Identification of Multichannel AR Models with Additive Noise: a Frisch Scheme Approach

Roberto Diversi
Department of Electrical, Electronic and Information Engineering
University of Bologna
Viale del Risorgimento 2, 40136 Bologna, Italy
Email: roberto.diversi@unibo.it

***Abstract*—A new approach for estimating multichannel AR (M-AR) models from noisy observations is proposed. It relies on the so-called Frisch scheme, whose rationale consists in finding the solution of the identification problem within a locus of solutions compatible with the second order statistics of the noisy data. Once that the locus of solutions has been defined, it is necessary to introduce a suitable selection criterion in order to identify a single solution. The criterion proposed in the paper is based on the comparison of the theoretical statistical properties of the residual of the noisy M-AR model with those computed from the data. The results obtained by means of Monte Carlo simulations show that the proposed algorithm outperforms some existing methods.**

I. Introduction

The identification of autoregressive (AR) models finds many applications, for instance, in spectral estimation [1], [2], speech analysis [3]–[5], biomedical signal processing [6], [7], structural health monitoring [8], radar signal processing [9]–[11], model-of-signal based fault detection [12]. These applications refer to both scalar AR models and multichannel AR models.

As is well known, when the AR process is corrupted by additive noise classical identification methods like least squares and Yule-Walker equations lead to biased estimates [13], [14]. Since this is a very common situation, many approaches have been proposed for identifying AR models in the presence of additive noise. Among them, the high-order Yule-Walker equations [1], [15], the prediction error method [16], the bias-compensated least squares [14], [17], the errors-in-variables approach [5], [18], [19].

With reference to the identification of multichannel AR (M-AR) models in the presence of additive noise one of the first published papers is [20], that extends to the multivariate case the high-order Yule-Walker equations. A Newton-Raphson algorithm is proposed in [21], where a set of nonlinear and a set of linear equations are solved iteratively. In [22], the bias-compensated least squares approach of [14] is extended to M-AR models. The model parameters and the noise variances are estimated in an iterative manner. The errors-in-variables method described in [23] exploits, on the one hand, some properties of the Frisch scheme and, on the other hand, a set of low-order and high-order Yule-Walker equations. In particular, the Frisch scheme is used to find an estimate of the noise variances that can, in turn, be used "to compensate" the effect of noise in the Yule-Walker equations. In [24], the authors suggest using two mutually interactive Kalman fiters. The first filter computes an estimate of the noise-free AR signal starting from the estimate of the AR parameters while the second filter updates the AR parameters starting from the estimated signal. A variant of the algorithm introduced in [22] has been suggested in [25]. Recently, a steepest descent method has been described in [26]. It combines Yule-Walker equations and inverse filtering to iteratively estimate the M-AR parameters and the noise variances.

In this paper, the M-AR model is considered as an errors-in-variables models and it is shown how the so-called Frisch scheme [27] can be fully exploited to identify both the AR parameters and the driving noise and additive noise covariance matrices. The rationale behind the Frisch scheme consists in finding the solution of the identification problem within a locus of solutions compatible with the second order statistics of the noisy data. In the specific case, it is first shown that the m channels of the multivariable AR models can be associated with m locus of solutions described by hypersurfaces belonging to the positive orthant of $\mathcal{R}^{m+1}$. Then, it is shown how to define a locus of solutions for the whole M-AR model by associating models with directions in $\mathcal{R}^{2m}$. In particular, a direction ρ belonging to the first orthant of $\mathcal{R}^{2m}$ leads to a set of M-AR parameters and noise variances. To estimate a single model among the set of possible solutions, it is necessary to define a suitable selection criterion. The criterion proposed in the paper is based on the comparison of the theoretical statistical properties of the residual of the noisy M-AR model with those computed from the data and the estimated model. It is worth highlighting that the proposed estimation method is quite different with respect to that described in [23], although both are related to the Frisch scheme. Indeed, the Frisch scheme properties are exploited in [23] only to identify the noise variances while the AR coefficients are estimated by using the Yule-Walker equations.

The remainder of the paper is organized as follows. Section II defines the identification problem. Section III describes some asymptotic properties of the M-AR process while Section IV shows how to exploit the Frisch scheme for M-AR models. The performance of the method is evaluated by means of Monte Carlo simulations and compared with those of [23] and [26]. The obtained results are shown in Section V.

ISBN 978-90-827970-1-5 © EURASIP 2018

II. Problem statement

Consider the m-dimensional multichannel AR process described by the equation

$$x(t) + A_1\, x(t-1) + \cdots + A_p\, x(t-p) = e(t) \tag{1}$$

where the AR signal $x(t)$ and the driving noise $e(t)$ are the vectors

$$x(t) = [\, x_1(t)\, x_2(t)\, \ldots\, x_m(t)\,]^T \tag{2}$$

$$e(t) = [\, e_1(t)\, e_2(t)\, \ldots\, e_m(t)\,]^T, \tag{3}$$

and $A_1, A_2, \ldots, A_p$ are $m \times m$ parameter matrices. The output of the AR process is corrupted by additive noise, hence the available measurements are given by

$$y(t) = x(t) + w(t), \tag{4}$$

where

$$y(t) = [\, y_1(t)\, y_2(t)\, \ldots\, y_m(t)\,]^T \tag{5}$$

$$w(t) = [\, w_1(t)\, w_2(t)\, \ldots\, w_m(t)\,]^T. \tag{6}$$

We introduce the following assumptions.

A1. The roots of the determinant of $A(z^{-1})$ lie inside the unit circle where

$$A(z^{-1}) = I_m + A_1\, z^{-1} + \cdots + A_p\, z^{-p}. \tag{7}$$

z^{-1} is the backward shift operator ($z^{-1}\, x(t) = x(t-1)$) whereas I_m denotes the $m \times m$ identity matrix.

A2. The driving noise $e(t)$ and the additive noise $w(t)$ are mutually uncorrelated zero–mean white processes with unknown diagonal covariance matrices Σ_e^* and Σ_w^*

$$E[\, e(t) e^T(t-\tau)\,] = \Sigma_e^*\, \delta(\tau) \tag{8}$$

$$E[\, w(t) w^T(t-\tau)\,] = \Sigma_w^*\, \delta(\tau) \tag{9}$$

$$E[\, e(t) w^T(t-\tau)\,] = 0, \forall \tau \tag{10}$$

where $\delta(\tau)$ denotes the Kronecker delta function and

$$\Sigma_e^* = \mathrm{diag}\,[\, \sigma_{e1}^{2*}, \sigma_{e2}^{2*}, \ldots, \sigma_{em}^{2*}\,], \tag{11}$$

$$\Sigma_w^* = \mathrm{diag}\,[\, \sigma_{w1}^{2*}, \sigma_{w2}^{2*}, \ldots, \sigma_{wm}^{2*}\,]. \tag{12}$$

A3. The order p is assumed as known.

The problem to be solved can be stated as follows.

Problem 1. Let $y(1), y(2), \ldots, y(N)$ be a set of noisy measurements generated by model (1), (4) under Assumptions A1-A3. Determine an estimate of the parameter matrices $A_1, \ldots, A_p$ and of the covariance matrices Σ_e^*, Σ_w^*.

III. Asymptotic properties of the M-AR process

The multichannel AR process (1) can also be represented by the set of m equations

$$x_i(t) + \sum_{j=1}^{m} \sum_{k=1}^{p} a_{ijk}\, x_j(t-k) = e_i(t), \quad i = 1, \ldots, m. \tag{13}$$

By comparing (1) to (13) it follows that

$$A_i = \begin{bmatrix} a_{11i} & a_{12i} & \cdots & a_{1mi} \\ a_{21i} & a_{22i} & \cdots & a_{2mi} \\ \vdots & \vdots & & \vdots \\ a_{m1i} & a_{m2i} & \cdots & a_{mmi} \end{bmatrix}, \quad i = 1, \ldots, p. \tag{14}$$

Model (1) can thus be seen as a set of m interconnected multi-input single-output subsystems. In fact, the output $x_i(t)$ of the i-th subsystem depends not only on its past p samples $x_i(t-1), \ldots, x_i(t-p)$ but also on the past p samples of $x_1(t), \ldots, x_{i-1}(t), x_{i+1}(t), \ldots, x_m(t)$, that play the role of inputs.

Let us consider the generic i-th subsystem of the M-AR model, described by the i-th equations in (13) and (4):

$$x_i(t) + \sum_{j=1}^{m} \sum_{k=1}^{p} a_{ijk}\, x_j(t-k) = e_i(t) \tag{15}$$

$$y_i(t) = x_i(t) + w_i(t). \tag{16}$$

By defining the vectors

$$\begin{aligned} \varphi_i^x(t) = [\, & x_1(t-1)\, \ldots\, x_1(t-p)\, \ldots\, x_i(t)\, x_i(t-1) \ldots \\ & \ldots x_i(t-p) \ldots x_m(t-1) \ldots x_m(t-p)\,]^T \end{aligned} \tag{17}$$

$$\varphi_i^e(t) = [\, \underbrace{0\, \ldots\, 0}_{(i-1)p}\, e_i(t)\ \ \underbrace{0\, \ldots\, 0}_{(m-i+1)p}\ ]^T \tag{18}$$

and the parameter vector

$$\begin{aligned} \theta_i^* = \big[\, & a_{i11}\, \ldots\, a_{i1p}\, a_{i21}\, \ldots\, a_{i2p}\, \cdots \\ & 1\, a_{ii1}\, \cdots\, a_{iip}\, \cdots\, a_{im1}\, \cdots\, a_{imp}\, \big]^T, \end{aligned} \tag{19}$$

equation (15) can be rewritten as

$$\left(\varphi_i^x(t) - \varphi_i^e(t)\right)^T\, \theta_i^* = 0. \tag{20}$$

By introducing also the vectors

$$\begin{aligned} \varphi_i^y(t) = [\, & y_1(t-1)\, \ldots\, y_1(t-p)\, \ldots\, y_i(t)\, y_i(t-1) \ldots \\ & \ldots y_i(t-p) \ldots y_m(t-1) \ldots y_m(t-p)\,]^T \end{aligned} \tag{21}$$

$$\begin{aligned} \varphi_i^w(t) = [\, & w_1(t-1)\, \ldots\, w_1(t-p)\, \ldots\, w_i(t)\, w_i(t-1) \ldots \\ & \ldots w_i(t-p) \ldots w_m(t-1) \ldots w_m(t-p)\,]^T, \end{aligned} \tag{22}$$

and using (16), it follows easily that

$$\varphi_i^y(t) = \varphi_i^x(t) + \varphi_i^w(t). \tag{23}$$

Consider now the covariance matrices

$$R_i^x = E\,[\varphi_i^x(t)\, \varphi_i^{xT}(t)] \tag{24}$$

$$R_i^y = E\,[\varphi_i^y(t)\, \varphi_i^{yT}(t)] \tag{25}$$

$$\bar{R}_i = E\left[\left(\varphi_i^x(t) - \varphi_i^e(t)\right)\left(\varphi_i^x(t) - \varphi_i^e(t)\right)^T\right], \tag{26}$$

where $E[\cdot]$ denotes the expectation operator. From (23) and Assumption A2 we get

$$\begin{aligned} R_i^y &= R_i^x + E\,[\varphi_i^w(t)\, \varphi_i^{wT}(t)] \\ &= R_i^x + \mathrm{diag}\left[\, \sigma_{w1}^{2*} I_p,\, \ldots, \sigma_{wi}^{2*} I_{p+1},\, \ldots,\, \sigma_{wm}^{2*} I_p\right]. \end{aligned} \tag{27}$$

Since $E[x_i(t)\, e_i(t)] = \sigma_{ei}^{2*}$ we also have

$$\bar{R}_i = R_i^x - \mathrm{diag}\ [\, \underbrace{0\, \ldots\, 0}_{(i-1)p} \sigma_e^{2*}\ \ \underbrace{0\, \ldots\, 0}_{(m-i+1)p}\ ], \tag{28}$$

therefore

$$R_i^y = \bar{R}_i + \tilde{R}_i^*, \tag{29}$$

ISBN 978-90-827970-1-5 © EURASIP 2018

where $\tilde{R}_i^*$ is the diagonal matrix

$$\tilde{R}_i^* = \text{diag}\left[\sigma_{w1}^{2*} I_p, \ldots, \sigma_{w(i-1)}^{2*} I_p, \sigma_{ei}^{2*} + \sigma_{wi}^{2*}, \sigma_{wi}^{2*} I_p, \sigma_{w(i+1)}^{2*} I_p \ldots, \sigma_{wm}^{2*} I_p\right]. \quad (30)$$

From (20) and (26) it follows that

$$\bar{R}_i\, \theta_i^* = 0, \quad (31)$$

hence, because of (29)

$$\left(R_i^y - \tilde{R}_i^*\right) \theta_i^* = 0. \quad (32)$$

The autocorrelation matrix of the noisy output R_i^y can thus be decomposed into a sum of a positive semidefinite matrix $\bar{R}_i$ and a diagonal matrix $\tilde{R}_i^*$. Note that if $\tilde{R}_i^*$ were known the true parameter vector θ_i^* could be determined from any basis of the null space of $\bar{R}_i$ by normalizing the appropriate entry to 1, see (19). However, as the driving noise variance σ_{ei}^{2*} and the additive noise variances $\sigma_{w1}^{2*}, \sigma_{w2}^{2*}, \ldots, \sigma_{wm}^{2*}$ are unknown, Eq. (32) cannot be directly used to find an estimate of the parameter vector θ_i^*.

The idea behind the Frisch scheme consists in finding the true noise variances $\sigma_{ei}^{2*}, \sigma_{w1}^{2*}, \sigma_{w2}^{2*}, \ldots, \sigma_{wm}^{2*}$ within a locus of solutions leading to a decomposition of the autocorrelation matrix R_i^y like the one in (29). This locus will be described in the next section.

IV. The Frisch scheme for M-AR models

Consider now the problem of determining the set of points $P = (\sigma_{ei}^2, \sigma_{w1}^2, \sigma_{w2}^2, \ldots, \sigma_{wm}^2)$ belonging to the positive orthant of $\mathcal{R}^{m+1}$ that lead to diagonal matrices

$$\tilde{R}_i(P) = \text{diag}\left[\sigma_{w1}^2 I_p, \ldots, \sigma_{w(i-1)}^2 I_p, \sigma_{ei}^2 + \sigma_{wi}^2, \sigma_{wi}^2 I_p, \sigma_{w(i+1)}^2 I_p \ldots, \sigma_{wm}^2 I_p\right] \quad (33)$$

such that

$$R_i^y - \tilde{R}_i(P) \geq 0, \quad \det(R_i^y - \tilde{R}_i(P)) = 0. \quad (34)$$

Every point P of this set can thus be associated to a parameter vector $\theta_i(P)$ that can be obtained by normalizing to 1 the appropriate entry of any basis of the null space of $R_i^y - \tilde{R}_i(P)$:

$$\left(R_i^y - \tilde{R}_i(P)\right) \theta_i(P) = 0 \quad (35)$$

where

$$\theta_i(P) = \left[a_{i11}(P) \ldots a_{i1p}(P)\, a_{i21}(P) \ldots a_{i2p}(P) \cdots 1\, a_{ii1}(P) \cdots a_{iip}(P) \cdots a_{im1}(P) \cdots a_{imp}(P)\right]^T. \quad (36)$$

It is possible to prove the following theorem.

Theorem 1: The set of points $P = (\sigma_{ei}^2, \sigma_{w1}^2, \sigma_{w2}^2, \ldots, \sigma_{wm}^2)$ satisfying (34) defines an hypersurface $\mathcal{S}(R_i^y)$ belonging to the positive orthant of $\mathcal{R}^{m+1}$ (the noise space). Every point P of $\mathcal{S}(R_i^y)$ defines a set of $m+1$ noise variances (the driving noise variance σ_{ei}^2 and the additive noise variances $\sigma_{w1}^2, \ldots, \sigma_{wm}^2$) and can be associated with a vector of parameters $\theta_i(P)$ (36) satisfying relation (35). The point $P_i^* = (\sigma_{ei}^{2*}, \sigma_{w1}^{2*}, \ldots \sigma_{wm}^{2*})$ defined by the true noise variances belongs to $\mathcal{S}(R_i^y)$ and is associated with the true parameter vector (19), i.e. $\theta_i(P_i^*) = \theta_i^*$.

The proof of Theorem 1 is omitted here due to space limitations. It could be proved by following similar reasoning as used for the proof of Theorem 1 in [28], although the problem under investigation is different.

A useful way to parameterize the hypersurface $\mathcal{S}(R_i^y)$ is based on the following result, whose proof is similar to that of Theorem 3 in [28].

Theorem 2: Let $\xi = (\eta, \xi_1, \xi_2, \ldots, \xi_m)$ be a generic point of the positive orthant of $\mathcal{R}^{m+1}$ and ρ the straight line from the origin through ξ. Its intersection with $\mathcal{S}(R_i^y)$ is the point P_i given by

$$P_i = \frac{\xi}{\lambda_{Mi}}, \quad \lambda_{Mi} = \max \text{ eig}\left((R_i^y)^{-1}\, \tilde{R}_i^\xi\right) \quad (37)$$

where

$$\tilde{R}_i^\xi = \text{diag}\left[\xi_1 I_p, \ldots, \xi_{i-1} I_p, \eta + \xi_i, \xi_i I_p, \xi_{i+1} I_p, \ldots, \xi_m I_p\right]. \quad (38)$$

Section III and Theorem 1 refer to the generic i–th subsystem, so that the whole M-AR model is described by m hypersurfaces $\mathcal{S}(R_1^y), \mathcal{S}(R_2^y), \ldots, \mathcal{S}(R_m^y)$, associated with the autocorrelation matrices $R_1^y, R_2^y, \ldots, R_m^y$ of the noisy outputs $y_1(t), y_2(t), \ldots, y_m(t)$. It is worth noting that the points $P_1^*, P_2^*, \ldots, P_m^*$ have the common entries $\sigma_{w1}^{2*}, \sigma_{w2}^{2*}, \ldots, \sigma_{wm}^{2*}$ since $P_i^* = (\sigma_{ei}^{2*}, \sigma_{w1}^{2*}, \ldots \sigma_{wm}^{2*})$ for $i = 1, 2, \ldots, m$. A parameterization of the whole M-AR model can be determined by exploiting Theorem 2, as it will be shown in the following.

Consider a straight line ρ passing through the origin and belonging to the positive orthant of $\mathcal{R}^{2m}$. This line can be represented by a point $P_\rho = (\eta_1, \ldots, \eta_m, \xi_1, \ldots, \xi_m)$ whose entries are the coordinates of the unit vector in the direction of ρ. Starting from ρ it is possible to define the m directions $\rho_1, \rho_2, \ldots, \rho_m$ in $\mathcal{R}^{m+1}$ which are represented by the points $P_{\rho 1}, P_{\rho 2}, \ldots, P_{\rho m}$ where $P_{\rho i} = (\eta_i, \xi_1, \ldots, \xi_m)$. The direction ρ_i is associated to the i–th subsystem. It is then possible to consider the intersections of $\rho_1, \rho_2, \ldots, \rho_m$ with the hypersurfaces $\mathcal{S}(R_1^y), \mathcal{S}(R_2^y), \ldots, \mathcal{S}(R_m^y)$, by using Theorem 2. Let these intersections be given by the points $P_1(\rho), P_2(\rho), \ldots, P_m(\rho)$ where

$$P_i(\rho) = (\sigma_{ei}^2, \sigma_{w1i}^2, \sigma_{w2i}^2, \ldots, \sigma_{wmi}^2), \quad i = 1, \ldots, m. \quad (39)$$

From Theorem 1, the point $P_i(\rho)$ can be associated with a parameter vector $\theta_i(P_i(\rho))$. According to these considerations, it is possible to associate a direction ρ in $\mathcal{R}^{2m}$ with the M-AR model

$$\mathcal{M}(\rho) \triangleq \{\theta_1(P_1(\rho)), \theta_2(P_2(\rho)), \ldots, \theta_m(P_m(\rho))\}. \quad (40)$$

It is worth noting that the point $P^* = (\sigma_{e1}^{2*}, \ldots, \sigma_{em}^{2*}, \sigma_{w1}^{2*}, \ldots \sigma_{wm}^{2*})$ whose coordinates are the true driving noise and additive noise variances defines the "true" direction ρ^* and then the true directions $\rho_1^*, \rho_2^*, \ldots, \rho_m^*$. It follows from Theorem 1 that $P_i(\rho_i^*) = P_i^*, i = 1, \ldots, m$. As a consequence, ρ^* is associated with the true M-AR model $\mathcal{M}(\rho^*) = \{\theta_1(P_1^*), \theta_2(P_2^*), \ldots, \theta_m(P_m^*)\} = \{\theta_1^*, \theta_2^*, \ldots, \theta_m^*\}$.

ISBN 978-90-827970-1-5 © EURASIP 2018

Remark 1: From the model $\mathcal{M}(\rho)$ it is easy to obtain a model like (1) by taking into account (14) so that (40) can be rewritten as

$$\mathcal{M}(\rho) = \{A_1(\rho), A_2(\rho), \ldots, A_p(\rho)\}. \tag{41}$$

Remark 2: Note that the generic direction ρ and then the model $\mathcal{M}(\rho)$ are associated with a set of driving noise variances $\sigma^2_{e1}, \sigma^2_{e2}, \ldots, \sigma^2_{em}$, see (39). On the contrary, the additive noise variances are not univocally defined as, from (39), we have m different values for each $\sigma^2_{wj}, j = 1, \ldots, m$. To define a unique set of additive noise variances we will consider the arithmetic means of the values given by $P_1(\rho), P_2(\rho), \ldots, P_m(\rho)$:

$$\bar{\sigma}^2_{wi} = \frac{1}{m} \sum_{j=1}^{m} \sigma_{wij}, \qquad i = 1, \ldots, m. \tag{42}$$

In this context, the identification problem (Problem1) consists in finding the true direction ρ^* in order to retrieve the true model $\mathcal{M}(\rho^*)$. The search for ρ^* in the positive orthant of $\mathcal{R}^{2m}$ can be performed by means of the selection criterion described in the next subsection, that is based on the statistical properties of the residual of the noisy M-AR model.

A. A selection criterion

The multivariable AR model (1) can be rewritten in the polynomial form

$$A(z^{-1})\, x(t) = e(t) \tag{43}$$

where $A(z^{-1})$ is the matrix polynomial (7). By inserting (4) in (43) we get

$$A(z^{-1})\, y(t) = A(z^{-1})\, w(t) + e(t). \tag{44}$$

Define the m-dimensional stochastic process

$$\varepsilon(t) = A(z^{-1})\, y(t) = A(z^{-1})\, w(t) + e(t), \tag{45}$$

that can be considered as the residual of the noisy AR model (1), (4). Because of Assumption A2, $\varepsilon(t)$ is the sum of a moving average process and a white noise so that it is characterized by a finite number of autocorrelations $R_\varepsilon(\tau) = E\left[\varepsilon(t)\, \varepsilon^T(t-\tau)\right]$:

$$R_\varepsilon(0) = \sum_{i=0}^{p} A_i\, \Sigma^*_w\, A_i^T + \Sigma^*_e \tag{46}$$

$$R_\varepsilon(\tau) = \sum_{i=0}^{p-\tau} A_{i+\tau}\, \Sigma^*_w\, A_i^T, \qquad \tau = 1, \ldots, p \tag{47}$$

$$R_\varepsilon(\tau) = 0, \qquad \tau > p, \tag{48}$$

where $A_0 = I_m$. The properties of the autocorrelation function (46)–(48) can be exploited to define a suitable selection criterion for identifying the true model $\mathcal{M}(\rho^*)$.

To this end, let us consider a generic direction ρ in $\mathcal{R}^{2m}$, the associated model $\mathcal{M}(\rho)$ (41) and the associated driving noise and additive noise variances (see Remark 2). By using the parameters and the noise variances of (41) it is possible to compute the estimates $\hat{R}_\varepsilon(0,\rho), \hat{R}_\varepsilon(1,\rho), \ldots, \hat{R}_\varepsilon(p,\rho)$ as in (46)–(48). On the other hand, by using the available data $y(1), y(2), \ldots, y(N)$ and the parameters of (41) it is possible to compute first a sequence of residuals

$$\hat{\varepsilon}(t) = \hat{A}(z^{-1}, \rho)\, y(t), \tag{49}$$

where

$$A(z^{-1}, \rho) = I_m + A_1(\rho)\, z^{-1} + \cdots + A_p(\rho)\, z^{-p} \tag{50}$$

and then the sample estimates $\hat{R}^N_\varepsilon(0,\rho)$, $\hat{R}^N_\varepsilon(1,\rho)$, $\ldots$, $\hat{R}^N_\varepsilon(p,\rho)$ given by

$$\hat{R}^N_\varepsilon(\tau,\rho) = \frac{1}{N-\tau} \sum_{t=\tau+1}^{N} \hat{\varepsilon}(t)\, \hat{\varepsilon}^T(t-\tau), \quad \tau = 0, \ldots, p. \tag{51}$$

The proposed selection criterion consists in comparing the theoretical statistical properties of $\varepsilon(t)$ with those computed from the data. To this end, define the vectors

$$\hat{r}_\varepsilon(\rho) = \left[\, \mathrm{vec}\left(\hat{R}_\varepsilon(0,\rho)\right)^T \ \mathrm{vec}\left(\hat{R}_\varepsilon(1,\rho)\right)^T \ \cdots \ \cdots \ \mathrm{vec}\left(\hat{R}_\varepsilon(p,\rho)\right)^T \right]^T \tag{52}$$

$$\hat{r}^N_\varepsilon(\rho) = \left[\, \mathrm{vec}\left(\hat{R}^N_\varepsilon(0,\rho)\right)^T \ \mathrm{vec}\left(\hat{R}^N_\varepsilon(1,\rho)\right)^T \ \cdots \ \cdots \ \mathrm{vec}\left(\hat{R}^N_\varepsilon(p,\rho)\right)^T \right]^T \tag{53}$$

and the loss function

$$f(\rho) = \|\hat{r}_\varepsilon(\rho) - \hat{r}^N_\varepsilon(\rho)\|^2_2. \tag{54}$$

The estimation of ρ^* and of the associated model $\mathcal{M}(\rho^*)$ can thus be performed by minimizing $f(\rho)$ in the positive orthant of $\mathcal{R}^{2m}$:

$$\hat{\rho} = \arg\min_{\rho \in \mathcal{R}^{2m}} f(\rho). \tag{55}$$

V. Simulation results

The performance of the proposed Frisch scheme-based identification approach has been evaluated by means of Monte Carlo simulation and compared with those of the following approaches:

- the errors-in-variables (EIV) approach described in [23]. In particular, Algorithm 2 in [23] has been considered;
- the improved least-squares algorithm based on inverse filtering (IFILSM) introduced in [26].

It is worth noting that the above mentioned methods outperforms other existing methods like, for instance, [21], [22], [25], as shown in [23], [26].

Let the available data be generated by the two-channel AR model of order $p = 2$

$$x(t) + A_1\, x(t-1) + A_2\, x(t-2) = e(t) \tag{56}$$

$$y(t) = x(t) + w(t), \tag{57}$$

where

$$A_1 = \begin{bmatrix} -0.71 & 0.32 \\ -0.88 & -0.24 \end{bmatrix} \qquad A_2 = \begin{bmatrix} 0.57 & -0.15 \\ -0.49 & -0.30 \end{bmatrix}. \tag{58}$$

ISBN 978-90-827970-1-5 © EURASIP 2018

TABLE I
TRUE AND ESTIMATED VALUES OF THE COEFFICIENTS OF A_1, A_2 AND OF THE NOISE VARIANCES $\sigma_{e1}^{2*}, \sigma_{e2}^{2*}, \sigma_{w1}^{2*}, \sigma_{w2}^{2*}$.

	a_{111}	a_{112}	a_{211}	a_{221}	σ_{e1}^{2*}	$\tilde{\sigma}_{e2}^{2*}$
true	-0.71	0.32	-0.88	-0.24	1	1
Frisch	-0.6985 ± 0.0548	0.3332 ± 0.1305	-0.8716 ± 0.0951	-0.2283 ± 0.1088	1.0343 ± 0.1622	1.0718 ± 0.5278
EIV [23]	-0.7057 ± 0.0363	0.3628 ± 0.1952	-0.8792 ± 0.0847	-0.2313 ± 0.1253	0.5086 ± 0.3441	0.9078 ± 0.4246
IFILSM [26]	-0.6894 ± 0.4802	0.4929 ± 1.6020	-0.8547 ± 0.6528	-0.6558 ± 2.8465	1.0219 ± 0.9118	1.0354 ± 9.3821
	a_{211}	a_{212}	a_{221}	a_{222}	$\tilde{\sigma}_{w1}^{2*}$	$\tilde{\sigma}_{w2}^{2*}$
true	0.57	-0.15	-0.49	-0.30	2.1	2.7
Frisch	0.5522 ± 0.1239	-0.1521 ± 0.1098	-0.5045 ± 0.1103	-0.3009 ± 0.1176	2.0774 ± 0.1123	2.6243 ± 0.4023
EIV [23]	0.5273 ± 0.1878	-0.1789 ± 0.1498	-0.5018 ± 0.1234	-0.3064 ± 0.1275	2.0858 ± 0.0898	2.6952 ± 0.3706
IFILSM [26]	0.4579 ± 2.1403	-0.2671 ± 0.9396	0.0445 ± 3.7279	0.0445 ± 2.2497	1.9925 ± 0.3666	1.6349 ± 3.7507

This model was also used in [21], [23], [24]. The driving processes $e_1(t), e_2(t)$ are white noise with unit variance so that $\Sigma_e^* = I_2$. The additive noise variances are set to $\sigma_{w1}^{2*} = 2.1$, $\sigma_{w1}^{2*} = 2.7$. This leads to a signal to noise ratio of 5 dB on both channels. A Monte Carlo simulation of 200 independent runs has been carried out by considering $N = 4000$ available samples in each run. The results are summarized in Table I that reports the true values of AR parameters, driving noise and measurement noise variances as well as their estimated means and standard deviations obtained in the simulation. We can note that the performance of the IFILSM method is quite bad with respect to those of Frisch and EIV. The best accuracy is achieved by the proposed Frisch scheme-based identification method for both the AR coefficients and the noise variances.

REFERENCES

[1] S. M. Kay, *Modern Spectral Estimation*, Prentice–Hall, Englewood Cliffs, New Jersey, 1988.

[2] P. Stoica and R. Moses, *Spectral Analysis of Signals*, Pearson Prentice–Hall, 2005.

[3] J. Lim and A. Oppenheim, “All-pole modeling of degraded speech,” *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 26, pp. 197-210, 1978.

[4] E. Grivel, M. Gabrea, and M. Najim, “Speech enhancement as a realization issue”, *Signal Processing*, vol. 82, pp. 1963-1978, 2002.

[5] W. Bobillet, R. Diversi, E. Grivel, R. Guidorzi, M. Najim and U. Soverini, “Speech enhancement combining optimal smoothing and errors-in-variables identification of noisy AR processes”, *IEEE Transactions on Signal Processing*, vol. 55, pp. 5564-5578, 2007.

[6] I. Güler, M. K. Kiymik, M. Akin, and A. Alkan, “AR spectral analysis of EEG signal by using maximum likelihood estimation,” *Computers in Biology and Medicine*, vol. 31, pp. 441-450, 2001.

[7] H. Hytti, R. Takalo, and H. Ihalainen, “Tutorial on multivariate autoregressive modelling”, *Journal of Clinical Monitoring and Computing*, vol. 20, pp. 101-108, 2006.

[8] R. Guidorzi, R. Diversi, L. Vincenzi, C. Mazzotti and V. Simioli, “Structural monitoring of a tower by means of MEMS-based sensing and enhanced autoregressive models,” *European Journal of Control*, vol. 20, pp. 4-13, 2014.

[9] S. Haykin and A. Steinhardt (Eds.), *Adaptive Radar Detection and Estimation*, Wiley–Interscience, 1992.

[10] Y. I. Abramovich, N. K. Spencer and M. D. E. Turley, “Time-varying autoregressive (TVAR) models for multiple radar observations,” *IEEE Transactions on Signal Processing*, vol. 55, pp. 1298–1311, 2007.

[11] T. Rouffet, E. Grivel, P. Vallet, C. Enderli and S. Kemkemian, “Combining two phase codes to extend the radar unambiguous range and get a trade-off in terms of performance for any clutter,” *Proc. of ICASSP 2015*, Brisbane, Australia, April 2015, pp. 1822–1826.

[12] R. Isermann, *Fault Diagnosis Systems – An Introduction from Fault Detection to Fault Tolerance*, Springer, 2006.

[13] S. M. Kay, “The effects of noise on the autoregressive spectral estimator,” *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 27, pp. 478–485, 1979.

[14] W. X. Zheng, “A least–squares based method for autoregressive signals in the presence of noise”, *IEEE Transactions on Circuits and Systems II*, vol. 46, pp. 81–85, 1999.

[15] Y. T. Chan and R. P. Langford, “Spectral estimation via the high–order Yule–Walker equations,” *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 30, pp. 689–698, 1982.

[16] A. Nehorai and P. Stoica, “Adaptive algorithms for constrained ARMA signals in the presence of noise,” *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 36, pp. 1282–1291, 1988.

[17] A. Mahmoudi and M. Karimi, “Inverse filtering based method for estimation of noisy autoregressive signals”, *Signal Processing*, vol. 91, pp. 1659–1664, 2011.

[18] R. Diversi, U. Soverini, and R. Guidorzi, “A new estimation approach for AR models in presence of noise”, *IFAC Proceedings Volumes*, vol. 38, pp. 160–165, 2005.

[19] R. Diversi, R. Guidorzi and U. Soverini, ”Identification of autoregressive models in the presence of additive noise”, *International Journal of Adaptive Control and Signal Processing*, vol. 22, pp. 465-481, 2008.

[20] T. Sen Lee, “Large sample identification and spectral estimation of noisy multivariate autoregressive processes”, *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 31, pp. 76–82, 1983.

[21] M. K. Hasan, M. J. Hossain, and M. A. Haque, “Parameter estimation of multichannel autoregressive processes in noise”, *Signal Processing*, vol. 83, pp. 603–610, 2003.

[22] A. Mahmoudi, and M. Karimi, “Estimation of parameters of multichannel autoregressive signals from noisy observations”, *Signal Processing*, vol. 88, pp. 2777–2783, 2008.

[23] J. Petitjean, E. Grivel, W. Bobillet, and P. Roussilhe, “Multichannel AR parameter estimation from noisy observations as an errors-in-variables issue”, *Signal, Image and Video Processing*, vol. 4, pp. 209–220, 2010.

[24] A. Jamoos, E. Grivel, N. Shakarneh, and H. Abdel-Nour, “Dual optimal filters for parameter estimation of a multivariate autoregressive process from noisy observations”, *IET Signal Processing*, vol. 5, pp. 471–479, 2011.

[25] X. Qu, J. Zhou, and Y. Luo, “A new noise-compensated estimation scheme for multichannel autoregressive signals from noisy observations”, *Journal of Supercomputing*, vol. 58, pp. 34–49, 2011.

[26] A. Mahmoudi, “Adaptive algorithm for multichannel autoregressive estimation in spatially correlated noise”, *Journal of Stochastics*, vol. 2014, Article ID 502406, 2014.

[27] R. Guidorzi, R. Diversi, and U. Soverini, “The Frisch scheme in algebraic and dynamic identification problems”, *Kybernetika*, vol. 44, pp. 585–616, 2008.

[28] R. Diversi, and R. Guidorzi, “The Frisch scheme in multivariable errors-in-variables identification”, *European Journal of Control*, vol. 37, pp. 43–53, 2017.