



HAL
open science

An Identification Algorithm for FIR Systems from Binary Output Measurements

Ali Mestrah, Hicham Oualla, Mathieu Pouliquen, Eric Pigeon

► **To cite this version:**

Ali Mestrah, Hicham Oualla, Mathieu Pouliquen, Eric Pigeon. An Identification Algorithm for FIR Systems from Binary Output Measurements. International Conference on Control, Decision and Information Technologies, Jul 2023, Rome, Italy. hal-04131670

HAL Id: hal-04131670

<https://normandie-univ.hal.science/hal-04131670v1>

Submitted on 27 Jun 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

An Identification Algorithm for FIR Systems from Binary Output Measurements

Ali Mestrah, Hicham Oualla, Mathieu Pouliquen, Eric Pigeon
Laboratoire d'Ingénierie des Systèmes - UR 7478
Normandie Univ, UNICAEN, ENSICAEN, LIS
14050 Caen Cedex, France
ali.mestrah@unicaen.fr

Abstract—The present study deals with a new identification algorithm from binary output measurements. The study focuses on the class of Finite Impulse Response (FIR) systems. The proposed algorithm is based on the estimation of correlation functions. A geometric interpretation is proposed and leads to a formulation of the algorithm using a Singular Value Decomposition (SVD). A convergence analysis is proposed showing the mean-square convergence with a rate of $\mathcal{O}(N^{-1})$, Monte Carlo simulations are proposed to confirm performance.

Index Terms—System Identification, Binary Measurements, FIR system.

I. INTRODUCTION

The system parameters identification is a process that requires usually some information on input, output and noise. The input is often known by the user, but the output not always as one might want. In this study, our interest is the identification with binary output, in which the real output of the system is unknown, we only have access to binary measurements. This type of identification is widely used in the domain of micro-electronics and telecommunications for instance. The choice of binary measurements is due to the facility of measuring, e.g knowing the sign of the output in binary measurements is much easier than knowing the real output, binary measurements are also easy to store and occupy less memory. An other important aspect is the fact that binary sensors have lower cost and are more robust against noise than ordinary sensors.

Several researches focused on the identification of systems from binary measurements, we cite for instance [1], [2], [3], [4], [5], [6], [7], [8], [9], [10], [11], [12], [13], [14] and [15]. What differ these studies are the process of implementation (online or offline), the impulse response of the system (finite or infinite), the representation of the system (as a state representation or a transfer function representation), the condition on the input and noise (statistical characteristics, periodicity, distribution function, etc..). In this study, our contribution focuses on Finite Impulse Response (FIR) systems. In most of studies developed for this type of systems, the input and the noise are subject to strict conditions: in [1], [16] and [17] the input must be periodic, in [18] the input must be binary, in [3] a known dithering signal is added on the input of the binary sensor, in [6] the algorithm is developed in a noise free framework, in [7] the noise distribution function must

be known, in [12] the input must be an Independent and Identically Distributed (IID) random variable sequence, in [17] the noise distribution function must be estimated before the identification process or it must be known. Sometimes some user defined parameters are difficult to adjust as in [10] where there are two parameters that must be well adjusted to have consistent estimates, sometimes the threshold must be variable as in [19] and sometimes there is no theoretically convergence proof as in [2] and [14].

Our first contribution in this paper is to propose a simplest identification algorithm with respect to other algorithms, simplest in the sense that the practical framework is less restrictive: simplest in terms of condition on the input signal, simplest in terms of required knowledge on the noise, simplest in terms of user-defined parameter. Roughly speaking, our algorithm is based on the estimation of correlation functions. These correlation functions can be estimated from knowledge of the input and the binary output data, and this leads to the estimation of the system parameters. Our second contribution is to propose an algebraic formulation, allowing an implementation more computationally robust.

The article is organized as follows: the identification problem is described and some notations are introduced in section II. The proposed identification algorithm is introduced, interpreted and analyzed in section III. Numerical simulations are proposed in section IV. Section V concludes the paper.

II. PROBLEM STATEMENT

The discrete time system considered in this paper is depicted in Fig.1. The system is described by $H(z^{-1})$ defined by:

$$H(z^{-1}) = \sum_{i=0}^n h_i z^{-i} \quad (1)$$

where n is the order and $\{h_i\}_{i \in [0;n]}$ is the impulse response. y_t and u_t are respectively the system output and input at the time instant tT , T is the sampling period.

From (1) and Fig.1 the output is expressed as follows:

$$y_t = \sum_{i=0}^n h_i u_{t-i} + v_t \quad (2)$$

where v_t is the noise on the output.

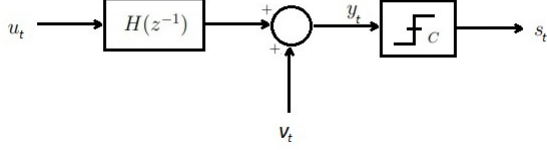


Fig. 1. Block diagram of the system model.

As shown in Fig.1 the output of the system is unknown but is measured by a 1-bit ADC (Analog to Digital Converter). We thus define the known binary output s_t as follows:

$$\begin{cases} s_t = 1, & \text{if } y_t \geq C \\ s_t = -1, & \text{if } y_t < C \end{cases} \quad (3)$$

Remark 1: It was possible to choose the values 1 and 0 for the binary output, but we choose the values 1 and -1 , in order to simplify the algebraic interpretation in section III-B. In the following, let us denote the parameter vector θ as

$$\theta^T = (h_0 \quad h_1 \quad \dots \quad h_n) \quad (4)$$

The **objective** of the paper is to estimate the parameter vector θ from N available data $\{u_t, s_t\}_{t=1}^N$.

Some assumptions are taken throughout the document in order to solve the identification problem:

- **A.1** The input is a stationary zero mean sequence with gaussian distribution.
- **A.2** The noise is zero mean sequence of stationary random variables uncorrelated with the input.
- **A.3** The order n of the system is known.
- **A.4** The L_2 norm of the parameter vector θ is known and different from 0.

Note that assumption **A.4** is a normalization assumption. Such an assumption is commonly used in system identification from binary measurements ([6] and [10] for instance). An other similar assumption can be an assumption on the static gain of the process which is the sum of the values of θ .

III. IDENTIFICATION ALGORITHM

The identification algorithm is presented in this section. First, the principle is described in subsection III-A, then an algebraic formulation is proposed in subsection III-B, finally a convergence analysis is provided in subsection III-C.

A. Principle

Let us define $\mathbb{E}[\cdot]$ as the expected value. From (2) it can easily be shown that

$$R_{uy}(i) = \sum_{j=0}^n h_j R_{uu}(i-j) + R_{uv}(i) \quad (5)$$

where $R_{uy}(i) = \mathbb{E}[u_t y_{t+i}]$ and $R_{uv}(i) = \mathbb{E}[u_t v_{t+i}]$ are respectively the cross-correlations between the variables u_t , y_t and u_t , v_t with lag equal to i and $R_{uu}(i) = \mathbb{E}[u_t u_{t+i}]$ is the

auto-correlation of the variable u_t with lag equal to i . From Assumptions **A.1** and **A.2** the noise is uncorrelated with the input and zero mean, it follows that $R_{uv}(i) = 0$. Using (5) for i from 0 to n we get

$$\begin{pmatrix} R_{uy}(0) \\ R_{uy}(1) \\ \vdots \\ R_{uy}(n) \end{pmatrix} = \begin{pmatrix} R_{uu}(0) & R_{uu}(1) & \dots & R_{uu}(n) \\ R_{uu}(1) & R_{uu}(0) & \dots & R_{uu}(n-1) \\ \vdots & \vdots & \ddots & \vdots \\ R_{uu}(n) & R_{uu}(n-1) & \dots & R_{uu}(0) \end{pmatrix} \theta. \quad (6)$$

In order to use (6) to estimate θ note that $\forall i \in [1; n]$

$$\mathbb{E} \left[\left(u_t - \frac{R_{uy}(i)}{R_{yy}(0)} y_{t+i} \right) y_{t+i} \right] = 0 \quad (7)$$

for each t . This implies that $u_t - \frac{R_{uy}(i)}{R_{yy}(0)} y_{t+i}$ and y_{t+i} are uncorrelated and it follows that

$$\mathbb{E} \left[\left(u_t - \frac{R_{uy}(i)}{R_{yy}(0)} y_{t+i} \right) \middle| y_{t+i} \right] = \mathbb{E} \left[\left(u_t - \frac{R_{uy}(i)}{R_{yy}(0)} y_{t+i} \right) \right]. \quad (8)$$

From Assumption **A.1** the input is a stationary zero mean sequence with gaussian distribution, it follows that the output too and consequently $\mathbb{E} \left[\left(u_t - \frac{R_{uy}(i)}{R_{yy}(0)} y_{t+i} \right) \right] = 0$. Then from (8) we have

$$\begin{aligned} \mathbb{E} [u_t | y_{t+i}] &= \frac{R_{uy}(i)}{R_{yy}(0)} \mathbb{E} [y_{t+i} | y_{t+i}] \\ &= \frac{R_{uy}(i)}{R_{yy}(0)} y_{t+i} \end{aligned} \quad (9)$$

The cross-correlation between u_t and s_t is defined by $R_{us}(i) = \mathbb{E}[u_t s_{t+i}]$, it can then be expressed by

$$\begin{aligned} R_{us}(i) &= \mathbb{E}[u_t s_{t+i}] \\ &= \mathbb{E} \left[\mathbb{E} [u_t s_{t+i} | y_{t+i}] \right] \\ &= \mathbb{E} \left[s_{t+i} \mathbb{E} [u_t | y_{t+i}] \right] \\ &= \frac{R_{uy}(i)}{R_{yy}(0)} \mathbb{E} [s_{t+i} y_{t+i}] \\ &= \alpha R_{uy}(i) \end{aligned} \quad (10)$$

where $\alpha = \frac{R_{yy}(0)}{R_{yy}(0)}$ is a constant. It follows from (5) and (10) that θ satisfies

$$\begin{pmatrix} R_{us}(0) \\ R_{us}(1) \\ \vdots \\ R_{us}(n) \end{pmatrix} = \alpha \begin{pmatrix} R_{uu}(0) & R_{uu}(1) & \dots & R_{uu}(n) \\ R_{uu}(1) & R_{uu}(0) & \dots & R_{uu}(n-1) \\ \vdots & \vdots & \ddots & \vdots \\ R_{uu}(n) & R_{uu}(n-1) & \dots & R_{uu}(0) \end{pmatrix} \theta. \quad (11)$$

Define $\bar{\theta}$ by

$$\bar{\theta} = \begin{pmatrix} R_{uu}(0) & R_{uu}(1) & \dots & R_{uu}(n) \\ R_{uu}(1) & R_{uu}(0) & \dots & R_{uu}(n-1) \\ \vdots & \vdots & \ddots & \vdots \\ R_{uu}(n) & R_{uu}(n-1) & \dots & R_{uu}(0) \end{pmatrix}^{-1} \begin{pmatrix} R_{us}(0) \\ R_{us}(1) \\ \vdots \\ R_{us}(n) \end{pmatrix}, \quad (12)$$

then θ is such that

$$\theta = \frac{1}{\alpha} \bar{\theta} \quad (13)$$

where α is given above. $\frac{1}{\alpha}$ corresponds to a normalization term, if $R_{sy}(0)$ and $R_{yy}(0)$ are not known (which is probably the case in practice), then it is possible to express α using Assumption **A.4** as follows $\alpha = \frac{\|\bar{\theta}\|_2}{\|\theta\|_2}$.

The algorithm proposed in this paper is then based on (13): we propose the estimation of θ as follows

$$\hat{\theta} = \frac{1}{\hat{\alpha}} \hat{\bar{\theta}} \quad (14)$$

with

$$\hat{\theta} = \begin{pmatrix} \hat{R}_{uu}(0) & \hat{R}_{uu}(1) & \dots & \hat{R}_{uu}(n) \\ \hat{R}_{uu}(1) & \hat{R}_{uu}(0) & \dots & \hat{R}_{uu}(n-1) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{R}_{uu}(n) & \hat{R}_{uu}(n-1) & \dots & \hat{R}_{uu}(0) \end{pmatrix}^{-1} \begin{pmatrix} \hat{R}_{us}(0) \\ \hat{R}_{us}(1) \\ \vdots \\ \hat{R}_{us}(n) \end{pmatrix} \quad (15)$$

where $\hat{R}_{us}(i)$ and $\hat{R}_{uu}(i)$ are computed as follows

$$\hat{R}_{us}(i) = \frac{1}{N} \sum_{t=1}^N u_t s_{t+i}, \quad \hat{R}_{uu}(i) = \frac{1}{N} \sum_{t=1}^N u_t u_{t+i} \quad (16)$$

and

$$\hat{\alpha} = \frac{\|\hat{\bar{\theta}}\|_2}{\|\theta\|_2}. \quad (17)$$

Note that the implementation of (14) and (15) requires two conditions. First the Toeplitz matrix of the estimated auto-correlation $\hat{R}_{uu}(i)$ must be full rank in order to compute $\hat{\theta}$. From Assumption **A.1** this should be the case. However, for a finite number of available data this condition can be not satisfied. In the following we consider the implementation of the above algorithm under the following complementary assumption:

- **A.5** The smallest eigen-value of the Toeplitz matrix of the estimated auto-correlation $\hat{R}_{uu}(i)$ is greater than or equal to a strictly positive user-defined value (10^{-9} for instance).

Second, $\|\hat{\bar{\theta}}\|_2$ must be different from 0 in order to realize the normalization. This condition leads to the question of the value of the threshold C . Indeed, for a finite number of available data, if C is too low (respectively too high) with respect to y_t , then $s_t = 1$ (respectively $s_t = -1$) for all t and then $\hat{\bar{\theta}} = 0$ from the fact that the input is zero mean. In the following, in order to avoid such a pathological case, we consider the implementation of the above algorithm only if $\|\hat{\bar{\theta}}\|_2$ is greater than or equal to an user-defined value (10^{-9} for instance). In practice, it is sometimes suggested (as in [1] and [3]) to add a dithering signal on the input of the binary sensor. Such a "complementary" noise allows to avoid the pathological case described above.

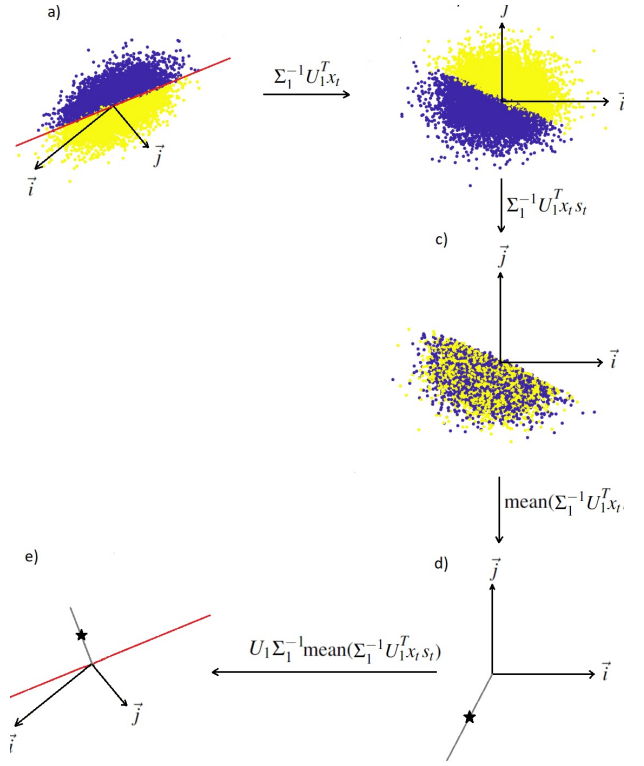


Fig. 2. 2D Geometric interpretation: scatter plot of $\begin{pmatrix} u_t \\ u_{t-1} \end{pmatrix}$ in different basis and estimation of a vector orthogonal to $h_0 u_t + h_1 u_{t-1} = 0$.

B. Geometric interpretation and algebraic formulation

In this subsection we exhibit a geometric interpretation of the proposed method which leads to an algebraic formulation of (15). This geometric interpretation is divided into several steps as detailed below. Some illustrations are depicted in Fig. 2 where, for simplicity, we assume that the number of parameters to be estimated is $n + 1 = 2$ and $C = 0$.

- Fig. 2. a) shows the scatter plot of the dataset $x_t = \begin{pmatrix} u_t \\ u_{t-1} \end{pmatrix}$ for t from 1 to N , the yellow and blue points represent respectively the dataset for s_t equal 1 and -1 . The red line in Fig. 2. a) corresponds to the equation $h_0 u_t + h_1 u_{t-1} = 0$. Visual inspection shows that the scatter plot has two main directions: \vec{i} is the primary vector and \vec{j} is the secondary vector.
- Let X be the regression matrix defined by

$$X = (x_1 \quad x_2 \quad \dots \quad x_N). \quad (18)$$

Applying the Singular Values Decomposition (SVD) on X we get

$$X = U_1 (\Sigma_1 \quad 0) \begin{pmatrix} V_1^T \\ V_2^T \end{pmatrix} = U_1 \Sigma_1 V_1^T \quad (19)$$

where $\Sigma_1 \in \mathbb{R}^{2 \times 2}$ is a diagonal square matrix, $U_1 \in \mathbb{R}^{2 \times 2}$ and $V_1^T \in \mathbb{R}^{2 \times N}$ such that $U_1^T U_1 = I_2$ and $V_1^T V_1 = I_2$.

Multiplying the data by $\Sigma_1^{-1}U_1^T$ we get the scatter plot of Fig. 2. b) (U_1 provides the rotation and Σ_1^{-1} realizes the decompression of data).

- Multiplying each vector in Fig. 2. b) by its corresponding binary output, we get Fig.2.c).
- Computing the mean of the scatter plot of Fig.2.c), we get the black star in Fig.2.d). The vector corresponding to this black star is given by $\frac{1}{N}\sum_{t=1}^N(\Sigma_1^{-1}U_1^T x_t) s_t = \Sigma_1^{-1}U_1^T \frac{1}{N}\sum_{t=1}^N x_t s_t$. Note that this vector is orthogonal to the separating line in Fig.2.b).
- Multiplying the previous vector $\Sigma_1^{-1}U_1^T \frac{1}{N}\sum_{t=1}^N x_t s_t$ by $U_1 \Sigma_1^{-1}$, we obtain the vector $U_1 \Sigma_1^{-2} U_1^T \frac{1}{N}\sum_{t=1}^N x_t s_t$. This vector is the vector with the black star depicted on Fig.2.e) and it is orthogonal to the red line $h_0 u_t + h_1 u_{t-1} = 0$. This vector is consequently proportional to $\begin{pmatrix} h_0 \\ h_1 \end{pmatrix}$.

Note that $\frac{1}{N}U_1 \Sigma_1^2 U_1^T = \frac{1}{N}X X^T$ corresponds to the matrix of the estimated auto-correlation $\hat{R}_{uu}(i)$ used in (15). It follows that $\hat{\theta}$ in (15) corresponds to

$$\begin{aligned} \hat{\theta} &= \left(\frac{1}{N}U_1 \Sigma_1^2 U_1^T \right)^{-1} \frac{1}{N}\sum_{t=1}^N x_t s_t \\ &= N U_1 \Sigma_1^{-2} U_1^T \frac{1}{N}\sum_{t=1}^N x_t s_t \end{aligned} \quad (20)$$

This shows that $\hat{\theta}$ is proportional to the vector with the black star depicted on Fig.2.e).

From the above steps we obtain the following formulation on $\hat{\theta}$:

$$\hat{\theta} = N U_1 \Sigma_1^{-2} U_1^T \text{mean}(U_1 \Sigma_1 V_1^T \odot S) \quad (21)$$

where $\text{mean}(U_1 \Sigma_1 V_1^T \odot S)$ denotes the vector built with the mean of each line of $U_1 \Sigma_1 V_1^T \odot S$, \odot denotes the Hadamard product and S the matrix defined by

$$S = \begin{pmatrix} s_1 & s_2 & \dots & s_N \\ s_1 & s_2 & \dots & s_N \end{pmatrix}. \quad (22)$$

Finally (21) provides the following formulation for the computation of $\hat{\theta}$:

$$\hat{\theta} = N U_1 \Sigma_1^{-1} \text{mean}(V_1^T \odot S). \quad (23)$$

This algebraic formulation is equivalent to (15), its interest is the use of a robust linear algebra tool (i.e. the SVD) for the estimation of $\hat{\theta}$ ([20]).

C. Analysis

An analysis of the algorithm is provided in this subsection. It requires a complementary assumption. Let us first define F_t^k the σ -algebra of events generated by the random variables $\{u_t, t \leq l \leq k\}$ and let $L_2(F_a^b)$ denotes the collection of all second-order random variables which are F_a^b -measurable. Define $\rho(k)$ the maximal correlation coefficient as

$$\rho(k) = \sup_{A \in L_2(F_{-\infty}^0), B \in L_2(F_k^\infty)} \frac{\text{cov}\{A, B\}}{\sigma_A \sigma_B} \quad (24)$$

where σ_A denotes the standard deviation and $\text{cov}\{A, B\}$ the covariance. The stationary process $\{u_t\}$ is called ρ -mixing ([21]) if it satisfies

$$\lim_{k \rightarrow \infty} \rho(k) = 0 \quad (25)$$

Our complementary assumption is thus the following:

- **A.6** $\{u_t\}$ is a ρ -mixing process.

Roughly speaking, Assumption **A.6** means that samples of $\{u_t\}$ sufficiently well separated in time are approximately independent. By now we can establish the following convergence theorem.

Theorem 1: Assume **A.1**, **A.2**, **A.3**, **A.4**, **A.5** and **A.6** hold, then $\hat{\theta}$ defined by (14) is asymptotically unbiased, mean-square convergent and there exists Σ such that

$$\lim_{N \rightarrow \infty} N \mathbb{E} \left[\|\hat{\theta} - \theta\|_2^2 \right] \leq \Sigma. \quad (26)$$

■

Proof 1: The proof is divided in two parts: first we focus on $\hat{\theta} - \bar{\theta}$, then we focus on $\hat{\theta} - \theta$.

• To begin let us consider $\hat{\theta} - \bar{\theta}$. $\hat{\theta}$ is defined by (15). Denote

$$\mathbf{M} = \begin{pmatrix} R_{uu}(0) & R_{uu}(1) & \dots & R_{uu}(n) \\ R_{uu}(1) & R_{uu}(0) & \dots & R_{uu}(n-1) \\ \vdots & \vdots & \ddots & \vdots \\ R_{uu}(n) & R_{uu}(n-1) & \dots & R_{uu}(0) \end{pmatrix} \quad (27)$$

and

$$\mathbf{N} = \begin{pmatrix} R_{us}(0) \\ R_{us}(1) \\ \vdots \\ R_{us}(n) \end{pmatrix}, \quad (28)$$

we have $\hat{\theta} = \hat{\mathbf{M}}^{-1} \hat{\mathbf{N}}$ and then

$$\hat{\theta} - \bar{\theta} = \hat{\mathbf{M}}^{-1} \hat{\mathbf{N}} - \mathbf{M}^{-1} \mathbf{N} \quad (29)$$

This gives

$$\begin{aligned} \hat{\theta} - \bar{\theta} &= \hat{\mathbf{M}}^{-1} (\hat{\mathbf{N}} - \mathbf{N}) + (\hat{\mathbf{M}}^{-1} - \mathbf{M}^{-1}) \mathbf{N} \\ &= \hat{\mathbf{M}}^{-1} (\hat{\mathbf{N}} - \mathbf{N}) + \hat{\mathbf{M}}^{-1} (\mathbf{M} - \hat{\mathbf{M}}) \mathbf{M}^{-1} \mathbf{N} \\ &= \hat{\mathbf{M}}^{-1} (\hat{\mathbf{N}} - \mathbf{N}) + \hat{\mathbf{M}}^{-1} (\mathbf{M} - \hat{\mathbf{M}}) \bar{\theta} \end{aligned} \quad (30)$$

Using the triangular inequality and the Frobenius norm this gives

$$\|\hat{\theta} - \bar{\theta}\|_2 \leq \|\hat{\mathbf{M}}^{-1}\|_F \left(\|\hat{\mathbf{N}} - \mathbf{N}\|_2 + \|\hat{\mathbf{M}} - \mathbf{M}\|_F \|\bar{\theta}\|_2 \right). \quad (31)$$

It is considered in Assumption **A.5** that the smallest eigenvalue of $\hat{\mathbf{M}}$ is greater than or equal to a strictly positive user defined value. Denote a this user defined value, this gives

$$\|\hat{\mathbf{M}}^{-1}\|_F \leq \frac{\sqrt{n+1}}{a} \quad (32)$$

and then

$$\|\hat{\theta} - \bar{\theta}\|_2 \leq \frac{\sqrt{n+1}}{a} \left(\|\hat{\mathbf{N}} - \mathbf{N}\|_2 + \|\hat{\mathbf{M}} - \mathbf{M}\|_F \|\bar{\theta}\|_2 \right) \quad (33)$$

• By now let us consider $\widehat{\theta} - \theta$. $\widehat{\theta}$ is defined by (14), then we have

$$\begin{aligned}\widehat{\theta} - \theta &= \frac{1}{\widehat{\alpha}} \widehat{\theta} - \frac{1}{\alpha} \bar{\theta} \\ &= \frac{1}{\widehat{\alpha}} (\widehat{\theta} - \bar{\theta}) + \left(\frac{1}{\widehat{\alpha}} - \frac{1}{\alpha} \right) \bar{\theta} \\ &= \frac{\|\theta\|_2}{\|\widehat{\theta}\|_2} (\widehat{\theta} - \bar{\theta}) + \frac{\|\theta\|_2}{\|\widehat{\theta}\|_2 \|\bar{\theta}\|_2} (\|\bar{\theta}\|_2 - \|\widehat{\theta}\|_2) \bar{\theta} \quad (34) \\ &= \frac{\|\theta\|_2}{\|\widehat{\theta}\|_2} (\widehat{\theta} - \bar{\theta}) + \frac{1}{\|\widehat{\theta}\|_2} (\|\bar{\theta}\|_2 - \|\widehat{\theta}\|_2) \theta\end{aligned}$$

Using the triangular inequality and the fact that $\|\bar{\theta}\|_2 - \|\widehat{\theta}\|_2 \leq \|\bar{\theta} - \widehat{\theta}\|_2$ we get

$$\|\widehat{\theta} - \theta\|_2 \leq 2 \frac{\|\theta\|_2}{\|\widehat{\theta}\|_2} \|\widehat{\theta} - \bar{\theta}\|_2. \quad (35)$$

It has been considered that $\|\widehat{\theta}\|_2$ is greater than or equal to an a priori user defined value. Denote b this user defined value, this gives

$$\|\widehat{\theta} - \theta\|_2 \leq 2 \frac{\|\theta\|_2}{b} \|\widehat{\theta} - \bar{\theta}\|_2. \quad (36)$$

Together with (33) we obtain

$$\|\widehat{\theta} - \theta\|_2 \leq 2 \frac{\|\theta\|_2}{b} \frac{\sqrt{n+1}}{a} \left(\|\widehat{\mathbf{N}} - \mathbf{N}\|_2 + \|\widehat{\mathbf{M}} - \mathbf{M}\|_F \|\bar{\theta}\|_2 \right) \quad (37)$$

which gives

$$\|\widehat{\theta} - \bar{\theta}\|_2^2 \leq 8 \frac{\|\theta\|_2^2}{b^2} \frac{n+1}{a^2} \left(\|\widehat{\mathbf{N}} - \mathbf{N}\|_2^2 + \|\widehat{\mathbf{M}} - \mathbf{M}\|_F^2 \|\bar{\theta}\|_2^2 \right) \quad (38)$$

and then

$$N \mathbb{E} \left[\|\widehat{\theta} - \bar{\theta}\|_2^2 \right] \leq 8 \frac{\|\theta\|_2^2}{b^2} \frac{n+1}{a^2} \left(N \mathbb{E} \left[\|\widehat{\mathbf{N}} - \mathbf{N}\|_2^2 \right] + N \mathbb{E} \left[\|\widehat{\mathbf{M}} - \mathbf{M}\|_F^2 \right] \|\bar{\theta}\|_2^2 \right). \quad (39)$$

Note that from Assumption A.6 $\{u_t\}$ is ρ -mixing process, it follows from ([22]) that for all j , $\widehat{R}_{uu}(j)$ is mean-square convergent and there exists Σ_j^{uu} such that

$$\lim_{N \rightarrow \infty} N \mathbb{E} \left[(\widehat{R}_{uu}(j) - R_{uu}(j))^2 \right] = \Sigma_j^{uu}. \quad (40)$$

It follows that there exists $\Sigma^{\mathbf{M}}$ such that

$$\lim_{N \rightarrow \infty} N \mathbb{E} \left[\|\widehat{\mathbf{M}} - \mathbf{M}\|_F^2 \right] = \Sigma^{\mathbf{M}}. \quad (41)$$

Note also that $\{y_t\}$ is the output of a FIR filter with $\{u_t\}$ as input, then $\{u_t; y_t\}$ is also a ρ -mixing process. $\{s_t\}$ being the output of a 1-bit ADC with $\{y_t\}$ as input, then it can be shown that $\{u_t; s_t\}$ is a ρ -mixing process too. It follows from ([22]) that for all j $\widehat{R}_{us}(j)$ is mean-square convergent and there exists Σ_j^{us} such that

$$\lim_{N \rightarrow \infty} N \mathbb{E} \left[(\widehat{R}_{us}(j) - R_{us}(j))^2 \right] = \Sigma_j^{us}. \quad (42)$$

It follows that there exists $\Sigma^{\mathbf{N}}$ such that

$$\lim_{N \rightarrow \infty} N \mathbb{E} \left[\|\widehat{\mathbf{N}} - \mathbf{N}\|_2^2 \right] = \Sigma^{\mathbf{N}}. \quad (43)$$

Using (41) and (43) in (39) gives

$$\lim_{N \rightarrow \infty} N \mathbb{E} \left[\|\widehat{\theta} - \theta\|_2^2 \right] \leq 8 \frac{\|\theta\|_2^2}{b^2} \frac{n+1}{a^2} (\Sigma^{\mathbf{N}} + \Sigma^{\mathbf{M}} \|\bar{\theta}\|_2^2). \quad (44)$$

Taking $\Sigma = 8 \frac{\|\theta\|_2^2}{b^2} \frac{n+1}{a^2} (\Sigma^{\mathbf{N}} + \Sigma^{\mathbf{M}} \|\bar{\theta}\|_2^2)$ we get conclusions of the theorem. ■

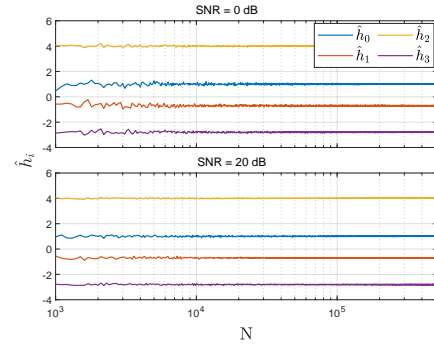


Fig. 3. First experiment: Plots of estimated components of θ as a function of N for $\text{SNR} = 0\text{dB}$ and 20db .

IV. NUMERICAL EXAMPLE

In this section some numerical simulations are depicted so as to show the efficiency of the proposed algorithm and confirm the analysis presented in subsection III-C. We choose the same example as in [23]. The system is then described by $h_0 = 1$, $h_1 = -0.7$, $h_2 = 4$, $h_3 = -2.8$ and $n = 3$. In this example, we assume $\|\theta\|_2$ to be known (Assumption A.4).

The input $\{u_t\}$ is a zero mean colored gaussian noise. It follows that $\{u_t\}$ satisfies Assumption A.1 and $\{u_t\}$ is a ρ -mixing process satisfying Assumption A.5. The noise $\{v_t\}$ is a zero mean white gaussian noise uncorrelated with $\{u_t\}$ (Assumption A.2). The threshold is $C = 1$.

A. First experiment: influence of N

In a first experiment, we investigate the behavior of the algorithm as a function of N . In this experiment the variance of the noise is adjusted so as to have a Signal to Noise Ratio (SNR) equal to 0dB pr 20db. Estimates of all h_i are depicted in Fig. 3 as function of N . These results show that the estimates of system parameters are asymptotically unbiased and are well estimated even if N is not very large.

B. second experiment: influence of the noise level

In a second experiment, we investigate the noise effect on the estimates. A Monte Carlo simulation is carried out with 100 runs for several SNR (from 0dB to 20dB with a step of 5dB) and with $N = 5000$ available data.

The Mean Square Error $\mathbb{E} \left[\|\widehat{\theta} - \theta\|_2^2 \right]$ is depicted in Fig. 4 as function of SNR. We notice that the higher the SNR is, the more accurate the estimate is. We conclude from this experiment that the algorithm performs well even if the noise range is wide.

C. Third experiment: confirmation of the analysis

In a third experiment we confirm the analysis of subsection III-C. A Monte Carlo simulation is carried out with 100 runs for several N (from 10^3 to 10^5 with a step of $5 \cdot 10^3$) and with $\text{SNR} = 0\text{dB}$ and $\text{SNR} = 20\text{dB}$.

Fig.5 depicts the variation of $\mathbb{E} \left[\|\widehat{\theta} - \theta\|_2^2 \right]$ in logarithmic scale as function of N . We notice that these results confirm the previous conclusion with respect the noise impact.

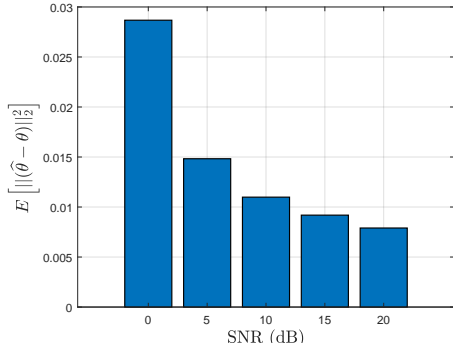


Fig. 4. Second experiment: The Mean Square Error as function of SNR for $N = 5000$.

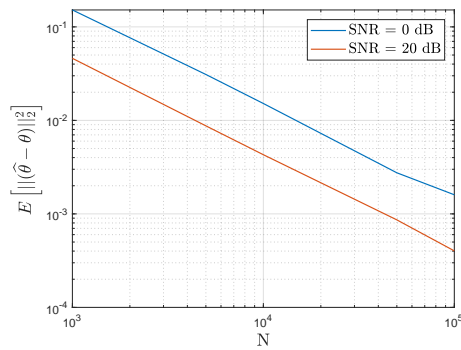


Fig. 5. Third experiment: The Mean Square Error as a function of N for SNR=0dB and SNR=20dB.

Moreover it appears that $\mathbb{E} \left[\|\hat{\theta} - \theta\|_2^2 \right]$ decreases as N^{-1} which confirms (26) in our theorem and the mean-square convergence rate of $\mathcal{O}(N^{-1})$.

V. CONCLUSION

An identification algorithm from binary output measurements is proposed in this paper. The algorithm is adapted to the identification of FIR systems. It is based on the estimation of correlation functions. From an implementation point of view, with respect to other algorithms, the algorithm requires few conditions on the input signal and on the noise, a main interest of the proposed algorithm is then the fact that the practical framework is less restrictive. An algebraic formulation using a SVD is also proposed in order to facilitate the implementation through a robust algebra tool (robust from a numerical implementation point of view). A convergence analysis is provided, it is demonstrated that the algorithm is asymptotically unbiased and mean-square convergent. Numerical simulations confirm the analysis. The extension of the approach to Infinite Impulse Response systems or to an online formulation are subjects for future works.

REFERENCES

[1] L.Y. Wang, J.F. Zhang, and G. Yin. System identification using binary sensors. *IEEE Transactions on Automatic Control*, 48(11):1892–1907, 2003.

[2] Kian Jafari, Jérôme Juillard, and Eric Colinet. A recursive system identification method based on binary measurements. *IEEE Conference on Decision and Control, Atlanta*, 2010.

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

[4] Bruno Depraetere, Julian Stoev, Gregory Pinte, and Jan Swevers. Identification of linear systems with binary outputs using short independent experiments. *IFAC Symposium on System Identification, Brussels*, 2012.

[5] Tianshi Chen, Yanlong Zhao, and Lennart Ljung. Impulse response estimation with binary measurements: A regularized FIR model approach. *IFAC Symposium on System Identification, Brussels*, 2012.

[6] K. Jafari, J. Juillard, and M. Roger. Convergence analysis of an online approach to parameter estimation problems based on binary observations. *Automatica*, 48(11):2837–2842, 2012.

[7] J. Guo and Y. Zhao. Recursive projection algorithm on FIR system identification with binary-valued observations. *Automatica*, 49(11):3396–3401, 2013.

[8] D. Marelli, K. You, and M. Fu. Identification of ARMA models using intermittent and quantized output observations. *Automatica*, 49(2):360–369, 2013.

[9] Abdelhak Goudjil, Mathieu Poulliquen, Eric Pigeon, Olivier Gehan, and Mohammed M’Saad. Identification of systems using binary sensors via support vector machines. *IEEE Conference on Decision and Control, Osaka*, 2015.

[10] Mathieu Poulliquen, Tomas Menard, Eric Pigeon, Olivier Gehan, and Abdelhak Goudjil. Recursive system identification algorithm using binary measurements. *IEEE European Control Conference, Aalborg*, 2016.

[11] Mathieu Poulliquen, Abdelhak Goudjil, Olivier Gehan, and Eric Pigeon. Continuous-time system identification using binary measurements. *IEEE Conference on Decision and Control, Las Vegas*, 2016.

[12] Q. Son. Recursive identification of systems with binary-valued outputs and with ARMA noises. *Automatica*, 93:106–113, 2018.

[13] M. Poulliquen, E. Pigeon, O. Gehan, and A. Goudjil. Identification using binary measurements for IIR systems. *IEEE Transactions on Automatic Control*, 65(2):786–793, 2019.

[14] A. Mestrah, M. Poulliquen, and E. Pigeon. Adaptation of recursive least squares identification for binary output systems. *IEEE International Conference on Systems and Control, Caen*, 2021.

[15] A. Mestrah, M. Poulliquen, E. Pigeon, and H. Oualla. Subspace identification from binary output measurements. *IEEE IFAC International Conference on Control, Decision and Information Technologies, Istanbul*, 2022.

[16] L.Y. Wang, G.G. Yin, Y. Zhao, and J. Zhang. Identification input design for consistent parameter estimation of linear systems with binary-valued output observations. *IEEE Transactions on Automatic Control*, 53(4):867 – 880, 2008.

[17] T. Wang, H. Min, and Z. Yanlong. Adaptive tracking control of FIR systems under binary-valued observations and recursive projection identification. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 51(9):5289–5299, 2021.

[18] Mathieu Poulliquen, Eric Pigeon, Olivier Gehan, Abdelhak Goudjil, and Romain Auber. Impulse response identification from input/output binary measurements. *Automatica*, 123:109307, 2021.

[19] Ying Wang, Yanlong Zhao, Ji-Feng Zhang, and Jin Guo. "a unified identification algorithm of fir systems based on binary observations with time-varying thresholds". *Automatica*, 135:109990, 2022.

[20] Lloyd N Trefethen and David Bau. *Numerical linear algebra*, volume 181. Siam, 1997.

[21] A. Kolmogorov and Y. Rozanov. On strong mixing conditions for stationary gaussian processes. *Theory of Probability and Its Applications*, 5(2):222–227, 1960.

[22] E. Masry. The estimation of the correlation coefficient of bivariate data under dependence: Convergence analysis. *Statistics Probability Letters*, 81(8):1039–1045, 2011.

[23] T. Wigren. Adaptive filtering using quantized output measurements. *IEEE transactions on signal processing*, 46(12):3423–3426, 1998.