Identification of ARMAX models with noisy input and output

Roberto Diversi ∗ Roberto Guidorzi ∗ Umberto Soverini ∗

∗ Department of Electronics, Computer Science and Systems,
University of Bologna, Viale del Risorgimento 2, 40136 Bologna, Italy
(e-mail: roberto.diversi@unibo.it, roberto.guidorzi@unibo.it,
umberto.soverini@unibo.it).

Abstract: ARMAX models are widely used in identification and are a standard tool in control engineering for both system description and control design. These models, however, can be non realistic in many practical contexts because of the presence of measurement errors that play an important role in applications like fault diagnosis and optimal filtering. ARMAX models can be enhanced by introducing also additive error terms on the input and output observations. This scheme, that can be denoted as “ARMAX + noise”, belongs to the errors–in–variables family and allows taking into account the presence of both process disturbances and measurement noise. This paper proposes a three–step procedure for identifying “ARMAX + noise” processes. The first step of the identification algorithm in based on an iterative search procedure while the second and third ones rely on simple least–squares formulas. The paper reports also the results of some Monte Carlo simulations that underline the effectiveness of the proposed approach.

Keywords: System identification; ARMAX models; errors–in–variables models; linear models.

1. INTRODUCTION

ARMAX models are widely used in identification because they allow to describe in an accurate way a large class of real processes with relatively low complexity (Ljung, 1999; Söderström and Stoica, 1989). The stochastic context of ARMAX models is the same considered in Kalman filtering and these models are a standard tool in control engineering for both system description and control design (Goodwin and Sin, 1984).

ARMAX processes belong to the family of equation–error models. In these models the observed output is described as the sum of three regression terms concerning past inputs, outputs and white noise samples. It is possible to interpret ARMAX processes as shown in Fig. 1 (dashed block), i.e. by considering the parallel connection of a deterministic part driven by the observed input $u_0(t)$ and of a stochastic part driven by a remote white process $v(t)$ (Guidorzi, 2003). The deterministic part is characterized by the transfer function $B(z^{-1})/A(z^{-1})$ and its output $y_0(t)$ is not accessible. The stochastic part is characterized by the transfer function $C(z^{-1})/A(z^{-1})$ and its output is a colored noise $v(t)$ that can represent the effect of non accessible disturbances on the state of the deterministic part. The observed output is then $\hat{y}(t) = y_0(t) + v(t)$.

Despite its wide use, this scheme can be non realistic in many practical contexts because of the obvious presence of measurement errors that play an important role in applications like fault diagnosis and optimal filtering. It is thus possible to enhance ARMAX models by introducing also a description of the observation errors on the input and output (see Fig. 1). This scheme, that can be denoted as “ARMAX + noise”, belongs to the errors–in–variables (EIV) family (see (Söderström, 2007) and the references therein) and allows taking into account the presence of both process disturbances and measurement noise.

This paper proposes a three–step identification procedure for identifying ARMAX + noise processes. First the MA part of the ARMAX model is approximated by means of an high–order autoregressive model so that the ARMAX + noise model is mapped into an ARX + noise model of higher order, whose identification has been analyzed in (Diversi et al., 2010). In particular, the technique described in (Diversi et al., 2010) is based on the extension of the dynamic Frisch scheme considered in (Beghelli et al., 1990). Once an high–order ARX model has been identified, the polynomials $A(z^{-1})$, $B(z^{-1})$ are estimated by using the properties of polynomials with common factors. Finally, the MA part $C(z^{-1})$ is estimated by considering the link with its high–order autoregressive approximation. The first
step of the identification algorithm in based on an iterative search procedure while the second and third ones rely on simple least–squares formulas. The effectiveness of the whole procedure has been tested by means of Monte Carlo simulations.

The contents of the paper are organized as follows. Section 2 describes the stochastic context of ARMAX + noise models and defines the identification problem. Section 3 shows how the noisy ARMAX model can be approximated by an auxiliary ARX model of higher order and reports a brief description of the three–step identification procedure. Section 4 summarizes the procedure used for identifying the auxiliary ARX + noise model while Section 5 completes the description of the three–step identification procedure for ARMAX + noise models. The results of some Monte Carlo simulations are described in Section 6 while short concluding remarks are finally given in Section 7.

2. PROBLEM STATEMENT

The upper part of Figure 1 (dashed block) represents an ARMAX model whose output \( y(t) \) is described by the relation
\[
A(z^{-1}) y(t) = B(z^{-1}) u(t) + C(z^{-1}) e(t),
\]
where
\[
A(z^{-1}) = 1 + a_1 z^{-1} + \cdots + a_n z^{-n}
\]
\[
B(z^{-1}) = b_0 + b_1 z^{-1} + \cdots + b_n z^{-n}
\]
\[
C(z^{-1}) = 1 + c_1 z^{-1} + \cdots + c_n z^{-n}
\]
are polynomials in the backward shift operator \( z^{-1} \), i.e. \( z^{-1} x(t) = x(t-1) \). Both the input and output of the ARMAX model are corrupted by additive noise so that the available observations are given by
\[
u(t) = u_0(t) + \tilde{u}(t)
\]
\[
y(t) = \tilde{y}(t) + \tilde{y}(t).
\]

The noisy ARMAX model (1)–(6) can be viewed as an errors–in–variables (EIV) model where:

- the true system, whose input and output are \( u_0(t) \) and \( y_0(t) \), is described by the equation
\[
y_0(t) = \frac{B(z^{-1})}{A(z^{-1})} u_0(t);
\]
- the noise–free input \( u_0(t) \) is affected by the measurement error \( \tilde{u}(t) \);
- the noise–free output \( y_0(t) \) is affected by two noise contributions, a measurement error \( \tilde{y}(t) \) and a process disturbance \( v(t) \)
\[
v(t) = \frac{C(z^{-1})}{A(z^{-1})} e(t).
\]

In fact, relation (6) can be rewritten as
\[
y(t) = y_0(t) + v(t) + \tilde{y}(t).
\]

This model is of particular interest in fault diagnosis and optimal filtering problems. The following assumptions will be considered as satisfied in the sequel.

A1. All zeros of \( A(z) \) and \( C(z) \) are outside the unit circle.
A2. \( A(z) \) and \( B(z) \) have no common factors.
A3. The order \( n \) is a priori known.
A4. The known input \( u_0(t) \) is a zero–mean ergodic random signal or a deterministic quasi–stationary signal (Ljung, 1999) and is persistently exciting of sufficient order.
A5. \( e(t), \tilde{u}(t) \) and \( \tilde{y}(t) \) are zero–mean ergodic white processes with unknown variances \( \sigma^2_e, \sigma^2_u \) and \( \sigma^2_y \).
A6. \( e(t), \tilde{u}(t) \) and \( \tilde{y}(t) \) are mutually uncorrelated and uncorrelated with \( u_0(t) \).

Under these assumptions, the ARMAX + noise identification problem can be stated as follows.

**Problem 1.** Given the set of noisy observations \( u(1), \ldots, u(N), y(1), \ldots, y(N) \), estimate the coefficients \( a_k (k = 1, \ldots, n), b_k (k = 0, \ldots, n), c_k (k = 1, \ldots, n) \) and the variances \( \sigma^2_e, \sigma^2_u, \sigma^2_y \).

**Remark 1.** It is well known that EIV models may not be uniquely identifiable when only the second order statistics are considered, see (Anderson and Deistler, 1984; Aguero and Goodwin, 2008) for a comprehensive treatment of this topic. Note that the ARMAX + noise model (1)–(6) belongs to the class of EIV models considered in (Aguero and Goodwin, 2008). In fact, from (9)
\[
y(t) = \frac{B(z^{-1})}{A(z^{-1})} u_0(t) + \frac{C(z^{-1})}{A(z^{-1})} e(t) + \tilde{y}(t)
\]
\[
= \frac{B(z^{-1})}{A(z^{-1})} u_0(t) + e_y(t),
\]
where \( e_y(t) \) is an additive disturbance that, by using the spectral factorization theorem (Söderström and Stoica, 1989), can be uniquely represented by means of the ARMA model
\[
e_y(t) = \frac{F(z^{-1})}{A(z^{-1})} e(t),
\]
where the stable polynomial \( F(z^{-1}) \) of degree \( n \) and the variance \( \sigma^2_e \) of the white process \( e(t) \) are given by
\[
\sigma^2_e F(z^{-1}) F(z) = \tilde{\sigma}^2_y A(z^{-1}) A(z) + \sigma^2_u C(z^{-1}) C(z).
\]

3. MAPPING THE NOISY ARMAX MODEL INTO A NOISY ARX MODEL

The identification problem under investigation will be solved by mapping the ARMAX + noise model into an ARX + noise model, whose identification has been analyzed in (Diversi et al., 2010). To this end, the following approximation will be introduced
\[
C(z^{-1}) e(t) \approx \frac{1}{D(z^{-1})} e(t),
\]
where
\[
D(z^{-1}) = 1 + d_1 z^{-1} + \cdots + d_{n_d} z^{-n_d}
\]
i.e., the moving average process \( C(z^{-1}) e(t) \) is approximated by an autoregressive model of suitably high order \( n_d \) (Durbin, 1959). Note that this technique is frequently used in many ARMA and ARMAX identification procedures (Guidorzi, 2003; Söderström and Stoica, 1989).

By defining the polynomials of degree \( \tilde{n} = n + n_d \)
\[
\tilde{A}(z^{-1}) = A(z^{-1}) D(z^{-1})
\]
\[
\tilde{B}(z^{-1}) = B(z^{-1}) D(z^{-1}),
\]
with coefficients
$\hat{A}(z^{-1}) = 1 + \alpha_1 z^{-1} + \cdots + \alpha_n z^{-n}$ \hspace{1cm} (17)

$\hat{B}(z^{-1}) = \beta_0 + \beta_1 z^{-1} + \cdots + \beta_n z^{-n}$, \hspace{1cm} (18)

it is possible to rewrite (1), (5), (6) as

$\hat{A}(z^{-1}) \hat{y}(t) = \hat{B}(z^{-1}) u_0(t) + e(t)$ \hspace{1cm} (19)

$a(t) = u_0(t) + \hat{u}(t)$ \hspace{1cm} (20)

$y(t) = \hat{y}(t) + \hat{g}(t)$, \hspace{1cm} (21)

which is an ARX model with noisy input and output.

The identification procedure solving Problem 1 consists in the following steps.

(1) Estimation of the high–order ARX model (19)–(21) and of the variances $\sigma_{g^2}^2, \sigma_u^2, \sigma_y^2$.

(2) Estimation of $A(z^{-1})$ and $B(z^{-1})$ by using the estimates of $\hat{A}(z^{-1}), \hat{B}(z^{-1})$.

(3) Estimation of $C(z^{-1})$ from the estimates of $\hat{A}(z^{-1}), A(z^{-1})$ obtained in steps (1) and (2).

Note that the above procedure does not require the estimation of $D(z^{-1})$.

**Remark 2.** When $\sigma_{g^2}^2 = 0$ the model under investigation becomes an ARMAX model with two distinct sources of noise on the output. The identification of this model can be performed by using an ad hoc technique (Diversi et al., 2009).

### 4. IDENTIFICATION OF THE HIGH–ORDER AUXILIARY ARX MODEL

The first step of the identification procedure, i.e. the estimation of the auxiliary ARX + noise model (19)–(21), will be performed by using the approach described in (Diversi et al., 2010).

By defining the vectors

$\varphi_0(t) = [-\hat{y}(t) \ldots -\hat{y}(t-n) u_0(t) \ldots u_0(t-n)]^T$ \hspace{1cm} (22)

$\varphi(t) = [-y(t) \ldots -y(t-n) u(t) \ldots u(t-n)]^T$ \hspace{1cm} (23)

$\varphi(t) = [-\hat{y}(t) \ldots -\hat{y}(t-n) \hat{u}(t) \ldots \hat{u}(t-n)]^T$ \hspace{1cm} (24)

$\varphi_e(t) = [e(t) 0 \ldots 0]_T$ \hspace{1cm} (25)

and the parameter vector

$\gamma_0^* = [1 \alpha_1 \cdots \alpha_n \beta_0 \cdots \beta_n]^T = [1 \gamma^+]^T$, \hspace{1cm} (26)

model (19)–(21) can also be written in the form

$(\varphi_0^T(t) + \varphi_e^T(t)) \gamma_0^* = 0, \quad \varphi(t) = \varphi_0(t) + \varphi_e(t)$. \hspace{1cm} (28)

Define also the covariance matrices

$\Sigma = E[\varphi(\varphi^T)]$ \hspace{1cm} (29)

$\Sigma = E\left[(\varphi_0(t) + \varphi_e(t))(\varphi_0(t) + \varphi_e(t))^T\right]$, \hspace{1cm} (30)

where $E[\cdot]$ denotes the mathematical expectation. From (28) and Assumptions A5–A6 it follows that

$\Sigma = E[\varphi_0(t) \varphi_0^T(t)] + E[\varphi(t) \varphi^T(t)]$, \hspace{1cm} (31)

where

$E[\varphi(t) \varphi^T(t)] = \left[\begin{array}{cc} \sigma_{g^2}^2 I_{n+1} & 0 \\ 0 & \sigma_u^2 I_{n+1} \end{array}\right]$. \hspace{1cm} (32)

Since $E[\hat{y}(t) e(t)] = \sigma_{g^2}^2$ it is also easy to show that

$\Sigma = E[\varphi_0(t) \varphi_0^T(t)] = \left[\begin{array}{cc} \sigma_{g^2}^2 & 0 \\ 0 & \sigma_u^2 \end{array}\right]$. \hspace{1cm} (33)

and, because of (27)

$\Sigma \gamma_0^* = 0$. \hspace{1cm} (34)

Finally, by combining (31) and (33) it is possible to write

$\Sigma = \Sigma + \tilde{\Sigma}^*$, \hspace{1cm} (35)

where

$\tilde{\Sigma}^* = \left[\begin{array}{cc} \sigma_{g^2}^2 + \sigma_u^2 & 0 \\ 0 & \sigma_u^2 \end{array}\right]$. \hspace{1cm} (36)

Consider now the following problem, which is an extension of the dynamic errors–in–variables problem considered in (Beghelli et al., 1990).

**Problem 2.** Determine the family of all non–negative definite diagonal matrices $\Sigma$ of the type

$\Sigma = \text{diag} \left[\sigma_{y^2}^2 + \sigma_e^2, \sigma_{y^2}^2, \cdots, \sigma_{y^2}^2, \sigma_{e^2}^2, \cdots, \sigma_{e^2}^2\right]$ \hspace{1cm} (37)

such that

$\Sigma - \tilde{\Sigma} \geq 0, \quad \min \text{eig} (\Sigma - \tilde{\Sigma}) = 0$. \hspace{1cm} (38)

The solution of this problem is described by the following results (Diversi et al., 2010).

**Theorem 1.** The set of all diagonal matrices of type (37) satisfying condition (38) defines the points $P = (\sigma_{y^2}^2, \sigma_{y^2}^2, \sigma_{e^2}^2)$ of a surface $S(\Sigma)$ belonging to the first orthant of the noise space $\mathbb{R}^3$. Every point $P$ of $S(\Sigma)$ can be associated with a unique coefficient vector $\gamma_0(P)$ satisfying the relation

$\Sigma(P) \gamma_0(P) = 0$, \hspace{1cm} (39)

where

$\Sigma(P) = \Sigma - \text{diag} \left[\sigma_{y^2}^2 + \sigma_e^2, \sigma_{y^2}^2, \cdots, \sigma_{y^2}^2, \sigma_{e^2}^2, \cdots, \sigma_{e^2}^2\right]$. \hspace{1cm} (40)

$\gamma_0(P) = [1 \alpha_1(P) \cdots \alpha_n(P) \beta_0(P) \cdots \beta_n(P)]^T$. \hspace{1cm} (41)

**Corollary 1.** The point $P^* = (\sigma_{y^2}^2, \sigma_{y^2}^2, \sigma_{e^2}^2)$, associated with the true variances of $\hat{u}(t), y(t)$ and $e(t)$ belongs to $S(\Sigma)$ and the coefficient vector $\gamma_0(P^*)$ is characterized by the true parameters, i.e. $\gamma_0(P^*) = \gamma_0^*$.\hspace{1cm} (42)

Figure 2 shows a typical shape of $S(\Sigma)$.
To identify the point $P^*$ (the model $\gamma^*_0$) on $S(\Sigma)$ it is necessary to introduce a suitable selection criterion. To this end, define the $\eta \times 1$ ($\eta \geq 1$) vectors
\[
\varphi^p_u(t) = [u_0(t - \bar{n} - 1) \ldots u_0(t - \bar{n} - \eta)]^T, \\
\varphi^q_u(t) = [u(t - \bar{n} - 1) \ldots u(t - \bar{n} - \eta)]^T, \\
\varphi^q_y(t) = [\bar{u}(t - \bar{n} - 1) \ldots \bar{u}(t - \bar{n} - \eta)]^T,
\]
that, because of (20), satisfy the condition
\[
\varphi^q_y(t) = \varphi^p_u(t) + \varphi^q_y(t).
\]
Define also the covariance matrix
\[
\Sigma^\eta E = E[\varphi^p_u(t) \varphi^p(t)].
\]
Because of (45) and assumptions A5–A6, we have
\[
\Sigma^\eta = E[(\varphi^p_u(t) + \varphi^q_u(t)) (\varphi^p(t) + \varphi^q(t))^T] = E[\varphi^p_u(t) \varphi^p(t)] = E[\varphi^p_u(t) (\varphi^p(t) + \varphi^q(t))^T],
\]
so that from (27)
\[
\Sigma^\eta \gamma^*_0 = 0.
\]
Because of the above relation, the search for the point $P^*$ on $S(\Sigma)$ can be performed by minimizing the cost function
\[
J(P) = ||\Sigma^\eta \gamma(P)||^2 = \gamma^T(P) \Sigma^\eta \gamma(P), \quad P \in S(\Sigma),
\]
that exhibits the following properties
\[
i) J(P) \geq 0, \\
ii) J(P^*) = 0.
\]
In practice, since only a finite number $N$ of data is available, the matrices $\Sigma$ and $\Sigma^\eta$ must be replaced by the sample estimates
\[
\hat{\Sigma} = \frac{1}{N - \bar{n} - \eta} \sum_{t=\bar{n}+\eta+1}^{\bar{n}+N} \varphi^p(t) \varphi^p(t)^T, \\
\hat{\Sigma}^\eta = \frac{1}{N - \bar{n} - \eta} \sum_{t=\bar{n}+\eta+1}^{\bar{n}+N} \varphi^p_u(t) \varphi^p(t).
\]
The high–order ARX model $\gamma^*_0$ and the noise variances $\sigma^2_{\varphi^p}$, $\sigma^2_{\varphi^q}$ are thus identified by minimizing $J(P)$ on $S(\hat{\Sigma})$, see (Diversi et al., 2010) for an efficient implementation of the identification algorithm.

### 5. IDENTIFICATION OF THE ARMAX MODEL

Once an estimate $\hat{\gamma}_0$ of $\gamma^*_0$ has been obtained, the second step of the procedure, i.e. the estimation of $A(z^{-1})$ and $B(z^{-1})$, can be performed by taking advantage of the properties of polynomials with common factors.

Multiplying (15) by $B(z^{-1})$ and (16) by $A(z^{-1})$ it is easy to obtain
\[
\hat{A}(z^{-1}) B(z^{-1}) - \hat{B}(z^{-1}) A(z^{-1}) = 0.
\]
The above relation can be written in the matrix form
\[
S^T \theta_0 = 0,
\]
where
\[
\theta_0 = [a_1 \ldots a_n b_0 \ldots b_n]^T = [1 \theta^T]^T
\]
and $S$ is the $(2n + 2) \times (\bar{n} + n + 1)$ Sylvester matrix
\[
S = \begin{bmatrix}
\beta_0 & \beta_1 & \ldots & \beta_n & 0 & \ldots & 0 \\
0 & \beta_0 & \beta_1 & \ldots & \beta_n & 0 & \ldots & 0 \\
\vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\
0 & \ldots & 0 & \beta_0 & \beta_1 & \ldots & \beta_n \\
-1 & -\alpha_1 & \ldots & -\alpha_n & 0 & \ldots & 0 \\
0 & -1 & -\alpha_1 & \ldots & -\alpha_n & 0 & \ldots & 0 \\
\vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\
0 & \ldots & 0 & -1 & -\alpha_1 & \ldots & -\alpha_n \\
\end{bmatrix}.
\]

By partitioning $S^T$ as
\[
S^T = [m \ M],
\]
where $m$ is the first column of $S^T$ and taking into account (54) it follows that (Stoica and Söderström, 1997)
\[
m + M \theta = 0.
\]
An estimate of $\theta$ can thus be computed as
\[
\hat{\theta} = -\hat{M}^T M^{-1} \hat{M}^T \hat{m},
\]
where $\hat{M}$ and $\hat{m}$ are constructed with the entries of $\hat{\gamma}$.

To estimate $C(z^{-1})$ (third step of the identification procedure) consider now the approximate relation linking $A(z^{-1})$, $A(z^{-1})$ and $C(z^{-1})$
\[
C(z^{-1}) = \frac{1}{A(z^{-1})},
\]
By expanding the relation $A(z^{-1}) C(z^{-1}) = A(z^{-1})$ and equating the coefficients of equal powers, it is possible to obtain the following relation (Graupe et al., 1975)
\[
\psi = \hat{G} \theta_C
\]
and
\[
\theta_C = [c_1 \ldots c_n]^T
\]
where
\[
\psi = \begin{bmatrix}
a_1 - \alpha_1 \\
a_2 - \alpha_2 \\
\vdots \\
-a_n \\
-a_{n+1} - \alpha_{n+1} \\
\vdots \\
a_{n+1} - \alpha_1 \\
\vdots \\
-a_n \\
\vdots \\
-a_n - \alpha_n - \ldots - \alpha_n - n + n
\end{bmatrix},
\]
\[
\hat{G} = \begin{bmatrix}
1 & 0 & \ldots & 0 \\
\alpha_1 & 1 & 0 & \ldots & 0 \\
\vdots & \vdots & \ddots & \ddots & \ddots \\
\alpha_{n-2} & \alpha_{n-1} & \ldots & 1 & 0 \\
\alpha_{n-1} & \alpha_{n-2} & \ldots & \alpha_1 & 1 \\
\vdots & \vdots & \ddots & \ddots & \ddots \\
\alpha_n & \ldots & \ldots & \ldots & \ldots \\
\end{bmatrix}.
\]
The coefficients of $C(z^{-1})$ can thus be estimated as follows
\[
\hat{\theta}_C = (\hat{G}^T \hat{G})^{-1} \hat{G}^T \hat{\psi},
\]
where $\hat{\psi}$ and $\hat{G}$ have been constructed with the entries of $\hat{\gamma}$ and $\hat{\theta}$.

The following algorithm summarizes the whole identification procedure.

**Algorithm 1.**

1. Compute, on the basis of the available observations, the sample estimates of $\Sigma$ and $\Sigma^\eta$ given by (50) and (51).
2. Compute an estimate $\hat{\gamma}_0$ of the high–order ARX model $\gamma^*_0$ and of the noise variances $\sigma^2_{\varphi^p}$, $\sigma^2_{\varphi^q}$, $\sigma^2_{\varphi^y}$ by means of the EIV identification procedure summarized in Section 4.
3. Construct, with the entries of $\hat{\gamma}_0$ the matrix (55) and partition it as in (56). Compute then an estimate $\hat{\theta}$ of $\theta$ by using (58).
The proposed identification algorithm has been tested on sequences generated by the following ARMAX model:

\[ A(z^{-1}) B(z^{-1}) C(z^{-1}) \]

Remark 3. An alternative approach for identifying the noisy ARMAX model can be based on relation (48). It constitutes, in fact, a set of \( \eta \) high–order Yule–Walker equations that can be directly used to estimate the parameter vector \( \hat{\gamma}_0 \). By partitioning \( \hat{\gamma}^T \) as follows

\[ \hat{\gamma}^T = [ \hat{\gamma}_0 \hat{R} ] \]

where \( \hat{\gamma} \) is a column and using (26), it is possible to compute the estimate

\[ \hat{\gamma}^{IV} = - \left( \hat{R}_T \hat{R} \right)^{-1} \hat{R}_T \hat{\gamma} \]

Remark 4. Taking into account (10) it is possible to consider the ARMAX + noise model as an errors–in–variables model with white input noise and colored output noise so that EIV techniques like those proposed in (Zheng, 2002; Söderström, 2008) could be used. When dealing with EIV models, the aforementioned methods are indeed more general than the proposed one since they can be applied to arbitrarily colored noises on the output. Nevertheless, they do not take into account the specific structure of noisy ARMAX models and lead to the identification of \( A(z^{-1}), B(z^{-1}) \) and \( \sigma^2_e \) but not of \( C(z^{-1}), \sigma^2_e \) and \( \sigma^2_y \).

6. NUMERICAL EXAMPLE

The proposed identification algorithm has been tested on sequences generated by the following ARMAX model:

\[ A(z^{-1}) B(z^{-1}) C(z^{-1}) \]

Table 1. True and estimated values of the coefficients of \( A(z^{-1}), B(z^{-1}) \). For each value of \( \eta \) a Monte Carlo simulation of 100 runs has been performed.

<table>
<thead>
<tr>
<th>( \eta )</th>
<th>( A_1 )</th>
<th>( A_2 )</th>
<th>( B_1 )</th>
<th>( B_2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alg (1)</td>
<td>-0.4983 ± 0.0193</td>
<td>0.2996 ± 0.0149</td>
<td>1.1955 ± 0.0282</td>
<td>-0.6973 ± 0.0312</td>
</tr>
<tr>
<td>Alg (2)</td>
<td>-0.4982 ± 0.0185</td>
<td>0.3000 ± 0.0142</td>
<td>1.1875 ± 0.0246</td>
<td>-0.6926 ± 0.0273</td>
</tr>
<tr>
<td>IV (29)</td>
<td>-0.1511 ± 0.2576</td>
<td>0.2016 ± 0.1977</td>
<td>1.0755 ± 0.2204</td>
<td>-0.3192 ± 0.2745</td>
</tr>
<tr>
<td>IV (60)</td>
<td>-0.3706 ± 0.1490</td>
<td>0.2224 ± 0.1301</td>
<td>1.1108 ± 0.1002</td>
<td>-0.5258 ± 0.1897</td>
</tr>
<tr>
<td>IV (100)</td>
<td>-0.4533 ± 0.1028</td>
<td>0.2796 ± 0.0786</td>
<td>1.1249 ± 0.0688</td>
<td>-0.6257 ± 0.1367</td>
</tr>
</tbody>
</table>

Table 2. True and estimated values of the coefficients of \( C(z^{-1}) \). For each value of \( \eta \) a Monte Carlo simulation of 100 runs has been performed.

<table>
<thead>
<tr>
<th>( \eta )</th>
<th>( c_1 )</th>
<th>( c_2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alg (1)</td>
<td>-0.9095 ± 0.2536</td>
<td>0.2579 ± 0.2448</td>
</tr>
<tr>
<td>Alg (2)</td>
<td>-0.9021 ± 0.1526</td>
<td>0.1632 ± 0.1468</td>
</tr>
<tr>
<td>IV (29)</td>
<td>-0.1478 ± 0.3767</td>
<td>-0.1073 ± 0.3064</td>
</tr>
<tr>
<td>IV (60)</td>
<td>-0.5401 ± 0.2102</td>
<td>-0.0587 ± 0.2471</td>
</tr>
<tr>
<td>IV (100)</td>
<td>-0.6630 ± 0.1654</td>
<td>0.0088 ± 0.1737</td>
</tr>
</tbody>
</table>

The ARMAX + noise models have been identified by using both Algorithm 1 and the IV estimator (65). The MA part \( C(z^{-1}) e(t) \) of the ARMAX model has been approximated by means of an AR model \( e(t)/D(z^{-1}) \) of order \( n_d = 12 \). Monte Carlo simulations of 100 independent runs have been performed by using different values of the user-chosen parameter \( \eta \). Note that the IV estimator (65) requires \( \eta \geq 29 \). The obtained results are summarized in Tables 1–3 that report the true values of parameters and variances, the means of their estimates and the associated standard deviations.

It can be observed that the performance of Algorithm 1 is remarkably better than that of the IV estimator. Moreover, the IV estimator leads to satisfactory results only when a very large number of equations \( \eta \) is used. It has also been observed that the use of larger values of \( n_d \) increases the computational burden without improving the estimation accuracy.
Table 3. True and estimated values of $\sigma_0^2$, $\tilde{\sigma}_u^2$ and $\tilde{\sigma}_y^2$. For each value of $\eta$ a Monte Carlo simulation of 100 runs has been performed.

<table>
<thead>
<tr>
<th>true value</th>
<th>$\sigma_0^2$</th>
<th>$\tilde{\sigma}_u^2$</th>
<th>$\tilde{\sigma}_y^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alg. 1 ($\eta = 10$)</td>
<td>0.0970 ± 0.0452</td>
<td>0.0250 ± 0.0218</td>
<td>0.0303 ± 0.0098</td>
</tr>
<tr>
<td>Alg. 1 ($\eta = 20$)</td>
<td>0.1119 ± 0.0394</td>
<td>0.0184 ± 0.0191</td>
<td>0.0094 ± 0.0095</td>
</tr>
</tbody>
</table>

7. CONCLUSIONS

By considering additive error terms on the input–output observations of a traditional ARMAX model the family of “ARMAX + noise” models can be introduced. These models can offer a more realistic choice than simple equation error models since they can be seen as errors–in–variables models where the noise–free input is affected by a measurement error while the noise–free output is affected by two noise contributions, a measurement error and a process disturbance.

The identification of these models has been performed by means of a three–step procedure whose first step relies on the solution of the ARX + noise identification problem analyzed in (Diversi et al., 2010). The second and third steps consist in the application of simple least–squares formulas. The Monte Carlo simulations that have been carried out show the effectiveness of the whole procedure.

REFERENCES


