

# Recursive Estimation of ARX Systems Using Binary Sensors with Adjustable Thresholds

Balázs Csanád Csáji<sup>\*,\*\*</sup> Erik Weyer<sup>\*</sup>

<sup>\*</sup> Department of Electrical and Electronic Engineering, The University of Melbourne, Parkville, Melbourne, Victoria, 3010, Australia, email: [bcsaji@unimelb.edu.au](mailto:bcsaji@unimelb.edu.au) and [ewey@unimelb.edu.au](mailto:ewey@unimelb.edu.au)

<sup>\*\*</sup> Computer and Automation Research Institute (SZTAKI), Hungarian Academy of Sciences, Kende utca 13-17, H-1111, Budapest, Hungary

---

**Abstract** We consider the identification of ARX systems which are observed via a *binary* sensor. Previous solutions typically assumed the complete knowledge of the noise distribution and that the inputs can be chosen by the user. Here, we only make mild assumptions on the noise and we make no assumptions that we can control the inputs. However, we assume that we can choose the *threshold* of the binary sensor or, equivalently, we can apply a *dither* signal to the system. We propose two recursive algorithms for this problem. The first one is based on an FIR approximation of the ARX system and requires post-processing. We prove that it provides a *strongly consistent* estimator. The other algorithm estimates the parameters directly, without post-processing, by simultaneously estimating the parameters and the outputs of the system. Numerical experiments which show that both algorithms work effectively are also presented.

---

## 1. INTRODUCTION

Estimating parameters of dynamical systems based on noisy observations is one of the fundamental problems of system identification (Goodwin and Sin [1984], Ljung [1999]). In some situations only *quantized* measurements are available, due to, e.g., limited sensor capacities or signal coding. Quantized variants of estimation problems were studied, for example, by Colinet and Juillard [2010], Wang et al. [2010], Casini et al. [2011], Godoy et al. [2011], but, with a few exceptions (Gerencsér and Vágó [1999], Weyer et al. [2009]), only for the case when the characteristics of the noise are *known*, e.g., its cumulative distribution function (CDF) is invertible and known to the user and its density function is continuously differentiable.

However, making strong assumptions on the noise is undesirable as the noise characteristics are often unknown.

Here, we try to reduce the statistical assumptions on the system and do not presuppose the knowledge of the noise distribution or that we can choose the inputs. We follow our previous framework (Csáji and Weyer [2011]) and assume that the threshold of the binary sensor can be changed at each time step or, equivalently, we can apply a carefully chosen dither signal to the system.

We propose two recursive estimators which, as we will see, can also be seen as “sign-error” type *stochastic approximation* algorithms (Chen [2009]). The idea of Algorithm I is to approximate the system with an FIR model and estimate the ARX parameters via a post-processing step. We will show that this estimator is strongly consistent and its error is asymptotically normally distributed.

Our other proposed recursive estimator, Algorithm II, does not require any post-processing, it directly estimates the true parameters by simultaneously estimating the outputs

(the hidden states of the system), too. We will also present results of numerical experiments which demonstrate the effectiveness of both recursive estimation methods.

### 1.1 Outline

After this introductory Section 1, in Section 2 we precisely define the problem and state our general assumptions. Then, in Section 3, we give the fundamental form of the solution algorithms, describe in details the two recursive estimators and prove that the first one is strongly consistent. In Section 4 we present numerical experiments on a 2nd-order ARX system which demonstrate that both algorithms work well. Finally, Section 5 summarizes the paper and highlights some further research directions.

### 1.2 Notations

If  $\dots, X_{t-1}, X_t, X_{t+1}, \dots$  is a (one- or two-sided) sequence (e.g., of random variables), the whole sequence is denoted by  $(X_t)_t$ . Symbol  $\mathbb{R}$  stands for the field of real numbers, while  $\mathbb{P}$  denotes probability and  $\mathbb{E}$  denotes expected value. Variables with capital letters usually indicate random variables, while variables with small letters are deterministic (constants). Symbol  $\triangleq$  represents a defining equation, i.e., the left hand side is defined to be equal to the right hand side. Term “independent and identically distributed” is abbreviated by “i.i.d.” and  $(\cdot)^T$  denotes matrix transpose. Function  $\mathbb{I}(\cdot)$  denotes the indicator function, whose value is 1 if its argument (a formula) is true and 0 otherwise.

## 2. PRELIMINARIES

In this section, we introduce the problem of binary identification of ARX models as well as present the general form of the proposed recursive estimation algorithms.

### 2.1 Problem Setting

Let us consider the problem of identifying the parameters of an ARX system based on binary observations:

$$X_t \triangleq \sum_{i=1}^p a_i^* X_{t-i} + \sum_{i=1}^q b_i^* U_{t-i} + N_t, \quad (1)$$

$$Y_t \triangleq \mathbb{I}(X_t \leq C_t), \quad (2)$$

where  $X_t$ ,  $U_t$  and  $N_t$  denote the output (hidden state), the input and the noise at time  $t$ . The user can only *observe* the inputs,  $(U_t)_t$ , and the *binary* outputs,  $(Y_t)_t$ , while the thresholds of the binary sensor,  $(C_t)_t$ , can be *controlled* at each time-step  $t$ . The aim is to identify (estimate) the “true” parameters,  $\theta^* = (a_1^*, \dots, a_p^*, b_1^*, \dots, b_q^*) \in \mathbb{R}^{p+q}$ .

The binary output can be rewritten as follows

$$Y_t = \mathbb{I}(\varphi_t^T \theta^* + N_t \leq C_t), \quad (3)$$

where the (random) vector  $\varphi_t \in \mathbb{R}^{p+q}$  is defined as

$$\varphi_t = (X_{t-1}, \dots, X_{t-p}, U_{t-1}, \dots, U_{t-q}). \quad (4)$$

Note that “choosing the threshold” can also be seen as a special form of “dithering”, namely, adding a carefully chosen noise term to the output  $X_t$ , i.e.,

$$Y_t = \mathbb{I}(X_t - C_t \leq 0) = \mathbb{I}(\varphi_t^T \theta^* + N_t - C_t \leq 0), \quad (5)$$

therefore, changing the threshold is equivalent to using a system with a fixed threshold, but applying an additional term,  $-C_t$ , at each time step  $t$ . The two perspectives on the problem are illustrated by Figure 1.

### 2.2 System Assumptions

In this paper, we will apply the following assumptions on the system, the (observed) inputs and the (hidden) noises:

*Assumption 1.*  $(N_t)_t$  is i.i.d., continuous, zero mean, zero median, has a finite variance:  $\sigma_n^2 \triangleq \mathbb{E}[N_t^2] < \infty$ , and has a continuous and positive density at zero (thus  $\sigma_n^2 > 0$ ).

*Assumption 2.*  $(U_t)_t$  is i.i.d., zero mean,  $(U_t)_t$  and  $(N_t)_t$  are independent, and  $0 < \sigma_u^2 < \infty$ , where  $\sigma_u^2 \triangleq \mathbb{E}[U_t^2]$ .

*Assumption 3.* The system is stable, namely, the roots of the polynomial  $A^*(z)$  lie strictly inside the unit circle; additionally, the transfer function  $B^*(z)/A^*(z)$  is irreducible, i.e., there are no pole-zero cancellations, where

$$A^*(z) \triangleq 1 - a_1^* z^{-1} - a_2^* z^{-2} - \dots - a_p^* z^{-p}, \quad (6)$$

$$B^*(z) \triangleq b_1^* z^{-1} + b_2^* z^{-2} + \dots + b_q^* z^{-q}, \quad (7)$$

where  $z^{-1}$  is the backward shift operator,  $z^{-i} x_t \triangleq x_{t-i}$ . We also assume that the orders  $p$  and  $q$  are known.

### 2.3 Underlying Probability Space

Let  $(\Omega, \mathcal{F}, \mathbb{P})$  be the underlying *probability space*, where  $\Omega$  is the sample space,  $\mathcal{F}$  is the  $\sigma$ -algebra of events and  $\mathbb{P}$  is the measure. We denote the  $\sigma$ -algebra generated by the past of the process up to time  $t \geq 0$  by  $\mathcal{F}_t \subseteq \mathcal{F}$ ,

$$\mathcal{F}_t \triangleq \sigma \left\{ N_t, N_{t-1}, \dots, U_t, U_{t-1}, \dots, \hat{\theta}_t, \dots, \hat{\theta}_0 \right\}, \quad (8)$$

where  $(\mathcal{F}_t)_{t \in T}$  is a *filtration*, viz., a nondecreasing sequence of  $\sigma$ -fields. Random variables  $(Y_k)_{k=0}^t$  are  $\mathcal{F}_t$  measurable and we only need to include  $(\hat{\theta}_t)_{t > 0}$  in (8) if want to apply some additional randomization for the estimates.

## 3. RECURSIVE ESTIMATION

In this paper we propose two recursive identification methods (Ljung and Söderström [1987]) for finding  $\theta^*$ . First, we will define the general form of the algorithms, then, we will specify the step-sizes and the expanding truncation of the iterations. Finally, we will present and analyze the proposed recursive estimation algorithms.

Algorithm I applies a FIR approximation of the system. This method requires post-processing in the form of a matrix inversion, in order to get an estimate of parameter  $\theta^*$ . We will prove that, under the assumptions above, Algorithm I provides a *strongly consistent* estimator.

Algorithm II does not require any post-processing, it works by simultaneously recursively estimating  $\theta^*$  and  $\varphi_t$ .

Both algorithms generate a sequence,  $(\hat{\theta}_t)_t$ , such that at time  $t$  either  $\hat{\theta}_t$  itself is the estimate of  $\theta^*$  (Algorithm II) or there is a measurable function  $f(\cdot)$ , such that  $f(\hat{\theta}_t)$  is the estimate of  $\theta^*$  (Algorithm I). In both algorithms, the threshold of the sensor at time  $t$  is set according to  $C_t = \hat{\varphi}_t^T \hat{\theta}_t$ , where  $\hat{\varphi}_t$  is an appropriately chosen regressor.

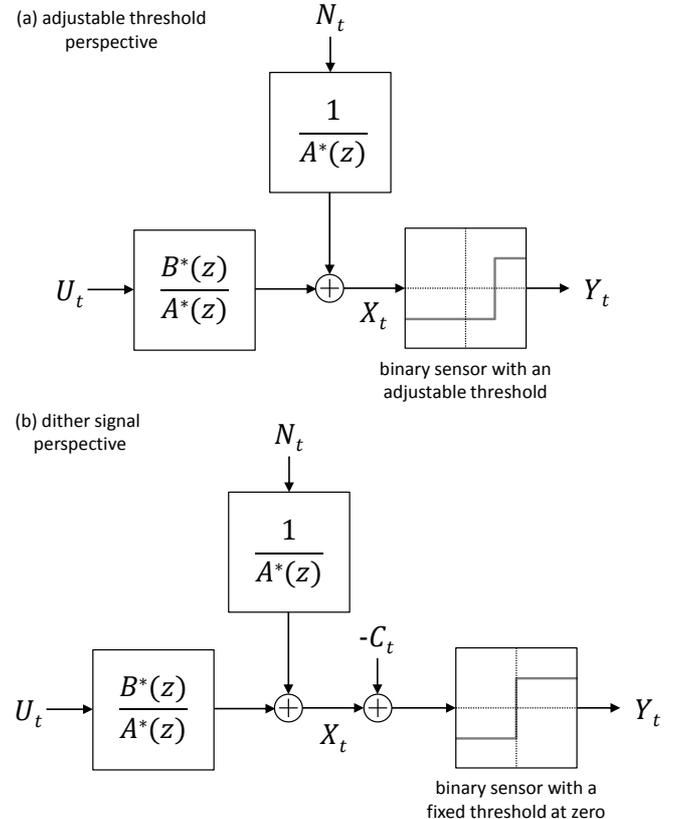


Figure 1. The perspectives of (a) applying an adjustable threshold and (b) inserting a dither signal.

### 3.1 General Form of the Algorithms

The general form of both proposed algorithms is

$$\hat{\theta}_{t+1} = \Pi_{M_{\mu(t)}} \left[ \hat{\theta}_t + \alpha_t \hat{\varphi}_t \left( 1 - 2\mathbb{I}(X_t \leq \hat{\varphi}_t^T \hat{\theta}_t) \right) \right], \quad (9)$$

where  $\hat{\varphi}_t$  is a regression vector defined differently in the two algorithms,  $(\alpha_t)_t$  is a sequence of step-sizes and  $\Pi_{M_{\mu(t)}}$  is an appropriately chosen sequence of projections, to prevent the iterations from growing too fast.

Assuming that  $N_t$  is continuous, we ( $\mathbb{P}$ -a.s.) have that

$$\text{sign}(X_t - \hat{\varphi}_t^T \hat{\theta}_t) = 1 - 2\mathbb{I}(X_t \leq \hat{\varphi}_t^T \hat{\theta}_t), \quad (10)$$

where  $\text{sign}(\cdot)$  is the *signum* function, i.e.,  $\text{sign}(x) = -1$  if  $x < 0$ ,  $\text{sign}(x) = 1$  if  $x > 0$  and  $\text{sign}(x) = 0$  if  $x = 0$ . Therefore, the algorithm (9) will behave almost surely as

$$\hat{\theta}_{t+1} = \Pi_{M_{\mu(t)}} \left[ \hat{\theta}_t + \alpha_t \hat{\varphi}_t \text{sign}(X_t - \hat{\varphi}_t^T \hat{\theta}_t) \right], \quad (11)$$

which is a “sign-error” type algorithm with expanding truncation bounds (Chen and Yin [2003]). This kind of stochastic approximation algorithms typically arise if the  $L_1$  optimization criterion is applied. They can also be seen as variants of the least-mean squares (LMS) algorithm.

### 3.2 Step-sizes

Algorithm (9) is a *stochastic approximation* algorithm (Kushner and Yin [2003]) for which it is typical to assume that the (possibly randomized) step-sizes satisfy ( $\mathbb{P}$ -a.s.)

$$\sum_{t=0}^{\infty} \alpha_t = \infty, \quad (12)$$

$$\sum_{t=0}^{\infty} \alpha_t^2 < \infty, \quad (13)$$

$$\forall t \geq 0 : \alpha_t \geq 0. \quad (14)$$

Condition (13) can often be weakened to  $\lim_{t \rightarrow \infty} \alpha_t = 0$  (Kushner and Yin [2003]). Here, we will simply apply

*Assumption 4.*  $\alpha_0 = 1$  and  $\forall t > 0 : \alpha_t = 1/t$ .

which is the archetypical choice of step-sizes. Note that in practice, a small, constant step-size could also be applied, in order to achieve tracking of time-varying parameters.

### 3.3 Expanding Truncation Bounds

If  $(M_t)_t$  is a sequence of (strictly) monotone increasing positive real numbers with  $M_t \rightarrow \infty$  as  $t \rightarrow \infty$ , then we define a sequence of random variables  $\mu(t)$  as follows

$$\mu(t) \triangleq \sum_{i=1}^{t-1} \mathbb{I}(|\hat{\theta}_i + \Delta \hat{\theta}_i| > M_{\mu(i)}), \quad (15)$$

where  $\mathbb{I}(\cdot)$  is the indicator function, as before, and  $\Delta \hat{\theta}_i$  is

$$\Delta \hat{\theta}_i \triangleq \alpha_i \hat{\varphi}_i (1 - 2\mathbb{I}(X_i \leq \hat{\varphi}_i^T \hat{\theta}_i)), \quad (16)$$

where the notations are the same as previously.

Given a positive real  $M$ , projection  $\Pi_M$  is defined as

$$\Pi_M(x) \triangleq \begin{cases} x & \text{if } \|x\| \leq M, \\ 0 & \text{otherwise,} \end{cases} \quad (17)$$

where  $\|\cdot\|$  is an (arbitrary) norm, chosen by the user.

Therefore, the estimates are projected back to the origin, each time they leave a given closed norm ball; furthermore, the radius of this ball is increased each time the estimates leave it. Since this radius goes to infinity, the estimates will eventually almost surely stay inside the ball.

The use of the origin to which we project the iterations back is only for convenience; any other interior point of the bounding sphere could be applied (Chen and Yin [2003]).

### 3.4 Algorithm I: FIR Approximation

In the first algorithm, we will consider a finite impulse response (FIR) approximation of the ARX system and estimate the parameters of this FIR approximation. As we will see, by assuming the knowledge of some (finitely many) input weights, the “true” system parameters,  $\theta^*$ , can be easily computed by a simple matrix inversion.

First, note that (1) can be written in the form of

$$X_t = \sum_{i=1}^{\infty} c_i^* U_{t-1} + \sum_{i=0}^{\infty} d_i^* N_{t-i}, \quad (18)$$

where  $(c_i^*)_{i=1}^{\infty}$  and  $(d_i^*)_{i=0}^{\infty}$  are the impulse responses. We will approximate this system with an FIR system of order  $p+q$ , more precisely we will consider the representation

$$X_t = \bar{\varphi}_t^T \bar{\theta}^* + W_t, \quad (19)$$

where the vectors  $\bar{\varphi}_t$  and  $\bar{\theta}^*$  are defined as follows

$$\bar{\varphi}_t \triangleq (U_{t-1}, \dots, U_{t-p-q})^T, \quad (20)$$

$$\bar{\theta}^* \triangleq (c_1^*, \dots, c_{p+q}^*)^T, \quad (21)$$

and  $W_t$  is simply the unmodelled part of the system

$$W_t \triangleq \sum_{i=p+q+1}^{\infty} c_i^* U_{t-i} + \sum_{i=0}^{\infty} d_i^* N_{t-i}. \quad (22)$$

An easy computation shows that if we can estimate  $\bar{\theta}^*$ , we can also estimate the true parameter vector  $\theta^*$ .

We can observe that given model (1) parametrized by the vector  $(a_1, \dots, a_p, b_1, \dots, b_q)$ , we can compute all input parameters,  $(c_i^*)_{i=1}^{\infty}$ , of problem (18) by the recursion

$$c_i = b_i + \sum_{j=1}^{i-1} a_j c_{i-j}, \quad (23)$$

with the convention that  $a_j \triangleq 0$  if  $j > p$  and  $b_j \triangleq 0$  if  $j > q$ . We treat the empty sum as zero, hence,  $c_1 = b_1$ .

Then, we obtain parameter  $\theta^*$  from  $\bar{\theta}^*$  by writing the above recursion in a matrix form and inverting it. More precisely, there is a measurable function  $f$ , such that

$$\theta^* = f(\bar{\theta}^*), \quad (24)$$

where  $f : \mathbb{R}^{p+q} \rightarrow \mathbb{R}^{p+q}$  is defined for a vector  $\bar{\theta}$  as

$$f(\bar{\theta}) \triangleq \begin{cases} G(\bar{\theta})^{-1} \bar{\theta} & \text{if } \det(G(\bar{\theta})) \neq 0, \\ 0 & \text{otherwise,} \end{cases} \quad (25)$$

where  $G(\bar{\theta})$  is a  $(p+q) \times (p+q)$  matrix parametrized by vector  $\bar{\theta} = (c_1, c_2, \dots, c_{p+q})$ . Matrix  $G(\bar{\theta})$  is defined as

$$G(\bar{\theta}) \triangleq \begin{bmatrix} G_1(\bar{\theta}) & I \\ G_2(\bar{\theta}) & 0 \end{bmatrix}, \quad (26)$$

where  $I$  is the  $p \times p$  identity matrix and  $G_1(\bar{\theta})$  is

$$G_1(\bar{\theta}) \triangleq \begin{bmatrix} 0 & \cdots & \cdots & \cdots & 0 \\ c_1 & 0 & & & \vdots \\ c_2 & c_1 & \ddots & & \vdots \\ \vdots & \ddots & \ddots & 0 & \vdots \\ c_{p-1} & \cdots & c_2 & c_1 & 0 \end{bmatrix}, \quad (27)$$

a  $p \times p$  matrix, while  $G_2(\bar{\theta})$  is a  $q \times p$  matrix,

$$G_2(\bar{\theta}) \triangleq \begin{bmatrix} c_p & \cdots & c_2 & c_1 \\ c_{p+1} & \cdots & c_3 & c_2 \\ \vdots & & \vdots & \vdots \\ c_{p+q-1} & \cdots & c_{q+1} & c_q \end{bmatrix}. \quad (28)$$

Note that  $G(\bar{\theta}^*)$  is always invertible because of the assumptions on the system. Moreover, if  $\hat{\theta}_t$  is the estimate provided by recursion (9) at time  $t$ , then it is easy to see that random variable  $f(\hat{\theta}_t)$  is  $\mathcal{F}_t$ -measurable.

Now, we specify Algorithm I by defining  $\hat{\varphi}_t \triangleq \bar{\varphi}_t$ , cf. equation (20). The next theorem shows that, after post-processing the iterations with function  $f$ , Algorithm I provides a strongly consistent recursive estimate for  $\theta^*$ .

*Theorem 1.* Let  $(\hat{\theta}_t)_{t=0}^{\infty}$  be the sequence generated by recursion (9) with  $\hat{\varphi}_t = \bar{\varphi}_t$ . Then, under Assumptions 1-4,  $f(\hat{\theta}_t)$  converges ( $\mathbb{P}$ -a.s.) to  $\theta^*$ , as  $t \rightarrow \infty$ , for any  $\hat{\theta}_0 \in \mathbb{R}^{p+q}$ .

*Proof.* First, note that it is enough to prove that  $\hat{\theta}_t \rightarrow \bar{\theta}^*$  almost surely as  $t \rightarrow \infty$ , since function  $f$  is continuous in a (small enough) neighbourhood of  $\theta^*$ . This local continuity of  $f$  simply follows from the continuity of matrix inversion and the fact that  $G(\bar{\theta}^*)$  is invertible (cf. Assumptions 3).

We will prove that  $\hat{\theta}_t \rightarrow \bar{\theta}^*$  as  $t \rightarrow \infty$  ( $\mathbb{P}$ -a.s.) by applying Theorem 3.5 of Chen and Yin [2003], which states that if process  $(X_t, \bar{\varphi}_t)$  is stationary and ergodic as well as the following matrix is (well-defined and) positive definite

$$\Gamma \triangleq \begin{bmatrix} \mathbb{E}[X_0^2] & \mathbb{E}[X_0 \bar{\varphi}_0^T] \\ \mathbb{E}[X_0 \bar{\varphi}_0] & \mathbb{E}[\bar{\varphi}_0 \bar{\varphi}_0^T] \end{bmatrix} > 0, \quad (29)$$

then  $d(\hat{\theta}_t, S) \rightarrow 0$  as  $t \rightarrow \infty$  ( $\mathbb{P}$ -a.s.), where  $(\hat{\theta}_t)_t$  is the sequence generated by (11),  $\alpha_t = 1/t$ , set  $S$  is defined as

$$S \triangleq \{ \bar{\theta}^\circ \in \mathbb{R}^{p+q} : \mathbb{E}[\bar{\varphi}_0 \text{sign}(X_0 - \bar{\varphi}_0^T \bar{\theta}^\circ)] = 0 \}, \quad (30)$$

and  $d(\cdot, \cdot)$  is a metric based on an (arbitrary) norm.

Our theorem is about sequences generated by iteration (9), but as it was noted earlier, if  $N_t$  is continuous (Assumption 1), then the sequences generated by iterations (9) and (11) are ( $\mathbb{P}$ -a.s.) the same (assuming the same initial estimates,  $\hat{\theta}_0 = \hat{\theta}'_0$ ). Thus, it is enough to study iteration (11).

Therefore, in order to prove Theorem 1, we have to prove the following three claims: (i) process  $(X_t, \bar{\varphi}_t)$  is stationary

and ergodic; (ii) matrix  $\Gamma$  is positive definite; and (iii) the vector  $\bar{\theta}^*$  is the only element of the set  $S$ , i.e.,  $S = \{\bar{\theta}^*\}$ .

Part (i): First, we will argue that the process  $(X_t, \bar{\varphi}_t)_t$  is stationary and ergodic. It is easy to see that  $(\bar{\varphi}_t)_t$  is stationary, since the input sequence  $(U_t)_t$  is an i.i.d. process. Its ergodicity follows from its  $m$ -dependence, with  $m = n-1$ , i.e.,  $\bar{\varphi}_i$  and  $\bar{\varphi}_j$  are independent if  $|i-j| > n-1$ .

Now, we have to show that  $(X_t)_t$  also has these properties. If we consider the representation (18), it can be seen that  $(X_t)_t$  is the sum of two processes and both of these processes are filtered i.i.d. sequences, where the filters are linear and stable. Therefore, both of these are stationary and ergodic, moreover, since they are independent, their sum, i.e.,  $(X_t)_t$ , is also (strictly) stationary and ergodic.

Part (ii): Next, we will prove that matrix  $\Gamma$  is positive definite. In our case (Assumptions 1, 2),  $\Gamma$  takes the form

$$\Gamma = \begin{bmatrix} P & Q \\ Q^T & R \end{bmatrix}, \quad (31)$$

where  $P = \mathbb{E}[X_0^2]$  is a scalar defined as

$$P \triangleq \sigma_u^2 \sum_{i=1}^{\infty} c_i^{*2} + \sigma_n^2 \sum_{j=0}^{\infty} d_j^{*2}, \quad (32)$$

where  $\sigma_n^2$  and  $\sigma_u^2$  are the variances of the noise and the input, respectively. Vector  $Q = \mathbb{E}[X_0 \bar{\varphi}_0^T]$  is

$$Q \triangleq (c_1^* \sigma_u^2, c_2^* \sigma_u^2, \dots, c_{p+q}^* \sigma_u^2). \quad (33)$$

Finally, matrix  $R = \mathbb{E}[\bar{\varphi}_0 \bar{\varphi}_0^T]$  is simply  $\sigma_u^2 I$ , where  $I$  is the  $p+q \times p+q$  identity matrix. Since  $\Gamma$  is symmetric and  $R$  is positive definite, the positive definiteness of  $\Gamma$  follows from the positive definiteness of the Schur complement of  $R$  (see, e.g., Corollary 14.8.6 of Harville [1997]). We have

$$P - QR^{-1}Q^T = \sigma_u^2 \sum_{i=p+q+1}^{\infty} c_i^{*2} + \sigma_n^2 \sum_{j=0}^{\infty} d_j^{*2} \geq \sigma_n^2 > 0, \quad (34)$$

since  $d_0^* = 1$  and  $\sigma_n^2 > 0$ . Therefore,  $\Gamma$  is positive definite.

Part (iii): Now, we will show that (a) vector  $\bar{\theta}^* \in S$  and, moreover, (b) this is its only element,  $S = \{\bar{\theta}^*\}$ .

We can observe, that  $\bar{\theta}^\circ \in S$  if and only if  $\bar{\theta}^\circ$  minimizes the  $L_1$  (or least absolute deviations) error criterion,

$$\bar{\theta}^\circ \in \arg \min_{\bar{\theta}} \mathbb{E} |X_0 - \bar{\varphi}_0^T \bar{\theta}|, \quad (35)$$

since function  $g(\bar{\theta}) \triangleq -\mathbb{E}[\bar{\varphi}_0 \text{sign}(X_0 - \bar{\varphi}_0^T \bar{\theta})]$  can be seen as its (sub)gradient, which must be zero to minimize it.

First, we show that  $g(\bar{\theta}^*) = 0$ , which proves that  $\bar{\theta}^* \in S$ . This is indeed the case, since we have

$$\begin{aligned} g(\bar{\theta}^*) &= -\mathbb{E}[\bar{\varphi}_0 \text{sign}(X_0 - \bar{\varphi}_0^T \bar{\theta}^*)] = -\mathbb{E}[\bar{\varphi}_0 \text{sign}(W_0)] = \\ &= -\mathbb{E}[\bar{\varphi}_0] \mathbb{E}[\text{sign}(W_0)] = 0, \end{aligned} \quad (36)$$

where variable  $W_0$  is defined according to equation (22).

The uniqueness of the solution follows from the classical theory of  $L_1$  linear regression. More precisely, assume that there is a vector  $\bar{\theta}^\circ \in S$  such that  $\bar{\theta}^\circ \neq \bar{\theta}^*$ . Then, this solution also minimizes the  $L_1$  loss function, an thus

$$l(\bar{\theta}^\circ) = \mathbb{E} |X_0 - \bar{\varphi}_0^T \bar{\theta}^\circ| = \mathbb{E} |N_0 + \bar{\varphi}_0^T (\bar{\theta}^* - \bar{\theta}^\circ)| =$$

$$= \mathbb{E} \left[ \mathbb{E} \left[ N_0 + \bar{\varphi}_0^T (\bar{\theta}^* - \bar{\theta}^\circ) \mid \mathcal{G} \right], \quad (37)$$

where  $\mathcal{G}$  is the  $\sigma$ -algebra generated by random variables  $U_{-1}, \dots, U_{-p-q}$ . Thus, the loss function  $l$  is minimized if

$$\mathbb{E} \left[ \left| N_0 + \bar{\varphi}_0^T (\bar{\theta}^* - \bar{\theta}^\circ) \right| \mid \mathcal{G} \right] \quad (38)$$

is minimized. Since  $N_0$  is independent of  $\mathcal{G}$  and it has a unique median zero (Assumption 1),  $\bar{\varphi}_0^T (\bar{\theta}^* - \bar{\theta}^\circ) = 0$  on  $\mathcal{G}$  almost surely. But (using Assumption 2) this can only be the case if  $\bar{\theta}^\circ = \bar{\theta}^*$  (see, e.g., Lemma 3 of Bloomfield and Steiger [1983]). Hence, we have showed that  $S = \{\theta^*\}$ .

Therefore, by Theorem 3.5 of Chen and Yin [2003], ( $\mathbb{P}$ -a.s.)

$$\lim_{t \rightarrow \infty} \|\hat{\theta}_t - \bar{\theta}^*\| = 0. \quad (39)$$

Moreover, as it was highlighted,  $f(\cdot)$  is continuous around  $\bar{\theta}^*$  and thus there is an index  $t_0$ , such that for all  $t > t_0$ :  $G(\hat{\theta}_t)$  is invertible and as  $\hat{\theta}_t \rightarrow \bar{\theta}^*$  we also have that  $f(\hat{\theta}_t) \rightarrow f(\bar{\theta}^*) = \theta^*$ , which completes the proof.  $\square$

An interesting question is the rate of convergence of such algorithms. By the ‘‘rate of convergence’’ we mean the asymptotic properties of the normalized estimation errors around point  $\bar{\theta}^*$ . We should find a scaling factor,  $\alpha \in (0, 1)$ , such that  $t^\alpha (\hat{\theta}_t - \bar{\theta}^*)$  has a non-trivial limit distribution. As it was shown by Chen and Yin [2003] for sign-error algorithms, the scaling factor is  $\alpha = 1/2$  and, furthermore, for large enough  $t$ ,  $\sqrt{t}(\hat{\theta}_t - \bar{\theta}^*)$  is approximately normal.

### 3.5 Algorithm II: Simultaneous State Estimation

In the previous section we saw that Algorithm I provided a strongly consistent estimator, moreover, the estimation error was approximately normal. However, Algorithm I requires post-processing in the form of a matrix inversion, which could become numerically sensitive if matrix  $G(\hat{\theta})$  is ill-conditioned. Additionally, the noise that Algorithm I needs to handle is ‘‘larger’’ than  $N_t$ , more precisely, it is

$$W_t = \sum_{i=p+q+1}^{\infty} c_i^* U_{t-i} + \sum_{i=0}^{\infty} d_i^* N_{t-i}, \quad (40)$$

which can also slow down the convergence of the algorithm.

In this section we will present another recursive estimator that does not need any post-processing and does not require to treat the unmodelled part of the system as part of the noise. Hence, it has numerical advantages.

The main idea of our Algorithm II is to achieve a direct estimate of  $\theta^*$  by simultaneously maintaining an estimate for the output,  $\hat{X}_t$  and for the parameter,  $\hat{\theta}_t$ , at time  $t$ . The algorithm starts with setting the estimate of the past outputs of the hidden process to zero, and it constructs a sequence of output estimates recursively as follows

$$\hat{X}_t \triangleq \begin{cases} \sum_{i=1}^p \hat{a}_{t,i} \hat{X}_{t-1} + \sum_{i=1}^q \hat{b}_{t,i} U_{t-i} & \text{if } t \geq 0 \\ 0 & \text{otherwise,} \end{cases}$$

where  $(\hat{a}_{t,i})_{i=1}^p$  and  $(\hat{b}_{t,i})_{i=1}^q$  are the estimates of the true parameters at time  $t$ . Algorithm II is also defined as (9),

$$\hat{\theta}_{t+1} = \Pi_{M_{\mu(t)}} \left[ \hat{\theta}_t + \alpha_t \hat{\varphi}_t \left( 1 - 2 \mathbb{I}(X_t \leq \hat{\varphi}_t^T \hat{\theta}_t) \right) \right], \quad (41)$$

by setting the components of this general recursion to

$$\hat{\varphi}_t \triangleq (\hat{X}_{t-1}, \dots, \hat{X}_{t-p}, U_{t-1}, \dots, U_{t-q})^T, \quad (42)$$

$$\hat{\theta}_t \triangleq (\hat{a}_{t,1}, \dots, \hat{a}_{t,p}, \hat{b}_{t,1}, \dots, \hat{b}_{t,q})^T. \quad (43)$$

Though, we cannot expect our estimates  $(\hat{X}_{t-1}, \dots, \hat{X}_{t-p})$  to eventually reach the true outputs (states) of the system, since the process progresses to a new state in each iteration and we only have binary measurements about its past, our numerical experiments support that this algorithm is also an efficient way of estimating parameter  $\theta^*$ .

## 4. EXPERIMENTAL RESULTS

The two recursive estimation methods were tested on a second order ARX system. More precisely, we investigated the identification of the following ARX process

$$X_t \triangleq a_1 X_{t-1} + a_2 X_{t-2} + b_1 U_{t-1} + b_2 U_{t-2} + N_t, \quad (44)$$

where the inputs,  $(U_t)_t$ , were i.i.d., standard normal, and the noises,  $(N_t)_t$ , were i.i.d., zero mean Gaussian with variance  $1/4$ . The true system parameters were set to

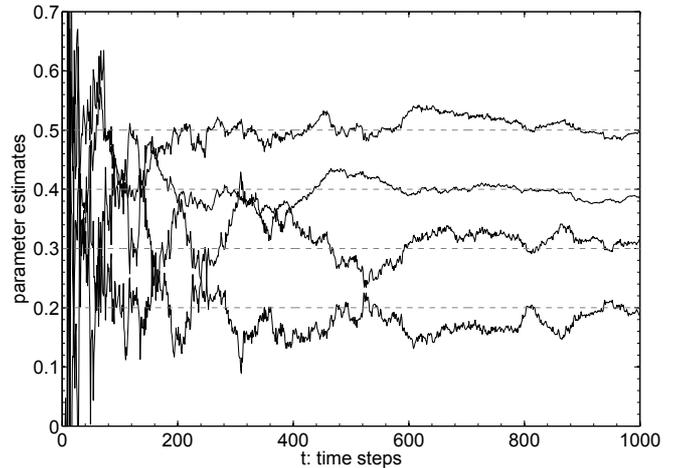


Figure 2. Recursive estimation with Algorithm I (based on FIR approximation); dashed lines: location of the true parameters, solid lines: evolution of their estimates.

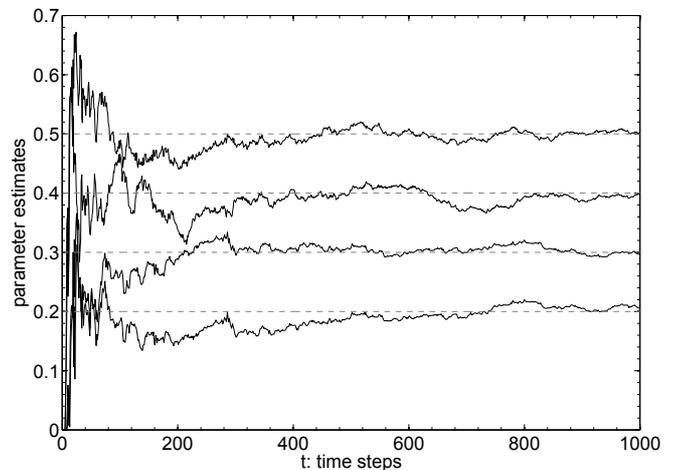


Figure 3. Identification with Algorithm II (based on simultaneous estimation of the parameters and the states); dashed lines: true parameters, solid lines: estimates.

$$a_1 = 0.2, \quad a_2 = 0.3, \quad b_1 = 0.4, \quad b_2 = 0.5, \quad (45)$$

and the initial conditions were  $X_{-1} = 0$  and  $X_{-2} = 0$ .

The initial parameter estimates were all set to zero, as well,  $\hat{\theta}_0 = (0, 0, 0, 0)$ , and the step-sizes were  $\alpha_t = 1/t$ . We applied  $M_0 = 10^3$  as the (first) truncation bound, which was never reached, hence it did not affect the iterations.

Figures 2 and 3 demonstrate the progress of Algorithms I and II, respectively. They show that after 300 iterations, the estimated parameters became “close” to the real ones. The convergence was quick for both algorithms, however, it can also be seen from the figures that, in these experiments, Algorithm II exhibited smaller fluctuations than Algorithm I did. An explanation of this could be the larger noise terms that Algorithm I needed to handle.

The estimates approached the true system parameters for both algorithms. These experiments agree with the theoretical results regarding Algorithm I and support our expectation that Algorithm II is also (strongly) consistent.

## 5. CONCLUDING REMARKS

In this paper the identification of ARX systems via binary sensors has been analysed and two recursive estimation methods have been proposed. In contrast to the previous approaches to this problem, which typically presupposed the complete knowledge of the noise distribution and that the inputs can be chosen by the user, we only made mild assumptions on the noise and assumed that the inputs can only be observed, but they cannot be changed. However, we made the assumption that the threshold of the binary sensor is adjustable, it can be set in each step.

Note that our problem setting can be reformulated in the framework of binary identification with a non-adjustable, fixed threshold, in case a dither signal is allowed.

This paper can be seen as a generalization of our earlier work (Csáji and Weyer [2011]) in which we investigated the identification of constants via binary sensors with adjustable thresholds. In that paper we showed that under minimal assumptions (i.e., the noise process is a mediangale and its conditional CDFs almost surely do not converge to  $1/2$  at any nonzero point), a strongly consistent estimator can be given. The proposed recursive estimator was based on stochastic approximation. It also had an easy extension to FIR systems, in case the inputs can be chosen by the user. In this paper we generalized these ideas to the binary identification of ARX systems without assuming the controllability of the input.

We have suggested two recursive identification algorithms to estimate the unknown true parameters of the process. Algorithm I applied an FIR approximation and required a matrix inversion to get an estimate of the original system parameters. We argued that it was a strongly consistent estimator and its error was approximately normal.

Then, we also introduced a second recursive estimator, Algorithm II. This method simultaneously estimates the outputs of the system together with the unknown parameters. It is an open question whether Algorithm II can also almost surely find the true parameters, asymptotically.

Finally, we presented results of numerical experiments about the suggested recursive estimators. These experiments demonstrated that the parameter estimates produced by both Algorithms I and II efficiently approached the unknown system parameters of a 2nd order ARX system after “small” number data points.

Possible future research directions are (i) to prove (or refute) the (strong) consistency of Algorithm II; (ii) to analyze its rate of convergence; and (iii) to extend both algorithms to more general systems, for example, to generalize the methods to Box-Jenkins models.

## REFERENCES

- P. Bloomfield and W. L. Steiger. *Least Absolute Deviations*. Birkhäuser, 1983.
- M. Casini, A. Garulli, and A. Vicino. Input design in worst-case system identification using binary sensors. *IEEE Trans. on Automatic Control*, 56:1186–1191, 2011.
- H-F. Chen. Recursive system identification. *Acta Mathematica Scientia*, 29:650–672, 2009.
- H-F. Chen and G. Yin. Asymptotic properties of sign algorithms for adaptive filtering. *IEEE Transactions on Automatic Control*, 48(9):1545–1556, 2003.
- E. Colinet and J. Juillard. A weighted least-squares approach to parameter estimation problems based on binary measurements. *IEEE Transactions on Automatic Control*, 55:148–152, 2010.
- B. Cs. Csáji and E. Weyer. System identification with binary observations by stochastic approximation and active learning. In *Proceedings of the 50th IEEE CDC and ECC, Orlando, Florida, December 12-15, 2011*.
- L. Gerencsér and Zs. Vágó. Stochastic approximation for function minimization under quantization error. In *Proceedings of the 38th IEEE CDC, Phoenix, Arizona, USA, pages 2373–2377, 1999*.
- B. I. Godoy, G. C. Goodwin, J. C. Agüero, D. Marelli, and T. Wigren. On identification of FIR systems having quantized output data. *Automatica*, 47:1905–1915, 2011.
- G. C. Goodwin and K. S. Sin. *Adaptive Filtering: Prediction and Control*. Prentice-Hall, 1984.
- D. A. Harville. *Matrix Algebra From a Statistician’s Perspective*. Springer, 1997.
- H. J. Kushner and G. Yin. *Stochastic Approximation and Recursive Algorithms and Applications*. Springer, 2003.
- L. Ljung. *System Identification: Theory for the User*. Prentice-Hall, Upper Saddle River, 2nd edition, 1999.
- L. Ljung and T. Söderström. *Theory and Practice of Recursive Identification*. The MIT Press, 1987.
- Le Yi Wang, G. George Yin, Ji-Feng Zhang, and Yanlong Zhao. *System Identification with Quantized Observations*. Birkhäuser, 2010.
- E. Weyer, S. Ko, and M. C. Campi. Finite sample properties of system identification with quantized output data. In *Procs. of the 48th IEEE CDC, Shanghai, China, December 16-18, pages 1532–1537, 2009*.