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.

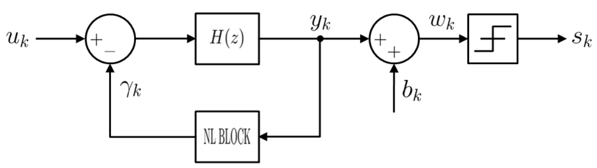

Figure 1: Block diagram of the system model

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):

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

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.

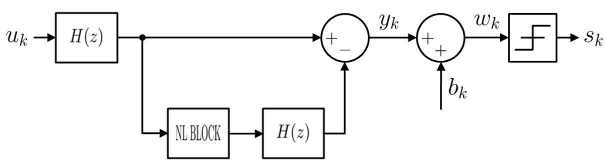

Figure 2: New block diagram of the system model

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: $\boldsymbol{u}, \boldsymbol{s}, L, N$

$\widehat{\boldsymbol{\chi}}_1 \leftarrow [1 \quad 0 \quad \ldots \quad 0]^T$
$\widehat{\boldsymbol{\theta}}_1 \leftarrow \frac{\widehat{\chi}_1}{\|\widehat{\chi}_1\|_2}$
$\hat{\eta}_1 \leftarrow 0$
$\hat{\alpha}_1 \leftarrow 0$
for $k = 1$ to N **do**
  $\boldsymbol{\varphi}_{k,L} \leftarrow (u_l)_{l=k}^{k-L+1}$
  $\boldsymbol{\psi}_{k,L} \leftarrow \left(\left(\widehat{\boldsymbol{\theta}}_k^T\boldsymbol{\varphi}_{l,L}\right)^3\right)_{l=k}^{k-L+1}$
  $\hat{y}_k \leftarrow \widehat{\boldsymbol{\theta}}_k^T\boldsymbol{\varphi}_{k,L} - \hat{\alpha}_k\widehat{\boldsymbol{\theta}}_k^T\boldsymbol{\psi}_{k,L}$
  $\hat{s}_k \leftarrow S(\hat{y}_k)$
  $\widehat{\boldsymbol{\chi}}_{k+1} \leftarrow \widehat{\boldsymbol{\theta}}_k - \mu_k(s_k - \hat{s}_k)^2\hat{y}_k\left(\boldsymbol{\varphi}_{k,L} - 4\hat{\alpha}_k\boldsymbol{\psi}_{k,L}\right)$
  $\widehat{\boldsymbol{\theta}}_{k+1} \leftarrow \frac{\widehat{\chi}_{k+1}}{\|\widehat{\chi}_{k+1}\|_2}$
  $\hat{\eta}_{k+1} \leftarrow \hat{\alpha}_k + \mu_k(s_k - \hat{s}_k)^2\hat{y}_k\widehat{\boldsymbol{\theta}}_k^T\boldsymbol{\psi}_{k,L}$
  $\hat{\alpha}_{k+1} \leftarrow \hat{\eta}_{k+1}\|\widehat{\boldsymbol{\chi}}_{k+1}\|_2^3$
end for
return $\widehat{\boldsymbol{\theta}}_{k+1}, \hat{\alpha}_{k+1}$

In algorithm 1, the normalization step on line 11 ensures that the norm of $\widehat{\boldsymbol{\theta}}_k$ 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.

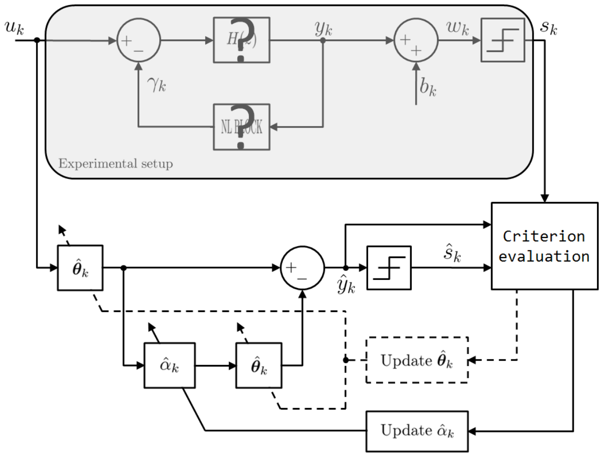

Figure 3: Block diagram of nonlinear LIMBO

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 $\widehat{\boldsymbol{\theta}}_k$ is defined as $1 - \nu_k$ where ν_k is the cosine of the angle made by $\widehat{\boldsymbol{\theta}}_k$ and $\boldsymbol{\theta}$. Since both vectors are normalized, we have $\nu_k = \boldsymbol{\theta}^T \widehat{\boldsymbol{\theta}}_k$ and the following equivalence relation:

$$\begin{aligned} \lim_{k\to\infty} (1 - \nu_k) = 0 &\Leftrightarrow \lim_{k\to\infty} \nu_k = 1 \\ &\Leftrightarrow \lim_{k\to\infty} \widehat{\boldsymbol{\theta}}_k = \boldsymbol{\theta} \end{aligned} \tag{15}$$

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.

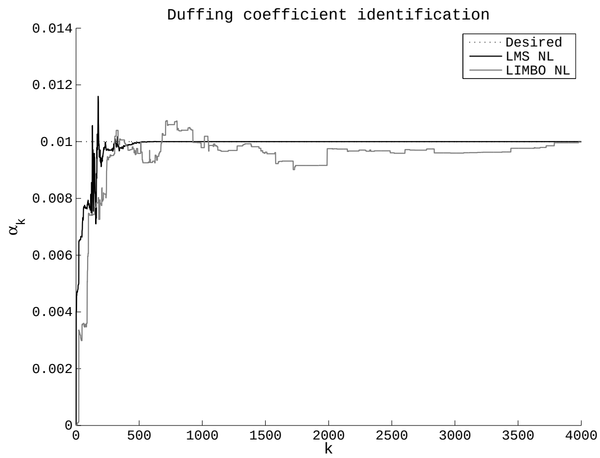

Figure 4: Comparison of nonlinear LMS and LIMBO methods for Duffing coefficient identification

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.

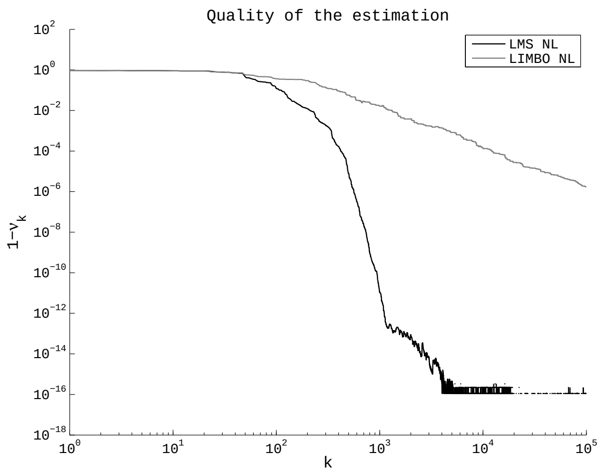

Figure 5: Comparison of nonlinear LMS and LIMBO methods in terms of convergence speed and estimation quality (SNR = ∞ dB)

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.

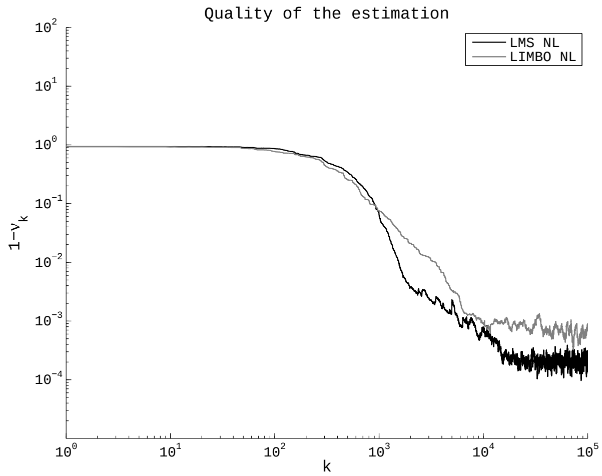

Figure 6: Comparison of nonlinear LMS and LIMBO methods in terms of convergence speed and estimation quality ($\mathrm{SNR} = 20$ dB)

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 $\boldsymbol{\theta}$ 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.

ACKNOWLEDGMENTS

This work was supported by funding from the Catrene project TOETS (Towards One European Test Solution).

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.