Convergence Analysis of an Online Approach to Parameter Estimation Problems Based on Binary Noisy Observations

Laurent Bourgois, Jérome Juillard

Convergence Analysis of an Online Approach to Parameter Estimation Problems Based on Binary Noisy Observations

Laurent Bourgois and Jérôme Juillard

***Abstract*— The convergence analysis of an online system identification method based on binary-quantized observations is presented in this paper. This recursive algorithm can be applied in the case of finite impulse response (FIR) systems and exhibits low computational complexity as well as low storage requirement. This method, whose practical requirement is a simple 1-bit quantizer, implies low power consumption and minimal silicon area, and is consequently well-adapted to the test of microfabricated devices. The convergence in the mean of the method is studied in the presence of measurement noise at the input of the quantizer. In particular, a lower bound of the correlation coefficient between the nominal and the estimated system parameters is found. Some simulation results are then given in order to illustrate this result and the assumptions necessary for its derivation are discussed.**

I. INTRODUCTION

Over the past decades, microfabrication of electronic devices such as micro-electro-mechanical systems (MEMS) has considerably developed. As their characteristic dimensions become smaller, these devices become increasingly afflicted with dispersion and become increasingly sensitive to changes in their operating conditions. Typical sources of dispersions and uncertainty are variations in the fabrication process or environmental disturbances such as temperature, pressure and humidity fluctuations [1]. It is then usually not possible to guarantee *a priori* that a given device will work properly under all operating conditions, and expensive tests must be run before the commercialization decision is made. To cut these costs, it is desirable to implement self-test (and self-tuning) features such as parameter estimation routines, so that devices can compensate the variations in the fabrication process and adapt to changing conditions.

Unfortunately, standard identification methods based on parameter estimation [2], [3] do not lend themselves easily to implementation at a microscopic scale. Their integration requires the implementation of high-resolution analog-to-digital converters (ADCs), which may require long design times and result in large silicon areas and increased power consumption. On the other hand, parameter estimation routines based on binary observations are very attractive since they only involve the integration of a 1-bit ADC [4], which requires minimal design and results in minimal silicon area and power consumption and, consequently, in minimal added costs.

Several identification methods based on binary or roughly quantized observations can be found in the literature [5], [6], [7], [8], [1], [9], [10], [11], [12], [13], [14]. For example, Wigren [5], [6] has developed a least-mean-square (LMS) approach to the problem of online parameter estimation from quantized observations, based on an approximation of the quantizer. The proof of convergence, which uses the ordinary differential equation (ODE) approach [15], relies on the assumption that at least one threshold of the quantizer is known and different from zero. Under these hypotheses, it is possible to guarantee the asymptotic convergence of this method to the nominal parameters. Wang and his co-authors [7], [8] have considered that the unknown system is excited by a periodic signal and the threshold of the quantizer is randomly specified by a dithering signal. They have proved that the cumulative distribution function of the threshold does not have to be known *a priori* and can be estimated simultaneously with the system parameters. More recently, Jafari and his co-authors [9] have studied a recursive identification method which does not rely on a pseudo-gradient of a least-squares criterion and requires neither a known non-zero threshold value, nor a varying threshold. This online LMS-like identification method based on binary observations (LIMBO) has little storage requirements and low computational complexity. Although LIMBO has already been put in practice for testing MEMS sensors [16], the convergence of this method has so far only been established in the case when no noise exists at the input of the quantizer. It is interesting to note that, in this noise-free context, LIMBO is similar to the relaxation method proposed and studied in [17], [18] for solving consistent sets of linear inequalities.

In this paper, we analyze the convergence of LIMBO in a more general context: we suppose an unknown measurement noise is present at the input of the quantizer and study the influence of this noise on the convergence of the method. More specifically, the convergence rate of the method is investigated and a lower bound of the expected value of the correlation coefficient between the nominal and the estimated system parameters is found and expressed as a function of the variance of the measurement noise. It is shown that the derived lower bound can be safely used as an accurate prediction of the expected value of the correlation coefficient. The structure of the article is the following. In section II, the system and its model are introduced. In section III, the LIMBO algorithm is derived under its general form. Then, the convergence of the proposed algorithm is studied in section IV and graphically illustrated in section V. Finally, concluding remarks and perspectives are given in section VI.

L. Bourgois is with Supélec E3S, 3 rue Joliot-Curie, 91192 Gif-sur-Yvette Cedex, France. laurent.bourgois@supelec.fr

J. Juillard is with Supélec E3S, 3 rue Joliot-Curie, 91192 Gif-sur-Yvette Cedex, France. jerome.juillard@supelec.fr

II. FRAMEWORK AND NOTATIONS

Let us consider the system illustrated in Fig. 1.

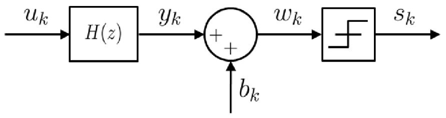

Fig. 1. Block diagram of the studied system.

The input signal u_k is filtered by a linear time-invariant discrete-time system H to produce the system output y_k, where subscript k denotes the discrete time. We assume that the transfer function has a finite impulse response of length L, *i.e.* the impulse response can be represented by a column vector $\boldsymbol{\theta} = (\theta_l)_{l=1}^{L}$. Consequently, the scalar value of the system output at time k is given by $y_k = \boldsymbol{\theta}^T \varphi_{k,L}$ where $\varphi_{k,L} = (u_l)_{l=k}^{k-L+1}$ is the regression column vector of dimension L. The system output is then measured via a 1-bit ADC so that only its sign $s_k = S\,(y_k + b_k)$ is available at time k. Here, b_k corresponds to the additive measurement noise at time k, and the function S of a real number x is characterized by:

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

Our purpose is to develop a recursive estimation method to find a good estimate of the parameter vector $\boldsymbol{\theta}$, starting from N observations of the binary output knowing the input. Let $\hat{\boldsymbol{\theta}}_k$ be the estimated parameter vector at time k. Let us also introduce $\hat{y}_k = \hat{\boldsymbol{\theta}}_k^T \varphi_{k,L}$ the estimated system output at time k and $\hat{s}_k = S\,(\hat{y}_k)$. Without loss of generality, we suppose that $\|\boldsymbol{\theta}\|^2 = 1$.

III. PROPOSED LMS APPROACH

The non-relaxed LIMBO method [9] consists in the following iteration:

if $s_k \neq \hat{s}_k$

$$\hat{\boldsymbol{\theta}}_{k+1} = \hat{\boldsymbol{\theta}}_k - 2\hat{y}_k \frac{\varphi_{k,L}}{\left\|\varphi_{k,L}\right\|^2} \tag{2}$$

else

$$\hat{\boldsymbol{\theta}}_{k+1} = \hat{\boldsymbol{\theta}}_k$$

Or, more compactly:

$$\hat{\boldsymbol{\theta}}_{k+1} = \hat{\boldsymbol{\theta}}_k - 2\hat{y}_k \frac{\varphi_{k,L}}{\left\|\varphi_{k,L}\right\|^2} [s_k \neq \hat{s}_k] \tag{3}$$

In this compacted expression, the notation $[s_k \neq \hat{s}_k]$ stands for a variable that is equal to unity if $s_k \neq \hat{s}_k$, *i.e.* if $y_k + b_k$ and $\hat{y}_k$ have opposite signs, and equal to zero otherwise. This non-relaxed iteration ensures that the norm of $\hat{\boldsymbol{\theta}}_k$ remains constant. One may then assume that $\left\|\hat{\boldsymbol{\theta}}_k\right\| = 1$.

Next, by projecting (3) onto the nominal parameter vector $\boldsymbol{\theta}$ and considering the sequence $\nu_k = \hat{\boldsymbol{\theta}}_k^T \boldsymbol{\theta}$, we obtain:

$$\nu_{k+1} = \nu_k - \frac{2\hat{y}_k y_k}{\left\|\varphi_{k,L}\right\|^2} [s_k \neq \hat{s}_k] \tag{4}$$

Note that ν_k is the cosine of the angle made by $\hat{\boldsymbol{\theta}}_k$ and $\boldsymbol{\theta}$ since both vectors are normalized, and we have $-1 \leq \nu_k \leq 1$, so that $\lim_{k\to\infty} \hat{\boldsymbol{\theta}}_k = \boldsymbol{\theta}$ is equivalent to $\lim_{k\to\infty} \nu_k = 1$.

IV. CONVERGENCE ANALYSIS IN THE PRESENCE OF NOISE

In [9], the almost sure convergence of the algorithm presented in the previous section is demonstrated under some specific assumptions. In particular, the proof is established for a relaxed version of the algorithm by supposing $b_k = 0$ (although a proof in the noise-free non-relaxed case could also be obtained by following the approach in [18]). Our purpose here is to study the convergence of LIMBO, without additional relaxation step, and taking into account measurement noise. To this end, we aim to evaluate the conditional expectation $\mathrm{E}\,(\nu_{k+1}|\nu_k)$ under the three following assumptions:

- y_k and $\hat{y}_k$ are two centered Gaussian random variables.
- u_k is white and centered, with a Bernoulli distribution and takes two values: 1 or -1.
- b_k is white and centered, with a uniform distribution in the interval $[-\beta, \beta]$ where $\beta > 0$. Furthermore, b_k is independent of y_k and $\hat{y}_k$.

The last two assumptions are made to keep the calculations which follow as simple and straightforward as possible and should not be seen as strict limitations to the validity of our results. First of all, note that the first assumption is verified in practice regardless of the distribution of the input signal, provided the impulse response $\boldsymbol{\theta}$ does not vanish too quickly, as is the case in many applications (for a more detailed discussion on the validity of this assumption, please refer to [19], [20]). By construction, ν_k corresponds to the correlation coefficient between the variables y_k and $\hat{y}_k$ whose means are equal to zero and whose variances are equal to one. In this case, their joint probability density function is defined for any $-1 < \nu_k < 1$ by:

$$f_1\,(y_k, \hat{y}_k) = \frac{1}{2\pi\sqrt{1-\nu_k^2}} \exp\left[-\frac{y_k^2 + \hat{y}_k^2 - 2\nu_k y_k \hat{y}_k}{2\left(1-\nu_k^2\right)}\right] \tag{5}$$

The reason for assuming a binary input, *i.e.* $u_k \in \{-1, 1\}$, is that this simplifies (3) and (4), because in that case, by construction, $\left\|\varphi_{k,L}\right\|^2 = L$. As already mentioned, this has little influence on the Gaussian nature of y_k and $\hat{y}_k$ in practical cases [19].

Now (4) can be rewritten as:

$$\nu_{k+1} = \nu_k - \frac{2}{L} [s_k \neq \hat{s}_k]\, y_k \hat{y}_k \tag{6}$$

The probability density function of b_k is defined by:

$$f_2(b_k) = \begin{cases} \frac{1}{2\beta} & \text{if } -\beta \leq b_k \leq \beta \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

Although the calculations which follow can be conducted with other measurement noise distributions, bounded or not, they are made much simpler by assuming a distribution with a compact support.

Taking the conditional expectation of (6) yields:

$$\mathrm{E}(\nu_{k+1}|\nu_k) = \nu_k - \frac{2}{L} \underbrace{\int_{y_k=-\infty}^{+\infty} \int_{\hat{y}_k=-\infty}^{+\infty} \int_{b_k=-\beta}^{\beta} \frac{1}{2\beta} [S(y_k + b_k) \neq S(\hat{y}_k)]\, y_k \hat{y}_k\, f_1(y_k, \hat{y}_k)\, \mathrm{d}b_k\, \mathrm{d}\hat{y}_k\, \mathrm{d}y_k}_{=I} \quad (8)$$

Let us focus on the integral over b_k. To this end, we define the following function:

$$F(y_k, \hat{y}_k) = \int_{b_k=-\beta}^{\beta} [S(y_k + b_k) \neq S(\hat{y}_k)] \frac{\mathrm{d}b_k}{2\beta} \quad (9)$$

The function $F(y_k, \hat{y}_k)$ is graphically represented in Fig. 2 in the two cases $\hat{y}_k > 0$ and $\hat{y}_k < 0$.

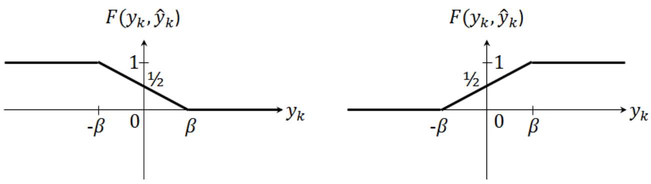

Fig. 2. $F(y_k, \hat{y}_k)$ when $\hat{y}_k > 0$ (left) and $\hat{y}_k < 0$ (right).

We may synthetically sum this up as:

$$F(y_k, \hat{y}_k) = G(y_k, \hat{y}_k) + S(\hat{y}_k)\, T(y_k) \quad (10)$$

where $G(y_k, \hat{y}_k) = [S(y_k) \neq S(\hat{y}_k)]$ and

$$T(y_k) = \begin{cases} -(y_k + \beta)\frac{1}{2\beta} & \text{if } y_k \in [-\beta, 0] \\ -(y_k - \beta)\frac{1}{2\beta} & \text{if } y_k \in [0, \beta] \\ 0 & \text{otherwise} \end{cases} \quad (11)$$

are represented in Fig. 3.

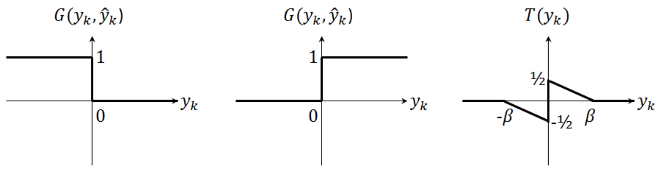

Fig. 3. $G(y_k, \hat{y}_k)$ when $\hat{y}_k > 0$ (left) and $\hat{y}_k < 0$ (center) and $T(y_k)$ (right).

Thus, the triple integral I defined in (8) can be expressed as the sum $I = I_1 + I_2$ where:

$$I_1 = \int_{y_k=-\infty}^{+\infty} \int_{\hat{y}_k=-\infty}^{+\infty} y_k \hat{y}_k\, G(y_k, \hat{y}_k)\, f_1(y_k, \hat{y}_k)\, \mathrm{d}\hat{y}_k\, \mathrm{d}y_k \quad (12)$$

And:

$$I_2 = \int_{y_k=-\infty}^{+\infty} \int_{\hat{y}_k=-\infty}^{+\infty} y_k \hat{y}_k\, S(\hat{y}_k)\, T(y_k)\, f_1(y_k, \hat{y}_k)\, \mathrm{d}\hat{y}_k\, \mathrm{d}y_k \quad (13)$$

Let us consider first the double integral I_1. By breaking both integrals into positive and negative parts, the following expression is obtained:

$$I_1 = 2 \int_{y_k=0}^{+\infty} \int_{\hat{y}_k=-\infty}^{0} y_k \hat{y}_k\, f_1(y_k, \hat{y}_k)\, \mathrm{d}\hat{y}_k\, \mathrm{d}y_k \quad (14)$$

for which an analytical expression is found by a cartesian to polar coordinate transformation:

$$I_1 = \frac{\nu_k \arccos(\nu_k) - \sqrt{1 - \nu_k^2}}{\pi} \quad (15)$$

Now consider the double integral I_2. By breaking the integral over $\hat{y}_k$ into positive and negative parts and noting that T is odd, the following relation can be established:

$$I_2 = 2 \int_{y_k=-\beta}^{\beta} \int_{\hat{y}_k=0}^{+\infty} y_k \hat{y}_k\, T(y_k)\, f_1(y_k, \hat{y}_k)\, \mathrm{d}\hat{y}_k\, \mathrm{d}y_k \quad (16)$$

An analytical expression of the integral over $\hat{y}_k$ can also be obtained, which yields:

$$I_2 = \int_{y_k=0}^{\beta} \frac{y_k(\beta - y_k)\sqrt{1 - \nu_k^2}}{\beta\pi} \exp\left(-\frac{y_k^2}{2\left(1 - \nu_k^2\right)}\right) \mathrm{d}y_k + \int_{y_k=0}^{\beta} \frac{y_k^2(\beta - y_k)\,\nu_k}{\beta\sqrt{2\pi}} \exp\left(-\frac{y_k^2}{2}\right) \mathrm{erf}\left(\frac{\nu_k y_k}{\sqrt{2\left(1 - \nu_k^2\right)}}\right) \mathrm{d}y_k \quad (17)$$

Finally, (15) and (17) are introduced into (8) to derive the conditional expectation:

$$\mathrm{E}\left(\nu_{k+1}|\nu_k\right) \quad = \quad \nu_k - \frac{2}{L}\left(I_1 + I_2\right) \tag{18}$$

or equivalently, writing the right-hand side of (18) as $f\left(\nu_k\right)$, we have, $\forall k$:

$$\mathrm{E}\left(\nu_{k+1}|\nu_k\right) \quad = \quad f\left(\nu_k\right) \tag{19}$$

Taking the expected value of (19) then yields:

$$\mathrm{E}\left(\nu_{k+1}\right) \quad = \quad \mathrm{E}\left(f\left(\nu_k\right)\right) \tag{20}$$

Now, provided f is convex, Jensen's inequality can be applied to get:

$$\mathrm{E}\left(f\left(\nu_k\right)\right) \quad \geq \quad f\left(\mathrm{E}\left(\nu_k\right)\right) \tag{21}$$

Since f is (infinitely) continuously differentiable, the best way to prove that the function is convex is to show that $f^{(2)}\left(\nu_k\right) \geq 0$ for all ν_k in $]-1,1[$. An analytical expression of this second derivative can be established as follows:

$$\begin{aligned} f^{(2)}\left(\nu_k\right) \quad = \quad & \frac{2\sqrt{2\pi}}{L\beta\pi}\,\mathrm{erf}\left(\frac{\beta}{\sqrt{2\left(1-\nu_k^2\right)}}\right) \\ - \quad & \frac{2}{L\pi\sqrt{1-\nu_k^2}}\exp\left(-\frac{\beta^2}{2\left(1-\nu_k^2\right)}\right) \end{aligned} \tag{22}$$

To study the monotony of $f^{(2)}\left(\nu_k\right)$, we check the sign of its derivative, for which an analytical expression can also be calculated:

$$f^{(3)}\left(\nu_k\right) \quad = \quad \frac{2\nu_k\left(1-\nu_k^2+\beta^2\right)}{L\pi\left(1-\nu_k^2\right)^{5/2}}\exp\left(-\frac{\beta^2}{2\left(1-\nu_k^2\right)}\right) \tag{23}$$

In the interval $]-1,1[$, the unique zero of $f^{(3)}\left(\nu_k\right)$ is obtained when $\nu_k = 0$ and the third derivative is negative whenever $\nu_k < 0$ and positive whenever $\nu_k > 0$. Thus, the minimum of $f^{(2)}\left(\nu_k\right)$ is obtained for $\nu_k = 0$ and we have:

$$\min_{-1<\nu_k<1}\left(f^{(2)}\left(\nu_k\right)\right) \quad = \quad \frac{2}{L\pi}\left[\frac{\sqrt{2\pi}}{\beta}\,\mathrm{erf}\left(\frac{\beta}{\sqrt{2}}\right) - \exp\left(-\frac{\beta^2}{2}\right)\right] \tag{24}$$

Now, since $\beta > 0$ by hypothesis, it is straightforward to show that the minimum in (24) is positive by studying the monotony of its product by β. Consequently, $f^{(2)}\left(\nu_k\right) \geq 0$ and f is convex.

Finally, (20) and (21) are gathered to yield:

$$\mathrm{E}\left(\nu_{k+1}\right) \quad \geq \quad f\left(\mathrm{E}\left(\nu_k\right)\right) \tag{25}$$

At this point, we aim to find an upper bound for I_2. We proceed in two steps. First, provided $y_k \geq 0$, we have:

$$\nu_k\,\mathrm{erf}\left(\frac{\nu_k y_k}{\sqrt{2\left(1-\nu_k^2\right)}}\right) \quad \leq \quad \frac{2}{\sqrt{\pi}}\left(\frac{\nu_k^2 y_k}{\sqrt{2\left(1-\nu_k^2\right)}}\right) \tag{26}$$

Then, we notice that the exponentials in (17) are less or equal to unity on $[0,\beta]$. Hence, I_2 is bounded from above by a sum of two integrals of polynomials which can easily be calculated, to yield the following inequality:

$$I_2 \quad \leq \quad \frac{\beta^2\sqrt{1-\nu_k^2}}{6\pi} + \frac{\beta^4\nu_k^2}{20\pi\sqrt{1-\nu_k^2}} \tag{27}$$

And the following relation can be established:

$$\begin{aligned} f\left(\mathrm{E}\left(\nu_k\right)\right) \quad \geq \quad & \mathrm{E}\left(\nu_k\right) \\ - \quad & \frac{2}{L}\left[\frac{\mathrm{E}\left(\nu_k\right)\arccos\left(\mathrm{E}\left(\nu_k\right)\right) - \sqrt{1-\mathrm{E}\left(\nu_k\right)^2}}{\pi}\right] \\ - \quad & \frac{2}{L}\left[\frac{\beta^2\sqrt{1-\mathrm{E}\left(\nu_k\right)^2}}{6\pi} + \frac{\beta^4\,\mathrm{E}\left(\nu_k\right)^2}{20\pi\sqrt{1-\mathrm{E}\left(\nu_k\right)^2}}\right] \end{aligned} \tag{28}$$

or equivalently, writing the right-hand side of (28) as $\tilde{f}\left(\mathrm{E}\left(\nu_k\right)\right)$, we have, $\forall k$:

$$f\left(\mathrm{E}\left(\nu_k\right)\right) \quad \geq \quad \tilde{f}\left(\mathrm{E}\left(\nu_k\right)\right) \tag{29}$$

Now, this more convenient lower bound of $\mathrm{E}\left(\nu_{k+1}\right)$ can be used to get:

$$\mathrm{E}\left(\nu_{k+1}\right) \quad \geq \quad \tilde{f}\left(\mathrm{E}\left(\nu_k\right)\right) \tag{30}$$

so that $\forall k$, $\mathrm{E}\left(\nu_k\right) \geq \mu_k$, where the sequence μ_k is defined by:

$$\left\{\begin{array}{lcl} \mu_1 & = & \mathrm{E}\left(\nu_1\right) \\ \mu_{k+1} & = & \tilde{f}\left(\mu_k\right) \end{array}\right. \tag{31}$$

Assuming $\mu_1 = 0$ (which corresponds to the case of a randomly chosen initial vector $\hat{\boldsymbol{\theta}}_1$) and $\beta^2 > 0$, it can be shown that the sequence μ_k increases monotonously. Since it is also bounded from above (by 1), it converges to μ_∞ such that $\mu_\infty = \tilde{f}\left(\mu_\infty\right)$. This shows that $\mathrm{E}\left(\nu_k\right)$ is at least equal to μ_k. This result is illustrated in the next section. Note that the limiting case $\beta^2 = 0$ corresponds to $\mu_\infty = 1$, which implies the convergence in the mean of the non-relaxed LIMBO method to the nominal system parameters.

V. RESULTS AND DISCUSSION

In this section, the convergence of the algorithm is graphically illustrated. The objective of this work is to show that the bound μ_k derived in section IV under some rather stringent hypotheses is in fact a good approximation of $\mathrm{E}(\nu_k)$. Furthermore, we aim to show that our hypotheses can be relaxed and our results extended to more general cases.

For that, we consider a set of 50 realizations of the binary input signal. Based on these 50 realizations, the empirical mean of ν_k is calculated and compared to the sequence μ_k, for different values of noise variance. $\boldsymbol{\theta}$ is a randomly generated impulse response of length $L = 50$. The identification procedure detailed in section III is applied starting from $N = 10^5$ observations of the binary output. The additive noise is uniformly distributed in the interval $[-\beta, \beta]$. The value of β changes between 10^{-3} and 10^0. Thus, the signal-to-noise ratio (SNR) lies in average between 64.77 dB, *i.e.* an almost absence of noise, and 4.77 dB.

The results corresponding to these conditions are represented in Fig. 4, which displays the empirical estimate of the quantity $1 - \mathrm{E}(\nu_k)$ for each β. Let us bear in mind that this specific quantity corresponds to the quality of the online estimation.

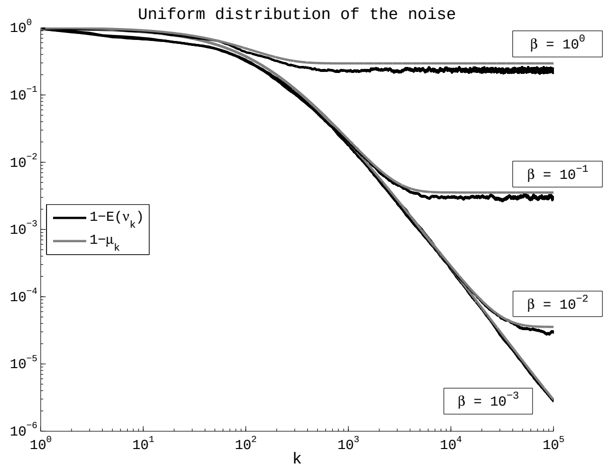

Fig. 4. Empirical estimate of $1 - \mathrm{E}(\nu_k)$ and $1 - \mu_k$ for various values of β by considering an uniform distribution of the noise.

The numerical simulations comfort our theoretical analysis and show that the upper bound given by $1 - \mu_k$ accurately predicts the value of $\mathrm{E}(\nu_k)$. Furthermore, this result seems to hold for many other distributions of the measurement noise, provided they are centered. For a given distribution with variance σ^2, it suffices to replace β^2 in (28) by $3\sigma^2$ to derive the corresponding analytical bound. This point is illustrated in Fig. 5.

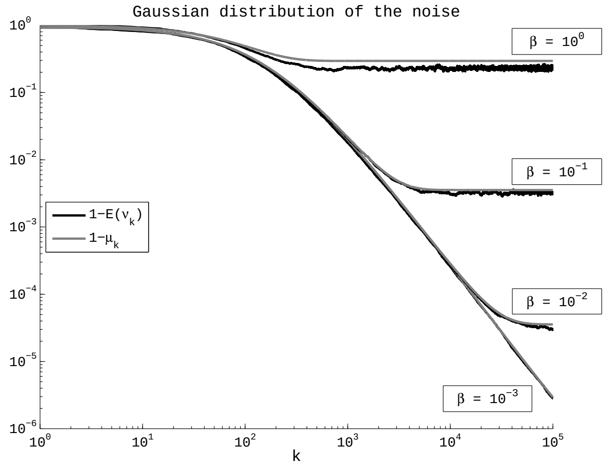

Fig. 5. Empirical estimate of $1 - \mathrm{E}(\nu_k)$ and $1 - \mu_k$ for various values of β by considering a Gaussian distribution of the noise of variance $\sigma^2 = \beta^2/3$.

VI. CONCLUSION

In this paper, we extended the analysis of the LIMBO method [9] to a more general context involving measurement noise and no relaxation step. We demonstrated the convergence in the mean of the non-relaxed version of the method, in the absence of noise. In the presence of noise, a lower bound of the correlation coefficient between the estimated and nominal parameters was analytically derived and verified by simulations. This lower bound is useful for predicting the convergence rate of the method. We also showed that the simplifying assumptions made in our demonstration could probably be relaxed. The variance of the correlation coefficient could probably be studied following the same lines, as well as the convergence rate of the relaxed version of LIMBO. This would be useful to determine some optimal relaxation strategies in the presence of measurement noise. Finally, it is interesting to note that an experimental application of LIMBO, in which the tested MEMS device was a micro-wire used as a heating resistor inserted in a Wheatstone bridge, had already been successfully developed in [16].

References

[1] 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, issue 1, 2010, pp. 148-152.

[2] E. Walter and L. Pronzato, *Identification of Parametric Models from Experimental Data*, Springer, 1997.

[3] L. Ljung, *System Identification: Theory for the User*, Prentice Hall, 1999.

[4] R. Von de Plassche, *CMOS Integrated Analog-to-Digital and Digital-to-Analog Converters*, Springer-Verlag, 2005.

[5] T. Wigren, ODE Analysis and Redesign in Blind Adaptation, *IEEE Transactions on Automatic Control*, vol. 42, issue 12, 1997, pp. 1742-1747.

[6] T. Wigren, Adaptive Filtering Using Quantized Output Measurements, *IEEE Transactions on Automatic Control*, vol. 46, issue 12, 1998, pp. 3423-3426.

[7] L. Wang, J. Zhang and G. Yin, System Identification Using Binary Sensors, *IEEE Transactions on Automatic Control*, vol. 48, issue 11, 2003, pp. 1892-1907.

[8] L. Wang, G. Yin and J. Zhang, Joint Identification of Plant Rational Models and Noise Distribution Functions Using Binary-Valued Observations, *Automatica*, vol. 42, issue 4, 2006, pp. 535-547.

[9] K. Jafari, J. Juillard and M. Roger, Convergence of an Online Approach to Parameter Estimation Problems Based on Binary Observations, *To be published in Automatica,* 2012.

[10] K. Jafari, J. Juillard and E. Colinet, ”A Recursive System Identification Method Based on Binary Measurements”, *in 49th IEEE Conference on Decision and Control*, Atlanta, Georgia (USA), 2010, pp. 1154-1158.

[11] M. Negreiros, L. Carro and A. Susin, ”Ultimate Low Cost Analog BIST”, *in 40th Design Automation Conference*, New York, New York (USA), 2003, pp. 570-573.

[12] E. Rafajlowicz, Linear Systems Identification from Random Threshold Binary Data, *IEEE Transactions on Signal Processing*, vol. 44, issue 8, 1996, pp. 2064-2070.

[13] E.-W. Bai and J. Reyland Jr., Towards Identification of Wiener Systems with the Least Amount of a priori Information on the Nonlinearity, *Automatica*, vol. 44, issue 4, 2008, pp. 910-919.

[14] F. Gustafsson and R. Karlsson, Statistical Results for System Identification Based on Quantized Observations, *Automatica*, vol. 45, issue 12, 2009, pp. 2794-2801.

[15] L. Ljung, Analysis of Recursive Stochastic Algorithms, *IEEE Transactions on Automatic Control*, vol. 22, issue 4, 1977, pp. 551-575.

[16] O. Legendre, H. Bertin, O. Garel, H. Mathias, S. Megherbi, K. Jafari, J. Juillard and E. Colinet, ”A Low-Cost, Built-In Self-Test Method for Resistive MEMS Sensors”, *in Eurosensors XXV (Eurosensors’11)*, Athens, Greece, 2011, pp. 182-185.

[17] S. Agmon, The Relaxation Method for Linear Inequalities, *Canadian Journal of Mathematics*, vol. 6, issue 3, 1954, pp. 382-392.

[18] T. S. Motzkin and I. J. Schoenberg, The Relaxation Method for Linear Inequalities, *Canadian Journal of Mathematics*, vol. 6, issue 3, 1954, pp. 393-404.

[19] J. Juillard and E. Colinet, ”Initialization of the BIMBO Self-Test Method using Binary Inputs and Outputs”, *in 46th IEEE Conference on Decision and Control*, New Orleans, Louisiana (USA), 2007, pp. 578-583.

[20] J. C. Taylor, *An Introduction to Measure and Probability*, Springer, 1997.