

A Recursive System Identification Method Based on Binary Measurements

Kian Jafari, Member, IEEE, Jerome Juillard, Member, IEEE and Eric Colinet

Abstract—An online approach to parameter estimation problems based on binary observations is presented in this paper. This recursive identification method relies on a least-mean squares approach which makes it possible to estimate the coefficients of a finite-impulse response system knowing only the system input and the sign of the system output. The impulse response is identified up to a positive multiplicative constant. The role of the regulative coefficient is investigated thanks to simulated data. The proposed method is compared with another online approach: it is shown that the proposed method is competitive with the other one in terms of estimation quality and of calculation complexity.

I. INTRODUCTION

Parameter estimation methods based on binary or quantized output observations have known an increasing interest in the past few years. The use of such methods is mostly motivated by the fact that low-resolution observations are usually much cheaper to obtain than high-resolution ones. Several applications based on binary sensors or on binary data are given in [1], [2].

The method presented in this paper is aimed at the test of micro-electronic devices, such as MEMS (Micro-electromechanical systems). It is known that as characteristic dimensions become smaller, the dispersions afflicting electronic devices tend to become larger. Typical sources of dispersion are variations in the fabrication process or changes in the operating conditions, such as temperature, pressure or humidity [3], [4]. As a consequence, it is not usually possible to guarantee a priori that a given device will function properly. The tests that are run after fabrication, under different operating conditions, ensure that only suitable devices are commercialized. However, these tests are very costly: a typical figure is that one third of the cost of a micro-fabricated device is due to testing [5]. An alternative is to integrate self-test (and self-tuning) features, such as

parameter estimation routines, in each device, so that it can compensate the fabrication dispersions and adapt to changing conditions. Ideally, this self-test feature should be as low-cost as possible, especially in terms of silicon area: thus, parameter estimation routines based on binary observations are very appealing, because they only require the integration of a 1-bit ADC (analog-to-digital converter). Furthermore, the proposed self-test approach should require as little data storage as possible and have a low computational complexity, so as to be easily implemented.

In this paper, we establish a recursive identification method based on binary observations which is well-adapted to the context of micro-electronic devices. This LMS (least-mean squares) parameter estimation method is based on the offline WLS (weighted-least squares) approach developed in [3], [4]. The convergence of this recursive algorithm is illustrated by simulation and the choice of an efficient regulative coefficient for convergence to the nominal system parameters is investigated. Our results are also compared with those obtained with the method developed by Wigren [6], [7], which is also well adapted to the context of micro electronics.

The structure of the article is the following. In section II, an overview of existing identification methods based on binary or quantized observations is given. In section III, the system and its model are introduced. In section IV, the LMS-like algorithm is derived. In section V, the role of the regulative coefficient of LMS algorithm is investigated by simulations. Furthermore, the proposed method is compared with an appropriate online method [7] in terms of estimation quality and of calculation complexity. Concluding remarks and perspectives are given in section VI.

II. STATE OF THE ART

Some important contributions in the field of parameter estimation based on binary (or quantized) data were presented in [1]–[4] and [6]–[13]. These methods can be divided in two categories, depending on whether they rely on the use of a dithering signal.

Although the methods using a dithering signal at the input of the quantizer [1], [8], [9], [12], [13] may be useful in several macro-scale applications [1], they do not scale down easily to the context of integrated micro-electronics: all these methods rely on the fact that the cumulative probability

This work was supported by the research grant of French Ministry of Higher Education and Research.

K. Jafari is with SUPELEC - Department of Signal Processing & Electronic Systems, 3 rue Joliot-Curie, 91192 Gif-sur-Yvette Cedex, France (Phone: + 33 169851424, e-mail: Kian.Jafari@ieee.org).

J. Juillard is with SUPELEC - Department of Signal Processing & Electronic Systems, 3 rue Joliot-Curie, 91192 Gif-sur-Yvette Cedex, France (Phone: +33 169851408, e-mail: Jerome.Juillard@supelec.fr).

E. Colinet is with CEA-LETI - Minatec, 17 rue des Martyrs, 38054 Grenoble Cedex, France (Phone: +33 438781140, e-mail: eric.colinet@cea.fr).

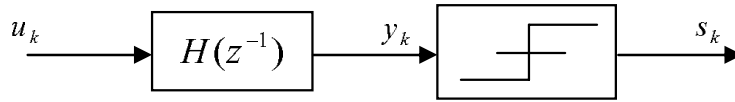


Fig. 1. Block diagram of the system model. Signal y_k is not available for measurement.

distribution function of the dithering signal must be invertible, which may only be achieved by using a high-resolution digital-to-analog converter (DAC).

The methods which are most easy to downscale are summed up hereafter. In [7], Wigren developed an LMS approach to the problem of online parameter estimation from quantized observations. This method relies on an approximation of the quantizer, which makes it possible to define an approximate gradient of the least-squares criterion. Under some hypotheses [14], [15], it is possible to guarantee the asymptotical convergence of this method to the nominal parameters.

In 2003, Negreiros [10] suggested to use 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 transfer function of the unknown system can be analytically derived. However, it is not possible to obtain any information concerning the phase of the transfer function. Juillard and Colinet [11] overcame this limitation by deriving an analytical relationship between the coefficients of the impulse response of the system and the cross-covariance of its binary input and output. Although this approach is fairly simple to implement [16], it relies on the mixing properties of the linear system, which may not be guaranteed a priori.

Recently, the theoretical framework of BIMBO (Basic Identification Method using Binary Observations), an offline WLS approach to parameter estimation problems based on binary data, was introduced in [3]. This method 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-squares criterion. It is possible to guarantee the consistency of this approach even in the presence of measurement noise provided the signal at the input of the quantizer is Gaussian and centred. Furthermore, the estimation quality of BIMBO was investigated in the sense of correlation coefficient between the nominal system parameters and the estimated system parameters [17].

An alternative WLS criterion was also presented in [4] which is easier to implement than the first one in the context of microelectronics. This approach is as efficient as the one proposed in [3] in the absence of measurement noise. However, it leads to a systematic error in the presence of measurement noise.

III. FRAMEWORK AND NOTATIONS

The known input signal u_k is filtered by a discrete-time invariant linear system $H(z^{-1})$ to produce the (scalar) value of the system output y_k at time k (Fig. 1). H has a finite impulse response of length L , i.e. the impulse response can be represented by a column vector $\theta_0 = (\theta_k)_{k=1}^L$.

Let $\hat{\theta}_k$ be the estimated vector of parameters (of length L) at time k . The output and the estimated output can be expressed as:

$$\begin{aligned} y_k &= \theta_0^\top \Phi_k \\ \hat{y}_k &= \hat{\theta}_k^\top \Phi_k \end{aligned} \quad (1)$$

where $\Phi_k = (u_l)_{l=k-L+1}^k$ is the (column) vector of inputs at time k . The system output then goes through a 1-bit ADC so that only the sign of the system output is known. Thus, we define $s_k = S(y_k)$ and $\hat{s}_k = S(\hat{y}_k)$ where

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

In [3] and [4], we showed that, given N observations, the problem of parameter estimation from binary measurements could be treated offline by minimizing a WLS criterion of the form:

$$G_1^N(\hat{\theta}) = \frac{1}{4N} \sum_{k=1}^N \hat{y}_k^2 (s_k - \hat{s}_k)^2. \quad (2)$$

It is also established that this criterion is continuously differentiable with respect to $\hat{\theta}$ and is thus amenable to simple optimization algorithms (e.g. gradient) [3], [4].

Our goal is to develop a recursive estimation method to estimate θ_0 based on observations of the binary output s_k knowing u_k . Note that $\hat{\theta} = \lambda \theta_0$ where $\lambda > 0$, is a minimizer of G_1^N , i.e. the sign of y_k is not changed by multiplying it with a positive constant. Without loss of generality, we assume in the rest of the paper that $\|\theta_0\| = 1$.

IV. PROPOSED LMS APPROACH

Starting from a general LMS algorithm, we derive a practical LMS-like method (that is called LIMBO, for LMS-BIMBO, in the rest of paper) in order to estimate the system parameter online from binary observations (Fig. 2). A general LMS algorithm can be written:

$$\hat{\theta}_{k+1} = \hat{\theta}_k - \alpha_k \frac{\partial E_k^2}{\partial \hat{\theta}_k},$$

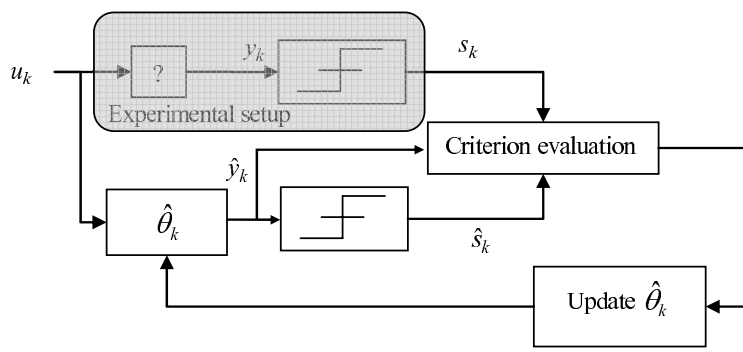


Fig. 2. Block diagram of LIMBO. A parametric model of the real system is built so as to maximize the similarity between θ and $\hat{\theta}_k$.

where α_k must satisfy some conditions to guarantee the stability and convergence of the algorithm. In classical problems, where the output is available, the instantaneous error is defined as:

$$E_k = (y_k - \hat{y}_k).$$

In the framework of system identification based on binary observations, we only have access to s_k and the unweighted instantaneous error,

$$E_k = (s_k - \hat{s}_k), \quad (3)$$

is not differentiable with respect to $\hat{\theta}_k$. However, the following form of instantaneous error can be derived from (2) which is differentiable with respect to $\hat{\theta}_k$ ([3], [4]) as opposed to (3):

$$E_k = |s_k - \hat{s}_k| \hat{y}_k. \quad (4)$$

As a consequence, we propose the following algorithm:

$$\hat{\theta}_1 = [1 \ 0 \ 0 \ \dots \ 0 \ 0], \quad (5a)$$

$$\hat{\Psi}_{k+1} = \hat{\theta}_k - \alpha_k \frac{\partial E_k^2}{\partial \hat{\theta}_k}, \quad (5b)$$

$$\hat{\theta}_{k+1} = \frac{\hat{\Psi}_{k+1}}{\|\hat{\Psi}_{k+1}\|}. \quad (5c)$$

Equation (5c) ensures that the norm of $\hat{\theta}_k$ is equal to unity. From (1) and (4), (5b) can be explicitated as:

$$\hat{\Psi}_{k+1} = \hat{\theta}_k - 2\alpha_k (s_k - \hat{s}_k)^2 \hat{y}_k \Phi_k. \quad (6)$$

The quality of the online estimation $\hat{\theta}_k$ is defined as:

$$1 - v_k,$$

where v_k is the cosine of the angle made by $\hat{\theta}_k$ and θ_0 . Since both vectors are normalized, we have

$$v_k = \theta_0^\top \hat{\theta}_k,$$

and $(1 - v_k) \rightarrow 0 \iff v_k \rightarrow 1 \iff \hat{\theta}_k \rightarrow \theta_0$.

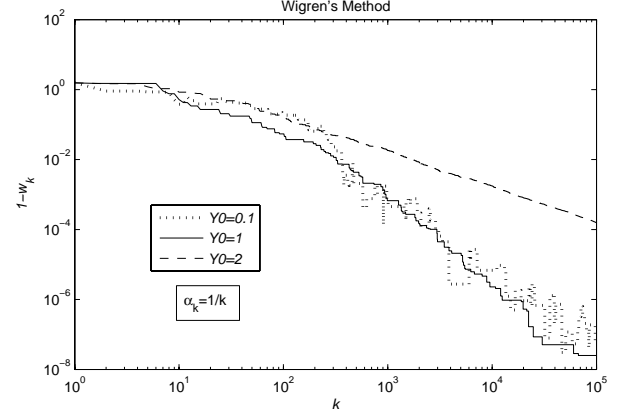


Fig. 3. Estimation quality of Wigren's method with $\alpha_k = 1/k$ for different values of y_0 .

V. RESULTS AND DISCUSSION

In this section, the results obtained with LIMBO are compared with those obtained with Wigren's approach [7] which is an efficient online method to parameter estimation problems using quantized output measurements [3]. Note that when used in the context of binary observations, Wigren's method [7] reduces, in its simplest form, to:

$$\tilde{\theta}_1 = [1 \ 0 \ 0 \ \dots \ 0 \ 0], \quad (7a)$$

$$\tilde{\Psi}_{k+1} = \tilde{\theta}_k - 2\alpha_k (\hat{s}_k - s_k) \frac{\partial \tilde{s}_k}{\partial \tilde{\theta}_k}, \quad (7b)$$

$$\tilde{\theta}_{k+1} = \frac{\tilde{\Psi}_{k+1}}{\|\tilde{\Psi}_{k+1}\|}, \quad (7c)$$

where

$$\begin{aligned} \hat{s}_k &= S(\tilde{y}_k), \\ \tilde{y}_k &= \tilde{\theta}_k^\top \Phi_k, \\ \tilde{s}_k &\approx \frac{2}{\pi} \arctan\left(\frac{\tilde{y}_k}{y_0}\right), \end{aligned}$$

and y_0 is an arbitrary constant [7]. The quantity $\partial \tilde{s}_k / \partial \tilde{\theta}_k$ is the so-called "pseudo-gradient" of \hat{s}_k . The quality of the

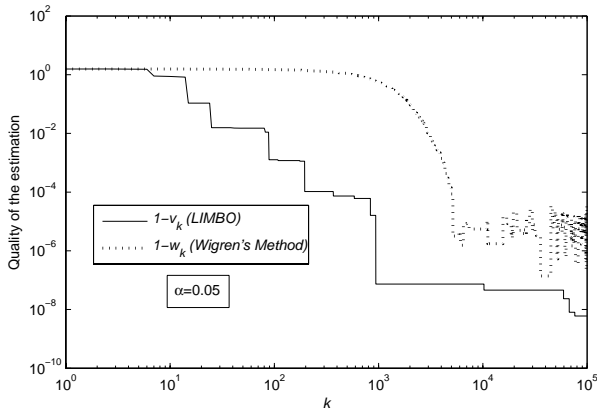


Fig. 4. Comparison of LIMBO and Wigren's method using a constant regulative coefficient ($\alpha = 0.05$) for two methods.

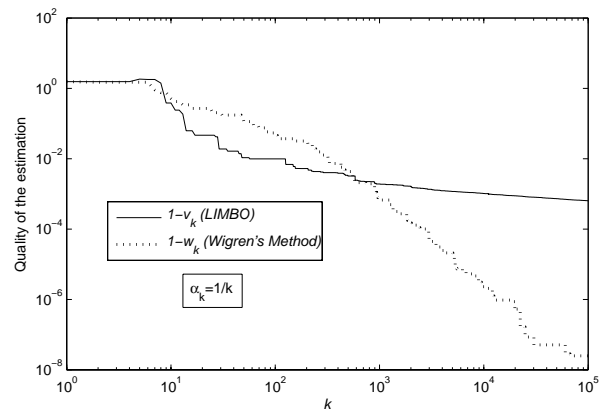


Fig. 5. Comparison of LIMBO and Wigren's method using $\alpha_k = 1/k$ for two methods.

estimation can be defined as $1 - w_k$ where,

$$w_k = \theta_0^\top \tilde{\theta}_k.$$

Furthermore the role of the regulative coefficient (α_k) is addressed for two methods. We also discuss the efficiency and the complexity of each method and their ease of implementation in the field of micro-devices.

In (5) or (7), the regulative coefficient (α_k) can be chosen with different strategies (constant or adaptive step size [18]). For instance, in [7], the convergence of the pseudo-gradient LMS is proved with $\alpha_k = 1/k$. However, it should be noted that this particular choice of regulative coefficient ($\alpha_k = 1/k$) is not perfectly adjusted to the integration context of micro-devices, because division, being one of the most complicated and expensive arithmetic operations, is costly to implement. This is the reason why we investigate the use of a constant regulative coefficient in (5) and (7).

The methods are compared, without measurement noise, on the same test case as in [7]:

$$\theta_0 = [0.2 \quad -0.14 \quad 0.8 \quad -0.56],$$

u_k is a Gaussian white noise with zero mean and unit standard deviation. The value of y_0 in Wigren's method is set by trial-and-error to $y_0 = 1$ when $\alpha_k = 1/k$ and to $y_0 = 10^3 \alpha_k$ while $\alpha_k = \alpha$ (using a constant regulative coefficient), so that a good compromise between convergence speed and estimation quality is reached. Typical results are shown in Fig. 3.

When a constant regulative coefficient is used, we find that $\alpha < 0.2$ yield reasonable results for both methods, in terms of convergence speed and estimation quality (table I), with a notable advantage for LIMBO in terms of estimation quality. The reason for this difference is that when a constant regulative coefficient is chosen, Wigren's approach applied to binary observations stops converging after reaching a

TABLE I

COMPARISON OF ESTIMATION QUALITY BETWEEN LIMBO AND WIGREN'S METHOD FOR DIFFERENT REGULATIVE COEFFICIENTS.

	$1 - v_{10^5}$	$1 - w_{10^5}$
$\alpha_k = 1/k$	$6.4e - 4$	$2.5e - 8$
$\alpha = 0.01$	$7.04e - 8$	$7.7e - 6$
$\alpha = 0.05$	$5.99e - 9$	$1.74e - 5$
$\alpha = 0.1$	$3.54e - 6$	$1.93e - 5$
$\alpha = 0.15$	$1.18e - 6$	$1.77e - 5$
$\alpha = 0.2$	$3.5e - 3$	$7.7e - 5$

certain accuracy (dependent on the choice of y_0). On the other hand, the LIMBO approach converges to the nominal parameters within the limits of finite machine precision. This is illustrated in Fig. 4.

When a monotonously decreasing regulative coefficient is used (in the present case, when $\alpha_k = 1/k$), there is a distinct advantage to using Wigren's approach: (5) converges very slowly, as is shown in Fig. 5. On the other hand, (7) converges faster to the nominal parameters (Fig. 5). However, the convergence speed and the estimation quality of LIMBO when an appropriate constant regulative coefficient is used are completely competitive to those obtained with Wigren's approach while $\alpha_k = 1/k$ (Fig. 6).

Qualitatively similar results can be obtained for other test cases: it means that the parameters of both methods (y_0 for Wigren's approach, α for LIMBO) can always be adjusted so that they converge to the nominal system parameters with a similar speed, using $\alpha_k = 1/k$ for (7) and a constant regulative coefficient (α) for (5). However, the computational cost of (5) is always smaller than that of (7): only one division (5-c) is required in LIMBO with a constant coefficient, whereas Wigren's approach with a monotonously decreasing coefficient requires at least two or three divisions, depending on whether the pseudogradient $\partial \tilde{s}_k / \partial \tilde{\theta}_k$ is actually computed or tabulated.

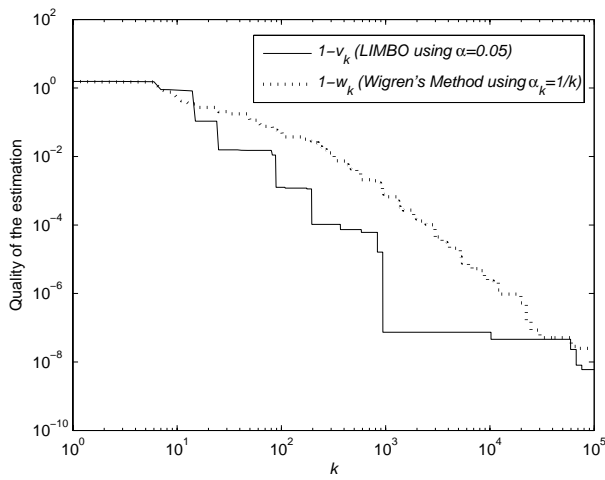


Fig. 6. Comparison of LIMBO and Wigren's method using $\alpha_k = 1/k$ for Wigren's approach and $\alpha = 0.05$ for the proposed method.

VI. CONCLUSION AND FUTURE WORKS

In this paper, LIMBO, an LMS-like method for estimating system parameters based on binary observations, was derived from the offline WLS approach presented in [3], [4]. Simulations showed that the results obtained when using a constant regulative coefficient in LIMBO were comparable, in terms of convergence speed and of estimation quality, to those obtained with the approach presented in [7], with a lesser computational complexity. This makes LIMBO an inexpensive test method, requiring only little memory storage (as opposed to an offline method) and amenable to implementation on FPGAs or small processor cores as used in System-on-Chip applications.

Proof of the convergence of LIMBO to the nominal system parameters will be given in an oncoming paper.

REFERENCES

- [1] L. Wang, J. Zhang, and G. Yin, "System identification using binary sensors," *IEEE Transactions on Automatic Control*, vol. 48, no. 11, pp. 1892–1907, 2003.
- [2] T. Wigren, "Approximate gradients, convergence and positive realness in recursive identification of a class of non-linear systems," *International Journal of Adaptive Control and Signal Processing*, vol. 9, pp. 325–354, 1995.
- [3] E. Colinet and J. Juillard, "A weighted least-squares approach to parameter estimation problems based on binary measurements," *IEEE Transactions on Automatic Control*, vol. 55, no. 1, pp. 148–152, 2010.
- [4] J. Juillard, K. Jafari, and E. Colinet, "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, 2009.
- [5] B. Charlot, S. Mir, F. Parraïn, and B. Courtois, "Electrically Induced Stimuli For MEMS Self-Test," *Proceedings of the 19th IEEE VLSI Test Symposium*, pp. 60–66, 2001.
- [6] T. Wigren, "Convergence analysis of recursive identification algorithms based on the nonlinear Wiener model," *IEEE Transactions on Automatic Control*, vol. 39, no. 11, pp. 2191–2206, 1994.

- [7] —, "Adaptive filtering using quantized output measurements," *IEEE Transactions on Signal Processing*, vol. 46, no. 12, pp. 3423–3426, 1998.
- [8] E. Rafajlowicz, "Linear systems identification from random threshold binary data," *IEEE Transactions on Signal Processing*, vol. 44, no. 8, pp. 2064–2070, 1996.
- [9] —, "System identification from cheap, qualitative output observations," *IEEE Transactions on Automatic Control*, vol. 41, no. 9, pp. 1381–1385, 1996.
- [10] M. Negreiros, L. Carro, and A. Susin, "Ultimate low cost analog BIST," *Proceedings of the 40th Conference on Design Automation*, pp. 570–573, 2003.
- [11] J. Juillard and E. Colinet, "Initialization of the BIMBO self-test method using binary inputs and outputs," *Proceedings of the 46th IEEE Conference on Decision and Control*, pp. 578–583, 2007.
- [12] L. Wang, G. Yin, and J. Zhang, "Joint identification of plant rational models and noise distribution functions using binary-valued observations," *Automatica*, vol. 42, no. 4, pp. 535–547, 2006.
- [13] Y. Zhao, L. Wang, G. Yin, and J. Zhang, "Identification of Wiener systems with binary-valued output observations," *Automatica*, vol. 43, no. 10, pp. 1752–1765, 2007.
- [14] L. Ljung, "Analysis of recursive stochastic algorithms," *IEEE Transactions on Automatic Control*, vol. 22, no. 4, pp. 551–575, 1977.
- [15] —, *System identification - theory for the user*. Upper Saddle River: Prentice Hall, 1999.
- [16] C. Le Blanc, E. Colinet, and J. Juillard, "Digital implementation of a BIST method based on binary observations," *Proceedings of the 11th Euromicro Conference on Digital System Design: Architectures, Methods and Tools (DSD08)*, pp. 709–713, 2008.
- [17] J. Juillard, K. Jafari, and E. Colinet, "Estimation Quality of a Weighted Least-Square Parameter Estimation Method Based on Binary Observations," *Proceedings of the 18th European Signal Processing Conference (EUSIPCO10)*, pp. 1043–1047, 2010.
- [18] B. Farhang-Boroujeny, *Adaptive Filters: Theory and Applications*. NY: John Wiley, 1998.