PREDICTION ERROR METHOD TO ESTIMATE THE AR PARAMETERS WHEN THE AR PROCESS IS DISTURBED BY A COLORED NOISE

Roberto Diversi[⋆†] *Hiroshi Ijima*[⋆] *Eric Grivel*[†]

[⋆†] DEI – University of Bologna, Viale del Risorgimento 2, 40136 Bologna, Italy
E-mail: roberto.diversi@unibo.it
[⋆] Faculty of Education, Wakayama University, 640-8510 Wakayama, Japan
E-mail: ijima@center.wakayama-u.ac.jp
[†] Université de Bordeaux, IPB, UMR CNRS 5218 IMS, 33400 Talence, France
E-mail: eric.grivel@ims-bordeaux.fr

ABSTRACT

Estimating the autoregressive parameters from noisy observations has been addressed by various authors for the last decades. Although several on-line or off-line approaches have been proposed when the additive noise is white, few papers deal with the additive moving average noise. In this paper, we suggest estimating the model parameters by using the prediction error method. Despite its high computational cost, the method has the advantage of being efficient in the Gaussian case. A comparative study with existing methods is then carried out and points out the efficiency of our approach especially when the number of samples is small.

Index Terms— Autoregressive processes, prediction error method, moving average noise, estimation.

1. INTRODUCTION

A great deal of interest has been paid to autoregressive (AR) models. Indeed, this model is one of the simplest and is defined by a finite set of parameters, namely the AR parameters. It has been used in a wide range of signal processing issues, from spectral analysis to signal analysis [1]. It is very popular in various applications such as speech processing (for coding, enhancement, etc.), radar processing (for clutter modeling, etc.), digital communications (for channel estimation, etc.), statistics, econometrics and forecasting. Its variants such as the multichannel AR process (M-AR) or the time-varying AR process (TVAR) are also often used respectively when there are several sensors and when a non-stationary signal must be processed.

Many ways to estimate the model parameters have been proposed and the most known are the least squares methods. They can be either off-line and based on the Yule–Walker (YW) equations or on-line and based on adaptive filters such as the LMS. However, when the observations are disturbed by an additive noise, the AR parameter estimates are biased. More particularly, when the noise is white, several authors have focused their attention on the way to take into account the influence of the noise and get rid of it. If the additive noise variance is known, the noise-compensated approach can be used. It consists in subtracting the noise-variance to the diagonal of the noisy-observation correlation matrix in the YW equations [2]. Otherwise, instrumental variable techniques, among which the most known method is the modified Yule–Walker equations, can be considered [1, 3]. As an alternative, an "off-line" bias correction scheme and some variants have been proposed by Zheng [4, 5]. Davila [6] suggests mapping this estimation issue into a quadratic eigenvalue problem. The errors-in-variables (EIV) approach [7, 8] consists in jointly estimating the AR parameters, the variance of the additive noise and the variance of the driving process. Concerning "on-line" methods, the ρ-LMS [9] or the γ-LMS [10] can be used. Otherwise, providing that the noise variances are known, the AR process and their parameters can be jointly estimated. This non-linear estimation issue can be addressed by using Kalman (or H_∞) algorithm such as the Extended Kalman (or EH_∞) filter [11], the Second-Order EKF (SOE H_∞) [12], the Sigma-Point H_∞ filter [13], etc. or by using coupled filters.

For the last years, extensions of most of the above methods have been proposed to estimate the M-AR matrices of M-AR processes disturbed by additive white noises. For instance, YW equations are considered in [14] whereas EIV is studied in [15]. In [16], an extension of Zheng's method to the multichannel case has been proposed. In [17], the authors suggest solving two set of equations that must be satisfied by the coefficients of the AR matrices and the noise variances. The first one is the noise-compensated Yule-Walker equations. In the second one, the noise variances are expressed from the coefficients of the AR matrices and the autocorrelation of the observations filtered by the inverse filter. The noise variances hence satisfy a set of non-linear equations that can be solved by means of a Newton-Raphson algorithm. Their estimations are then used in the noise-compensated Yule-Walker equations to deduce the AR matrix parameters. Similarly, the estimations of the TVAR parameters have been studied in [18].

In this paper, we focus our attention on the estimation of the AR parameters when the AR process is disturbed by a colored noise. Unlike the additive white noise case, few papers deal with this issue. To our knowledge, this problem has been investigated by looking at two cases: when the additive noise is a moving average (MA) process [19, 20] or when it is a first-order AR process [21]. Note that the first resulting algorithms are based on [4] while the second algorithm is based on [6]. Our contribution is twofold:

1) We propose a new method to estimate the AR parameters when the AR process is disturbed by a MA process. It is based on the prediction error method (PEM), which is known to be asymptotically unbiased and efficient in the Gaussian case [22, 23].
2) We compare our approach with [20].

The remainder of this paper is organized as follows. Section 2 defines the identification problem. Section 3 recalls the main concepts of the prediction error method and describes how the PEM can be applied to the case of AR models corrupted by MA processes. In Section 4, a comparative study between our approach and that introduced in [20] is carried out by means of Monte Carlo simulations.

2. PROBLEM STATEMENT

Let us assume that the signal is modeled by an autoregressive process as follows:

$$s(k) + a_1\, s(k-1) + \cdots + a_p\, s(k-p) = u(k) \tag{1}$$

where $\{a_i\}_{i=1,\dots,p}$ are the AR parameters and $u(k)$ is a zero-mean white noise process with variance σ_u^2. The noise–free AR signal $s(k)$ is affected by the additive noise $b(k)$ so that the available observation is given by

$$y(k) = s(k) + b(k). \tag{2}$$

We assume that $b(k)$ is a q–th order MA process defined as follows:

$$b(k) = w(k) + c_1\, w(k-1) + \cdots + c_q\, w(k-q) \tag{3}$$

where $\{c_i\}_{i=1,\dots,q}$ are the MA parameters and $w(k)$ is a zero-mean white process with variance σ_w^2, uncorrelated with $u(k)$. By introducing the polynomials

$$A(z^{-1}) = 1 + a_1\, z^{-1} + \cdots + a_p\, z^{-p} \tag{4}$$

$$C(z^{-1}) = 1 + c_1\, z^{-1} + \cdots + c_q\, z^{-q} \tag{5}$$

where z^{-1} is the unit delay operator, relations (1) and (3) can be rewritten as

$$A(z^{-1})\, s(k) = u(k) \tag{6}$$

$$b(k) = C(z^{-1})\, w(k). \tag{7}$$

Inserting (6) and (7) in (2) leads to

$$y(k) = \frac{u(k)}{A(z^{-1})} + C(z^{-1})\, w(k), \tag{8}$$

which is an alternative representations of the noise-corrupted AR process.

The problem to be solved consists in estimating the AR parameters $\{a_i\}_{i=1,\dots,p}$ and the driving noise variance σ_u^2 given the set of noisy observations $\{y(k)\}_{k=1,\dots,N}$. The orders p and q are assumed as *a priori* known.

3. A PREDICTION ERROR METHOD FOR ESTIMATING AR MODELS DISTURBED BY MA PROCESSES

Consider a single-output linear system described by the following state space representation

$$x(k+1) = A(\vartheta)\, x(k) + G(\vartheta)\, e_1(k) \tag{9}$$

$$y(k) = C(\vartheta)\, x(k) + e_2(k) \tag{10}$$

where $e_1(k)$ and $e_2(k)$ are mutually uncorrelated white processes with covariances

$$E\left[e_1(k)\, e_1^T(k)\right] \triangleq Q(\vartheta) \tag{11}$$

$$E\left[e_2^2(k)\right] \triangleq r(\vartheta) \tag{12}$$

and ϑ is a vector of unknown parameters that completely characterize the model. By considering the Kalman predictor associated with model (9)–(10), it is possible to obtain the so-called innovation form [22, 23]:

$$\hat{x}(k+1|k,\vartheta) = A(\vartheta)\, \hat{x}(k|k-1,\vartheta) + K(\vartheta)\, \varepsilon(k,\vartheta) \tag{13}$$

$$y(k) = C(\vartheta)\, \hat{x}(k|k-1,\vartheta) + \varepsilon(k,\vartheta) \tag{14}$$

where (omitting the argument ϑ for simplicity)

$$K = A\, P\, C^T \left(C\, P\, C^T + r\right)^{-1}, \tag{15}$$

and P is the solution of the algebraic Riccati equation

$$P = A\, P\, A^T + G\, Q\, G^T - A\, P\, C^T \left(C\, P\, C^T + r\right)^{-1} C\, P\, A^T. \tag{16}$$

$\varepsilon(k,\vartheta)$ is the innovation, i.e. the one step-ahead prediction error of $y(k)$:

$$\varepsilon(k,\vartheta) = y(k) - \hat{y}(k|k-1,\vartheta) = y(k) - C(\vartheta)\, \hat{x}(k|k-1,\vartheta). \tag{17}$$

The prediction error method consists in estimating ϑ by minimizing the cost function

$$V_N(\vartheta) = \frac{1}{N} \sum_{k=1}^{N} \varepsilon^2(k,\vartheta), \tag{18}$$

where N is the number of available data [22]–[25]. The PEM parameter estimates are consistent and asymptotically Gaussian distributed under weak conditions. Moreover, if $e_1(k)$ and $e_2(k)$ are Gaussian processes, the obtained estimation is also asymptotically efficient [22, 23].

To obtain a state space representation of the model (8), let us define the following stochastic processes

$$\xi_j(k) = \sum_{i=1}^{q+1-j} c_{i+j-1}\, w(k-i), \quad j = 1, \dots, q. \tag{19}$$

From (2) and (3) it follows that

$$y(k) = s(k) + w(k) + \sum_{i=1}^{q} c_i\, w(k-i) = s(k) + w(k) + \xi_1(k). \tag{20}$$

By considering the state vector

$$x(k) = \left[\, s(k) \cdots s(k-p+1)\ \xi_1(k) \cdots \xi_q(k) \,\right]^T \tag{21}$$

it is possible to define a state space model of type (9)–(10) where

$$A(\vartheta) = \begin{bmatrix} -a_1 & \cdots & \cdots & -a_p & & & & \\ 1 & 0 & \cdots & 0 & & & & \\ 0 & \ddots & 0 & \vdots & & 0_{p\times q} & & \\ 0 & & 1 & 0 & & & & \\ & & & & 0 & 1 & & 0 \\ & & & & 0 & \ddots & \ddots & 0 \\ & 0_{p\times q} & & & \vdots & & \ddots & 1 \\ & & & & 0 & \cdots & \cdots & 0 \end{bmatrix} \tag{22}$$

$$G(\vartheta) = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 & \cdots & \cdots & 0 \\ 0 & \cdots & \cdots & 0 & c_1 & c_2 & \cdots & c_q \end{bmatrix}^T \tag{23}$$

$$C = \left[\, 1\ \underbrace{0 \cdots 0}_{p-1}\ 1\ \underbrace{0 \cdots 0}_{q-1} \,\right] \tag{24}$$

$$e_1(k) = \left[\, u(k)\ w(k-1) \,\right]^T \qquad e_2(k) = w(k) \tag{25}$$

$$Q(\vartheta) = \begin{bmatrix} \sigma_u^2 & 0 \\ 0 & \sigma_w^2 \end{bmatrix} \qquad r(\vartheta) = \sigma_w^2 \tag{26}$$

and the parameter vector characterizing the model is

$$\vartheta = \left[\, a_1\ a_2 \cdots a_p\ c_1 \cdots c_q\ \sigma_u^2\ \sigma_w^2 \,\right]^T. \tag{27}$$

Table 1. Example 1: true and estimated values of AR parameters and driving noise variance and NRMSE for PEM and YWILS. Monte Carlo simulation of $M = 300$ runs have been performed with $N = 250, 500, 1000$.

	a_1	a_2	a_3	σ_u^2	NRMSE
true	-1.9363	1.7233	-0.7050	1	–
PEM ($N = 1000$)	-2.0067 ± 0.0448	1.8147 ± 0.0662	-0.7485 ± 0.0338	0.8071 ± 0.5265	0.0561
YWILS ($N = 1000$)	-2.0015 ± 0.3767	1.8583 ± 0.9124	-0.8062 ± 0.6660	0.6267 ± 2.1220	0.4476
PEM ($N = 500$)	-2.0574 ± 0.0345	1.9114 ± 0.0536	-0.7961 ± 0.0304	0.7556 ± 0.5102	0.0937
YWILS ($N = 500$)	-2.0894 ± 0.3668	2.1497 ± 0.9995	-1.1463 ± 1.1951	-0.1504 ± 2.9206	0.6397
PEM ($N = 250$)	-2.0298 ± 0.3961	1.8518 ± 0.5297	-0.7325 ± 0.2418	0.9492 ± 0.5901	0.2685
YWILS ($N = 250$)	-2.0599 ± 0.5966	1.9574 ± 1.2001	-1.0853 ± 1.0796	-0.0374 ± 7.3903	0.6624

The identification problem under investigation can thus be solved by finding the parameter vector which minimizes (18)

$$\hat{\vartheta}_N = \arg\min_{\vartheta} V_N(\vartheta). \quad (28)$$

It should be noted that, by using the spectral factorization theorem [23], model (8) can be represented by means of the ARMA process

$$y(t) = \frac{F(z^{-1})}{A(z^{-1})}\varepsilon(t), \quad (29)$$

where the stable polynomial $F(z^{-1})$ of degree $p+q$ and the variance σ_ε^2 of the white process $\varepsilon(t)$ can be obtained from the relation

$$\sigma_\varepsilon^2\, F(z^{-1})\, F(z) = \sigma_u^2 + \sigma_w^2\, A(z^{-1})\, A(z)\, C(z^{-1})\, C(z). \quad (30)$$

It is thus possible to apply PEM by considering the standard ARMA model (29). Nevertheless, since the number of parameters to be identified is $2p + q + 1$ instead of $p + q + 2$, an approach of this kind presents the following drawbacks:

- the computational load increases;
- the estimation accuracy decreases (the coefficients of $A(z^{-1})$ and $F(z^{-1})$ are treated as independent parameters) [26].

Remark: given the state space representation of the system introduced above, an alternative would consist in explicitly estimating both the AR process, the MA parameters and the AR parameters from the noisy observations. In that case, one basic solution would consist in storing p consecutive samples of the AR process, the q variables introduced in equ. (19), the p AR parameters $\{a_i\}_{i=1,\dots,p}$ and the q MA parameters $\{c_j\}_{j=1,\dots,q}$. This would lead to a state space representation of the system where the extended state vector would be updated by using a non-linear function of both the model noise and the extended vector at the previous instant. Therefore, a general extended Kalman filtering would be necessary. Nevertheless, the noise variance would remain undefined as well as the model noise covariance matrix. In addition, the direct estimation of the MA parameters would be difficult. For this reason, we have focused our attention on an alternative identification approach based on the PEM. Another approach would consist in approximating the MA process by a high-order AR process. Once again, a non-linear Kalman based-approach could be considered, but the estimation of the driving process covariance matrix has to be addressed. Note that, many authors have focused their attention on the MA parameter estimation such as Stoica and Broersen [27]–[29].

4. SIMULATION RESULTS

In this section, the performance of the proposed PEM approach is tested by means of Monte Carlo simulations and compared with that of the method described in [20], that will be denoted as YWILS. It should be noted that the YWILS method relies on a iterative bias–compensated least squares algorithm where, at each step, an estimation of the autocovariances of the MA process is used to compensate the bias in the least squares estimation of the AR parameters. The simulations will be performed on data generated by two different noisy AR models already considered in [20].

In the first example, we consider a third–order AR model described by the coefficients

$$a_1 = -1.9363, \qquad a_2 = 1.7233, \qquad a_3 = -0.7050,$$

corresponding to the poles $p_{1,2} = 0.89\, e^{\pm j(0.3\pi)}$, $p_3 = 0.89$. The driving process $u(k)$ is a Gaussian white noise with variance $\sigma_u^2 = 1$. The additive noise $b(k)$ is a second–order MA process characterized by the coefficients

$$c_1 = -1, \qquad c_2 = 0.2$$

and $w(t)$ is a Gaussian white process with variance $\sigma_w^2 = 2.175$, corresponding to a SNR of about 5 dB. Monte Carlo simulations of $M = 300$ independent runs have been carried out by considering different numbers of data, namely $N = 250, 500, 1000$. The parameters m and δ characterizing the YWILS algorithm (see [20]) have been set to $m = 5$ and $\delta = 0.001$. The obtained results are summarized in Table 1, which reports the true and estimated values of AR coefficients and driving noise variance, the mean of their estimates, the associated standard deviation and the normalized root mean square error defined as

$$\text{NRMSE} = \frac{1}{\|\theta\|}\sqrt{\frac{1}{M}\sum_{i=1}^{M}\|\hat{\theta}^i - \theta\|^2}, \quad (31)$$

where $\hat{\theta}^i$ denotes the estimate of $\theta \triangleq [a_1\ a_2\ \cdots\ a_p]^T$ obtained in the i–th trial of the Monte Carlo simulation and M denotes the number of Monte Carlo runs. Figures 1 and 2 report the true and estimated poles for the case $N = 500$.

In the second example, we consider a fourth–order AR model described by the coefficients

$$a_1 = -2.4863,\ a_2 = 3.0909,\ a_3 = -1.9694,\ a_4 = 0.6274,$$

leading to the poles $p_{1,2} = 0.89\, e^{\pm j(0.3\pi)}$ and $p_{3,4} = 0.89\, e^{\pm j(0.2\pi)}$. The driving process $u(k)$ is a Gaussian white noise with variance $\sigma_u^2 = 1$. The additive noise $b(k)$ is the same MA process considered in the first example and $w(k)$ is a Gaussian white process with

Table 2. Example 2: true and estimated values of AR parameters and driving noise variance and NRMSE for PEM and YWILS. Monte Carlo simulation of $M = 300$ runs have been performed with $N = 250, 500, 1000$.

	a_1	a_2	a_3	a_4	σ_u^2	NRMSE
true	-2.4863	3.0909	-1.9694	0.6274	1	$-$
PEM ($N = 1000$)	-2.5122 ± 0.0321	3.1486 ± 0.0704	-2.0202 ± 0.0645	0.6508 ± 0.0238	1.2261 ± 0.3978	0.0298
YWILS ($N = 1000$)	-2.1860 ± 0.5386	2.5289 ± 0.9953	-1.5048 ± 0.8186	0.4934 ± 0.2555	4.2010 ± 8.9856	0.3639
PEM ($N = 500$)	-2.5650 ± 0.0270	3.2787 ± 0.0604	-2.1403 ± 0.0563	0.6950 ± 0.0218	1.1580 ± 0.3950	0.0645
YWILS ($N = 500$)	-2.4120 ± 0.6311	3.0416 ± 1.5004	-1.9769 ± 1.4739	0.6699 ± 0.6230	3.5105 ± 10.0348	0.5099
PEM ($N = 250$)	-2.6632 ± 0.1801	3.4617 ± 0.3506	-2.2807 ± 0.3041	0.7333 ± 0.0900	1.0009 ± 0.4433	0.1631
YWILS ($N = 250$)	-2.2820 ± 0.9300	2.8053 ± 2.0684	-1.7789 ± 1.9534	0.6133 ± 0.7719	7.4344 ± 13.0745	0.6957

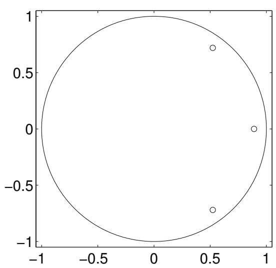


Fig. 1. True poles: Example 1, $N = 500$.

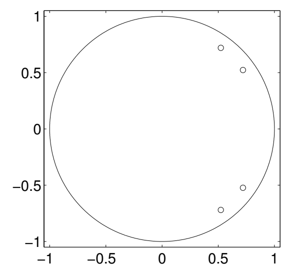
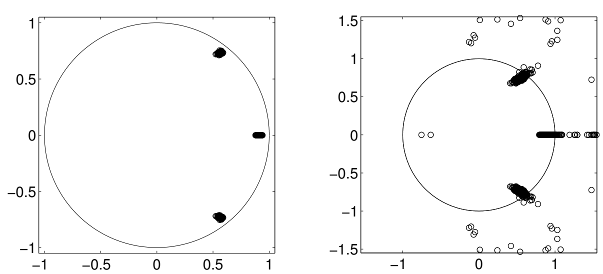
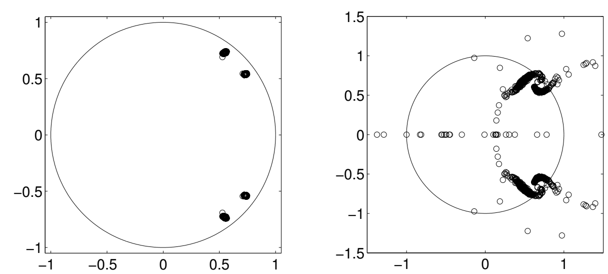


Fig. 3. True poles: Example 2, $N = 500$.



PEM YWILS

Fig. 2. Estimated poles for PEM and YWILS: Example 1, $N = 500$. Some poles estimated with YWILS fall outside the figure area.



PEM YWILS

Fig. 4. Estimated poles for PEM and YWILS: Example 2, $N = 500$. Some poles estimated with YWILS fall outside the figure area.

variance $\sigma_w^2 = 0.7834$, leading to a SNR of about 15 dB. Monte Carlo simulations of $M = 300$ independent runs have been carried out by considering $N = 250, 500, 1000$. The parameters m and δ in the YWILS algorithm have been set as in the first example. The obtained results are summarized in Table 2. Figures 3 and 4 report the true and estimated poles for the case $N = 500$.

For both models, the YWILS algorithm does not converge in some Monte Carlo runs. In these cases, the simulation has been stopped after 200 iterations. It can be observed that the performance of PEM is significantly better than that of YWILS. Moreover, the YWILS algorithm can be affected by convergence problems. The PEM approach does not suffer from convergence problems, but the algorithm can converge to a local minimum. To overcome this problem, it is possible to start the iterative minimization procedure at different initial estimates and to compare the obtained results. As the computational load of PEM is quite high with respect to that of YWILS, the YWILS algorithm is preferable when the number of data is high. In fact, in this case the YWILS performance is good and there are no convergence problems, see [20] where $N = 4000$ has been considered.

5. CONCLUSIONS AND PERSPECTIVES

In this paper, we have proposed an alternative approach to estimate the AR parameters from observations disturbed by an additive MA process. Our approach has the advantage of outperforming the existing method proposed by Mahmoudi *et al.* [20], especially when the number of samples is small (< 1000). We are currently investigating other solutions based on on-line EM based method or on the approximation of the MA part of model by a high-order AR model.

6. REFERENCES

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

[2] S. M. Kay, "Noise compensation for autoregressive spectral estimates," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 28, pp. 292–303, June 1980.

[3] 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, Oct. 1982.

[4] 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, Jan. 1999.

[5] W. X. Zheng, "Fast identification of autoregressive signals from noisy data," *IEEE Transactions on Circuits and Systems II: express briefs*, vol. 52, pp. 43–48, Jan. 2005.

[6] C. E. Davila, "A subspace approach to estimation of autoregressive parameters from noisy measurements," *IEEE Transactions Signal Processing*, vol. 46, pp. 531–534, Feb. 1998.

[7] 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, Dec. 2007.

[8] 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, June 2008.

[9] J. R. Treichler, "Transient and convergent behavior of the adaptive line enhancer," *IEEE Transactions on Acoustics, Speech, Signal Processing*, vol. 27, pp. 53–62, Feb. 1979.

[10] W. R. Wu and P. C. Chen, "Adaptive AR modeling in white Gaussian noise," *IEEE Transactions on Signal Processing*, vol. 45, pp. 1184–1191, May 1997.

[11] W. Li and Y. Jia, "H-infinity filtering for a class of nonlinear discrete-time systems based on unscented transform," *Signal Processing*, vol. 90, pp. 3301-3307, Dec. 2010.

[12] J.-S. Hu and C.-H. Yan, "Second-Order extended H_∞ filter for nonlinear discrete-time systems using quadratic error matrix approximation," *IEEE Transactions on Signal Processing*, vol. 59, pp. 3110–3119, July 2011.

[13] J.-H. Wang, C.-L. Song, X.-T. Yao and J.-B. Chen, "Sigma-Point H-infinity filter for initial alignment in marine strapdown inertial navigation system," in *Proc. of the 2nd IEEE International Conference on Signal Processing Systems (ICSPS 10)*, vol. 1, July 2000, pp. V1-580-V1-584.

[14] 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, Feb. 1983.

[15] J. Petitjean, E. Grivel, W. Bobillet and P. Roussilhe, "Errors-in-Variables for multichannel AR processes estimation from noisy observations," *Signal, Image and Video Processing (SIVIP)*, vol. 4, no. 2, pp. 209-220, 2010.

[16] A. Mahmoudi, M. Karimi, "Estimation of the parameters of multichannel autoregressive signals from noisy observations," *Signal Processing*, vol. 88, pp. 2777-2783, Nov. 2008.

[17] K. Hasan, J. Hossain, A. Haque, "Parameter estimation of multichannel autoregressive processes in noise," *Signal Processing*, vol. 83, pp. 603-610, Mar. 2003.

[18] H. Ijima and E. Grivel, "Deterministic regression methods for unbiased estimation of time-varying autoregressive parameters from noisy observations," *Signal Processing*, vol. 92, pp. 857-87, Apr. 2012.

[19] W. X. Zheng, "Estimation of the parameters of autoregressive signals from colored noise-corrupted measurements," *IEEE Signal Processing Letters*, vol. 7, pp. 201–204, July 2000.

[20] A. Mahmoudi and M. Karimi, "Parameter estimation of autoregressive signals from observations corrupted with colored noise," *Signal Processing*, vol. 90, pp. 157-164, Jan. 2010.

[21] A. Mahmoudi, M. Karimi and H.Amindavar, "Parameter estimation of autoregressive signals in presence of colored AR(1) noise as a quadratic eigenvalue problem," *Signal Processing*, vol. 92, pp. 1151-1156, Apr. 2012.

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

[23] T. Söderström and P. Stoica, *System Identification*, Prentice Hall, Cambridge, UK, 1989.

[24] T. Söderström, "Prediction error methods", in *Enciclopedia of Life Supportt Systems (EOLSS)*, EOLSS Publishers, Oxford, UK, 2003.

[25] J. Hu, K. Hirasawa and K. Kumamaru, "A homotopy approach to improving PEM identification of ARMAX models", *Automatica*, vol. 37, pp. 1323-1334, Sep. 2001.

[26] M. Pagano, "Estimation of models of autoregressive signal plus white noise," *The Annals of Statistics*, vol. 2, pp. 99–108, 1974.

[27] P. Stoica, T. McKelvey and J. Mari, "MA Estimation in Polynomial Time", *IEEE Transactions on Signal Processing*, vol. 48, pp. 1999-2012, July 2000.

[28] P. M. T. Broersen and S. de Waele, "Automatic identification of time-series models from long autoregressive models", *IEEE Transactions on Instrumentation and Measurement*, vol. 54, pp. 1862-1868, Oct. 2005.

[29] N. Sandgren, P. Stoica, and P. Babu, "On moving average parameter estimation", in *Proc. of the 20th European Signal Processing Conference (EUSIPCO 2012)*, Bucharest, Romania, August 2012, pp. 2348–2351.