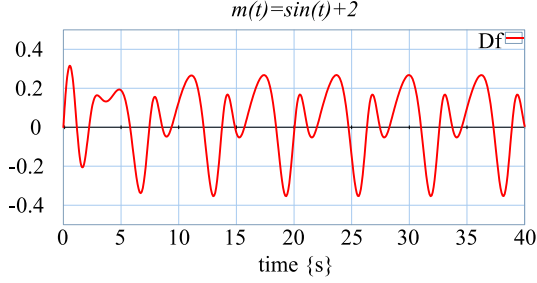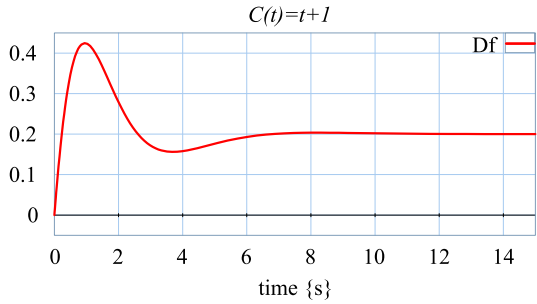
Fig. 10. The output Df curve with $m(t) = \frac{1}{t+1}$

Fig. 11. $m(t) = \sin(t) + 2$

$\Delta_{\mathbb{C}(t)}\left(-\frac{1}{t+1}\right)$ according to equation (7). So there is no null invariant zero for the system. The curve of output Df is shown in figure (12).

Fig. 12. The output Df curve with $C = t + 1$

3). Two elements TF and C are time varying and proportional

Let $C(t) = m(t) = t^2 + 1$, the invariant zero of system is $\Delta_{\mathbb{C}(t)}(0)$ according to equation (7). So there is a null invariant zero for the system. The curve of output Df is shown in figure (13).

4). Two elements TF and C are time varying and non proportional

Let $m(t) = t + 2, C(t) = t^2 + 1$, the invariant zero of system is $\Delta_{\mathbb{C}(t)}\left(\frac{-t^2-4t+1}{(t^2+1)(t+2)}\right)$ according to equation (7). So there is no null invariant zero for the system. The curve of output Df is shown in figure (14).

From these four aforementioned examples, one can find that there exists invariant zero in spite of the existence of time varying elements in the causal path of BGD model. It means that the rule to determine null invariant zero from the BGD model must be completed with a study of the BGB
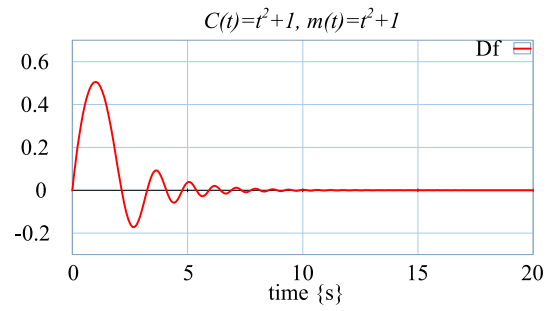
Fig. 13. The output Df curve with $C(t) = m(t) = t^2 + 1$
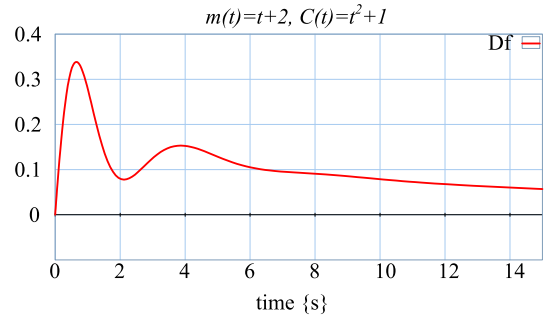
Fig. 14. The output Df curve with $m(t) = t + 2, C(t) = t^2 + 1$

model and an algebraic criteria. The invariant polynomials of system matrices give the intrinsic implementation of the existence of null invariant zeros.

3. With time varying elements

The model with two time varying elements *I* and *C* is considered here. One is in the input-output causal path and another is not. Let $I = t + 1, C = t^2 + 1$, the invariant zero of system is $\Delta_{\mathbb{C}(t)}\left(\frac{-2t}{t^2+1}\right)$ according to equation (7). So there is no null invariant zero. Figure (15) shows the output Df curve.
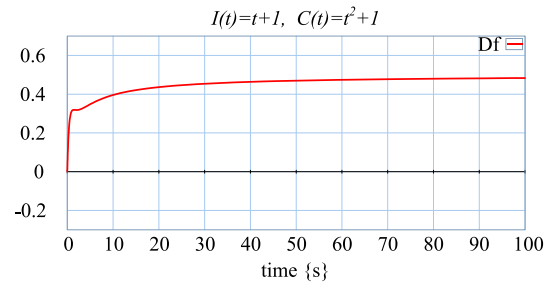
Fig. 15. The output Df curve with $I = t + 1, C = t^2 + 1$

According to the simulation results of aforementioned cases, one can find that there is always one invariant zero. Only the cases where elements $C(t)$ and $m(t)$ are time varying separately or simultaneously but nonproportional have one non null invariant zero. In other cases, there is always a null invariant zero which can not be affected by the time varying parameters of elements. These results verify the algebraic procedure in the article.

5. CONCLUSION

In this paper, invariant zeros structure of LTV systems is studied. The LTI bond graph criteria is extended to the time varying case. The rule for determining systems null invariant zeros is developed by using bicausal bond graph models and torsion module notion. The algebraic calculations supported by Maple show the consensus. 20-sim verified the counterparts from simulations.

REFERENCES

1. A. Achir and C. Sueur. Non commutative ring bond graphs: application to flatness. *International Conference on Bond Graph Modelling ICBGM, WMC SCS, New Orleans (USA), F. E. Cellier, J.J Granda, editors, Simulation Series, January 23-27, 2005*, 1 (37):59–64, 2005.
2. J. M. Bertrand, C. Sueur, and G. Dauphin-Tanguy. Bond graph for modeling and control: structural analysis tools for the design of input-output decoupling state feedbacks. *Proceedings of International Conference on Bond Graph Modeling and Simulation, Phoenix*, pages 103–108, 1997.
3. H. Bourles and B. Marinescu. Poles and zeros at infinity of linear time-varying systems. *IEEE Transactions on Automatic Control*, 44:1981–1985, 1999.
4. H. Bourlès and B. Marinescu. *Linear Time-Varying Systems: Algebraic-Analytic Approach*, volume 410. Springer, 2011.
5. Z. Chalh, C. Andaloussi, and C. Sueur. Structural properties of bond graph rings. *ICBGM*, 38:20–26, 2007.
6. J. M. Dion and C. Commault. Smith-mcmillan factorizations at infinity of rational matrix functions and their control interpretation. *System & Control Letters*, 1:312–320, 1982.
7. J. M. Dion and C. Commault. Feedback decoupling of structured systems. *IEEE Transactions on Automatic Control*, 38:1132–1135, 1993.
8. P. L. Falb and W. A. Wolovich. On the decoupling of multivariable systems. *Preprints JACC, Philadelphia*, pages 791–796, 1967.
9. M. Fliess. Some basic structural properties of generalized linear systems. *Systems & Control Letters*, 15:391–396, 1990.
10. M. Fliess, J. Lévine, P. Martin, and P. Rouchon. Flatness and defect of nonlinear systems: introductory theory and examples. *Int. Journal Control*, 61:1327–1361, 1995.
11. S. Lichiardopol and C. Sueur. Decoupling of linear time-varying systems with a bond graph approach. *ECMS 2006, Bonn, Germany*, 2006.
12. S. Lichiardopol and C. Sueur. Duality in system analysis for bond graph models. *Journal of the Franklin Institute*, 347:377–414, 2010.
13. B. Malgrange. Systemes différentiels à coefficients constants. *Séminaire Bourbaki*, 246:63, 1962.
14. B. Marinescu and H. Bourlès. An intrinsic algebraic setting for poles and zeros of linear time-varying systems. *Systems and Control Letters*, 58(4):248 – 253, 2009.
15. J. Rudolph. Duality in time-varying linear systems: A module theoretic approach. *Linear Algebra and its Applications*, 245:83–106, 1996.
16. C. Sueur and G. Dauphin-Tanguy. Bond-graph approach for structural analysis of mimo linear systems. *Journal of the Franklin Institute*, 328:55–70, 1991.
17. D. Yang, N. Palitzyne, and C. Sueur. Finite structure of ltv systems modelled by bond graph: an algebraic approach. *4th IFAC Symposium on System Structure and Control, Ancona, Italy*, October 15-17, 2010.

COMPONENT FAULT DETECTION AND ISOLATION COMPARISON BETWEEN BOND GRAPH AND ALGEBRAIC APPROACH.

Samir BENMOUSSA, Belkacem OULD BOUAMAMA, Rochdi MERZOUKI

LAGIS, FRE CNRS 3303, Polytech'Lille, Avenue Paul Langevin,
59655 Villeneuve d'Ascq, FRANCE
Samir.benmoussa@polytech-lille.fr

ABSTRACT

This paper makes a comparison of component fault detection and isolation between an algebraic approach and a bond graph one. The conditions of component fault detection and isolation are viewed in algebraic approach as an observation problem of the fault with respect to the input and the output. In Bond graph approach, these conditions are performed by analyzing the causal paths from faults to outputs using the notion of bicausality. It is shown that the use of bicausal bond graph helps to integrate many mathematical approaches particularly the algebraic one. The component fault detection and isolation performed from bond graph is much simpler as compared to the algebraic approaches in which an analytical model is needed and complex computations are performed to determine the diagnosability conditions. An illustrative example is given to show the efficiency and the simplicity of the bond graph approach compared to the algebraic one.

Keywords: Fault detection and isolation, Bond graph, Algebraic approach.

1. INTRODUCTION

Fault Detection and Isolation (FDI) has become an important tool in the ingredients of a modern automatic system. Its significance is based on enhancement in terms of safety, reliability, dependability and operating costs of the plant. And for that, FDI has been a widely exploited research topic in the recent years, and several methods have been developed among them ones based on the model of the system such as: parity equations (Gertler, 1997), observer model-based (Patton and Chen, 1997), and analytical redundancy relations (Staroswiecki and Comet-Varga, 2001).

Several FDI model-based approaches can be found in the literature. Among them, the algebraic approach (Fliess and Join, 2003; Fliess and Join, 2004; Cruz-Victoria, Martinez-Guerra and Rincon-Pasaye, 2008) consists of the ability of the detection, the identification, and the estimation of the fault variable. It is viewed as an observation problem of the latter with respect to the input and the output variables. These approaches are applied to linear and a class of nonlinear systems to detect sensor and actuator faults.

Other approaches are based on structured and graphical models. They intend to create a graph that describes the mathematical model of the system. These approaches are based on a digraph $G(S;A)$, where nodes S represent the state, the input, and the measurement output variables, and edges A are the interaction between these nodes. In (Commault, Dion, and Agha, 2008), the authors study the sensors location problem for internal faults in term of separator in the associated graph of the structured system. This separator gives the necessary and sufficient sensors to be added. Fault detection and isolation based on digraph has limits that it concerns only actuator and sensor faults.

The Bond Graph (BG) which is also a graph, can be an alternative for plant fault detection and isolation, since the nodes S represent not only state, input, and output variables but also the physical components of the system. It describes the power transfer between the passive and active components of multi-physical systems. It is the interface between the physical system and the mathematical model of the last.

Several problems have been solved structurally using this graphical approach, such as: observability and controllability (Sueur and Dauphin-Tanguy, 1991), system inversion (Ngwompo and Gawthrop, 1999), and FDI (El-Osta, Ould Bouamama and Sueur, 2006; Samantary and Ghosal, 2008). FDI BG model-based uses the analytical redundancy relations (ARR) which are generated from the BG model of the system using its causal and structural proprieties. A fault signature matrix (FSM) is defined to study the fault monitorability and isolability. This matrix depends mainly on the number of sensors, and the unknown variables to be eliminated, which is complex and difficult to apply for a nonlinear system.

This paper makes a comparison of component fault detection and isolation between an algebraic approach and a BG one, in which plant faults are modeled as additional modulated inputs to the system. The conditions of the detectability and isolability of plant faults are performed on the BG model using the notion of bicausality and disjoint causal path analysis. In this work, the detectability and the isolability of the fault are performed directly on the BG model without computations, which are usually complex. Since BG model is a graphical representation of the physical

system, the concept of fault is much more physical than mathematical, contrary to the algebraic approach.

The paper is structured as follows: in Section 2 the problem statement. Conditions of detectability and isolability by algebraic approach are given in Section 3. And by BG approach in section 4. An illustrative example is introduced in Section 5. The paper ends with general synthesis and remarks.

2. PROBLEM STATEMENT

Let us consider a class of linear time invariant systems described by the following state space format as:

$$\begin{cases} \dot{x}(t) = Ax(t) + B(u) \\ y(t) = Cx(t) \end{cases} \tag{1}$$

where $x(t) \in R^n$ is the state vector, $u(t) \in R^m$ is the input vector, $y(t) \in R^p$ is the output vector, and A, B, C are matrices of appropriate dimensions.

The vector $F \in R^q$ such that $q \leq p$, represent internal components failure in the plant, are introduced in the state space format as:

$$\begin{cases} \dot{x}(t) = (A+F)x(t) + B(u) \\ y(t) = Cx(t) \end{cases} \tag{2}$$

Note that faults on inputs are not considered in this paper and we assume that all sensors are fault free. The system of Eq.(2) can be rewritten as:

$$\begin{cases} \dot{x}(t) = Ax(t) + B(u) + MF(x,t) \\ y(t) = Cx(t) \end{cases} \tag{3}$$

where $M \in R^{n \times q}$ is a known matrix, and F is an unknown vector that needs to be detected.

In the next sections, the conditions of the detectability and isolability of the fault F are given.

3. FAULT DETECTION AND ISOLATION BASED ON AN ALGEBRAIC APPROACH

3.1. Basics in algebraic approach

System Λ^{pert} is defined as a finitely generated free $k[s]$-module, where $k[s]$ is a commutative principal ideal domain of linear differential operators of the form $\sum_{finite} c_v s^v, c_v \in k$; s is the usual symbol of derivation, and k is the field of real or complex numbers.

In Λ^{pert} two finite subsets are distinguished, the fault variable F and the perturbation variable π, which do not 'interact', i.e., $span_{k[s]}(F) \cap span_{k[s]}(\pi) = \{0\}$.

The nominal system is defined by:
$$\Lambda = \Lambda^{pert} / span_{k[s]}(\pi).$$

Definition: *An input output system* is a linear system, Λ^{pert} equipped with an input u and an output y, such that:

- The input of the linear system Λ^{pert} is a finite sequence $u = (u_i), 1 \leq i \leq m$ of elements of Λ^{pert} such that $\Lambda^{pert} / [u]_{k[s]}$ is torsion, the input u is assumed independent.
- The output of the linear system Λ^{pert} is a finite sequence $y = (y_i), 1 \leq i \leq p$ of elements of Λ^{pert}.

Assumptions: the following propreties are assumed to be satisfied:

- $span_{k[s]}(u) span_{k[s]}(\pi) = \{0\}$,
- $span_{k[s]}(u) span_{k[s]}(F) = \{0\}$.

This means that the control variable u does not interact with the perturbation and the fault variables.

3.2. Fault detection and isolation

The system described by Eq.(3) is an input output system Λ^{pert}.

Theorem 1: The system of Eq.(3) is observable (in the sense that the state is observable with respect to u and y), then it is diagnosable if, and only if, F is observable with respect to u, y, and x (Diop and Martinez-Guerra, 2001).

Definition (Algebraic Detectability): a fault F is said to be detectable if, it is observable over u and y (Fliess and Join, 2003).

Definition (Algebraic Isolability): any fault variable in F is said to be isolable if, and only if, there exists a system of parity equation (Fliess and Join, 2003)

$$M \begin{pmatrix} F_1 \\ \vdots \\ F_q \end{pmatrix} = Q \begin{pmatrix} u_1 \\ \vdots \\ u_m \end{pmatrix} + S \begin{pmatrix} y_1 \\ \vdots \\ y_p \end{pmatrix} \tag{4}$$

where $M \in k[s]^{q \times q}, Q \in k[s]^{q \times m}, S \in k[s]^{q \times p}, \det P \neq 0$.

In other words, it is required that:
- The system must be observable: the states of the system can be expressed as a function of outputs and their derivatives,
- Each fault variable has to be written under a polynomial equation format F_i and finitely many time derivatives of u and y with coefficients in $k[s]$.

$$\varphi(F_i, u, \dot{u}, ..., y, \dot{y}, ...) = 0 \tag{5}$$

4. FAULT DETECTION AND ISOLATION BASED ON BICAUSAL BOND GRAPH

4.1. Graph and Bond Graph

A graph theory approach is used to study and analyze structured systems which are independent of the system parameter values. This approach requires a low computation which allows dealing with large scale systems. The existing contributions related to the graph analysis proved that observability, controllability, input-output decoupling ... etc, can be simply deduced from the structural properties of the graph. There are different graphical methods used in the literature: Digraph (Dion, 2003), signed digraph (Maurya, Rengaswamy, and Venkatasubramanian, 2004), bipartite graph (Blanke, Kinnaert, Lunze, and Staroswiecki, 2003), and BG (Samantary, and Ghoshal, 2008).

Definition: The digraph, denoted by $G(S;A)$, is deduced from state space equations. It is composed by a set of nodes (S), $S = \{U \cup Y \cup X\}$ which corresponds to the system inputs, outputs, and states. The interactions between these nodes are represented by directed edges (A).

Definition: A graph $G(S;A)$ is bipartite, if its vertices can be partitioned into two disjoint subsets Z (set of variables that defines the dynamic behavior of the system), and C (set of equations that defines the relations among the variable set), $S = \{C \cup Z\}$ .The relations between these two subsets are represented by edges (A).

Definition: The BG which is also a graph $G(S;A)$, is a unified graphical language for multi-physical domains. Unlike the others graphs mentioned above, the nodes S represent physical components, subsystems, and other basic elements called junctions. While the edges A, called power bonds represent the power exchanged between nodes.
In BG, the possible set of components $S = \{R \cup C \cup I \cup TF \cup GY \cup Se \cup Sf \cup De \cup Df \cup J\}$.
The R- element represents passive energy dissipation phenomena, while C, and I model the passive energy storage elements. (Se), and (Sf) are the sources of effort and flow, respectively. Sensors are represented by flow (Df), and effort (De) detectors. Finally, J (which can be a 0 or a 1 junction), is used to connect the elements having the same effort (1-junction), or flow (0-junction). The conservation of energy laws are obtained from the latter. TF, and GY are used to represent transformers and gyrators, respectively.
The difference between BG and the other graphical approaches is that the former is directly generated from the physical system, and not from state space equations. In addition, from the BG model state space equations can be generated manually using systematic way (Mukherjee, and Karmakar, 2000) or automatically from dedicated software such as *Symbol2000* (Ould-Bouamama, Medjaher, Samantaray, and Dauphin-

Tanguy, 2005) or *20-sim* (Twentesim, 1996). Furthermore, system components are clearly represented in the BG model because of its graphical architecture.

4.2. Multiplicative fault modeling on BG

The component fault is modeled using bond graph as an additional modulated input (*MSe* for 1- junction, *MSf* for 0-junction), placed at the same junction of the component element. Let consider as an illustrative example, a resistive (R) element in resistive causality with the following characteristic equation

$$e = R.f \tag{6}$$

If the element R is faulty, then an additional value R_f is added to the nominal value R_n, Fig.(1)-(a), so Eq.(6) can be rewritten as

$$e = R_n.f + R_f.f = e_n + e_d \tag{7}$$

where e_d is the effort brought by the fault, which is unknown time function, can be considered as an additional modulated input added to the 1-junction as given in Fig.1-(b).
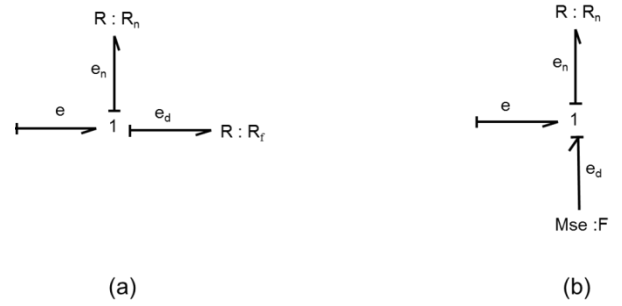
Figure 1 : (a) The BG interpretation of Eq.(7); (b) Component fault modeling in BG.

4.3. Bicausal Bond Graph

The Bicausal BG is introduced to study control problems such as system inversion, state estimation, and unknown parameter estimation (Gawthrop, 2000). It overcomes the assignment statements that cannot be derived from the constraint equations of a so-called 'unicausal' bond graph model, Fig.2-(a), which implies two assignment statements :

$$e_1 := e_2; \quad f_2 := f_1 \tag{8}$$

the effort e_2 and the flow f_1 are used to determine the effort e_1 and the flow f_2 respectively, and they have an opposite direction. Contrary to 'unicausal' BG, bicausal one imposes the same direction to the effort and the flow, Fig.2-(b), which implies two assignment statements

$$e_2 := e_1; \quad f_2 := f_1 \tag{9}$$

Causal half strokes indicate the fixed or known variables of the bond, and determine the right hand side of the assignment statements form.
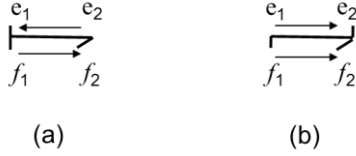
Figure 2 : (a) A bond with causality; (b) a bicausal bond.

Bicausal introduces also some additional BG elements (Ngwompo and Gawthrop, 1999), among which SS (Source-Sensor), *AE* (Amplifier of effort), and *AF* (Amplifier of flow). Table.1 gives causality assignment for Source-Sensor element.

Table 1: Source / Sensor causality assignement

'unicausal stroke'	SS element type	'Bicausal stroke'	SS element type
SS	Flow source/ Effort sensor	SS	Flow source/ Effort source
SS	Effort source/ Flow sensor	SS	Flow sensor/ Effort sensor

The concept of bicausality allows fixing or imposing at the same time a variable and its conjugate as a bicausal bond thus decoupling the effort and flow causalities.

In the context of fault detection and isolation, imposing the output variable without modifying the energy structure (or constraint equations) of the system can be effectuated with an *SS* element having a flow source/effort source causality (Table 1). Then the output to be imposed plays the role of input variable of that *SS* element while its conjugate is set to a null value leading to null power propagation on that bond. Similarly, the fault variable to be isolated will be observed on another *SS* element with flow sensor/effort sensor causality.

4.4. Fault detection and isolation based on bicausal Bond Graph

The bicausal BG used to study fault detection and isolation is obtained by applying Sequential Causality Assignment Procedure for Inversion (SCAPI) algorithm as for system inversion (Ngwompo, Scavarda and Thomasset, 1996) on BG model in integral causality by replacing the input variable by the fault one, and assigning a derivative causality instead of an integral one.

Consider *n* as the number of the storage elements in integral causality, and *a* as the number of storage elements on the shortest path between the input and the output, which will take the derivative causality when the causality assignment algorithm for FDI is applied. $n-a$ is the number of storage elements which are not on the shortest path, and which keep the same causality as they were in the BG in preferred integral causality.

Proposition 1: (fault detectability)
If all storage elements take the derivative causality even ones that are not on the shortest path ($n-a=0$), so the fault variables modeled by modulated inputs are observable (detectable) with respect to the input *u* and the output *y* variables.

Proof: It was shown in (Sueur and Dauphin-Tanguy, 1991) that a system modeled by BG is observable when all the storages elements are derivative causality, so the inputs (sources or faults), and the states variables can be expressed in term of the output variable and their derivative.

On the Bicausal Bond graph, the fault isolability is done by analyzing the causal path from the SS element associated to the output variable to the SS element associated to the fault variable through the storage elements.

Definition: Two causal paths are said to be disjoint if, and only if, they do not share a common variable (Ngwompo and Gawthrop, 1999).

Proposition 2: (fault isolability)
q faults are structurally isolable if, and only if, there are *q* disjoint causal paths linking the sensor to the fault through all storage elements *C*, and *I* that exist in the path.

Proof: Let consider *p* the number of sensors and *q* the number of component faults, such as $q \le p$. The ith causal path, where $i = 1...q$, links the component fault (F_i) which is represented in BG model by SS element to the y_i sensor, has *d* storage elements $C_{l,i}$ and $I_{l,i}$, where $l = 1...d$. The covering causal path from the sensor (SS element) to the fault input (SS element) leads to the following oriented graph where $x_{l,i}$ is the state variable, corresponds to the storage elements in BG model. From Fig.3, the algebraic equations Eq.(10) can be derived.
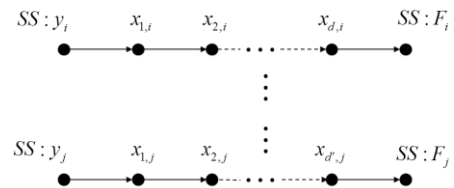
Figure 3 : Digraph representation of disjoint causal paths.

$$F_i = h_i \left(y_i, \dot{y}_i, \ldots, y_i^{(d)} \right)$$
$$\vdots \qquad\qquad (10)$$
$$F_j = h_j \left(y_j, \dot{y}_j, \ldots, y_j^{(d')} \right)$$

The number of derivation *d* and *d'* corresponds to the number of the storage elements that belong to i^{th} and j^{th} causal path respectively.

It is clear that Eq.(10) satisfies Eq.(4), so the q component faults are isolable.

Remark: if one storage element remains in the integral causality when the assignment causality algorithm for FDI is applied on the BG model of the system, the q faults are not isolable even if there is a q disjoint causal paths.

5. ILLUSTRATIVE EXAMPLE

In this section, we will apply the presented methodology to detect and isolate component faults on the electromechanical system of the Fig.4.
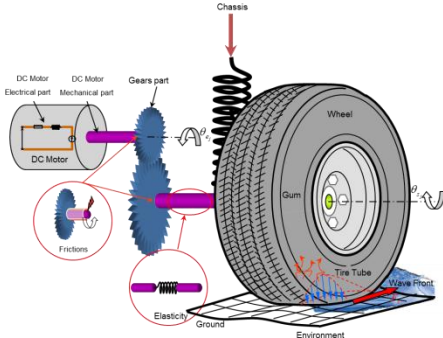

Figure 4: The electromechanical system

The electromechanical system is composed of three principal parts: A DC motor, a gear, and a wheel part. The DC motor is a combination of electrical and mechanical domains. The gear is concerned with connecting the mechanical par with the load one (wheel + ground).

5.1. Case 1

In the first case, two faulty components R: Re and R: f_e are considered. The faults are modeled by a modulated source of effort, Fig.5. The available sensors for the current, and the wheel velocity, represented on the BG model in integral causality by Df: y_1 and Df: y_2.
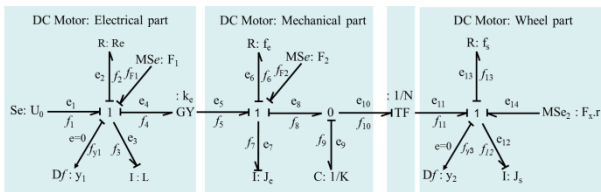
Figure 5 : The BG model of the faulty system of case 1 in integral causality.

5.1.1. Algebraic isolability:
The state space form of the system can be derived from the BG model of the system in integral causality, it is given by the Eq.(11)

The faults F_1 and F_2 are detectable and isolable if, and only if, they satisfy the parity equation Eq.(4). As presented in the Eq.(12), the latter is verified, so the faults F_1 and F_2 are detectable and isolable.

5.1.2. Bicausal BG detectability:
The system given in bicausal BG of the Fig.6 is diagnosable since all I and C elements are in derivative causality which means that the fault F can be expressed in terms of the input u and the output y. We can add that all states are observable with respect to u and y and their derivatives.

$$\begin{cases} \dot{x}_1 = -\dfrac{Re}{L}x_1 - \dfrac{k_e}{J_e}x_2 + u_1 + F_1 \\[2mm] \dot{x}_2 = \dfrac{k_e}{L}x_1 - \dfrac{f_e}{J_e}x_2 - Kx_3 + F_2 \\[2mm] \dot{x}_3 = \dfrac{1}{J_e}x_2 - \dfrac{N}{J_s}x_4 \\[2mm] \dot{x}_4 = NKx_3 - \dfrac{f_s}{J_s}x_4 + u_2 \\[2mm] y_1 = \dfrac{x_1}{L} \\[2mm] y_2 = \dfrac{x_4}{J_s} \end{cases} \quad (11)$$

$$\begin{cases} F_1 = (sL + Re)\,y_1 + \dfrac{k_e}{NK}\left(s^2 J_s + s f_s + N^2 K\right)y_2 \\[2mm] \qquad\qquad - u_1 - s\dfrac{k_e}{NK}u_2 \\[2mm] F_2 = -k_e y_1 + \Bigg[s^3 \dfrac{J_e J_s}{NK} + s^2 \dfrac{1}{NK}\left(f_s J_e + f_e J_s\right) \\[2mm] \quad + s\left(NJ_e + \dfrac{f_e f_s}{NK} + \dfrac{J_s}{N}\right) + \left(Nf_e + \dfrac{f_s}{N}\right)\Bigg]y_2 \\[2mm] \quad - \left[s^2 \dfrac{J_e}{NK} + s\dfrac{f_e}{NK} + \dfrac{1}{N}\right]u_2 \end{cases} \quad (12)$$
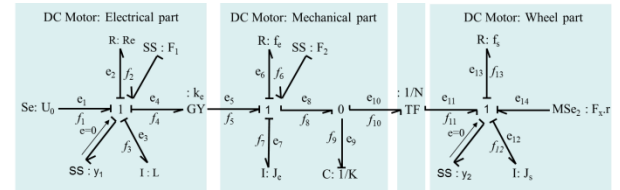
Figure 6 : The bicausal BG model of the faulty system of the case 1.

5.1.3. Bicausal BG isolability:
The Bicausal BG isolability is performed by analyzing the the causal paths from the outputs to the faults. For the faults to be isolable, it is necessary that there are disjoint causal paths equal to number of faults. In other words, q component faults are isolable if, and only if, there are q disjoint causal paths from the output to the fault passing through the storage elements (I, C) that exist in the causal path.

From the Fig.6, there are only two disjoint causal paths from the SS: y_1 to SS: F_1 and, SS: y_2 to SS: F_2

$$SS:y1 \rightarrow I:L \rightarrow SS:F_1$$
$$SS:y2 \rightarrow I:J_s \rightarrow C:\frac{1}{K} \rightarrow I:J_e \rightarrow SS:F_2 \qquad (13)$$

then, the two faults F_1 and F_2 are isolable.

From the bicausal bond graph, the order of derivation is also obtained, it is equal to the number of the storage elements in the path that links the output to the fault: in the expression of the fault F_1, the output y_1 is derived one time $n = 1$, while for the fault F_2, the output y_2 is derived three times $n = 3$.

By following the previous conditions, one is able to state the same conclusion as the algebraic approach (F_1 and F_2 are isolable) without any calculations.

Now, we are going to show how from BG model, the same equations as Eq.(12) are obtained.

5.1.4. Fault indicator determination:
The fault indicator is derived by writing causal relations between BG variables (constraints and characteristic equations). For fault F_1, it is given by:

From the first 1-junction, the following equation can be derived
$$e_{F1} = -e_1 + e_2 + e_3 + e_4 \qquad (14)$$

The efforts e_2, e_3, and e_4 are expressed in function of the input, and the output by following the causal path between the former and the latter. The effort e_2 is linked to the output $SS: y_1$ by the causal path of the Eq.(15), and it is given by $e_2 = R_e f_{y1}$

$$f_3 \rightarrow f_2 \rightarrow e_2 \qquad (15)$$

The same reasoning for the effort e_3, it is given by:
$$e_3 = L\frac{d}{dt}f_{y1}.$$

The effort e_4 is linked to the output $SS: y_2$ (There is no path between e_4 and $SS: y_1$) by

$$f_{y2} \rightarrow f_{12} \rightarrow e_{12} \rightarrow e_{11} \rightarrow e_{10} \rightarrow e_9 \rightarrow f_8 \rightarrow f_7 \rightarrow f_5 \rightarrow e_4 \qquad (16)$$

So, the effort e_4 is given by:

$$e_4 = \frac{k_e}{NK}\left(J_s \frac{d^2}{dt^2} f_{y2} + f_s \frac{d}{dt} f_{y2} + N^2 K f_{y2} - \frac{d}{dt} u_2 \right) \qquad (17)$$

The fault indicator e_{F1} takes the following expression:

$$e_{F1} = R_e f_{y1} + L\frac{d}{dt} f_{y1} + \frac{k_e}{NK}\left(J_s \frac{d^2}{dt^2} f_{y2} + f_s \frac{d}{dt} f_{y2} + N^2 K f_{y2} \right)$$
$$- u_1 - \frac{k_e}{NK}\frac{d}{dt} u_2 \qquad (18)$$

From the second 1- junction, the e_{F2} is given by:

$$e_{F2} = -e_5 + e_6 + e_7 + e_8 \qquad (19)$$

Same reasoning as for fault F_1, the effort e_5, e_6, e_7, and e_8 are given by following the causal path between the former and a given output. They are given by:

$$e_5 = k_e f_{y1}$$
$$e_6 = \frac{fe}{KN}\left(Js\frac{d^2}{dt^2} f_{y2} + fs\frac{d}{dt} f_{y2} + N^2 K f_{y2} - \frac{d}{dt} u_2 \right)$$
$$e_7 = \frac{J_e}{KN}\left(Js\frac{d^3}{dt^3} f_{y2} + fs\frac{d^2}{dt^2} f_{y2} + N^2 K \frac{d}{dt} f_{y2} - \frac{d^2}{dt^2} u_2 \right)$$
$$e_8 = \frac{1}{N}\left(J_s \frac{d}{dt} f_{y2} + f_s f_{y2} - u_2. \right)$$
$$(20)$$

Thus, the expression of the fault variable F_2 is:

$$e_{F2} = -k_e f_{y1} + \left[s^3 \frac{J_e J_s}{NK} + s^2 \frac{1}{NK}\left(f_s J_e + f_e J_s \right) \right.$$
$$+ s\left(NJ_e + \frac{f_e f_s}{NK} + \frac{J_s}{N} \right) + \left(Nf_e + \frac{f_s}{N} \right) \right] f_{y2} \qquad (21)$$
$$- \left[s^2 \frac{J_e}{NK} + s\frac{f_e}{NK} + \frac{1}{N} \right] MSe_2.$$

To get the same algebraic equations as Eq.(12), e_{F1}, e_{F2}, Se_1, MSe_2, f_{y1}, and f_{y2} are substituted by F_1, F_2, u_1, u_2, y_1, and y_2 respectively.

5.2. Case 2
In this case, three faults are considered, these faults are on $R: R_e$, $R: f_e$, and $R: f_s$. The available sensors are the current, the rotor velocity, and the wheel velocity.
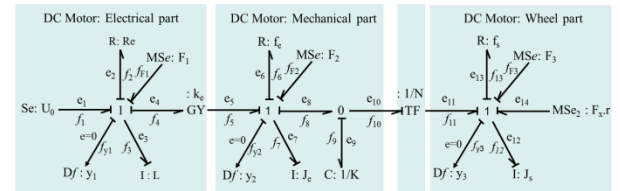
Figure 7 : The faulty system of the case 2.

5.2.1. Algebraic detectability and isolability:
The three faults cannot be expressed with respect to input and the output as shown in Eq(22).

The system is not diagnosable since the state x_3 is not observable with respect to the input and the output, and the three faults are neither detectable nor isolable.

$$\begin{cases} F_1 = (sL + R_e)y_1 + k_e y_2 - u_1 \\ F_2 = -k_e y_1 + (sJ_e + f_e)y_2 + Kx_3 \\ F_3 = -NKx_3 + (sJ_s + f_s)y_3 - u_2 \end{cases} \qquad (22)$$

5.2.2. Bicausal BG detectability and isolability:

The BG model of the system with the faulty elements is given in integral causality in Fig.7, and in bicausal one in Fig.8
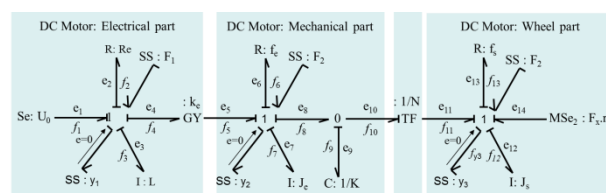

Figure 8 : The bicausal BG of the faulty system of case 2.

From Fig.8, one storage element remains in integral causality when the causality assignment algorithm for FDI is applied on the system in integral causality of Fig.7, so the system is not diagnosable since proposition 1 is not satisfied.

The faults are not also isolable even if there are three disjoint causal paths, which link the outputs to the faults.

CONCLUSIONS

It is shown in this paper that the detection and isolation of plant faults based on bond graph model is simpler than an algebraic approach. The latter is due to the fact that it is based on structural and graphical analysis of the model. The two approaches have similar results expect that bond graph model needs no complex computation compared to an algebraic approach.

REFERENCES

Gertler, J., 1997. Fault detection and isolation using parity relations. Control Engineering Practice, 5:653–661.

Patton, R.J., and Chen, J., 1997. observer-based fault detection and isolation : robustness and applications. Control Engineering Practice, 5:671-682.

Staroswiecki, M., and Comtet-Varga, G., 2001. Analytical redundancy relations for fault detection and isolation in algebraic dynamic systems. automatica, 37:687–699.

Fliess, M., and Join, C., 2003. An algebraic approach to fault diagnosis for linear systems. In CESA.

Fliess, M., Join, C., and SIRA-RAMIREZ, H., 2004. Robust residual generation for linear fault diagnosis: an algebraic setting with examples. International journal of control, 77:1223–1242.

Cruz-Victoria, J.C., Martinez-Guerra, R., and Rincon-Pasaye, J.J., 2008. On nonlinear systems diagnosis using differential and algebraic methods. Journal of the Franklin Institute, 345:102–117.

Commault, C., Dion, J.M., and Agha, S.Y., 2008. Structural analysis for the sensor location problem in fault detection and isolation. Automatica, 44:2074–2080.

Sueur, C. and Dauphin-Tanguy, G., 1991. Bond-graph approach for structural analysis of mimo linear systems. Journal of the Franklin Institute, 328:55–70.

Ngwompo, R.F. and Gawthrop, P.J., 1999. Bond graph-based simulation of non-linear inverse systems using physical performance specifications. Journal of the Franklin Institute, 336:1225–1247.

El-Osta, W., Ould Bouamama, B. and Sueur, C., 2006. Monitorability indexes and bond graphs for fault tolerance analysis. In Safeprocess IFAC, Bejing, Chine, 29-1 Sept.

Samantary, A.K., and Ghoshal, S.K., 2008. Bicausal bond graph for supervision : from fault detection and isolation to fault accommodation. Journal of the Franklin Institute, 345:1–28.

Diop, S. and Martinez-Guerra, R., 2001. On an algebraic and differential approach of nonlinear systems diagnosis. In proceedings of the 40th IEEE : Conference On Decision and Control, 585–589, Orlando, USA.

Dion, J. M.,2003. Generic properties and control of linear structured systems: a survey. Automatica, 39:1125–1144.

Maurya, M. R., Rengaswamy, R., and Venkatasubramanian, V., 2004. Application of signed digraphs-based analysis for fault diagnosis Of chemical process flowsheets. Engineering Applications of Artificial Intelligence, 17(5):501–518.

Blanke, M., Kinnaert, M., Lunze, J., and Staroswiecki, M., 2003. Diagnosis and Fault-Tolerant Control. Springer.

Samantary, A.K., and Ghoshal, S.K., 2008. Bicausal bond graph for supervision : from fault detection and isolation to fault accommodation. Journal of the Franklin Institute, 345:1–28.

Mukherjee, A., and Karmakar, R., 2000. Modelling and Simulation of Engineering Systems through Bondgraphs. Narosa, 1st edition.

Ould-Bouamama, B., Medjaher, K., Samantaray, A.K., and Dauphin-Tanguy, G., 2005. Model builder using functional and bond graph tools for FDI design. Control Engineering Practice (CEP) journal, 13(7):875-891.

Twentesim, 1996. Users Manual of Twentesim (20sim). Comtollab Products Inc, Box 217 NL-7500AEEnschede, Nederland.

Gawthrop, P.J., 2000. Estimating physical parameters of nonlinear systems using bond graph models. In SYSID.

Ngwompo, R.F., Scavarda, S. and Thomasset, D., 1996. Inversion of linear time-invariant siso systems modelled by bond graph. Journal of the Franklin Institute, 336:157–174.

FAULT DETECTION AND ISOLATION IN PRESENCE OF INPUT AND OUTPUT UNCERTAINTIES USING BOND GRAPH APPROACH

Youcef TOUATI, Rochdi MERZOUKI, Belkacem OULD BOUAMAMA

Ecole Polytechnique de Lille, LAGIS, Boulevard Paul Langevin,
59655 Villeneuve d'Ascq Cedex, France

youcef.touati@polytech-lille.fr

ABSTRACT

In this paper, a bond graph model based approach for robust diagnosis in presence of input and output uncertainties is presented. Based on the structural and causal proprieties of the bond graph tool, a procedure of input and output uncertainties modeling is proposed in order to generalize the threshold generation. The proposed procedure is applied to the graphical model in preferred derivative causality, used for analytical redundancy relations. Simulation results are presented in order to validate the proposed procedure of thresholds and residuals generation.

Keywords: Robust fault detection and isolation, measurement uncertainty, Bond graph modeling.

1. INTRODUCTION

Robust fault diagnosis has been the subject of several researches in order to increase the sensibility of the diagnosis systems, to avoid false alarms and to insure the systems safety. Fault detection and isolation (FDI) is based essentially on the comparison between the real behavior of the system and a reference behavior describing the normal situation. The existing approaches to FDI in the literature can be classified on qualitative and quantitative approaches. Qualitative or non-model based methods, are principally based on the artificial intelligence and form recognition such as neuronal and Bayesian approaches (Hsing-Chia K. 2004), (Rothstein A. P. 2005). Qualitative or model based approaches, such as observers and parity relations, are based on the generation of residuals (Frank P.1990), (Iserman R. 1994). The latter are used as indicators in order to detect and isolate faults.In normal operation the residuals are close to zero and different from zero in faulty situations.

Several papers have been devoted to the FDI task in these last years. A survey of classical methods can be found in (Frank P. M. 1997). Many solutions to the robust diagnosis problem have been developed. For example, in (Casavola A. 2008), a solution to robust FDI problem is proposed using deconvolution filters (H_∞ and H_-) with a quasi-convex linear time invariant formulation.

In (Guo J. 2009), the author studied the robust fault detection filter design problem for linear time invariant (LTI) systems with unknown inputs and modeling uncertainties using the formulation of the robust fault detection filter design, as H_∞ model-matching problem, where a solution of the optimal problem is presented using the Linear Matrix Inequality (LMI) technique. Another approach has been developed using the parity space methodology (Han Z. 2002). In the latter, a new scheme of sensor and actuator fault detection and isolation for multivariate dynamic system in presence of parameter and measurement uncertainties is proposed. The measurement and parameter uncertainties are considered bounded and represented by bounded variables in the discrete time state space format.

Most of consulted papers try to eliminate the effect of the parameter or measurement uncertainties on the residuals, such as the filtering approach, which can cause the non-detection of certain faults. However, the parameter uncertainties is not associated to physical parameters but to the state matrix A. The bond graph can be an alternative for uncertainties modeling using the graphical representation. Furthermore, it is well suited for FDI. Indeed, the bond graph tool is used not only for modeling but also for diagnosis, due to its structural and causal proprieties. It is a unified graphical language for multi-physics domains (Karnopp D. 2000), (Borutzky W. 1999). The basic idea of diagnosis using bond graph approach is to generate the residuals, which represent the equations of energy conservation (junction equations), using directly the graphical model (Sueur C. 1991), (Low C. b. 2008). The causal proprieties of the bond graph are used to eliminate systematically the unknown variables, and the structural ones allow the generation of Analytical Redundancy Relations (ARR), using a covering causal paths methodology. The residuals are the evaluation of these ARRs and are used for real time diagnosis. More details about the structural proprieties of the bond graph model can be found in (Sueur C.1989).

2. BOND GRAPH THEORY

A bond graph $G(S,A)$ is a unified graphical language for multi-physic domains. Where the nodes S represent

physical components, subsystems, and other basic elements called junctions. While the edges A, called power bonds represent the power exchanged between nodes.

The set of components named bond graph element is: $S = \{R \cup C \cup I \cup TF \cup GY \cup Se \cup Sf \cup De \cup Df \cup J\}$.

The R element represents a passive energy dissipation phenomena, while C, and I model the passive energy storage elements. (Se), and (Sf) are the sources of effort and flow, respectively. Sensors are represented by flow (Df), and effort (De) detectors. Finally, J (which can be a zero or a one junction), is used to connect the elements having the same effort (0 junction), or flow (1 junction). The conservative energy laws are obtained from the latter. TF, and GY are used to represent transformers and gyrators, respectively.

The passive elements are described by generic constitutive equations: dissipative R-elements (electrical resistor, hydraulic friction) are described by algebraic relationship $F_R(e, f) = 0$, potential storage energy C-element (capacitor, tank, spring) are modeled by an integral equation linking effort and integral of flow $\Phi_c\left(e, \int f dt\right) = 0$ and kinetic storage energy I-element (mechanical inertia, electric coil) is quantified by integral equation between integral of effort and flow $\Phi_I\left(f, \int e dt\right) = 0$.

Definition Dualization of detectors is the replacing of the detector De and Df (Figure 1-a) by a signal source SSe and SSf respectively (Figure 1-b).
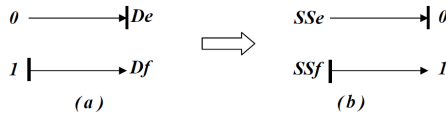
Figure 1: Dualization of detectors

Definition In a bond graph model, a causal path is an alternating of bonds and elements ($R, C, I...$) called nodes such that all nodes have a complete and correct causality. Causal paths of two bonds which have the same node have opposite causal direction. Depending on the causality, the passed variable is the effort or the flow. To change this variable, the causal path must pass through a junction element GY, or a passive element (I, C or R). Definition A system is under-constrained if its dynamic bond graph elements cannot accept the derivative causality when the detectors are dualized.

3. ROBUST FDI TO INPUT AND OUTPUT UNCERTAINTIES

3.1. Output uncertainty modeling
In this section, we describe a generalized method to model measurement uncertainties in order to generate uncertain ARRs, which can be used to obtain both residuals and thresholds. The dualized detectors on the bond graph model, used for diagnosis, in derivative

causality impose the information signal to its associated junctions. Hence, the following equations can be obtained from Figure 2-(a) and (b), respectively:
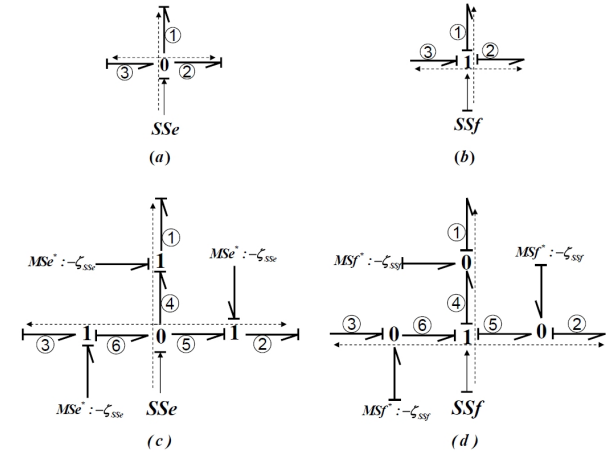
Figure 2: Measurement uncertainties modeling.

$$(a)\begin{cases} e_1 = SSe; \\ e_2 = SSe; \\ e_3 = SSe; \end{cases}$$
$$(b)\begin{cases} f_1 = SSf; \\ f_2 = SSf; \\ f_3 = SSf. \end{cases} \tag{1}$$

The modeling of measurement uncertainties can be done as shown in Fig. 2-(c, d), which are based on the following equation (2):

$$(c)\begin{cases} e_1 = e_4 - \zeta_{SSe} = SSe - \zeta_{SSe}; \\ e_2 = e_5 - \zeta_{SSe} = SSe - \zeta_{SSe}; \\ e_3 = e_6 - \zeta_{SSe} = SSe - \zeta_{SSe}; \end{cases}$$
$$(d)\begin{cases} f_1 = f_4 - \zeta_{SSf} = SSf - \zeta_{SSf}; \\ f_2 = f_5 - \zeta_{SSf} = SSf - \zeta_{SSf}; \\ f_3 = f_6 - \zeta_{SSf} = SSf - \zeta_{SSf}. \end{cases} \tag{2}$$

Where SSf and SSe represent the measured signal ζ_{SSe} and ζ_{SSf} represent the measurement error respectively on SSf and SSe.

3.2. Input uncertainty modeling
The inputs in bond graph tool are represented by a source of effort Se, or by a source of flow Sf, depending on the physical nature of the input components. As a mathematical point of view, the error on the inputs can be expressed as follow:

$$\begin{cases} Se_r = Se + \zeta_{Se}; \\ Sf_r = Sf + \zeta_{Sf}. \end{cases}$$

Where Se and Sf are the predicted input of the effort and flow source respectively, Se_r and Sf_r are the real effort and flow input. ζ_{Se} and ζ_{Sf} represent the uncertainties on the effort and flow source respectively.
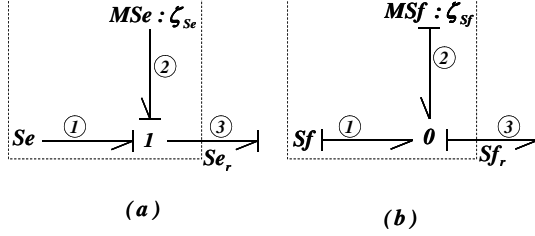
Figure 3: Input uncertainties modeling.

This can be represented in a bond graph form as shown in Figure 3. For the ARR and threshold generation, the errors ζ_{Se} and ζ_{Sf} are considered bounded (3):

$$\begin{cases} -\Delta_{Se} \le \zeta_{Se} \le \Delta_{Se}; \\ -\Delta_{Sf} \le \zeta_{Sf} \le \Delta_{Sf}. \end{cases} \qquad (3)$$

3.3. ARRs and thresholds generation
The thresholds generation can be done after the ARR generation using the following rules:

- Put the model in preferred derivative causality if possible.
- Model the measurement uncertainties directly on the bond graph model.
- Write the ARRs of the model using the equations of energy conservation, and use the causal path to eliminate the unknown variables.
- Write the ARRs of detectors redundancy.
- For all ARRs derived from the equations of energy conservation, the threshold is obtained by adding the maximal absolute values of the different parts of the ARRs containing the measurement errors.
- For the ARRs generated from the redundant detectors, the threshold is equal to the sum of the maximum measurement errors of the two redundant detectors.

Let us consider the linear system illustrated in Figure 4.
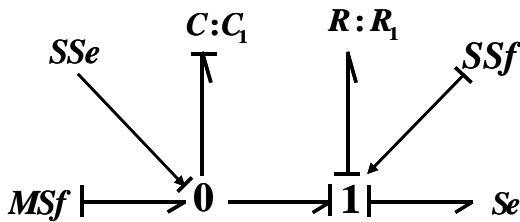
Figure 4: Linear system.

Two ARRs can be generated from this model, one from the 0-junction and the other from the 1-junction. The two ARRs are used as residuals, and theoretically equal to zero in normal situation, without considering the uncertainties and model errors. In presence of uncertainties on the sensors measurement, and if we know that the measurement error is an additive and bounded error, the residual can be bounded by a two thresholds, which can be calculated using the bond graph model directly. Applying the procedure of measurement uncertainties modeling, the model of the linear system become as shown in Figure 5. The model can be used to generate the uncertain part directly, by using the causal paths. For example in Figure 5: $MSf^* : \zeta_{Sf} \to 15 \to 3 \to 8 \to ARR_1$. These causal paths are used to generate the uncertain part of the ARR. We must start from virtual sources that represent the measurement errors to observed junctions (junctions which are connected to a detector). We remark that the virtual source of effort connected with the bond 13, and the virtual source of flow connected with the bond 7 have no causal path to an observed junction. So the measurement uncertainty on these bonds can be removed.

The measurement error on the bond 12 can be modeled directly on the observed junction like shown in Figure 6. This modeling procedure can be used with the R, C and I element connected to an observed junction. Using these rules the model of Figure 5 can be simplified to the model of Figure 7.
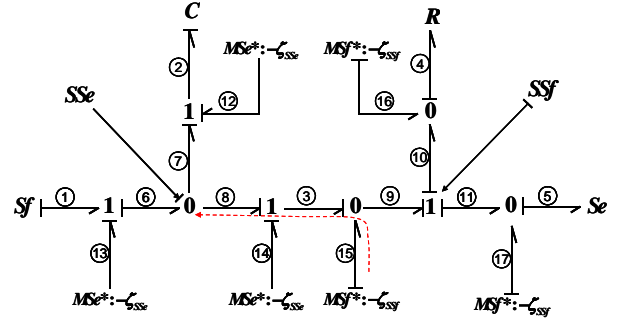
Figure 5: The model of the linear system with measurement uncertainties.

The robust ARRs of this system (Figure7) can be written as follows:

$$ARR_1 : MSf - C_1 \frac{dSSe}{dt} - SSf + \zeta_{SSf} + \Phi_{C_1}\left(\zeta_{SSe}\right) = 0;$$

$$ARR_2 : SSe - R_1 SSf - Se + \zeta_{SSe} + \Phi_{R_1}\left(\zeta_{SSf}\right) + \Delta_{Se} = 0.$$

These two ARR can be decomposed to two residuals r_1 and r_2 (equation 4), and two thresholds a_1 and a_2 (equation 5).

$$\begin{aligned} r_1 &= MSf - C_1 \frac{dSSe}{dt} - SSf; \\ r_1 &= SSe - R_1 SSf - Se. \end{aligned} \qquad (4)$$

$$a_1 = max\left(\zeta_{SSf} + \Phi_{C_1}\left(\zeta_{SSe}\right)\right);$$
$$a_2 = max\left(\zeta_{SSe} + \Phi_{R_1}\left(\zeta_{SSf}\right) + \zeta_{Se}\right).$$

$$a_1 = max\left(\zeta_{SSf}\right) + max\left(\Phi_{C_1}\left(\zeta_{SSe}\right)\right);$$
$$a_2 = max\left(\zeta_{SSe}\right) + max\left(\Phi_{R_1}\left(\zeta_{SSf}\right)\right) + max\left(\zeta_{Se}\right). \quad (5)$$
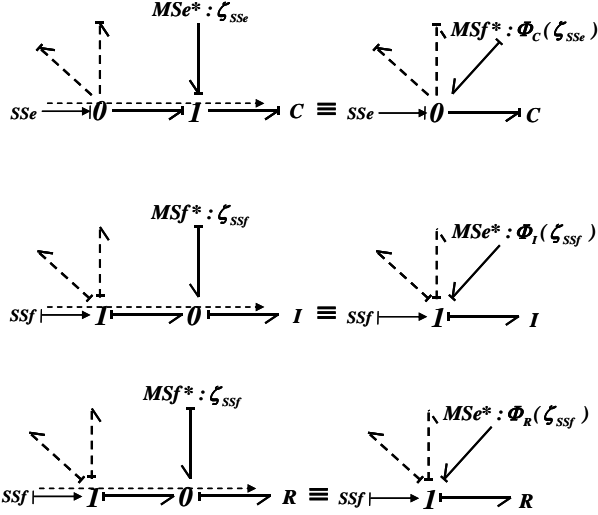
Figure 6: Simplification.
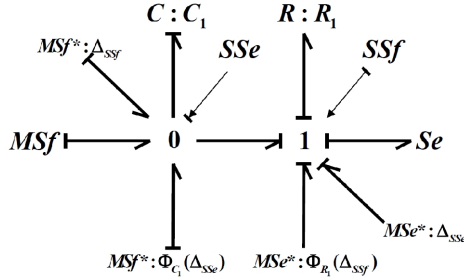
Figure 7: Simplified model of the linear system with measurement uncertainties.

$max\left(\Phi_{C_1}\left(\zeta_{SSe}\right)\right)$ can be bounded using the maximum and the minimum value of ζ_{SSe}.

$$max\left(\Phi_{C_1}\left(\zeta_{SSe}\right)\right) = max\left(C_1 \frac{d\zeta_{SSe}}{dt}\right).$$

In real applications, the differentiation is done in discrete time, by using some methods; the simplest is to use two values in different time as follows:

$$\frac{d\zeta_{SSe}}{dt} = \frac{\zeta_{SSe}^i - \zeta_{SSe}^{i-1}}{t_i - t_{i-1}}.$$

Where ζ_{SSe}^i is the measurement error in the time t_i. The measurement error is ζ_{SSe}^i bounded by Δ_{SSe}, so the quantity $\frac{\zeta_{SSe}^i - \zeta_{SSe}^{i-1}}{t_i - t_{i-1}}$ can be bounded as follows:

$$\frac{\zeta_{SSe}^i - \zeta_{SSe}^{i-1}}{t_i - t_{i-1}} \le \frac{\Delta_{SSe} - \Delta_{SSe}}{t_i - t_{i-1}} \le \frac{2\Delta_{SSe}}{t_i - t_{i-1}}.$$

So we can write:

$$max\left(\Phi_{C_1}\left(\zeta_{SSe}\right)\right) = C_1 \frac{2\Delta_{SSe}}{t_i - t_{i-1}}.$$

4. ROBUST FAULT ISOLATION

The fault isolation can be done using the fault signature matrix (*FSM*), which can be directly deduced either from the analytical redundancy relations or from the bond graph model directly, noting that the ARRs can be deduced directly from the graphical model. In this section, we propose a

Robust FSM (*S*) shown in Table I, where the columns are the residuals (r_i, $i = 1, 2,...,n$), and the rows are the parameters that model the components (c_i, $i = 1, 2, ...m$).

Table 1:ROBUST FAULT SIGNATURE MATRIX.

	r_1	r_1		r_n	I_b	D_b
c_1	$s_{1,1}$	$s_{1,2}$	$\cdots$	$s_{1,n}$	g_1	e_1
c_2	$s_{2,1}$	$s_{2,2}$	$\cdots$	$s_{2,n}$	g_2	e_2
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
c_m	$s_{m,1}$	$s_{m,2}$	$\cdots$	$s_{m,n}$	g_m	e_m

The $S^{m \times n}$ is a Boolean matrix, where:

$$\begin{cases} 1 \text{ if the } i^{th} \text{ ARR contains } c_j. \\ 0 \text{ otherwise.} \end{cases}$$

G is defined as the isolability vector, where:

$$\begin{cases} 1 \text{ if } s_{i,j=1 \text{ to n}} \text{ is unique,} \\ 0 \text{ otherwise.} \end{cases}$$

If the signature of the fault is unique, then the fault is isolable when its effect on the residuals is bigger than all the thresholds associated with these residuals.
E is the detectability vector calculated as follows:

$$e_i = s_{i,1} \vee s_{i,2} \vee \cdots \vee s_{i,n}.$$

The fault is detectable when its effect on the residuals is bigger than one of the sensible residual,

5. APPLICATION

In this section, the presented procedure of robust ARRs generation with respect to output and input uncertainties is applied to an electromechanical subsystem of a robot named Robotino. The latter is composed of three electromechanical subsystems (three omni-directional wheels with three DC motors). Each Subsystem contains two detectors that measure the current and the angular speed of the DC motor.

5.1. Modeling

The considered subsystem is modeled by bond graph tool as shown in Figure 8. To obtain the model, the different physical components and phenomenon are considered. The electrical part of the *DC* motor is modeled by R_a, L_a and U which represent the electrical resistance, inductance and the voltage, respectively. The energy transfer between the electrical part and the mechanical part (mechanical resistance R_e and inertia J_e) of the *DC* motor is represented by a gyrator (*GY*). The reducer is modeled by a transformer *TF* and the inertia of the wheel is represented by J_s and the resistance with R_s.
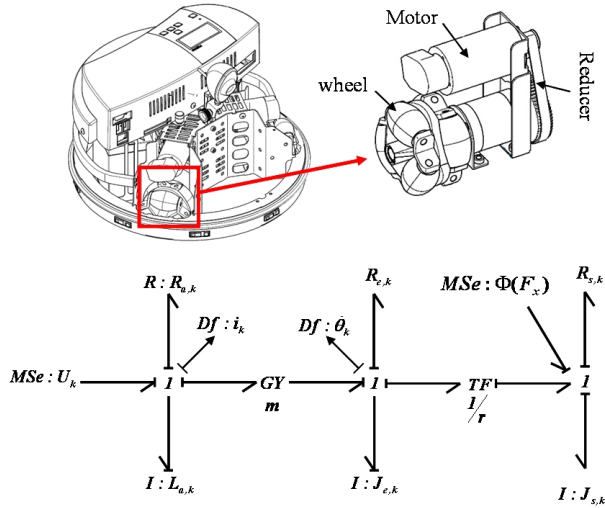
Figure 8: Bond graph model of the electromechanical subsystem.

From the model of Figure 9, the following ARRs can be obtained:

$$ARR_1 = U - L_a \frac{di}{dt} - R_a i - m \cdot \dot{\theta}_m + a_1;$$

$$ARR_2 = mi - J_e \frac{d\dot{\theta}_m}{dt} - R_e \dot{\theta}_m - \frac{J_s}{r^2} \frac{d\dot{\theta}_m}{dt} - \frac{R_s}{r^2} \dot{\theta}_m + \frac{1}{r} \Phi(F_x) + a_2;$$

$$a_1 = -\zeta_{R_a/i} - \zeta_{L_a/i} - m\zeta_{\dot{\theta}_m};$$

$$a_2 = -m \cdot \zeta_i - \zeta_{R_e/\dot{\theta}_m} - \zeta_{J_e/\dot{\theta}_m} - \frac{J_s}{r^2} \frac{d\zeta_{\dot{\theta}_m}}{dt} - \frac{R_s}{r^2} \zeta_{\dot{\theta}_m} + \frac{1}{r} \zeta_{\Phi(F_x)};$$

With

$$\zeta_{R_a/i} = \Phi_{R_a}(\zeta_i) = R_a \zeta_i;$$

$$\zeta_{L_a/i} = \Phi_{L_a}(\zeta_i) = R_a \frac{d\zeta_i}{dt};$$

$$\zeta_{R_e/\dot{\theta}_m} = \Phi_{R_e}(\zeta_{\dot{\theta}_m}) = R_e \zeta_{\dot{\theta}_m};$$

$$\zeta_{J_e/\dot{\theta}_m} = \Phi_{J_e}(\zeta_{\dot{\theta}_m}) = R_e \frac{d\zeta_{\dot{\theta}_m}}{dt}.$$

Where ζ_i and $\zeta_{\dot{\theta}_m}$ are the measurement errors on the current and velocity detectors respectively. $\zeta_{\Phi(F_x)}$ is the error on the input of the impact effort torch $\Phi(F_x)$, All the output and the input errors are considered bounded as follows:

$$|\zeta_i| \le 0.01A; \quad |\zeta_{\dot{\theta}_m}| \le 0.1 rad/s; \quad |\zeta_{\Phi(F_x)}| \le 0.001Nm;$$
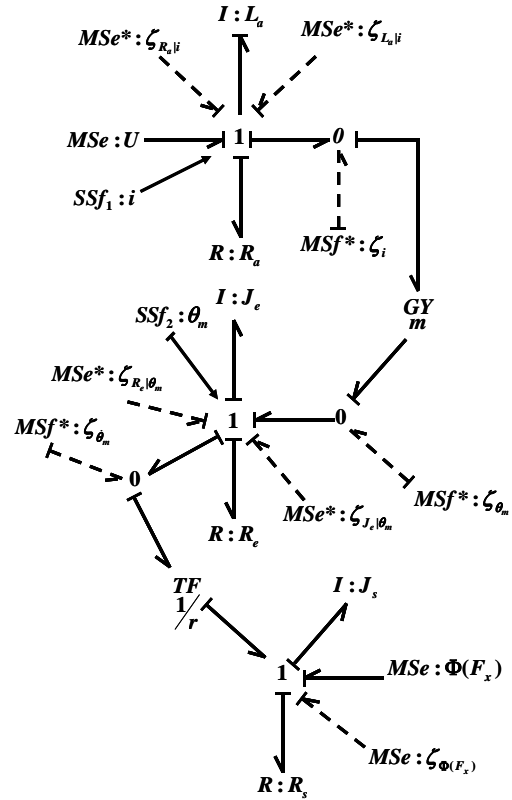
Figure 9: Bond graph model of the electromechanical subsystem with input

5.2. Simulation

In this subsection, simulations are performed in order to validate the developed diagnosis procedure in presence of input and output (measurement) uncertainties. The parameters (Table 2) of the model are obtained from a real system. The torque of the impact effort $\Phi(F_x)$ is considered as a constant and a bounded variable. The used sampling interval time is 0.2s.

Table 2: Parameters.

$$
\begin{array}{c|l}
L_a & = 8.9mH \\
R_a & = 8.13\Omega \\
m & = 43.1mV \\
R_e & = 47\mu Nm.sec.rad^{-1} \\
J_e & = 7.95 \times 10^{-6}Kg.m^2 \\
r & = 16 \\
J_s & = 63 \times 10^{-3}Kg.m^2 \\
R_s & = 0.02Nm.sec.rad^{-1}
\end{array}
$$

Fig. 10 shows the input voltage signal, the output current of the electrical part of the system, and the output signal of the velocity of the wheel. In Fig. 11, the residuals and thresholds are represented. In Fig. 12, the residuals in faulty case are presented, where the additive fault on the input signal is around 12 %.

As the results show, the threshold depends on the precision of the detector and parameter values, and we remark that the derivative of the signal amplifies the noise. Consequently the thresholds also increase, because the estimation of the derivative depends on the sampling time.

6. CONCLUSION

In this paper, a procedure of robust fault detection and isolation using the bond graph tool is proposed, toward output (measurement) and input uncertainties, in order to avoid false alarms. This approach clearly shows that the performance of the diagnosis system depends on the sensors and actuators precision. In this graphical approach, the measurement and input uncertainties are associated, to the sensors (*Df* and *De*) and actuators (*Se, Sf*) respectively. In addition, this approach can be generalized and automated for *LTI* systems and some nonlinear systems.
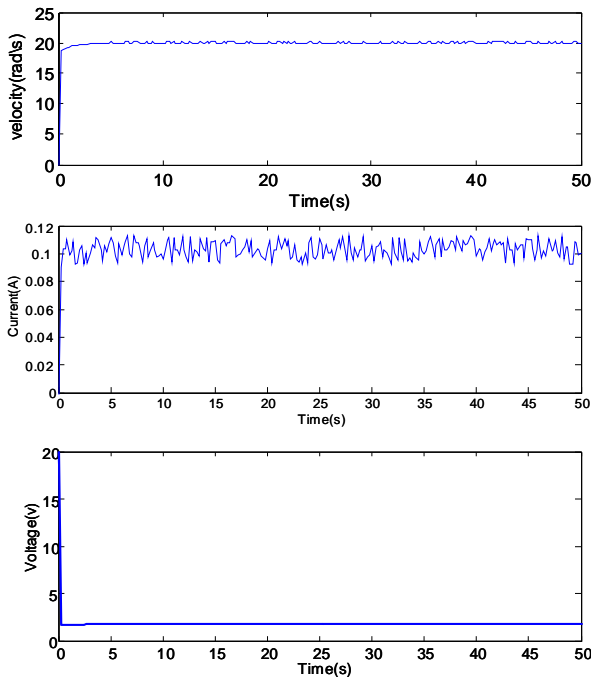
Figure 10: Voltage, current and velocity signals.
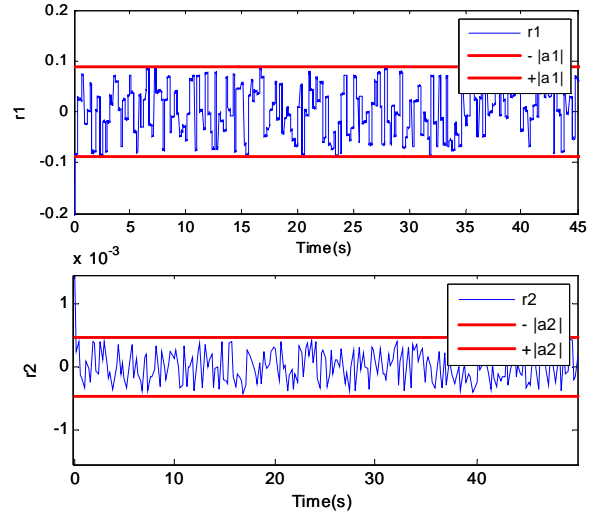
Figure 11: Residuals in normal situation.
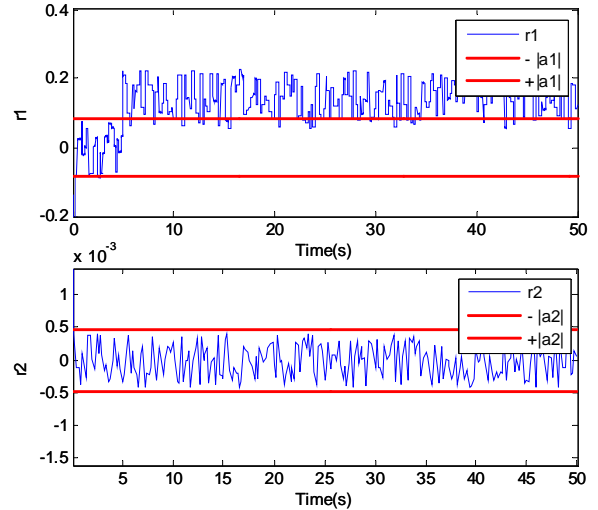
Figure 12: Residuals in faulty situation.

REFERENCES

Hsing-Chia K., Hui-Kuo C. 2004. "A newsymbiotic evolution-based fuzzyneural approach to fault diagnosis of marine propulsion systems". Engineering Applications of Arti cial Intelligence,Vol.17. pp. 919-930.

Rothstein A. P., Posner M., Rakytyanska H.B. 2005. "Cause and effect analysis by fuzzy relational equations and a genetic algorithm". Reliability Engineering and System Safety.

Frank P., 1990 "Fault diagnosis in dynamic systems using analytical and knowledge-based redundancy- a survey and some new results", Automatica26 (3), 459-474.

Iserman R., 1994 Fault Detection based on modelling and estimation methods, a survey, Automatica 20 387-404.

Frank P. M. and Ding X. 1997 "Survey of robust residual generation and evaluation methods in observer- based fault detection systems" J. Proc. Cont. Vol. 7, N°. 6, pp. 403-424. Elsevier Science

Ltd.

Casavola A., Famularo D., Franz G. 2008, "Robust fault detection of uncertain linear systems via quasi-LMIs". Automatica 44 289 - 295.

Guo J., Huang X., Cui Y., 2009 "Design and analysis of robust fault detection filter using LMI tools". Computers and Mathematics with Applications 57 1743 1747.

Z. Han ,W. Li , S. L. Shah. 2002. "Fault detection and isolation in the presence of process uncertainties". 15th IFAC Wold Congress. pp.1887-1892.

Karnopp D.C., Margolis D.L., Rosenberg R.C. , 2000 System Dynamics, Modeling and simulation of mechatronic systems. John Wiley and Sons Inc. ISBN 0-471-33301-8.

Borutzky W. , 1999, Relations between Graph based and object-oriented physical systems modeling, CBGM'99 International Conference on Bond Graph Modelling and Simulation, San Fransisco, CA, Jan. 17-20, pp.11-17.

Sueur C., Dauphin-Tanguy G., 1991, Bond graph approach for structural analysis of MIMO linear systems. J. Franklin Inst. 328 (1) 55-70.

Low C. B., Wang D., Arogeti S., Zhang J. B., 2008 Monitoring ability analysis and qualitative fault diagnosis using hybrid bond graph. Proceedings of the 17th World Congress The International Federation of Automatic Control Seoul, Korea, July 6-11.

Sueur C., Dauphin-Tanguy G. 1989. Structural Controllability and Observability of linear Systems Represented by Bond Graphs. Journal of Franklin Institute. Vol 326, Issue 6, 1989, Pages 869-883.

MONITORING STEADY-STATE GAINS IN LARGE PROCESS SYSTEMS

Germán A. Bustos [(a)], José L. Godoy [(b)], Alejandro H. González[(c)], Jacinto L. Marchetti [(d)]

[(a,b,c,d)] INTEC (CONICET and Universidad Nacional del Litoral), Güemes 3450 (3000), Santa Fe, Argentina
[(a)]gbustos, [(b)]jlgodoy, [(c)]alejgon, [(d)]jlmarch@santafe-conicet.gov.ar

ABSTRACT

Stationary process gains are critical model parameters to determine targets in commercial MPC technologies. Consequently, important savings can be reached by acceding to an early prevention system capable of indicating when the actual process moves away from the modeled dynamics, particularly when the actual process gains are no more represented by those included in the model identified during commissioning stages. A subspace identification method is used here to determine the state-space model matrices that help to define a gain matrix estimator. The pursued goal is a monitoring method capable to identify those gains of a multivariable model that start a mismatching condition, or that show tendencies to change already known mismatches with the actual process. The anticipated knowledge of these events should prevent process engineers about the eventual reliability of targeting optimal process conditions with wrong gain estimations, and thus, help to localize the dynamic relationships for which an updating identification would be necessary.

Keywords: Steady-State Gains, Multivariable Estimator, Subspace Identification, LP-MPC

1. INTRODUCTION

Model predictive control (MPC) has a wide application in the chemical process industry and other industrial sectors. Commercial MPC systems are typically implemented in conjunction with a steady-state linear (LP) or quadratic programming (QP) optimizers (Ying and Joseph 1999), whose main function is to track the economic optimum and provide feasible set-points or targets to the predictive controller. However, despite the widespread adoption of these two-level control systems, occurrences of poor performance have been reported. A frequent claim has been that model mismatches lead the operation away from the real optimum, and that large variations in the computed input and output targets have been observed (Nikandrov and Swartz, 2009). Since the stationary process gains are critical model parameters to determine the MPC targets, important savings can be obtained by acceding to an early prevention system capable to indicate when the actual process moves away from the modeled behavior, particularly by indicating when the actual process gains are no more represented by those included in the model identified during early commissioning stages.

This work is the first stage of a research project attempting to contribute to the development of a method for the online estimation of multivariable process gains directly from row data, and creating the basis for a monitoring tool capable of detecting significant gain changes, particularly when working simultaneously with multivariable control applications like the combined LP (or QP) and MPC. The analysis presented here is still far from the final goal but highlights questions and dificulties that need to be answered and overcome to obtain the reliable estimations.

This paper is organized as follows: after the short introduction exposing the motivations and goals of this project, the subspace identification method is described in Section 2. The gain matrix estimator is defined in Section 3 base on the state space model matrices and Section 4 comments about monitoring alternatives for multivariable cases; however, the simpler Shewhart chart strategy was adopted for a closer inspection of the individual estimations. In Section 5, some simulation results are presented and the conclusions are given in Section 6, together with comments about future work associated to this subject.

2. SUBSPACE IDENTIFICATION (SID)

Early in the nineties, a new identification method for dynamic systems received the attention of many scholars and practitioners. The subspace identification method (SID) has the appealing feature of allowing the direct use of row data with scarce preprocessing needs and the ability of being applicable to multivariable process systems. Several analysis and applications were reported since then, to name a few: Van Overschee and De Moor (1996); Favoreel et al. (2000) and Katayama (2005).

Most subspace approaches fall into the unification theorem proposed by Van Overschee and De Moor (1996), being the followings the three better known: **N4SID** (Van Overschee and De Moor, 1994); **CVA** (Larimore, 1990) and **MOESP** (Verhaegen and Dewilde, 1992). These algorithms can be viewed as singular values decompositions of a weighted matrix. They provide reliable state-space models of multi-variable LTI systems directly from input-output data and, do not require iterative optimization procedures; this basically means that there are no problems of local minima, convergence, or initialization.

A main advantage of the subspace approaches is that preliminary or previous parameterization - a

complex task when dealing with MIMO systems - is not necessary before the identification process. Another advantage is the computational efficiency and robustness given by linear algebra tools like QR and SVD decompositions.

In this paper, our goal heads mainly to the online identification of steady-state gains of MIMO dynamic systems using a standard subspace approach (Overschee and De Moor, 1996) where specific matrices are obtained from projections of subspaces generated by the input and output data. These projections are useful to eliminate noisy components.

2.1. State space model

The method considers the following stochastic state-space model in an innovation form:

$$x(k+1) = Ax(k) + Bu(k) + Ke(k) \qquad (1)$$

$$y(k) = Cx(k) + Du(k) + e(k) \qquad (2)$$

where $x(k) \in \mathbb{R}^n$ stands for a n-dimensional state, $u(k) \in \mathbb{R}^m$ represents the m inputs to the system, $y(k) \in \mathbb{R}^l$ is the l-dimensional output, K is the steady state Kalman gain and, $e(k) \in \mathbb{R}^l$ is an unknown innovation with covariance matrix $R = E\{e(k)e(k)'\}$.

The problem solved by the SID method can be described as follows: given a large enought data set $\{y_k, u_k\}$ from an unknown system, find the model order n, the model matrices A, B, C and D for a state-space representation similar to (1) and (2), and estimate K.

2.2. Estimation of model matrices

An iterative substitution of Eqns. (1) yields the expresión

$$Y_f = \Gamma_i X_f + H_i^d U_f + H_i^s E_f \qquad (3)$$

where the subscripts f and p denote future and past horizons respectively. The matrices $Y_f \in \mathbb{R}^{li \times j}$, $U_f \in \mathbb{R}^{mi \times j}$ and $E_f \in \mathbb{R}^{li \times j}$ are the output, input and noise block Hankel matrices respectively. H_i^d and H_i^s are low triangular Toeplitz matrices composed from the impulse responses of deterministic and stochastic subsystems respectively.

Notice that the system information is mainly in the first term, which includes the extended observability matrix and the state sequence X_f. Assuming the noise E_f is independent from the past input U_p, the past output Y_p and the future input U_f, then the following relationship can be obtained (see Appendix A for nomenclature usage)

$$Y_f /_{U_f} W_p = \Gamma_i . X_f \qquad (4)$$

According with Van Overschee and De Moor (1996), this result tells the space column of Γ_i is the same as the space column of $Y_f /_{U_f} W_p$, which can be estimated from input-output data. Then, applying

singular value decomposition the left side can be written as

$$Y_f /_{U_f} W_p = U S V^T = U S^{1/2} S^{1/2} V^T \qquad (5)$$

from where the extended observability matrix Γ_i can estimated by

$$\hat{\Gamma}_i = U S^{1/2} \qquad (6)$$

Once this matrix is obtained, the model matrices in (1) and (2) can estimated in the following order: (C, A) and (B, D).

The matrix C is obtained directly from the first row block of Γ_i, i.e.,

$$C = \Gamma_i(0:l-1, 0:n-1). \qquad (7)$$

The matrix A is then obtained using the invariant displacement property of Γ_i, this is:

$$\overline{\Gamma}_i = \underline{\Gamma}_i A , \qquad (8)$$

where $\overline{\Gamma}_i \in \mathbb{R}^{l(i-1) \times n}$ and $\underline{\Gamma}_i \in \mathbb{R}^{l(i-1) \times n}$ are matrices obtained from the last $l(i-1)$ and the first $l(i-1)$ row blocks of Γ_i respectively. Hence, matrix A can be estimated by the mean squared solution as follows:

$$A = \underline{\Gamma}_i^\dagger \overline{\Gamma}_i \qquad (9)$$

where $\underline{\Gamma}_i^\dagger = \left(\underline{\Gamma}_i' \underline{\Gamma}_i\right)^{-1} \underline{\Gamma}_i'$ is a pseudo inverse matrix.

Now, left and right multiplying Eqn. (2) by $\Gamma_i^\perp$ and $U_f^\dagger$ respectively

$$\Gamma_i^\perp Y_f U_f^\dagger = \Gamma_i^\perp \Gamma_i X_f U_f^\dagger + \\ \Gamma_i^\perp H_i^d U_f U_f^\dagger + \Gamma_i^\perp H_i^s E_f U_f^\dagger \qquad (10)$$

Notice $\Gamma_i^\perp \in \mathbb{R}^{(li-n) \times li}$ is a complete range matrix satisfying $\Gamma_i^\perp . \Gamma_i = 0$ and $U_f U_f^\dagger = I$. Hence, assuming negligible noise, the above expression simplifies to

$$\underbrace{\overbrace{\Gamma_i^\perp}^{Data} Y_f U_f^\dagger}_{\in \mathbb{R}^{(li-n) \times mi}} = \overbrace{\Gamma_i^\perp}^{Data} H_i^d \qquad (11)$$

This is an over determinate system of linear equations where matrices B and D are the unknowns and that we can rewrite as follows:

$$(M_1 \; M_2 \ldots M_i) = (L_1 \; L_2 \ldots L_i) \begin{pmatrix} D & 0 & \cdots & 0 \\ CB & D & \cdots & 0 \\ CAB & CB & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ CA^{i-2}B & CA^{i-3}B & \cdots & D \end{pmatrix}$$

and rearranging

$$\underbrace{\begin{pmatrix} M_1 \\ M_2 \\ \vdots \\ M_i \end{pmatrix}}_{\in \mathbb{R}^{i(li-n) \times m}} = \underbrace{\begin{pmatrix} L_1 & L_2 & \cdots & L_{i-1} & L_i \\ L_2 & L_3 & \cdots & L_i & 0 \\ L_3 & L_4 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ L_i & 0 & 0 & \cdots & 0 \end{pmatrix}}_{\in \mathbb{R}^{i(li-n) \times li}} \underbrace{\begin{pmatrix} I_l & 0 \\ 0 & \Gamma_i \end{pmatrix}}_{\in \mathbb{R}^{li \times (l+n)}} \begin{pmatrix} D \\ B \end{pmatrix} \qquad (12)$$

where $M_k \in \mathbb{R}^{(li-n) \times m}$ and $L_k \in \mathbb{R}^{(li-n) \times l}$. This system can then be solved with a least squared approach.

Once the model matrices *A*, *B*, *C* and *D* in equations (1) and (2) are estimated, the process model gain matrix can be computed by using the estimator

$$\hat{G}(1) = \hat{C}[I - \hat{A}]^{-1}\hat{B} + \hat{D}. \tag{13}$$

3. MONITORING ALGORITHMS

Modern data acquisition allowed process control systems to become multivariable and provided the technological base to develop monitoring applications capable of a simultaneous surveillance of several related characteristics. This motivated the challenge for extending several single variable statistical methods to applications in multivariable systems. Today there are at least three multivariable control statistics that rise as extensions of previous single-variable versions: the statistical distance or Hotelling's T^2 proposed by Hotelling (1947), multivariate acumulated sum or MCUSUM proposed by Woodall and Ncube (1985) and Crosier (1988), and the multivariate exponetially weighted moving average MEWMA proposed by Lowry et al. (1992). All of them have received attention from the industry and the academy; thus, many authors and references should be mentioned for barely covering the later contributions. In this context, it is worth to recall that a main difference between the T^2 statistic and MCUSUM or MEWMA is that the first one defines a sequence of hypothesis tests based just on the last multivariate observation while the others use a collection of past data of the process employing a different forgetting mechanism.

Despite the fact that important algorithms like those mentioned above are available, the application example exposed in the following section emphasizes the inspection of plain gain estimators of a known multivariable linear system, with the purpose of highlighting main difficulties appearing during the block estimation through the subspace approach method.

In general, monitoring systems are based on statistics defined such to estimate main parameters of the model representing the system to be watched. If the distribution density function of this estimator is known, then, a Shewhart-type control chart can be easily set by defining an interval with a given confidence level. Appendix B presents a short review of concepts providing fundaments to the well known two-side control chart of the Shewhart type. In this framework and considering the final motivation of this work, the MPC model gains would be taken as plausible parameter values (θ_0) to set the hypotesis tests implicit in plain Shewhart control charts for monitoring significant changes.

4. SIMULATION RESULTS

The gain estimation method proposed here is tested by numerical simulation where the process plant is represented by the 5x3 LTI system given in Table 1. This arbitrary example is defined with more outputs than inputs becouse this is a frequent characteristic of actual process systems where MPC is applied.

$\dfrac{-0.8}{15s^2 + 3s + 1}$	$\dfrac{4.7}{10s^2 + 7.8s + 1}$	$\dfrac{-2.0}{9.14s + 1}$
$\dfrac{1.4}{8s^2 + 5s + 1}$	$\dfrac{0.8}{10s^2 + 3s + 1}$	0.0
$\dfrac{1.5}{7.5s + 1}$	$\dfrac{-0.3}{10s^2 + 4s + 1}$	$\dfrac{0.9}{15s^2 + 10s + 1}$
$\dfrac{4.0}{7.8s + 1}$	0.0	$\dfrac{2.5}{5s + 1}$
$\dfrac{1.9}{6s^2 + 4s + 1}$	$\dfrac{5.0}{16s^2 + 5s + 1}$	$\dfrac{-2.8}{9.8s^2 + 10s + 1}$

Table 1: Multivariable linear system used for testing the proposed gain estimation method.

Like most identification procedures, the experience here starts by introducing a PRBS signal in the manipulated input variables (Gaikwad and Rivera, 1996) as shown if Figure 1. In order to approach the conditions of typical signals coming from actual process systems, Gaussian noise is added to the output variables with such intensity that the noise-to-signal ratio reaches 0.1 approximately (Huang and Kadali, 2008). The exciting random signals are conditioned to comply with the persistent excitation requirements and starting with a different seed for each input to avoid simultaneous similar sequences, i.e. to avoid correlations between inputs. The adopted sampling time is 2 min., but the estimations are calculated every 400 min only, each one using the last 400 observations of the input and output vectors. Hence, these are "moving" estimations calculated with 200 "new" observations and 200 "old" observations. This procedure was adopted to have a gain-matrix estimation every 6.6 hours without resigning a conveniently large number of observations.
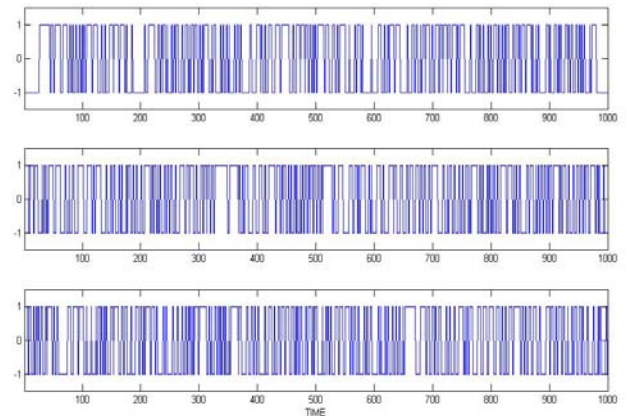

Figure 1: PRBS signals used in the input variables.

Two sequences of data were simulated. The first one assumes the system is stable and operating without receiving disturbances or gain changes; this set of data serves to desing the control limits for the individual

gains according with the analysis in Apendix B. Thus, for every gain $\theta_{ij} \equiv G_{ij}$, $i = 1$ to l; $j = 1$ to m, the limits are defined by

$$LS_{\theta_{ij}} = \overline{\theta}_{ij} + d_{(\alpha/2)}\hat{\sigma}_{\hat{\theta}_{ij}} = \overline{\theta}_{ij} + 3\hat{\sigma}_{\hat{\theta}_{ij}}$$

$$LI_{\theta_{ij}} = \overline{\theta}_{ij} - d_{(1-\alpha/2)}\hat{\sigma}_{\hat{\theta}_{ij}} = \overline{\theta}_{ij} - 3\hat{\sigma}_{\hat{\theta}_{ij}}$$

where subscripts i and j represent the output and input respectively, and the "3-sigma" convention is adopted for simplicity. The central line for these control charts are calculated using $N = 49$ estimations as follow:

$$\overline{\theta}_{ij} = \frac{\sum_{k=1}^{N} \hat{\theta}_{ij}(k)}{N},$$

and the standard deviation of the estimator is calculated by

$$\hat{\sigma}_{\hat{\theta}_{ij}} = \sqrt{\frac{\sum_{k=1}^{N} \left(\hat{\theta}_{ij}(k) - \overline{\theta}_{ij} \right)^2}{N-1}}.$$

The second data set is used for verification and demonstration of the ability to detect slow gain changes. Figure 2 shows the gain matrix estimations along 49 intervals of 200 mins; most of them show quite good accuracy though with different dispersion patterns. For instance, both gains G_{12} and G_{13} receive a 5% ramp change from time 3000 to 8000 min, i.e., from 7.5 to 20 in terms of the estimation number used in the horizontal axis in Figs. 2 to 4. Besides, G_{31} have also a 5% ramp change from time 4000 to 5000 (or equivalently 10 to 12.5 estimations time).
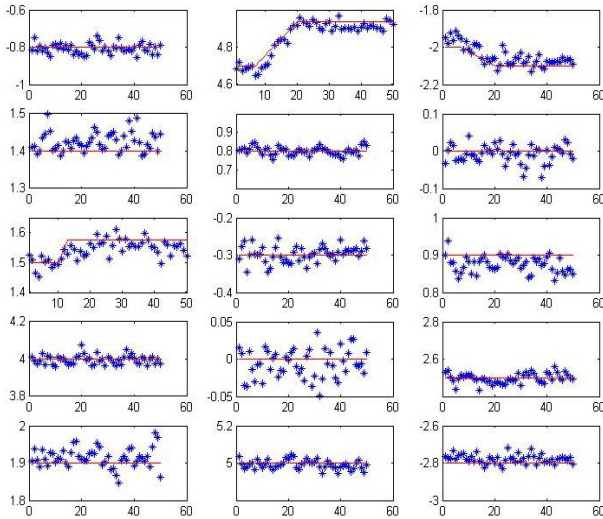
Figure 2: Responses of the gain matrix estimator vs. estimation number (one every 200 min) showing the effect of ramp changes in G_{12} G_{13} and G_{31}. Red lines indicate true gain values.
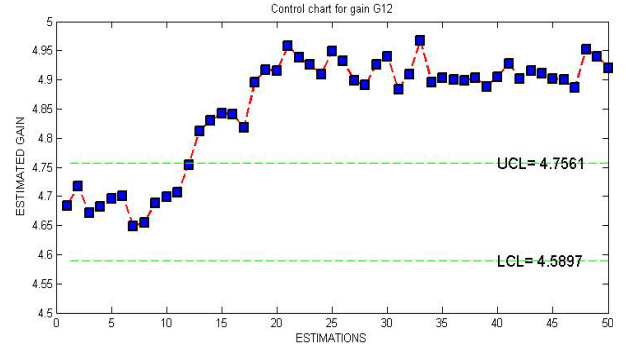
Figure 3: G_{12}-estimator response to a ramp change augmenting 5% the gain from 7.5 to 20 estimations. Green dash lines indicate 3-sigma control limits.
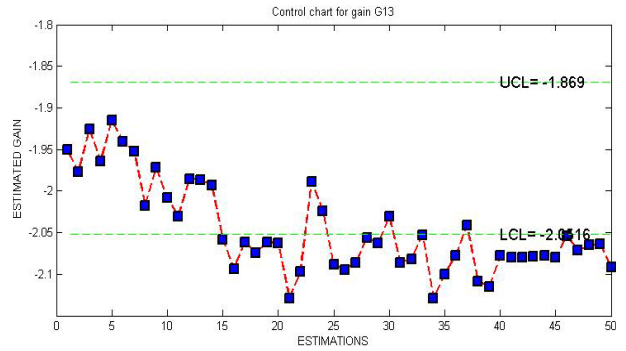
Figure 4: G_{13}-estimator response to a ramp change augmenting 5% the gain from 7.5 to 20 estimations. Green dash lines indicate 3-sigma control limits.

Though the specific changes are successfully detected – this is clearly observed in Figs. 3 to 5 – Figure 2 revels certain tendency for several individual estimators to present a bias. This exposes a problem to be analyzed in future works.
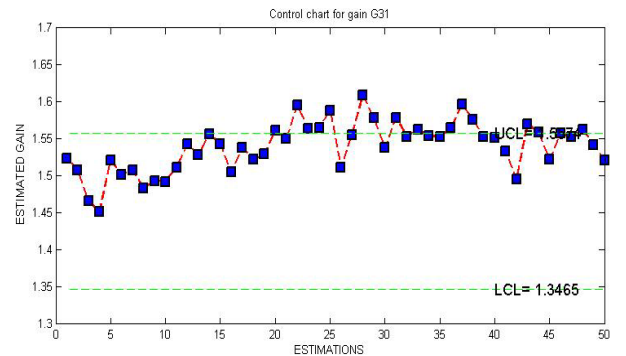
Figure 5: G_{31}-estimator response to a ramp change augmenting 5% the gain from 10 to 12.5 estimation time. Green dash lines indicate 3-sigma control limits.

5. CONCLUSIONS AND FUTURE WORK

The methodology presented here for online following of process gains shows promising results. However, several problems were highlighted by the presented analysis: i) slightly biased distribution functions with shapes doubtfully normal prevent from determining control limits associated to accurate detection rates or

false alarm frequencies; ii) several parameters used by the subspace identification method need to be optimized accordingly with the objectives; iii) the minimum necessary intensity of the persistent excitation when dealing with complex multivariable process systems is a pending duty to be face from a practical point of view; iv) the extension to closed loop identification appears as a mandatory alternative to be analyzed and developed.

APPENDIX A: Notation

The notation used in this paper follows that commonly used in the extensive literature available about subspace identification methods [Refs]. For instance, the block Hankel matrix of a single input signal is defined and written as

$$
U_{0|2i-1} \triangleq \left(\begin{array}{ccccc} u_0 & u_1 & u_2 & ... & u_{j-1} \\ u_1 & u_2 & u_3 & ... & u_j \\ ... & ... & ... & ... & ... \\ u_{i-1} & u_i & u_{i+1} & ... & u_{i+j-2} \\ \hline u_i & u_{i+1} & u_{i+2} & & u_{i+j-1} \\ u_{i+1} & u_{i+2} & u_{i+3} & & u_{i+j} \\ & & & & \\ u_{2i-1} & u_{2i} & u_{2i+1} & & u_{2i+j-2} \end{array} \right)
$$

$$
= \left(\frac{U_{0|i-1}}{U_{i|2i-1}} \right) = \left(\frac{U_p}{U_f} \right) = \left(\frac{U_{0|i}}{U_{i+1|2i-1}} \right) = \left(\frac{U_p^+}{U_f^-} \right)
$$

where the indexes i and j are such that $i \geq n$ and $j \gg n$, and n is the assumed order of the system. Note that indexes p and f stand for "past" and "future" respectively; this is becouse each column of matrix U_p composes of i elements previous to the following i elements being part of U_f. This matrix structure is also extended to the block Hankel matrix for a single output variable, in this case using the notation $Y_{0|2i-1}, Y_p, Y_f$, Y_p^+ and Y_p^-. The above notation serves also to define the block Hankel matrix of past input and output data, i.e.,

$$
W_p \triangleq \left(\begin{array}{c} U_p \\ Y_p \end{array} \right).
$$

In a similar way, the state sequence

$$
X_i = \left(\begin{array}{cccc} x_i & x_{i+1} & \cdots & x_{i+j-2} \quad x_{i+j-1} \end{array} \right) \in \mathbb{R}^{n \times j},
$$

is partitioned by setting $X_p = X_0$ and $X_f = X_i$.

The subspace identification algorithm used in this paper needs of other two important matrices: the observability matrix $\Gamma_i \in \mathbb{R}^{li \times n}$,

$$
\Gamma_i = \left(\begin{array}{c} C \\ CA \\ CA^2 \\ \vdots \\ CA^{i-1} \end{array} \right)
$$

and two low-triangular block Toeplitz, $H_i^d \in \mathbb{R}^{li \times mi}$ and $H_i^s \in \mathbb{R}^{li \times mi}$, which are written as follows:

$$
H_i^d = \left(\begin{array}{ccccc} D & 0 & 0 & \cdots & 0 \\ CB & D & 0 & \cdots & 0 \\ CAB & CB & D & \cdots & 0 \\ ... & ... & ... & ... & ... \\ CA^{i-2}B & CA^{i-3}B & CA^{i-4}B & \cdots & D \end{array} \right)
$$

and

$$
H_i^s = \left(\begin{array}{ccccc} I & 0 & 0 & \cdots & 0 \\ CK & I & 0 & \cdots & 0 \\ CAK & CK & I & \cdots & 0 \\ ... & ... & ... & ... & ... \\ CA^{i-2}K & CA^{i-3}K & CA^{i-4}K & \cdots & I \end{array} \right)
$$

Assuming the pair $\{A, C\}$ is observable and $i \geq n$, then Γ_i is a full column range matrix, i.e., $Rank(\Gamma_i) = n$.

Finally, the notation A/B stands for the orthogonal proyection of the row space of A into the row space of B,

$$
A/B \triangleq A \Pi_B
$$

$$
= A B^T \left(B B^T \right)^\dagger B
$$

Besides, $A/_C B$ denotes the oblique proyection of the row space of A on the direction of the row space of C into the row space of B,

$$
A/_B C \triangleq \left[A/B^\perp \right] \left[C/B^\perp \right]^\dagger C
$$

APPENDIX B: Confidence interval for a single estimator.

Assume θ is a parameter to be estimated using the estimator $\hat{\theta}$ whose random behavior is described by the probability density function $f(\hat{\theta})$. Then, the probability of $\hat{\theta}$ for being lower or grater than a lower and upper limit (*UL* and *LL*) respectively, can be defined as follows:

$$
\alpha/2 = \int_{-\infty}^{LL} f(\hat{\theta}) d\hat{\theta} = P(\hat{\theta} < LL)
$$

$$
1 - \alpha/2 = \int_{-\infty}^{UL} f(\hat{\theta}) d\hat{\theta} = P(\hat{\theta} \leq UL)
$$

This is equivalent to say that the probability for $\hat{\theta}$ to fall inside the given interval is given by

$$
P\left\{ LL \leq \hat{\theta} \leq UL \right\} = 1 - \alpha
$$

For convenience, the estimator is normalized to have cero mean and unit variance, then rewriting the above expression we can arrive to

$$
P\left\{ -d_{(\alpha/2)} \leq \frac{\hat{\theta} - E\{\hat{\theta}\}}{\sigma_{\hat{\theta}}} \leq d_{(1-\alpha/2)} \right\} = 1 - \alpha \qquad (B1)
$$

where the quantities

$$d_{\alpha/2} = \frac{E\{\hat{\theta}\} - LL}{\sigma_{\hat{\theta}}} \qquad d_{1-\alpha/2} = \frac{UL - E\{\hat{\theta}\}}{\sigma_{\hat{\theta}}}$$

are the "$\alpha/2$ percent points" of the normalized distribution. Note that these quantities are completely defined by α and the normalized distribution function (with 0 mean and unit deviation). Let us analyze now the general case in which we deal with a biased estimator, i.e.,

$$E\{\hat{\theta}\} = \theta + b, \quad |b| \ge 0 . \qquad (B2)$$

where b is the bias. Note that if for a given estimator $\hat{\theta}$, $b = 0$ can not be analytically demonstrated, the true population parameter θ can never be accurately estimated becouse the true bias b remains also uncertain. Then, in this case, all what we can do is to work in terms of an assumed, or expected, value θ_0, so we write

$$E\{\hat{\theta}\} = \theta_0 + b_0, \quad |b_0| \ge 0 . \qquad (B3)$$

where all the uncertainty has been sent into the bias b_0. The above reasoning tells that through $\hat{\theta}$ we can only monitor the parameters characterizing the population of $\hat{\theta}$ values, and detect operating conditions producing significant changes in the observed values. This is equivalent to following parameter changes by using the bias estimation $\hat{b}_0 = \hat{\theta} - \theta_0$ referred to the specific value θ_0. In other words, two equivalent confidence intervals can be written, one directly in terms of the parameter estimator, and the other in terms of the bias estimator, as follows:

$$LL_{\hat{\theta}} = \hat{\theta} - d_{(1-\alpha/2)}\hat{\sigma}_{\hat{\theta}} \le E\{\hat{\theta}\} \le \hat{\theta} + d_{\alpha/2}\hat{\sigma}_{\hat{\theta}} = UL_{\hat{\theta}} \qquad (B4)$$

$$LL_{\hat{b}} = \hat{b}_0 - d_{(1-\alpha/2)}\hat{\sigma}_{\hat{\theta}} \le E\{\hat{b}_0\} \le \hat{b}_0 + d_{\alpha/2}\hat{\sigma}_{\hat{\theta}} = UL_{\hat{b}} , \qquad (B5)$$

both defined for a $(1-\alpha)$% confidence level.

REFERENCES

Crosier, R.B., 1988. Multivariate generations of cumulative sum quality control schemes, *Technometrics*, 30, pp. 219-303.

Favoreel W., De Moor B. and Van Overschee P., 2000. Subspace state space system identification for industrial processes', *Journal of Process Control*, 10, pp. 149-155.

Gaikwad, S.V. and D.E. Rivera, 1996. Control-relevant input signal design for multivariable system identification. Application to high-purity distillation, in Proceedings of the 13th IFAC World Congress, pp. 349-354.

Hotelling, H., 1947. Multivariate quality control, in Techniques of Statistical Analysis. Eds. C. Eisenhart, M. Hastay and Wallis, W. A., McGraw-Hill.

Katayama, T., 2005. Subspace methods for system identification, Springer.

Larimore, W.E., 1990. Canonical variate analysis in identification, filtering and adaptive control, Proceedings of the 29th Conference on Decision and Control, pp. 596–604.

Lowry, C.A., Woodall, W.H., Champ, C.W. and Rigdon, S.E. 1992. A Multivariate exponentially weighted moving average control chart, *Technometrics*, 34, pp. 46-53.

Nikandrov, A. and Swartz, C.L.E., 2009. Sensitivity analysis of LP-MPC cascade control systems, *Journal of Process Control*, 19 pp. 16-24.

Overschee, P.V. and Moor, B.D., 1994. N4SID: Subspace algorithms for the identification of combined deterministic-stochastic systems, *Special Issue on Statistical Signal Processing and Control, Automatica*, 30, 1, pp. 75-93.

Pignatello, J. J. and Runger, G. C., 1990. Comparisons of multivariate CUSUM charts. *Journal of Quality Technology*. 22, pp. 173- 186.

Van Overschee P. and De Moor, B., 1996. Subspace Identification for linear systems: theory, implementation, applications. Kluwer Academic Publishers, 254 p.

Verhaegen, M. and P. Dewilde, 1992. Subspace model identification, part I: The output-error state-space model identification class of algorithms. *International Journal of Control*, 56, pp. 1187–1210.

Woodwall, W.H. and Ncube, M.M., 1985. Multivariate CUSUM quality control process. *Technometrics*, 27, pp. 285-292.

Ying, C.M. and Joseph, B., 1999. Performance and stability analysis of LP-MPC and QPMPC cascade control systems, *AIChE J*. 45 pp. 1521–1534.

STABILITY ANALYSIS OF 1 DOF HAPTIC INTERFACE: TIME DELAY AND VIBRATION MODES EFFECTS

Q. V. Dang[a], A. Dequidt[b], L. Vermeiren[c], M. Dambrine[d]

[a,b,c,d] Univ Lille Nord de France, F-59000 Lille, France
[a,b,c,d] UVHC, LAMIH, F-59313 Valenciennes, France
[a,b,c,d] CNRS, UMR 8530, F-59313 Valenciennes, France
[a]dqviet2212@gmail.com, [b]antoine.dequidt@univ-valenciennes.fr,
[c]laurent.vermeiren@univ-valenciennes.fr, [d]michel.dambrine@univ-valenciennes.fr

ABSTRACT
Stability analysis continues to be a major challenge in the haptic field. This paper addresses the issue of stability for haptic interfaces in contact with a virtual wall. The feedback contact force is calculated from the impedance-based virtual environment model including a linear stiffness and a damping. A mechanical model includes two vibration modes used to characterize the overall device dynamics. Their effects on the stability and those of the time delay are examined. The stability boundary is derived from a sampled data model of the haptic interface by the gain margin method.

Keywords: Haptic device, Stability region, Time delay, Vibration mode

1. INTRODUCTION
Haptic system includes the human operator and the haptic interface, c.f. Fig.1. Haptic interface is the robotic system that allows human operator to manipulate objects in a Virtual Environment (VE) by a haptic device with the sense of touch and the kinesthetic perception.
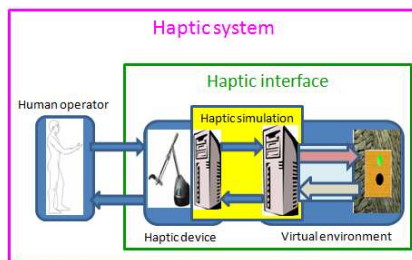
Figure 1: Parts of a haptic system

However, unlike robot manipulators, the haptic interfaces should have both high force rendering and maximum dynamic transparency (low inertia, low friction, etc.). Therefore, the stability of a haptic interface is a key issue. Any unstable behavior arising during the dynamic interaction can damage the haptic device or injure the human operator.

Nowadays, there are two main approaches to study the stability of a haptic system. In the first approach (Gil and al. 2004; Diolaiti and al. 2006; Hulin and al. 2006; Hulin and al. 2008), the haptic system is described by a sampled-data model; the VE is seen as a virtual impedance characterized by a linear spring and a damper. The stability region is represented by the stability boundaries in the stiffness-damping plan of the virtual impedance. The influence of parameters such as physical damping and time delay on the stability boundaries is evaluated through the transfer function. The human hand impedance can be included in the model to study the human dynamic effects. In the second approach, the two-port network theory is applied to separate the virtual environment and the human operator into two independent parts (Colgate and Brown 1994; Colgate and Schenkel 1997; Adams and Hannaford 1998; Adams and al. 1998; Adams and Hannaford 1999; Adams and Hannaford 2002). Then, the virtual coupling is used and the stability is ensured through the passivity or unconditional stability criteria.

There are a lot of works concerning the interaction of the human operator, the haptic device and the virtual environment. However, there still remain many challenges such as expanding the boundary of stability region, modeling the human operator, designing the control law, etc. In previous studies, the haptic device was modeled as a mechanical system including a mass and a physical damper. The existence of vibration modes in the haptic device has been studied recently (Díaz and Gil 2008; Díaz and Gil 2008), but these works have not fully explained the influence of the vibration mode parameters.

The contribution of this paper consists in analyzing the mechanical frequency effects of the haptic device's vibration modes and the time delay influence on the interface stability based on the first approach. The results are carried out on the physical parameters of the PHANTOM® haptic device.

2. HAPTIC DEVICE MODELING
In previous studies, a rigid model of haptic device is used to analyze the stability. In this paper, the elastic deformations are considered in the mechanical transmission between motor and human hand. Consequently, the resulting vibration modes can modify

the stability boundaries. That's why a more complete model of haptic device with two most significant vibration modes is shown in fig.2. These vibration modes result from 1) the elastic joint between the motor (position X_1) and the link shaft (position X_2) and 2) the flexible link between the shaft and the tip (position X_3). The motor applies a force F_1 and the human hand laid on the link tip applies a force F_h. With the mechanical parameters pointed out on fig.2, the linear model can be derived in Laplace domain as:

$$\left(\mathbf{M}s^2 + \mathbf{B}s + \mathbf{K}\right)\mathbf{X}(s) = \mathbf{F}(s) \tag{1}$$

with the position and force vector and the mass, damping and stiffness matrices, respectively:

$$\mathbf{X}(s) = \left[\mathrm{X}_1(s)\,\mathrm{X}_2(s)\,\mathrm{X}_3(s)\right]^T \tag{2}$$

$$\mathbf{F}(s) = \left[\mathrm{F}_1(s)\,F_h(s)\right]^T \tag{3}$$

$$\mathbf{M}(s) = diag(m_1, m_2, m_h) \tag{4}$$

$$\mathbf{B}(s) = \begin{bmatrix} b_{c1} + b_1 & -b_{c1} & 0 \\ -b_{c1} & b_{c1} + b_{c2} + b_2 & -b_{c2} \\ 0 & -b_{c2} & b_{c2} + b_3 \end{bmatrix} \tag{5}$$

$$\mathbf{K}(s) = \begin{bmatrix} k_{c1} & -k_{c1} & 0 \\ -k_{c1} & k_{c1} + k_{c2} & -k_{c2} \\ 0 & -k_{c2} & k_{c2} \end{bmatrix} \tag{6}$$

The relationship between output positions and input forces is presented on the form of a transfer function matrix:

$$\mathbf{X}(s) = \mathbf{G}(s)\,\mathbf{F}(s) \tag{7}$$

with:

$$\mathbf{G}(s) = \begin{bmatrix} \mathrm{G}_{11}(s) & \mathrm{G}_{1h}(s) \\ \mathrm{G}_{21}(s) & \mathrm{G}_{2h}(s) \\ \mathrm{G}_{31}(s) & \mathrm{G}_{3h}(s) \end{bmatrix} = \left(\mathbf{M}s^2 + \mathbf{B}s + \mathbf{K}\right)^{-1} \tag{8}$$

Consequently, haptic device's dynamic characteristic is described by six transfer functions $G_{11}(s)$, $G_{1h}(s)$, $G_{21}(s)$, $G_{2h}(s)$, $G_{31}(s)$, $G_{3h}(s)$. These functions have the same denominator. The transfer function $G_{11}(s)$ can be rewritten as following:

$$G_{11}(s) = G_r(s).G_{f1}(s).G_{f2}(s)$$

$$G_r(s) = \frac{1}{m.s^2 + b.s}$$

$$G_{f1}(s) = \frac{w_{r1}^2.(s^2 + c_{a1}.s + w_{a1}^2)}{w_{a1}^2.(s^2 + c_{r1}.s + w_{r1}^2)} \tag{9}$$

$$G_{f2}(s) = \frac{w_{r2}^2.(s^2 + c_{a2}.s + w_{a2}^2)}{w_{a2}^2.(s^2 + c_{r2}.s + w_{r2}^2)}$$

where m and c are the total mass and the total viscous friction parameters of the device, w_{ri} and c_{ri} ($i = 1,2$)

are the parameters of the ith resonant mode, w_{ai} and c_{ai} are the parameters of the ith antiresonant mode.
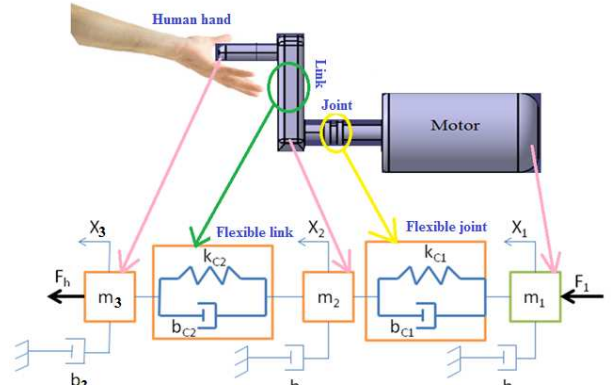
Figure 2: Haptic device model with two vibration modes

The resonant and anti-resonant frequencies are indexed by r and a, respectively. $G_r(s)$, $G_{f1}(s)$ and $G_{f2}(s)$ are the transfer functions of the rigid mode, the first vibration mode and the second vibration mode of the haptic device, respectively. The physical parameters of the PHANTOM haptic device are given in Table 1.

Table 1 (Gil and al. 2010): Physical parameters of the PHANTOM

Parameter	Variable	Value
First vibration mode		
Inertia	m	1.168 gm^2
Physical damping	b	0.0054 Nms/rad
Anti-resonant frequency	w_{a1}	417.612 rad/s
Damping coefficient	c_{a1}	80 Nms/rad
Resonant frequency	w_{r1}	479.166 rad/s
Damping coefficient	c_{r1}	83 Nms/rad
Second vibration mode		
Anti-resonant frequency	w_{a2}	546.626 rad/s
Damping coefficient	c_{a2}	90 Nms/rad
Resonant frequency	w_{r2}	1159.31 rad/s
Damping coefficient	c_{r2}	352 Nms/rad

3. METHODOLOGY

In order to study the stability of haptic interface, a dynamic model including different parts is established. The more critical case for haptic interface stability occurs when the manipulated object is in contact with a virtual wall. Impedance-based virtual environment for constrained motion is a coupling system including linear spring with virtual stiffness (K) and a virtual damping coefficient (B). The actuator position X_1 is measured by an encoder. The resulting force F_1 is calculated from this impedance model. It is assumed that some nonlinear phenomena (like sensor resolution, actuator saturations, etc.) are negligible. In particular, dry friction and sensor quantization can be ignored (Gil

and al. 2007). The sampling and the time delay due to the computer part (controller, virtual environment computation, and communication) are included in the haptic interface model. With the sampling process and the zero-order holder (ZOH), the overall model is a hybrid one with both continuous and discrete form, cf. Fig.3.
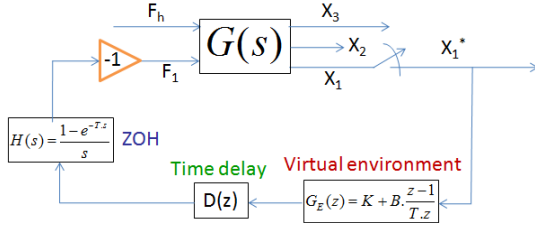

Figure 3: Dynamic model of haptic interface

From (2), the output position X_1 can be written:

$$X_1(s) = G_{11}(s).F_1(s) + G_{1h}(s).F_h(s) \qquad (10)$$

Combining to the dynamic model in Fig.3, the discrete form of the sampled position $X_1^*(z)$ becomes:

$$X_1^*(z) = \frac{Z\left[G_{1h}(s).F_h(s)\right]}{1+D(z).G_E(z).Z\left[G_{11}(s).H(s)\right]} \qquad (11)$$

where Z[.] denotes the Z-transform operation of a transfer function within brackets, and * designed a discrete-time signal. The virtual interface $G_E(z)$, the ZOH H(s) and the time delay D(z) are formulated as:

$$G_E(z) = K+B.\frac{z\text{-}1}{T.z}, \ H(s) = \frac{1\text{-}e^{\text{-}T.s}}{s}, \ D(z) = z^{\text{-}d} \qquad (12)$$

Notice that T is the sampling period and d is the time delay ratio ($d=T_d/T$, with T_d the total time delay). The haptic interface stability depends on the following characteristic equation:

$$1+D(z).G_E(z).Z\left[G_{11}(s).H(s)\right] = 0 \qquad (13)$$

Equation (13) can be rewritten as following:

$$1+K.\frac{D(z).Z\left[G_{11}(s).H(s)\right]}{1+B.\dfrac{z-1}{T.z}.D(z).Z\left[G_{11}(s).H(s)\right]} = 0 \qquad (14)$$

Therefore, the stability boundary obtained by calculating the gain margin $K_{critical}$ of the transfer function is defined by (Gil and al. 2007):

$$K_{critical} = Gm\left\{\frac{D(z).Z\left[G_{11}(s).H(s)\right]}{1+B.\dfrac{z-1}{T.z}.D(z).Z\left[G_{11}(s).H(s)\right]}\right\} \qquad (15)$$

where Gm{.} represents the gain margin of the transfer function within brackets. Matlab® is used to compute the gain margin. This method is very useful to reconstruct the stability boundaries of haptic interfaces even if their transfer functions are very complex.

In the following parts, the gain margin method is used to find the stability boundaries for different values of the time delay and the anti-resonant frequency. The presented results are based on the physical parameters of the PHANTOM® (cf. Table 1).

4. TIME DELAY INFLUENCE

Stability boundaries of the haptic interface model that includes two vibration modes can be found for different time-delay values by calculating (15) over a range of virtual damping values. The results for some time delay ratio values ($d=0, 1, 2, 4, 8$) are shown in fig.4, with a sampling period equals to T = 1 ms.
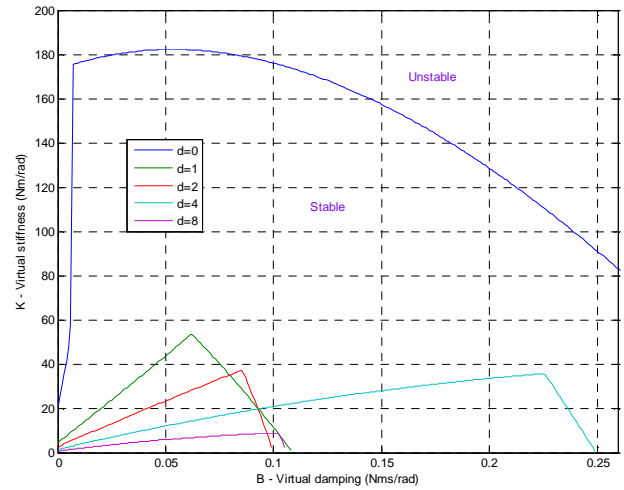
Figure 4: Stability boundaries of haptic interface model with two vibration modes for different time delay values

Let us denote $K_{max}(B,d)$ the stability boundary that is the maximal value of K for which the system is stable for a given virtual damping coefficient B and a given time delay d. As it can be intuitively guessed, the maximal value $K^*(d) = \max_B K_{max}(B,d)$ is a decreasing function of the time-delay value. This fact can be explained by the reduction of the phase crossover frequency (frequency at which the phase is -180°) in the Bode diagram of $Z[G_{11}(s).H(s)].D(z)$, cf. Fig.5. However, the stability region doesn't always decrease for rising time delay value. It is interesting to find that the stability region with a time delay of 4 ms is larger along the B axis than the stability regions with 2ms and 8ms. This behavior is explained by introducing the notion of *critical frequency*, denoted w_{cri}, which is defined as the maximal value of the phase crossover frequency. Obviously, the higher the critical frequency is, the larger the stability region is, cf. Fig.4 & Fig.6. These results are completely consistent with the experimental results shown by Gil and al. 2010, where

the critical frequency values of the PHANTOM were experimentally obtained by using the relay method.
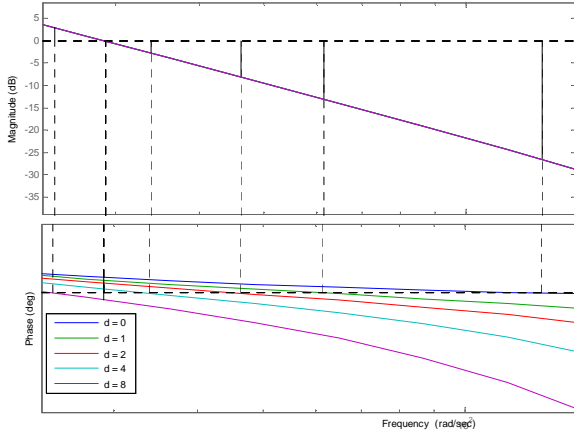
Figure 5: Bode diagram of $Z[G_{11}(s).H(s)].D(z)$ for different time delay values
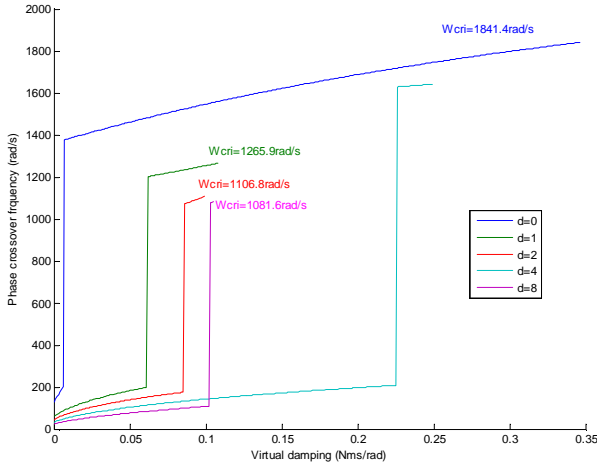
Figure 6: Phase crossover frequencies for different time delay values

5. ANTI-RESONANT FREQUENCY EFFECT

The linear model parameters of the PHANTOM haptic device including the two most significant vibration modes were given in the Table 1. The theoretical Bode diagram of this device is shown in fig.7.
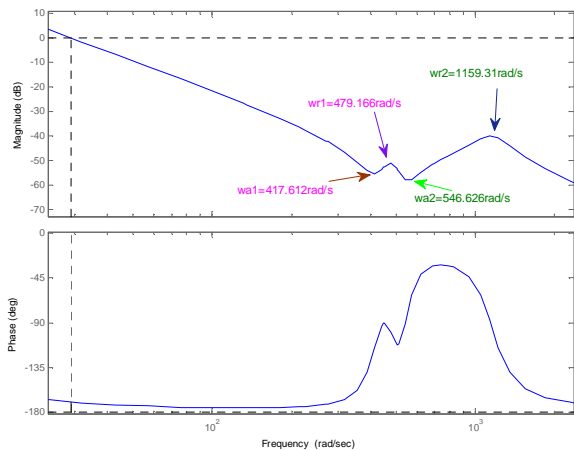

Figure 7: Theoretical Bode diagram of the PHANTOM

In the next sections, the anti-resonant frequency effects of each vibration mode are examined. Obviously the other physical parameters are considered to be constant.

5.1. Effect of the first vibration mode

In this part, the anti-resonant frequency effect of the first vibration mode on the haptic interface stability is examined. This frequency has to satisfy the condition (16), which allows assuring the distinction of two vibration modes as seen in fig.7. The others parameters are given in Table 1.

$$0 < w_{a1} < w_{r1} = 479.166 \text{ rad/s} \qquad (16)$$

The figure 8 shows the stability boundaries of the haptic interface with the time delay ratio $d = 1$ ($T_d = T = 1$ms) for some of different anti-resonant frequency values of the first vibration mode. The stability region shape changes when the anti-resonant frequency rises in the range of the condition (16). From the Bode diagram of the transfer function $Z[G_{11}(s).H(s)].z^{-1}$, as seen in fig.9, it is obvious that the phase crossover frequency changes from the right side to the left side of the first vibration mode.
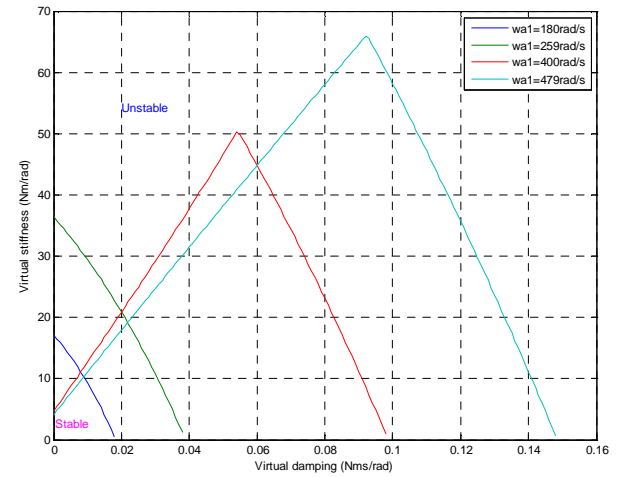
Figure 8: Stability boundaries for different values of w_{a1}
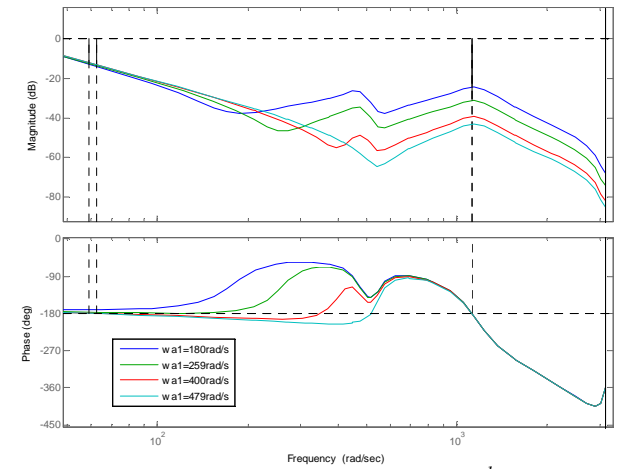
Figure 9: Bode diagram of $Z[G_{11}(s).H(s)].z^{-1}$ for different values of w_{a1}

These studies shown out the existence of a transition frequency at which there is a qualitative modification of the phase crossover frequency in the Bode diagram, and so that, it changes the stability region shape. As seen in Figs. 10 and 11, this transition frequency, denoted w_{at1}, is in the range of $264\text{rad/s} < w_{at1} < 265\text{rad/s}$. For $w_{a1} \leq 264\text{rad/s}$, the stability region increases when the anti-resonant frequency rises in this range; the phase crossover frequency is located on the right side of the first vibration mode and depends very slightly on w_{a1}. For $265\text{rad/s} \leq w_{at1} < 479.166\text{rad/s}$, the phase crossover frequency jumps to the left side of the first vibration mode; the stability region has a reverse V-shape including two parts: the part 1 corresponding with small values of B, the stability region decreases along the K-axe, the part 2 corresponding with higher values of B, the stability region increases along the B-axe when the anti-resonant frequency rises in its range.
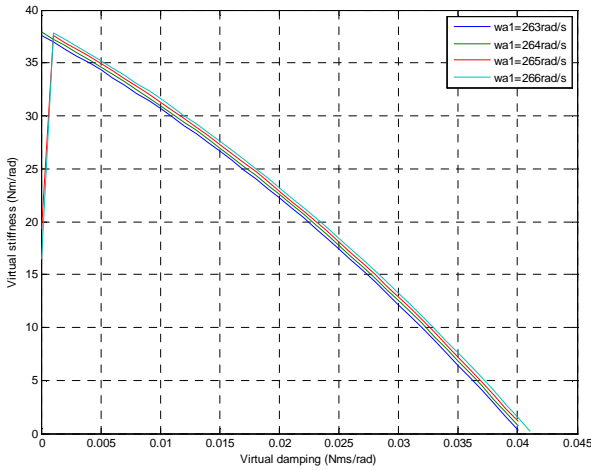
Figure 10: Stability boundaries for different values of w_{a1} around the transition frequency w_{at1}.
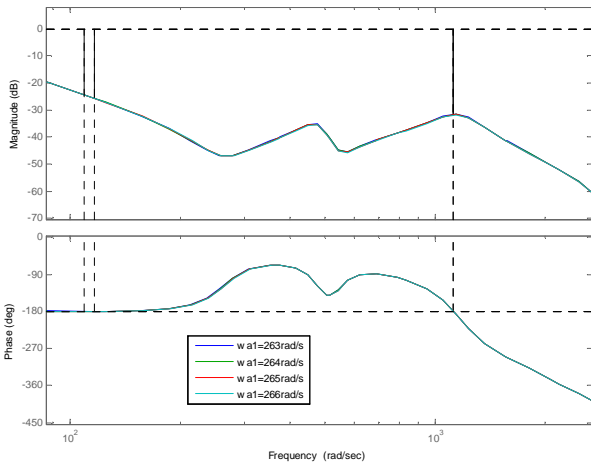
Figure 11: Bode diagram of $Z[G_{11}(s).H(s)].z^{-1}$ for different values of w_{a1} around the transition frequency w_{at1}.

5.2. Second vibration mode effect
The anti-resonant frequency effect of the second vibration mode on the haptic interface stability is studied with the condition (17) in order to assure the distinction of two vibration modes as seen in Fig. 7. The others parameters can be found in Table 1.

$$w_{r1} = 479.166\text{rad/s} < w_{a2} < w_{r2} = 1159.31\text{rad/s} \qquad (17)$$

As shown in Figs. 12-13, although the phase crossover frequencies in the Bode diagram of the transfer function $Z[G_{11}(s).H(s)].z^{-1}$ are always located on the left side of two vibration modes, the stability regions of the haptic interface with the time delay ratio $d=1$ ($T_d=T=1\text{ms}$) for some of anti-resonant frequency values of the second vibration mode are very different shapes.
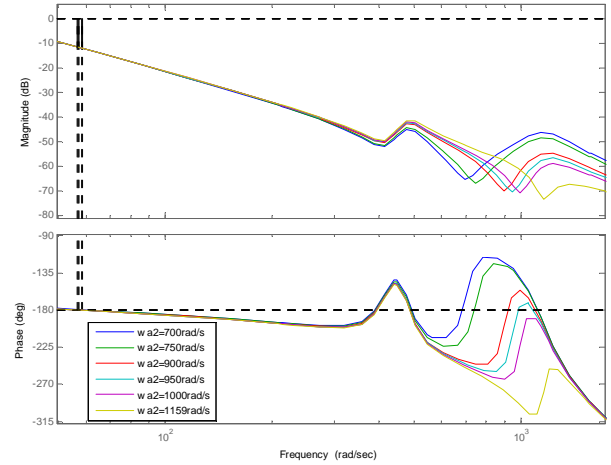
Figure 12: Bode diagram of $Z[G_{11}(s).H(s)].z^{-1}$ for different values of w_{a2}
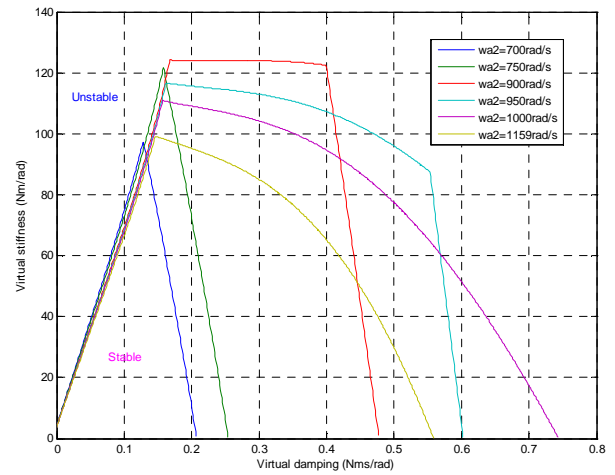
Figure 13: Stability boundaries for different values of w_{a2}

Obviously, for small values of B, the stability region is most linear and decreases gradually along the K-axe with the increasing values of w_{a2} (cf. Fig. 14). It is noteworthy that this linear region can be explained by the Bode diagram of the transfer function $Z[G_{11}(s).H(s)].z^{-1}$. The larger linear stability region corresponds with the bigger value of the phase crossover frequency (cf. Fig. 15).
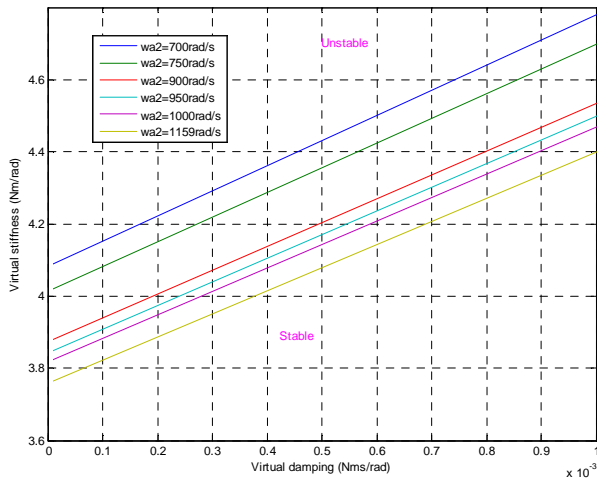
Figure 14: Partial view of the stability regions for different values of w_{a2}

6. CONCLUSIONS

This paper has introduced an overview on the stability analysis with respect to the time delay and the two first vibration modes of the haptic interface model. The impedance model has been used to compute the force feedback including a virtual stiffness and a virtual damping. The stability boundaries of both virtual parameters have been derived from a sampled-data model with the gain margin method. The stability analysis is based on the linear model with two vibration
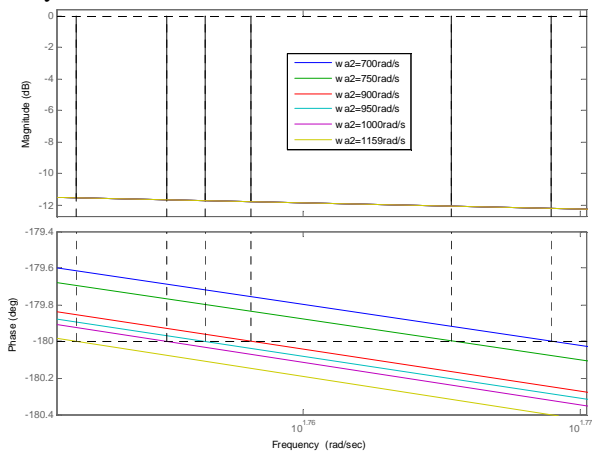
Figure 15: Partial view of the Bode diagram of $Z[G_{11}(s).H(s)].z^{-1}$ for different values of w_{a2} (the left side of two vibration modes)

modes of a haptic device and the effects on stability of each mode have been detailed. The presented results allow us to conclude that the haptic device's vibration modes have an intricate effect on the interface stability. The analysis of the anti-resonant frequency effect leads to explanations about qualitative modification of stability boundaries. These results may provide some basic engineering guidelines in choosing design specifications of a haptic system.

Haptic devices have stability limitations that can be overcome by incorporing feedback controllers or

signals shapers (Kuchenbecker 2006). While the uncertainties on model parameters is a key issue for flexible mechanism (Dieulot and Colas 2009), robust control design can play an important role in haptic interface design.

REFERENCES

Adams, R.J., Hannaford, B., 1998. A Two-Port Framework for the Design of Unconditionally Stable Haptic Interfaces. *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, Victoria, BC.

Adams, R.J., Moreyra, M.R., Hannaford, B., 1998. Stability and Performance of Haptic Displays: Theory and Experiments. *Proceedings of the ASME International Mechanical Engineering Congress and Exhibition*, Anaheim, CA.

Adams, R.J., Hannaford, B., 1999. Stable Haptic Interaction with Virtual Environments. *IEEE Trans. on Robotics and Automation*, 15(3), 465–474.

Adams, R.J., Hannaford, B., 2002. Control Law Design for Haptic Interfaces to Virtual Reality. *IEEE Trans. on Control Systems Technology*, 10(1), 3–13.

Colgate, J.E, Brown, J.M., 1994. Factors affecting the Z-Width of a Haptic Display. *Proceedings of the IEEE International Conference on Robotics and Automation*, May, San Diego, CA.

Colgate, J.E., Schenkel, G., 1997. Passivity of a class of sampled-data systems: Application to haptic interfaces. *Journal of Robotic Systems*, 14(1), 37–47.

Diolaiti, N., Niemeyer, G., Barbagli, F., Salisbury, J. K., 2006. Stability of haptic rendering: Discretization, quantization, time-delay and Coulomb effects. *IEEE Trans. on Robotics*, 22(2), 256–268.

Díaz, I., Gil, J.J., 2008. Influence of Internal Vibration Modes on the Stability of Haptic Rendering. *Proceedings of the IEEE International Conference on Robotics and Automation*, May, Pasadena, CA.

Díaz, I., Gil, J.J., 2010. Influence of Vibration Modes and Human Operator on the Stability of Haptic Rendering. *IEEE Trans. on Robotics*, 26(1), 160–165.

Dieulot J.-Y., Colas F., 2009. Robust PID control of a linear mechanical axis: A case study. *Mechatronics* 19(2), 269-273

Gil, J.J, Avello, A., Rubio, Á., Flórez, J., 2004. Stability analysis of a 1 DOF haptic interface using the Routh–Hurwitz criterion. *IEEE Trans. on Control Systems Technology*, 12(4), 538–588.

Gil, J.J., Sanchez, E., Hulin, T., Preusche, C., Hirzinger, G., 2007. Stability boundary for haptic rendering: Influence of damping and delay. *Proceedings of the IEEE International Conference on Robotics and Automation*, Roma (Italy).

Gil, J.J., Puerto, M.J., Díaz, I., Sánchez, E., 2010. On the Z-Width Limitation due to the Vibration Modes of Haptic Interfaces. *IEEE Trans. on Robotics*, 26(1), 160-165.

Hulin, T., Preusche, C., Hirzinger, G., 2006. Stability boundary for haptic rendering: Influence of physical damping. *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems,* October 9-15, Beijing (China).

Hulin, T., Preusche, C., Hirzinger, G., 2008. Stability boundary for haptic rendering: Influence of human operator. *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems,* September 22-26.

Kuchenbecker K. J., 2006.Characterizing and controlling the high-frequency dynamics of haptic interfaces, Thesis (PhD). Stanford University.

BOND GRAPH MODELING OF REVERSE-FLOW MONOLITH REACTOR FOR AIR POLLUTANT ABATEMENT

A. Riaud [(a)], S. Paul [(a,b,c)], G. Dauphin-Tanguy [(a,b,d)]

[(a)] Ecole Centrale de Lille, F-59651 Villeneuve d'Ascq Cedex, France.
[(b)] Univ. Lille Nord de France, F-59000, Lille, France.
[(c)] CNRS UMR8181, UCCS, F-59655 Villeneuve d'Ascq Cedex, France.
[(d)] CNRS FRE 3303, LAGIS, Ecole Centrale de Lille, F-59651 Villeneuve d'Ascq Cedex, France.

[(a)]antoine.riaud@centraliens-lille.org, [(b)]sebastien.paul@ec-lille.fr, [(c)]genevieve.dauphin-tanguy@ec-lille.fr

ABSTRACT
A bond graph model of a reverse-flow monolith reactor has been developed. It aims at offering a tool for the design of such devices, optimizing control strategies and generating code for this control. The construction of the model is detailed and compared with previous results issued from the literature. The influence of some driving parameters on the dynamic behavior of the system and the pollutant abatement is discussed from simulation results.

Keywords: bond-graph, reverse-flow monolith reactor, combustion, VOC, pollutant abatement

1. INTRODUCTION
The Volatile Organic Compounds (VOC) emission in the atmosphere by industrial processes (as for instance by a painting workshop of the car industry) is currently a major environmental issue. As a matter of fact, these compounds cause respiratory disorders, greenhouse effect and are even sometimes carcinogens. Therefore, highly efficient low cost solutions need to be implemented for the abatement of these air pollutants. The most attractive end-of-the-pipe treatments involve thermal or catalytic oxidation of these compounds. In the latter case the use of a solid catalyst allows working in the continuous mode at relatively moderate temperature. However, even if total oxidation reactions are exothermic, conventional combustion reactors are energy demanding particularly in the case of the treatment of highly diluted effluents. Therefore energy consumption has to be reduced by recovering the heat released during combustion (process integration). Reverse-Flow Reactors (RFR) have proved to be efficient in such a case even if the dilution of the VOC is high and if the temperature of the feed is low (Van de Beld and Westerterp 1994a, Van de Beld et al 1994b). To decrease the pressure drop induced by the catalytic bed, monolithic support with a so-called honeycomb structure which walls are coated with the catalyst are often used to pack this kind of reactor (Matros and Bunimovich 1996). Basically, the RFR is a fixed-bed reactor where the direction of the feed is alternatively switched from one way to another. In this case the heat of the reaction is partially stored in the monolith and then restored to heat the feed to be treated. When the pseudo steady-state is reached the RFR is sometimes capable to work auto-thermally, which is, of course, the most efficient way to carry out pollutants abatement.

Obviously, the temperature and the concentrations profiles inside the reactor change continually because of the intrinsic unsteady-state of this process. To calculate these profiles, RFR are generally modeled by a set of differential-algebraic equations corresponding to the non-stationary mass and energy balances (Marin et al 2009). In the present work, a Bond Graph modeling of such a process is developed in order to study the influence of the switching time on transient and pseudo steady-state reactor behavior.

Some publications concern the study of the physical parameters behavior inside such a reverse-flow reactor using detailed models. For instance, Balcaen (2011) used a 2D model to calculate temperature and pollutant distribution in the reactor. Ramdani et al (2001) developed a linear control model for short switching periods reactors (with a very small catalytic zone). However, because they are time-consuming and often too restrictive, none of these models open prospects to integrate the reactor in the overall plant design (pumps, pipes, power supply...), and to adapt control strategies during the development.

The bond graph language can bring a solution because it allows the coupling of several physical fields, simplifies the study of non-linear systems, automatically finds system variables, is easy to simplify or to make more accurate, and, through the physics, suggests sizing optimization and control strategies.

Some papers in the literature deal with bond graph models of chemical reactors such as continuously stirred tank reactors (Heny et al 1999), perfectly mixed gas autoclaves and membranes (Couenne et al 2006), batteries (Karnopp 1990, MacKenzie et al 1993, Thoma and Ould-Bouamama 1999) but, as far as we know, none of them concern directly a reverse-flow reactor.

This paper is organized as follows. After this short introduction, the second section deals with the description of the general structure of the reverse-flow reactor and the definition of the model. In the third section, after defining the variables used for the modeling in pseudo-bond graph (BG), and the chosen hypotheses, two BG models of a once-through reactor are described with an increasing complexity. In the fourth section, it is shown how to extend this model to the reverse-flow reactor. Some comparisons are made using literature results in the fifth section. Thus the model is used to study the influence on the efficiency of air pollutant abatement of some parameters as for instance the semi-cycle time, the dynamic variation of the pollutant inlet concentration and the reactor start-up.

2. REVERSE-FLOW REACTOR DESCRIPTION

2.1. Structure of the RFR
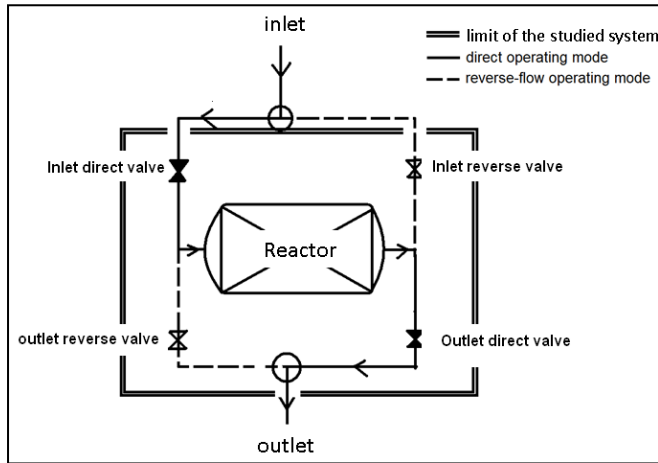
Figure 1: General Description of a Reverse-Flow Reactor

Figure 1 gives a general description of a RFR. This reactor is packed with a monolith whose walls are coated with an appropriate combustion catalyst. In the first semi-cycle the polluted air is introduced to the left and if the temperature reached on the catalyst surface is sufficient the reaction of combustion occurs. The heat released is then partly evacuated by the downstream gas and partly stocked in the monolith. Then in the second half-cycle the direction of circulation of the gas is inverted. The heat stocked in the monolith is then used to heat-up the cold polluted air entering the reactor at the right side this time.

The studied reactor is made of several tubular channels. As shown in Figure 2, each channel is composed of four layers: the polluted air in the middle, the washcoat and catalyst layers, the monolith by itself and finally an equivalent virtual insulation coating modeling the real insulation coating around the overall reactor as shown in equation (50).
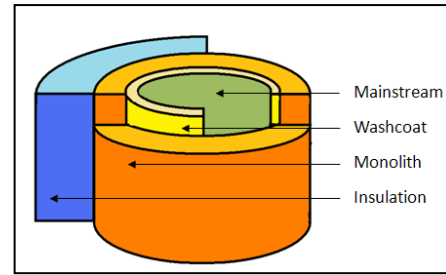
Figure 2: Cross-Section of a Single Channel of the Reactor

The proposed model represents only one channel but can be easily extended to the whole monolith. Pollutants are transported through the channel, and burn on the catalyst's surface. A part of the heat generated is released in the air, and another part is stocked in the monolith.
Air flow direction is periodically inverted; half-period is called the semi-cycle time.

2.2. Model Assumptions
For simulation purpose, the channel is axially subdivided in n cells. In the cells, the following assumptions are made:
- The temperature, pressure and composition are axially uniform
- The gas phase is considered as ideal (low pressure)
- No mass-transfer occurs from the catalyst to the monolith (airtight material)
- The combustion is a first-order reaction and follows the Arrhenius law (Ramanathan 2004)
- Each channel in a cross-section of the reactor has the same thermal profile
- Axial mass and heat transfer within the washcoat are neglected
- Axial diffusive heat and mass transfer within the channel are neglected (compared to convective heat and mass transfer)
- Thermal capacities are non-temperature-dependent
- The gas phase is only composed of O_2, N_2, propane C_3H_8 (considered as a model molecule for the pollutant), CO_2 and H_2O.
- The combustion enthalpy is non-temperature-dependent (Kirchhoff law validates this hypothesis within 1% in the considered temperature range).
- Inertial phenomena are neglected compared to pressure phenomena (Re << 2300)
- The data provided for modeling are inlet gas mass flow rate $q_{m,in}$, the inlet temperature T_{in}, the ambient air temperature T_{atm} and the inlet molar fraction x_i of each component i.
- The diameter of the channel is much smaller than the length of the reactor so the entrance

effect of fluid mechanics are neglected (Nu, Sh = Nu_∞, Sh_∞)

3. MODEL CONSTRUCTION IN CASE OF SINGLE MONO-DIRECTIONAL FLOW

3.1. Choice of Variables

Previous works in this field have shown that equations are simpler when written with molar flow rates, partial pressures, heat flow rates or enthalpy flow rates and temperature. Therefore, the proposed model is built using these variables, it is hence a pseudo-bond graph. Three types of power bonds are used:

Table 1: Bonds and state variables of the bond graph model

Field	Mass Transfer (Gas Variable Composition)	Mass Transfer (Mixture, variable Composition)	Heat Transfer (Convective)	Heat Transfer (Conducive)
Symbol	———▶	——⁵—▶	– – –▶	———▶
Flow	$\dot{n}$	$\dot{n}_i$	$\dot{H}$	$\dot{Q}$
Effort	P	P_i	T	T

The mass transfer multi-bonds carries one flow for each chemical species.

3.2. Three Cells Model

In a first step, a simple mono-directional flow is modeled. Several material layers storing matter and heat are represented by multiport C elements. Some pressure drop and convection heat transfer occurring between the layers are represented using multiport R elements.

If the feed gas has a temperature lower than the washcoat, then a heat transfer will occur upstream and the air will get hotter. Then, in the second cell, the combustion can start if the gas is hot enough. The products are released in the same cell because we assume no convection occurs close to the walls. The reaction is exothermal and a part of the heat generated by the combustion is exchanged with the flow. This hot fluid warms-up the downstream cell.

For clarity purpose, the model is presented here with only 3 cells: inlet, combustion and outlet cell. As shown in Figure 3, polluted air is considered as gas with known composition x_{i-in}, mass flow rate $q_{m,in}$ and temperature T_{in}. For the inlet cell, molar flow rate $\dot{n}_{in,i}$ and enthalpy flow rate $\dot{H}_{in}$ are calculated using equations (1) and (2).

$$\dot{n}_{in,i} = \frac{q_{m,in}}{M} x_{i,in} \tag{1}$$

$$\dot{H}_{in} = \frac{q_{m,in}}{M} C_p T_{in} \tag{2}$$

One single bond is used for the outlet pneumatic part, because the mass flow rate does not depend on the atmosphere composition but on the total pressure.
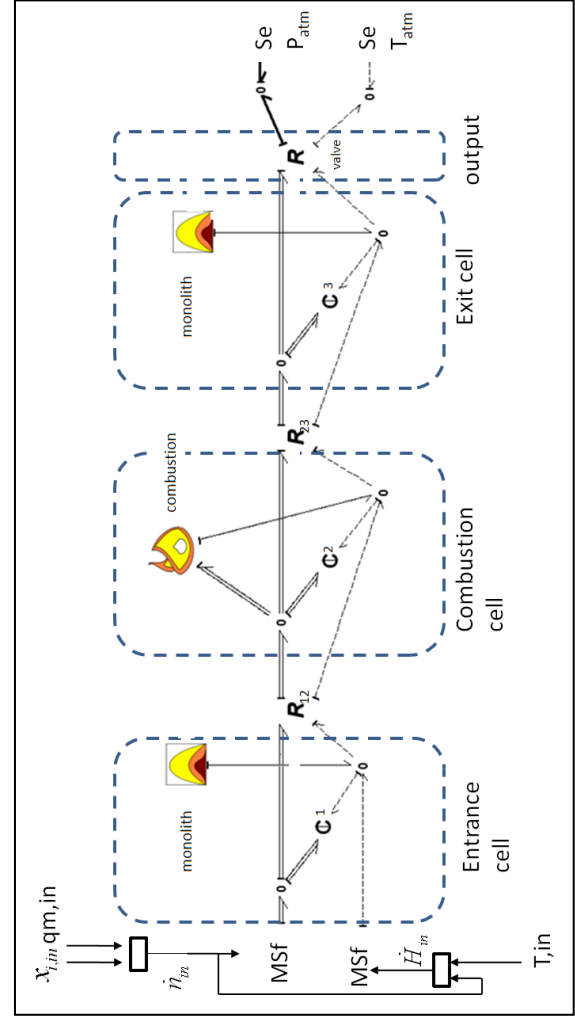
Figure 3: 3 Cells Channel Model

- Equations for the C-elements

Consider for example C2-element.

The number of moles of species i is given by equation (3) coupled with (4) where the different terms correspond to the species molar flow rates through the R-elements and to the combustion zone.

For C1 and C3, no combustion occurs: $\dot{n}_{i1\to combustion} = 0$

$$n_{iC2} = \int \dot{n}_{iC2}\, dt \tag{3}$$

$$\dot{n}_{iC2} = \dot{n}_{i1\to2} - \dot{n}_{i2\to3} - \dot{n}_{i2\to combustion} \tag{4}$$

The temperature in each cell i is given by equation (5) coupled with (6) and (7). The different terms in (6) correspond to the convective heat flow rates through the R-elements, the conduction heat flow from the monolith and the combustion heat (only for C2).

$$T_{C2} = \frac{1}{C_{th}} \int \dot{H}_{C2}\, dt \tag{5}$$

$$\dot{H}_{C2} = \dot{H}_{1 \to C2} + \dot{Q}_{combustion \to C2} - \dot{H}_{C2 \to R23} + \dot{Q}_{monolith} \qquad (6)$$

$$C_{th} = nC_p \qquad (7)$$

The partial pressures in each cell i are given by equation (8) and the total pressure and total number of moles by equations (9) and (10). The volume of each cell is constant.
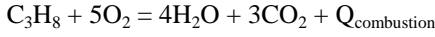
$$P_{i_{C2}} = \frac{n_{i_{C2}}RT_{C2}}{V_{cell}} \qquad (8)$$

$$P_{C2} = \sum_i P_{i_{C2}} \qquad (9)$$

$$n_{C2} = \sum_i n_{i_{C2}} \qquad (10)$$

- The reaction block is a kinetic formula which has to link the Gibbs energy with the reaction rate.

The chemical reaction is

$$C_3H_8 + 5O_2 = 4H_2O + 3CO_2 + Q_{combustion}$$

As it was written before, the bond-graph is not restrictive to Arrhenius kinetic model. It was a choice to model the combustion as a first-order reaction.

$$r_v = A_r \exp\left(\frac{-E}{RT_{C2}}\right) \qquad (11)$$

$$\dot{n}_{C_3H_8} = r_v \frac{P_{C_3H_8}}{RT_{C2}} V_{cat} \qquad (12)$$

Where V_{cat} is the volume of catalyst per cell

$$\dot{n}_{O_2} = 5\dot{n}_{C_3H_8} \qquad (13)$$

$$\dot{n}_{H_2O} = -4\dot{n}_{C_3H_8} \qquad (14)$$

$$\dot{n}_{CO_2} = -3\dot{n}_{C_3H_8} \qquad (15)$$

$$\dot{Q}_{combustion} = -\dot{n}_{C_3H_8} \Delta_r H° \qquad (16)$$

- Equations for R-elements

Consider for instance R23-element.

The volume flow rate in R23, given by equation (18), allows the calculation of the different flow rates needed for the C-elements.

$$Q_{v2 \to 3} = \frac{\pi D^4}{128 \mu L_{cell}} (P_{C2} - P_{C3}) \qquad (17)$$

$$\dot{n}_{2 \to 3} = \frac{P_{C2} Q_{v2 \to 3}}{RT_{C2}} \qquad (18)$$

$$\dot{H}_{2 \to 3} = \dot{n}_{2 \to 3} C_p T_{C2} \qquad (19)$$

$$\mu = \mu_a T_{C2} + \mu_b \qquad (20)$$

$$\dot{n}_{i_{2 \to 3}} = \frac{P_{i_{C2}}}{P_{C2}} \dot{n}_{2 \to 3} \qquad (21)$$

The R-element corresponding to matter and heat transfer to environment is described hereafter.

- Equations for R_{output} (in once-through operating mode, $D_g = D$)

$$Q_{v3 \to out} = \frac{\pi D_g^4}{128 \mu L_g} (P_{C3} - P_{out}) \qquad (22)$$

$$\dot{n}_{3 \to out} = \frac{P_{C3} Q_{v3 \to out}}{RT_{C3}} \qquad (23)$$

$$\dot{H}_{3 \to out} = \dot{n}_{3 \to out} C_p T_{C3} \qquad (24)$$

$$\mu = \mu_a T_{C3} + \mu_b \qquad (25)$$

At this point, a basic model for once-through operating mode has been built.

This model is extended in the next section by introducing the dynamic of heat transfer between the channel and the monolith.

3.3. Extension of the 3 cells-model

The heat transfer and mass transfer between combustion sub-system and channel sub-system involve a boundary layer. It is a 1D correlation of heat and mass-transfer between the gas and the washcoat, which temperature and composition are supposed to be homogeneous. It is a relevant parameter because during simulation, a thermal gradient of about 10K was calculated within this boundary layer. It is represented by an R-element. Figure 4 shows the detailed and extended model of a cell.
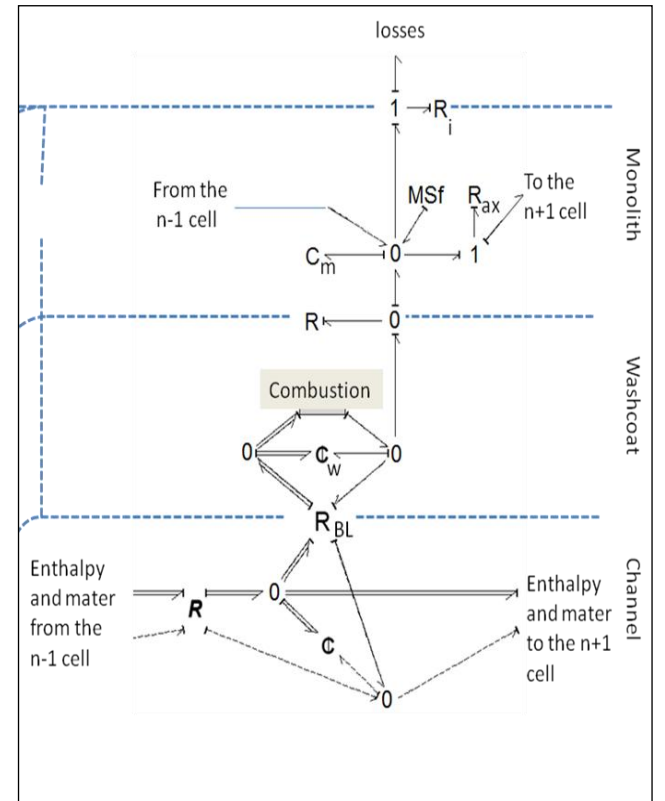

Figure 4: Detailed Bond Graph Model of a Cell

- Equations for the boundary layer R_{BL}

The heat flow rate is calculated using washcoat and mainstream temperatures in (26)

$$\dot{Q}_{BL} = \frac{1}{R_{th}} (T_w - T_{ms}) \qquad (26)$$

Where $R_{th} = \frac{1}{\pi h D}$ $\qquad (27)$

With $h = \frac{kNu}{D}$ (28)

For the same reason as viscosity, thermal conductivity and diffusivity were linearly approximated. The temperature was taken in the washcoat capacitive element.

$$k = k_a T_w + k_b \quad (29)$$

The species molar flow rate is given by equation (30)

$$\dot{n}_i = \frac{1}{R_n}\left(P_{i_{ms}} - P_{i_w}\right) \quad (30)$$

Where $R_n = \frac{1}{\pi \beta R T_w D}$ (31)

With $\beta = \frac{\alpha Sh}{D}$ (32)

and $\alpha = \alpha_a T_w + \alpha_b$ (33)

The washcoat layer is the catalytic layer between the monolith and the mainstream. It was previously neglected as far as mass and heat transfer is concerned because it is very thin. Let's now improve the model considering that this material is porous and can store heat in the catalyst or gas in the pores. A C- element is added in the model to take this into accounts.

- Equations for the washcoat heat and matter storage element Cw

$$\dot{Q}_{Cw} = \dot{Q}_{combustion \to w} - \dot{Q}_{BL} - \dot{Q}_{w \to m} \quad (34)$$
$$\dot{n}_{i_{Cw}} = \dot{n}_{i_{ms}} - \dot{n}_{i_{w \to combustion}} \quad (35)$$

The heat capacity of the air is neglected: heat storage occurs only in the catalyst.

$$T_w = \frac{1}{C_{th}} \int \dot{Q}_w \, dt \quad (36)$$

With $C_{th} = (1 - \varepsilon)\rho_c V_w C_c$ (37)

The matter accumulates inside the pores of the catalyst. The partial pressures are given by equations (38) and (39)

$$P_{i,w} = \frac{RT}{\varepsilon V_w} n_{i,w} \quad (38)$$

With $n_{i,w} = \int \dot{n}_{i,w} \, dt$ (39)

When the reactants diffuse inside the catalyst, part of them is consumed. Concentration in the catalyst is not equal to surface concentration as it can be calculated with ideal gas equation. Hence, equation (12) is slightly modified: η factor makes a correlation between the partial pressures of species on the surface of the catalyst and a medium concentration for the reaction rate. According the simulations η is about 0.8.

Using Thiele factor equations (combustion model accuracy improvement), it leads to

$$\dot{n}_{C_3H_8} = \eta r_v \frac{P_{C_3H_8}}{RT_{Cw}} V_{cat} \quad (40)$$

With $\eta = \frac{\tanh(\theta)}{\theta}$ (41)

Where $\theta = t_w \sqrt{\frac{r_v}{D_e}}$ (42)

Most of the heat is stored in the monolith.

- Equations of the monolith (Cm-element)

The temperature in the monolith is given by

$$T_m = \frac{1}{C_{th}} \int \dot{Q}_{Cm} \, dt \quad (43)$$

With $C_{th} = \rho_m V_m C_m$ (44)

The monolith stores heat, but also exchanges it axially and radially. Equation (45) gives the heat balance of the nth cell:

$$\left[\dot{Q}_{Cm}\right]_n = \left[\dot{Q}_{w \to m}\right]_n + \left[\dot{Q}_{ax}\right]_{n-1 \to n} - \left[\dot{Q}_{ax}\right]_{n \to n+1} - \left[\dot{Q}_{m \to loss}\right]_n \quad (45)$$

The thermal gradient between the washcoat and the monolith was modeled with an R element.

We make the assumption the thermal conductivity of the washcoat is independent of its porosity. It is legitimized because this resistance is not relevant, and does not require a detailed model. Maxwell or Nielsen correlations are often used for porous materials.

$$\left[\dot{Q}_{w \to m}\right]_n = \frac{1}{R_{th}}\left(T_w - T_m\right) \quad (46)$$

With $R_{th} = \frac{\ln\left(1 + \frac{2t_w}{D}\right)}{2\pi k_w L_{cell}}$ (47)

thermal resistance for a cylindrical material

The monolith can also exchange heat axially.

$$\left[Q_{ax}\right]_{n \to n+1} = \frac{1}{R_{th,axial}}\left(\left[T_m\right]_n - \left[T_m\right]_{n+1}\right) \quad (48)$$

With $R_{th_{axial}} = \frac{L_{cell}}{S_m k_m}$ (49)

The MSf models a heater which pre-heats the reactor before reaching autothermal state, or warms-up it enough to burn the gas if the pollutant is too much diluted. A small amount of heat can also be exchanged with other channels or with the outside of the reactor. It is the "to the ambient" port.

The ultimate layer is a virtual layer called insulation layer. It is just an R element associated with an effort source Se representing ambient temperature.

- Equations of the insulation coating

The temperature distribution of all the channels of the reactor was assumed to be the same. It means that the heat carried and removed by the air is much higher than the heat loss. As a consequence all the channels have more or less the same contribution to heat loss. These losses can be equally distributed among the

channels. Moreover, with this assumption, the temperature of the wall of every channel is the same.

$$\left[\dot{Q}_{m\rightarrow loss}\right]_{1\_channel} = \frac{\left[\dot{Q}_{loss}\right]_{cross-section}}{N_{channels}} = \frac{T_i - T_{atm}}{N_{channels}R_{insulation}} \quad (50)$$

In the simulations, we assume the reactor's cross section to be square:

$$R_{insulation} = \frac{t_i}{4k_i L_{cell}(D+2t_m+2t_w)\sqrt{N_{channels}}} \quad (51)$$

4. REVERSE- FLOW REACTOR MODEL

To get a more precise model, the central section (combustion cell) is duplicated several times to increase the space convergence of the model. Each duplication of the number of cells multiplies by 4 the calculation time (variable step-solver Vode&Adams). After several simulation tests, it appeared that 82 sections model might be a good compromise between precision and computation time.

Using the graphical conventions given in Table 2, Figure 5 shows the reverse flow reactor global model.

Table 2: Graphical Conventions for Direct and Reverse Flows

	Direct or shared bonds		Exclusive to reverse-flow bonds	
Information				
Mass transfer	gas	mixture	gas	mixture
Enthalpy				
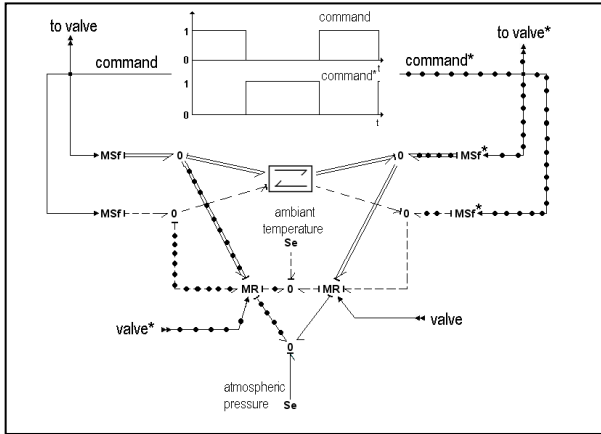
Figure 5: Reverse Flow Reactor Model

For once-through operating mode, the flow is generated by the MSf and "valve" is always opened: the flow follows the continuous bonds. For reverse-flow operating mode, MSf* is added and "gate*" is opened while "gate" is closed and MSf flow generated drops to zero. Hence, the flow in the channel is reversed.

To observe the reverse-flow phenomena, a Boolean was added in the convective heat transfer elements.

The inlet and outlet cells are modified to take into account the reversal of the flow. The inlet flow, modeled with MSf element enters into a C-element, the outlet flow goes out to the environment through a R-element. These R- and C-element have to appear in these cells, but they are used only when needed.

Causality ensuring pressure and heat drop is known so the direction of the flows is deduced. Then, the temperature of the original flow and its partial pressures$[T, P_i]$ can be known from upstream T and P_i.

Then equations (17) to (21) are used. For equations (18) to (21), T_{C2} must be replaced by the temperature of the upstream, and the partial pressures are also given by the upstream capacitance element.

Gates are symbolized by MR elements, controlled by squared signals. It is the R_{output} described in equation (22), where D_g is controlled ($0 < D_g < D$).

The abatement rate $y(t)$ is measured by adding the pollutant molar fractions at the exits of MR_{valve} and $MR_{valve}*$

$$y(t) = 1 - \frac{x_{C_3H_8,valve}(t) + x_{C_3H_8,valve*}(t)}{x_{C_3H_8,inlet}(t) + x_{C_3H_8,inlet*}(t)} \quad (53)$$

5. SIMULATION RESULTS

5.1. Comparison of the BG Model with the Literature

The reverse-flow operating mode was compared to V. Balcaen (2011) detailed simulation, showing the temperature spacial distribution in a reverse flow reactor in quasi-steady state.

The reactor dimensions and materials are given in Table 3. The experiment starts with a hot monolith at a homogenous temperature of 725K. Then, an air flow of $1.436*10^{-6}$ kg/s/channel is introduced at T=300K, $x_{C3H8}=0.0035$, $x_{O2} = 0.21$, $x_{N2}=0.78$, $x_{H2O} = x_{CO2} = 0$.

The comparison model results are represented by the dotted curve, the simulation results of the current model by the continuous line.
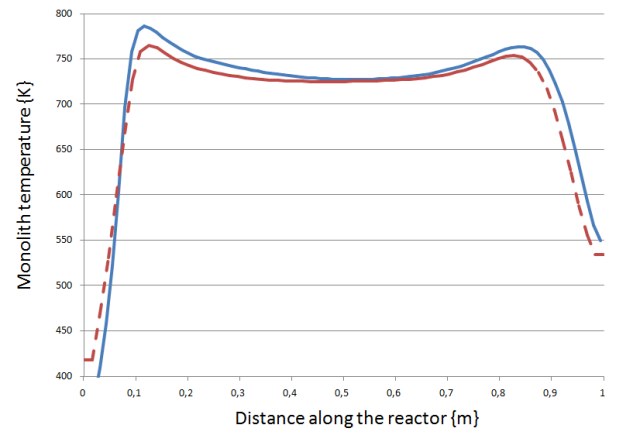

Figure 6: comparison with previous work (250s after the 8th semi-cycle)

Balcaen results are based on Heynderickx (2010) kinetic model whereas the bond graph calculations are

based on Ramanathan's approximation. Moreover, Balcaen simulation was run with a 64 sections model.

The current model presents a difference of about 5% with Balcaen's one which is satisfying.

In Balcaen 's thesis, no transient state investigations were presented.

5.2. Some Case Studies

1. Reactor simulation under standard conditions:

The reactor starts-up with an initial monolith temperature of 300K. The air flow is introduced with a pollutant concentration of 3500 ppmv and a temperature of 300K.

The central part of the reactor [0.384-0.616m] is warmed-up by an external heat source. The heater is constituted of several independent sub-systems with a closed-loop control, each cell beneficiates of one sub-system and its control.

The closed loop system has a set point of 600K, and cannot supply more power than 10W/channel. An hysteresis with amplitude of 50K is added to prevent small oscillations.

On Figure 7 the power delivered by the heater, the abatement rate and the temperature of a middle section are reported.
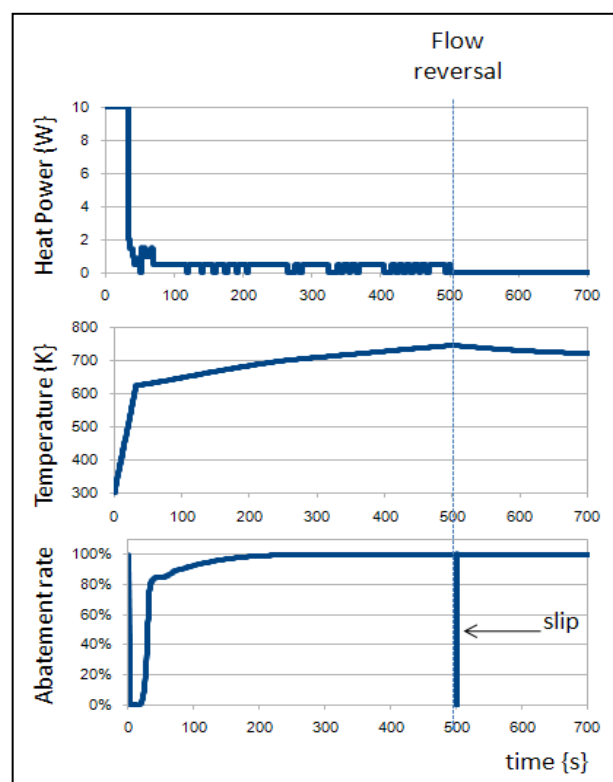
Figure 7: Heater Power, Abatement Rate and Temperature versus Time

In the current study, the whole heating-step occurs during the first semi-cycle.

After about 50 s the system is very close to the set point temperature. However, the temperature continues to increase as a consequence of the heat released by the combustion.

The small amounts of power released during the first semi-cycle are liberated by the upstream cell's heater. Indeed, this cell is directly in contact with cold cells and receives fresh air (at a temperature close to 300K). Hence, it dissipates a lot of heat and requires more power than downstream cells. The advantage of RFR is here demonstrated because for the time interval 400-500s, the set-up dissipates 0.25W/channel which means about 21kW for an industrial installation of 100L/s, 504KWh/day ; when the flow is reversed, the energetic consumptions drops to zero.

Just before the first semi-cycle ends-up, the downstream part of the reactor is hot, while the upstream remains cold.

When the flow is reversed, the autothermal process starts: the air is warmed-up by the hot sections ignites and liberates heat in the cold sections.

During the flow reversal, the abatement rate falls to zero. It is due to the pollutant repartition within the channel just before the flow reversal (see Figure 8).
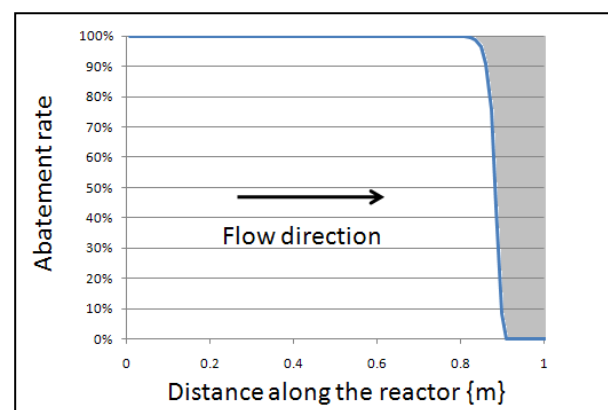
Figure 8: Pollutant Concentration Profile just after Flow Reversal (16th flow reversal)

A few amount of pollutant remains in the cold part of the reactor (here on the right). Usually, the air flow ensures the pollutant is going to cross the reactor and burn. However, during the flow reversal, the pollutant is ejected and does not pass through the channel. All this fresh and polluted air is liberated. This is an important phenomenon in the study of reverse-flow reactors, as it will be shown in the next case.

2. Influence of the switching time

To check the influence of the switching time on transient and pseudo steady-state reactor behavior, a simulation scenario was designed.

The reactor starts with an homogenous monolith temperature T=500K, and polluted air is introduced under standard conditions (3500ppmv, 300K).

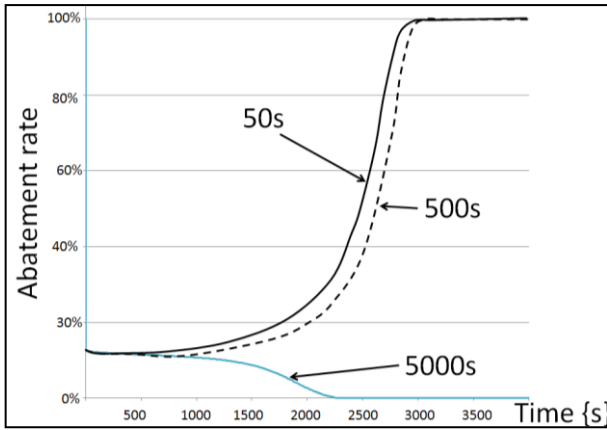Several semi-cycle times are tested (tsc = 50 s, 500 s, 5000 s).

Figure 9: Abatement Rate versus Time for several Semi-Cycle Times

The ignition process occurs in the time interval 0 to 3000 s. The abatement rate shows this process mostly depends on physical parameters of the reactor. After 1000 s, the 5000 s semi-cycle time reactor starts loosing ignition. An interpretation is if the switching time is too high (tsc=5000s), the reaction front quits the channel, and the pollutant abatement goes down to zero. The transient state is shorter of about 100s for short semi-cycle time. As a result, the cumulative pollutant emissions are lower of 6% at t=4000s.

However if the switching time is too short, slips emissions rise-up. Figure 10 below shows the abatement rate and the slip emission around t=17500s. Slip emission is defined as :

$$ slip\_emissions = \frac{\int_{To}^{t} \dot{n}_{C3H8}\ dt}{\int_{17250}^{17750} \dot{n}_{C3H8} dt} \quad (54) $$
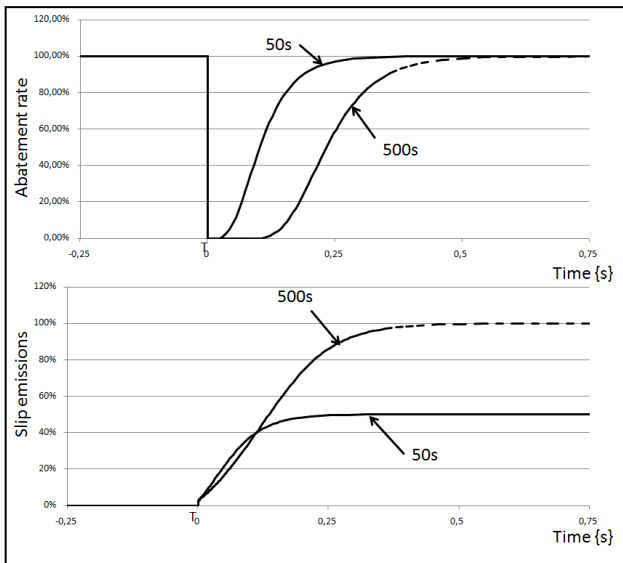
Figure 10: Abatement Rate and Slip Emissions for 50s and 500s Semi-Cycle Times at T_0=17500s Flow Reversal

The amplitude of the emissions of the slip in the case of $t_{sc} = 50s$ is about half of the amplitude for t_{sc}=500s, because the penetration of the heat-front is deeper in the 500s case. However, the flow is reversed 10 times more frequently in the case of $t_{sc} = 50s$ and hence, during 500s the $t_{sc} = 50s$ reactor emits 5 times more pollutant in the atmosphere. As a result, it is preferable to work on large semi-cycle time during the steady state. However, according to Figure 9 it is worth to start-up with short semi-cycle time. Thus, this study shows the interest of a dynamic variation of switching time frequency during the reactor start-up.

6. CONCLUSION AND FUTURE WORKS

In this paper, a bond graph model for monolithic reverse-flow reactors has been developed. Its results match with previous studies on this subject.
On the one hand, simulations of an 82 cells model were run with a ratio of about 10s of simulation per second of calculation on a laptop. This rate can clearly be improved by, for instance, simplifying the reaction rate expression or reducing the number of cells through a variable space-step computation.
On the other hand, the reverse-flow reactor can be handled through several parameters like purge time, mass-flow rate or heat exchange in the monolith.
Hence, supervision strategy field is opened for monolithic reverse-flow-reactor. Moreover, the bond-graph permits to simulate this end-of pipe device as a part of a bigger system.
Future works will study critical conditions of use under dynamic constraints like a dynamic variation of pollutant concentration or inlet mass-flow rate fluctuation.

APPENDIX

Table 3: standard values of parameters

Solver : Vode& Adams	Step size absolute precision relative precision	0.001 10^{-8} 10^{-8}
Length of the reactor	L {m}	1.000
Geometry of the channel : circular	Nu Sh	4.364 4.364
Length of inert section	L_g $= 0.05 * L$ {m}	0.05
Hydraulic diameter of one channel	D {m}	0.002
Number of channels	Nchannel	100
Number of cells	N	82
Thickness of the washcoat	tw {mm}	25
Thickness of the monolith	tm {mm}	150
Thickness of the insulation	ti {mm}	1

Thermal conductivity of the insulation	λ_i {W/m.K}	0.002
density of the catalyst	ρ_c {kg/m^3}	4000
Porosity of the washcoat	ε	0.5
Density of the monolith	ρ_m{kg/m^3}	8000
Thermal conductivity of the air ($\lambda = \lambda_a*T+\lambda_b$)	λ {W/m.K}	$\lambda_a = 5.91*10^{-5}$ $\lambda_b = 0.01$
Diffusivity of the air ($\alpha = \alpha_a*T + \alpha_b$)	α {m²/s}	$\alpha_a = 2.08*10^{-7}$ $\alpha_b = -4.56*10^{-5}$
Viscosity of the air ($\mu = \mu_a*T + \mu_b$)	μ {Pa.s}	$\mu_a = 3.30*10^{-8}$ $\mu_a = 3.30e^{-8}$
Heat capacity of the air	Cp {J/mol.K}	29.1
Thermal conductivity of the washcoat	λ_w{W/m.K}	16
Heat capacity of the catalyst	Cc { J/kg.K}	947
Thermal conductivity of the monolith	λ_m{W/m.K}	16
Heat capacity of the monolith	Cm {J/kg.K}	502
Reaction enthalpy	ΔrH {J/mol}	- 2040240*10^3
Frequency rate of the reaction	Ar {s^{-1}}	$9.6*10^{10}$
Reaction activation energy	E {J/mol}	89791
Semi-cycle time	t_{sc} {s}	500s
Atmospheric pressure	P$_0$ {Pa}	101325
Inlet temperature	T$_{in}$ {K}	300

NOTATIONS:

Q_v: Volume flow rate
h: Newton's heat transfer coefficient
β: Fick's heat transfer coefficient
$q_{m,in}$: Mass flow rate
M: Molecular weight
P_i: Partial pressure in specie i
$x_{i,in}$: Inlet molar fraction of specie i
τ: Specific time of reactor thermal behavior
η: Thiele factor
θ: non-dimensional number for Thiele factor
r_v: volumetric reaction rate
Layers :
- ms : main-stream
- w : washcoat (catalyst+air bubbles)
- m : monolith
- i : insulation
- c : catalyst

REFERENCES

Balcaen, V., 2011. The Total Oxidation of Propane over Metal Oxyde Catalyst: Transient Kinetics and Monolith Reactors, Thesis (PhD). Gent University.

Couenne, F., Jallut, C., Maschke, B., Breedveld, P. C., Tayakout, M., 2006. Bond Graph modeling for chemical reactors. *Mathematical and Computer Modeling of Dynamical Systems*, 12, 159.

Heny, C., Simanca, D., Delgado, M., 2006. Pseudo-bond graph model and simulation of a continuous stirred tank reactor. *Journal of the Franklin Institute*, 12, 159.

Heynderickx, M.P., Thybaut J.W., Poelman, H., Poelman, D., Marin, G.B., 2010. Kinetic modeling of the total oxidation of propane over CuO-CeO$_2$/γ-Al$_2$O$_3$. *Applied Catalysis B: Environmental*, 95, 26-38.

Karnopp, D., 1990. Bond graph models for electrochemical energy storage : electrical, chemical and thermal effects. *Journal of the Franklin Institute*, 327(6), 983-992.

MacKenzie, S.A., Gawthrop, P.J., Jones, R.W., 1993. Modeling chemical process with pseudo bond graphs. *Proceedings of the International Conference in Bond Graph Modeling*, pp. 327-332. January 17-20, La Jolla (California, USA).

Matros, Y.S., Bunimovich, G.A., 1996. Reverse-flow operation in fixed bed catalytic reactors. *Catalysis Reviews: Science and Engineering,* 38(1), 1-68.

Marin, P., Ordonez, S., Diez, F.V., 1990. Simplified design methods of reverse-flow catalytic combustors for the treatment of lean hydrocarbon-air mixtures. *Chemical Engineering and Processing*, 48, 229-238.

Ramanathan, K., West, D.H., Balakotaiah, V., 2004. Optimal design of catalytic converters for minimizing cold-start emissions. *Catalysis Today, 98,* 357–373

Ramdani, K., Pontier, R., Schweich, D., 2001. Reverse flow reactor at short switching periods for VOC combustion. *Chemical Engineering Science,* 56(4), 1531-1539

Thoma, J. and Ould-Bouamama, B., 1999. Modeling and Simulation in Thermal and Chemical Engineering. London: Pringer Engineering Edition.

Van de Beld, B., Westerterp, KR., 1994a. Air purification by catalytci oxidation in a reactor with periodic reverse flow. *Chemical Engineering & Technology*, 17, 217-226

Van de Beld, B., Borman, R.A., Derkx, O.R., van Woezik, B.A.A., Westerterp, K.R., 1994b. Removal of volatile organic compounds from polluted air in a reverse flow reactor: An experimental study. *Industrial and Engineering Chemistry Research,* 33, 2946-2956

MODELING MULTIPHASE FLOW DYNAMICS IN OFFSHORE PETROLEUM PRODUCTION SYSTEMS

Rafael Horschutz Nemoto, Jorge Luis Baliño

Departmento de Engenharia Mecânica, Escola Politécnica, Universidade de São Paulo
Av. Prof. Mello Moraes, 2231, CEP 05508-900, Cidade Universitária, São Paulo-SP, Brazil

rafael.nemoto@usp.br, jlbalino@usp.br

ABSTRACT

In this study, a mathematical model is proposed to investigate the dynamics of gas, oil and water flow in a pipeline-riser system. The pipeline is modeled as a lumped parameter system and considers two switchable states, one in which the gas is able to penetrate into the riser, and another in which there is a liquid accumulation front, preventing the gas from penetrating the riser. The riser model considers a distributed parameter system, in which movable nodes are used to evaluate local conditions along the subsystem. The method of characteristics allows the simplification of the differentiation in the hyperbolic system of equations obtained. The resulting equations are discretized and integrated using an implicit method with a predictor-corrector scheme for the treatment of the nonlinearities. Simulations corresponding to severe slugging conditions are presented. As result of the simulations, pressure, volume fractions and superficial velocities for the phases along the tubes are calculated.

Keywords: multiphase flow, petroleum production systems, lumped and distributed parameter systems, switched systems.

1. INTRODUCTION

In offshore petroleum production systems, the fluids that leave the well are often transported to platforms by means of flexible pipes. The pipes are composed of a pipeline (or flowline), which conducts the fluids over the seabed topography, and a riser, which elevates the fluids to the separator vessel located at the platform, as shown in Fig. 1. Usually, the transported fluids are composed of gas, oil and water, but due to the severe conditions of pressure and temperature, it is possible the formation of emulsions, hydrates and wax. These features make the modeling of the multiphase flow dynamics a complex task.

Some models were proposed for the air-water flow in a pipeline-riser system. Schmidt *et al.* (1979), Fabre *et al.* (1987), Taitel *et al.* (1990), Sarica and Shoham (1991) and Baliño *et al.* (2010) are some of the authors that researched the behavior of this biphasic flow and proposed different methods to determine the system stability. Using air-water as flowing fluids, it is possible to investigate basic mechanisms of the flow behavior; however, there are many limitations when trying to extrapolate these results to petroleum production systems.

Pipeline lengths and riser heights in petroleum production systems are much larger (order of kilometers long) than the values for air-water experimental facilities. The high pressure ratios between the bottom and top of the riser give rise to important expansion effects in the gas phase, invalidating models based on the assumption of a mean void fraction.

Petroleum is a multicomponent system in which both liquid and gas phases coexist at operating conditions (McCain, 1990). Mass transfer between the phases are dependent on pressure and temperature through the pressure, volume and temperature (PVT) data. With the high pressure variations in the riser, mass transfer effects cannot be ignored. Besides, the fluid coming from the reservoir has a water content, so three phases can coexist in the general case.

In the present work, a mathematical model is proposed to investigate the dynamics of gas, oil and water flow in a pipeline-riser system. The pipeline is modeled as a lumped parameter system and considers two switchable states, one in which the gas is able to penetrate into the riser, and another in which there is a liquid accumulation front, preventing the gas from penetrating the riser. The riser model considers a distributed parameter system, in which movable nodes are used to evaluate local conditions along the subsystem. For both subsystems continuity equation for gas, oil and water phases are taken into account. Oil and water phases are considered to have the same velocity and are homogenized. Mass transfer between the oil and gas phases is calculated using the black oil approximation. The properties of fluids are calculated by analytical correlations based on experimental results and field data.
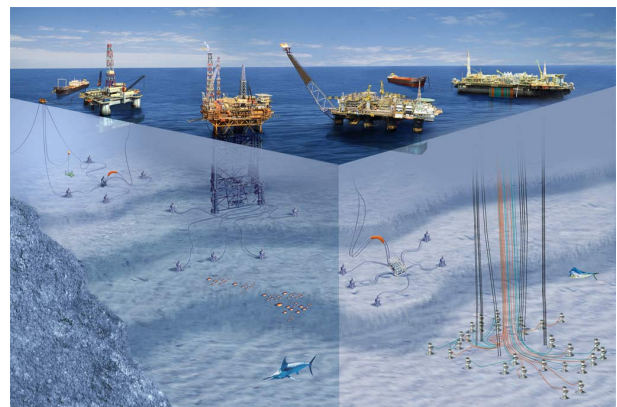

Figure 1: Typical offshore petroleum production systems (source: Petrobras).

The pipeline model considers a stationary linear mo-

mentum balance for the smooth stratified flow regime and the riser model uses a simplified momentum equation without inertia terms for the mixture. A drift-flux model, evaluated for the local conditions in the riser, is used as a closure law. The method of characteristics allows the simplification of the differentiation in the hyperbolic system of equations obtained. The resulting equations are discretized and integrated using an implicit method with a predictor-corrector scheme for the treatment of the non-linearities.

2. Riser model

The model is based on a distributed parameter system, in which nodes are used to evaluate the local condition along the subsystem. It considers a one-dimensional three-phase isothermal flow. Continuity equation for gas, oil and water phases and a simplified momentum equation without inertia terms for the phases flowing together are the governing equations. Oil and water phases are considered to have the same velocity and are homogenized. Slip between the liquid and gas phases are taken into account by using a drift flux model. Mass transfer between the oil and gas phases are calculated using the black oil model. The liquid and gas phases are assumed to be compressible and the gas behaves as a real gas. Solubility of gas and vaporization are neglected for water.

2.1 Riser geometry

The riser geometry is characterized by a function expressing the coordinates of the points belonging to the riser; from this, it is possible to determine the local position along the riser and the local inclination angle. For a catenary riser, geometry is characterized by the coordinates X and Z, corresponding to the abscissa and the height of the top of the riser (see Fig. 2). It is assumed that the inclination angle at the bottom is zero.
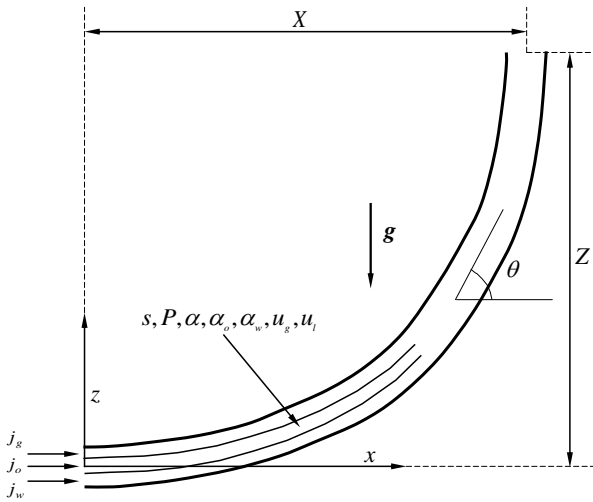
Figure 2: Variables at the riser.

The local height z of a point belonging to catenary can be written as:

$$z = \varphi \left[\cosh \left(\frac{x}{\varphi} \right) - 1 \right] \tag{1}$$

where the dimensional catenary constant φ is obtained from the solution of the following transcendental equation:

$$Z = \varphi \left[\cosh \left(\frac{X}{\varphi} \right) - 1 \right] \tag{2}$$

The local position s along the catenary results:

$$s = \varphi \sinh \left(\frac{x}{\varphi} \right) \tag{3}$$

The local inclination angle θ can be written as:

$$\theta = \arctan \left[\sinh \left(\frac{x}{\varphi} \right) \right] \tag{4}$$

Knowing the position s, the local abscissa x can be calculated from Eq. (3):

$$x = \varphi \operatorname{arcsinh} \left(\frac{s}{\varphi} \right) \tag{5}$$

2.2 Conservation equations

Considering continuity equations for the phases gas, oil and water, we get:

$$\frac{\partial}{\partial t} (\rho_g \, \alpha) + \frac{\partial j_g}{\partial s} = \Gamma \tag{6}$$

$$\frac{\partial}{\partial t} (\rho_o \, \alpha_o) + \frac{\partial j_o}{\partial s} = -\Gamma \tag{7}$$

$$\frac{\partial}{\partial t} (\rho_w \, \alpha_w) + \frac{\partial j_w}{\partial s} = 0 \tag{8}$$

where s is the coordinate along the flow direction, t is time, ρ_g, ρ_o and ρ_w are the densities of the phases (correspondingly gas, oil and water), j_g, j_o and j_w are the superficial velocities, α, α_o and α_w are the volume fractions and Γ is the vaporization source term.

In most of the transients occurred in oil and gas transport, for instance in severe slugging, the response of the system proves to be relatively slow, showing that pressure waves do not have a strong effect on the initiation and transport of void waves. In the no-pressure-wave (NPW) model (Masella *et al.*, 1998), acoustic waves are ruled out by neglecting inertia terms from the momentum equation, resulting an algebraic relation for the pressure gradient:

$$\frac{\partial P}{\partial s} = -\frac{4 \, \tau_w}{D} + \rho_m \, g_s \tag{9}$$

$$\rho_m = \rho_g \, \alpha + \rho_o \, \alpha_o + \rho_w \, \alpha_w \tag{10}$$

where P is pressure, ρ_m is the density of the mixture, D is the pipe diameter, g_s is the gravity component in the s-direction and τ_w is the mean shear stress at the pipe wall. The volume fractions are related by:

$$\alpha_o + \alpha_w + \alpha = 1 \tag{11}$$

2.3 Closure laws

In order to close mathematically the problem, some simplifications must be made.

2.3.1 Homogenization of liquid phases

Assuming equal velocities for oil and water, we obtain:

$$j_o = j_l \frac{\alpha_o}{1 - \alpha} \qquad (12)$$

$$j_w = j_l \frac{\alpha_w}{1 - \alpha} \qquad (13)$$

$$j_l = j_o + j_w = u_l (1 - \alpha) \qquad (14)$$

where j_l and u_l are correspondingly the superficial velocity and the velocity of the liquid (oil plus water) phase.

2.3.2 Shear stress at the wall

The shear stress at the wall is estimated using a homogeneous two-phase model and the correlation from Chen (1979) for the Fanning friction factor f.

2.3.3 Real gas

As the pressures involved are high, the constitutive relation for the gas phase is considered as:

$$\rho_g = \frac{\gamma_g M_a}{\Lambda T} \frac{P}{Z} \qquad (15)$$

where $\gamma_g = M_g/M_a$ is the gas specific gravity, M_g and $M_a = 28.966$ are respectively the molar masses of gas and air, Z is the gas compressibility factor (dependent on pressure, temperature and gas composition) and $\Lambda = 8.314 \, m^2 s^{-2} K^{-1}$ is the gas universal constant.

2.3.4 Drift flux model

The superficial velocities for the liquid and gas phases are determined by using a drift flux model (Zuber and Findlay, 1965):

$$j_g = \alpha (C_d j + U_d) \qquad (16)$$

$$j_l = (1 - \alpha C_d) j - \alpha U_d \qquad (17)$$

$$j = j_l + j_g \qquad (18)$$

where the parameters C_d and U_d depend on the local geometric and flow conditions (Bendiksen, 1984; Chexal *et al.*, 1992). In a general form, it will be assumed that $C_d = C_d (\alpha, P, j, \theta)$ and $U_d = U_d (\alpha, P, j, \theta)$.

2.3.5 Black oil model

The vaporization term can be calculated by using the black oil model (McCain, 1990). According to this model, the gas specific gravity does not change with variations of pressure and temperature:

$$\gamma_g \cong \gamma_{g0} \qquad (19)$$

$$\gamma_{dg} \cong \gamma_{g0} \qquad (20)$$

where γ_g is the gas specific gravity at local conditions, γ_{g0} is the gas specific gravity at standard conditions and γ_{dg} is the dissolved gas specific gravity.

In this way, many properties corresponding to the phases at operating conditions can be estimated based on parameters at standard condition (1 atm and 60 ^oF for API, American Petroleum Institute) and a set of correlations depending on pressure, temperature and composition, which will be considered as locally and instantaneously valid.

The vaporization term can be expressed as:

$$\Gamma = -\frac{\rho_{g0} \, \alpha_o}{B_o} \left(\frac{\partial R_s}{\partial t} + \frac{j_o}{\alpha_o} \frac{\partial R_s}{\partial s} \right) \qquad (21)$$

where ρ_{g0} is the gas density at standard condition, B_o is the oil formation volume factor and R_s is the solution gas-oil ratio. It is worth noting that for $\Gamma > 0$ must be $\alpha_o > 0$ (there must exist oil for vaporization), while for $\Gamma < 0$ must be $\alpha > 0$ (there must exist gas for condensation).

2.4 Well-posedness and method of characteristics

For a model to describe physical phenomena correctly it must be well-posed, this is, the solution must exist, must be uniquely determined and must depend in a continuous fashion on the initial and boundary conditions (Drew and Passman, 1999). This property is particularly important in multiphase flows, where partial differential equations of hyperbolic nature can be found; in this case, well-posedness implies that the characteristic values (eigenvalues or characteristic wave velocities) must be real.

The characteristic values of the presented system of conservation equations are given by Nemoto and Baliño (2009):

$$e_1 = \frac{\partial j_g}{\partial \alpha} \quad e_2 = \frac{j_o}{\alpha_o} = u_l \quad e_3 = \infty \quad e_4 = \infty \qquad (22)$$

where u_l is the liquid velocity. If the parameters C_d and U_d are not dependent of α, i.e. $C_d = C_d(P, j, \theta)$ and $U_d = U_d(P, j, \theta)$ (as in the correlation developed by Bendiksen (1984)) we have:

$$\frac{\partial j_g}{\partial \alpha} = \frac{j_g}{\alpha} = u_g \qquad (23)$$

where u_g is the gas velocity.

There exists an algebraically-double eigenvalue equal to ∞, these eigenvalues are related to the pressure wave velocities. The pressure wave is propagated in negative and positive directions with an infinite velocity, meaning that any pressure change is felt by the entire system instantaneously.

The method of characteristics will be applied to solve the system of equations. This method is the natural numerical procedure for hyperbolic systems. By an appropriate choice of coordinates, the original system of hyperbolic partial differential equations can be replaced by a system of ordinary differential equations expressed in the characteristic coordinates. Characteristic coordinates are the natural coordinates of the system in the sense that, in terms of these coordinates, differentiation is simpler (Ames, 1992).

The resulting system of equations in the characteristic coordinates, or compatibility conditions, is given by:

$$b_{11}^* \frac{D_g \alpha}{Dt} + b_{13}^* \frac{D_g P}{Dt} + d_1^* = 0 \qquad (24)$$

$$b_{21}^* \frac{D_l \alpha}{Dt} + b_{22}^* \frac{D_l \alpha_o}{Dt} + b_{23}^* \frac{D_l P}{Dt} = 0 \tag{25}$$

$$\frac{\partial j}{\partial s} = f_1 \frac{D_g P}{Dt} + f_2 \frac{D_l P}{Dt} \tag{26}$$

$$\frac{\partial P}{\partial s} = -\frac{4\,\tau_w}{D} + \rho_m\, g_s \tag{27}$$

where the coefficients b_{11}^*, b_{13}^*, b_{21}^*, b_{22}^*, b_{23}^*, d_1^*, f_1 and f_2 are function of the state variables and dependent variables.

The directional derivatives are defined as:

$$\frac{D_g}{Dt} = \frac{\partial}{\partial t} + u_g \frac{\partial}{\partial s} \tag{28}$$

$$\frac{D_l}{Dt} = \frac{\partial}{\partial t} + u_l \frac{\partial}{\partial s} \tag{29}$$

3. Pipeline model

The pipeline model is based on a lumped parameter system, assuming two switchable states, one in which the gas is able to penetrate into the riser, and another in which there is a liquid accumulation front, preventing the gas from penetrating the riser. The model considers one-dimensional three-phase isothermal flow. The governing equations are the continuity equation for gas, oil and water phases and a momentum balance equation evaluated at the stationary state (Taitel and Dukler, 1976), which is used to determine the void fraction in the pipeline.

It is assumed that water and oil phases have the same velocity and are homogenized, the flow pattern is the smooth stratified, the pressure is constant in the gas cavity, mass transfer between the oil and gas phases is calculated using the black oil model and variations in the void fraction α_p are neglected during the transient.

The pipeline model is composed of two control volumes (see Fig. 3). The control volume $\forall_1$ comprises the gas and liquid in the stratified region, while the control volume $\forall_2$ comprises the liquid penetration region.
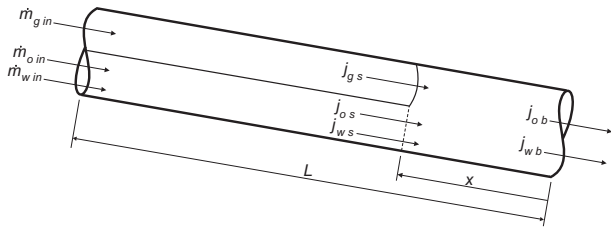
Figure 3: Control volumes in the pipeline model.

3.1 State $x = 0$

The state $x = 0$ is observed when there is no liquid penetration in the pipeline. In this case, there exist only one control volume in the pipeline, the control volume $\forall_1$, and the pressure in the gas cavity is constant and equal to the pressure at the bottom of the riser (P_b).

The state equations are obtained by applying continuity equations for the gas, oil and water phases, resulting in expressions that allows the calculation of the superficial velocity of gas (j_{gb}), oil (j_{ob}) and water (j_{wb}) at the bottom of the riser.

3.2 State $x > 0$

In this state, the liquid front penetrates the pipeline and two control volumes are taken into account. The pressure in the control volume $\forall_1$, P_s, is constant throughout its length and the pressure in the control volume $\forall_2$ depends on position. In order to evaluate fluid properties and its derivatives, a representative pressure P_m for the control volume $\forall_2$ is used:

$$P_s = P_b - \rho_l\, g\, x\, \sin\beta \tag{30}$$

$$P_m = \frac{P_b + P_s}{2} \tag{31}$$

where ρ_l is the density of the liquid phase, g is the gravity, β is the pipeline inclination, x is the liquid penetration length, the index s refers to the interface between the control volumes and the index m refers to the mean value at the control volume $\forall_2$.

3.2.1 Volume control $\forall_1$

From the continuity equations for the gas, oil and water, it is possible to obtain expressions for the velocity of the liquid accumulation front (dx/dt) and superficial velocity of oil (j_{os}) and water (j_{ws}) at the interface of the control volumes.

3.2.2 Volume control $\forall_2$

From the continuity equations for the gas, oil and water, we obtain the superficial velocity of oil (j_{ob}) and water (j_{wb}) at the bottom of the riser.

3.3 Switching condition

Whenever the gas superficial velocity at the bottom of the riser becomes zero, commutation exists between the set of equations determined by the state $x = 0$ to the set of equations determined by the state $x > 0$. And the commutation from the state $x > 0$ to the state $x = 0$ happens when the liquid penetration length becomes zero.

4. FLUID PROPERTIES

The properties of fluids are calculated by analytical correlations based on experimental results and field data.

4.1 Gas formation volume factor and gas density

The gas formation volume factor is calculated by the following expression:

$$B_g = \frac{P_0}{T_0} \frac{Z\,T}{P} \tag{32}$$

where P_0 is the pressure at standard conditions, T_0 is the absolute temperature at standard conditions, Z is the compressibility factor and P and T are the pressure and the absolute temperature at local conditions.

The compressibility factor is determined using the correlation of Dranchuk and Abu-Kassem (1975). For the evaluation of the compressibility factor is also necessary to calculate the pseudocritical temperature and pressure, which can be determined using the correlation of Standing (1981).

Considering the black oil approximation, which assumes a approximately constant gas specific gravity, it can be shown that Eq. (15) reduces to:

$$\rho_g \cong \frac{\rho_{g0}}{B_g} \tag{33}$$

4.2 Water formation volume factor and water density

The correlation for water formation volume factor is presented in the work of McCain (1990). Water density at local condition is determined by:

$$\rho_w = \frac{\rho_{w0}}{B_w} \tag{34}$$

4.3 Gas-oil solubility and bubble point pressure

If the local pressure is above the bubble point pressure, the gas-oil solubility is equal to the gas-oil ratio GOR, otherwise the gas-oil solubility is calculate according to the correlation of Standing (1981).

The bubble point pressure is determined based on the correlation of Velarde *et al.* (1999).

4.4 Oil formation volume factor and oil density

Based on the definition of oil formation volume factor:

$$B_o = \frac{v_o}{v_{o0}} \tag{35}$$

where v_o and v_{o0} are, respectively, the oil volume at local conditions and at standard conditions, the following material balance relation results:

$$B_o = \frac{\rho_{o0} + \dfrac{P_0\, M_a}{\Lambda\, T_0}\, R_s\, \gamma_{dg}}{\rho_o} \tag{36}$$

Assuming that the black oil approximation is valid and substituting Eq. (20) in Eq. (36), we obtain:

$$B_o \cong \frac{\rho_{o0} + \rho_{g0}\, R_s}{\rho_o} \tag{37}$$

The oil density is calculated based on the correlation of Velarde *et al.* (1999).

4.5 Gas, oil and water viscosities

The gas viscosity is calculated using the correlation of Lee *et al.* (1966).

The dead oil viscosity at standard pressure is calculated using the correlation of Ng and Egboah (1994). The dead oil viscosity is necessary to calculate the saturated and subsaturated oil viscosity. The former is calculated using the correlation of Beggs and Robinson (1975), while the latter is calculated using the correlation of Vasquez and Beggs (1980).

The water viscosity is calculated using the results of Collins (1987). The first step is the determination of the water viscosity at standard pressure, then it is possible to evaluate the water viscosity at local conditions.

4.6 Vaporization term

Considering isothermal flow, B_o, α_o and R_s depend only on pressure, so Eq. (21) reduces to:

$$\Gamma = -\frac{\rho_{g0}\, \alpha_o}{B_o} \frac{\partial R_s}{\partial P} \frac{D_l P}{Dt} \tag{38}$$

5. STATIONARY STATE

The stationary state is important since it is used as the initial condition for the transient simulations. The stationary state can be obtained by setting to zero the time derivatives in the dynamic equations. Applying this procedure to Eqs. (6), (7) and (8), we obtain respectively:

$$j_g = \frac{Q_{o0}\,(GOR - R_s)\,B_g}{A} \tag{39}$$

$$j_o = \frac{Q_{o0}\,B_o}{A} \tag{40}$$

$$j_w = \frac{Q_{o0}\,WOR\,B_w}{A} \tag{41}$$

where Q_{o0} is the oil flow rate at standard conditions, R_s is the gas-oil solubility ratio, B_o, B_w and B_g are the formation factor for the corresponding fluids, A is the flow passage area of the riser tube and WOR is the water-oil ratio.

The void fraction α along the riser is calculated using Eq. (16). The oil volume fraction is calculated using the following expression, obtained from Eqs. (11), (12) and (13):

$$\alpha_o = \frac{1 - \alpha}{\dfrac{j_w}{j_o} + 1} \tag{42}$$

The water volume fraction is calculated from Eq. (11):

$$\alpha_w = 1 - \alpha - \alpha_o \tag{43}$$

6. TRANSIENT STATE

To evaluate the transient state, the initial conditions are taken from the stationary state results. A moving grid method was adopted, in which node i ($1 \le i \le N-1$) moves with the gas characteristic velocity. Last node N moves with the liquid velocity if the liquid level falls bellow the top of the riser, or remain fixed at the top otherwise. The time step (Δt) is calculated as the time step such that the node $N-1$ intersects the node N. As the gas velocity is positive, a node disappears at the liquid level or top of the riser and a node is created at the bottom of the riser, keeping constant the number of nodes.

As the nodes move with the gas velocity, lines are defined in the space-time plane (s, t), as shown in Fig. 4. These lines are the gas characteristic path (red lines) on which the integration of the equations projected in the gas direction is performed. The liquid characteristic path (blue lines) must also be known in order to evaluate the integration of the equations projected in the liquid direction. For this purpose, it is considered that at time $t_2 = t_1 + \Delta t$ the liquid characteristic path crosses the nodes. The characteristic velocities for the gas and liquid and the time step Δt are determined from the nodal values at time t_2, as an implicit scheme is used. Once the liquid characteristic path is known, it is possible to interpolate the values for the liquid directional derivatives from the known solution at time t_1.
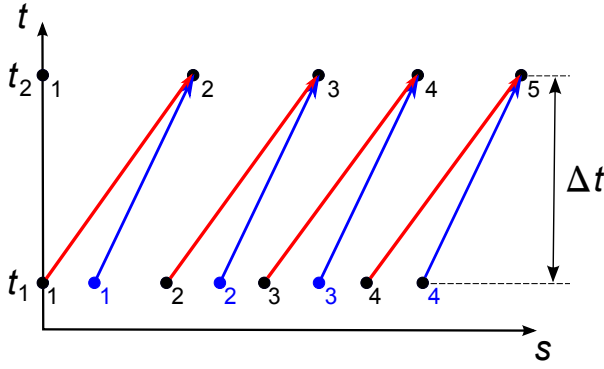
Figure 4: Characteristic lines in the (s, t) plane.

The transient calculations can be summarized as follows:

1. The predictor values of the variables are assumed as the corresponding values obtained at time t_1.

2. The time step Δt is calculated.

3. The fluid properties are calculated at the nodal positions.

4. The predictor values of the fluids superficial velocities at the bottom of the riser are used to evaluate the pressure and gas, oil and water volume fractions along the riser at time t_2, using an implicit scheme.

5. The pressure at the bottom of the riser is used to determine the fluids superficial velocity at the riser base, using the state equations of the pipeline model.

6. The convergence of the variables is verified and in the case it is not achieved, new predictor values for the variables are determined using an under-relaxation factor. A new iteration starts with step 2.

7. If convergence is achieved, time is increased and a new time step is calculated, beginning with step 1.

7. SIMULATIONS

Based on the presented model, a computational program for transient simulations was developed using MATLAB (MATLAB, 2011).

Table 1 presents the input data used to simulate a case. The following figures show the transient response of important variables: void fraction at the bottom of the riser (Fig. 5(a)), oil volume fraction at the bottom of the riser (Fig. 5(b)), water volume fraction at the bottom of the riser (Fig. 5(c)), gas (red), oil (black) and water (blue) superficial velocity at the bottom of the riser (Fig. 5(d)), gas (red), oil (black) and water (blue) superficial velocity at the liquid level in the riser (Fig. 5(e)), pressure at the bottom of the riser (Fig. 5(f)), position of liquid accumulation front at pipeline (Fig. 5(g)) and height of the liquid level at the riser (Fig. 5(h)).

Table 1: Input data for simulation.

Symbol	Variable	Values
API	API-grad	19
γ_g	Gas specific gravity	0.6602
Q_{g0}	Gas volumetric flow rate at standard conditions	$0.1\ Sm^3/s$
GOR	Gas-oil ratio	145
WOR	Water-oil ratio	0.3
T	Temperature	$323\,K$
D	Inner diameter	4"
X	Horizontal length of the top of the riser	$845\ m$
Z	Height of the top of the riser	$1300\ m$
ϵ	Roughness	$4,6 \cdot 10^{-5}\ m$
P_{sep}	Pressure at the separator	$25\ bara$
Y	Salinity	0
N	Number of nodes	101
L	Pipeline length	$1000\ m$
β	Pipeline inclination angle	2°

The simulations correspond to a phenomenon known in petroleum technology as severe slugging (Taitel *et al.*, 1990), in which the flow destabilizes and reaches a limit-cycle. The flow destabilization results from two competing mechanisms: pressure drop across the riser (mainly influenced by the volume fraction distribution) and pipeline compressibility.

The cycle begins with the blockage of the gas passage at the bottom of the riser, that is represented by a void fraction at the bottom of the riser (α) equals to zero. As the liquid coming from pipeline continues to flow in, as show the oil and water superficial velocities at the bottom of the riser (j_{ob} and j_{wb} respectively), and the gas that was already in the riser continues to flow out, as show the gas superficial velocity at the liquid level (j_{gu}), the riser column becomes heavier. Thus, the pressure at the bottom of the riser (P_b) increases continuously, causing the same behavior to the liquid penetration length (x), compressing the gas in the pipeline. The liquid level at the riser (z_u), which was at the lowest level of the cycle, also begins to increase since the liquid supply to the riser remains approximately constant. This stage is known as slug formation.

As the liquid level (z_u) reaches the top while the gas passage is kept blocked at the bottom, pressure reaches a maximum and there is only liquid and vaporized gas flowing in the riser. This is the slug production stage.

Observe that during the slug formation stage no vaporization occurs until the liquid level reaches the top, as can be observed from the gas superficial velocity at the liquid level (j_{gu}). From Eq. (38), as $\partial R_s/\partial P > 0$, the necessary condition for vaporization is $D_l P/Dt < 0$. As the inclination angle increases with position in the riser, the vertical velocity component of the liquid in the riser is smaller than the vertical velocity component of the liquid level. As a consequence, the liquid particles in the riser undergo higher pressures (mainly hydrostatic) as they flow upwards. When the liquid level reaches the top of the riser, the liquid particles undergo lower pressures as they flow upwards and vaporization occurs in the riser. The gas vaporization leads to a pressure decrease at the bottom of the riser due to the decrease in hydrostatic pressure drop; thus, in the slug production stage, vapor flowing out the riser comes from vaporization in the riser. This phenomenon prevents the pressure at the
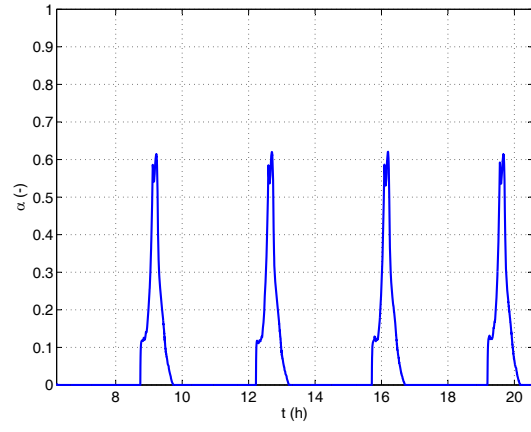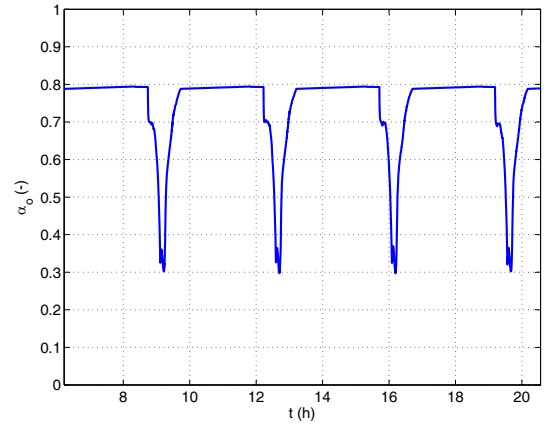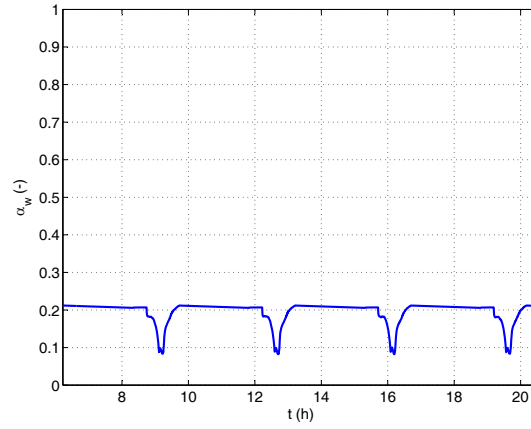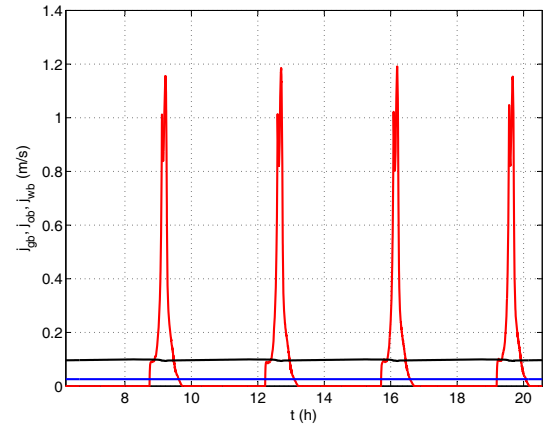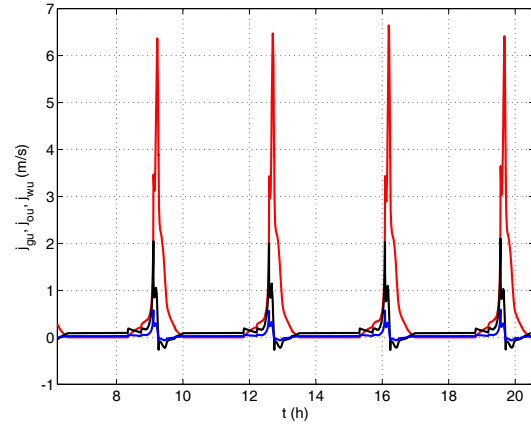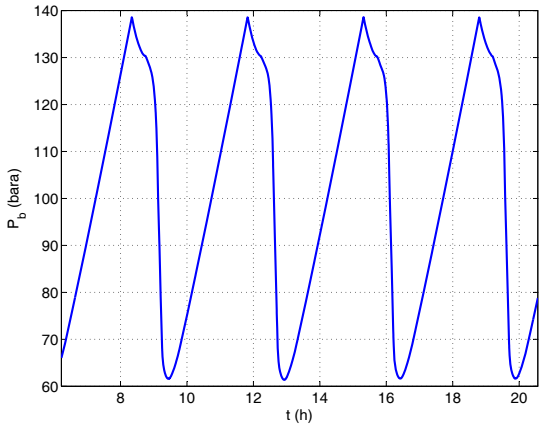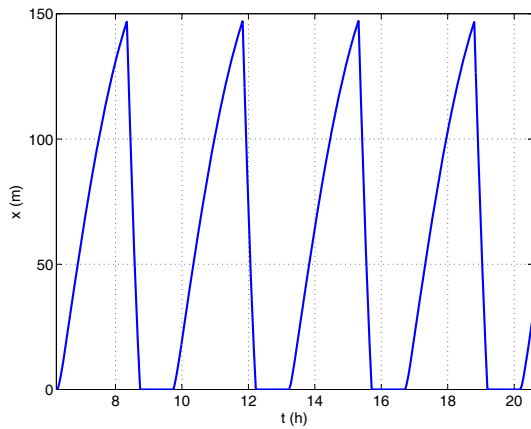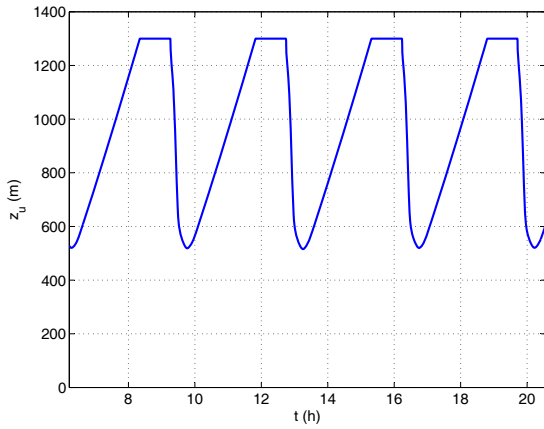
Figure 5: Simulation results.

bottom of the riser from keeping the maximum value during the production stage, as observed in severe slugging cycles for air-water systems.

As gas keeps flowing in the pipeline, the gas pressure eventually equals and overcomes the riser column pressure, so that the liquid accumulation front is pushed back until it reaches the bottom of the riser, starting the blowout stage. This stage is characterized by a peak in the gas superficial velocity (j_{gb}) and in the void fraction at the bottom of the riser (α).

As the gas penetrates into the riser the column becomes lighter, decreasing the pressure and then rising the gas flow. When gas reaches the top of the riser, gas passage is free through the stratified flow pattern in the pipeline and the intermittent/annular flow pattern in the riser, causing a violent expulsion and rapid decompression that brings the process to the slug formation again. This stage is known as gas blowdown. Observe that at the end of the decompression the pressure at the bottom of the riser (P_b) and the liquid level at the riser (z_u) reach the lowest levels in the cycle.

As can be seen from the figures, the periods associated to severe slugging are of the order of hours in offshore facilities; besides, fluctuations of all the variables are usually very high, compared to the stationary values. These long periods in which there is only liquid phase flowing, followed by relatively shorter periods in which there is a gas blowdown, cause instabilities in the liquid control systems of the separators and usually lead to platform shutdown. It is interesting to notice that the stationary state used as the initial condition does not exist, so any prediction made by a steady state program would be useless.

The simulation results shown in Fig. 5 characterize a severe slugging type 1 (SS1), in which the liquid slug has a length equal or greater than the riser length. Depending on the input parameters, other less dangerous severe slugging types exist: SS2 (in which the liquid slug length is smaller than the riser length) and SS3 (in which there is a continuous gas penetration at the bottom of the riser). The existence of SS1, SS2 and SS3 show a richness in the configuration states that can be attained by a pipeline-riser system, as well as poses a challenge in the model simulation capabilities.

Simulations with the same input variables shown in Table 1 were performed using the comercial software OLGA (OLGA, 2011). OLGA is a computational program developed to simulate multiphase flow in pipeline networks, with processing equipment included. The program solves separate continuity equations for the gas, liquid bulk and liquid droplets, two momentum equations, one for the continuous liquid phase and one for the combination of gas and possible liquid droplets and one mixture energy equation, considering that both phases are at same temperature. The equation are solved using the finite volume method and a semi-implicit time integration. Although the results obtained from the simulations using OLGA are not presented in this paper, they show a very

good agreement with the ones obtained with the present, simpler model.

8. CONCLUSIONS

The proposed pipeline-riser model was able to simulate the dynamics of the gas, oil and water flow in a typical offshore petroleum production system. It is important to notice that the model captures important physical effects in a highly nonlinear system such as moving liquid level at the riser and liquid penetration front at the pipeline. The model represents an advance compared to other works in literature (for air-water systems) in the sense that it considers oil-gas mass transfer, real gas behavior, fluids compressibility and local thermodynamic properties based on the black oil model. Further developments are being researched, such as the consideration of a variable void fraction in the pipeline, the inclusion of a choke valve at the top of the riser, the insertion of inertia terms in the momentum equations and the modeling of the gas lift. This is a work in progress.

REFERENCES

Ames, W.F., 1992. *Numerical methods for partial diferential equations*. Academic Press, Rio de Janeiro.

Baliño, J.L., Burr, K.P. and Nemoto, R.H., 2010. "Modeling and simulation of severe slugging in air-water pipeline-riser systems". *International journal of multiphase flow*, Vol. 36, pp. 643–660.

Beggs, H.D. and Robinson, J.R., 1975. "Estimating the viscosity of crude oil systems". *JPT*, pp. 1140–1141.

Bendiksen, K.H., 1984. "An experimental investigation of the motion of long bubbles in inclined tubes". *International journal of multiphase flow*, Vol. 10, No. 4, pp. 467–483.

Chen, N.H., 1979. "An explicit equation for friction factor in pipe". *Ind. Eng. Chem. Fundam.*, Vol. 18, pp. 296–297.

Chexal, B., Lellouche, G., Horrowitz, J. and Healer, J., 1992. "A void fraction correlation for generalized applications". *Progress in nuclear energy*, Vol. 27, No. 4, pp. 255–295.

Collins, A.G., 1987. *Petroleum Engineering Handbook*, SPE, Dallas, chapter Properties of produced waters.

Dranchuk, P.M. and Abu-Kassem, J.H., 1975. "Calculation of z-factors for natural gases using equations of state". *JCPT*, pp. 34–36.

Drew, D.A. and Passman, S.L., 1999. *Theory of Multicomponent Fluids*. Springer-Verlag, New York.

Fabre, J., Presson, L.L., Corteville, J., Odello, R. and Bourgeois, T., 1987. "Severe slugging in

pipeline/riser systems". *SPE 16846*, Vol. 5, No. 3, pp. 299–305.

Lee, A.L., Gonzalez, M.H. and Eakin, B.E., 1966. "The viscosity of natural gases". *Trans. AIME 237*, pp. 997–1000.

Masella, J.M., Tran, Q.H., Ferre, D. and Pauchon, C., 1998. "Transient simulation of two-phase flows in pipes". *Int. J. Multiphase Flow*, Vol. 24, pp. 739–755.

MATLAB, 2011. "Matlab - the language of technical computing". URL `http://www.mathworks.com`.

McCain, W.D., 1990. *The properties of petroleum fluids*. PennWell Books, Tulsa.

Nemoto, R.H. and Baliño, J.L., 2009. "Characteristic values and compatibility conditions for the no-pressure-wave model applied to petroleum systems". In *Proceeding of the 20th International Congress of Mechanical Engineering*. Gramado, RS, Brazil, paper code COB09-0748.

Ng, J.T.H. and Egboah, E.O., 1994. "An improved temperature-viscosity correlation for crude oil systems". CIM, Banff.

OLGA, 2011. "Olga - multiphase flow simulator". URL `http://www.sptgroup.com/Products/olga`.

Sarica, C. and Shoham, O., 1991. "A simplified transient model for pipeline-riser systems". *Chemical Engineering Science*, Vol. 46, No. 9, pp. 2167–2179.

Schmidt, Z., Brill, J. and Beggs, H., 1979. "Experimental study of severe slugging in a two phase flow pipeline-riser system". *SPE, Paper 8306*, Vol. 20, pp. 407–414.

Standing, M.B., 1981. *Volumetric and Phase Behavior of Oil Field Hydrocarbon Systems*. SPE, 9a. edição, Dallas.

Taitel, Y. and Dukler, A.E., 1976. "A model for predicting flow regime transitions in horizontal and near horizontal gas-liquid flow". *AIChE Journal*, Vol. 22, No. 1, pp. 47–55.

Taitel, Y., Vierkand, S., Shoham, O. and Brill, J.P., 1990. "Severe slugging in a riser system: experiments and modeling". *International journal of multiphase flow*, Vol. 16, No. 1, pp. 57–58.

Vasquez, M. and Beggs, H.D., 1980. "Correlations for fluid physical properties prediction". *JPT*, Vol. 32, pp. 968–970.

Velarde, J., Blasingame, T.A. and McCain, W.D., 1999. "Correlation of black oil properties as pressures below bubblepoint pressure – a new approach". *Journal of Canadian Petroleum Technology*, Vol. 38, No. 13, pp. 1–6.

Zuber, N. and Findlay, J., 1965. "Average volumetric concentration in two-phase flow system". *Journal of Heat Transfer*, Vol. 87, p. 453.

AUTHORS BIOGRAPHY

Rafael Horschutz Nemoto was born in Campinas, Brazil. He graduated from *Universidade de São Paulo*, Brazil (Mechatronics Engineering, 2008), with part of the credits fulfilled at *Technische Universität Darmstadt*, Germany. Since 2009 he is PhD student in Mechanical Engineering at *Universidade de São Paulo*. His research interests are Fluid Mechanics and Multiphase Flow.

Jorge Luis Baliño was born in Buenos Aires, Argentina. He graduated from *Instituto Balseiro*, Argentina (Nuclear Engineering, 1983, PhD in Nuclear Engineering, 1991). He worked for Techint S.A. (1983-1984), *Centro Atómico Bariloche* and *Instituto Balseiro* (1985-2000) in Argentina, *Instituto de Pesquisas Energéticas e Nucleares* (2001-2003) at São Paulo, Brazil. Since 2004 he is Professor at *Universidade de São Paulo*. His research interests are Fluid Dynamics, Heat Transfer and Multiphase Flow.

BOND GRAPH MODEL BASED FOR ROBUST FAULT DIAGNOSIS

R. El Harabi[(a,b)], B. Ould Bouamama[(a)], Mohamed Naceur Abdelkrim[(b)], Mohamed Ben Gayed[(b)]

[(a)] LAGIS, UMR CNRS 8146, Ecole Polytechnique Universitaire de Lille,
Cité Scientifique, 59655 Villeneuve d'Ascq Cedex, France.
[(b)] Unité de Recherche Modélisation, Analyse et Commande des Systèmes,
Ecole Nationale d'Ingénieurs de Gabès, 6029, Tunisie.

[(a)]rafika.El-harabi@polytech-lille.fr, [(b)] Belkacem.Ouldboumama@polytech-lille.fr

ABSTRACT

In this paper, robust Fault Detection and Isolation (FDI) design in nonlinear uncertain dynamic system, with chemical and thermodynamic phenomenon, is addressed. The methodology using a Bond Graph (BG) representation in linear fractional transformation (LFT) form is shown to be a valuable tool for developing dynamic threshold generators and achieving robustness against model uncertainty in combination with sensitivity to faults. The proposed FDI method is illustrated through an equilibrated reaction occurred in a continuous reactor coupled with a heat exchanger. Simulations are given to support the theoretical development and demonstrate the potential of the developed procedure.

Keywords: bond graph, chemical reactors, FDI design, dynamic threshold generators

1. INTRODUCTION

Due to the growing complexity of automatic control systems, there is an increasing demand for fail-safe operation, fault diagnosis (FD) and fault tolerance (FT). The early detection of system malfunctions and faults as well as the isolation of their origin have become an important issue in advanced control system design. Much attention has been paid to the design of robust fault detection and isolation systems (see for instance Blanke, Kinnaert, Lunze, and Staroswiecki (2006)).
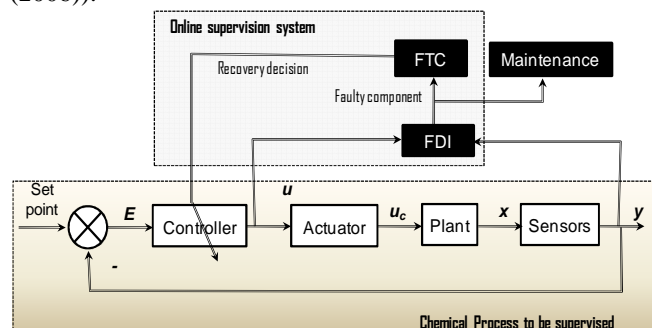
Figure 1: Supervision scheme in process engineering

Supervision of chemical reactors is a difficult task (as shown in Figure 1). This is due to several factors, such as the transient operation conditions, the various uses of these reactors and the evolution of the state variables which is nonlinear. The evolution of some parameters (the activation energy, pre-exponential factor, specific enthalpy) is non-stationary, which changes according to the condition variation inside the reactor Levenspiel (1999). It is this fact that has motivated our research in this paper.

Furthermore, due to the strong nonlinearities and parameters uncertainties in the chemical systems, their modelling is often complex and therefore less developed in the literature. The graphical modelling such as the bond graph tool becomes significant in this case, because it is appropriate for multiphysics modelling of complex and uncertain systems, as it is given in El Harabi, Ould Bouamama, M. El Koni Ben Gayed, and Abelkrim (2009). However, this tool can be used for residuals generation and monitorability analysis of uncertain systems Djeziri, Merzouki, Ould Bouamama and Dauphin Tanguy (2007).

The aim of presented paper is the design and analysis of a robust diagnosis scheme for nonlinear chemical processes taking into account the parameter uncertainties, described by coupled pseudo Bond Graph models using LFT form, when the secondary events (secondary reaction, hazard event of thermal runaway…ect) appear in chemical reaction. Thus, due to the energetic and multi physical properties of the Bond Graph, the whole of nonlinear model, structural analysis, residual with adaptive thresholds generations, and residual sensitivity analysis, can be synthesized using only one tool.

Section 2 gives a brief review of based element of coupled Bond Graph. The third section presents uncertain bond graph modelling and linear fractional transformations using in the chemical processes. In the fourth section, the bond graph LFT modelling of the chemical reaction in presence of parameter uncertainties is given. This section describes also the robust ARRs generation algorithm and the residual analysis. The developed methodology is applied for pseudo bond graph model based FDI of a continuous reactor coupled with a heat exchanger in section five. Finally some conclusions are drawn.

2. BASIC ELEMENT OF COUPLED BOND GRAPH

Bond graph models are network type models which are composed of multiports related by power bonds representing the (acausal) identity between pair of conjugated variables (named effort and flow) whose product is the instantaneous energy flow between the multiport elements. The multiport elements represent storage (**C-element**) (as compliance for instance or volume), inertia (**I-element**) (electrical inductance and mechanical inertia), energy dissipation (**R-element**) (electrical, mechanical or thermal friction), balance and continuity equations (the **0- and 1-junctions**) or inter-domain coupling (the **TF transformer** and **GY gyrator** elements). Finally to reproduce the architecture of the global system to be modelled, bond graph elements (R, C, I,..) are interconnected by a "0" junctions when they have a common effort and by "1" junction if their flow is the same.

In process engineering processes, several phenomena (chemical, thermal and fluidic) are coupled. In addition to matter transformation phenomena, chemical and electrochemical processes involve additional complexity in the modelling task, since the mass that flows through the process carries the internal energy which is stored in it, and which is thus transported from one location to another in a non-dissipative fashion. Power variables are thus in vectorial form:

$$E = \begin{bmatrix} e_h & e_t & e_c \end{bmatrix}^t, \ F = \begin{bmatrix} f_h & f_t & f_c \end{bmatrix}^t \qquad (1)$$

where e_h, e_t and e_c represent respectively the thermal effort (specific enthalpy h or the temperature T), the hydraulic effort (the pressure P), and the chemical effort (the chemical potential μ, chemical affinity A or the concentration c). f_h, f_t and f_c represent respectively the thermal (or entropy) flow (by conduction $\dot{Q}$ or by convection $\dot{H}$ i.e. enthalpy flow), hydraulic flow (mass flow $\dot{m}$ or volume flow $\dot{V}$) and chemical flow (molar flow $\dot{n}$).

Consider a thermofluid process (Figure 2 (a)) which consists of a pump (considered as a flow source) fulfilling a heated tank where a bottom pressure is measured by the sensor P_m, and the average temperature of the fluid is indicated by T_m. The coupled bond graph model in integral causality is given by Figure 2 (b). The two ports C_{ht} represents the coupled thermal and hydraulic energy of the stored fluid (considered here in under saturated state) is decoupled into thermal and hydraulic capacity C_t, and C_h. $Sf : \dot{Q}_{in}$, $Sf : \dot{m}_{in}$, and $Se : T_{in}$ represent, respectively, thermal flow source, inlet mass flow, and the temperature of the incoming fluid (considered constant). The coupling is modelled by the fictive R_c element in the thermal bond.

Another complexity can be added taking into account transformation of matter in chemical phenomena. The corresponding bond graph model is

given by Figure 3. The mixture of mass flow $\dot{m}_{in}$ is considered multicomponent with n species. The n transformers with $1 / Mi \ \left[Kg.mole^{-1} \right]$ as modulus used to transform massic flow $\dot{m}_{in}$ to molar flow $\dot{n}_{i,in}$ of ith specie:

$$\dot{n}_{i,in} = \frac{\dot{m}_{in}}{M_i} \qquad (2)$$

Figure 2: Heated tank (a) and its BG model in integral causality (b)
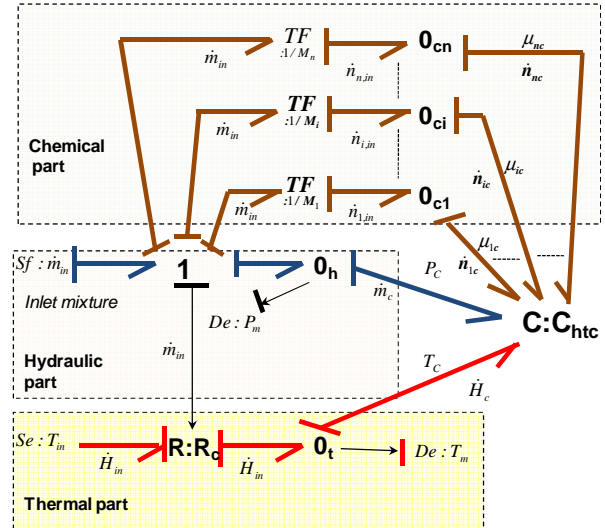
Figure 3: Bond graph model with three coupled energies

The state equation form $\dot{x} = f(x, u)$ well suited for control analysis, can be systematically deduced from a bond graph in a linear or non linear form. The input vector u is represented in a bond graph by the sources (Se and Sf), the measured variables are effort and flow detectors. The state vector is composed by the energy variables stored by C (general displacement) and I elements (impulse). The state vector does not appear on the Bond Graph, but only its derivative: The dimension of the state vector is equal to the number of C and I elements in integral causality. In the given dynamic model, there are $n + 2$ state variables: $x = \begin{bmatrix} m_c & H_c & n_{cn}...n_{cn} \end{bmatrix}$. They represent storage of number of mole for $n + 2$ species, total mass m_c and internal energy of the mixture H_c.

3. UNCERTAIN COUPLED BOND GRAPH

3.1. Uncertain Bond Graph interest

Various bond graph based qualitative and quantitative Maurya, Rengaswamy, and Venkatasubramanian (2007), FDI approaches have been developed to detect and isolate faults in single or piece-wise single energy domains, but none deal with FDI of coupled (energetic and transformation phenomena) nonlinear systems.

Among recent works that deal with parameter uncertainties modelling using bond graph approach, in Borutzky and Dauphin-Tanguy (2004), the authors proposed to construct in a systematic manner a bond graph from another bond graph using standard interconnection form, which is called the associated incremental bond graph (IBG).

In Sié Kam and Dauphin-Tanguy (2005), Djeziri, Ould Bouamama, and Merzouki (2009), authors proposed two methods for modelling uncertainties by using bond graph approach, applied on Electromechanical and thermodynamic systems (vehicle, test bench and steam generator) . The first method is based on describing parameter uncertainties as bond graph elements, and the second method introduces the LFT form for uncertainties modelling. Here, this problem is addressed using the linear fractional transformation (LFT) paradigm.

After the pioneering work of Oster and Perelson, it has been mainly used for membrane processes some reaction processes and some electrochemical processes Couenne, Jallut, Maschke, Breedveld, and Tayakout (2006), Khaled, Bouamama, and Nakrachi (2006). Bond graph modeling has been used for hydraulic and thermal domain in chemical reactor but not for monitoring and observing kinetic and thermodynamic evolution of chemical mixture. Thus, uncertain bond graph modeling of chemical reaction are not treated until now in literature and diagnosis of chemical reaction is an open research work.

3.2. BG-LFT form

The principle of the uncertainties representation using LFT consists in building the uncertain system in the form of a looping between the increased invariant system M, whose parameters are defined perfectly, and a block of uncertainty, noted Δ, gathering various uncertainties, Figure 4. Setting of LFT form requires that the system must be reachable and observable. These properties are necessary conditions for the monitoring ability of the system [41].
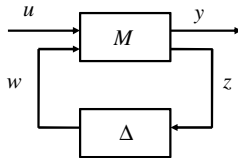
Figure 4: LFT representation

The interconnection structure induces the following state equations:

$$\begin{cases} w = -\Delta\, z \\ \dot{x} = Ax + B_1 w + B_2 u \\ z = C_1 x + D_{11} w + D_{12} u \\ y = C_2 x + D_{21} w + D_{22} u \end{cases} \quad (3)$$

where $x \in \Re^n$ the state vector, $u \in \Re^m$ the inputs vector, $y \in \Re^p$ the outputs vector. $w \in \Re^l$ and $z \in \Re^l$ are ,respectively, the auxiliary input and output vectors. n, m, l et p are positive entireties. A, B_1, B_2, C_1, C_2, D_{11}, D_{12}, D_{21}, D_{22} are appropriate ranks matrices.
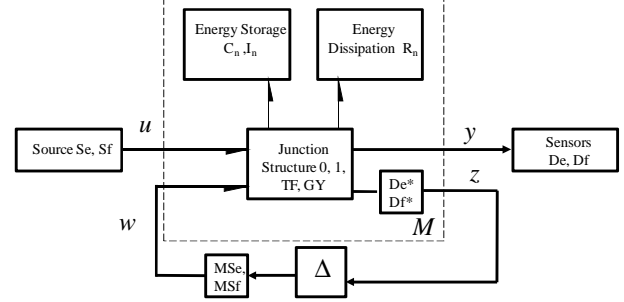
Figure 5: BG-LFT representation

In LFT bond graph representation, parameters uncertainties are represented under multiplicative form at the level of bond graph component. The method consists in replacing each uncertain element by its BG-LFT. BG-LFT representation is shown in Figure 5.

The advantage of approach BG-LFT compared to an approach of LFT state is summarized in two points: complexity in the model construction and the uncertainties structure on the model Djeziri, Merzouki, Ould Bouamama and Dauphin Tanguy (2007).

Modeling of bond graph elements $i \in \{R, I, C, TF, GY\}$ in the LFT form consists in decoupling the nominal element $i_n \in \{R_n, I_n, C_n, TF_n, GY_n\}$ part from its uncertain part $\delta_i i_n \in \{\delta_R R_n, \delta_I I_n, \delta_C C_n, \delta_{TF} TF_n, \delta_{GY} GY_n\}$, with δ_i is a multiplicative uncertainty on the parameter i. In the combined BG-LFT representation, the parameter uncertainties are explicitly represented under multiplicative form for each bond graph element. The additive uncertainties of the parameters are related to their multiplicative values by the following relations: $\delta_i = \dfrac{\Delta i}{i_n}$, where Δi is the additive uncertainty values on the bond graph element i.

The principle of this modeling consists in representing the influence of the parameter uncertainty, by a fictive effort or flow input ($MSe : w_i$ or $MSf : w_i$), modulated by $\delta_i (i_n e_i)$ or $\delta_i (i_n f_i)$. Details on this modelling procedure are given in Djeziri, Ould Bouamama, and Merzouki (2009).

In chemical processes, to explain the modelling in LFT form using the bond graph, let us consider the

Multiport $\mathbb{R}$, we know that the R-elements dissipate power and that this power comes out as heat. So including thermal effects, an R-element becomes an irreversible and power conserving structure. It is denoted as multiport $\mathbb{R}$ (see Figure 6 (a)). So power can flow only as indicated by the half arrows, and not backwards. In other words it cannot become negative. So, when we are not interested in thermal effects, we speak of R-elements and multiport-R, otherwise of multiport-$\mathbb{R}$. Regarding the multiport $\mathbb{R}$, it can have bonds with several strands as shown on Figure 6 (b). With multiport $\mathbb{R}$, irreversibility and energy conservation of multiport R are as follows: with several strands, only the sum of the non thermal bonds must be positive, but in single strands power can become negative as long as it is more positive in others. One can also say that power in the thermal bond must by always positive.
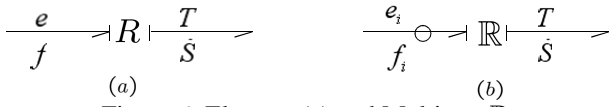
$$(a) \qquad (b)$$
Figure 6: Element (a) and Multiport $\mathbb{R}$

In a chemical reaction, the product of chemical affinity A by the global reaction rate J is a power. The thermal loss (transformation of chemical to thermal energy) is modeled by an active resistance Multiport $\mathbb{R}$ (a resistance which generates entropy) Thoma and Ould Bouamama (1999). The multiport absorbs chemical power $A \times J$ and product an equivalent quantity in thermal power $T \times S$. Thus, a RS-field is used as a link between the mass and energy parts of the reactor vessel subsystem. It is a two-port element connecting the molar and energy balances.

The characteristic law of Multiport $\mathbb{R}$ in resistance causality with uncertainty can be written as follows:

$$
\begin{aligned}
\dot{S} &= \Phi(\mathbb{R}_n, A)(1 + \delta_{1/\mathbb{R}}) \\
&= \Phi(\mathbb{R}_n, A) + \delta_{1/\mathbb{R}} \Phi(\mathbb{R}_n, A) \\
&= \dot{S}_n + \dot{S}_{inc}
\end{aligned} \qquad (4)
$$

The effort A is known at the entry of the multiport $\mathbb{R}$. $\dot{S}_n$, $\dot{S}_{inc}$, $\delta_{1/\mathbb{R}}$ represent, respectively, the nominal value, the multiplicative uncertainty. $\Phi(\mathbb{R}_n, A)$ is the global reaction rate and is written as

$$
\Phi(\mathbb{R}_n, A) = r_f \left(1 - \exp\left(\frac{A}{RT_r}\right)\right) V \qquad (5)
$$

where the reaction $r_f = k_0 \exp\left(\frac{-E_a}{RT_r}\right) \prod C_i^{v_i}$ represents non linear term which depends on reactional temperature T_r, according to Arrhenius equation, and concentrations ($C_i = \frac{n_i}{V}$). R is a universal gas constant, E_a is the activation energy of the reaction and k_0 is the pr-exponential factor.
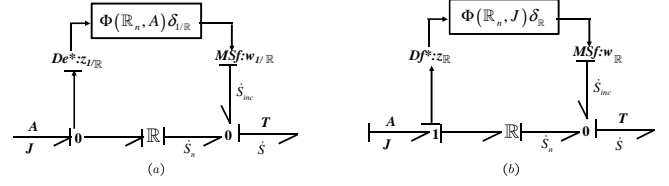
Figure 7: Multiport $\mathbb{R}$ in resistance causality using the LFT form(a) Multiport $\mathbb{R}$ in conductance causality using the LFT form (b)

The characteristic law of R element in conductance causality is given as follows:

$$
\begin{aligned}
\dot{S} &= \Phi(\mathbb{R}_n, J)(1 + \delta_{\mathbb{R}}) \\
&= \Phi(\mathbb{R}_n, J) + \delta_{1/\mathbb{R}} \Phi(\mathbb{R}_n, J) \\
&= \dot{S}_n + \dot{S}_{inc}
\end{aligned} \qquad (6)
$$

The flow J is known at the entry of the multiport $\mathbb{R}$.

In the next section will be considered the use of coupled bond graph for FDI design.

4. FAULT INDICATORS GENERATION FROM BOND GRAPH

In this paper, a bond graph methodology is used to synthesis a robust FDI method for nonlinear system in presence of parameter uncertainties, Figure 8.

Parametric uncertainties are explicitly appears on the BG, one can automatically generate the robust ARRs for the uncertain system by decoupling the nominal and the uncertain parts; residuals correspond to the ARR nominal part, while the residual thresholds represents the ARR uncertain parts.

The main advantages of the bond graph model in LFT form for robust diagnosis are given as follows:
- Introduction of the uncertainties on the nominal model, does not affect the causality and the structural properties of the BG elements;
- Representation of all uncertainties (i.e. structured and unstructured);
- Uncertain part is perfectly separated from the nominal part;
- Parameter uncertainties are easily evaluated.

This FDI method is summarized by the following steps:
i) Modeling of studied system using bond graph tool with standard LFT form; ii) Generation of Analytical Redundancy Relations (ARRs) from the uncertain model by decoupling the nominal and the uncertain parts. Residuals correspond to the ARR nominal part, while their adaptive thresholds represent the ARR uncertain parts; ii) Residual' sensitivity analysis is done by using the ARR uncertain part.
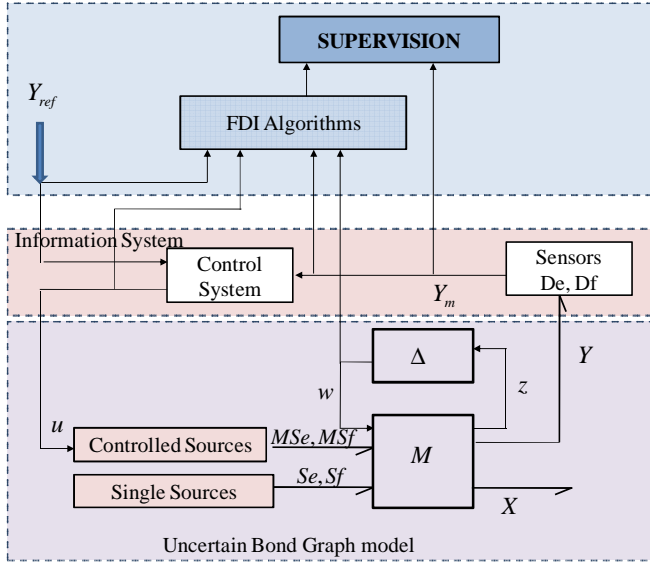
Figure 8: Representation of the robust FDI scheme using bond graph tool

5. CASE STUDY: A CHEMICAL REACTOR

5.1. Process Description

Let us consider an adiabatic Continuous Stirred Tank Reactor, where the exothermic reversible reaction is occurred. This reaction is defined as follows:

$$\nu_A A \underset{A_r}{\overset{A_f}{\rightleftharpoons}} \nu_C C \qquad (7)$$

where ν_i (for i=A, C) are the stoichiometric coefficients. In our case these coefficients are equal to one.

The technological diagram of reactor system is depicted in Figure 9. The supply system (component A) consists of a storage tank and a pump. The level regulation is guaranteed by the means of a PI regulator acting on a centrifugal pump which supplies continuously the tank.
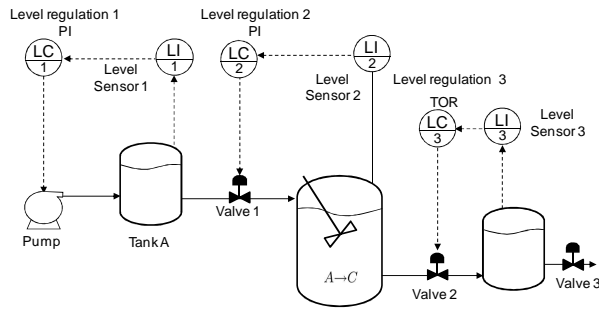
Figure 9: Technological diagram of the process

The level controller in the reactor is ensured by a regulator which acts on a valve at the reactor input. The tank containing the components (C, A) is controlled in level by a regulator which acts a self-closing valve on the outlet side of the reactor.

5.2. Word Bond Graph model

The modelling hypotheses are, the reactor is perfectly stirred so that temperature and concentrations of different chemical species are homogeneous in all the reaction mixture, the reaction mixture is composed of one homogeneous liquid phase, and no phase change is considered, the volume of the liquid in the tank is constant. For illustration of developed method and because of limited space, we consider only the main component of the system: reactor vessel especially chemical domain.

The word bond graph model is presented in Figure 10. This model is decomposed into several modules, linked by a pair of pseudo power variables (effort-flow). To simplify the process modelling, we introduced bond graph model of reactor vessel which is composed of several parts corresponding to multi-energy domains.

The used pseudo power variables (effort-flow) are: pressure-mass flow $(P, \dot{m})$, temperature-enthalpy flow $(T, \dot{H})$ in the case of convection, and temperature-thermal flow $(T, \dot{Q})$ in the case of conduction, chemical potential-molar flow $(\mu, \dot{n})$, chemical affinity-reaction velocity (A, J).
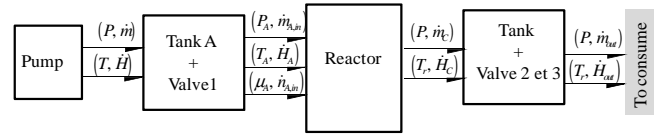
Figure 10: Word pseudo-bond graph of the chemical reactor

5.3. Pseudo Bond Graph model

The bond graph model is given (Figure 11). This part includes chemical subsystem in reactor vessel. The bond graph transformers $TF : 1/\nu_A$ and $TF : \nu_C$ represent a chemical transformation. Their modulus is the stoichiometric coefficients (the chemical affinity A represent the driving force in reactor vessel).

In the chemical domain the 0-junctions represent the molar balance of each component (A, C). 1-junction is used to represent the equality of the molar reaction flows of the different substances involved.

Thus, a $\mathbb{R}$-field is used as a link between the mass and energy parts of the reactor vessel subsystem. It is a two-port element connecting the molar and energy balances.
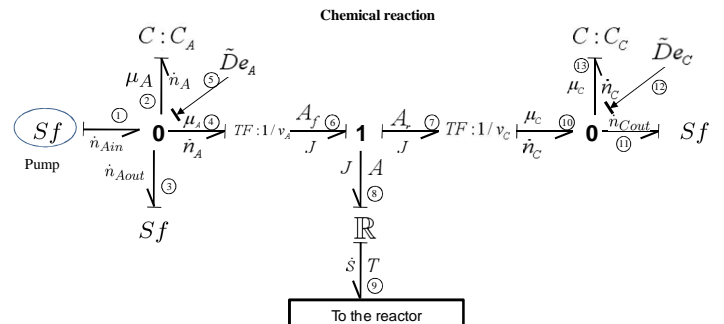
Figure 11: BG determinist model in preferred derivative causality of chemical domain

Modeling of bond graph elements $i \in \{C_A, C_C\}$ and multiport $\{\mathbb{R}\}$ in the LFT form consists in decoupling the nominal element $i_n \in \{C_{A,n}, C_{C,n}, \mathbb{R}_n\}$ part from its uncertain part $\delta_i i_n \in \{\delta_1 C_{A,n}, \delta_2 C_{C,n}, \delta_3 \Phi(\mathbb{R}_n, A)\}$, with δ_i is a multiplicative uncertainty on the parameter i.

The determinist and uncertain bond graph model of the chemical domain are respectively given in Figures 11 and 12. The symbols $\tilde{D}e$ and $\tilde{D}f$ correspond to virtual sensors. They are used to distinguish the real measurements from the fictive ones.

The storage of chemical energies is modelled by the bond graph elements $C : C_A$ and $C : C_B$. Then the following equation is deduced from the junction 0 of the bond graph determinist model in derivative causality:

$$\begin{cases} \dot{n}_A = \dfrac{V}{RT} \exp\left(\dfrac{\mu_A - \mu_A^0}{RT} \right) \dot{\mu}_A \\ \dot{n}_C = \dfrac{V}{RT} \exp\left(\dfrac{\mu_C - \mu_C^0}{RT} \right) \dot{\mu}_C \end{cases} \tag{7}$$

where μ_A^0 is standard chemical potential.

In general case, the previous equations becomes

$$\dot{n}_i = C_i \frac{d\mu_i}{dt} \tag{8}$$

where $\dot{n}_i$ is the reaction output' molar flow. μ_i is chemical potential inside the reaction. C_i represents the chemical capacity of the reaction and can be expressed as follows:

$$C_i = \frac{V}{RT} \exp\left(\frac{\mu_i - \mu_i^0}{RT} \right) \tag{9}$$

The relation between $C_{i,n}$ and δ_{C_i} is given by the following expression:
$$C_i = C_{i,n} + \delta_{C_i} C_{i,n} \tag{10}$$
where $C_{i,n}$ is the nominal value of C_i.

The modulated input $w_i (i = 2,3)$ in Figure 12 corresponds to an effort variable deduced from δ_{C_i} and expressed by the following equation:

$$w_i = -\delta_{C_i} C_{i,n} \frac{d\mu_i}{dt} \tag{11}$$

w_i is taken with a negative sign, because it is considered as a fictive flow input' source.
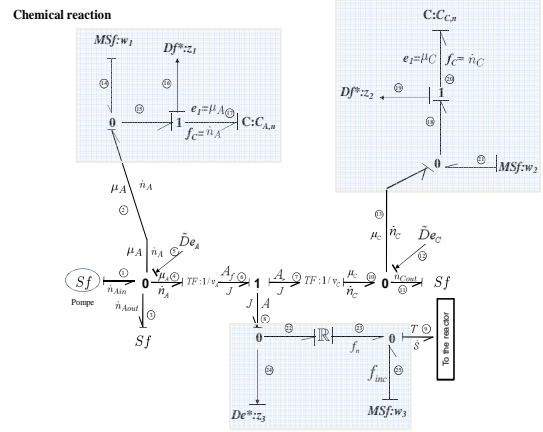
Figure 12: BG-LFT uncertain model in preferred derivative causality

5.4. Design of supervision system

A method to derive ARR from bond graph models by applying the causality inversion algorithm, have been presented in Djeziri, Merzouki, Ould Bouamama and Dauphin Tanguy (2007), which use structural and causal properties.

5.4.1. Determinist ARRs Generation:

The ARRs are deduced from junctions 0 that contain detectors on the nominal bond graph model of Figure 12. The unknown variables f_2 and f_4 are eliminated using covering causal paths from detectors to unknown variables.

From first junction 0
$$r_1 = f_1 + f_4 - f_3 - f_2 = 0 \tag{12}$$

with $f_1 = Sf = \dot{n}_{Ain}$, $f_3 = Sf_{out}$, $f_2 = C_A \dfrac{de_2}{dt}$

$$f_4 = v_A f_6 = v_A J = v_A \Phi(\mathbb{R}, A) = v_A r_f \left(1 - \exp\left(\frac{A}{RT_r} \right) \right) V$$

where $\Phi(\mathbb{R}, A)$ is given by the equation (5).

f_2 is calculated (eliminated) from the following causal paths $f_2 \to \Phi_{C_A} \to e_2 \to \tilde{D}e_A : \mu_{A,m}$

The first ARR, r_1, is deduced from equation (12) and is given as

$$\begin{aligned} r_1 &= \dot{n}_{Ain} + v_A \Phi(\mathbb{R}, A) - Sf_{out} - C_A \frac{d\tilde{D}e_A}{dt} \\ &= \dot{m}_{in} \frac{C_{Ain}}{\rho_A} + v_A \Phi_R(n_A, m, H) - \dot{m}_{out} \frac{n_A}{m_{mé}} - \dot{n}_{A,m} = 0 \end{aligned} \tag{13}$$

with $m_m = \rho V = \rho S L_m$ (S sectional surface of the reactor) and $\dot{n}_{A,m} = \dfrac{V}{RT_r} \exp\left(\dfrac{\mu_{A,m} - \mu_A^0}{RT_r} \right) \dot{\mu}_{A,m}$.

From second junction 0
$$r_2 = f_{10} - f_{11} - f_{13} = 0 \tag{14}$$

and from the constraint in equation (14), the second ARR, r_2, is given by

$$r_2 = f_{10} - f_{11} - f_{13} = v_C \Phi(\mathbb{R}, A) - Sf_{out} - C_C \frac{d\tilde{D}e_C}{dt}$$

$$= \dot{m}_{out} \frac{n_C}{m_{m\acute{e}}} - v_C r_f \left(1 - \exp\left(\frac{A}{RT_r}\right)\right) V - \dot{n}_{C,m} = 0 \tag{15}$$

The fault in chemical domain (appearance of secondary event: release of toxic or explosive material, etc.) related to transformer phenomenon can be detected by using the first and the second ARR.

5.4.2. Robust ARRs' Generation

In this section, the ARRs are generated for nonlinear systems, using bond graph approach in the LFT form. The aim of the robust diagnosis for the presented chemical reaction is to detect and isolate a chemical fault situation (appearance of secondary reaction when the reaction takes place; undesirable product and runway of the reaction) in presence of parameter uncertainties. This fault corresponds to the increase of the reaction velocity and chemical affinity, which is distinguished from the parameter uncertainties.

The chemical reaction model in the LFT form with derivative causality, after sensors dualization is given in Figure 12. The fictive inputs $w_i (i = 1,...,3)$ are related with the fictive outputs $z_i (i = 1,...,3)$ and expressed in the system of (16)

$$\begin{cases} w_1 = -\delta_1 z_1; z_1 = \dot{n}_A = C_A \frac{d\tilde{D}e_A}{dt} = C_A \frac{d\mu_A}{dt} = \frac{V}{RT} e^{\left(\frac{\mu_A - \mu_A^0}{RT}\right)} \frac{d\mu_A}{dt} \\[2mm] w_2 = -\delta_2 z_2; z_2 = \dot{n}_C = C_C \frac{d\tilde{D}e_C}{dt} = C_C \frac{d\mu_C}{dt} = \frac{V}{RT} e^{\left(\frac{\mu_C - \mu_C^0}{RT}\right)} \frac{d\mu_C}{dt} \\[2mm] w_3 = \delta_3 \Phi(\mathbb{R}_n, A) z_3; z_3 = e_6 = A_f = v_A \mu_A \\[2mm] \qquad = v_A \left(\mu_A^0 + RTLog\left(\frac{\int \dot{n}_A dt}{V}\right) \right) \end{cases}$$

where δ_1 and δ_2 represent, respectively, multiplicative uncertainties on the energy accumulation of reactant A and product C (leads to uncertainties in heat-storage capacity). δ_3 is the multiplicative uncertainty on the reaction velocity (leads to uncertainties in activation energy, pre-exponential factor, enthalpy...).

The two parts of the ARRs generated from chemical reaction model with parameter uncertainties of Figure 12 are given by equations (12) and (15) where r_1 and r_2 represent the ARRs nominal parts that describe the system operating. a_1 and a_2 (the ARRs uncertain part) represent the intake reduced by the parameter uncertainties such as flow or effort which affect the residuals. It is described by the sum of fictive input values and is used to calculate the normal operating thresholds.

$$\begin{cases} a_1 = |w_1| + |v_A w_3| \\ a_2 = |w_2| + |v_C w_3| \end{cases} \tag{17}$$

Uncertain ARR part cannot be quantified perfectly, it is evaluated to generate a normal operation' threshold which satisfies the following inequality:
$$-a \leq ARR \leq a \tag{18}$$

5.5. Simulations results

The chemical system is instrumented with the following sensors. The mixture temperature inside the tank De : Tm1, the level inside the tank De : Lm1, the flow sensor (Df : Fm1) is used to measure the amount of mixture leaving the tank. The water flow in the cooling circuit can be measured using the flow sensor (Df : Fm2). The output control signal of each controller is considered as a known value.

Figure 13 show respectively the residuals r_1 and r_2 without faults.

Fault scenario: Appearance of the secondary reaction: It is supposed now for example that the cooling system is never failing and that the exits of the regulators and the sensors are always correctly measured. A sudden appearance of secondary product occurs between 30 and 60 min. Indeed, to stop the evolution of the secondary reaction and to eliminate these effects in real-time, it is necessary to add a reagent able to eliminate the undesirable products. As can be seen in Figure 14, the appearance of undesirable product is detected perfectly by the residual evolution. The fault is detected perfectly, as it is alarmed by two residuals r_1 and r_2, and not by the other residuals. The thresholds of normal operation are given with dot lines.
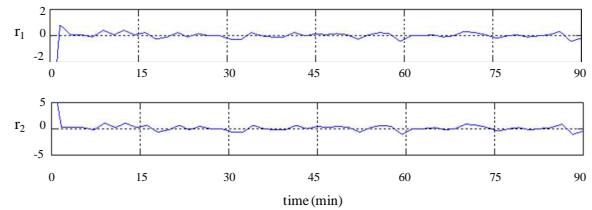

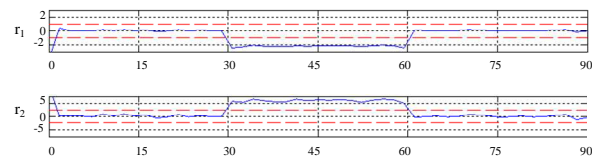Figure 13. Residual evolutions of normal system

Figure 14. Residual evolutions of faulty system

In order to explain the appearance of secondary reaction; for example, appearance of undesirable product E which modify reaction dynamics. Namely an unmodeled side reaction, is added to the simulation model; in detail the reaction scheme becomes:

$$\nu_A A \underset{A_r}{\overset{A_f}{\rightleftharpoons}} \nu_C C + \nu_E E \tag{19}$$

and consequently the mass balance and the fault indicator will be modified. The RRAs (r_1 and r_2) should add the term $v_E \mu_{E,m}$ and can not be equal to zero.

6.   CONCLUSION

In this paper, a robust FDI with respect to parameter uncertainties is given using bond graph modelling approach in the LFT configuration. The robust ARR are generated directly from a bond graph model. This approach is study for complex systems where numerical values of parameters are not available. The obtained results are validated using real process (continuous reactor). The proposed FDI method witch can detect kinetic and thermodynamic drift of chemical reactors due to appearance of secondary reaction. The performances obtained are acceptable.

ACKNOWLEDGMENTS

REFERENCES

Blanke, M., Kinnaert, M., Lunze, J. and Staroswiecki, M., 2006 *Diagnosis and Fault-Tolerant Control*. Springer Verlag.

Borutzky, W. and Dauphin-Tanguy, G., 2004. *Incremental bond graph approach to the derivation of state equations for robustness study*. simulation. ModellingPractice Theory, 12:41-60.

Couenne, F., Jallut, C., Maschke, B., Breedveld, P.C., and Tayakout, M., 2006. *Bond graph modelling for chemical reactors*. Math. Comput. Modell. Dyn. Syst, 12:159.

Djeziri M.A., Merzouki, R., Ould Bouamama, B. and Dauphin Tanguy, G., 2007. Bond graph model based for robust fault diagnosis. *in: Proceedings of the 2007 American Control Conference*, New York, USA, p. 3017–3022.

Djeziri, M.A., Merzouki, R., Ould Bouamama, B., and Dauphin Tanguy, G., 2007. Bond graph model based for robust fault diagnosis. *Proceedings of the 2007 American Control Conference*, New York, USA:3017-3022.

Djeziri, M. A., Ould Bouamama, B., and Merzouki, R., 2009. *Modeling and robust FDI of steam generator using uncertain bond graph model*. Journal of Process Control, 19:149-162.

El Harabi, R., Ould Bouamama, B., El Koni Ben Gayed, M., and Abelkrim, M.N., 2009. FDI of chemical process using bond graph tool. *In 14th IEEE IFAC International Conference on Methods and Models in Automation and Robotics*. August 19-21, Miedzyzdroje, Poland.

Khaled, A.R., Bouamama, B.O., and Nakrachi, A., 2006. Generic bond graphs procedure for chemical reactions modelling , *computational engineering in systems applications*. IMACS IEEE "CESA'06", 1:412-417.

Levenspiel, O., 1999. *Chemical Reaction Engineering*. John Wiley and Sons.

Maurya, M.R., Rengaswamy, R., and Venkatasubramanian, V., 2007. *A signed directed graph and qualitative trend analysis-based framework for incipientfault diagnosis*. Chemical Engineering Research and Design, 85(10):1407-1422.

Sié Kam, C. and Dauphin-Tanguy, G., 2005. *Bond graph models of structured parameter uncertainties*. Journal of the Franklin Institute, 342:379-399.

Thoma, J. and Ould Bouamama, B., 1999. *Modelling and Simulation in Thermal and Chemical Engineering: A Bond Graph Approach*. Springer Engineering, ISBN 3-540-66388-6.

AUTHORS BIOGRAPHY

Rafika EL HARABI is PhD. student at the Ecole Polytechnique de Lille in France and Ecole Nationale d'Ingénieurs de Gabès in Tunisia. She is preparing her PhD. thesis in LAGIS (Laboratory of Control Engineering, Computer Science & Signal), UMR CNRS 8146, Lille, France and in MACS (Modelling, Analyze and Systems Control), UR 11-12, Gabès, Tunisia.

Belkacem OULD BOUAMAMA is full Professor and head of the research at « Ecole Polytechnique Universitaire de Lille, France) » His main research areas developed at the Laboratoire d'Automatique Génie Informatique et Signal de Lille (UMR CNRS8146) concern *Integrated Design for Supervision of System Engineering*. Their application domains are mainly nuclear, petrochemical, and mechatronic systems. He is the author of several international publications in this domain. He is co-author of three books in bond graph modeling and Fault Detection and Isolation area. Research and teaching activities can be consulted at: http://sfsd.polytech-lille.net/BelkacemOuldBouamama

Mohamed Kony BEN GAYED is a Professor at ISET Gabès and he is searcher in MACS (Modeling, Analyse and systems Control), UR 11-12, Gabès, Tunisia.

Momamed Naceur ABELKRIM is a Professor and head of the «Ecole Nationale d'Ingénieurs de Gabès in Tunisia » and he is searcher in MACS (Modeling, Analyze and Systems Control), UR 11-12, Gabès, Tunisia.

GEOMETRIC PROPERTIES OF KERNEL PARTIAL LEAST SQUARES FOR NON-LINEAR PROCESS MONITORING

José L. Godoy[a], Germán Bustos[b], Alejandro H. Gonzalez[c] and Jacinto L. Marchetti[d]

[a,b,c,d] INTEC (CONICET and Universidad Nacional del Litoral), Güemes 3450 (3000), Santa Fe, Argentina

[a]jlgodoy, [b]gbustos, [c]alejgon, [d]jlmarch@santafe-conicet.gov.ar

ABSTRACT

This work proposes a new strategy for monitoring non-linear processes based on Kernel Partial Least Squares (KPLS). When strongly non-linear process are considered, a PLS regression model could not be enough accurate. So, the first stage of the proposed method is to map the input data to a high-dimension space, where a linear regression model can be obtained. Then, an *implicit* linear regression model relating the high-dimension space with output space (output data) is obtained. This model implicitly induces a decomposition of the high-dimension space into the Model subspace and the complementary Residual subspace, being the vectors in the first subspace the effective domain of the linear regression model. Finally, once the space decomposition is understood, new statistics (metrics) for each subspace are proposed to monitor the process and detect possible abnormal behaviours. The effectiveness of the method is tested by means of a synthetic simulation example from the literature.

Keywords: KPLS, space decomposition, multivariate process monitoring, fault detection indexes.

1. INTRODUCTION

KPLS is a promising regression method for deal with nonlinear problems because it can efficiently compute regression coefficients in high-dimensional space by means of the nonlinear kernel function. It is also an efficient method for estimating and predicting quality variables in the strongly nonlinear process by mapping data from the original space into a high-dimensional space. It only requires the use of linear algebra, making it as simple as linear multivariate projection methods, and it can handle a wide range of nonlinearities because of its ability to use different kernel functions. Its application results from a simple example show that the proposed method can effectively capture the nonlinear relationship among variables.

The need for associating input and output data sets obtained by online data login of complex process variables constitutes a problem that requires increasing attention. Lately, KPLS has become a powerful approach to find multivariable non-linear structures, mainly because large non-linear correlations implicit in the data can be adequately overcome.

Fault detection makes use of the so-called fault detection indices based on model. A fault is detected when one of the fault detection indices is beyond its control limit. Once a fault is detected, it is necessary to diagnose its cause.

In this work, the implicit space decomposition (made by KPLS) is obtained and their main geometric properties are used to design a non-linear monitoring strategy using new fault detection indices actuating on each subspace. This allows us to diagnose the fault type. The main contributions of this work are then the introduction of new fault detection indices derived from the KPLS decomposition, and a diagnosis tool based on the statistics that triggers the alarm condition.

2. KERNEL PLS REGRESSION

As usual in data driven method, we first should collect a set of N samples of the predictor vector, $\left\{ \mathbf{x}_i \right\}_{i=1}^{N}$, with $\mathbf{x}_i \in \Re^m$ for $i, \cdots, N$, and N samples of the response vector, $\left\{ \mathbf{y}_i \right\}_{i=1}^{N}$, with $\mathbf{y}_i \in \Re^p$ for $i, \cdots, N$, which are called Identification Data Set (IDS). Once these data are collected, the Kernel Partial Least Squares (KPLS) method maps the predictor vectors, $\mathbf{x}_i$, from $\Re^m$ to a high-dimension space $\Re^c$ (with $c \gg m$), where a linear PLS regression model can be created to relate vectors in $\Re^c$ with the response vectors, $\mathbf{y}_i$, in $\Re^p$. In this way, a latent model can be formulated in $\Re^c$, in order to extend linear PLS to non-linear kernel PLS (Rosipal and Trejo, 2001). The non-linear transformation that maps vectors from $\Re^m$ to $\Re^c$ is not made by means of an explicit non-linear function $\boldsymbol{\varphi}$, but by means of a kernel function k selected to compute the following inner products:

$$k\left(\mathbf{x}_i, \mathbf{x}_j\right) = \boldsymbol{\varphi}\left(\mathbf{x}_i\right)' \boldsymbol{\varphi}\left(\mathbf{x}_j\right),$$
$$\text{for } i = 1, \cdots, N, \text{ and } j = 1, \cdots, N. \quad (1)$$

Notice that trough the introduction of the kernel trick (Rosipal and Trejo, 2001), $\boldsymbol{\varphi}\left(\mathbf{x}_i\right)' \boldsymbol{\varphi}\left(\mathbf{x}_j\right) = k\left(\mathbf{x}_i, \mathbf{x}_j\right)$, one can avoid both, performing explicit nonlinear mapping and computing dot products in the high-dimensional space. Furthermore, as it is known and will be seen later, the PLS regression method need only dot products to perform the regression. Now, from Eq. (1) it is obtained the *Gram Kernel* matrix, $\mathbf{K} \in \Re^{N \times N}$, defined as follows:

$$\mathbf{K} = \boldsymbol{\Phi}\boldsymbol{\Phi}', \quad \boldsymbol{\Phi} = \left[\boldsymbol{\varphi}\left(\mathbf{x}_1\right) \cdots \boldsymbol{\varphi}\left(\mathbf{x}_N\right) \right]'. \quad (2)$$

The entries of this last matrix are the cross inner products of all mapped predictor vectors, $\{\varphi(\mathbf{x}_i)\}_{i=1}^{N}$.

As in the PLS method, it is assumed here a non-linear KPLS model with zero mean. To center the mapped data in the high-dimensional space, $\Re^c$, matrix $\mathbf{K}$ should be substituted by the centered matrix $\bar{\mathbf{K}}$, given by:

$$\bar{\mathbf{K}} = \bar{\mathbf{\Phi}}\bar{\mathbf{\Phi}}' = \mathbf{K} - \mathbf{KE} - \mathbf{EK} + \mathbf{EKE}, \qquad (3)$$

where $\mathbf{E}$ is a matrix with unitary entries, and $\bar{\mathbf{K}}$ and $\bar{\mathbf{\Phi}}$ are the centered versions of $\mathbf{K}$ and $\mathbf{\Phi}$.

Now, from matrices $\bar{\mathbf{K}}$ and $\mathbf{Y} = [\mathbf{y}_1, \cdots, \mathbf{y}_N]' \in \Re^{N \times p}$, the training KPLS algorithm is derived as a sequence of modified NIPALS steps, as follows (Rosipal and Trejo, 2001):

First, set the index $a=1$, $\bar{\mathbf{K}}_1 = \bar{\mathbf{K}}$ and $\mathbf{Y}_1 = \mathbf{Y}$. Then:

1. Set the vector of $\mathbf{Y}_a$-scores $\mathbf{u}_a \in \Re^N$ as the maximum-variance column of $\mathbf{Y}_a$.
2. Calculate the vector of $\bar{\mathbf{\Phi}}_a$-scores $\mathbf{t}_a \in \Re^N$ as:
   $$\mathbf{t}_a = \bar{\mathbf{K}}_a\mathbf{u}_a, \quad \mathbf{t}_a \leftarrow \mathbf{t}_a/\|\mathbf{t}_a\|$$
3. Regress columns of $\mathbf{Y}_a$ on $\mathbf{t}_a$: $\mathbf{c}_a = \mathbf{Y}_a'\mathbf{t}_a$, where $\mathbf{c}_a$ is a weight vector.
4. Calculate the new score vector $\mathbf{u}_a$ for $\mathbf{Y}_a$ as:
   $$\mathbf{u}_a = \mathbf{Y}_a\mathbf{c}_a, \quad \mathbf{u}_a \leftarrow \mathbf{u}_a/\|\mathbf{u}_a\|$$
5. Repeat steps 2-3 until convergence of $\mathbf{t}_a$.
6. Deflate the matrices:
   $$\bar{\mathbf{K}}_{a+1} \leftarrow (\mathbf{I} - \mathbf{t}_a\mathbf{t}_a')\bar{\mathbf{K}}_a(\mathbf{I} - \mathbf{t}_a\mathbf{t}_a'),$$
   $$\mathbf{Y}_{a+1} \leftarrow \mathbf{Y}_a - \mathbf{t}_a\mathbf{t}_a'\mathbf{Y}_a,$$
7. Set $a=a+1$ and return to step 1. Stop when $a>A$, where A is the number of latent variables selected in the high-dimensional space, $\Re^c$.

Notice that the selection of A is determined by supervising the deflation of $\mathbf{Y}_a$. Again, as in linear PLS, the prediction on training data has the following from (Rosipal and Trejo, 2001):

$$\hat{\mathbf{Y}} = \bar{\mathbf{\Phi}}\mathbf{B}_{PLS} = \bar{\mathbf{\Phi}}\bar{\mathbf{\Phi}}'\mathbf{U}(\mathbf{T}'\bar{\mathbf{K}}\mathbf{U})^{-1}\mathbf{T}'\mathbf{Y}$$
$$= \bar{\mathbf{K}}\mathbf{U}(\mathbf{T}'\bar{\mathbf{K}}\mathbf{U})^{-1}\mathbf{T}'\mathbf{Y} = \mathbf{T}\mathbf{T}'\mathbf{Y} = \mathbf{T}\mathbf{C}', \qquad (4)$$

where the matrices $\mathbf{T} = [\mathbf{t}_1 ... \mathbf{t}_A]$, $\mathbf{U} = [\mathbf{u}_1 ... \mathbf{u}_A]$ and $\mathbf{C} = [\mathbf{c}_1 ... \mathbf{c}_A]$ are orthogonal by columns. Notice that the regression coefficient matrix $\mathbf{B}_{PLS}$ exists but is never computed by the KPLS algorithm, since the kernel substitution avoids the necessity of an explicit computation.

Equality (4) shows that the output response can be obtained from the inner products of the mapped predictor vectors. So, for a new observation $\mathbf{x}$ of the predictor vector, the response vector will be given by:

$$\hat{\mathbf{y}}' = \bar{\varphi}(\mathbf{x})'\mathbf{B}_{PLS} = \bar{\mathbf{k}}'\mathbf{U}(\mathbf{T}'\bar{\mathbf{K}}\mathbf{U})^{-1}\mathbf{T}'\mathbf{Y}$$
$$= \bar{\mathbf{k}}'\mathbf{R}\mathbf{C}' = \bar{\mathbf{k}}'\mathbf{M}, \qquad (5)$$

where $\bar{\mathbf{k}}$ is the vector of centered kernel functions evaluated in the pairs $(\mathbf{x}, \mathbf{x}_j)$ with $j=1,\dots,N$. From Eq. (3), this vector is given by:

$$\bar{\mathbf{k}} = [\bar{k}_1(\mathbf{x}) \cdots \bar{k}_N(\mathbf{x})]' = \mathbf{k} - \mathbf{Ke} - \mathbf{Ek} + \mathbf{EKe}, \qquad (6)$$

where $\mathbf{e}$ is an unitary vector and $\mathbf{k} = [k_1(\mathbf{x}) \cdots k_N(\mathbf{x})]' \in \Re^{N \times N}$ is the kernel functions vector, where the element j is given by the kernel function evaluated at $(\mathbf{x}, \mathbf{x}_j)$, i.e. $k_j(\mathbf{x}) = k(\mathbf{x}_j, \mathbf{x})$. Each of the elements of vector $\bar{\mathbf{k}}$ is computed as follows:

$$\bar{k}_j(\mathbf{x}) = \bar{k}(\mathbf{x}_j, \mathbf{x}) = k(\mathbf{x}, \mathbf{x}) - \frac{2}{N}\sum_{j=1}^{N} k(\mathbf{x}_j, \mathbf{x})$$
$$+ \frac{1}{N^2}\sum_{j=1}^{N}\sum_{n=1}^{N} k(\mathbf{x}_j, \mathbf{x}_n) \qquad (7)$$

Now, we focus our attention on the latent vectors $\mathbf{t} \in \Re^A$. From Eq. (4), it follows that for a new observation $\mathbf{x}$, this vector will be given by:

$$\mathbf{t}' = [t_1 \cdots t_A]' = \bar{\mathbf{k}}'\mathbf{R}, \qquad (8)$$

where $t_a = \bar{\mathbf{k}}'\mathbf{r}_a$ $a=1,\dots.A$, with $\mathbf{R} = [\mathbf{r}_1 \cdots \mathbf{r}_A]$.

Therefore, for a new observation $\mathbf{x}$, the prediction can be computed from $\mathbf{t}$, as (Eq. (5)):

$$\hat{\mathbf{y}}' = \mathbf{t}'\mathbf{C}'. \qquad (9)$$

Now, given a new measurement predictor vector $\mathbf{x}$ in original units, the response vector (also in original units) is predicted by:

$$\hat{\mathbf{y}} = \mathbf{D}_y\mathbf{M}'\bar{\mathbf{k}}\left(\mathbf{D}_x^{-1}(\mathbf{x} - \bar{\mathbf{x}})\right) + \bar{\mathbf{y}}, \qquad (10)$$

where the sample standard deviations $\mathbf{D}_x = diag(\hat{\sigma}_{x_1} ... \hat{\sigma}_{x_m})$, $\mathbf{D}_y = diag(\hat{\sigma}_{y_1} ... \hat{\sigma}_{y_p})$, and the means $\bar{\mathbf{x}}$, $\bar{\mathbf{y}}$, are determined in the training procedure.

3. KPLS-BASED MODELING FOR PROCESS MONITORING

The Kernel PLS algorithm induces an external model which decomposes $\bar{\mathbf{\Phi}}$ and $\mathbf{Y}$ in score vectors $\mathbf{t}_a$, weight vectors ($\mathbf{p}_a$ and $\mathbf{c}_a$), and residual error matrices ($\widetilde{\bar{\mathbf{\Phi}}}$ and $\widetilde{\mathbf{Y}}$), as follows:

$$\bar{\mathbf{\Phi}} = \mathbf{T}\mathbf{P}' + \widetilde{\bar{\mathbf{\Phi}}}, \qquad (11)$$
$$\mathbf{Y} = \mathbf{T}\mathbf{C}' + \widetilde{\mathbf{Y}}, \qquad (12)$$

In agreement with the standard linear PLS model (Godoy et al, 2011) it is assumed that the score variables t_a are good predictor of $\mathbf{Y}$. Furthermore, it is also assumed a linear inner relation between the scores of $\mathbf{t}_a$ and $\mathbf{u}_a$, that is,

$$\mathbf{U} = \mathbf{T}\mathbf{B} + \mathbf{H}, \qquad (13)$$

where $\mathbf{B}$ is the $A \times A$ diagonal matrix and $\mathbf{H}$ denotes the matrix of residuals. Equivalently, we have $\mathbf{UB}^{-1} - \mathbf{HB}^{-1} = \mathbf{T}$, where $\mathbf{C}' = \mathbf{BQ}'$ with $\mathbf{Q}$ orthonormal by columns. Then, the following regression kernel-based model is obtained:

$$\mathbf{Y} = \bar{\mathbf{K}}\mathbf{RC}' + \mathbf{HB}^{-1}\mathbf{C}' + \tilde{\mathbf{Y}}_2 = \hat{\mathbf{Y}} + \tilde{\mathbf{Y}}_1 + \tilde{\mathbf{Y}}_2,$$
$$\mathbf{Y} = \mathbf{TC}' + \tilde{\mathbf{Y}}_1 + \tilde{\mathbf{Y}}_2 \quad (14)$$

where $\tilde{\mathbf{Y}} = \tilde{\mathbf{Y}}_1 + \tilde{\mathbf{Y}}_2$. We call $\mathbf{V}$ the pseudo-inverse of $\mathbf{P}'$ ($\mathbf{V}'\mathbf{P} = \mathbf{P}'\mathbf{V} = \mathbf{I}$); then, the prediction of $\mathbf{T}$ is directly obtained from Eq. (11), as:

$$\mathbf{T} = \bar{\mathbf{\Phi}}\mathbf{V}, \quad (15)$$

because the row space of $\tilde{\bar{\mathbf{\Phi}}}$ belongs to the null-space of the linear transformation $\mathbf{P}'$, i.e. $\tilde{\bar{\mathbf{\Phi}}}\mathbf{V} = \mathbf{0}$ (Godoy et al, 2011). If $\mathbf{P}' = \mathbf{W}_A \mathbf{\Sigma}_A \mathbf{V}_A'$ is the compact singular value decomposition (SVD) of $\mathbf{P}'$, then $\mathbf{V} \triangleq (\mathbf{P}')^{-} = \mathbf{V}_A \mathbf{\Sigma}_A^{-1} \mathbf{W}_A'$ where $^{-1}$ denote a generalized inverse (Meyer, 2000). Equivalently, from Eq. (14), $\mathbf{D}$ is the pseudo-inverse of $\mathbf{C}'$ ($\mathbf{C}'\mathbf{D} = \mathbf{D}'\mathbf{C} = \mathbf{I}$) and, since $\tilde{\mathbf{Y}}_2 \mathbf{D} = \mathbf{0}$, then: $\mathbf{T} + \mathbf{HB}^{-1} = \mathbf{YD}$.

An interesting point of the propose procedure is that the matrices $\mathbf{V}$, $\mathbf{P}$, $\mathbf{B}$ and $\mathbf{Q}$ will never be estimated (otherwise it would be impractical). They are only defined to develop the proof that follows, in order to find metrics based on kernel substitution trick.

3.1. Underlying KPLS decomposition of the input mapped and output spaces

After synthesizing an *in-control* KPLS model, the measurement vectors $\bar{\mathbf{\phi}}(\mathbf{x}) \in \mathfrak{R}^c$ and $\mathbf{y} \in \mathfrak{R}^p$ are (underlying) decomposed as described below.

First the latent vector is computes as $\bar{\mathbf{k}}'\mathbf{R} = \bar{\mathbf{\phi}}(\mathbf{x})'\mathbf{V} = \mathbf{t}'$ (see Eq. 15), where $\mathbf{V} = [\mathbf{v}_1 \cdots \mathbf{v}_A]$ is the matrix of PLS components, given by $\mathbf{v}_a = \sum_{j=1}^{N} \alpha_j \bar{\mathbf{\phi}}(\mathbf{x}_j)$ with $\alpha_j \in \mathfrak{R}$. Then, a new mapped vector $\bar{\mathbf{\phi}}(\mathbf{x})$ (associated to the measurements $\mathbf{x}$) can be decomposed as:

$$\bar{\mathbf{\phi}}(\mathbf{x}) = \hat{\bar{\mathbf{\phi}}}(\mathbf{x}) + \tilde{\bar{\mathbf{\phi}}}(\mathbf{x}) \quad \in \mathfrak{R}^c$$
$$\hat{\bar{\mathbf{\phi}}}(\mathbf{x}) = \mathbf{PV}'\bar{\mathbf{\phi}}(\mathbf{x}) = \mathbf{Pt} \quad \in W_M \subseteq \mathfrak{R}^c \quad (16)$$
$$\tilde{\bar{\mathbf{\phi}}}(\mathbf{x}) = (\mathbf{I} - \mathbf{PV}')\bar{\mathbf{\phi}}(\mathbf{x}) \in W_R \subseteq \mathfrak{R}^c$$

where $\mathbf{P} = \bar{\mathbf{\Phi}}'\mathbf{T} = \bar{\mathbf{\Phi}}'\bar{\mathbf{K}}\mathbf{R}$, with $\mathbf{P}$ and $\mathbf{V}$ orthogonal by columns. The oblique projections that appear in Eq. (16), decompose the high-dimensional space, $\mathfrak{R}^c$, into two complementary subspaces W_M and W_R, i.e. $W_M \oplus W_R \equiv \mathfrak{R}^c$ (Godoy et al, 2011). Matrix $\mathbf{PV}'$ is the projector onto the model subspace $W_M \equiv Span\{\mathbf{P}\} \subseteq \mathfrak{R}^c$, along the residual subspace $W_R \equiv Span\{\mathbf{V}\}^{\perp} \subseteq \mathfrak{R}^c$ direction (Godoy et al, 2011), where $^{\perp}$ denotes the orthogonal complement of the subspace (notice that range of $\mathbf{P}$ is W_M and the null-space is W_R). Hence, mapped predictor space is decomposed in an underlying form, by KPLS, into complementary oblique subspaces.

In addition, if we generalize the results of Godoy et al (2011) for the case KPLS, response space is also decomposed (by KPLS) in complementary oblique subspaces, which is related with the predictor modeled subspace according to:

$$\mathbf{y} = \hat{\mathbf{y}}^* + \tilde{\mathbf{y}}_2 \ \in \mathbb{R}^p,$$
$$\hat{\mathbf{y}}^* = \mathbf{CD}'\mathbf{y} \ \in S_{MY} \equiv Span\{\mathbf{C}\},$$
$$\tilde{\mathbf{y}}_2 = (\mathbf{I} - \mathbf{CD}')\mathbf{y} \in S_{RY} \equiv Span\{\mathbf{D}\}^{\perp} \quad (17)$$
$$\hat{\mathbf{y}}^* = \hat{\mathbf{y}} + \tilde{\mathbf{y}}_1,$$
$$\hat{\mathbf{y}} = \mathbf{CR}'\bar{\mathbf{k}}(\mathbf{x}) \ \in S_{MY},$$
$$\tilde{\mathbf{y}}_1 = \hat{\mathbf{y}}^* - \hat{\mathbf{y}} = \mathbf{CD}'\mathbf{y} - \mathbf{CR}'\bar{\mathbf{k}}(\mathbf{x}) \ \in S_{MY} \quad (18)$$

Figure 1 illustrates part of underlying space decomposition developed in this work. Each mapped measurement vector is decomposed and their projections are compared with their control limits.

4. PROCESS MONITORING BASED ON KPLS
4.1. Fault detection indexes

The multivariate process monitoring strategy uses statistical indexes associated to different subspaces for fault detection purposes. Based on the *in-control* KPLS model it can be analyzed every future behavior of the process by mapping the new observations $\mathbf{x}$ to the modeled subspace and to the (complementary) residual subspace, $\mathbf{x} \mapsto \hat{\bar{\mathbf{\phi}}}(\mathbf{x}) + \tilde{\bar{\mathbf{\phi}}}(\mathbf{x})$. In each of these subspaces it is possible to measure distances, independently. However, since no explicit formulas are available for the projections of Eq. (16), it must be found new statistics using the kernel substitution to obtain (estimated) norms for the projected vectors.
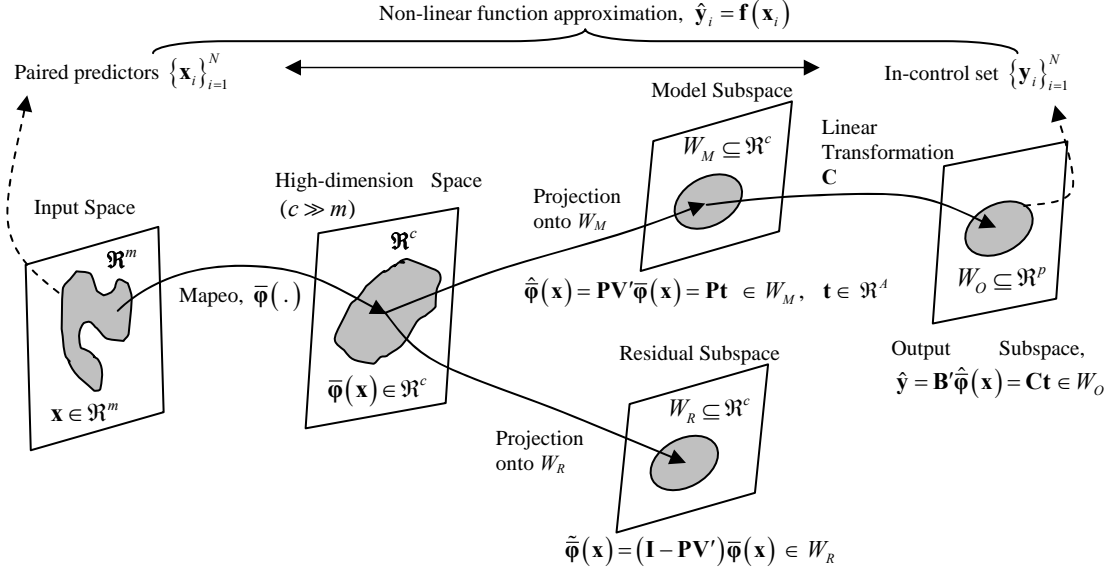
Figure 1: KPLS intrinsic architecture showing the input underlying decomposition with their relations and control limits.

For instance, to detect a significant change in W_M, the following Hotelling's T^2 statistic for $\mathbf{t}$ is defined:

$$T_t^2(\mathbf{x}) = \mathbf{t}'\mathbf{\Lambda}^{-1}\mathbf{t} = (N-1)\overline{\mathbf{k}}'\mathbf{R}\mathbf{R}'\overline{\mathbf{k}}, \qquad (19)$$

where $\mathbf{\Lambda} = (N-1)^{-1}\mathbf{T}'\mathbf{T} = (N-1)^{-1}\mathbf{I}$.

When a new special event (originally not considered by the in-control KPLS model) occurs, the new mapped observation $\overline{\mathbf{\phi}}(\mathbf{x})$ will move out from W_M, into W_R. The squared prediction error of $\overline{\mathbf{\phi}}(\mathbf{x})$ (SPE_X), or distance to the $\overline{\mathbf{\phi}}(\mathbf{x})$-model, is defined as:

$$
\begin{aligned}
SPE_X(\mathbf{x}) &= \left\|\tilde{\overline{\mathbf{\phi}}}(\mathbf{x})\right\|^2 = \left\|\overline{\mathbf{\phi}}(\mathbf{x}) - \mathbf{P}\mathbf{V}'\overline{\mathbf{\phi}}(\mathbf{x})\right\|^2 \\
&= \overline{\mathbf{\phi}}(\mathbf{x})'\overline{\mathbf{\phi}}(\mathbf{x}) - 2\overline{\mathbf{\phi}}(\mathbf{x})'\hat{\overline{\mathbf{\phi}}}(\mathbf{x}) + \hat{\overline{\mathbf{\phi}}}(\mathbf{x})'\hat{\overline{\mathbf{\phi}}}(\mathbf{x}) \\
&= \overline{k}(\mathbf{x},\mathbf{x}) - 2\overline{\mathbf{k}}'\overline{\mathbf{K}}\mathbf{R}\mathbf{t} + \mathbf{t}'\mathbf{t}
\end{aligned}
$$
$$(20)$$

where $\overline{\mathbf{\phi}}(\mathbf{x})'\hat{\overline{\mathbf{\phi}}}(\mathbf{x}) = \overline{\mathbf{\phi}}(\mathbf{x})'\mathbf{P}\mathbf{t} = \overline{\mathbf{\phi}}(\mathbf{x})'\overline{\mathbf{\Phi}}'\overline{\mathbf{K}}\mathbf{R}\mathbf{t} = \overline{\mathbf{k}}'\overline{\mathbf{K}}\mathbf{R}\mathbf{t}$.
Then, SPE_X can be used for detecting a change in W_R. When the process is in-control, the SPE_X index represents the fluctuations that can not be explained by the KPLS model. The distance from the regression model in S_{MY} is defined as:

$$SPE_{Y1} = \left\|\tilde{\mathbf{y}}_1\right\|^2 = \left\|\begin{bmatrix}\mathbf{C}\mathbf{D}' & -\mathbf{C}\mathbf{R}'\end{bmatrix}\begin{bmatrix}\mathbf{y} \\ \overline{\mathbf{k}}(\mathbf{x})\end{bmatrix}\right\|^2 \qquad (21)$$

and the distance from the $\mathbf{y}$-model in S_{RY} is defined as:

$$SPE_{Y2} = \left\|\tilde{\mathbf{y}}_2\right\|^2 = \left\|(\mathbf{I}-\mathbf{C}\mathbf{D}')\mathbf{y}\right\|^2 \qquad (22)$$

Frequently, $\mathbf{R}_{\hat{\overline{\phi}}}$ and $\mathbf{R}_{\hat{\mathbf{y}}}$ are singular. Then, the generalized Mahalanobis' distance for $\hat{\overline{\mathbf{\phi}}}$ and $\hat{\mathbf{y}}$ are:

$$D_{\hat{\overline{\phi}}} = \hat{\overline{\mathbf{\phi}}}'\mathbf{R}_{\hat{\overline{\phi}}}^-\hat{\overline{\mathbf{\phi}}}, \qquad (23)$$

$$D_{\hat{\mathbf{y}}} = \hat{\mathbf{y}}'\mathbf{R}_{\hat{\mathbf{y}}}^-\hat{\mathbf{y}}. \qquad (24)$$

where the correlation matrices are given by:

$$
\begin{aligned}
\mathbf{R}_{\hat{\mathbf{y}}} &= (N-1)^{-1}\hat{\mathbf{Y}}'\hat{\mathbf{Y}} \\
&= \mathbf{C}\left[(N-1)^{-1}\mathbf{T}'\mathbf{T}\right]\mathbf{C}' = (N-1)^{-1}\mathbf{C}\mathbf{C}'
\end{aligned}
$$
$$(25)$$

$$
\begin{aligned}
\mathbf{R}_{\hat{\overline{\phi}}} &= (N-1)^{-1}\hat{\overline{\mathbf{\Phi}}}'\hat{\overline{\mathbf{\Phi}}} \\
&= (N-1)^{-1}\mathbf{P}\mathbf{T}'\mathbf{T}\mathbf{P}' = (N-1)^{-1}\mathbf{P}\mathbf{P}'
\end{aligned}
$$
$$(26)$$

Since $\mathbf{C}$ is orthogonal by columns, the property of the generalized inverse of a SVD yields: $\mathbf{R}_{\hat{\mathbf{y}}}^- = \left(\mathbf{C}\left[(N-1)^{-1}\mathbf{I}\right]\mathbf{C}'\right)^- = (N-1)\mathbf{C}\mathbf{C}'$. Then, replacing this and Eq. (9) into the Eq. (24), result the following equality:

$$D_{\hat{\mathbf{y}}} = \hat{\mathbf{y}}'\mathbf{R}_{\hat{\mathbf{y}}}^-\hat{\mathbf{y}} = \mathbf{t}'\mathbf{C}'((N-1)\mathbf{C}\mathbf{C}')\mathbf{C}\mathbf{t} = (N-1)\mathbf{t}'\mathbf{t} = T_t^2. \quad (27)$$

Similarly, $\mathbf{R}_{\hat{\overline{\phi}}}^- = (N-1)\mathbf{P}\mathbf{P}'$ (see Eq. 26). Then, result:

$$D_{\hat{\overline{\phi}}} = \hat{\overline{\mathbf{\phi}}}'\left[(N-1)\mathbf{P}\mathbf{P}'\right]\hat{\overline{\mathbf{\phi}}} = (N-1)\mathbf{t}'\mathbf{t} = T_t^2 = D_{\hat{\mathbf{y}}}. \quad (28)$$

and consequently, the metrics on $\hat{\overline{\mathbf{\phi}}}(\mathbf{x})$, $\mathbf{t}$ or $\hat{\mathbf{y}}$ are equivalents.

Hence, the process output (quality variables) can be monitored through a PLSR-based input statistic. Then, we propose monitoring using four non-overlapped metrics (SPE_X, T_t^2, SPE_{Y1}, and SPE_{Y2}) which completely

cover both measurement spaces, each one on a different subspace.

4.2. Fault diagnosis by means of alarmed subspaces

An anomaly is a change in the measurements following or not the correlation structure captured by the PLSR model. If the change produces an out-of-control point, the anomaly source can be classified according to a) an excessively large operation change of the normal operation; b) a significant increase of variability; c) the alteration of cross-correlations, and d) sensor faults. Cases a) and b) involve changes in the measurement vector following the modeled correlation structure; while cases c) and d) involve changes in some variables altering the correlation pattern with the others. In fact, an abnormal process behavior involves a deviation of the modeled correlations, thus increasing the value of the proper metric being used. In order to classify the abnormalities, we analyze the effect on each subspace (see Table 1). Rows 1), 2), 3) and 6) feature complex process changes; while rows 4) and 5) represent localized sensor faults. By analyzing the contributions to an alarmed index (Alcala and Qin, 2011), such as SPE_X (or SPE_Y), it could be possible to discriminate changes in the **X**- / **Y**-outer part against sensor fault in **x** / **y** (see this ambiguity in Table 1).

In summary, the proposed monitoring strategy is based on an input and output space PLS decomposition, which classifies the type of process fault or anomaly according to the statistic that triggers the alarm condition.

Table 1. Fault diagnosis based on alarmed index.

Fault/Anomaly in	SPE_X	T_t^2	SPE_{Y1}	SPE_{Y2}
1- Inner part, $d\mathbf{B}$	–	–	×	–
2- **X** outer part, $d\mathbf{P}$	×	–	o	–
3- **Y** outer part, $d\mathbf{Q}$	–	–	o	×
4- **x** sensor	×	–	o	–
5- **y** sensor	–	–	o	×
6- latent space upset	–	×	–	–

×: high value. –: negligible value. o: high/low value.

5. APPLICATION STUDY
5.1. A Non-linear Numerical Simulation Example

A simulated case study is used to evaluate the performance of the proposed monitoring technique for a variety of fault scenarios. To understand the implicit decompositions and statistics as monitoring tools, we simulated different faults in a synthetic system. We use the nonlinear multidimensional simulation example devised in Refs. Zhang et al (2008) and Zhao et al (2006). It is defined as follows:

$$x_1 = t^2 - t + 1 + \varepsilon_1,$$
$$x_2 = \sin(t) + \varepsilon_2,$$
$$x_3 = t^3 + t + \varepsilon_3,$$
$$y = x_1^2 + x_1 x_2 + 3\cos(x_3) + \varepsilon_4$$

where t is uniformly distributed over [-1, 1] and ε_i, i=1,2,3,4 are noise components uniformly distributed over [-0.1, 0.1]. The generated data of 300 samples are segmented into training and testing data sets. They are illustrated in the Fig. 2. The first 200 samples are selected for training, and the subsequent 100 samples are used as testing data set. It is apparent that the input variables are driven by one latent variable t only in this case. From Fig. 2, we can easily see that the response variable is nonlinearly correlated with the input variables. In this work, we used the radial basis kernel,

$k(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\|\mathbf{x}_i - \mathbf{x}_j\|^2 / h\right)$, in our implementation.

When using this kernel function the value of parameter $h=2\sigma^2=0.06$ (σ=0.1732) has a significant influence on the KPLS prediction performance (Zhang et al, 2008).

The mean squared error (MSE) is used to evaluate the estimator. Figure 3a show the $\% RMSE = 100\sqrt{MSE}$ (standard error) for each a component using training data or testing data (or external validation data). Notice that $MSE(a)$ refers to MSE when the first a latent variables are used. The selection of an adequate number A of latent variables to be included in the KPLS model is crucial; if more than necessary variables are used, an undesirable over-fitting might reduce the predictive ability. We use the Wold's $R_{0.9}$ rule, which can be written as $R(a+1) = MSE(a+1))/MSE(a)$. For successive a values, the sequence is stopped when $R(a+1)>0.9$; and hence $A=a$. Conceptually, this criterion states that an additional latent variable will not be included in the KPLS model unless it provides a meaningful prediction improvement, and consequently it gives the maximum number on components to be included in the model. Figure 3b show the above indicator vs. the number of components for training data and testing data. The ratio for training data is larger than 0.9 in $a+1$=12, hence the most parsimonious model corresponds to A=11 (Zhang et al 2008). The same ratio with testing data is for monitor overtraining when is selecting a reliable A value. Figures 4 show the prediction results of the training and testing data using the Eq. 10. The upper part of Figure 4 shows the actual and predicted values, and the lower part shows the errors between both values.

Six faults were simulated, which are described in the Table 3. Table 1 shows the diagnosis expected in each simulated fault/anomaly. Figure 5 shows the time evolution of each statistic normalized by its control limit, which clearly exhibit the six simulated abnormalities, where it is possible to detect and classify each type of simulated fault, on the basis of the information given in Table 1. Since $y \in \Re^1$ (is not multivariable), then $SPE_{Y2} \equiv 0$ and $SPE_{Y1} = \|y - \hat{y}\|^2$. Hence, the y-sensor fault will occur in SPE_{Y1} (see Table 1). The simulation results show that the developed strategy is able to identify abnormalities attributed to sensor faults, process changes, process upsets and disturbances. The model supporting this monitoring approach is based on normal operating data. Process

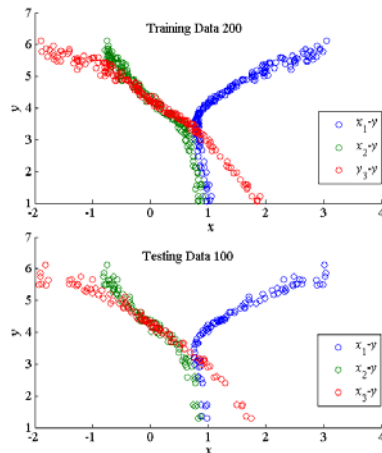data recollection constitutes a critical step when developing empirical models for monitoring.

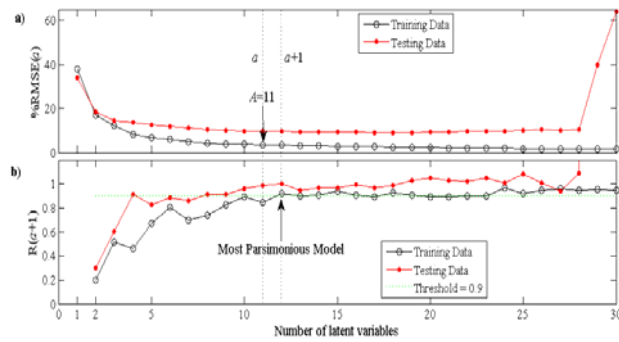Figure 2. Relationship between inputs and the response.

Figure 3. KPLS Model order determination. a) Residual error variance vs. number of latent variables. b) Ratio of residual error variances successive.
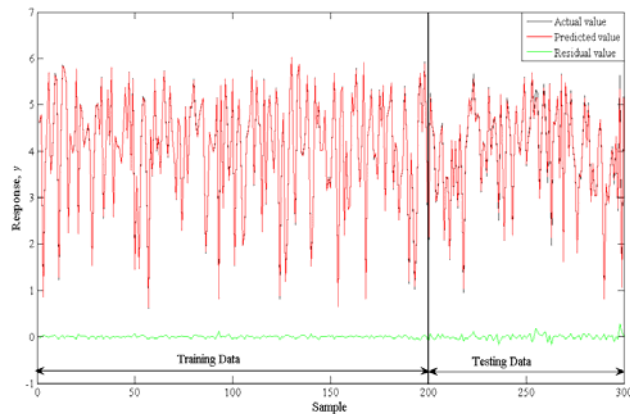

Figure 4. Prediction results via KPLS method.

Table 3: Simulated fault scenarios.

Type	Samples	Magnitude	Diagnosis
4	20..30	$dx_1 = -1.5$	2/4
4	50..60	$dx_2 = 1.5$	2/4
4	80..90	$dx_3 = 1.5$	2/4
2	110..120	$x_1 = 3t^3 - t + 1 + \varepsilon_1$	2/4
6	140..150	$t = 1.01$ (fixed)	6
5	170..180	$dy = 1.5$	5
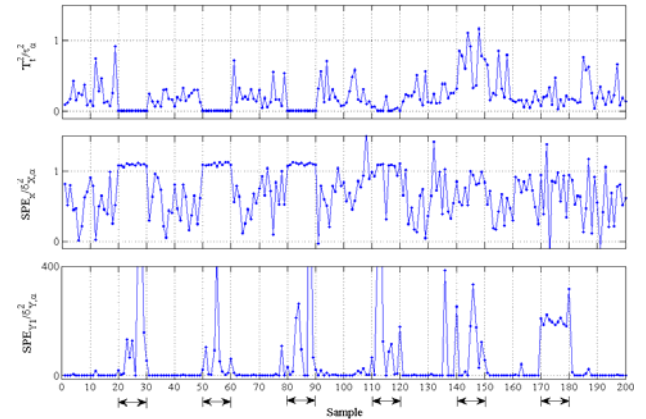
Figure 5: Time evolution of the combined index and each normalized KPLS-statistic.

6. CONCLUSIONS AND FUTURE WORKS

Many multivariate process monitoring systems could be based on a non-linear KPLS model that represents 'in-control' conditions. As in more traditional, a meaningful deviation of the variables from their expected trajectories serves for the detection and diagnosis of abnormal process behaviors. The results of a non-linear simulation example illustrates that the proposed strategy is efficient and accurate.

However, these results are preliminary, more realistic applications are necessary for a reliable validating of the method and to learn more about the proposed non-linear strategy.

REFERENCES
Godoy J.L., Vega J.R., Marchetti J.L., 2011. Geometric Properties of Partial Least Squares Regression Applied to Process Monitoring. *III MACI 2011*. Bahia Blanca (Argentina), 9 to 11 may.

Meyer D., 2000. *Matrix analysis and applied linear algebra*. SIAM, USA.

Zhao, S. J.; Zhang, J.; Xu, Y. M.; Xiong, Z. H., 2006. Nonlinear Projection to Latent Structures Method and its Applications. *Ind. Eng. Chem. Res.* 45, 3843-3852.

Zhang X., Yan W., and Shao H., 2008. Nonlinear multivariate quality estimation and prediction based on kernel partial least squares. *Ind. Eng. Chem. Res.*, 47 (4), 1120–1131.

Rosipal R. and Trejo L.J., 2001. Kernel Partial Least Squares Regression in Reproduciing Kernel Hilbert Space. J. of Machine Learning Research, 2, 97-123.

Diodes and an alternative Procedure to derive equations from Bond Graph models by exploting a Linear Graph approach.

Noé Villa-Villaseñor, Gilberto González-Ávalos and Jesús Rico-Melgoza
DEP-FIE
Universidad Michoacana
de San Nicolás Hidalgo
Morelia, Michoacán, México
contact Email: vilaslord@hotmail.com

Keywords- Algebraic equations, State Space Equations, Bond Graph Modelling, Linear Graph Modelling, Tree Theory, Causal Paths, Shockley Model of the diode.

Abstract— **This paper deal with methods to get equations from a Bond Graph (BG) Model. Mixed Techniques to get equations from both Linear Graph (LG) and BG Models are employed. A criterion that allows to identify the tree and the co-tree in a BG model is reviewed. Also, the advantages of causal paths inherent to BG are exploited. The equivalence between a LG model and its respective BG model is also used. A procedure to get equations from a given BG model is proposed. To show the functionality of the proposed methods, examples with diodes are solved.**

I. INTRODUCTION

Graphs are simple geometrical figures consisting of nodes and lines that connect some of these nodes; they are sometimes called "Linear Graphs" [1]. One of the most important applications of Graph Theory is its use in the formulation and solution of the electrical network problem by Kirchhoff [1]. Equations can be determined from a LG by finding a normal form based on an appropriate tree [1].

Bond graph is a highly structured modelling technique that allows to analyze different kinds of physical systems in an unique way determined by a basis of unified description [2]. This technique allows the determination of the State Space Equation (SSE) in several different ways by reviewing the interactions of the dynamic elements and the causal paths they determine [3], [4].

Both BG and LG treats with spatially discrete physical systems [5] even when only LG modelling preserves the spatial visualization related to the model being analyzed. The spatial constraints in a BG model are implicit in the location of the 0 and 1 junctions [5].

There exists several ways to get the SSE from a BG model [6], some of them exploiting the concept of causality and causal paths.

Mathematical manipulation of non linear semiconductor models could be difficult. Here, diodes are reviewed and modeled with a non linear constitutive relation.

In this paper a procedure to obtain equations from a BG model by exploiting the LG tree Theory is proposed. The proposed procedure can be particularly useful in the analysis of some systems involving algebraic loops.

The contents are ordered as follows:

Section II shows a reviewing of the Shockley diode model . Section III presents some current techniques to derive equations from BG and LG models. Section IV shows the main result of this work, First a criterion that allow to identify the LG related tree in a BG model is discussed.Then, a procedure that allow to derive the SSE from a BG model by employing the LG related tree theory is presented. Finally three examples are analyzed and discussed while they are solved with the proposed approach. In section V, the conclusions and final comments are written.

In the next Section a short review of the Shocley diode model is presented.

II. THE DIODE MODEL

The diode can be considered as a non controlled switching device [7]. Switched systems can be viewed from several approaches and two main classifications can be made: The variant topology approach and the invariant topology approach [6].

Most texts deal with the idealized model of the diode, i.e. a variant topology model whose representation changes between a short circuit and an open circuit, depending on its on-off state.

The Shockley model is a non linear invariant topology model that is presented in several power electronics textbooks. However, a procedure to derive equations from systems involving this Shockley model is rarely presented and authors start their analysis with the idealized on-off-switch approach. In this paper the Shockley model is used.

The constitutive relation of the Shockley model as appears in [8] is

$$I_D = I_S \left(e^{\frac{V_D}{n V_T}} - 1 \right)$$

where:

- I_D = Current through the diode.

273

- V_D = Voltage in the diode in the anode referred to the cathode.
- I_S = Reverse saturation current with a value of $1 \times 10^{-12}A$.
- n = Empirical constant known as emission coefficient, that depends on the construction of the element.
- V_T = Thermal voltage given by

$$V_T = \frac{kT}{q}$$

Where q is the electron charge equal to $1.6022 \times 10^{-19}C$; T is the absolute temperature in kelvins and k is the Boltzmann constant equal to $1.3806 \times 10 - 23 J/K$.

For simplicity, some values will be fixed and it will be assumed that

$$I_D = I_S \left(e^{gV_D} - 1 \right) \quad (1)$$

considering a $27°C$ temperature and $n = 2$ that is a value that can be assigned to silicon diodes. With these considerations the g value is

$$g = \frac{1}{\eta v_T} \simeq 19.2308$$

The characteristic $v - i$ curve for this constitutive relation is shown in Figure 1.
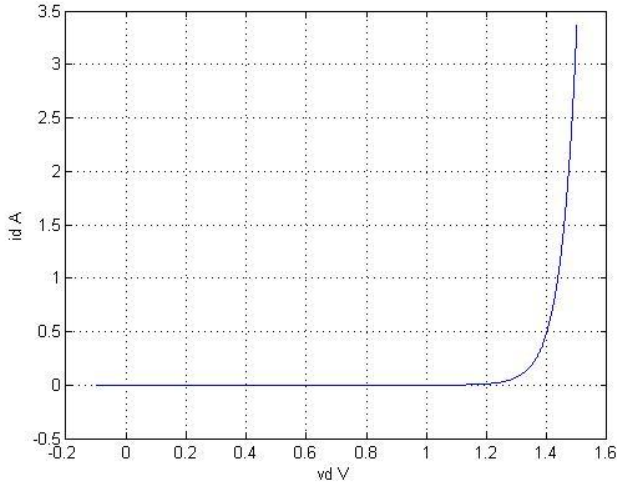
Fig. 1.   The caracteristic $v - i$ curve resulting of the constitutive relation (1).

The BG model that will be used in this paper is a resistive non linear R element whose constitutive relation is that presented in equation (1).This model appears in figure 2.

Fig. 2.   A resistive non linear diode model.

In the next section a reviewing of some of the current techniques to get the equations from both BG and LG models is performed.

III. CURRENT METHODS OF EQUATIONS DERIVATION.

There exist several techniques in order to derive equations from both LG and BG models, and some of them are reviewed in the following paragraphs.

A. SSE derivation from a BG models through junction structure matrix

There exists several approaches to develop the SSE from a BG model [6], one of the most powerful is the related with the junction structure matrix [3]. In Figure 3 The key vectors and the junction structure relationships are shown.
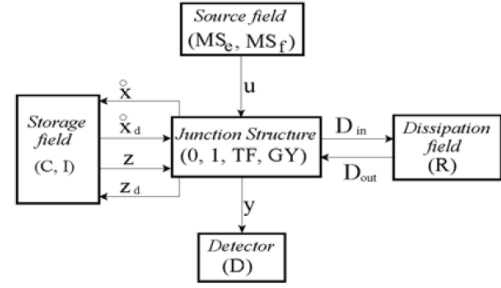
Fig. 3.   Key vectors and the junction structure.

In figure 3, the key vectors are formed as follows: (MS_e, MS_f) is the sources field; (L, C) is the storage field; (R) is the dissipative field; (D_e) is the field of detectors; finally, $(0, 1, TF, GY)$ are the elements of the junction structure. The vectors x and x_d represent the states of the system at integral and derivative causality, respectively. z denotes the co-energy vector and z_d the derivative co-energy vector. The vectors u and y are the input and the output, respectively. D_{in} and D_{out} show the relationships between efforts and flows in the dissipative field. The constitutive relationships are given by,

$$z = Fx \quad (2)$$

$$D_{out} = LD_{in} \quad (3)$$

$$z_d = F_d x_d \quad (4)$$

Also the relationships of the junction structure are specified by,

$$\begin{bmatrix} \dot{x} & D_{in} & y \end{bmatrix}^T = S \begin{bmatrix} z & D_{out} & u & \dot{x_d} \end{bmatrix}^T \quad (5)$$

where the matrix of junction structure S is given by,

$$S = \begin{bmatrix} S_{11} & S_{12} & S_{13} & S_{14} \\ S_{21} & S_{22} & S_{23} & S_{24} \\ S_{31} & S_{32} & S_{33} & S_{34} \end{bmatrix} \quad (6)$$

$$z_d = G_1 - S_{14}^T z$$

By employing the junction structure matrix S and the relationships between the different fields it is possible to write the model as a space-state equation of the form

$$\dot{x} = Ax + Bu \quad (7)$$

$$y = Cx + Du \qquad (8)$$

with

$$A = E^{-1}\left(S_{11} + S_{12}MS_{21}\right)F \qquad (9)$$

$$B = E^{-1}\left(S_{13} + S_{12}MS_{23}\right) \qquad (10)$$

$$C = \left(S_{31} + S_{32}MS_{21}\right)F \qquad (11)$$

$$D = S_{33} + S_{32}MS_{23} \qquad (12)$$

where

$$E = I + S_{14}F_d^{-1}S_{14}^T F \qquad (13)$$

$$M = \left(I - LS_{22}\right)^{-1}L \qquad (14)$$

The dynamic elements being placed in vectors x and x_d are selected depending on its causality assignment.
One disadvantage of this method arises in non linear systems modelling, since this matrix representation is not always possible to obtain.

In the next subsection, other approach employed to get the SSE from a BG model is cited.

B. Other SSE derivation method from Bond Graph

According to [6] and [2], given a BG model, the SSE can be achieved by proceeding in a systematic way, as follows:

1) Write the structure laws in the junctions by considering causality.
2) Write the constitutive relationships of the elements by considering causality.
3) Combine these different laws in order to put on explicitly the derivatives of the state variables as function of the state variables and the inputs.

The key concept in both methods is causality, because the formulation of S matrix (in the first method) depends totally on causal paths as well as the systematic (second) method needs the causality assignment already performed. Once the S matrix has been constructed, the procedure to derive the SSE of the model is just a step-by-step algorithm. In the next subsection a short review on causality and causal paths is done.

C. Causality and causal paths

One of the most powerful characteristics of bond graph is that a lot of information can be retrieved without writing any equations, just by analyzing the causality. Physical systems are full of interacting variable pairs [2]. If two elements are bonded, the effort causes one element to respond with flow, while the flow causes the first element to respond with effort. Thus, the cause-effect relationships for efforts and flows are represented in opposite directions.

A detailed description of causal paths can be found in [9] from where the next has been extracted:
Definition 1: A causal path in bond graph is an alternation of bonds and basic elements, called "nodes" such that:

1) For the acausal graph (before establishing causality), the sequence forms a single chain.

2) All of the nodes in the path have complete and correct causality.
3) Two bonds in a causal path have in the same node opposite causal orientations.

According to the variable being followed, there are two kinds of causal paths. The causal path is simple if it can be crossed by following always the same variable and the causal path is mixed if it is necessary to perform a variable change while the graphic is crossed. In addition, two elements P_1 and P_2, belonging to the set $\{R, C, I, S_e, S_f, D_e, D_f\}$ are *causally connected* if the input variable of one is influenced by the output variable of the other.

In the next section some of the techniques used to develop equations from a LG model are mentioned.

D. SEE derived from a LG model

The tree theory developed originally by Kirchhoff [1] is very useful in the formulation of equations in electrical networks. In purely resistive linear networks, two general approaches can lead to the formulation of a minimal set of equations: general node analysis and general loop analysis [10]. If a given network has dynamic elements, there exists procedures to derive the SSE by using the basis of a normal tree. According to [1] a normal tree can be defined as follows:

Definition 2: A normal tree of a connected directed graph representing a network is a tree that contains all the independent voltage (effort) sources edges, the maximum number (all) of capacitive edges, the minimum number (none) of inductive edges, and none of the independent current (flow) source edges.

The SSE equation of any linear time-invariant system represented by a LG model can be derived by following a systematic precedure [1]. In the case when a normal tree as described in definition 2 cannot be constructed, the LG model includes dynamical elements that cannot be expressed independently in the resulting SSE.

In the next section a mixed approach that exploits the LG tree theory is presented. This proposed approach works directly on the BG model.

IV. Deriving the Equations from a BG model by using a LG tree approach

The equivalence between a LG model and a BG model is reviewed in [12]. Authors establishes a relation between a BG model an the tree of the related LG model by considering causality and adyacent causal strokes. A slightly different cualitative analysis can be performed. Consider the electrical network shown in Figure 4.

In Figure 4 (a) an electric network is drawn. In Figure 4 (b) the corresponding tree (the only possible normal tree) of the system is shown. In Figure 4 (c) the unique related Bond Graph model with integral preferred causality (BGI) can be seen. It is very important to notice the branches of the tree and the orientation of the causal strokes in the corresponding elements on the BG model. The following concepts allow to
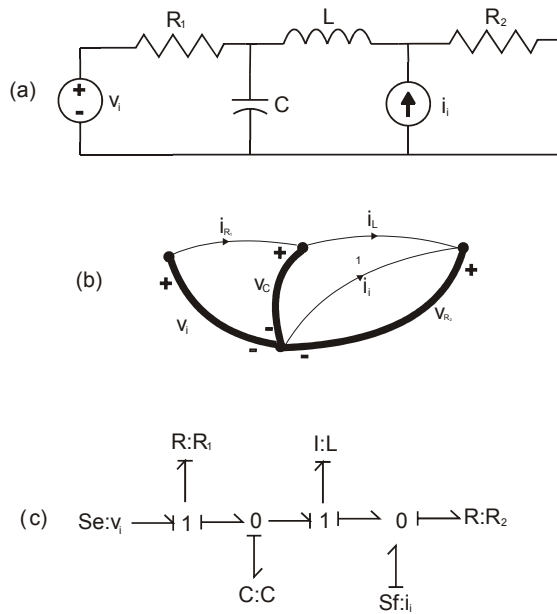
Fig. 4.   Electrical network (a) and their LG (b) and BGI (c) representation.

formalize the criterion employed to identify a tree on a BG model.

When a port-1 element is causally connected to a BG junction structure two situations are possible: (i) The causal stroke (effort) go from the junction structure toward the element or *outwards* as in Figure 5 (a), or (ii) the causal stroke go from the element toward the junction structure or *inwards*, as in Figure 5 (b). In Figure 5, the X element is assumed to be a port-1 element.
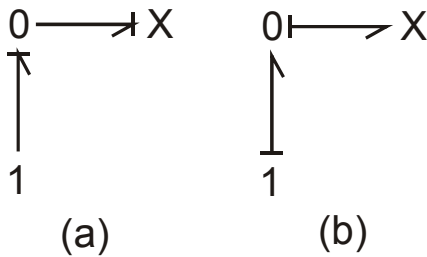
Fig. 5.   The two possible causal orientations of a port-1 element.

This direction on the casual strokes is important, because it bring the conditions to identify the spanning tree of a system in a BG model. This is summarized in the next criterion:

Criterion 1: In a BGI model, each port-1 element whose causal stroke is inwards (junction adjacent) determine a branch of the related spanning tree. The set of all these port-1 elements spans the related tree. In the other hand, each port-1 element whose causal stroke is outwards (element adjacent) determine a chord of the related cotree.

This criterion allow to find out a one to one relation between a BG model and its associated LG model.

In LG theory, Kirchhoff Current Law (KCL) is applied over tree branches in order to get their currents as a sum of currents of the cotree chords. This is completely equivalent to analyze the causal paths related to each "tree branch" on the BG model after Criterion 1 has been applied. With this consideration, a direct procedure to derive the SSE from a BG model can be written as follows:

Procedure 1:

1) Given an acausal BG model, assign causality as BGI.
2) Determine the tree branches and the cotree chords according to criterion 1.
3) Set the effort sources and capacitive elements in the tree as tree-main-effort (TME) variables. If resistive elements belong to the tree, set them as tree-helper-effort (THE) variables.
4) Set the flow sources and inductive elements in the cotree as key-flow (KF) variables. If resistive elements belog to the cotree, set them as nonkey-flow (NF) variables
5) By using the element constitutive relation, write each NF variable as a combination of all the TME variables that hold a direct causal path with it.
6) By using the element constitutive relation, write the each THE variable as a combination of all the flow variables that hold a direct causal path with it.
7) Write the state equations by determining each capacitor flow in the tree as a combination of the cotree flows that hold a direct causal path with it. And by determining each inductive effort in the cotree as a combination of the tree efforts that hold a direct causal path with it. Finally, write the state equation by using (in the linear case) the relations $f_C = Ce_C$ and $e_I = If_I$

Remark 1: If in step 1, Integral causality cannot be assigned in all elements, it implies that a normal tree cannot be constructed. In step 5 and 6 if it is found that direct causal paths are held between THE and NF elements, then algebraic loops will arise in the case of linear systems or differential algebraic equations must be constructed in the case of non linear systems.

The preceding Procedure allows to derive the SSE from a BG model. If only algebraic equations are needed (resistive systems), the general node analisys can be performed. General node analysis mixed with BG analysis brings a very systematic procedure to derive equations. Once a normal tree is established, it is always possible to find causal paths between tree elements and cotree elements. The key of the procedure is to determine the flows in the tree elements in terms of cotree variables by following the (unique) causal paths that are formed, at the same time cotree elements can be determined by their constitutive relationship and the (also unique) causal paths to tree elements. The next example show how this Procedure can be applied.

Example 1: Consider again the system in Figure 4. In Figure 6 the BGI model of Figure 4 has been redrawn and the bonds are now numbered, as well as the tree branches are
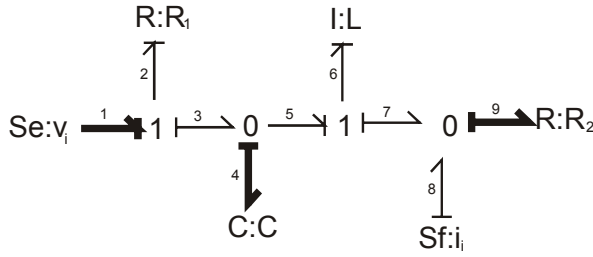
Fig. 6. The BGI model of system in Figure 4.

now remarked as black bonds. According to Procedure 1, the first step is already done. According to criterion 1, the second step is also performed and the bonds corresponding to the tree branches are darkened. In this terms, the bonds belonging to the tree are 1, 4 and 9. the bonds of the cotree are 2, 6 and 8. According with step 3, TME variables are e_1 and e_4 whereas e_9 is the unique THE variable. Step 4 allow to identify f_6 and f_8 as KF variables and f_2 as NF variable. Step 5 of Procedure 1 ask to write any cotree R element flow (NF variable) as a combination of main efforts on the tree and the related constitutive relationship. This is the case of R_1. The flow on R_1 can be written as,

$$
\begin{aligned}
f_2 &= \frac{1}{R_1} e_2 \\
f_2 &= \frac{1}{R_1} (e_1 - e_3) \\
f_2 &= \frac{1}{R_1} (e_1 - e_4)
\end{aligned}
\tag{15}
$$

Step 6 requires to write e_9 as a combination of flows in the cotree. It can be seen that R_2 holds direct causal paths with i_1 through $\{9, 8\}$ and with L through $\{9, 7, 6\}$, therefore it can be written

$$
\begin{aligned}
e_9 &= R_2 f_9 \\
e_9 &= R_2 (f_7 + f_8) \\
e_9 &= R_2 (f_6 + f_8)
\end{aligned}
\tag{16}
$$

Step 7 requires to write the flow in the capacitor as a combination of cotree variables causally connected with it. Reviewing the causal paths, it can be noticed that

$$
\begin{aligned}
f_4 &= f_3 - f_5 \\
f_4 &= f_2 - f_6
\end{aligned}
\tag{17}
$$

Replacing (15) into (17) the result is,

$$
\begin{aligned}
f_4 &= \frac{1}{R_2} (e_1 - e_4) - f_6 \\
f_4 &= -\frac{1}{R_2} e_4 - f_6 + \frac{1}{R_2} e_1
\end{aligned}
\tag{18}
$$

Following step 7 I effort equation can also be derived as

$$
\begin{aligned}
e_6 &= e_5 - e_7 = \\
e_6 &= e_4 - e_9
\end{aligned}
\tag{19}
$$

Replacing (16) into (19) the result is

$$
\begin{aligned}
e_6 &= e_4 - R_2(f_6 + f_8) \\
e_6 &= e_4 - R_2 f_6 - R_2 f_8
\end{aligned}
\tag{20}
$$

Finally, equations (18) and (20) can be modified as

$$
\begin{aligned}
\dot{e}_4 &= -\frac{1}{R_2 C} e_4 - \frac{1}{C} f_6 + \frac{1}{R_2 C} e_1 \\
\dot{f}_6 &= \frac{1}{L} e_4 - \frac{R_2}{L} f_6 - \frac{R_2}{L} f_8
\end{aligned}
\tag{21}
$$

Equation (21) is the SSE corresponding to model in Figure 4. In the next example, a diode circuit is analyzed.

Example 2: A purely resistive system: The central tap full wave rectifier with resistive load

Consider the electrical network shown in Figure 7.
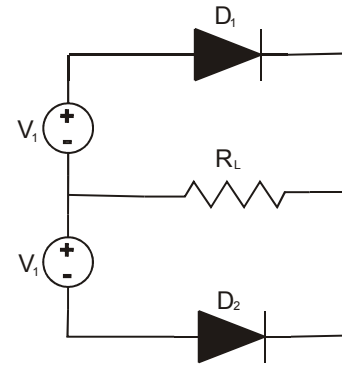
Fig. 7. A full wave rectifier with resistive load.

In Figure 8 the model with causality assigned is shown
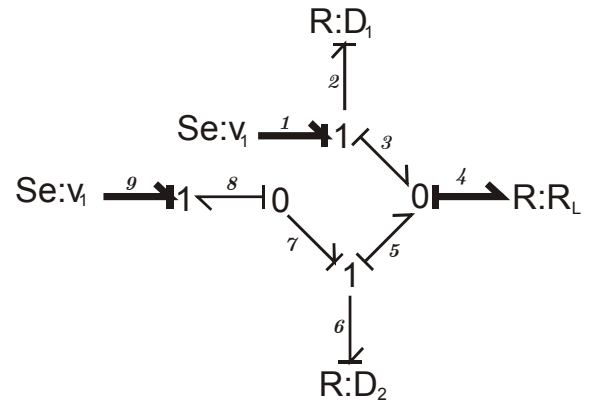
Fig. 8. Bond Graph model of the system in figure 7 with causality assigned.

Notice in figure 8 that bonds 1, 4 and 9 belong to the tree. The R elements in the cotree are D_1 and D_2. Their flows depending on the tree efforts and the constitutive relation (1) are as follows:

$$
f_2 = I_s \left(e^{g e_2} - 1 \right)
$$

or

$$
f_2 = I_s \left(e^{g(e_1 - e_4)} - 1 \right)
\tag{22}
$$

277

and

$$f_6 = I_s \left(e^{g e_6} - 1 \right)$$

or

$$f_6 = I_s \left(e^{g(-e_9 - e_4)} - 1 \right) \quad (23)$$

Writing one equation in this example can lead to the solution of the entire model. This equation is

$$f_4 = f_2 + f_6 \quad (24)$$

Substituting equations (22) and (23) in equation (24) it results

$$f_4 = I_s \left(e^{g(e_1 - e_4)} + e^{g(-e_9 - e_4)} - 2 \right) \quad (25)$$

Also the effort in R_L can be expressed as

$$e_4 = R_L f_4 \quad (26)$$

Substituting equation (26) in equation (25), the following equation can be written

$$f_4 = I_s \left(e^{g(e_1 - R_L f_4)} + e^{g(-e_9 - R_L f_4)} - 2 \right) \quad (27)$$

Equation (27) is an implicit function of the flow f_4 in R_L that can be simulated in Matlab as follows

```
t=[0:1e-4:32e-3];
ei=[36*sin(377*t)];
e9=ei;
n=2;
vt=26e-3; %@T=27 celsius
is=1e-12;
g=1/(n*vt)
R=500;
sixe=size(ei);
sais=sixe(:,2);
fl4=zeros(1,sais);
e2=zeros(1,sais);
e6=zeros(1,sais);
ye=zeros(1,sais);
for j=1:sais
funk=@(f4)((is*(exp(g*(ei(1,j)-(R*f4)))+exp(g*(-e7(1,j)-
(R*f4)))-2))-f4);
varr=fzero(funk,[-.0010 .1]);
fl4(1,j)=double(varr);
ye(1,j)=R*fl4(1,j);
end
plot(t,ei,t,ye)
xlabel('time, s');
ylabel('in and out Voltages, V');
```

And the graphic of the Figure 9 is obtained

The preceding example has shown how to apply the proposed Procedure in purely resistive systems containing diodes represented by the Shockley model (1).

Now a dynamic system is analized. Consider the electric network shown in Figure 10.

The corresponding BGI is shown in Figure 11.
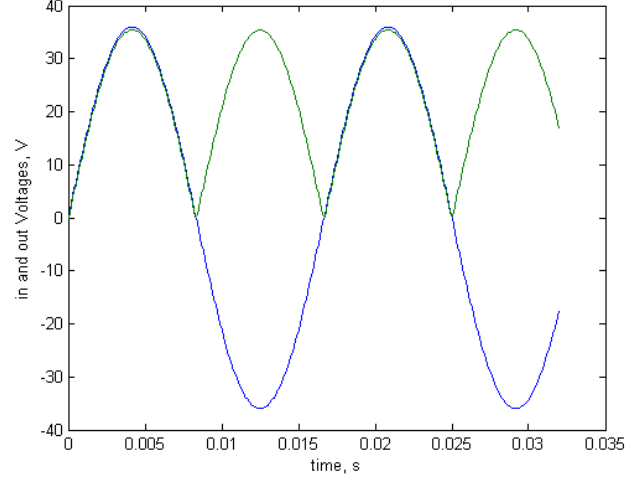
Fig. 9. The input and output voltages of the system in Figure 7.
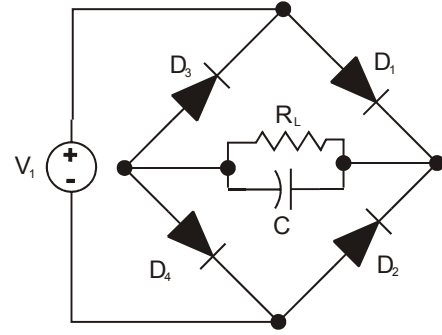
Fig. 10. A full wave rectifier with RC load.
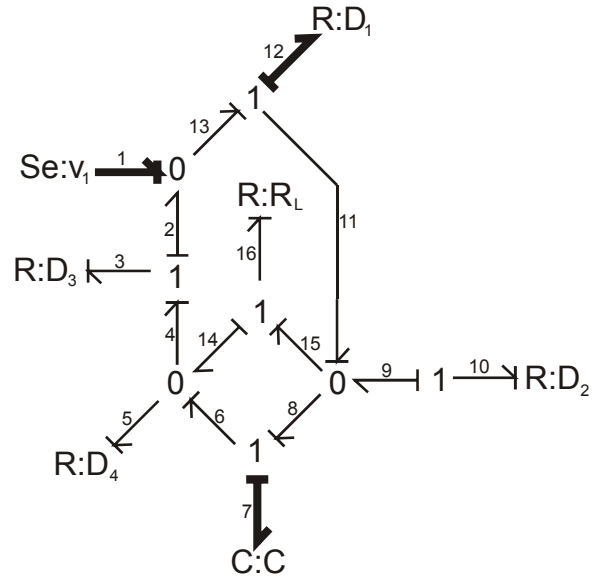
Fig. 11. BGI model of the system in Figure 10.

The tree is conformed by 2 TME variables an 1 THE varaible, (e_1, e_7 and e_{12}, respectively). All elements in the cotree are causally conected with elements in the tree, D_2 is causally connected with D_1 through the causal path $\{10, 9, 11, 12\}$, and with V_1 through the causal path $\{10, 9, 11, 13, 1\}$, and effort e_{10} can be expressed as

$$e_{10} = -e_1 + e_{12} \tag{28}$$

Similarly, R is causally connected with C, D_1 and V_1, and

$$e_{16} = e_7 \tag{29}$$

In a similar way, the following cotree efforts can be expressed in terms of the tree variables as:

$$e_3 = -e_{12} - e_7 \tag{30}$$
$$e_5 = e_1 - e_{12} - e_7 \tag{31}$$

According to general node analysis, and the given procedure, KCL must be applied twice, and in the BG model this means that the flow in the unknowns must be expressed in terms of the flows in the cotree. Starting with the flow in D_1, it can be seen that this element is causally connected with D_2, R_L, D_3 and D_4. This dependence can be written as

$$f_{12} = f_3 + f_5 - f_{10} \tag{32}$$

due to the non linear nature of the system, the e_{12} tree variable is very difficult to eliminate, so a differential algebraic equation must me constructed, and equation 32 shall be rewritten as

$$-f_{12} + f_3 + f_5 - f_{10} = 0$$

by applying the diode constitutive relations, it can be seen that

$$I_s \left[-e^{ge_{12}} + e^{ge_3} + e^{ge_5} - e^{ge_{10}} \right] = 0 \tag{33}$$

Substituting 28, 30 and 31 into 33, the first system equation is obtained as:

$$I_s \left[-e^{ge_{12}} + e^{g(-e_{12}-e_7)} + e^{g(e_1-e_{12}-e_7)} - e^{g(-e_1+e_{12})} \right] = 0 \tag{34}$$

Now, considering the flow in C, this element is causally connected with D_4, R_L and D_3, and its flow can be written as

$$f_7 = f_5 + f_3 - f_{16}$$

or

$$f_7 = I_s \left(e^{ge_5} - 1 \right) + I_s \left(e^{ge_3} - 1 \right) - \frac{1}{R_L} e_7 \tag{35}$$

Substituting 29, 30 and 31 into 35, and isolating the derivative of the state, the second system equation can be written as

$$\dot{e_7} = \frac{I_s}{C} \left(e^{g(e_1-e_{12}-e_7)} + e^{g(-e_{12}-e_7)} - 2 \right) - \frac{1}{R_L C} e_7 \tag{36}$$

equations 34 and 36 conform a differential algebraic system. This system is simulated in Matlab with the next code lines:

```
clear
clc
masa=[1 0;...
0 0];%mass matrix to indicate DAE
```

options = odeset('Initialstep',1e-30,...
'MStateDependence','none','Mass',masa,...
'RelTol',1e-10,'AbsTol',[1e-12 1e-12]);
[T,Y] = ode23t(@full_RC,[0:0.1e-5:32e-3],...
[0 0],options);
ent=100*sin(377*T);
plot(T,ent,T,Y(:,1),'-',T,Y(:,2),'-.')
%plotting the input, the D1 voltage and the output

With the fucntion full_RC defined as
function dy = full_RC(t,y)
%Shockley model parameters
n=2; %silice 2, germanium 1
vt=26e-3; %@T=27 degrees celsius
is=1e-12;
g=1/(n*vt);
%component values
c=47e-6;%load capacitance
r=100;%load resistence
ei=100.*sin(377*t);%input voltage
dy = zeros(2,1);
dy(1)=((is/c)*(exp(g*(-y(2)-y(1)))+...
exp(g*(ei-y(2)-y(1)))-2))-((1/(r*c))*y(1));
dy(2)=-exp(g*y(2))+exp(g*(-y(2)-y(1)))+...
exp(g*(ei-y(2)-y(1)))-exp(g*(y(2)-ei));

Once these codes are implemented in Matlab, the plotting in the following Figure is obtained:
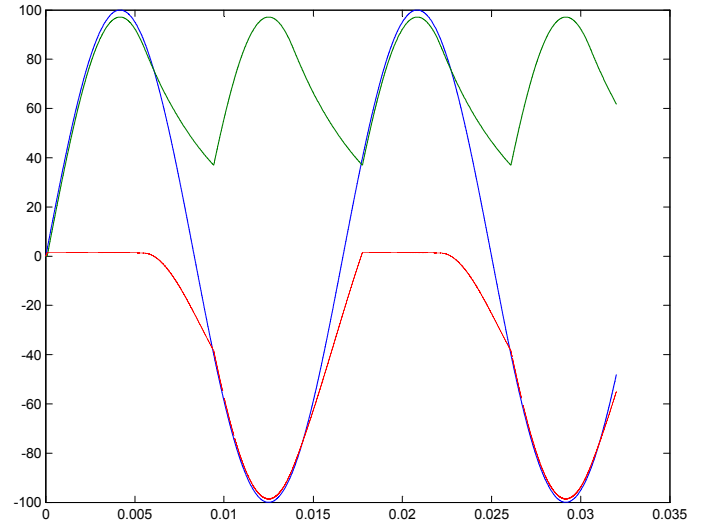
Fig. 12. Simulation results of the system in Figure 11.

V. CONCLUSIONS

An alternative approach to derive equations from a BG model has been presented. Two important tools are employed: The causality and causal paths involved in BG theory and the tree formulation of SSE employed in LG theory.
The proposed approach allow to derive the SSE in a systematic way by following a step by step procedure.

Causal connections between tree elements and cotree elements can be directly found with the suggested Procedure, allowing the equation derivation of systems containing elements with non linear constitutive relations, such as the well known Shockley diode model. General node analysis allows to analyze even purely resistive (algebraic) systems, and also dynamic systems with the proposed modification in BG models.

REFERENCES

[1] Wai-Kai Chen, "Applied Graph Theory" 2nd revised edition, North-Holland Publishing Company, 1976, Netherlands.

[2] Ronald C. Rosenberg and Dean C. Karnopp, "Introduction to physical system dynamics", McGraw-Hill 1983.

[3] C. Sueur and G. Dauphin-Tanguy, "Structural controllability/observability of linear systems represented by bond graphs", J. Franklin Inst., vol. 326 No. 6, P. 869-883, 1989.

[4] C. Sueur and G. Dauphin-Tanguy, "Bond-graph approach for structural analysis of MIMO linear systems", J. Franklin Inst. vol. 328 No. 1 P. 55-70, 1991

[5] Stephen Birkett Jean Thoma and Peter Roe, "A pedagogical analysis of bond graph and linear graph physical system models", Mathematical and computer modelling of dynamical systems, vol. 12. no. 2-3, april - june 2006, P. 107-125, Taylor & Francis.

[6] Dauphin-Tanguy Geneviève, et al., "Les Bond Graphs", Hermes Science Publications, 2000, France.

[7] Ned Mohan, et al. "Power electronics", John wiley and Sons, U.S.A. 2003

[8] Muhammad H. Rashid "Power electronics Circuits, Devices, and Applications", Prentice Hall, 1993

[9] G. Dauphin-Tanguy and S. Scavarda, "Bond-graph modeling of physical systems", Chapman & Hall. A. J. Fossard, D. Normand-Cyrot. 1995. Vol. 1, pp. 33-109.

[10] William Hayt et al. "Engineering Circuit Analysis", 7th ed. 2007, Mc. Graw Hill.

[11] Amalendu Mukherjee, et al, "Bond Graph in Modeling, Simulation and Fault Identification", CRC Press, 2006, India.

[12] Alan S. Perelson, George F. Oster, "Bond Graphs and Linear Graphs" J. of the Franklin Institute, 302 (2), pp.159-185. (1976)

BOND GRAPH MODELING OF A HYDRAULIC VIBRATION SYSTEM: SIMULATION AND CONTROL

Farid Arvani, Geoff Rideout, Nick Krouglicof, Steve Butt

Faculty of Engineering, Memorial University of Newfoundland, St. John's, Canada

(farvani, g.rideout, nickk, sdbutt) @mun.ca

ABSTRACT

Electro hydraulic drives are often used where the dynamic response requirements of the driven load in terms of speed, cycling, accuracy and stability are particularly severe. One such application is the control and drive system associated with a high frequency hydraulic vibrator. A hydraulic shaker controls a valve which directs high pressure hydraulic fluid to generate controlled vibration. Hydraulic shakers typically have long stroke and large force and are useful up to many hundreds of Hertz vibration frequency.

In this paper, bond graph modeling is used to study and simulate the effect of various components in a hydraulic circuit on a shaker performance. The intention is that the electrohydraulic system vibrates at a specified force or displacement amplitude and frequency for the controlled vibration of a load attached to the rod. This paper describes the simulation and control of such system. Feedforward PID control scheme is utilized in the simulations to control the system and to track a variable reference signal.

Keywords: Hydraulics Actuator, Vibrator, Control System, Bond graph modeling

1. INTRODUCTION

Hydraulic control systems are used to control the position, speed or force of resisting loads. A linear hydraulic cylinder or a rotary-motion hydraulic motor usually provides the final drive. A positive displacement pump at a relatively high pressure between 7 and 35 MPa delivers the hydraulic fluid to the system to power the actuators.

Hydraulic control systems are used where relatively large forces and torques accompanied with fast, stiff response of resisting loads are required such as in industrial presses, machine tools and flight simulators, to name a few. They are used for closed-loop control of a response variable(s) such as in aircraft and industrial robot control; and in manual control over a substantial powered motion such as mobile equipment and road vehicle steering (Merritt 1967).

Understanding the hydraulic system dynamics plays an important role in designing controllers for ensuring a good system response or in finding out why an existing system is not performing satisfactorily (Dransfield 1981).

A typical hydraulic control system consists of a power supply, a flow control valve, an actuator, a transducer and a servo controller for closed-loop control. The control valve adjusts the flow of hydraulic fluid into actuator chambers to move the actuator to the desired position based on the control signal generated by the servo controller. A force control scheme operates similarly by using proper load transducer(s).
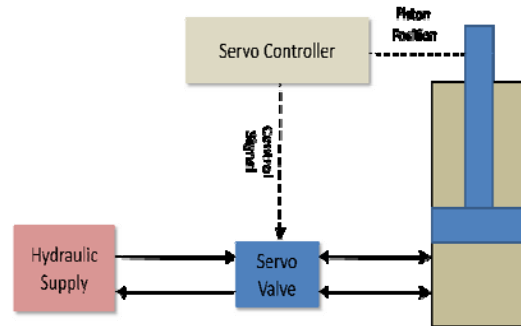
Figure 1: Schematics of the modeled hydraulic system

In this paper, the hydraulic dynamics of a valve-controlled hydraulic cylinder is developed and simulated, and a simplified model of a high frequency hydraulic system is presented. The models are derived from first principles and implemented using bond graph methodologies. Finally, a control system is proposed to perform a tracking control scheme for a hydraulic shaker. Figure 1 depicts the schematics of the modeled system. Solid lines show the bidirectional hydraulic signals while dashed lines designate the electrical control signals. Servovalves and actuators are the most important parts of this hydraulic system and therefore has been the main focus of this work. However other components have been modeled in detail as well.

2. MODELING METHODOLOGY

The simulated hydraulic circuit consists of the bond graph model of a closed–center four-way directional valve, hydraulic actuator, two pressure relief valves, three accumulators and four hydraulic hoses. 20-sim package is used to create and simulate the bond graphs. A tracking PID controller is implemented to perform position control on a linear actuator.

2.1. Servovalve modeling

This section describes the methodologies to model a servovalve which is generally comprised of a torque motor and a valve spool assembly (Jelali and Kroll 2003; Dransfield 1981; Poley 2005). The servovalve consists of a torque motor; application of electrical signal to its coil will generate a torque. The flapper of the torque meter is integrated with the spool assembly. The flapper acts as a mechanical feedback for the control of spool position (Maskrey 1978; Jones 1997; DeRose 2003).

The servovalves can be modeled using two approaches. The first method is to model the subcomponents of the valve. This requires the design parameters for the mechanical and electrical components of the valve. It is possible to derive meaningful transfer functions for electrohydraulic servovalves, and several papers have reported such work (Jelali and Kroll 2003). Unfortunately, servovalves are complex devices and have many nonlinear characteristics which are significant in their operation. These nonlinearities include: electrical hysteresis of the torque motor, change in torque-motor output with displacement, change in orifice fluid impedance with flow and with fluid characteristics, change in orifice discharge coefficient with pressure ratio, sliding friction of the spool, and others (DeRose 2003, Maskrey 1978).

Experience has shown that these nonlinear and non-ideal characteristics limit the usefulness of theoretical analysis of servovalve dynamics in systems design. Instead, the more meaningful approach is to approximate measured servovalve response with suitable transfer functions (Thayer 1965, Clarke 1969). Moog has performed sophisticated analyses of servovalve dynamic response including computer simulations of various nonlinear effects, and up to eight-order dynamics (excluding any load dynamics). The results indicate that these complex analyses have not contributed significantly to servovalve design due to uncertainties and inaccuracies associated with the higher-order effects.

2.1.1. Servovalve Hydraulic Modeling

The actual flow through the valve is dependent upon electrical command signal and valve pressure drop. The flow for a given valve pressure drop can be calculated using the square root function for sharp edge orifices:

$$Q = Q_N \sqrt{\frac{\Delta P}{\Delta P_N}} \tag{1}$$

where Q [gpm] = calculated flow, Q_N [gpm] = rated flow, ΔP [psi] = actual valve pressure drop and ΔP_N [psi] = rated valve pressure drop. The servovalve used in these simulations is high performance Moog- D765 series valves and the values for Q_N and ΔP_N are provided in the valve datasheet. The modeled valve's schematics (Johnson 2008) is shown in figure 2.
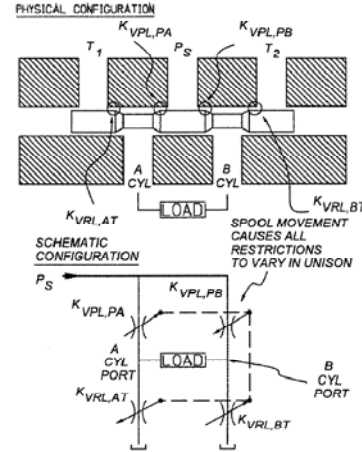
Figure 2: Physical and schematic configuration of the modeled valve (Johnson 2008)

The total valve coefficient is defined to be the flow coefficient at the flow rating conditions:

$$K_V = \frac{Q_N}{\sqrt{\Delta P_N}} \tag{2}$$

It is necessary to know the coefficient of each of the four lands. The valve ratio is defined as

$$\rho_V = \frac{K_{VPL}}{K_{VRL}} \tag{3}$$

where K_{VPL} and K_{VRL} have units of Holes and represent the valve coefficients for the pressure and return lines, respectively. It is safe to assume that valve ratio is unity in absence of evidence to the contrary as there is no reason to deliberately design a non-symmetrical valve. The relation between total valve coefficient and coefficients for valve lands is

$$K_{VPL} = K_V \sqrt{1 + \rho_V{}^2} \tag{4}$$

$$K_{VRL} = K_V \frac{\sqrt{1 + \rho_V{}^2}}{\rho_V} \tag{5}$$

The above equations are derived based on the fact that two orifice resistances are in series and are related by

$$K_V = \frac{K_{VPL} K_{VRL}}{\sqrt{K_{VPL}{}^2 + K_{VRL}{}^2}} \tag{6}$$

It is noteworthy that these parameters take into account the spool geometry, discharge coefficient and oil density (Johnson 2008).

Servovalve delivers a control flow proportional to the spool position which is also proportional to the square root of the pressure drop across the valve. Combining equations (1) and (2) and comparing with the well known formula for flow–pressure relation for turbulent flow,

$$Q = C_d A \sqrt{\frac{2}{\rho}} \sqrt{\Delta P} \qquad (7)$$

it is clear that K_V provides us with the coefficient in the above formula easily which contains parameters for the spool geometry, discharge coefficient and oil density.

Torque motor can be modeled as a series L-R circuit, neglecting any back-EMF effects generated by the load (Poley 2005). The transfer function is

$$\frac{I(s)}{V(s)} = \frac{1}{sL_c + R_c} \qquad (8)$$

where R_c is the combined resistance of the motor coil and the current sensor resistor of the servo amplifier and L_c is the inductance of the motor coil. Values of these parameters for series and parallel wiring configurations of the motor are published in the manufacturer's data sheet.

Control flow, input current and valve pressure drop are related by

$$Q = K_V \times i_V \times \sqrt{\Delta P} \qquad (9)$$

where i_V is the input command corresponding to a specific valve opening and ΔP, pressure drop across the valve, is given by $\Delta P = P_s - P_T - P_L$ where P_s, P_T and P_L are system pressure, return line pressure and load pressure, respectively. The change in flowrate with current and load pressure is shown in figure 3.
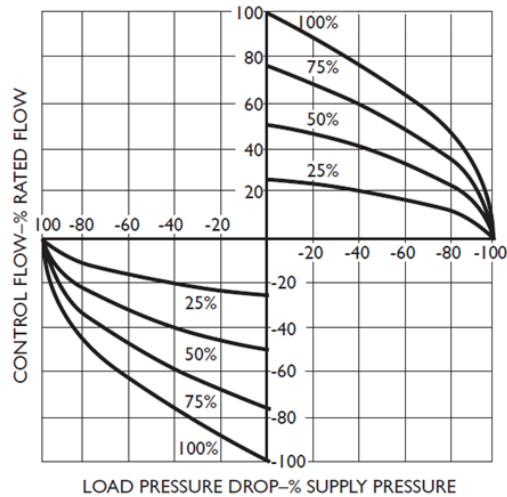
Figure 3: Relationship between flowrate and pressure drop across a servovalve for different input command percentages (Johnson 2008)

The servovalve is not the primary dynamic element in a typical hydraulic servo valve. The servo valves are sized in a way that its natural frequency is at least three times the natural frequency of the hydraulic actuator and therefore it is only necessary to model the valve for a low range of frequencies without major accuracy loss.

A typical performance graph for a high-response servo-valve Moog D765 is shown in figure 4 where the frequency response of the valve is plotted for different command percentages, taken from the manufacturer's datasheet.
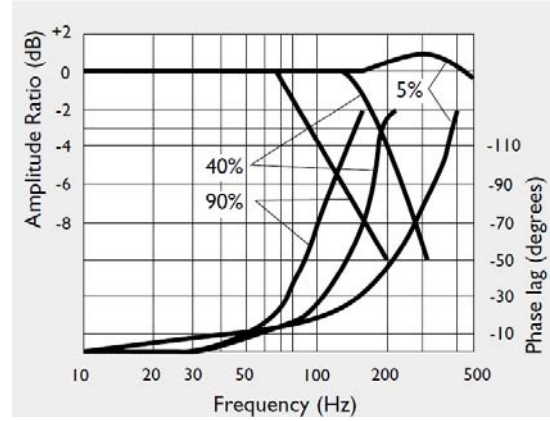
Figure 4: Frequency response of Moog D765 servovalve (Courtesy of Moog Inc.)

A second order model yields an approximation to actual behavior. Standard second order system identification of the frequency response can be performed where suitable values for natural frequency and damping ratio will need to be determined from the graph. Natural frequency (ω_V) can be read fairly accurately from the -3dB or 90 degree phase point of the 5% curve in figure 4. Damping can be determined from an estimate of the magnitude of the peaking present. For an under-damped second order system, the damping factor (S_V) can be shown to be related to peak amplitude ratio (M_V) by the formula (Poley 2005).

$$M_V = \frac{1}{2\zeta_V \sqrt{1 - \zeta_V^2}} \qquad (10)$$

A reasonable estimate of peaking based on the 5% response curve would be about 1.5 dB, which corresponds to an amplitude ratio of about 1.189. Solving the above equation, a value of damping is determined to be about $\zeta_V = 0.48$. Using these values, a model of the servo-valve spool dynamics may be constructed. The input to the model will be the torque motor current normalized to the saturation current obtained from the datasheet, and the output will be the normalized spool position. SIMULNK modeling is used to obtain the simulated frequency response of the transfer function for 5% opening (shown in figure 5).
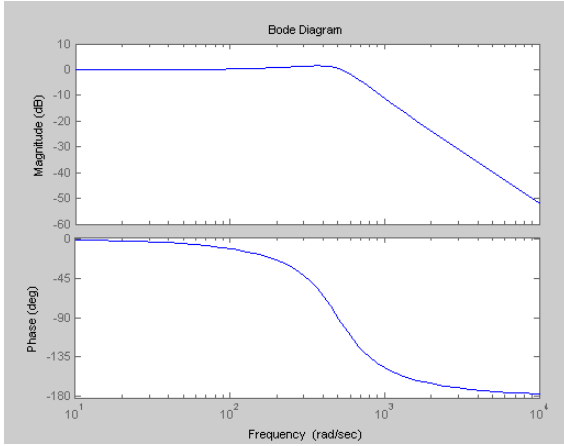
Figure 5: Simulated frequency response of Moog D765 valve

The step response of the servo valve is shown in figure 6. The open loop settling time is 0.04s.
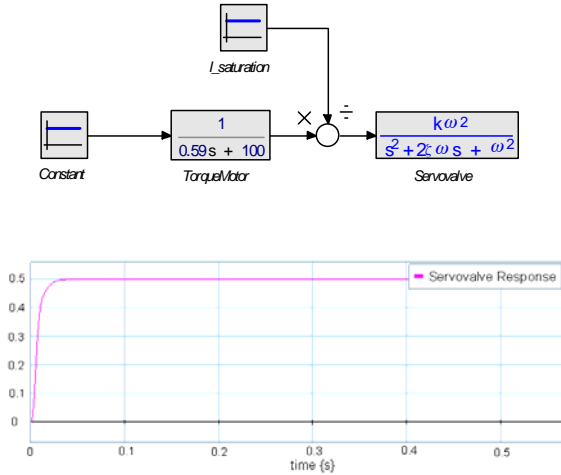
Figure 6: Transfer function for valve spool positioning system and its step response

2.1.2. Servovalve Bond Graph Modeling
The nonlinear pressure-flow relation is modeled using the *modulated R-elements* for each valve land utilizing the equations and formulation in section 2.1.1.
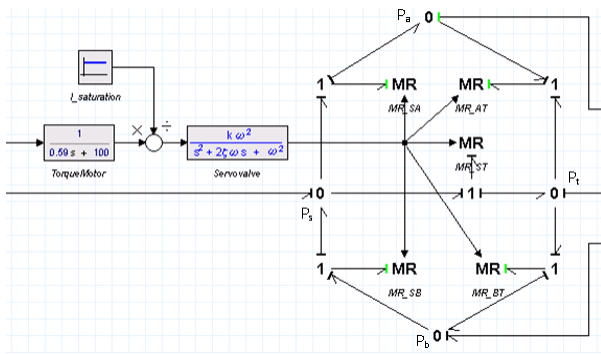

Figure 7: Bond graph model of the servo valve

The model of the torque motor and the servovalve dynamics is described in the previous section. The values for the parameters are taken from the manufacturer's data sheet. *Modulated resistors* are used to model the sharp edge orifices. As it is shown in figure 7, the modulated resistors are located between pressure drops between four valve ports. P_s, P_a, P_b and P_t represent supply pressure, port A and B pressures and return line pressure, respectively. The spool position is provided by the valve dynamics computed by the transfer functions and the *R-elements* are modulated by the value of the spool position. In this model, the pressures of port A and B of the valve are respectively connected to the actuator chambers A and B. Port S is connected to an effort source which models an ideal source of power capable of supplying constant pressure at any flow required (Muvengei and Kihiu 2010). Port T is the pressure for the return line. The infinitesimal leakage between port S and T is also modeled as a *Resistor*.

2.2. Actuator Modeling
The model of the actuator consists of two *C-elements* for each chamber. These capacitors model the oil compressibility in the chambers. It is notable that while the piston rod is retracting or extending, the chamber volume is changing and therefore two modulated capacitors are used to model this effect. The leakage is modeled using Hagen-Poisseille equation (Muvengei and Kihiu 2010). The position of the piston is controlled by the modulation of the *I–element*. The model is shown in figure 8.
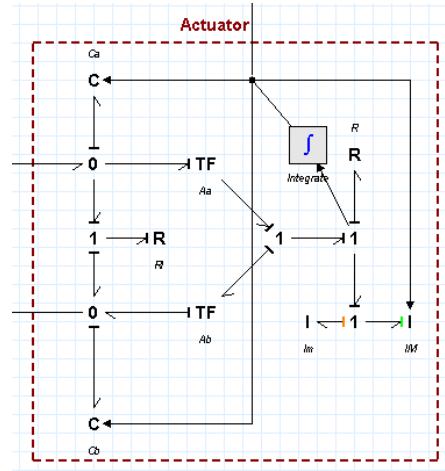
Figure 8: Bond graph model of the linear actuator

2.3. Pressure Relief Valve Model
The pressure relief valve is modeled (shown in figure 9) so that if the pressure is above a threshold, it will bleed the hydraulic fluid to generate balanced pressure.

284

Figure 9: Bond graph model of the pressure relief valve

2.4. Hoses Model

There are four hoses that run between the servovalve, pump and the actuator. Compliance elements are included to take the expansion of the tubes radially, and oil compressibility into account. An *R-element* for the pressure loss due to the wall friction in the hose is also considered (Hölcke 2002). An *I-element* is added to account for the fluid inertia (Karnopp, Margolis, and Rosenberg 2006).

2.5. Bond Graph Model of Hydraulic Circuit

The full bond graph model using the subcomponents of the previous subsections is shown in figure 10. This model includes the model of the servovalve, two pressure relief valves, three accumulators and four corresponding hydraulic hoses. This model is simplified by removing the models for the accumulators and the hoses for simplification in analysis and using a PID controller to perform position control.

3. CONTROLLER DESIGN AND SIMULATION

In this section, a feedforward PID is implemented to simulate the desired motion on the actuator (Caputo 1996). The controller utilizes the piston position of the actuator as a feedback signal which is obtained by the integration at the *I-element* associated with the piston. The generated model is used to perform PID control on two scenarios. The following figure shows the complete bond graph model with the implemented control.

First a constant reference position is used to carry out position control on the cylinder. The stroke length is 30cm and the initial position of the piston is 15cm. The controller will extend the piston 5 cm and stop (shown in figure 11).
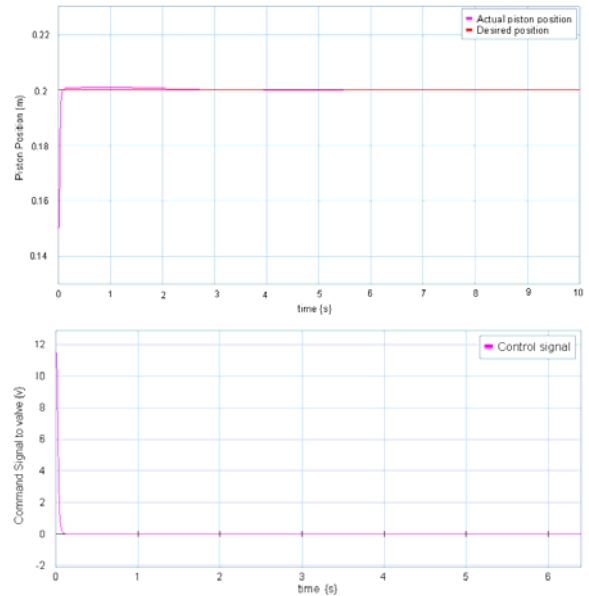
Figure 11: Step response of the system for 5cm movement – Desired and actual piston location (above) and corresponding control signal (below)

In the second scenario, a sinusoidal profile is used to perform target tracking to simulate the behavior of the system for oscillation generation. In this case, a sinusoidal signal, is used to perform position tracking and as it is shown in the figure 12, it can perform the tracking with a good performance. The control signal to the valve is shown in figure 12.
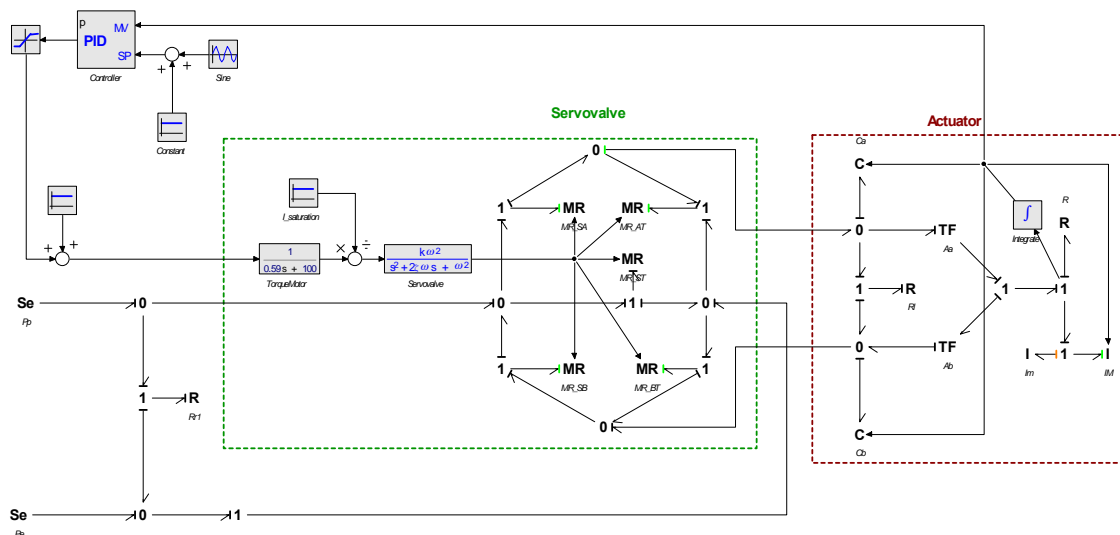
Figure 10: Bond graph model of the hydraulic system
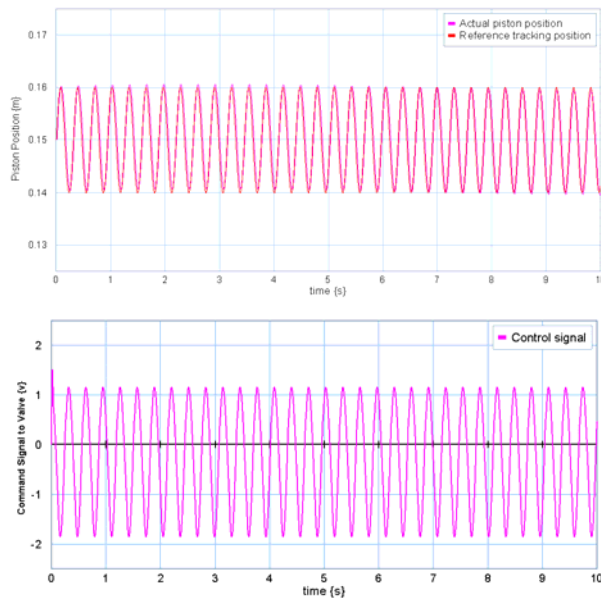
Figure 12: Response of the system for sinusoidal piston movement – Desired and actual piston location (above) and corresponding control signal (below)

Feedforward PID controller enables increased dynamic system responsiveness while decreasing positioning and velocity error. Feedforward gains are predictive gains that contribute to the control signal in the form of a linear combination of the target velocity and acceleration terms. The result is higher performance and durable machine control due to smoother motion. The control signal is applied to the valve and in both cases is appropriate in the sense that they do not excite the valve with jerk that usually causes immature failure of the components.

APPENDIX A - PARAMETERS AND SYMBOLS

Symbol	Description	Value
ρ	Oil Density	867 kg/m^3
A_{h1}	Hose 1 Area	0.000791 m^2 (1 ¼")
A_{h2}	Hose 2 Area	0.000506 m^2 (1")
L_{h1}	Hose 1 length	20 cm
L_{h2}	Hose 2 length	20 cm
υ	Kinematic Fluid Viscosity	27cSt
μ	Absolute Fluid Viscosity	0.023409 Pa.s
B	Bulk modulus of Oil	1.5 GPa
E	Elastic Modulus of hydraulic hoses	30 MPa
m	Rod and piston mass	2.5 kg
M	Load mass	500kg

Actuator parameters	
Description	**Value**
Rod Diameter	1 ¾"
Piston Diameter	2 ½"
Length of cylinder	0.3m
Length of piston	0.01m

REFERENCES

Merritt, H. E., 1967. *Hydraulic Control Systems*. New York: John Wiley and Sons.

Jelali, M. and Kroll, A., 2003. *Hydraulic Servo Systems- Modelling*, *Identification and Control*, Springer Verlag.

Dransfield, P., 1981. *Hydraulic Control systems- Design and Analysis of Their Dynamics*. Springer Verlag.

Muvengei, M. and Kihiu, J., 2010. Bond Graph Modeling of Inter-Actuator Interactions in a Multi-Cylinder Hydraulic System. *International Journal of Mechanical, Industrial and Aerospace Engineering*. *World Academy of Science, Engineering and Technology*, 32-41.

Johnson, J.L, 2008. *Designer`s Handbook for Electrohydraulic Servo and Proportional Systems*. 4th edition. IDAS Engineering Inc.

Karnopp, D. C., Margolis, D. L, and Rosenberg, R. C., 2006. *System Dynamics: Modelling and Simulation of Mechatronic Systems*. 4th edition. New York:John Wiley and Sons.

Maskrey, R.H. and Thayer, W.J., 1978. A Brief History of Electrohydraulic Servomechanisms. *ASME Journal of Dynamic Systems, Measurement and Control* Vol 100, No 2, pp 110-116.

Thayer, W.J., 1965. Transfer Functions for Moog Servovalves. *Technical Bulletin 103*. Moog Inc.

Jones, J.C., 1997. Developments in Design of Electrohydraulic Control Valves. *Technical Paper*. Moog Inc.

Clarke, D.C., 1969. Selection and Performance Criteria for Electrohydraulic Servovalves. *Technical Bulletin 103*. Moog Inc.

DeRose, D., 2003. Proportional and Servo Valve Technology. *Fluid Power Journal*. March/April, 8-15.

Caputo, D., 1996. Digital Motion Control for Position and Force Loops. Proc. National Conference Confreence. Fluid Power. Vol: Issue 47th:263 Vol 1-268.

Hölcke, J., 2002. *Frequency Response of Hydraulic Hoses*. Licentiate Thesis. Royal Institute of Technology, KTH, Stockholm.

Poley, R., 2005. DSP Control of Electro-Hydraulic Servo Actuators. *Application Report* SPRAA76. Texas: Texas Instruments.

D765 Series Servovalves. *Product Datasheet*. Moog Inc.

20sim version 4.1. Controllab Products B.V.. Enchede, Netherlands. http://www.rt.el.utwente.nl/20sim/.

NEW CONCEPT OF MODELING TO PREDICT TEMPERATURE IN OIL-HYDRAULIC CYLINDER CHAMBER CONSIDERING INTERNAL FLOW

Kazuhiro TANAKA[a], Kouki TOMIOKA[b], Masaki FUCHIWAKI[c], Katsuya SUZUKI[d]

[a)(c)(d]Kyushu Institute of Technology
[b]NITTO DENKO Co.

[a]kazuhiro@mse.kyutech.ac.jp, [b]kouki_tomioka@gg.nitto.co.jp, [c] futiwaki@mse.kyutech.ac.jp,
[d]kachandesu2002@yahoo.co.jp

ABSTRACT

It is important to study a precise and practical way to predict temperature rise of oil hydraulic systems. It is interesting whether internal flow patterns have any effect on temperature change in a cylinder. Nowadays commercial CFD (Computational Fluid Dynamics) codes give accurate solutions of temperature as well as flow patterns. The authors have calculated the flow patterns in a cylinder using a reliable CFD code and have found that three-dimensional effect is very strong there and the internal flow patterns can be classified into two flow regions with 3D vortex flow pattern and 1D parallel flow pattern. So, it becomes possible to predict precisely temperature change in a cylinder by incorporating the internal flow patterns with modeling method based on a lumped parameter system. A new method with prediction of temperature rise in a cylinder is proposed in this study.

Keywords: oil-hydraulic cylinder, temperature prediction, lumped parameter system, vortex flow pattern

1. INTRODUCTION

An oil-hydraulic system has become so compact and highly-pressurized recently. Its small heat capacity or small surface area for heat dissipation causes high temperature rise of a system while it works for long time under highly loaded condition. The high temperature rise causes deterioration of working oil and the system may become dangerous finally. From this viewpoint, it is important to study a precise and practical way to predict system temperature.

In a viewpoint that energy loss causes temperature rise in a system, modeling and simulation by Bondgraph method is very effective to predict the temperature rise because the method is based on energy balance. In the past some studies have been conducted to predict temperature of oil-hydraulic systems considering heat generation and heat transfer in components of the systems. Johansson analyzed heat generation from a hydraulic motor embedded in an aircraft actuator system and heat transfer from oil passages (Johansson

et al., 2001). Yamamoto analyzed heat generation and heat transfer from oil circuits of a mobile crane (Yamamoto et al., 1997). Tomioka proposed a new idea to predict temperature rise of a hydraulic pipe by considering heat transfer between working oil and a pipe-housing (Tomioka et al., 2005). In this study, one-dimensional Bondgraph modeling based on a lumped parameter system is applied to predict temperature inside the pipe and three-dimensional heat conduction analysis method is applied to calculate temperature distribution in the pipe-housing. By coupling the above two prediction methods the temperature distribution in the pipe component is well predicted time-dependently.

Following Tomioka's study (Tomioka et al., 2005), a new idea to predict temperature distribution in a cylinder is studied in the present study. While a cylinder works for long time, the temperature distribution appears there. In modeling with a lumped parameter system, it is generally believed that the more dividing number of elements becomes, the more precise calculation results will be obtained. It is interesting whether the idea is applicable to a cylinder or not and whether an internal flow in a cylinder has any effect on the temperature distribution.

Commercial CFD (Computational Fluid Dynamics) codes are developing rapidly accompanied with remarkable developments of computers in recent years. Not only the improvement of analysis precision, but also new functions such as moving body analysis have been introduced and sophisticated. The authors have calculated internal flow patterns in a cylinder using a reliable CFD code and have found that three-dimensional effect is very strong there. However, a prediction of time-dependent temperature change is not practical because the calculations with CFD codes are much time-consuming.

In this study, internal flow patterns in an oil-hydraulic cylinder has been calculated and verified that the flow patterns can be classified into two flow regions with 3D vortex flow pattern and 1D parallel flow pattern. So, it becomes possible to predict precisely the temperature distribution in the cylinder by incorporating the internal flow patterns with modeling based on a lumped parameter system.

NOMENCLATURE

A [m^2] heat dissipation surface area of cylinder
a [m] representative length of inlet head port
b [m] length from cylinder end to inlet port center
c_p [J/kgK] specific heat at constant pressure
D [m] cylinder diameter
E [J] internal energy
H [J] enthalpy
h [W/m^{2}K] heat transfer coefficient
Lx [m] location of dividing between 3D and 1D flow
N [s] reciprocating motion cycle of piston
P [Pa] pressure
Q [m^3/s] volumetric flow rate
R_c [m] standard diameter of inlet head port
$\dot{q}$ [W] amount of heat inflow
T [K] temperature
t [s] time
$\mathbf{U}$ [m/s] velocity vector
V [m^3] volume of cylinder head chamber
v [m/s] velocity
W [W] power consumed for outside
x [m] axial location of piston head
δ [-] Kronecker delta
λ [W/mK] thermal conductivity
μ [Pas] viscous coefficient
ν [m^2/s] kinematic viscous coefficient
ρ [kg/m^3] density
Suffix
c cylinder
in inflow
p piston
$1D$ one-dimensional
$3D$ three-dimensional

2. NUMERICAL SIMULATION METHOD
2.1. One-dimensional Model of Heat Balance
Figure 1 shows a schematic figure of heat balance in a cylinder. Temperature in a cylinder chamber is calculated as the sum of amount of heat inflow and outflow of working oil, volumetric change of working oil, and heat dissipation to cylinder housing. First law of thermodynamics derived the following equation.

$$\frac{dE}{dt} = \frac{dH}{dt} + \dot{q} - W \tag{1}$$

If pressure loss of working oil could be negligibly-small in the inlet and outlet port and working oil does not work against outside, changing rate of the internal energy is expressed by the following equation.

$$\frac{dE_c}{dt} = c_p \rho Q T_i + h A_c \left(T_{wall} - T_c\right) \tag{2}$$

$$if \quad Q \geq 0 (in) \; T_i = T_{in}$$
$$Q \leq 0 (out) \; T_i = T_c$$

When the cylinder volume changes time-dependently with piston motion the changing rate of internal energy is expressed by the following equation.

$$\frac{dE_c}{dt} = c_p \rho \frac{dV_c T_c}{dt} = c_p \rho \left(T_c \frac{dV_c}{dt} + V_c \frac{dT_c}{dt} \right) \tag{3}$$
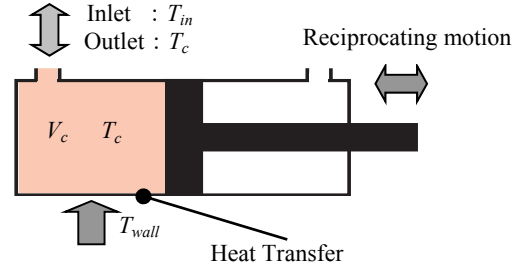
Figure 1: Heat Balance in Cylinder

Equation (4) is derived from Eq. (2) and Eq. (3) in the cylinder chamber.

$$\frac{dT_c}{dt} = \frac{1}{c_p \rho V_c} \left\{ c_p \rho Q T_i + h A_c (T_{wall} - T_c) - c_p \rho T_c \frac{dV_c}{dt} \right\} \tag{4}$$

In the above equation, the first term of right hand side indicates the inflow or outflow amount of heat with piston motion, the second term the amount of heat transfer between the cylinder chamber and cylinder housing, and the third term the effect of cylinder volume change. The temperature in the cylinder chamber can be calculated by solving Eq. (4) using Runge-Kutta method.

2.2 Governing Equations in CFD Calculations
A commercial CFD code, ANSYS-CFX(ANSYS, 2005) which is world-widely used today and gives reliable results even in case of moving bodies, is used to analyze flow patterns as well as temperature distribution time-dependently in a hydraulic cylinder in the present study. However, time-dependent CFD calculations consume so much time that the calculating method is not practical for system simulations despite giving precise results. In the present study the validity of temperature prediction method proposed as one-dimensional modeling method based on a lumped parameter system (hereafter 1D-Model) is verified by comparing numerical results obtained through three dimensional analyses with a CFD code (hereafter 3D-CFD).

The flow patterns as well as heat transport phenomena are solved using the following equations, where Eq. (5) is equation of continuity, Eq. (6) Navier-Stokes equation, and Eq. (7) energy equation.

$$\frac{\partial \rho}{\partial t} + \nabla \bullet (\rho U) = 0 \tag{5}$$

$$\frac{\partial \rho U}{\partial t} + \nabla \bullet (\rho U \otimes U) = \nabla \bullet \left(-P\delta + \mu \left(\nabla U + (\nabla U)^T \right) \right) \tag{6}$$

$$\frac{\partial \rho H}{\partial t} - \frac{\partial P}{\partial t} + \nabla \bullet (\rho U H) = \nabla \bullet (\lambda \nabla T) \tag{7}$$

The discretization way of the above equations is based on finite volume method and hexahedrally-bonded structured grid is generated as a numerical grid system.

3. CHARACTERISTICS OF 1D-MODEL
3.1. Analysis Conditions
In order to study characteristics of Eq. (4), a simple system shown in Fig. 2 is firstly set as an example model. This model indicates a piston head chamber of a

double-acting cylinder. The cylinder diameter is 100 mm, the initial length of the chamber 500 mm and the head port with rectangular configuration 10×10 mm². The pressure at the head port is set as 0 Pa as the boundary condition and the piston is moved along sine function with motion frequency 0.1 Hz and amplitude 100 mm. The initial temperature in the chamber is set 20 deg. as the initial condition and the temperature of working oil inflow is always set 40 deg. In the present study, cylinder walls are set adiabatic not to consider interference between the chamber inside and the cylinder housing but to consider energy balance only in the chamber inside. The grid number for 3D-CFD calculation is 61,000 nodes and 57,000 elements which is enough to calculate unsteady flow patterns as well as heat transport phenomena. The analysis conditions and properties of the working oil are shown in Table 1.

3.2. Results

Figure 3 shows the comparison between 1D-Model calculation and 3D-CFD simulation on the cylinder temperature. The horizontal axis indicates time and the vertical axis temperature of working oil in the cylinder chamber. The broken line indicates volume-average temperature of 3D-CFD simulation using Eq. (5)-(7) and the solid line the calculated temperature based on 1D-Model using Eq. (4). At inflow of working oil from Time=7.5 to Time=12.5, the temperature rise is almost same between 1D-Model calculations results and 3D-CFD results. However, at outflow of working oil from Time=12.5 to Time=17.5, the temperature obtained in 3D-CFD simulations decreases despite the constant temperature obtained in 1D-Model calculations.

This is because heat amount of working oil inflow is simultaneously diffused in the chamber and the inside temperature is fixed to the average temperature in 1D-Model and the temperature is kept constant there at outflow of working oil. On the other hand, in the results of 3D-CFD simulations at inflow the working oil with high temperature stays near head port and at outflow the working oil with high temperature flows out first and as a result the volume-average temperature in the chamber decreases. This difference of mechanism causes the error of 1D-Model. It is verified that the prediction precision of result based on Eq. (4) is not enough and the model should be improved to predict the chamber temperature more precisely.

3.3. Effect of Dividing Number of Elements in 1D-Model

In order to improve precision of temperature prediction based on 1D-Model lumped parameter system, cylinder chamber is divided into some partial elements in axial direction, as shown in Fig. 4. The effect of dividing number of elements is studied in case of 3, 5 and 10 divisions. The element volume is time-dependently changed at the element, which corresponds to amplitude of piston motion, and the temperature change is expressed by the last term of Eq. (4).

Figure 5 shows the result of comparison on the
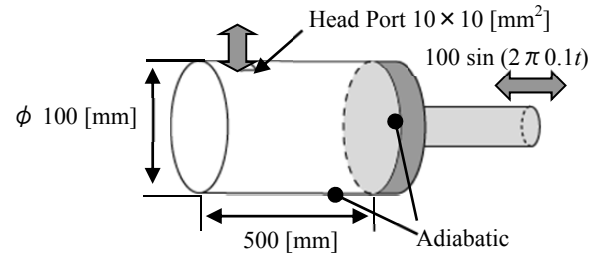
Figure 2: Test Model for Cylinder

Table 1: Analysis Conditions

Physical Properties	ρ =860 [kg/m³], μ =0.027 [Pas] λ =0.13[W/mK], c_p=1894 [J/kgK]
Turbulence Model	SST (Shear Stress Transport)[5]
Time Step	1D (Eq. (4)) : 1.0×10⁻⁵ [s]
	3D (CFD) : 1.0×10⁻² [s]
Boundary Conditions	Inlet : T_{in} = 40 [deg.], P=0.0 [Pa] Wall : Adiabatic, Non-Slip Wall
Cylinder Motion (sin wave)	Frequency :0.1 [Hz] Amplitude :100 [mm]
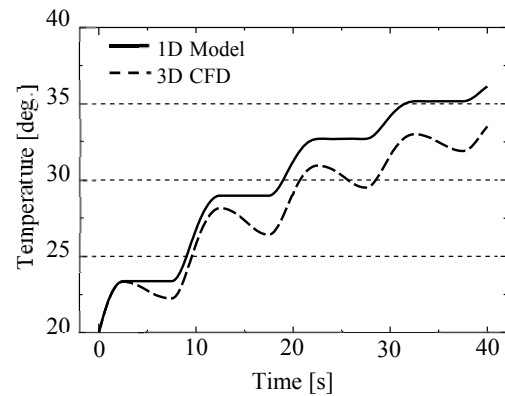
Figure 3: Temperature Change at Cylinder
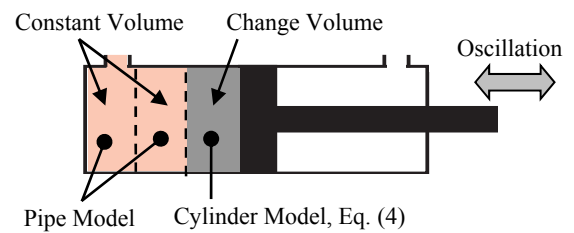
Figure 4: Divide Model for Cylinder (3 Divisions)

volume-average temperature of cylinder between the results in 1D-Model with 3, 5 and 10 divisions and the results in 3D-CFD simulation. The prediction becomes more precise. However, Fig. 5 shows strange tendency. The more precise the prediction becomes, the more the dividing number increases in case of 3 and 5 divisions. However, the difference from 3D-CFD results increases inversely in case of 10 divisions.

$$Error = \frac{1}{2N} \int_0^{2N} \frac{|T_{3D} - T_{1D}|}{T_{3D}} dt \qquad (8)$$

Here the inverse tendency is discussed. The difference between 1D-Model and 3D-CFD results is represented as the error defined in Eq. (8), where N, T_{3D} and T_{1D} indicate the piston motion cycle, the temperature predicted through 3D-CFD simulation, and the temperature predicted through 1D-Model calculation, respectively.

Figure 6 shows relationship between the error and the dividing numbers of elements in modeling of the cylinder. The error reaches a minimum at 6 divisions; however the error becomes higher as the dividing numbers of elements increases. Figure 7 shows a 3D-CFD result of the streamlines through the inlet port in the cylinder chamber when the piston is moving with the maximum speed at neutral position, x = 0 mm, from the maximum displacement, x = -100 mm. In this figure three arrows are described at intervals of 100 mm from the cylinder head end. 3D vortex flow is strongly generated near the inlet port and the flow becomes like parallel as the working oil flows further. Here it can be possible to classify the flow field in the cylinder chamber into two parts, one of which is 3D vortex flow region and another 1D parallel flow region. In the 3D vortex flow region the temperature distribution is so complicated and non-uniform by the vortex flow that the error becomes higher when the region with 3D vortex flow is automatically divided into some elements like the parallel flow region. Consequently the manner to divide automatically the region with complex flow pattern into some elements with the same volume causes imprecise prediction.

4. IMPROVEMENT OF CYLINDER 1D-MODEL

Here a new modeling concept is introduced to improve the 1D-Model of a cylinder based on lumped parameter model. The concept is that the region with 3D vortex flow should be separated from the region with 1D parallel flow in the modeling. Figure 7 shows an internal flow pattern inside a cylinder obtained by 3D-CFD calculation. The figure shows that a transition region between 3D vortex flow and 1D parallel flow can be seen near x= 200 mm, which is called Location of Dividing flow patterns (hereafter LoD).

Figure 8 shows the result obtained when 3D vortex flow region is modeled as one element and 1D parallel flow region is divided into 3, 5 and 10 elements. As the dividing number of elements increases, the result of 1D model asymptotically approaches 3D-CFD calculation results. As a result, it is verified that in order to obtain precise prediction using a lumped parameter system, it is better not to divide total region of a cylinder into elements automatically but to model 3D-vortex flow region as one element and the other region as some elements.

Figure 9 shows the relationship between the prediction error and the dividing number of elements when LoD is set different locations. When LoD is set x= 200 mm, the error is minimum at the respective dividing number of elements. This result means that it is important for precise prediction to set LoD suitably by comparing with 3D-CFD calculation results.
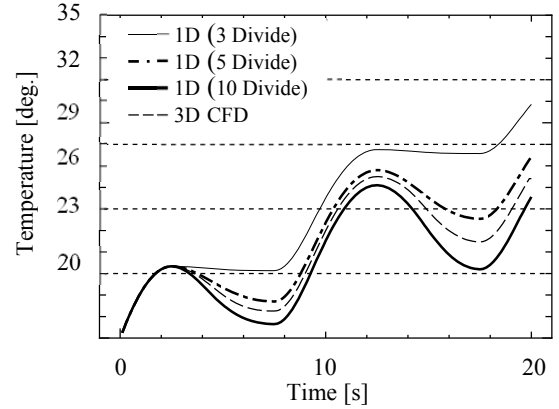
Figure 5: Results of 3D-CFD and 1D Analysis
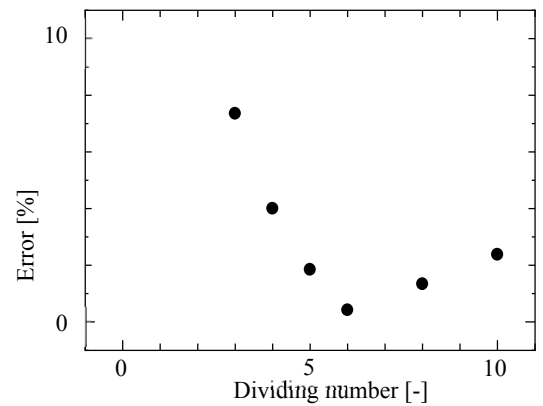
Figure 6: Prediction Error of Temperature at Cylinder
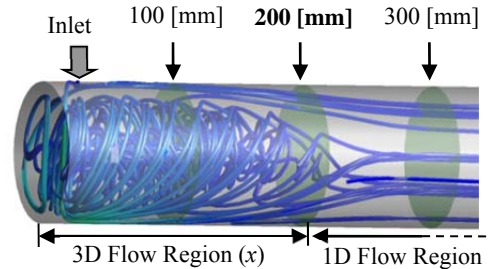
Figure 7: Streamlines at Cylinder Chamber
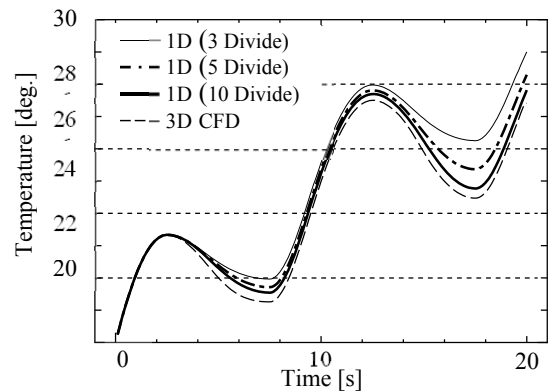
Figure 8: Temperature Change at Cylinder

5. GENERALIZATION OF NEW 1D-MODEL OF A CYLINDER

It is important to determine the size of 3D vortex flow region in order to set suitable LoD for precise prediction. However, because cylinder size as well as the properties of working oil change the suitable LoD, it is inevitable to generalize the way to determine LoD in order to apply the proposed model to any cylinder. Here generalization of the way to determine LoD has been performed through dimensional analyses considering directional dependency against any cylinder with a different size. The size of 3D vortex flow region may depend on the piston speed. Here it is supposed that the piston speed is constant when the piston moves from piston head side to rod side. Table 2 shows calculating conditions.

Figure 10 shows the maximum vorticity of rotation about the cylinder axis at any axial length from the cylinder head end. In this study, LoD is defined as the length from the cylinder head end to the location with 95% down of the maximum vorticity.

All of parameters affecting on LoD are shown in Fig. 11. Considering directional effects in this study, dimension of length in each direction is defined as $[L_x]$, $[L_y]$ and $[L_z]$ in three dimensional coordinate system. A cylinder diameter, piston velocity, velocity of inlet flow, length of inlet port, and length from the cylinder head end to the center of the inlet port are defined as D, $[L_z]$, v_p $[L_x/s]$, v_{in} $[L_z/s]$, a $[L_x]$, and b $[L_x]$, respectively. The viscous coefficient and the density are selected as the important parameters of the properties of working oil. They are defined as μ $[kgL_z/L_xL_ys]$ and ρ $[kg/L_xL_yL_z]$, respectively.

Through Buchingham's π theorem, the following equation is derived from these parameters.

$$F\left(\frac{D}{a} \cdot \frac{\rho D v_p}{\mu}, \frac{\rho D v_{in}}{\mu}, \frac{b}{a}, \frac{x}{a}\right) = 0 \qquad (9)$$

In the above equation, the non-dimensional length of LoD expressed by x/a is a function of Reynolds number expressed by $\rho D v_{in}/\mu$ with representative velocity v_{in}, the product of Reynolds number with representative velocity v_p and D/a, and the non-dimensional length b/a. After determining unknown parameter values so that LoD is the same as the result of 3D-CFD calculations, the following characteristic equations are derived crossing the dividing line of $\rho D v_{in}/\mu = 2.3 \times 10^4$.

LoD $[x/a]$:
(i) $\rho D v_{in}/\mu \geq 2.3 \times 10^4$

$$\frac{x}{a} = 1.44\left(\frac{D}{a} \cdot \frac{\rho D v_p}{\mu}\right)^{-0.014}\left(\frac{\rho D v_{in}}{\mu}\right)^{0.19}\left(\frac{b}{a}\right)^{0.27} \qquad (10)$$

(ii) $\rho D v_{in}/\mu < 2.3 \times 10^4$

$$\frac{x}{a} = 0.015\left(\frac{D}{a} \cdot \frac{\rho D v_p}{\mu}\right)^{-0.69}\left(\frac{\rho D v_{in}}{\mu}\right)^{0.96}\left(\frac{b}{a}\right)^{0.098} \qquad (11)$$

Figure 12 shows the comparison of LoD obtained between through the above characteristic equations and through 3D-CFD calculations. The abscissa and ordinate indicate Reynolds number with representative
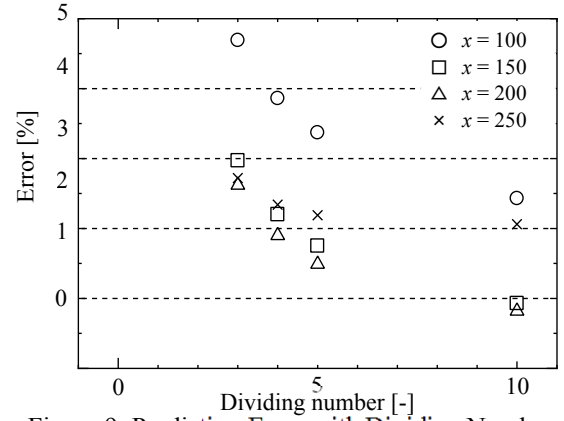
Figure 9: Prediction Error with Dividing Number

Table 2: Conditions for Dimensional Analysis

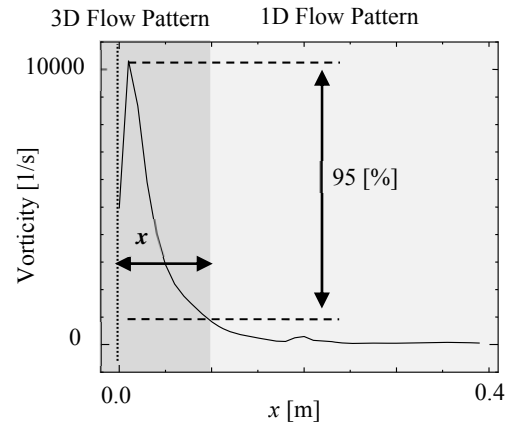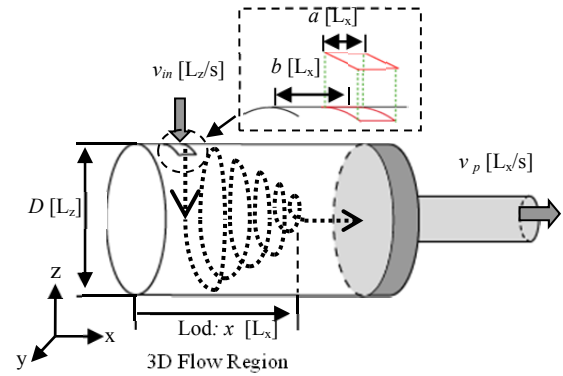| Cylinder Diameter	D=0.07, 0.1, 0.15 [m]
Piston Velocity	V_p= 0.05, 0.1, 0.15 [m/s]
Kinematic Viscosity	32×10^{-6}, 11.5×10^{-6} [m²/s]

Figure 10: Definition of LoD

Figure 11: Parameters of Dimension Analysis

velocity v_{in} and the non-dimensional length LoD $[x/a]$, respectively. The maximum difference is 14% and the both results agree well. It is verified that the suitable LoD for precise prediction can be calculated through the above characteristic equations. Moreover, even when the shape of inlet port is round the results are almost the same as the above results in case of rectangular inlet port by setting inlet port diameter as the representative length, a.

Figure 13 shows applicability of Eq. (10) and (11) to any cylinder contained in a catalogue of products (TAIYO, LTD, 2005). The abscissa and ordinate indicate cylinder diameter D and non-dimensional diameter of inlet port R_c/D, respectively. LoD of all practical products contained in the catalogue (Series A, B and C, TAIYO, LTD, 2005) can be represented well by the above equations. It is verified that the above equations can be applicable to any practical cylinder.

Table 3 shows comparison of calculating time between in 1D-Model calculation proposed in this study and in 3D-CFD calculation. In case of calculating temperature distributions inside a cylinder for practical 20 seconds, 1D-Model calculation takes only 10 [s] though 3D-CFD calculation takes $1.9\times10^{+5}$[s]. The difference of volume-average temperature of cylinder chamber between 1D-Model analysis according to Eq. (10) and (11) and 3D-CFD calculation is about 1% and precise prediction can be obtained successfully.

6. CONCLUSIONS

1D-Model predicting temperature of cylinder chamber is proposed and its applicability to any cylinder is verified by comparing with 3D-CFD calculation results.

(1) In a cylinder chamber, inflow generates 3D vortex flow region near the inlet port and does not mix with working oil previously fulfilled. And parallel flow region appears near a piston head. The working oil at this 3D vortex region flows out first at delivery time. When a piston moves with reciprocating motion, temperature has its distribution in axial direction inside a cylinder chamber and it becomes different from volume-average temperature.

(2) In modeling cylinder chamber based on a lumped parameter system in order to predict axial distribution of temperature, it is important to divide 3D vortex flow region and parallel flow region like 1D flow. If the 3D vortex flow region is divided automatically into some elements with the same volume, it may cause imprecise prediction. Inversely modeling 3D-vortex flow region as only one element and dividing another region like 1D parallel flow into some elements give precise prediction of temperature in a cylinder chamber.

(3) Dimensional analysis derives the characteristic equations to determine the suitable length of Location of Dividing between 3D vortex flow region and parallel flow region. The prediction model based on a lumped parameter system by use of Location of Dividing gives quickly simulation results close to 3D-CFD calculation results. The characteristic equations on Location of Dividing proposed in this study are applicable to any practical product.

REFERENCES

ANSYS, 2005, *ANSYS CFX-Solver, Release 10.0.*

Johansson, B., Anderson, J., Krus, P., 2001. Thermal Modeling of an Electro-Hydrostatic Actuation System,*Proceedings of the International Conference on Recent Advantages in Aerospace Actuation System and Components*, Toulouse, France.
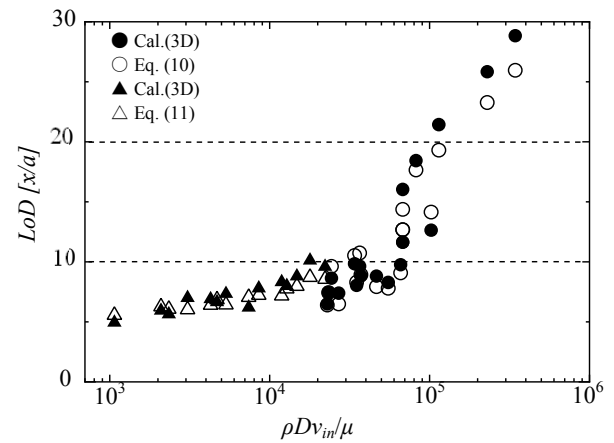
Figure 12: Non-dimensional Directional Analysis
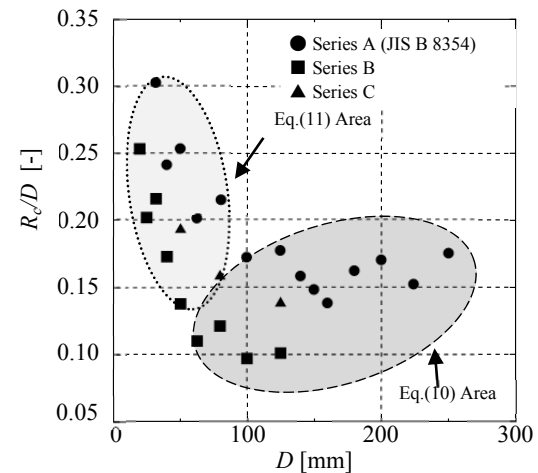
Figure 13: Area of Prediction Equation

Table 3: Calculation Time and Error

CPU Power	Calculation Time		Error
Pentium Xeon 3.6 GHz	1D Analysis	10 [s]	1 [%]
	3D-CFD	$1.9\times10^{+5}$ [s]	

Menter, F. R., 1994, Tow-Equation Eddy-Viscosity Turbulence Model for Engineering Applications, *AIAA JOURNAL*, **32**-8, 1598-1605.

Tomioka, K., Tanaka, K., Nagayama, K., Tokuda, K., 2005. Prediction method of heat generation and heat transfer in an oil-hydraulic system, *Proc. of IMAACA2005*, 89-94.

TAIYO, LTD, 2005, Products, Hydraulic Equipment, *http://www.taiyo-ltd.co.jp/eg/index.html.*

Yamamoto, K., Tanaka, K., Nakanishi, M., Tarumi, S., 1997, Analysis of Dynamic Behavior of a Hydraulic and Pneumatic Suspension Including Temperature Change Effect, *Proc. of 5th Tri. Int. Symp. on Fluid Control, Measurement and Visualization (FLUCOME97)*, Vol. 1, 457-462.

MODELING AND SIMULATION OF POWER HACKSAW MACHINE USING BOND GRAPH

Anand Vaz[a], Aman Maini[b]

[a] Department of Mechanical Engineering, Dr. B.R. Ambedkar National Institute of Technology, Jalandhar
[b] Department of Mechanical Engineering, DAV Institute of Engineering & Technology, Jalandhar

[a]anandvaz@ieee.org, [b]er_maini@yahoo.co.in

ABSTRACT

The mechanism of a machine involves a number of interconnected rigid bodies which interact with each other dynamically. Formulations of dynamics for such mechanisms available in literature, usually either do not reveal the physical aspects involved, or tend to be too mathematically inclined. Aspects of cause and effect are also obscured. In this paper, the multibond graph approach is used as an alternative but powerful method, for modeling the dynamics of mechanisms taking the specific example of a power hacksaw machine, which is quite common in industry. The mechanism is modeled as a combination of rigid links suitably constrained at joints. Bond graph offers the capability for pictorial representation of the dynamics of mechanisms, based on interaction of power between elements, and also depicts the cause and effect relationships. Coding for simulation can be done directly from the multibond graph without deriving system equations. In this work, the program coding is implemented in MATLAB. The simulation results are analyzed, plotted and discussed.

Keywords: Modeling, simulation, power hacksaw mechanism and Bond Graph.

1. INTRODUCTION

The mechanism of a machine involves a number of interconnected rigid bodies which interact with each other dynamically. Kinematics of such mechanisms is presented in most books on machines and mechanisms but the dynamics is rarely available. Modeling the dynamics of mechanisms is also necessary from the point of view of design, analysis and control. The dynamics of such mechanisms can be described using a number of formalized procedures like the Newton-Euler, Lagrange-Euler, Kane's dynamics etc (Fu, Gonzalez and Lee 1986; Murray, Li, and Sastry 1993). These formulations usually either do not reveal the physical aspects involved, or tend to be too mathematically inclined. Aspects of cause and effect are also obscured. The technique of Bond graph modeling successfully overcomes the difficulties encountered, and also provides several significant advantages not provided by the conventional formulation techniques (Karnopp, Margolis, and Rosenberg 2000). The Bond Graph technique has become widely accepted and several software tools are now commercially available for Bond graph modeling and simulation.

The Bond graph technique is especially convenient and computationally advantageous for dynamic analysis of mechanisms. Bond graph is the pictorial representation of the dynamics of the system and is convenient for modeling of physical system dynamics in multiple energy domains. It offers a unified framework for modeling the mechanism, and, the actuation and control systems due to its capability of handling multi-energy domains. The application of this technique results in a number of advantages. In addition, depiction of causality in Bond graph shows the *cause* and *effect* relationship between the *flow* and *effort* elements of the bond. The notion of causality, apart from aiding with the formulation of system equations which govern the behavior of the system, help in pointing out any physical impossibility or system property that may have failed to be taken into account at the modeling stage. Software are commercially available for multibody dynamics simulation. However, these do not offer sufficient freedom to the user to incorporate modeling of additional physical phenomena, actuator dynamics, fault diagnostics, or scope for study of *cause* and *effect*.

Multibond graph modeling of rigid body based mechanisms have been discussed in (Bos and Tiernego 1985; Tiernego and Bos 1985). Formulation of rotational dynamics in the inertial frame using multibond graphs has been presented in (Vaz and Hirai 2004; Vaz and Thommen 2010), and the computational advantages elaborated in (Vaz 2008). Ersal et al. (2004) have discussed multibond graph based formulations as applied to machine tool mechanisms.

In this paper, the multibond graph approach is used as an alternative but powerful method, for modeling the dynamics of mechanisms taking the specific example of a power hacksaw machine, which is quite common in industry (Norton 1996). The mechanism of power hacksaw machine offers an opportunity to demonstrate modeling of revolute as well as prismatic joints, and also the interface of the machine with work piece during the cutting operation. The power hacksaw mechanism is modeled as a combination of five rigid links suitably

constrained at joints, having relative movement with each other.

Reference frames are fixed on each link using the Denavit-Hartenberg convention (Craig 2005). Usually, the translational dynamics of a link is considered in the inertial frame and rotational dynamics is expressed in the moving frame fixed on the moving link. In this paper, both the translational and rotational dynamics are expressed in the inertial frame, as discussed in (Vaz and Hirai 2004, Vaz 2008). The computational advantages due to this formulation, therefore, have not been discussed again in this paper.

The problem of differential causality arises due to constraints imposed at joints; suitable stiffness and damping elements are added for rectifying this problem and converting the model into one with complete integral causality. These elements make the model more realistic by bringing in the effects of compliance and dissipation at the joints, within definable tolerance limits. At the interface between the cutting blade and work-piece, the resistive forces are different during forward stroke and cutting stroke which is in the backward direction.

Coding for simulation is done directly from the multibond graph without deriving system equations. In this work, the program coding is implemented using MATLAB (2010). Important dynamic quantities like reaction forces and torques at various joints on the system can be determined, plotted and analyzed. The simulation results are analysed, plotted and discussed.

2. MODELING

In this work, a multibond graph model of power hacksaw machine is developed. The mechanism is considered as a combination of five rigid links connected at joints. There is relative movement of links with each other and also with inertial frame. Frame 1 which is a fixed frame, is taken as inertial frame. The 0 frame and 1 frame coincide with each other. This inertial frame does not possess any translational or rotational motion. The power hacksaw machine is used for metal cutting in industries. The mechanism involves five rigid links; link 1 is taken as the fixed frame, link 2 is the crank which is connected to connecting rod (link 3). The connecting rod is connected to the frame (link 4), in which the cutting blade is adjusted. The link 4 has sliding motion in guide link 5. The link 5 has revolute movement about origin O_5. The origin is assigned as O_1, O_2 and O_5 to link 1, link 2 and link 5. The mass of the link 5 is more than link 4. The link 5 applies force due to gravity on link 4 resulting in downward movement of link 4, so that during cutting operation the cut may proceed in the downward direction. C_1, C_2, C_3, C_4 and C_5 are center of masses of respective links.
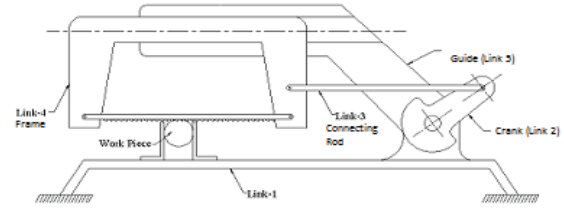
Fig. 1: The power Hacksaw machine

The dynamics of the system is modeled using multibond graph. The bond graph of the system is shown in Fig. 3. The thick lines in bond graph represent the multibonds while the thin lines show scalar bonds. The left side of the bond graph is associated with the rotational dynamics of the model while the right side shows the translational dynamics. The translational motion of the crank is restricted by applying S_f equal to zero between O_1 and O_2; the origin of fixed frame and that of frame 2 respectively. The relative movement between links at each origin is distinguished such as the point A_2 on link 2 and point A_3 on link 3. Point B_3 on link 3 and B_4 on link 4.
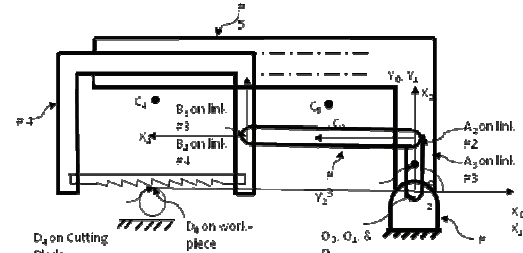

Fig. 2: The Power Hacksaw Machine (link diagram)

To eliminate the differential causality, viscoelastic elements introducing stiffness and damping are provided at revolute joints such as at the joint of frame and crank (link 1 and link 2) stiffness $C: k_{2r}$ and damping $R: r_{2r}$ at the joint of crank and connecting rod the stiffness $C: k_{23r}$ and damping $R: r_{23r}$ and at the joint of connecting rod and frame 4 the stiffness $C: k_{34r}$ and $R: r_{34r}$ are applied.

For simulation, a torque of 10 $N \cdot m$ is applied at the crank. The crank is made to rotate about point O_1 (about Z-axis). The movement of the crank in other two directions (X and Y axis) is constrained by applying S_f equal to 0. As the link 4 slides in the guide-way of link 5 so the relative angular movement between these two links is constrained equal to zero by equating S_f to 0. At the interface of cutting blade and work-piece, point D_4 is taken at cutting blade (i.e. blade is fixed in frame 4) while point D_0 is considered at work-piece. The distance of D_4 from O_1 remains fixed. Due to the reciprocating movement of frame 4, the relative distance between center of mass (C_4) of link 4 and point D_4 changes during cutting operation. The blade slides in X-direction of frame 4 only, so the movement of blade in other two directions is constrained.

$$\sideset{_0^0}{}{\overline{r}_{D_0}} = \sideset{_{D_4}^0}{}{\overline{r}_{D_0}} + \sideset{_{C_4}^0}{}{\overline{r}_{D_4}} + \sideset{_0^0}{}{\overline{r}_{C_4}} \tag{1}$$

Differentiating w.r.t. time *t*,

$$\sideset{_0^0}{}{\dot{\overline{r}}_{D_4}} = \sideset{_0^0}{}{\dot{\overline{r}}_{C_4}} + \left[\sideset{_0^0}{}{\overline{\omega}_4} \times \right] \sideset{_{C_4}^0}{}{\overline{r}_{D_4}} + \sideset{_4^0}{}{R} \sideset{_{C_4}^4}{}{\dot{\overline{r}}_{D_4}} \tag{2}$$

which can be written as,

$$\sideset{_0^0}{}{\dot{\overline{r}}_{D_4}} = \sideset{_0^0}{}{\dot{\overline{r}}_{C_4}} - \left[\sideset{_{C_4}^0}{}{\overline{r}_{D_4}} \times \right] \sideset{_0^0}{}{\overline{\omega}_4} + \sideset{_4^0}{}{R} \sideset{_{C_4}^4}{}{\dot{\overline{r}}_{D_4}} \tag{3}$$

$\sideset{_0^0}{}{\overline{r}_{D_0}}$ is the position vector from Origin O_1 to point D_0 expressed in frame *0*. $\sideset{_{D_4}^0}{}{\overline{r}_{D_0}}$ is the position of the interface of cutting blade and work-piece from point C_4. $\sideset{_{C_4}^0}{}{\overline{r}_{D_4}}$ is the position vector from center of mass of link *4* and cutting point at blade. This distance changes during cutting operation. $\sideset{_0^0}{}{\overline{r}_{C_4}}$ is the distance of center of mass of link *4* from origin O_1. $\sideset{_0^0}{}{\overline{\omega}_4}$ is the angular velocity of blade frame *4*, w.r.t. frame *0* and is expressed in frame *0*. The Bond graph of Fig. 3 incorporates the kinematics of (3).

The translational effect is concentrated at the center of mass of each link, while the rotational effect is considered in the inertial frame itself by considering the inertia tensor for each link about its respective center of mass and expressed in the inertial frame. This implies that the inertia tensor changes continuously with orientation. However, this aspect has been taken into consideration, and requires lesser number of computations in terms of multiplications and additions (Vaz 2008).

3. SIMULATION RESULTS
The simulation has been performed in MATLAB. The parameters used are shown in Table 1. Table 2 represents the values of stiffness and damping at different couplings.

Table-1: Link properties used for simulation

Crank (Link *2*)			
	$\sideset{^2}{}{l_x} = 0.075\ m$	$\sideset{^2}{}{l_y} = 0.04\ m$	$\sideset{^2}{}{l_z} = 0.02\ m$
Connecting Rod (Link 3)			
	$\sideset{^3}{}{l_x} = 0.7\ m$	$\sideset{^3}{}{l_y} = 0.04\ m$	$\sideset{^3}{}{l_y} = 0.03\ m$
Frame (Link 4)			
	$a_{41} = 0.03\ m$	$b_{41} = 0.13\ m$	$c_{41} = 0.03\ m$
	$a_{42} = 0.03\ m$	$b_{42} = 0.13\ m$	$c_{42} = 0.03\ m$
	$a_{43} = 0.4\ m$	$b_{43} = 0.03\ m$	$c_{43} = 0.03\ m$
Guide (Link 5)			
	$a_{51} = 0.1\ m$	$b_{51} = 0.09\ m$	$c_{51} = 0.03\ m$
	$a_{52} = 0.09\ m$	$b_{52} = 0.7\ m$	$c_{52} = 0.03\ m$

Table- 2: Stiffness and damping of various couplings for simulation.

Location	Stiffness (K)	Damping (R)
Translational coupling between		
fixed frame and crank	$k_2 = 10^7\ N/m$	$r_2 = 0.5\ N\text{-}s/m$
crank and connecting rod	$k_{23} = 10^7\ N/m$	$r_{23} = 10^2\ N\text{-}s/m$
connecting rod and frame 4	$k_{34} = 10^8\ N/m$	$r_{34} = 0.5\ N\text{-}s/m$
fixed frame and guide 5	$k_{15} = 10^6\ N/m$	$r_{15} = 10\ N\text{-}s/m$
Rotational coupling between		
fixed frame and crank	$k_{2rx} = 10^3$ *N.m/rad* $k_{2ry} = 10^3$ *N.m/rad*	$r_{2rx} = 10^2$ *N.m.s/rad* $r_{2ry} = 10^2$ *N.m.s/rad*
crank and connecting rod	$k_{23rx} = 10^3$ *N.m/rad* $k_{23ry} = 10^3$ *N.m/rad*	$r_{23rx} = 10^2$ *N.m.s/rad* $r_{23ry} = 10^2$ *N.m.s/rad*
connecting rod and frame 4	$k_{34rx} = 10^3$ *N.m/rad* $k_{34ry} = 10^3$ *N.m/rad*	$r_{34rx} = 80$ *N.m.s/rad* $r_{34ry} = 80$ *N.m.s/rad*
frame 4 and Guide 5	$k_{45r} = 10^4$ *N.m/rad*	$r_{45r} = 100$ *N.m.s/rad*
Resistive Forces between cutting blade and work-piece.	$k_{Rfy} = 10^2\ N/m$ $k_{Rfz} = 10^2\ N/m$	$r_{Rfx} = 10^2$ *N.m.s/rad* $r_{Rfy} = 10^2$ *N.m.s/rad* $r_{Rfz} = 10^2$ *N.m.s/rad*

3.1. Simulation graphs
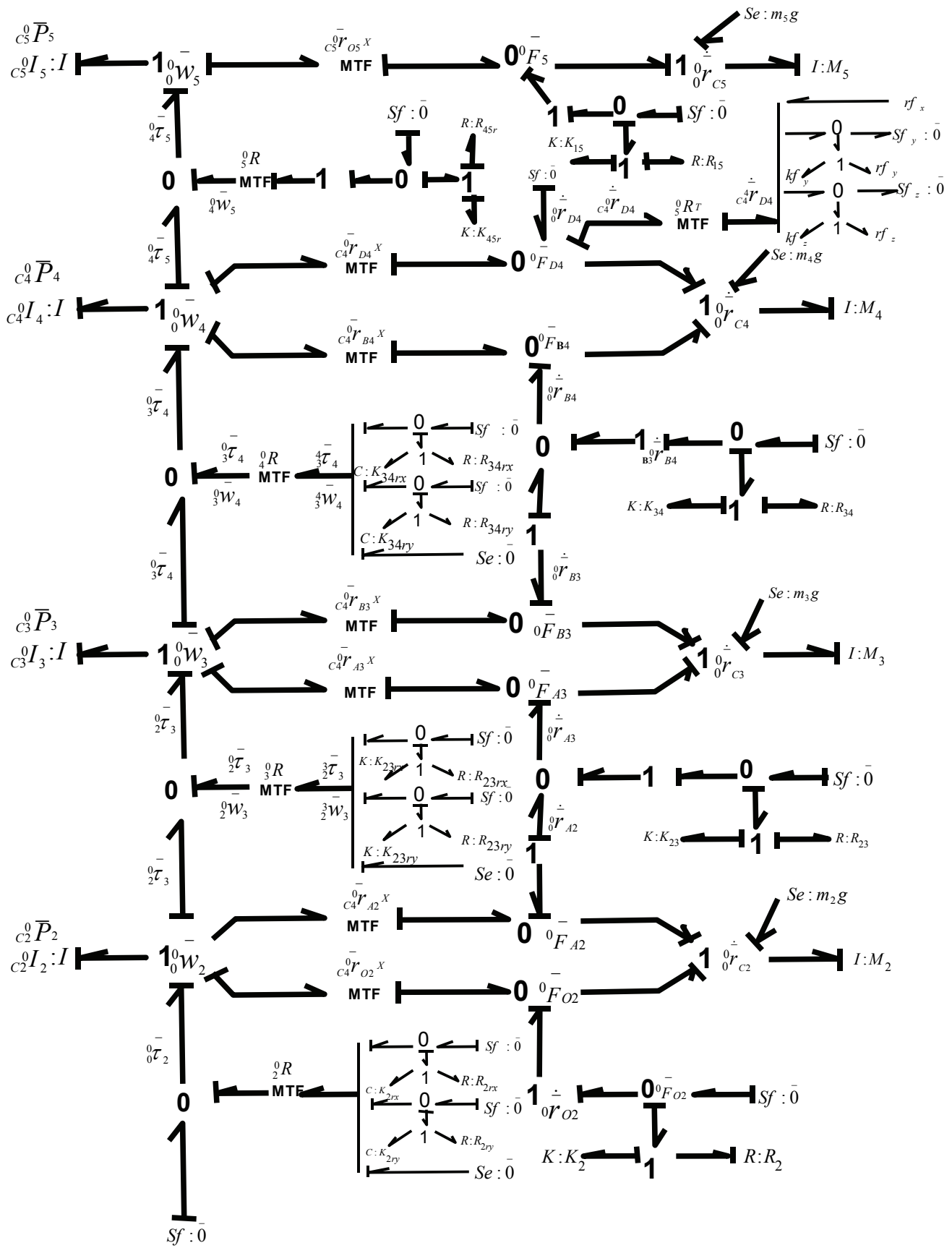The simulation plots for various links are shown in this section. The time span for simulation is taken as 5 seconds.

Fig.3: Bond Graph of Power Hacksaw Machine

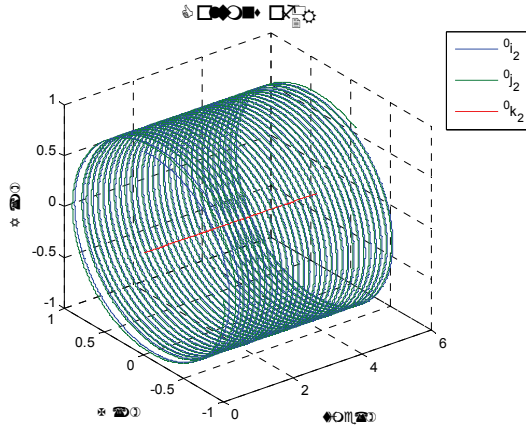### 3.1.1 Dynamics of the crank

Fig. 4: Variation of orientation matrix of link *2* with time.

The link *2*, that is the crank, rotates about origin O_2. Hence in Fig. *4*, the unit vectors in *x* and *y* direction describe circular paths while that in the *z* direction is stationary.
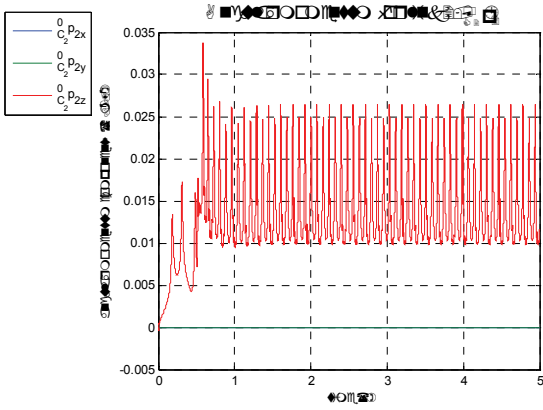

Fig. 5: Variation of angular momentum of link *2* with time.

In Fig. *5,* only the *Z* component of angular momentum varies. This is due to the constant torque applied at the crank.
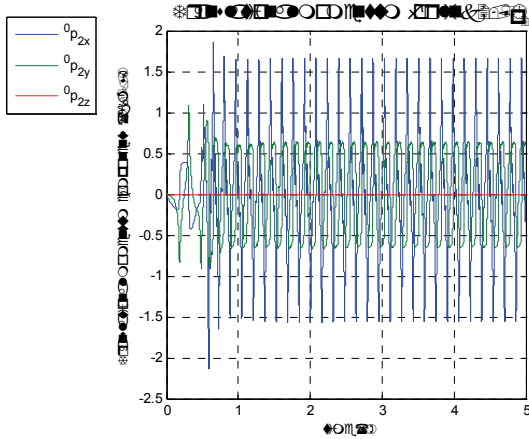
Fig. 6: Variation of Translational momentum of link *2* with time.
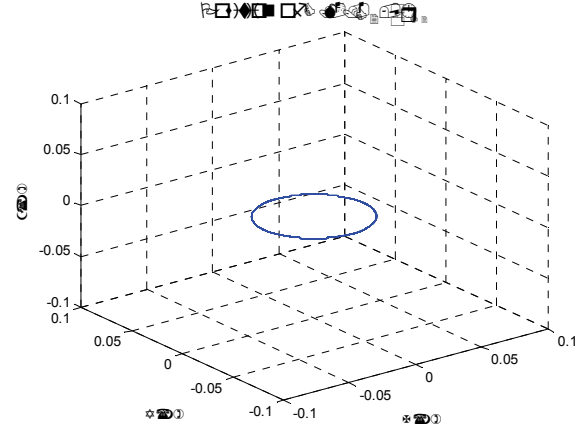
Fig. 7: Position of center of mass of link *2*.

The crank rotates at origin O_2 about *z*-axis. So the path traced by center of mass C_2 of the crank is a circle. The Fig. 7 shows the position trajectory of center of mass of crank.

3.1.2 Dynamics of connecting rod
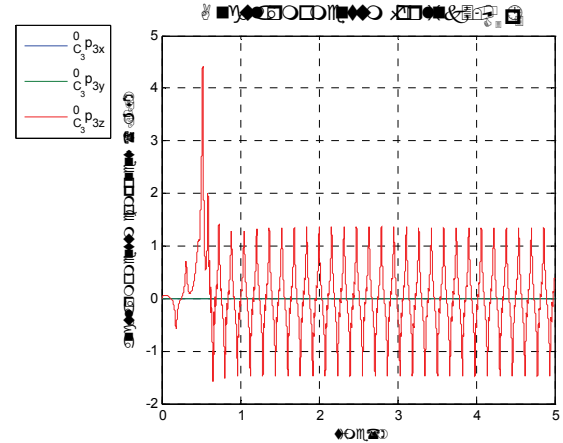
Fig. 8 : Variation of angular momentum of connecting rod with time.

The link *3*, which is connecting rod, is the link between frame *4* and crank *2*. The variation of its angular momentum with time is shown in Fig. *8*. The angular momentum increases due to the constant torque applied on crank.
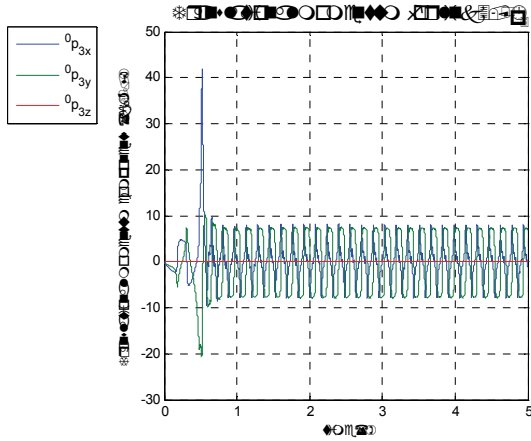
Fig. 9: Variation of Translational momentum of link *3* with time.
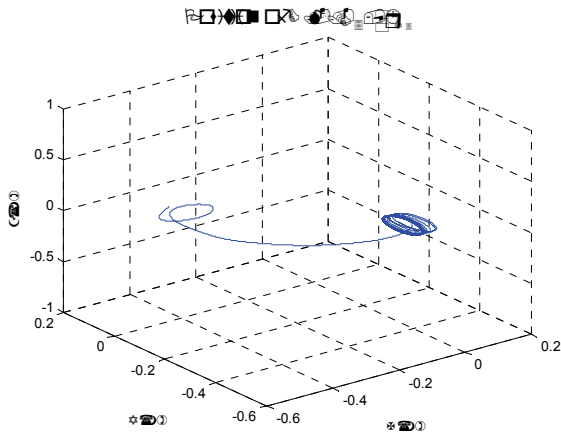
Fig. 10: Position of center of mass of link *3*.

The connecting rod oscillates and also moves up and down. The movement of the center of mass of link *3*, C_3, is represented by Fig.*10*.
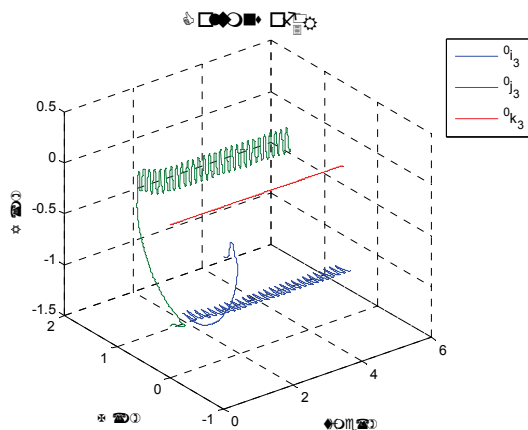
Fig. 11: Variation of orientation matrix of link *3* with time.

The movement of the connecting rod is only in *x* and *y* directions, but in *z* direction the position does not change. The connecting rod oscillates about *z* - axis. This variation is shown in graph Fig.11.

3.1.3 Dynamics of Frame

The link *4* is the frame in which the cutting blade is adjusted. The frame slides in guide ways of link *5*, in forward and backward direction and provides movement to the cutting blade for cutting operation.
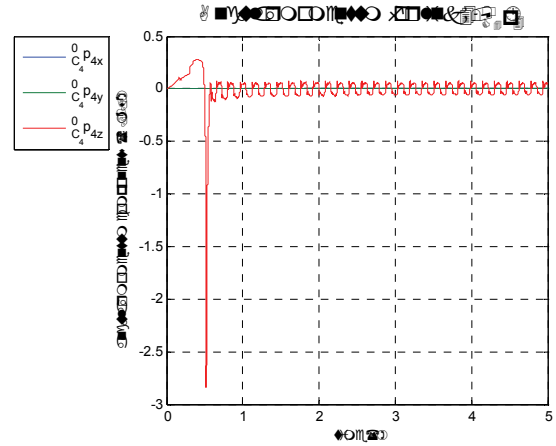
Fig. 12: Variation of angular momentum of link *4* with time.

The angular momentum changes with time at link *4*. This change is shown in Fig. *12*.
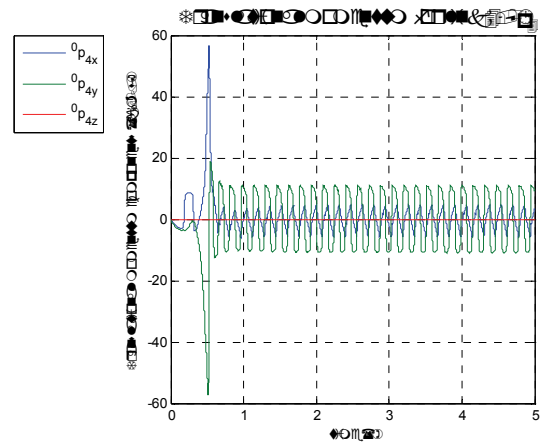
Fig. 13: Variation of Translational momentum of link *4* with time.
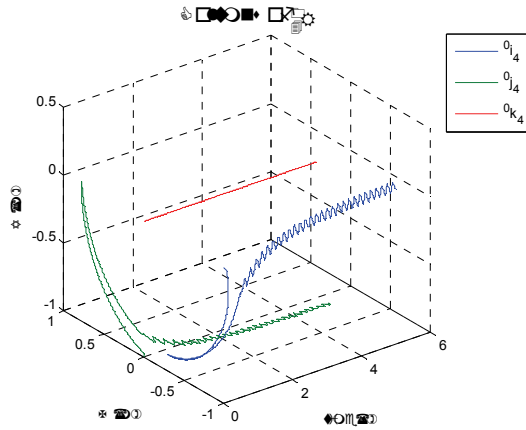
Fig. 14: Variation of orientation matrix of link *4* with time.

The frame *4* has sliding movement in the guide ways (link *5*) in *x* direction (forward and backward) and due to downward movement of link *5*, the frame *4* also moves downward. So the variation in *x* and *y* direction is shown in graph while the orientation in *z* direction does not change so the graph shows the *z* direction as a straight line.
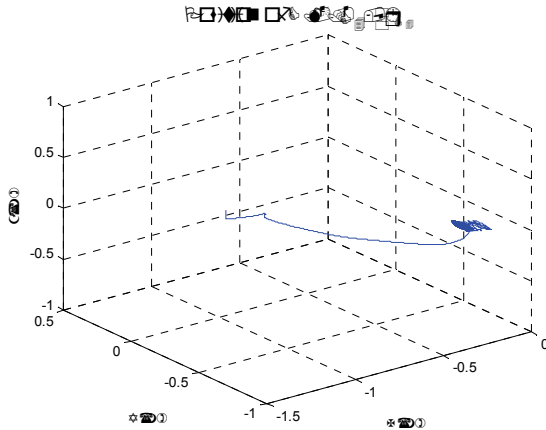
Figure 15: Position of center of mass of link *4*.

The movement of the frame *4* is changing in *x* and *y* direction and the position of its center of mass changes accordingly as shown in Fig. 15.

3.1.4 Dynamics of Guide-way (link 5)

The guide way provides downward movement to frame *4*, due to gravitational force. A slot is designed in this link for providing sliding movement to link *4*.
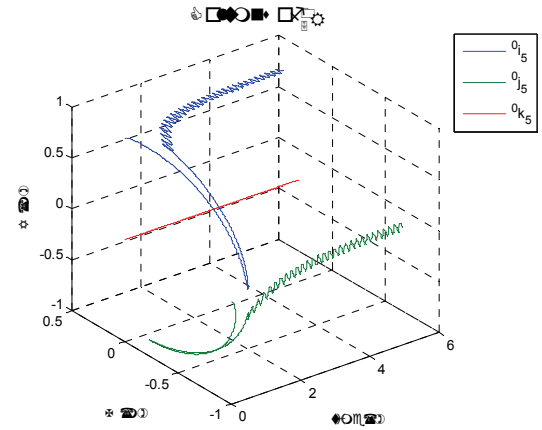
Fig. 16: Variation of orientation matrix of link *5* with time.
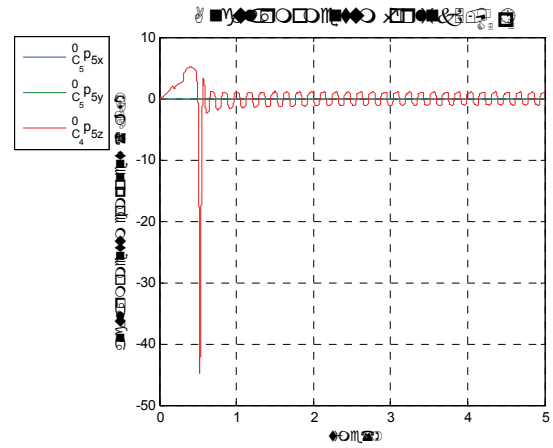
Fig. 17: Variation of angular momentum of link *5* with time.

The angular momentum of link *5* is represented in Fig. 17. The angular momentum changes sharply in the initial period due to high starting torque which is suddenly.
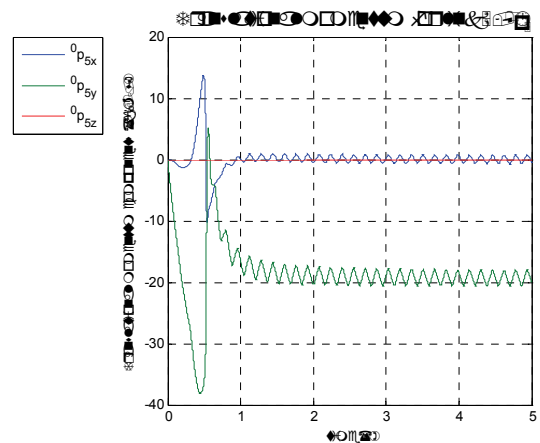
Fig. 18: Variation of Translational momentum of link *5* with time.

The position trajectory of the center of mass of the guide way link 5 is shown in Fig.19.
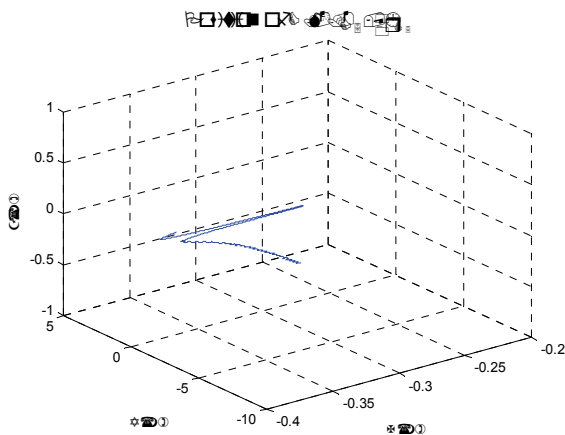
Figure 19: Position of center of mass of link *5*.

4. Conclusions

Dynamics of the power hacksaw machine has been modeled in this paper using bond graphs. The dynamics of the system is simulated in MATLAB, after coding directly from the Bond graph model, without deriving system equations. Simulation results agree with expected behavior. The mechanism of power hacksaw machine offers an opportunity to demonstrate multibond graph modeling of revolute as well as prismatic joints, and also the interface of the machine with work piece during the cutting operation.

REFERENCES

Bos, A. M., and Tiernego, M. J. L., 1985. Formula Manipulation in the Bond Graph Modelling and Simulation of Large Mechanical Systems," *Journal of The Franklin Institute*, 319(1/2), 51-65.

Craig, J. J., 2005. *Introduction to Robotics: Mechanics and Control*, 3rd ed. Pearson Education Inc.

Ersal, T., Stein, J. L., and Louca, L. S., 2004. A Bond Graph Based Modular Modeling Approach towards an Automated Modeling Environment for Reconfigurable Machine Tools. *The International Conference on Integrated Modeling & Analysis in Applied Control & Automation*. Paper ID IMAACA BG-06, October 28-31, Genoa, Italy.

Fu, K. S., Gonzalez, R. C., and Lee, C. S., 1987. *Robotics: Control, sensing, vision and intelligence*. McGraw Hill Book Publishing Company.

Jang, J., and Han, C., 1998. Proposition of a Modeling Method for Constrained Mechanical Systems Based on the Vector Bond Graph. *J. Franklin Institute*. 335B (3), 451-469.

Karnopp, D. C., Rosenberg, R. C., Margolis, D. L., 1990. *System Dynamics: A Unified Approach.* New York: John-Wiley and Sons Inc.

Karnopp, D. C., 1997. Understanding Multibody Dynamics Using Bond Graph representations. *J. Franklin Institute*, 334B (4), 641-642.

MATLAB, 2010. Available from: http://www.mathworks.com [accessed March, 2011]

Mukherjee, A., and Karmakar, R., 2000. *Modeling and Simulation of Engineering System through Bond Graphs*. New Delhi: Narosa Publishing House.

Murray, R. M., Li, Z., and Sastry, S. S., 1993. *A Mathematical Introduction to Robotic Manipulation*. Florida: CRC Press.

Tiernego, M. J. L., and Bos, A. M., 1985. Modelling the dynamics and Kinematics of Mechanical Systems with Multibond Graphs. *Journal of The Franklin Institute*, 319(1/2), 37-50.

Vaz, A., 2003. *Lecture Notes: Bond graph modeling for Rigid Body Dynamics*. NIT Jalandhar, India.

Vaz, A., 2004. Modeling a Hand Prosthesis with Word Bond Graph Objects. *The International Conference on Integrated Modeling & Analysis in Applied Control & Automation*. Paper ID IMAACA BG-08, October 28-31, Genoa, Italy.

Vaz, A., 2008. A Bond Graph Perspective of Computational issues in Multibody Mechanics. *Proceedings of National Conference on Mechanisms, Science and Technology from Theory to Application*, Nov. 13-14, Hamirpur, India.

Vaz, A., Sharma, P., and Atri, R., 2009. Modeling and Simulation of the Dynamics of Crankshaft-Connecting Rod-Piston-Cylinder Mechanism and a Universal Joint Using The Bond Graph Approach. 14*th National Conference on Machines and Mechanisms (NaCoMM09)*, December 17-18, NIT, Durgapur, India.

Vaz, A., and Thommen G. K., 2010. Modeling and Simulation of dynamics of the quick return Mechanisms: A Bond Graph Approach. *10th National Conference on Industrial Problems on Machines and Mechanisms (IPRoMM)*, December.

OPTIMIZETION OF ACTUATING SYSTEM FOR FLAPPING ROBOT DESIGN

T. Imura[a], M. Fuchiwaki[b], K. Tanaka[b]

[a] *Graduate School of Computer Science and Systems Engineering Kyushu Institute of Technology*
680-4 Kawazu, Iizuka-cit, Fukuoka 820-8502, Japan
Research Fellow of the Japan Society for the Promotion of Science
[b] *Department of Mechanical Information Science and Technology Kyushu Institute of Technology*
680-4 Kawazu, Iizuka-city, Fukuoka 820-8502, Japan

[a]imura@vortex.mse.kyutech.ac.jp , [b]futiwaki@mse.kyutech.ac.jp , [c]kazuhiro@mse.kyutech.ac.jp

ABSTRACT

In this paper, my purpose is suggestion for design method for actuating system of flapping robot. It is necessary to develop that. Considering the drag of flapping wings and the characteristic of motor revolution for its load, the actuating system of flapping robot is modeled. The motor whose resultant flapping frequency and wing loading required for flapping robot is selected, and flapping robot is developed by that result. As the result, the flapping robot can fly stably for about 15 minutes. Moreover, the design method is effective in case of attaching important to the increase in flight altitude by clarifying that flapping frequency play an important role for control of flight altitude.

Keywords: flapping robot design, tailless airplane, actuating system modeling, optimizing actuating system

1. INTRODUCTION

1990s, Defense Advanced Research Projects Agency (DARPA) led developing small airplane called Micro air vehicles (MAVs) with a maximum dimension of 15 [cm] or less and a flight speed of 10 [m/s]. MAVs are expected to both military and civilian applications, especially surveillance and reconnaissance at a remote or otherwise hazardous location and at the other planets. As rapid advance in power source miniaturization, structures, materials, control device and communication technology, many researchers have developed many MAVs of fixed wing, rotation wing and flapping wing (Yongsheng L., Wei Shyy, D. Viieru and B. Zhang 2003; Robert C. Michelson 2008; Zaeem A .Khan and Sunil K. Agrawal 2006).

As the result of evolution of life for several 100 millions of years, insects, birds and bats are recognized as the most effective flying object. In nature flying objects, the flight systems such as wing structure, wing kinematics, and flight controlling are optimized and harmonized well. It has been clarified that insects and birds fly with many mechanism, i.e. delayed stall, rotational circulation, wake capture, clap and fling and so on (F. O. Lehmann 2004; M.H.Dickinson, F.O.Lehman, S.P.Sane 1999; T.Maxworthy 1979). Many researchers have taken notice and reported for the force generation mechanism on the flights of insects and birds and so on.

In many kinds of insect which is taken notice by them, it is well known that a butterfly combines flapping wing motion with gliding and the figure of flying is very beautiful. Moreover, butterflies do not perform linearly flying motions similar to dragonfly, but they perform motion like dancing with low flapping frequency. Sunada et al. observed Pieris melete on a taking-off flight and reported on characteristics of behaviors of a butterfly wings, its center of gravity and its abdomen and so on (S. Sunada, K. Kawachi, I. Watanabe, and A. Azuma 1993).

The concept called Biomimetics by which nature creatures is mimicked has been taken notice. Based on this concept, a lot of flapping robots to which the insect and bird flight mechanisms, i.e. Flapping resonance, Traverse bending, Clap and fling and so on, were applied have reported (Robert J. Wood 2008; Zaeem A .Khan and Sunil K. Agrawal 2006; De Croon G.C.H.E., de Clerq K.M.E., Ruijsink R., Remes B., and de Wagter C. 2009). However, almost of that have taken notice on whether robot can fly or not, and unsteady aerodynamic forces generated by wing flapping. From the result, it is the present situation that the reports which take notice on design and developing method of flapping robot are a few.

In this paper, my purpose is suggestion for method of optimization for actuating system of flapping robot. It is necessary to develop that. Concretely speaking, considering the drag of flapping wings and the characteristic of motor revolution for its load, optimized motor and reduction ratio of actuating system for flapping robot is selected. Moreover, flapping robot is developed by that result, and its flight is evaluated by flight observation experiment.

2. MODELIZATION OF FLAPPING SYSTEM
2.1. Modeling of actuating system and wings

In this chapter, the actuating system of flapping robot is modeled due to estimate flapping frequency from motor specification. Fig. 1 shows a pattern diagram of actuating system for flapping. Flapping motion is powered by an electric motor through a

gearbox to reduce rotational speed and load at motor. A crank rotation conducted from a gearbox is transformed to linear motion with linkage mechanism. The linear motion actuates the wings which are made of thin Japanese paper and light carbon rods, and flapping robot performs flapping motion. This flapping robot operates flapping motion as wing deforms because the wing membrane is consist of Japanese paper. However it is expected that each value of wing changes by flapping because of wing tenderness, the author models with 1 period averages, not considering dynamic effects, due to design actuating system.

The revolution of motor ω_M [Hz] changes lineally for the load of motor. Therefore, ω_M is shown by Eq. 1. In the Eq. 1, k_t [Hz/(N*m)], N [Hz] and T [N*m] show the gradient of revolution for load at motor, revolution at no load and torque (load) at motor, respectively. In this way, a motor part of actuating system is modeled.

$$\omega_M = N - k_t T \tag{1}$$

The load (torque) generated at motor is shown simply by Eq. 2. In the Eq. 2, T_ω [N*m] and T_o [N*m] show load (near to drag) by wing flapping and load by such as friction except the load by flapping motion at whole system, respectively. G shows reduction ratio, and T [N*m] inversely relates to G.

$$T = \frac{T_\omega + T_0}{G} \tag{2}$$

Considering the inverse relation between ω_M and G, the revolution of gear connected to the wings through linkages (equal to flapping frequency) ω is shown by Eq.3 which is obtained from Eq.1 and 2. Eq.3 is function which shows relation between reduction ratio, torque, motor characteristics and flapping frequency, respectively. In this way, motor and gearbox parts of actuating system are modeled. Then, main motor load by wing drag generated with flapping motion is conducted.

$$\omega = \frac{NG - k_t(T_\omega + T_0)}{G^2} \tag{3}$$

Fig.2 shows a pattern diagram for calculating wing drag. The wing velocity changes for the position on the wing span direction, because the wings perform flapping motion which is similar to rotating. In this paper, the wing velocity is approximated to $r\omega$, simply. a [m], b [m] and r [m] show wing span length, wing chord length and position for wing span direction, respectively. The partial torque (load) fom r of wing position dT_ω [N*m] shows multiplication of partial drag at r of position dD [N] and r [m] (Eq.4). The partial torque generated at basis of wing $dT\omega$ from the partial area of wing dS is shown by Eq.5. In the Eq.5, C_D, ρ and U show drag coefficient of wing, density, flapping velocity at dS.
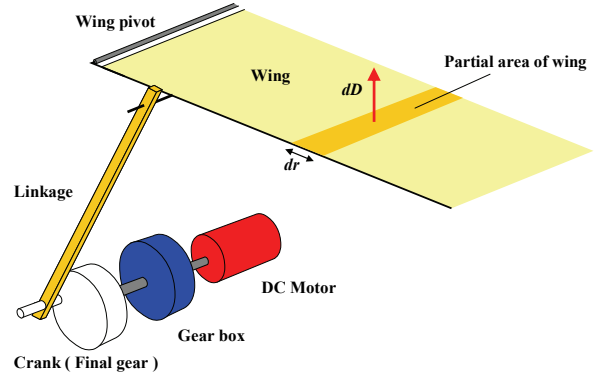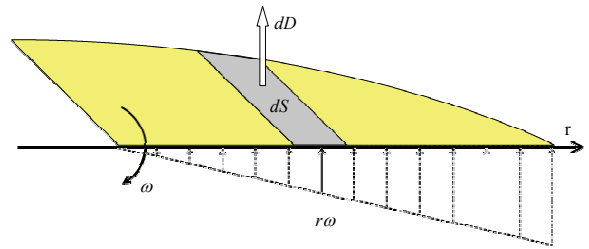

Figure 1: Pattern Diagram of Flapping Actuating System

Figure 2: Pattern Diagram for Calculating Wing Drag
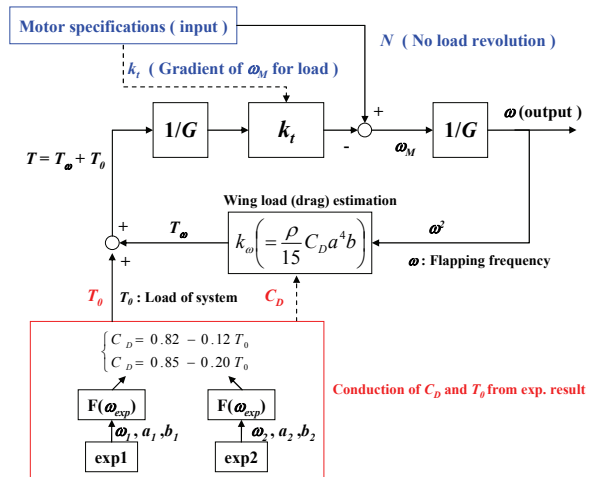

Figure 3: Model for Estimating Flapping Frequency

$$dT_W = rdD \tag{4}$$

$$dT_W = r\frac{\rho}{2}C_D U_{(r)}^2 dS \tag{5}$$

Considering two factors which are wing shape is a quarter of ellipse and U shows $r\omega$, the load (torque) from wings is shown by Eq.6.

$$T_\omega = \int_0^a r\frac{\rho}{2}C_D(r\omega)^2 b\sqrt{1 - \frac{r^2}{a^2}}dr$$
$$= \frac{1}{15}\rho C_D a^4 b\omega^2 \tag{6}$$

Flapping frequency ω is shown by Eq.7 from Eq.3 and 6. In this way, the model of actuating system considering the load at motor by wing drag with flapping is constructed.

This model is assumed that drag coefficient which is affect by wing deformation is constant because this model is recommended for using and calculation 1 period average. Moreover, this model is assumed that wings have uniform motion at mean velocity although actual wing velocity always oscillates. As the result, dynamic oscillation of wing deformation and wing velocity is not modeled in this model. On the other word, this model is nearly equal to a static model which motor revolution is conduct from torque which generates on the fixed rigid airfoil in different flow velocity distribution at wing tip and base of wing. For this reason, this model is not suitable for calculating dynamic values as wing flaps. Therefore, 1 period average is applied for estimating flapping frequency in this model.

Moreover, in this model, it seems to consider that the error between estimating value and real one when flapping frequency is very different from measuring value due to not model wing deformation as it changes by flapping frequency.

$$\omega = \frac{-G^2 + \sqrt{G^4 + 8k_T k_\omega NG + 8k_T^2 k_\omega T_0}}{4k_T k_\omega} \qquad (7)$$

$$k_w = \frac{\rho}{15} C_D a^4 b$$

Fig.3 shows a block diagram of this model. The input of this model is no load revolution N [Hz] and gradient of revolution for load at motor k_t [Hz/(N*m)], and the output is flapping frequency ω [Hz]. Therefore, it is possible that a suitable motor is selected for flapping robot by estimating flapping frequencies for each motor.

However flapping frequency can be estimated with inputting motor specifications (k_t [Hz/(N*m)] and N [Hz]), load at motor T [N*m] and reduction ratio G, the wing load with flapping is proportional to the square of flapping frequency ω^2. As the result, this block diagram has closed-loop. Because it is difficult to conduct the wing drag coefficient C_D and load except the load by flapping motion at whole system T_0, they are conducted by experimental results. In the chapter 2.2, C_D and T_0 are conducted from the result of experiment and Eq.7.

2.2. Derivation of coefficient value

In this chapter, the wing drag coefficient C_D and load at whole system T_0 are conducted by Eq.7 and experiment of measuring flapping frequency. Fig.4 shows the experimental apparatus for flapping frequency measuring. This apparatus consists of flapping robot and high speed camera (1125[fps]). The flapping frequency of flapping robot fixed on a shaft is measured with image measurement. In concrete terms,
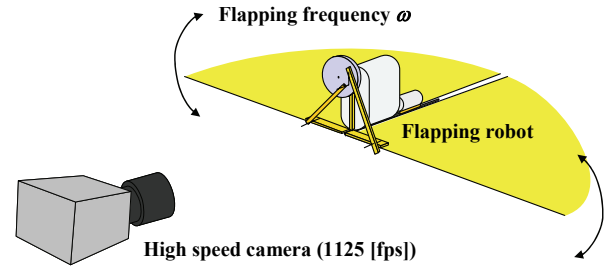

Figure 4: Experimental Apparatus for Measuring Flapping Frequencies

the time span for 10 periods flapping is measured with high speed camera, and the time span required for 1period flapping is calculated from measuring value. Flapping frequency is equal to the reciprocal of this calculating value. The wing drag coefficient is calculated with back calculation from the flapping frequency. Eq.8 shows a function for back calculation and is conducted from Eq.7. Because the wing drag coefficient C_D with back calculation contains the load at whole system T_0 (Eq.9), the wing drag coefficient C_D and load at whole system T_0 are conducted by measuring flapping frequencies of 2 kinds of wing which have different wing area. In this calculation, the author assumes that the wing drag coefficient C_D is constant for wing area and deformation and load at whole system T_0 is constant for flapping frequency.

$$C_D = \frac{NG - \omega G^2 - k_t T_0}{2k_{\omega 2} k_t \omega^2} \qquad (8)$$

$$k_{\omega 2} = \frac{\rho}{15} a^4 b$$

When a wing areas are 8.2 $*10^{-3}$ [m^2] and 6.5 $*10^{-3}$ [m^2], flapping frequencies are 7.51 [Hz] and 10.04 [Hz], respectively. The drag coefficient C_D of wings which have different wing area and flapping frequency is calculated by Eq.8 and flapping frequencies(Eq.9). As previously described, because the wing drag coefficient C_D with back calculation contains the load at whole system T_0 in mathematical expression which can calculate drag coefficient from flapping frequency with back calculation, the wing drag coefficient C_D = 0.690 and load at whole system T_0 = 0.192 [N*m] is conducted by two equations of Eq.9. The calculated wing drag coefficient C_D is smaller than drag coefficient C_D = about 2.3 (angle of attack of 90 [deg.]) which is estimated by steady aerodynamic force at the rigid wing mimicking wing shape of a kind of hawk moth (F. O. Lehmann 2004). In addition to different wing shape and experimental method between this study and hawk moth result, the wings deform largely with flapping motion because the wing membrane is Japanese paper in this study and its arc parts of wing are free end condition. It seems to consider that the drag coefficients C_D of this study is smaller than that of rigid

wing because of decrease in projected area of wing against wing motion direction by wing deformation.

When flapping frequency is different extremely from that of experimental result, it is expected that actual wing deformation value changes largely from value on the assumption. Not surprisingly, it is expected that the actual drag coefficients C_D is different from that of this study because of assumption which drag coefficient C_D is constant for wing deformation in this model. As the result, it seems that the error of estimated flapping frequency increases. From this result, it is not suitable for estimating flapping frequency for extremely different frequency from around flapping frequency of 10 [Hz].

In chapter 3, the optimized motor and reduction ratio is selected by C_D, T_0 conducted in this chapter and Eq.7, and the actuating system is designed.

$$\begin{cases} C_D = 0.82 - 0.12T_0 \\ C_D = 0.85 - 0.20T_0 \end{cases} \qquad (9)$$

3. DESIGN OF ACTUATING SYSTEM FOR FLAPPING ROBOT

3.1. Required values for flapping robot

In this study, flapping robot applied by butterfly flight is suggested and developed it. Fig.5 shows the relation between wing loading, flapping frequency and wing aspect ratio of Lepidoptera (mainly butterflies and moths) from literatures of butterfly. The horizontal axis shows wing loading, and the left and right vertical axes show flapping frequency and aspect ratio of wing, respectively. The red circles and black circles show that of butterflies and that of moths, respectively. The soild circles and the midair circles show flapping frequency and aspect ratio of wing, respectively. The flapping frequency, wing loading and aspect ratio of butterfly are approximately 7 - 15 [Hz], 0.5 - 1.5 [N/m²] and 1.5 – 3.0. Especially, it is well known that flapping frequency directly affects unsteady aerodynamic forces by wing flapping (Che-Shu Lin, Chyanbin Hwu and Wen-Bin Young 2006), and it is clear from Eq.6 which is used for modelization in this paper. In this paper, the author intends to make flapping frequency and wing loading of flapping robot to approximate to these data.

H. Tanaka et al. clarified that an appropriate position of the center of gravity makes the balance of longitudinal moment, and it leads to stable forward flight. In his experiment, when center of gravity was at $0.34C$ against wing chord length C [m], stable forward flight was realized (H. Tanaka, K. Hoshino, K. Matsumoto, and I. Shimoyama 2005). The flapping robot in this study has also similar position of center of gravity, and its position is $0.33C$.

3.2. Optimization of motor for actuating system of flapping robot

Fig.6 shows flapping frequencies and wing loadings of each motor. The horizontal and vertical axes show

flapping frequency and wing loading, respectively. The wing loading is estimated from total
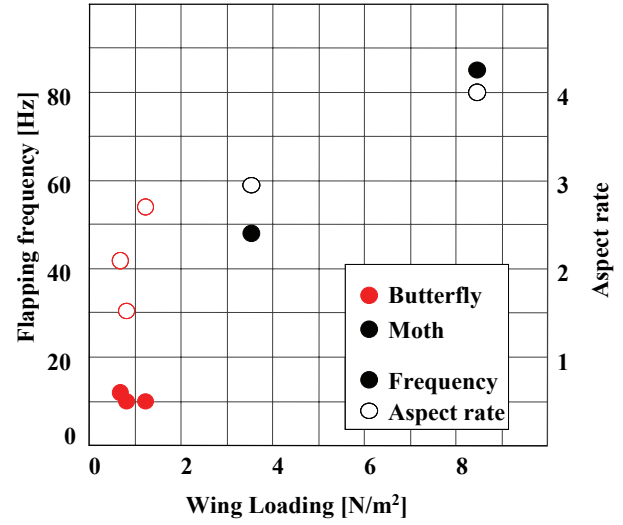

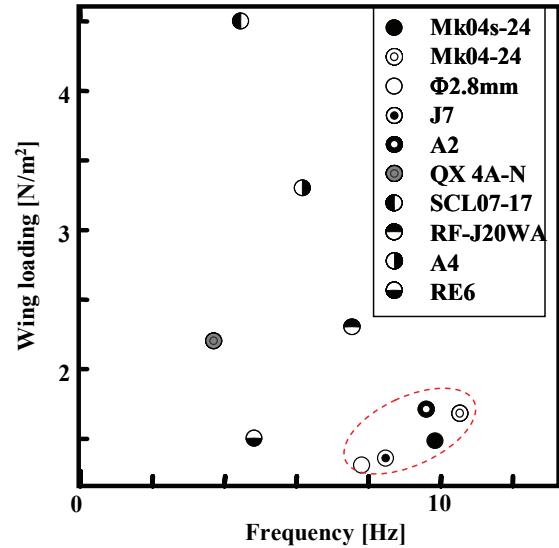Figure 5: Specification of butterflies and Moths
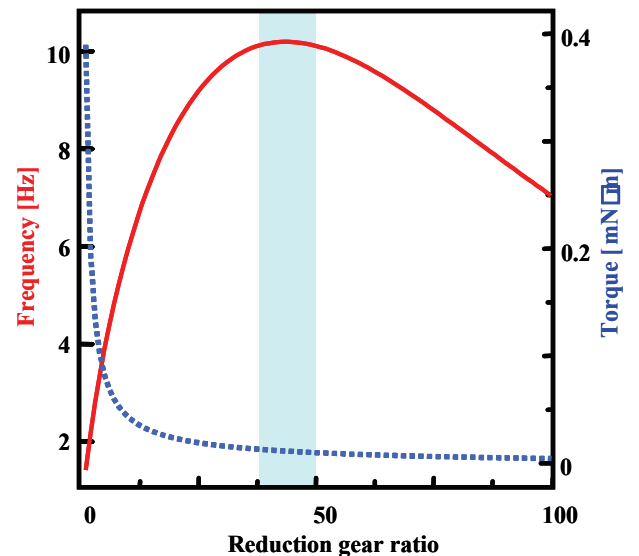

Figure 6: Optimized Motors for Actuating System

Figure 7: Optimized Reduction Ratio for Actuating System weight which is conducted by adding 1.5 [g] to motor weight, and aspect ratio, drag coefficient and reduction ratio are 3.0, 0.775 and 40.0, respectively.

As previously described, flapping frequency of 7-15 [Hz] and wing loading 0.5-1.5 [N/m2] are required to flapping robot. Therefore, the motors near red circle (Mk04s-24, Mk04-24, A2, ϕ 2.8mm and J7) are suitable for motor of actuating system. In this paper, Mk04-24 is used for flapping robot, because it is expected that flapping frequency which affect unsteady aerodynamic forces is important for flight. Not surprisingly, this judgmental standard can change by intention of developer. For instance, J7 is the optimized motor if the developer requires a light weight flapping robot.

3.3. Optimization of reduction ratio for actuating system of flapping robot

The reduction ratio of actuating system affects flapping frequency as it is clear from Eq.7. Therefore, it seems to consider that estimating optimized reduction ratio is important factor for developing flapping robot. In this chapter, the optimized reduction ratio is estimated for obtaining high flapping frequency with conducted model.

Fig.7 shows the relation between reeducation ratio and flapping frequency. The horizontal and vertical axes show reduction ratio and flapping frequency respectively. The solid and dashed lines show flapping frequency and load (torque) at motor. The selected motor, aspect ratio, drag coefficient and wing loading are Mk04-24, 3.0, 0.775 and 1.4 [N/m^2], respectively. The flapping frequency increases as reduction ratio enlarges, and it becomes peak at reduction ratio of 40-50. When reduction ratio is larger than 50, the flapping frequency decreases. Because actual reduction ratio depends on combination of gears, it is distributed discretely. The numbers of light, small and commercial gear teeth are 9, 12, 40, 60 and 90, considering these combinations, reduction ratio of 44.4 (Using 4 gears; the no. of teeth are 9, 9, 60, 60) can be realized. As the result, the reduction ratio of 44.4 is used for flapping robot in this paper.

4. FLAPPING ROBOT AND ITS FLIGHT

Fig.8 and Table 1 show a developed flapping robot and its specification. The developed flapping robot does not have tail plane as Fig.8 shows. Its wing span length, chord length and weight are 40 [mm], 80 [mm] and 1.97 [g], respectively, and the author achieves developing for small and light flapping robot. This flapping robot can fly with range of flapping frequency of 7 − 11 [Hz] by adjusting the voltage which battery applies to. The flapping frequency is higher than estimating flapping frequency of approximately 10 [Hz] because the flapping frequency on free flight tends to be higher than that on fixed flight. It has already been confirmed that the flapping frequency on fixed flight is approximately 10 [Hz].

Figure 8: Flapping Robot without Tail Plane

Table 1: Specifications of Flapping Robot

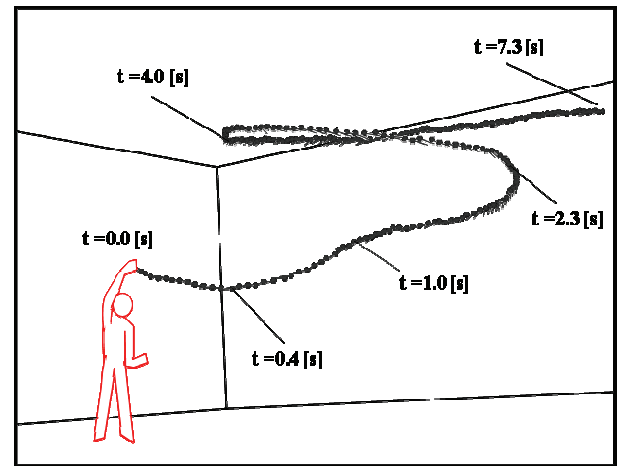| Wing span length	240 [mm]
Wing chord length	80 [mm]
Mass	1.9 [g]
Wing loading	1.5 [W/m^2]
Flapping frequency	7 ~ 11 [Hz]
Aspect ratio	3.0
battery endurance	15 [min]

Figure 9: Flight Trajectory of Flapping Robot

In this flapping robot, the mean pitch angle, vertical flight velocity and yaw angle can be adjusted by dihedral angle of wing, flapping frequency and mounting position of both wings. The flight can be controlled by adjusting these parameters. However these adjust is static, the flight can be controlled in case of dynamic change of flapping frequency and mounting position. Therefore, the flight altitude and direction can be commanded remotely by remote control of these parameter. On the other words, the flight of this flapping robot can be controlled remotely.

Fig.9 shows flight trajectory of flapping robot. The symbol *t* in this figure shows elapsed time from flight starting. Between flight time t = 0.0 [s] and t = 1.0 [s], the flight altitude once vibrates. After t = 1.0 [s], flapping robot continues to increase the flight altitude stably. At the around t = 2.3 [s], flapping robot begins to turn right. Between t = 2.3 [s] and t = 4.0 [s] approximately, flapping robot change turning direction from toward right to toward left. At the around t = 4.0 [s], flapping robot turns left and flies stably. After t = 7.3 [s], flapping robot goes out range of camera. Until t =7.3 [s], the flight altitude continues to increase despite of flight direction variation. From this result, developed flapping robot can fly upward with flight direction variation except flight starting 1 second, and it can fly for about 15 minutes.

Fig.10 shows the vertical velocity of flapping robot on the stable flight with different flapping frequencies. The horizontal and vertical axes show flapping frequency and vertical flight velocity. The red circles and the blue circles indicate flight which vertical velocity is lower than 0 [m/s] and higher than 0 [m/s], respectively. The red line indicates straight line of linear approximation. It is understood that vertical velocity increases linearly from -0.3 [m/s] to 0.3 [m/s] between flapping frequency of 7.8 [Hz] and 10.5 [Hz], and the gradient of that is 0.20 [m/s*Hz]. From this result, flapping frequency affects for vertical flight velocity, and it is confirmed that flapping frequency is important parameter for control of flight altitude. Moreover, from these results, it seems to consider that designing method which can select optimized motor and reduction ratio by estimating flapping frequency which is important parameter for control of flight altitude with modeling of wings and actuating system is effective for flapping robot developing.

5. CONCLUDING REMARKS

Considering wing drag and characteristics of motor, actuating system of flapping robot is modeled, and the designing method with its model for actuating system is suggested. With its designing method, actuating system is designed, and small and light flapping robot which can fly for about 15 minutes is developed. Moreover, from these results, it seems to consider that designing method which can select optimized motor and reduction ratio by estimating flapping frequency which is important parameter for control of flight altitude with modeling of wings and actuating system is effective for flapping robot developing.

REFERENCES

Yongsheng L., Wei Shyy, D. Viieru and B. Zhang, 2003. Membrane wing aerodynamics for micro air vehicles, *Progress in Aerospace Sciences*, 39, pp. 425–465

Robert C. Michelson, 2008. Test and Evaluation of Fully Autonomous Micro Air Vehicles, *The ITEA Journal*, 29(4), pp. 2-11
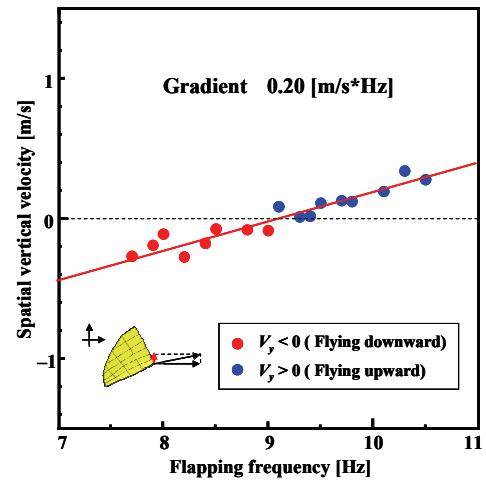
Figure 10: Relation between Flapping Frequency and Vertical Velocity

Zaeem A .Khan and Sunil K. Agrawal, 2006. Design of Flapping Mechanisms Based on TransverseBending Phenomena in Insects, *IEEE International Conference on Robotics and Automation*, pp. 2323-2328, May 2006, Orlando, Florida

F. O. Lehmann, 2004. The mechanisms of lift enhancement in insect flight, *Naturwissenschaften*, 91, pp. 101-122

M.H.Dickinson, F.O.Lehman, S.P.Sane, 1999. Wing Rotation and the Aerodynamic Basis of Insect Flight, *Science*, 284, pp. 1954-1960

T.Maxworthy, 1979. Experiments on the Weis-Fogh mechanism of lift generation by insects in hovering flight. Part. 1 Dynamics of the fling, *J.Fluid Mech.*, 93, part1, pp. 47-63

Robert J. Wood, 2008. The First Takeoff of a Biologically Inspired At-Scale, *IEEE TRANSACTIONS ON ROBOTICS*, 24 (2), pp. 341-1960

De Croon, G.C.H.E., de Clerq, K.M.E., Ruijsink, R., Remes, B., and de Wagter C., 2009. Design, aerodynamics, and vision-based control of the DelFly, *In the International Journal of Micro Air Vehicles*, 1(4), pp. 255-262

Sunada S, Kawachi K, Watanabe I, and Azuma A, 1993. Performance of a Butterfly in Take-off Flight, *Journal of Experimental Biology*, Vol. 183, pp. 249-277

H. Tanaka, K. Hoshino, K. Matsumoto, and I. Shimoyama, 2005. *The 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'05)*, pp. 310 – 315, August 2-6, Edmonton, Alberta, Canada

Che-Shu Lin, Chyanbin Hwu and Wen-Bin Young, 2006. The thrust and lift of an ornithopter's membrane wings with simple flapping motion, *Aerospace Science and Technology*, 10, pp. 111-119

A New Approach for Industrial Diagnosis by Neuro-Fuzzy systems: Application to Manufacturing System

[a] Rafik Mahdaoui, Hayet Mouss, [c] Djamal Mouss

[a,b,c] *Laboratoire d'Automatique et Productique (LAP) Université de Batna*

1, Rue Chahid Boukhlouf 05000 Batna,Algeria.

[a] mahdaoui.rafik@yahoo.fr, [b] hayet_mouss@yahoo.fr, [c] d_mouss@yahoo.fr

Abstract

In this study we propose a strategy for the follow-up of the process behavior and detection of failures. An approach of industrial diagnosis based on the statistical pattern recognition Neuro-Fuzzy being based on a digital representation and symbolic system of the forms is implemented. Within this framework, data-processing interactive software of simulation baptized NEFDIAG (NEuro Fuzzy DIAGnosis) version 1.0 is developed. This software devoted primarily to creation, training and test of a classification Neuro-Fuzzy system of industrial process failures. NEFDIAG can be represented like a special type of fuzzy perceptron, with three layers used to classify patterns and failures.The system selected is the workshop of SCIMAT clinker , cement factory in Algeria.

Keywords: Diagnosis; artificial neuronal networks; fuzzy logic; Neuro-fuzzy systems; pattern recognition; FMEAC.

1. Introduction

The function of diagnosis is a very complex task and can be only one part solved by the technique of pattern recognition (PR). The diagnosis by PR can be presented as an alternative solution at the model approach since the operating modes are modeled, in an analytical manner, but by using only one whole of measurements of these operating [8]. Therefore the human expert in his mission of diagnosing the cause of a failure of a whole system uses quantitative or qualitative information. On another side, in spite of the surprising results obtained by the ANN in monitoring, and precisely in diagnosis, the results remain far from equal human sensory capacities and reasoning. Fuzzy logic provides another very effective technique in industrial diagnosis.

Also, can we entirely replace the human expert for automating the task of diagnosis by using the neuro-fuzzy approach? In addition, how did the human expert gather all relevant information and permit him to make their decision? Our objective consists of the following: making an association (adaptation) between the techniques of fuzzy logic and the neural techniques (Neuro-fuzzy system), choosing the types of neural networks, determining the fuzzy rules, and finally determining the structure of the Neuro-Fuzzy system to maximize the automation of the diagnosis task

In order to achieve this goal we organize this article into three parts. The first part presents principal architectures and principles for Neuro-Fuzzy systems operation and their applications (sections 2 and 3). The second part is dedicated to the workshop of clinker on the level of cement factory (section 4). Lastly, in the third part we propose a Neuro-Fuzzy system for system of production diagnosis (section 5).

2. Neuro-Fuzzy systems

The hybrid systems which combine fuzzy logic, genetic algorithms, neural networks, and expert systems prove their effectiveness in a variety of real world problems and in industry. Each intelligent technique has particular properties (for example capacity of training, explanation of decisions). Each technique is appropriate to solve particular problems.

As an example, neural networks for example are used for the recognition of the models. However, they are unable to explain how they reach their decisions. Therefore for the fuzzy logic using information and can explain their decisions still cannot automatically acquire the rules they

use action to make these decisions. These limits were a reason behind the creation of intelligent hybrid systems where, two or more techniques are combined in order to overcome the limitations of only one technique.

In order to summarize the contribution of the Neuro-Fuzzy systems, Table 1 gathers the advantages and the disadvantages of fuzzy logic and ANN.

Table 1

Comparing neural networks and fuzzy systems

Neural network	Fuzzy system
Black box.	Rules must be available.
No mathematical process model required.	No mathematical process model required.
Rules cannot be extracted.	Rules must be available.
Prior knowledge cannot be used	Prior knowledge can be used.
(Learning from scratch).	Cannot learn.
No guarantee that leaning convergence.	Tuning may be not successful.
	Simple implementation and interpretation.
Different learning algorithms	

The applications show the advantages of fuzzy logic when the model of the systems is difficult to implement. Unfortunately, with the increase in the complexity of the process model, we encountered a difficulty to develop fuzzy rules and membership functions.

This difficulty led to the development of another approach which is similar to Neuro-fuzzy approaches. The integration of the neural networks and the fuzzy inference systems can be grouped in three principal categories: cooperative, concurrent and hybrid Neuro-Fuzzy [5]. As, one can say as the Neuro-Fuzzy systems are models of connection which allow the training like a Artificial Neural Network, but their structure can be interpreted like a set of fuzzy rules.

2.1 Definition

A hybrid Neuro-Fuzzy System is a neural network with a fuzzy signal, fuzzy weights, a fuzzy transfer function and a module of adaptation of the structure of NFS by a training of set of data [7].

2.2 Principle of function

The hybrid Neuro-Fuzzy Networks learn from scratch and formulate the models by using an algorithm of supervised training that examines the data in a data set which is comprised of examples of inputs and their associated outputs.

During the phase of training, a hybrid Neuro-Fuzzy system modifies its internal structure to reflect the relationship between the inputs and the outputs in the set of the measurements (knowledge bases). The accuracy of a Neuro-Fuzzy Network is checked after the cycle of training is complete by using a separate input set and outputs called the validation.

2.3 The Fuzzy Perceptron

The architecture of the fuzzy perceptron is identical to that of the usual multi-layer perceptron, but the weights are modeled by fuzzy sets. Activations, inputs, and the functions of propagation will be changed. The functioning of this model is interpretable by linguistic rules and uses knowledge bases of the rules a priori, therefore the training can not start at zero (the rule base is not empty), «Fig. 1» illustrates a fuzzy perceptron with N inputs, M fuzzy rules and P outputs. The neurons of the first layer carry out the first phase of the fuzzy inference. With each observation, the neurons of the first layer of input are tasked to calculate the degrees of membership of the fuzzy variables to the various subsets of linguistic terms. The function of membership uses the symmetrical triangular function. The neurons of the second layer calculate the degree of truth of the antecedents of the fuzzy rules via T-norm.

The number of neurons of this layer is equal to the basic size of rules. A total connection between the first layer and the second layer can exist because the connections are not constrained by the structure of the linguistic rule. The output values of the third layer are the maximum of the activation values of all the rules units that are associated within a class.

In order to implement this type of fuzzy perceptron architecture for diagnosing a dedicated production system we developed the data-processing software NEFDIAG.

3. Description of NEFDIAG.

3.1 Introduction

NEFDIAG is a data processing program for interactive simulation. The NEFDIAG development was carried out within LAP (University of Batna), written under DELPHI, and was primarily dedicated to the creation, the training, and the test of a Neuro-Fuzzy system for the classification of the breakdowns of a dedicated industrial process. NEFDIAG models a fuzzy classifier Fr with a whole of classes $C = \{c_1, c_2 ...... c_m\}$.

The structure and training of NEFDIAG can be represented using a special type of fuzzy perceptron, with three layers used to classify failures.

NEFDIAG makes it's training by a set of forms and each form will be affected (classified) using one of the preset classes. Next NEFDIAG generates the fuzzy rules by:

evaluating of the data, optimizing the rules via training and using the fuzzy subset parameters, and partitioned the data into forms «characteristic» and classified with parameters of the data. NEFDIAG can be used to classify a new observation. The system can be represented in the form of fuzzy rules

If symptom1 is A_1
Symptom2 is A_2
Symptom3 is A_3
Symptom $_N$ is A_n
Then the form $(x_1, x_2, x_3..., x_n)$ belongs to class «Failure mode 1».

For example A1 A_2 A_3 A_n are linguistic terms represented by fuzzy sets. This characteristic makes it possible to complete the analyses on our data, and to use this knowledge to classify them. The training phase of the networks of artificial Neuro-Fuzzy systems makes it possible to determine or modify the parameters of the network in order to adopt a desired behavior. The stage of training is based on the decrease in the gradient of the average quadratic error made by network RNF.
The NEFDIAG system typically starts with a knowledge base comprised of a collection partial of the forms, and can refine it during the training. Alternatively NEFDIAG can start with an empty base of knowledge. The user must define the initial number of the functions of membership for partitioning the data input fields. And it is also necessary to specify the number K, which represents the maximum number of the neurons for the rules which will be created in the hidden layer. The principal steps of the training algorithm.
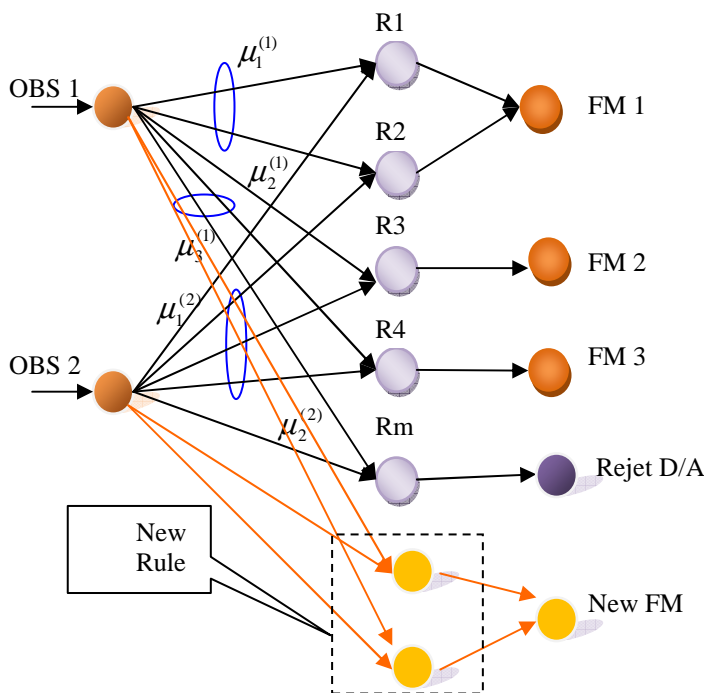
Fig. 1. *Neuro-Fuzzy Architecture*

Initialization: For each data originating from the sensors there is an input unit, and for each mode of failure there is an output unit. For each input unit there is an initial fuzzy partition specified with number of the of triangular membership functions.
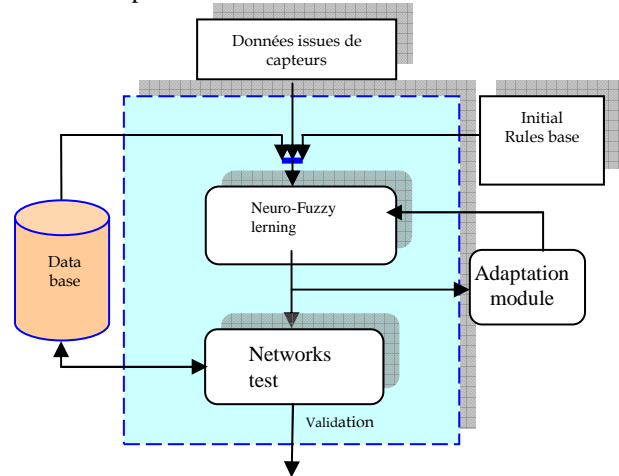
Fig. 2. *Neuro-Fuzzy learning phases.*

Training of the rules:
As mentioned, the NEFDIAG system starts with a base of knowledge comprised of a partial of forms, and can refine the knowledge base during the training «Fig. 3». The rule will be created by analyzing (for a given form F) the combination of the functions of membership, such as each combination producing the greatest function of membership «Fig.3 ». If this combination is not identical for the rules that exist in the rules knowledge base, and number of rules is not maximum, then new rules will be created and added to the rules knowledge base « Fig. 1»,

$$\varepsilon_r = \tau_r(1-\tau_r)\frac{1}{m}\sum_{j=1}^{m}(2v_r^{(j)}(t_j)-1)\left|E_j\right| \qquad (1)$$
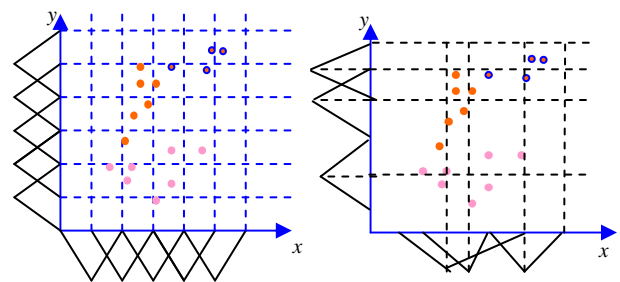

Fig. 3. *Neuro-Fuzzy learning rules.*

The algorithm of training detects determines all the antecedents of the rules, and then creates the list of the antecedents. Initially this list is empty, or contains antecedents of rules of knowledge a priori. The algorithm selects then antecedent A and seeks to create the basic list

of rule candidates. The best rules will be selected from the knowledge base of the rules candidates, based upon their measurement of performance [7].

$$P_R = \frac{1}{s} \sum_{(p,t)\in\Gamma} (-1)^c R(p) \qquad (1)$$

In this case some classes (mode of failure) would not be represented in the knowledge base of rules, especially if the rules for this mode of failure represent a value of very small performance.

Training of the Functions of Membership:
For the training of the membership functions, a simple reverse propagation technique is used. It depends on the error of output for each **unit** of rules. Each rule changes its membership functions by the amount of change in their supports « fig. 7 ».
It is necessary that the error of each rule is calculated [5].
τ_r is fulfillment of a rule *r*.
After the appearance of a new mode of failure in the training phase of our system, the Neuro-fuzzy network will make an adaptation or a reorganization of the system to adapt to the new situation «Fig. 2».
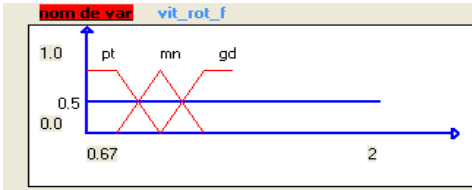
Fig. 4. *V_R_F before training*
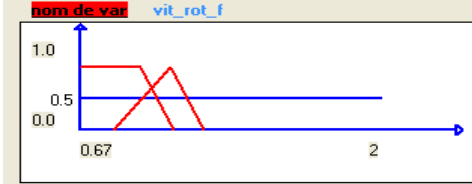
Fig. 5. *V_R_F after training*

Initially the layers of rules (or rules bases) will be added all rules for failure mode detected. Then in the layer of the modes of failure, another node will be connected to the Neuro-Fuzzy network «fig.7 »
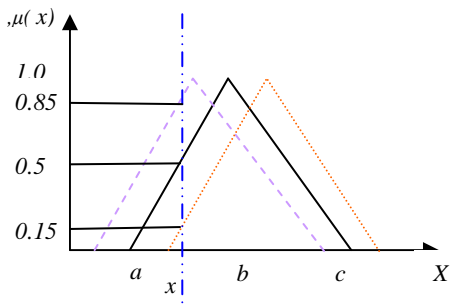
Fig. 7. *The training of membership function*

4. The workshop of clinker

Our application is illustrated on an industrial process for the manufacture of cement. This installation is at the cement factory. This cement factory has a capacity of 2.500.000 t/year Two furnaces comprised of several units determine the various phases of the manufacturing process of cement. The cooking workshop gathers two furnaces whose flow clinker is 1560 t/h. The cement crushing includes two crushers that are each rated 100t/hour. Forwarding of cement is carried out from two stations, one for the trucks and another for the coaches «Fig.6, ".
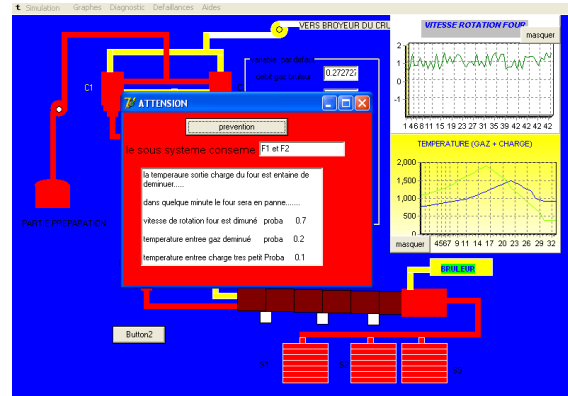

Fig.. 6. *Alarm message.*

5. Neuro-Fuzzy diagnosis

5.1 Dysfunctions analyses

This step has an objective of the identification of the dysfunctions which can influence the mission of the system. This analysis and recognition are largely facilitated using the structural and functional models of the installation. For the analysis of the dysfunctions we adopted the method of Failure Modes and Effects Analysis and their Criticality (FMEAC). While basing itself on the study carried out by [6], on the cooking workshop, we worked out an FMEAC by considering only the most critical modes of the failures (criticality >10), and for reasons of simplicity [6]. Therefore we have a Neuro-fuzzy system of 27 inputs and 4 outputs which were used to make a diagnosis of our system. The rules which are created with the system are knowledge a priori, a priori the base of rule. Each variable having an initial partition will be modified with the length of the phase of training (a number of sets fuzzy for each variable). The reasoning for the diagnosis is described in the form of fuzzy rules inside our Neuro-fuzzy system.
The principal advantage of the use of the base of fuzzy rules lies in its modularity and its facility of expansion

(suppression or addition of other rules). The initial rules base to establish the diagnosis of the failures is built by evaluating the model worked out during dysfunction phases of our system (FMEAC). Indeed, this analysis makes it possible to establish the bonds between the components failing and the symptoms observed. These bonds are represented as of fuzzy rules, in turn building the knowledge base that will be trained and later tested, and finally used to carry out the fuzzy reasoning necessary to lead to the results expressing the function of diagnosis.
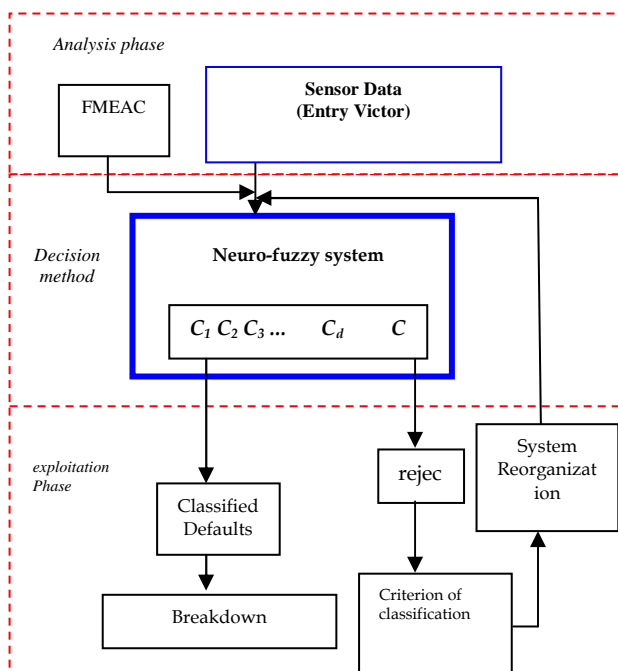
Fig. 7. *The diagnosis by NEFDIAG.*

The detection of anomalies is represented in the form of an alarm message announces to the operator (user) the appearance of an anomaly (or anomalies) and makes it possible to identify the component responsible using a data base containing the information provided by the FMEAC (mode of failure, possible causes, equipment, effects on the system).

Fig. 2 illustrates the system with the presence of a dysfunction in the cooking workshop. Let us note that in this study, the anomalies or dysfunctions indicate functional anomalies.

After the posting of the message, the operator can save the message or to remove it « Fig. 6». Additionally, NEFDIAG makes interventions to control the variables which are the origin of the current failure « Fig. 8»

6. conclusion

In this article, we presented a new tool for diagnosis by Neuro-Fuzzy systems using FMEAC; we detailed the implementation of an example of industrial application using the NEFDIAG tool. We illustrated our assistance tool to the diagnosis using a prototype NEFDIAG installed on a PC. We approached the various stages to be followed for the development of the assistance system for diagnosis, starting from the methods of classification and fuzzy recognitions of the forms. NEFDIAG is represented like a special type of fuzzy perceptron, with three layers used to classify failures (Type 3). NEFDIAG accomplishes training in two phases. The training phase generates the fuzzy rules using the collection of data, optimizes the rules using the parameters of the fuzzy sets, and finally uses the fuzzy sets for partitioning and classifying the form data and their parameters.

REFERENCES

[1] H.K. Kwan and Y.Cai, " A fuzzy neural network and its application to pattern recognition" IEEE Transactions on Fuzzy Systems, 3 pp. 185-193. 1994

[2] J.-S. Roger Jang" ANFIS: Adaptive-network-based fuzzy inference system " *IEEE* Trans. Syst., Man, and Cybernetics, 23(1993) 665-685.

[3] J.M. Keller and H.Tahani "Implementation of conjunctive and disjunctive fuzzy logic rules with neural networks" International Journal of Approximate Reasoning,6(1992) 221-240.

[4] G. Zweingelstein " Diagnostic des défaillances, théories et pratique pour les systèmes industriels " Col. Traité des nouvelles technologies, séries diagnostic et maintenance, Hermès ,1995.

[5] L.-X wang and J.M Mendel "Generation fuzzy rules by learning from examples" IEEE trans. Syst., Man, and Cybernetics, (1992) 22(6):1414-1427.

[6] D. Mouss "Diagnostic et conduite des systèmes de production par approche a base de connaissances " Thèse de doctorat Université de Batna, 2006.

[7] D. Nauck "Neuro-fuzzy systems: review and prospects"European congress on intelligent technique and sift computing (EUFIT'97), Aachen, sep.8.11, (1997), pp.1044-1053.

[8] D. Racoceanu "contribution à la surveillance des systèmes de production en utilisant les techniques de l'intelligence artificielle" habilitation à diriger des recherches, université de Franche Comté de Besançon, 2006.

Authors' Index