Proceedings of the 17th World Congress
The International Federation of Automatic Control
Seoul, Korea, July 6-11, 2008

Identification of ARARX models in presence of additive noise

Roberto Diversi * Roberto Guidorzi * Umberto Soverini *

* Dipartimento di Elettronica, Informatica e Sistemistica, 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: The identification of dynamic processes can be performed by means of different classes of models relying on different stochastic environments to describe the misfit between the model and process observations. This paper introduces a new class of models by considering additive error terms on the observations of the input and output of ARARX models and proposes a three–step identification procedure for their identification. ARARX + noise models extend the traditional ARARX or ARMAX ones and can be seen as errors–in–variables models where both measurement errors and process disturbances are taken into account. The results of Monte Carlo simulations show the good performance of the proposed identification procedure.

Keywords: System identification, errors–in–variables models, ARARX models, linear systems

1. INTRODUCTION

The modeling of a dynamic process on the basis of observed sequences i.e. its identification, can rely on many families of possible models, describing different stochastic environments, as well as on different selection criteria within a specified class of models. The choice of model families and criteria is often based more on the planned use of the model then on the adherence of the associated stochastic contexts to real ones because real processes are in general more complex than the representations used for their description.

Equation error models constitute a very useful category of models because of their applicability in prediction and control (Söderström and Stoica, 1989; Ljung, 1999); the description of the misfit between model and observations only by means of an error term on the output is, however, restrictive.

Errors–in–Variables (EIV) models are a class of models based on the assumption that the process behind the data can be described by means of a linear model whose observations are corrupted by additive errors, see (Söderström, 2007) and the references therein. These models are often more realistic because all measures are considered as affected by errors.

This paper considers a new family of models that derives from the integration of EIV models and ARARX ones. Inside the class of equation error models, ARARX are very peculiar since they can be considered as an extension of ARX models and can approximate, at any desired degree, the family of ARMAX models (Guidorzi, 2003; Söderström and Stoica, 1989). This characteristic leads to the use of ARARX processes also in model reduction (Söderström et al., 1991; Tjärnström and Ljung, 2003).

ARARX + noise models consider additive error terms on the observations of the input and output of an ARARX model. In this way, it is possible to obtain representations that take into account both measurement errors and process disturbances. This feature is particularly suitable for fault detection and filtering purposes.

This paper proposes a three–step identification procedure for identifying ARARX + noise models. The first step concerns the identification of an auxiliary high–order ARX model and is based on the results reported in (Diversi et al., 2007). The second and third steps take advantage of the properties of polynomials with common factors and consist in simple least–squares algorithms. The proposed method has been tested by means of Monte Carlo simulations and compared with an instrumental variable approach.

The organization of the paper is as follows. Section 2 contains a description of the considered stochastic context and the statement of the identification problem. Section 3 describes the steps to be performed in the identification procedure. Section 4 concerns the identification of an auxiliary high–order ARX model while the complete ARARX + noise identification procedure is described in Section 5. In Section 6 the ARARX + noise identification problem is solved by using an instrumental variable approach. Section 7 reports some numerical results while short concluding remarks are finally given in Section 8.

2. CONTEXT AND STATEMENT OF THE PROBLEM

Consider a linear, single input single output, discrete time ARARX model described by the equation

$$A(q^{-1})\bar{y}(t) = B(q^{-1})u_0(t) + \frac{e(t)}{D(q^{-1})}, \tag{1}$$

where $A(q^{-1})$, $B(q^{-1})$ and $D(q^{-1})$ are polynomials in the backward shift operator q^{-1}

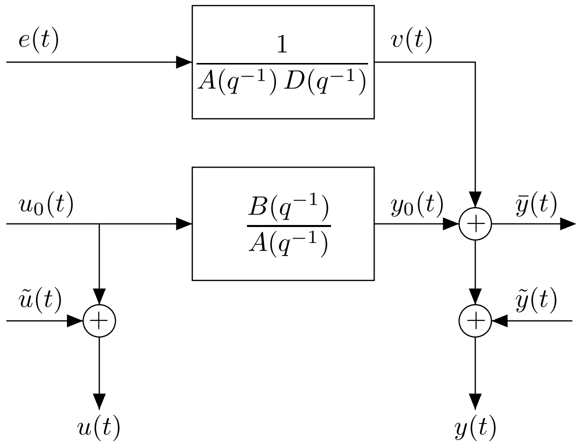

Fig. 1. ARARX model with noisy input and output.

$$A(q^{-1}) = 1 + a_1\, q^{-1} + \cdots + a_n\, q^{-n} \tag{2}$$
$$B(q^{-1}) = b_0 + b_1\, q^{-1} + \cdots + b_n\, q^{-n} \tag{3}$$
$$D(q^{-1}) = 1 + d_1\, q^{-1} + \cdots + d_{n_d}\, q^{-n_d}. \tag{4}$$

The ARARX structure admits the following interpretations:

- An equation error model, with input $u_0(t)$ and output $\bar{y}(t)$,whose equation error is given by the autoregressive process $e(t)/D(q^{-1})$ (see (1)).
- A "true" system $B(q^{-1})/A(q^{-1})$, with input $u_0(t)$ and output $y_0(t)$, whose output is affected by the additive colored noise $v(t) = e(t)/(A(q^{-1})\, D(q^{-1}))$ (see Fig. 1). In this case, $\bar{y}(t)$ denotes the observed output: $\bar{y}(t) = y_0(t) + v(t)$.

Remark 1. Since a moving average process driven by a white noise $\eta(t) = C(q^{-1})\, e(t)$ can be approximated by an autoregessive process of suitable high order, $\eta(t) \approx \frac{1}{D(q^{-1})}\, e(t)$, $v(t)$ can be seen as an approximation of a generic ARMA model $\frac{C(q^{-1})}{A(q^{-1})}\, e(t)$. As a consequence, an ARARX model can approximate an ARMAX structure.

In this paper, we will assume that $u_0(t)$ and $\bar{y}(t)$ are corrupted by the additive noises $\tilde{u}(t)$ and $\tilde{y}(t)$ so that the available signals $u(t), y(t)$ are given by

$$u(t) = u_0(t) + \tilde{u}(t) \tag{5}$$
$$y(t) = \bar{y}(t) + \tilde{y}(t) = y_0(t) + v(t) + \tilde{y}(t). \tag{6}$$

Model (1)–(6) can thus be seen as an errors–in–variables model where the noise–free input $u_0(t)$ is affected by the measurement error $\tilde{u}(t)$ while the noise–free output $y_0(t)$ is affected by two noise contributions, a measurement error $\tilde{y}(t)$ and a process disturbance $v(t)$ whose sum could also be considered as a single colored noise generated by an ARMA process. The separation of the output disturbance into a white noise and a colored one considered in this paper is however useful in the solution of specific problems like, for instance, diagnosis.

Remark 2. Note that the EIV model (1)–(6) can approximate the extended–noise Kalman filter context described in (Diversi et al., 2005), where input, output and state noises are present.

The following assumptions are introduced.

A1. $A(z)$ and $D(z)$ have all zeros outside the unit circle.
A2. $A(z)$ and $B(z)$ do not share any common factor.
A3. The orders n and n_d are assumed as *a priori* known.
A4. The noise–free input $u_0(t)$ is a zero–mean ergodic random signal and is persistently exciting of a suitably high order.
A5. $e(t)$, $\tilde{u}(t)$ and $\tilde{y}(t)$ are zero–mean ergodic white processes with unknown variances σ_e^{2*}, $\tilde{\sigma}_u^{2*}$ and $\tilde{\sigma}_y^{2*}$ respectively. These processes are mutually uncorrelated and uncorrelated with the noise–free input $u_0(t)$.

The problem under investigation is the following.

Problem 1. Given a set of noisy input–output observations $u(1), \ldots, u(N), y(1), \ldots, y(N)$, determine an estimate of the coefficients a_k $(k = 1, \ldots, n)$, b_k $(k = 0, \ldots, n)$, d_k $(k = 1, \ldots, n_d)$ and of the variances σ_e^{2*}, $\tilde{\sigma}_u^{2*}$, $\tilde{\sigma}_y^{2*}$.

3. A THREE–STEP IDENTIFICATION PROCEDURE

By defining the polynomials of degree $\bar{n} = n + n_d$

$$\bar{A}(q^{-1}) = A(q^{-1})\, D(q^{-1}) \tag{7}$$
$$\bar{B}(q^{-1}) = B(q^{-1})\, D(q^{-1}), \tag{8}$$

with coefficients

$$\bar{A}(q^{-1}) = 1 + \alpha_1\, q^{-1} + \cdots + \alpha_{\bar{n}}\, q^{-\bar{n}} \tag{9}$$
$$\bar{B}(q^{-1}) = \beta_0 + \beta_1\, q^{-1} + \cdots + \beta_{\bar{n}}\, q^{-\bar{n}}, \tag{10}$$

it is possible to rewrite (1) as

$$\bar{A}(q^{-1})\, \bar{y}(t) = \bar{B}(q^{-1})\, u_0(t) + e(t). \tag{11}$$

so that model (1)–(6) can be seen also as an ARX process with noisy input and output, whose identification has been treated in (Diversi et al., 2007).

On the basis of the above consideration, we will solve Problem 1 by means of the following steps.

Procedure 1.

(1) Estimation of the high–order ARX model (11) and of the variances σ_e^{2*}, $\tilde{\sigma}_u^{2*}$, $\tilde{\sigma}_y^{2*}$.
(2) Estimation of $A(q^{-1})$ and $B(q^{-1})$ by using the estimates of $\bar{A}(q^{-1}), \bar{B}(q^{-1})$.
(3) Estimation of $D(q^{-1})$ from the estimates obtained in steps (1) and (2).

Let us introduce the regressor vectors

$$\varphi_0(t) = [\,-y_0(t)\ \ldots\ -y_0(t-n)\ u_0(t)\ \ldots\ u_0(t-n)\,]^T \tag{12}$$
$$\varphi(t) = [\,-y(t)\ \ldots\ -y(t-n)\ u(t)\ \ldots\ u(t-n)\,]^T \tag{13}$$
$$\tilde{\varphi}(t) = [\,-\tilde{y}(t)\ \ldots\ -\tilde{y}(t-n)\ \tilde{u}(t)\ \ldots\ \tilde{u}(t-n)\,]^T \tag{14}$$
$$\varphi_v(t) = [\,-v(t)\ \ldots\ -v(t-n)\ \underbrace{0\ \ldots\ 0}_{n+1}\,]^T, \tag{15}$$

and the parameter vector

$$\theta_0 = \begin{bmatrix} 1\ a_1\ \cdots\ a_n\ b_0\ \cdots\ b_n \end{bmatrix}^T = \begin{bmatrix} 1\ \theta^{*T} \end{bmatrix}^T. \tag{16}$$

From $A(q^{-1})\, y_0(t) = B(q^{-1}) u_0(t)$ and (5)–(6) it is possible to rewrite model (1)–(6) as follows

$$\varphi_0^T(t)\, \theta_0 = 0, \tag{17}$$
$$\varphi(t) = \varphi_0(t) + \tilde{\varphi}(t) + \varphi_v(t). \tag{18}$$

Similarly, define the regressor vectors

$$\bar{\phi}(t) = [\,-\bar{y}(t)\ \ldots\ -\bar{y}(t-\bar{n})\, u_0(t)\ \ldots\ u_0(t-\bar{n})\,]^T \quad (19)$$
$$\phi(t) = [\,-y(t)\ \ldots\ -y(t-\bar{n})\, u(t)\ \ldots\ u(t-\bar{n})\,]^T \quad (20)$$
$$\tilde{\phi}(t) = [\,-\tilde{y}(t)\ \ldots\ -\tilde{y}(t-\bar{n})\, \tilde{u}(t)\ \ldots\ \tilde{u}(t-\bar{n})\,]^T \quad (21)$$
$$\phi_e(t) = [\,e(t)\ \underbrace{0\ \ldots\ 0}_{2\bar{n}+1}\,]^T \quad (22)$$

and the parameter vector

$$\mu^* = \left[1\ \alpha_1\ \cdots\ \alpha_{\bar{n}}\ \beta_0\ \cdots\ \beta_{\bar{n}}\,\right]^T. \quad (23)$$

By taking into account (11), model (1)–(6) can also be written in the form

$$\left(\bar{\phi}^T(t) + \phi_e^T(t)\right)\ \mu^* = 0, \quad (24)$$
$$\phi(t) = \bar{\phi}(t) + \tilde{\phi}(t). \quad (25)$$

The vector forms (17)–(18) and (24)–(25) will be both useful in the sequel.

4. IDENTIFICATION OF THE AUXILIARY ARX MODEL

Define the covariance matrices

$$\Sigma = E\,[\phi(t)\,\phi^T(t)] \quad (26)$$
$$\bar{\Sigma}_0 = E\,\left[\bar{\phi}(t)\left(\bar{\phi}^T(t) + \phi_e^T(t)\right)\right], \quad (27)$$

where $E[\,\cdot\,]$ denotes the mathematical expectation. From (25) and assumption A5 it follows that

$$\Sigma = E\,[\bar{\phi}(t)\,\bar{\phi}^T(t)] + E\,[\tilde{\phi}(t)\,\tilde{\phi}^T(t)], \quad (28)$$

where

$$E\,[\tilde{\phi}(t)\,\tilde{\phi}^T(t)] = \begin{bmatrix} \tilde{\sigma}_y^{2*}\, I_{\bar{n}+1} & 0 \\ 0 & \tilde{\sigma}_u^{2*}\, I_{\bar{n}+1} \end{bmatrix}. \quad (29)$$

Since $E\,[\bar{y}(t)\,e(t)] = \sigma_e^{2*}$ it is also easy to show that

$$\bar{\Sigma}_0 = E\,[\bar{\phi}(t)\,\bar{\phi}^T(t)] - \text{diag}\ [\sigma_e^{2*}\ \underbrace{0\ \cdots\ 0}_{2\bar{n}+1}], \quad (30)$$

and, because of (24)

$$\bar{\Sigma}_0\,\mu^* = 0. \quad (31)$$

Finally, by combining (28) and (30) it is possible to write

$$\Sigma = \bar{\Sigma}_0 + \tilde{\Sigma}^*, \quad (32)$$

where

$$\tilde{\Sigma}^* = \begin{bmatrix} \tilde{\sigma}_y^{2*} + \sigma_e^{2*} & & 0 \\ & \tilde{\sigma}_y^{2*}\, I_{\bar{n}} & \\ 0 & & \tilde{\sigma}_u^{2*}\, I_{\bar{n}+1} \end{bmatrix}. \quad (33)$$

The covariance matrix of the noisy data Σ can thus be decomposed into the sum of a positive semidefinite singular matrix $\bar{\Sigma}_0$, whose kernel defines the true parameter vector, and of a diagonal matrix $\tilde{\Sigma}^*$.

Consider now the problem of determining the family of all non–negative definite diagonal matrices $\tilde{\Sigma}$ of type

$$\tilde{\Sigma} = \text{diag}\ [\tilde{\sigma}_y^2 + \sigma_e^2, \tilde{\sigma}_y^2\, I_{\bar{n}}, \tilde{\sigma}_u^2\, I_{\bar{n}+1}] \quad (34)$$

such that

$$\Sigma - \tilde{\Sigma} \geq 0, \qquad \min \text{eig}\ (\Sigma - \tilde{\Sigma}) = 0. \quad (35)$$

This problem, which is an extension of the dynamic errors–in–variables problem considered in (Beghelli et al., 1990), consists in determining the set of points $P = (\tilde{\sigma}_u^2, \tilde{\sigma}_y^2, \sigma_e^2)$ belonging to the first orthant of $\mathcal{R}^3$ satisfying (34)–(35), i.e. leading to positive semidefinite matrices $\bar{\Sigma}_0(P) = \Sigma - \tilde{\Sigma}(P)$ with one eigenvalue equal to zero. This set is decribed by the following results (Diversi et al., 2007).

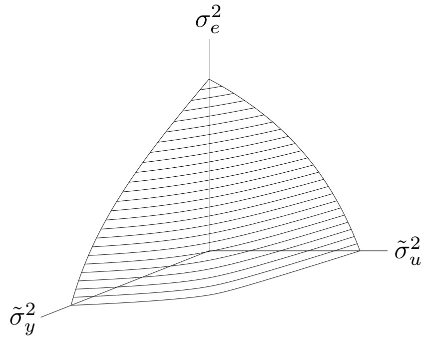

Fig. 2. Typical shape of $\mathcal{S}(\Sigma)$.

Theorem 1. The set of all diagonal matrices satisfying (34)–(35) defines the points $P = (\tilde{\sigma}_u^2, \tilde{\sigma}_y^2, \sigma_e^2)$ of a convex surface $\mathcal{S}(\Sigma)$ belonging to the first orthant of the noise space $\mathcal{R}^3$ whose concavity faces the origin. Every point P of $\mathcal{S}(\Sigma)$ can be associated with a coefficient vector $\mu(P)$ satisfying the relation

$$\bar{\Sigma}_0(P)\,\mu(P) = 0, \quad (36)$$

where

$$\bar{\Sigma}_0(P) = \Sigma - \text{diag}\ [\tilde{\sigma}_y^2 + \sigma_e^2, \tilde{\sigma}_y^2\, I_{\bar{n}}, \tilde{\sigma}_u^2\, I_{\bar{n}+1}] \quad (37)$$
$$\mu(P) = \left[1\,\alpha_1(P)\ \cdots\ \alpha_{\bar{n}}(P)\ \beta_0(P)\ \cdots\ \beta_{\bar{n}}(P)\,\right]^T. \quad (38)$$

A typical shape of $\mathcal{S}(\Sigma)$ is shown in Figure 2.

Corollary 1. The point $P^* = (\tilde{\sigma}_u^{2*}, \tilde{\sigma}_y^{2*}, \sigma_e^{2*})$, associated with the true variances of $\tilde{u}(t)$, $\tilde{y}(t)$ and $e(t)$ belongs to $\mathcal{S}(\Sigma)$ and the coefficient vector $\mu(P^*)$ is characterized (after normalizing its first entry to 1) by the true parameters (23), i.e. $\mu(P^*) = \mu^*$.

In this asymptotic context, the identification of μ^* consists thus in finding, by means of a suitable selection criterion, the point P^* among the set of possible solutions described by $\mathcal{S}(\Sigma)$. Define, for this purpose, the $k \times 1$ ($k \geq 1$) following vectors of delayed signals

$$\phi_{u_0}^k(t) = [\,u_0(t-\bar{n}-1)\ \ldots\ u_0(t-\bar{n}-k)\,]^T \quad (39)$$
$$\phi_u^k(t) = [\,u(t-\bar{n}-1)\ \ldots\ u(t-\bar{n}-k)\,]^T \quad (40)$$
$$\phi_{\tilde{u}}^k(t) = [\,\tilde{u}(t-\bar{n}-1)\ \ldots\ \tilde{u}(t-\bar{n}-k)\,]^T. \quad (41)$$

Because of (5), they satisfy the condition

$$\phi_u^k(t) = \phi_{u_0}^k(t) + \phi_{\tilde{u}}^k(t). \quad (42)$$

Define also the covariance matrix

$$\Sigma^k = E\,[\phi_u^k(t)\,\phi^T(t)]. \quad (43)$$

Because of (42) and assumption A5 we have

$$\Sigma^k = E\,[\phi_{u_0}^k(t)\,\bar{\phi}^T(t)] = E\,[\phi_{u_0}^k(t)\,(\bar{\phi}(t) + \phi_e(t))^T], \quad (44)$$

so that from (24)

$$\Sigma^k\,\mu^* = 0. \quad (45)$$

Relation (45) constitutes a set of high–order Yule–Walker equations that could be directly used to obtain the parameter vector μ^*. This approach can also be viewed as an instrumental variable method that uses delayed inputs as instruments (Söderström and Stoica, 1989). In this paper, equations (45) are used jointly with the results of Theorem 1 and Corollary 1 in order to solve the first step of Procedure 1, i.e. the identification of $\mu^*, \tilde{\sigma}_u^{2*}, \tilde{\sigma}_y^{2*}$ and σ_e^{2*}. In fact, the search for the point P^* on $\mathcal{S}(\Sigma)$ can be performed by minimizing the cost function

$$\begin{aligned} J(P) &= \|\Sigma^k \mu(P)\|_2^2 \\ &= \mu^T(P)(\Sigma^k)^T \Sigma^k \mu(P), \quad P \in \mathcal{S}(\Sigma), \end{aligned} \tag{46}$$

which exhibits the following properties

i) $J(P) \geq 0$
ii) $J(P^*) = 0$.

For the practical implementation of the search procedure it is useful to parameterize the surface $\mathcal{S}(\Sigma)$ in a different way, that allows to associate a solution of (34)–(35) with every straight line departing from the origin and lying in the first orthant of $\mathcal{R}^3$. This parameterization, introduced in (Guidorzi and Pierantoni, 1995), is described in the next theorem.

Theorem 2. Let $\xi = (\xi_1, \xi_2, \xi_3)$ be a generic point of the first orthant of $\mathcal{R}^3$ and r the straight line from the origin through ξ. Its intersection with $\mathcal{S}(\Sigma)$ is the point $P = (\tilde{\sigma}_u^2, \tilde{\sigma}_y^2, \sigma_e^2)$ given by

$$\tilde{\sigma}_u^2 = \frac{\xi_1}{\lambda_M}, \qquad \tilde{\sigma}_y^2 = \frac{\xi_2}{\lambda_M}, \qquad \sigma_e^2 = \frac{\xi_3}{\lambda_M}, \tag{47}$$

where

$$\lambda_M = \max \mathrm{eig}\Big(\Sigma^{-1} \,\mathrm{diag}\, [\xi_2 + \xi_3, \xi_2 I_n, \xi_1 I_{n+1}]\Big) . \tag{48}$$

Previous considerations lead to the following identification algorithm.

Algorithm 1.

(1) Compute, on the basis of the available observations, an estimate of Σ and Σ^k, i.e.

$$\hat{\Sigma} = \frac{1}{N - \bar{n} - k} \sum_{t=\bar{n}+k+1}^{t=N} \phi(t)\, \phi^T(t),$$

$$\hat{\Sigma}^k = \frac{1}{N - \bar{n} - k} \sum_{t=\bar{n}+k+1}^{t=N} \phi_u^k(t)\, \phi^T(t).$$

(2) Start from a generic direction r belonging to the first orthant of $\mathcal{R}^3$.
(3) Compute, by means of (47)–(48), the intersection $P = (\tilde{\sigma}_u^2, \tilde{\sigma}_y^2, \sigma_e^2)$ between r and $\mathcal{S}(\hat{\Sigma})$.
(4) Compute $\bar{\Sigma}_0(P)$ and $\mu(P)$ by means of the relations

$$\bar{\Sigma}_0(P) = \hat{\Sigma} - \mathrm{diag}\, [\tilde{\sigma}_y^2 + \sigma_e^2, \tilde{\sigma}_y^2 I_{\bar{n}}, \tilde{\sigma}_u^2 I_{\bar{n}+1}]$$
$$\bar{\Sigma}_0(P)\, \mu(P) = 0,$$

and normalize the first entry of $\mu(P)$ to 1.
(5) Compute the cost function

$$J(P) = \|\hat{\Sigma}^k \mu(P)\|_2^2. \tag{49}$$

(6) Move to a new direction $r \pm \Delta r$ corresponding to a decrease of $J(P)$.
(7) Repeat steps 3–6 until the point $\hat{P} = (\hat{\tilde{\sigma}}_u^2, \hat{\tilde{\sigma}}_y^2, \hat{\sigma}_e^2)$ associated with the minimum of $J(P)$ is found.
(8) The estimates of the model coefficients and of the noise variances are thus given by $\hat{\mu} = \mu(\hat{P})$ and $\hat{\tilde{\sigma}}_u^2, \hat{\tilde{\sigma}}_y^2, \hat{\sigma}_e^2$.

5. ARARX IDENTIFICATION

In this section, the second and third steps of Procedure 1 are solved starting from the estimate of the ARX coefficients obtained in Section 4. Multiplying (7) by $B(q^{-1})$ and (8) by $A(q^{-1})$ it is easy to show that

$$\bar{A}(q^{-1})\, B(q^{-1}) - \bar{B}(q^{-1})\, A(q^{-1}) = 0. \tag{50}$$

This expression can be written in the matrix form

$$S^T \theta_0 = 0, \tag{51}$$

where S is the $(2n+2) \times (\bar{n} + n + 1)$ Sylvester matrix

$$S = \begin{bmatrix} \beta_0 & \beta_1 & \dots & \beta_{\bar{n}} & 0 & \dots & 0 \\ 0 & \beta_0 & \beta_1 & \dots & \beta_{\bar{n}} & \dots & 0 \\ \vdots & & \ddots & \ddots & & \ddots & \vdots \\ 0 & \dots & 0 & \beta_0 & \beta_1 & \dots & \beta_{\bar{n}} \\ -1 & -\alpha_1 & \dots & -\alpha_{\bar{n}} & 0 & \dots & 0 \\ 0 & -1 & -\alpha_1 & \dots & -\alpha_{\bar{n}} & \dots & 0 \\ \vdots & & \ddots & \ddots & & \ddots & \vdots \\ 0 & \dots & 0 & -1 & -\alpha_1 & \dots & -\alpha_{\bar{n}} \end{bmatrix}. \tag{52}$$

By partitioning S^T as

$$S^T = [\, m \; M \,], \tag{53}$$

where m is the first column of S^T and taking into account (16) it follows that (Stoica and Söderström, 1997)

$$m + M\, \theta^* = 0. \tag{54}$$

An estimate of θ^* can thus be computed as

$$\hat{\theta} = -\big(\hat{M}^T \hat{M}\big)^{-1} \hat{M}^T \hat{m}, \tag{55}$$

where $\hat{M}$ and $\hat{m}$ are constructed with the entries of $\hat{\mu}$.

Finally, once that an estimate of θ^* is available, it is possible to solve step (3) of Procedure 1. In fact, relations (7) and (8) can be jointly written in the matrix form

$$\mu^* = G\, \theta_D^*, \tag{56}$$

where

$$G = \begin{bmatrix} 1 & a_1 & \dots & a_n & 0 & \dots & 0 \\ 0 & 1 & a_1 & \dots & a_n & \dots & 0 \\ \vdots & & \ddots & & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & a_1 & \dots & a_n \\ b_0 & \dots & b_n & 0 & \dots & 0 & \\ 0 & b_0 & \dots & b_n & \dots & 0 & \\ \vdots & & \ddots & & \ddots & \vdots & \\ 0 & \dots & 0 & b_0 & \dots & b_n & \end{bmatrix}^T, \tag{57}$$

and

$$\theta_D^* = \big[1 \; d_1 \; \cdots \; d_n\big]^T. \tag{58}$$

The coefficients of $D(q^{-1})$ can thus be estimated as follows

$$\hat{\theta}_D = \big(\hat{G}^T \hat{G}\big)^{-1} \hat{G}^T \hat{\mu}, \tag{59}$$

where $\hat{G}$ has been constructed with the entries of $\hat{\theta}$.

Procedure 1 can thus be summarized as follows.

Procedure 1.

(1) Estimate the high–order ARX model (11) and the variances σ_e^{2*}, $\tilde{\sigma}_u^{2*}$, $\tilde{\sigma}_y^{2*}$ by means of Algorithm 1. Let $\hat{\mu}$ be the estimate of μ^*.
(2) Construct, with the entries of $\hat{\mu}$, the vector $\hat{m}$ and the matrix $\hat{M}$ as in (52)–(53) and compute an estimate $\hat{\theta}$ of θ^* by using (55).
(3) Construct, with the entries of $\hat{\theta}$, the matrix $\hat{G}$ with structure (57) and compute an estimate of $D(q^{-1})$ by using (59).

6. AN INSTRUMENTAL VARIABLE APPROACH

The coefficients of $A(q^{-1}), B(q^{-1})$ can also be estimated by means of an instrumental variable (IV) approach. In fact, by defining the vector of delayed inputs

$$z(t) = [\,u(t-n-1) \,\dots\, u(t-n-k)\,]^T, \qquad (60)$$

and the covariance matrix

$$\Sigma_{z\varphi} = E\,[z(t)\,\varphi^T(t)], \qquad (61)$$

it follows, from (5), (18) and Assumption A5, that

$$\Sigma_{z\varphi} = E\left[z(t)\,(\varphi_0(t) + \tilde{\varphi}(t) + \varphi_v(t))^T\right] = E\,[z(t)\,\varphi_0^T(t)] \qquad (62)$$

and, because of (17)

$$\Sigma_{z\varphi}\,\theta_0 = 0. \qquad (63)$$

Relation (63) constitutes a system of equations in the unknowns $(a_1, \dots, a_n, b_0, \dots, b_n)$. The matrix $\Sigma_{z\varphi}$ can be estimated from the data as

$$\hat{\Sigma}_{z\varphi} = \frac{1}{N-n-k} \sum_{t=n+k+1}^{t=N} z(t)\,\varphi^T(t). \qquad (64)$$

By partitioning (64) as follows

$$\hat{\Sigma}_{z\varphi} = \left[\hat{r}\ \hat{R}\right], \qquad (65)$$

where $\hat{r}$ is a column, it is possible to obtain an estimate of θ^* by means of the least squares estimator

$$\hat{\theta}_{IV} = -(\hat{R}^T\hat{R})^{-1}\hat{R}^T\hat{r} \qquad (66)$$

which is an extended IV estimator (Söderström and Stoica, 1989). An estimate of $D(q^{-1})$ can be finally obtained by using (59).

This method is more simple from the computational point of view since it does not require an estimation of the auxiliary ARX model (11); however, it is characterized by the following drawbacks:

- IV approaches lead often to poor estimation accuracy (Soverini and Söderström, 2000; Söderström, 2007);
- An estimation of the noise variances cannot be directly obtained by means of IV estimators.

7. NUMERICAL RESULTS

The proposed identification procedure has been tested by means of numerical simulations performed on the following ARARX model, already used in (Tjärnström and Ljung, 2003)

$$A(q^{-1}) = 1 - 0.5\,q^{-1} + 0.06\,q^{-2}$$
$$B(q^{-1}) = q^{-1} - 0.7\,q^{-2}$$
$$D(q^{-1}) = 1 + 0.95\,q^{-1}.$$

The noise–free input is a pseudo random binary sequence with unit variance while the noises $e(t), \tilde{u}(t), \tilde{y}(t)$ are gaussian white noise sequences with variances $\sigma_e^{2*} = 0.1$, $\tilde{\sigma}_u^{2*} = 0.06$ and $\tilde{\sigma}_y^{2*} = 0.02$. These values correspond to signal to noise ratios on the input and output

$$\text{SNRI} \approx 12\text{dB} \qquad \text{SNRO} \approx 4dB,$$

where

$$\text{SNRI} = 20\,\log_{10}\sqrt{\frac{E[u_0^2(t)]}{E[\tilde{u}^2(t)]}} = 20\,\log_{10}\sqrt{\frac{E[u_0^2(t)]}{\tilde{\sigma}_u^{2*}}}$$

$$\text{SNRO} = 20\,\log_{10}\sqrt{\frac{E[y_0^2(t)]}{E[v^2(t)+\tilde{y}^2(t)]}} = 20\,\log_{10}\sqrt{\frac{E[y_0^2(t)]}{(E[v^2(t)]+\tilde{\sigma}_y^{2*})}}.$$

Monte Carlo simulations of 100 independent runs have been performed by setting $k = 5$ in Step (1) of Algorithm 1 and different numbers of data samples ($N = 500, 1000, 2000$).

Tables I and II report the true values of parameters and variances, the means of their estimates and the associated standard deviations for both Procedure 1 and the IV approach described in Section 6.

The obtained results confirm the good performance of the proposed procedure. As expected, the accuracy of the IV estimator is quite poor, especially for short observation sequences.

8. CONCLUDING REMARKS

This paper has introduced the new family of ARARX + noise models and proposed a three–step identification procedure for these models. These models can offer a more realistic choice than simple equation error models in many applications like fault detection and filtering. ARARX + noise models can, in fact, 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 effectiveness of the proposed identification procedure has been verified by means of Monte Carlo simulations that show how it leads to accurate results also in presence of low signal to noise ratios.

REFERENCES

S. Beghelli, R. Guidorzi and U. Soverini. The Frisch scheme in dynamic system identification. *Automatica*, 26:171–176, 1990.

R. Diversi, R. Guidorzi and U. Soverini. Kalman filtering in extended noise environments. *IEEE Transactions on Automatic Control*, 50:1396–1402, September 2005.

R. Diversi, R. Guidorzi and U. Soverini. Identification of ARX models with noisy input and output. *Proc. of the 9th European Control Conference*, Kos, Greece, pp. 4073–4078, July 2007.

R. Guidorzi. *Multivariable System Identification: From Observations to Models*. Bononia University Press, Bologna, Italy, 2003.

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

	a_1	a_2	b_1	b_2	d_1
true	-0.5	0.06	1	-0.7	0.95
Proc. 1 ($N = 2000$)	-0.4963 ± 0.0289	0.0610 ± 0.0122	0.9976 ± 0.0180	-0.6951 ± 0.0379	0.9487 ± 0.0079
IV ($N = 2000$)	-0.4654 ± 0.0959	0.0606 ± 0.0279	0.9472 ± 0.6233	-0.4346 ± 0.7905	0.7278 ± 0.3793
Proc. 1 ($N = 1000$)	-0.4932 ± 0.0371	0.0583 ± 0.0173	0.9949 ± 0.0229	-0.6921 ± 0.0480	0.9476 ± 0.0105
IV ($N = 1000$)	-0.4749 ± 0.3700	0.0540 ± 0.0548	0.9014 ± 1.6454	-0.4586 ± 0.7795	0.5319 ± 0.4001
Proc. 1 ($N = 500$)	-0.4781 ± 0.0520	0.0634 ± 0.0267	0.9898 ± 0.0353	-0.6728 ± 0.0683	0.9427 ± 0.0182
IV ($N = 500$)	-0.4586 ± 0.2917	0.0644 ± 0.1254	0.9776 ± 1.6914	-0.5669 ± 1.2672	0.4093 ± 0.4193

Table 2. True and estimated values of the variances of $\tilde{u}(t)$, $\tilde{y}(t)$ and $e(t)$. For each value of N a Monte Carlo simulation of 100 runs has been performed.

	$\tilde{\sigma}_u^{2*}$	$\tilde{\sigma}_y^{2*}$	σ_e^{2*}
true	0.06	0.02	0.1
Proc. 1 ($N = 2000$)	0.0583 ± 0.0156	0.0220 ± 0.0218	0.0985 ± 0.0159
Proc. 1 ($N = 1000$)	0.0540 ± 0.0196	0.0299 ± 0.0281	0.0924 ± 0.0244
Proc. 1 ($N = 500$)	0.0483 ± 0.0295	0.0352 ± 0.0332	0.0920 ± 0.0371

R. Guidorzi and M. Pierantoni. A new parametrization of Frisch scheme solutions. *Proc. of the 12th International Conference on Systems Science*, Wroclaw, Poland, pp. 114–120, 1995.

L. Ljung. *System Identification – Theory for the User.* Prentice Hall, Englewood Cliffs, NJ, 1999.

T. Söderström. Errors–in–Variables methods in system identification. *Automatica*, 43: 939–958, June 2007.

T. Söderström and P. Stoica. *System Identification.* Prentice Hall International, Hemel Hempstead, UK, 1989.

T. Söderström, P. Stoica and B. Friedlander. An indirect prediction error method for system identification. *Automatica*, 27:183-188, January 1991.

U. Soverini and T. Söderström. Identification methods of dynamic systems in presence of input noise. *Proc. of the 12th IFAC Symposium on System identification*, Santa Barbara, California, June 2000.

P. Stoica and T. Söderström. Common factor detection and estimation. *Automatica*, 33:985–989, 1997.

F. Tjärnström and L. Ljung. Variance properties of a two–step ARX estimation procedure. *European Journal of Control*, 9:422–430, 2003.