SIViP (2015) 9 (Suppl 1):S235–S244
DOI 10.1007/s11760-015-0803-3

ORIGINAL PAPER

Identifying an autoregressive process disturbed by a moving-average noise using inner–outer factorization

Ahmed Abdou[1] · Flavius Turcu[2] · Eric Grivel[2] · Roberto Diversi[3] · Guillaume Ferré[2]

Received: 18 March 2015 / Revised: 22 May 2015 / Accepted: 15 July 2015 / Published online: 2 August 2015

Abstract This paper deals with the identification of an autoregressive (AR) process disturbed by an additive moving-average (MA) noise. Our approach operates as follows: Firstly, the AR parameters are estimated by using the overdetermined high-order Yule–Walker equations. The variance of the AR process driving process can be deduced by means of an orthogonal projection between two types of estimates of AR process correlation vectors. Then, the correlation sequence of the MA noise is estimated. Secondly, the MA parameters are obtained by using inner–outer factorization. To study the relevance of the resulting method, we compare it with existing algorithms, and we analyze the identifiability limits. The identification approach is then combined with Kalman filtering for channel estimation in mobile communication systems.

✉ Ahmed Abdou
aabdou@eng.alquds.edu

Flavius Turcu
flavius.turcu@ims-bordeaux.fr

Eric Grivel
eric.grivel@ims-bordeaux.fr

Roberto Diversi
roberto.diversi@unibo.it

Guillaume Ferré
guillaume.ferre@ims-bordeaux.fr

1 Department of Electronics and Communications Engineering, Al-Quds University, P.O. Box 20002, Jerusalem, Palestine

2 Université Bordeaux-Bordeaux-INP-ENSEIRB-MATMECA, UMR CNRS 5218 IMS - 351 cours de la Libération, 33405 Talence Cedex, France

3 Department of Electrical, Electronic and Information Engineering, University of Bologna, Viale del Risorgimento 2, 40136 Bologna, Italy

Keywords Autoregressive process (AR) · Moving-avergae process (MA) · Inner–outer factorization · Overdetermined high-order Yule–Walker equations

1 Introduction

A great deal of interest has been paid to autoregressive (AR) models. They are used in many fields such as speech processing and digital communications. Several approaches either on-line [1–3] or off-line [4–7] have been proposed to estimate the model parameters when the additive measurement noise is a zero-mean white Gaussian sequence. However, this assumption is not always representative of the reality. Therefore, the additive colored noise case must be studied, but few authors have addressed this issue [8–11], and different scenarios may happen:

1 In some applications such as speech enhancement, silent frames can be used to estimate the measurement noise parameters [12]. In this case, if the noise is assumed to be AR, the Yule–Walker (YW) equations or any recursive least-squares (LS) method can be considered. If the noise is a moving-average (MA) process, there are various ways to deduce the MA parameters such as Durbin's method [13]. The MA parameters can be also estimated in the maximum likelihood sense. However, this is a highly nonlinear problem, and solving it leads to a high computational cost. Therefore, covariance fitting approaches have been proposed, more particularly in [14,15]. Nevertheless, as the estimated covariances have to form a "valid" MA covariance sequence, i.e., a sequence guaranteeing the positivity of the corresponding power spectral density (PSD), some authors aimed at modifying the estimated MA covariance sequence to obtain this property [16,17]. In [18], another approach consists in taking the inverse

Fourier transform of the inverse of the MA PSD in order to get to the so-called inverse covariance sequence.

2 In other applications such as mobile communications, there is no signal-free period, i.e., no periods during which there is only the additive noise, at the receiver. Therefore, noisy data must be processed to estimate the parameters of both the AR signal and the colored noise, and this is particularly the case when referring to [19]. To our knowledge, few people have addressed this issue: In [10], the authors consider a first-order AR noise. When the additive noise is modeled by a MA process, an improved LS-based method [8] (also denoted "ILS-CN") aims at removing the biased caused by the colored noise to yield the unbiased estimates of the AR parameters. In [9], an improved LS-based method (denoted "YWILS method") combines low-order and high-order YW equations. The algorithm proposed in [11], which is based on the prediction error method (PEM) [20,21], leads to accurate results even when the number of samples is small. However, its computational cost is much higher than the ones of the ILS-CN- or YWILS methods.

In this paper, as done in [8,9,11], the model orders are assumed to be known. The approach we propose operates with the two following steps:

1 After estimating the AR parameters by using the overdetermined high-order YW equations and obtaining the variance of the AR driving process by means of an orthogonal projection between two types of estimates of AR process correlation vectors, the correlation function of the noise and the corresponding PSD are deduced. If the PSD is not positive, then an additional step based on [16,17] should be added.

2 The MA parameters are estimated by using the inner–outer factorization. Our contribution is twofold: We propose a new way to compute it, different from the one we used in [22]. In addition, we analyze the identifiability limits. Indeed, a MA process can be seen as the output of a filter, the input of which is zero mean with a given variance. Several filters make it possible to obtain the same correlation function of the MA process. Among them, there is just one minimum-phase filter, also known as "outer factor" or spectrally equivalent minimum-phase (SEMP) filter. The other solutions can be deduced from the SEMP filter by combining it with all-pass filters. When the driving process of the MA process is assumed to be Gaussian, we show that one can derive the SEMP filter and deduce all the solutions. Nevertheless, there is no way to say which one leads to the true MA parameters. The theorem we propose is based on Lukacs and King extension of Bernstein's theorem [23].

The rest of this paper is organized as follows: The problem statement and the proposed identification method are described in Sect. 2. A comparative study with [8,9,11] is carried out in Sect. 3. In Sect. 4, our approach is used in the field of mobile communication system for Kalman-filter-based channel estimation using a training sequence. Finally, conclusions and perspectives are given in Sect. 5.

2 Problem statement and the proposed identification method

Let us consider a pth-order AR process $x(n)$ defined by:

$$x(n) = -\sum_{i=1}^{p} a_i x(n-i) + u(n) \tag{1}$$

where $\{a_i\}_{i=1,\ldots,p}$ are the AR parameters and the driving process $u(n)$ is a zero-mean white Gaussian process with variance σ_u^2. $x(n)$ can be seen as the output of an infinite impulse response filter driven by $u(n)$, whose transfer function is only defined by the poles $\{p_i\}_{i=1,\ldots,p}$. The AR process is then disturbed by an additive measurement noise:

$$y(n) = x(n) + b(n) \tag{2}$$

where $b(n)$ is a qth-order MA noise:

$$b(n) = \sum_{j=0}^{q} c_j w(n-j) \tag{3}$$

with $\{c_j\}_{j=0,\ldots,q}$ the MA parameters and $w(n)$ is a zero-mean white Gaussian process with variance equal to 1 uncorrelated with $u(n)$. Given $\{z_i\}_{i=1,\ldots,q}$ the zeros of the MA process and using the z-transform, its transfer function is given by:

$$C(z) = \frac{B(z)}{W(z)} = \sum_{i=0}^{q} c_i z^{-i} = \prod_{i=1}^{q} \left(1 - z_i z^{-1}\right) \tag{4}$$

2.1 Step 1: estimating the AR model parameters

Given (1)–(3), the overdetermined high-order Yule–Walker equations (HOYW) equations [24] can provide the AR parameter estimates $\{\hat{a}_i\}_{i=1,\ldots,p}$.

Let us now estimate σ_u^2. On the one hand, if θ is the normalized angular frequency, the AR process PSD can be estimated, up to a multiplicative factor by:

$$\hat{D}_{yy}(f) = \frac{1}{\left\{\left|\sum_{i=0}^{p} \hat{a}_i z^{-i}\right|^2\right\}_{z=e^{j\theta}}} \tag{5}$$

By taking its inverse Fourier transform of $\hat{D}_{yy}(f)$, an estimation of $R_{xx}(\tau)/\sigma_u^2 = E[x(n)x(n-\tau)]/\sigma_u^2$ can be obtained and the following correlation column vector can be defined:

$$\hat{\underline{\hat{\mathbf{r}}}}_{xx} = \left[\hat{\hat{R}}_{xx}(q+s)\cdots\hat{\hat{R}}_{xx}(l)\right]^T \text{ with } \quad l > q+s \geq q+1 \tag{6}$$

where the length $l-q-s+1$ of the vector is chosen by the user.

On the other hand, let us consider the autocorrelation vector directly estimated from the noisy data:

$$\hat{\underline{\mathbf{r}}}_{yy} = \left[\hat{R}_{yy}(q+s)\cdots\hat{R}_{yy}(l)\right]^T \tag{7}$$

By computing the orthogonal projection of $\hat{\underline{\hat{\mathbf{r}}}}_{xx}$ onto the $\hat{\underline{\mathbf{r}}}_{yy}$, the variance of the driving process can be retrieved as follows:

$$\hat{\underline{\hat{\mathbf{r}}}}_{xx} = \frac{\hat{\underline{\hat{\mathbf{r}}}}_{xx}^T\hat{\underline{\mathbf{r}}}_{yy}}{\hat{\underline{\mathbf{r}}}_{yy}^T\hat{\underline{\mathbf{r}}}_{yy}}\hat{\underline{\mathbf{r}}}_{yy} + \text{residual} \approx \frac{1}{\hat{\sigma}_u^2}\hat{\underline{\mathbf{r}}}_{yy} \tag{8}$$

Then, we suggest deducing the whole correlation sequence of the AR process and subtracting it from the noisy observation correlation function estimated from the data.

Indeed, we can iteratively obtain an estimation $\hat{R}_{xx}^{(l)}(\tau)$ of the AR process correlation function for lags lower than $q+s$, starting from lag l:

$$\begin{cases} \hat{R}_{xx}^{(l)}(l-p) = \frac{-1}{\hat{a}_p}\left(\hat{R}_{xx}^{(l)}(l) + \sum_{i=1}^{p-1}\hat{a}_i\hat{R}_{xx}^{(l)}(l-i)\right), \\ \hat{R}_{xx}^{(l)}(\tau) = \hat{R}_{yy}(\tau) \text{ for } \quad l \geq \tau \geq l-p+1 \geq q+s, \end{cases} \tag{9}$$

However, depending on the initial lag l, the estimation of the correlation function of the AR process varies. To reduce the estimate variance, we suggest computing several correlation function sequences obtained with different initial lags and combining them by using a median or a mean function. This leads to the correlation sequences $\hat{R}_{xx}^{\text{median}}(\tau)$ and $\hat{R}_{xx}^{\text{mean}}(\tau)$, respectively. Thus, the estimation of the correlation sequence of the MA process satisfies:

$$\hat{R}_{bb}(\tau) = \hat{R}_{yy}(\tau) - \hat{R}_{xx}^{\text{mean (or median)}}(\tau) \quad \forall\tau < q+s \tag{10}$$

Then, $\hat{R}_{bb}(\tau)$ is modified so that $\hat{R}_{bb}(\tau) = 0$ for $|\tau| > q$ and the PSD of the MA process can be deduced by using the Fourier transform. If the PSD is not positive, then an additional step based on [16,17] can be introduced.

2.2 Step 2: estimating the MA parameters from the estimated MA PSD

To obtain the MA parameters, we propose a spectral factorization approach based on the estimation of the outer factor in the power spectral density. Let us first recall the main theoretical ideas, playing a key role in function theory in Hardy spaces [25]. Thus, let us denote:

$$D_{bb}\left(e^{j\theta}\right) = \sum_{\tau=-q}^{q} R_{bb}(\tau)e^{-j\tau\theta} \tag{11}$$

the PSD of the MA process (3). As all the results from the function theory we consider in this section are defined in the unit disk instead of its complement in the z-plane, the one-sided power series are considered in the following. Instead of the z-transform $C(z)$ used in (4), we work with[1]:

$$\tilde{C}(z) = \sum_{i=0}^{q} c_i z^i \tag{12}$$

It follows from (3), (11) and (12) that:

$$D_{bb}\left(e^{j\theta}\right) = \left|\tilde{C}\left(e^{j\theta}\right)\right|^2, \qquad \forall\theta \in [0, 2\pi[\tag{13}$$

The computation of the MA parameters amounts to solving the functional equation (13), which is generally referred to as the spectral factorization problem. It is well known [25] that whenever D_{bb} is a positive and log-integrable function on the unit disk in the z-plane, (13) has infinitely many solutions. The generic solution can be written as follows:

$$\tilde{C}(z) = \tilde{I}(z)\tilde{C}_0(z) \tag{14}$$

where $\tilde{I}(z)$ is an arbitrary inner function—i.e., unimodular on the unit disk in the z-plane, and $\tilde{C}_0(z)$ is the unique outer function satisfying (13). This latter particular solution $\tilde{C}_0$ has the standard explicit form:

$$\tilde{C}_0(z) = e^{\Phi(z)} \tag{15}$$

where $\Phi(z)$ is the analytic function given by the Poisson kernel integral:

$$\Phi(z) = \frac{1}{2\pi}\int_0^{2\pi} \frac{e^{j\theta_1}+z}{e^{j\theta_1}-z}\log\left(D_{bb}^{\frac{1}{2}}\left(e^{j\theta_1}\right)\right)d\theta_1 \tag{16}$$

In [22], the model parameters were expressed by means of $\tilde{C}_0(z)$, computed for z on a circle of radius smaller than 1 in the z-plane. In this paper, we propose an alternative way and suggest expressing the power series of $\Phi(z)$ by replacing the complex Poisson kernel in (16) by its analytic expansion. Indeed, one has:

[1] In this case, the exterior of the unit disk in the z-plane plays the role of the unit disk when using the z-transform.

$$\frac{e^{j\theta_1}+z}{e^{j\theta_1}-z}=1+2\sum_{n=1}^{+\infty}e^{-jn\theta_1}z^n \qquad \forall\theta_1\in[0,2\pi[,\,|z|<1 \tag{17}$$

Combining (16) and (17) yields:

$$\Phi(z)=d(0)+2\sum_{n\geq 1}d(n)z^n, \tag{18}$$

where:

$$d(n)=\frac{1}{2\pi}\int_0^{2\pi}e^{-jn\theta_1}\log\left(D_{bb}^{\frac{1}{2}}\left(e^{j\theta_1}\right)\right)d\theta_1 \quad \forall n\in\mathbb{Z} \tag{19}$$

Given (19), the sequence $\{d(n)\}$ represents the Fourier coefficients of $\log(D_{bb}^{\frac{1}{2}}(e^{j\theta}))$, also known as the cepstral coefficients of $D_{bb}^{\frac{1}{2}}(e^{j\theta})$. Hence, they satisfy:

$$\log\left(D_{bb}^{\frac{1}{2}}\left(e^{j\theta}\right)\right)=\sum_{n=-\infty}^{+\infty}d(n)e^{jn\theta} \tag{20}$$

Therefore, the real part of $\Phi(e^{j\theta})$ coincides with $\log(D_{bb}^{\frac{1}{2}}(e^{j\theta}))$ on the unit disk in the z-plane. Indeed, from (18) and (20), one has:

$$\begin{aligned}\mathrm{Re}\left(\Phi\left(e^{j\theta}\right)\right)&=\frac{1}{2}\left(\Phi\left(e^{j\theta}\right)+\overline{\Phi\left(e^{j\theta}\right)}\right)\\&=\sum_{n=-\infty}^{+\infty}d(n)e^{jn\theta}\underset{(20)}{=}\log\left(D_{bb}^{\frac{1}{2}}\left(e^{j\theta}\right)\right)\end{aligned} \tag{21}$$

Then, from (15) and (21), one can see that:

$$\left|\tilde{C}_0\left(e^{j\theta}\right)\right|^2\underset{(15)}{=}e^{2\mathrm{Re}(\Phi(e^{j\theta}))}\underset{(21)}{=}e^{\log D_{bb}(e^{j\theta})}=D_{bb}\left(e^{j\theta}\right) \tag{22}$$

$\tilde{C}_0$ is hence a solution of the Eq. (13).

Given the above considerations, the proposed implementation to estimate the outer factor $\tilde{C}_0$ is the following:

1. Given $\hat{R}_{bb}(\tau)$, estimate $D_{bb}(e^{j\theta})$ over an N-point uniform sampling of the unit circle, i.e., such as $\theta=\theta_k=2k\pi/N, k=0,1,\ldots N-1$. This can be done by using the discrete Fourier transform (DFT) of the sequence $\hat{R}_{bb}(\tau)$, padded with zeros up to N values if necessary.
2. Estimate $d(n)$ in (19), by the inverse DFT (IDFT) of the sequence $\frac{1}{2}\log D_{bb}(e^{j\theta_k})$.
3. Estimate $\Phi(e^{j\theta_k})$ in (18), by means of the DFT of the sequence $d(0), 2d(1), 2d(2)$, etc.
4. Deduce $\tilde{C}_0(e^{j\theta_k})=e^{\Phi(e^{j\theta_k})}$ according to (15).
5. Estimate the coefficients of $\tilde{C}_0$ by the IDFT of the sequence $\tilde{C}_0(e^{j\theta_k})$.

2.3 Some insights on the limitations of MA identification with a Gaussian driving noise process

As mentioned above, any other solution of (13) can be obtained by multiplying $\tilde{C}_0(z)$ with any inner function $\tilde{I}(z)$, which produces an IIR filter in general. A particular class of inner functions consists of the Blaschke products of the form:

$$\tilde{K}(z)=\gamma\prod_j\frac{z-\alpha_j}{1-\bar{\alpha_j}z}=\sum_{l\geq 0}k_lz^l|\gamma|=1,\,|\alpha_j|<1 \tag{23}$$

where k_l are the coefficients of the Blaschke products.

They can be seen as transfer functions of discrete-time all-pass filters. In particular, if $\tilde{C}_0(z)$ is a polynomial corresponding to a finite impulse response filter (FIR),—as in our case—and if one looks for other polynomial solutions for (13) of the form $\tilde{C}(z)=\tilde{K}(z)\tilde{C}_0(z)$, then the complex numbers $1/\bar{\alpha_j}$ have to be roots of $\tilde{C}_0(z)$. Therefore, multiplying $\tilde{C}_0(z)$ by Blaschke products $\tilde{K}(z)$ amounts to replacing some roots of $\tilde{C}_0(z)$ with their inverse conjugates and hence producing a finite number of FIR solutions $\tilde{C}(z)$ in (14), which are not minimum phase.

Let us also remark that the inner–outer factorization (14) is related to, but distinct from the minimum-phase/maximum-phase factorization, see e.g., [26]. This latter is obtained by simply factorizing apart the roots of $\tilde{C}(z)$ inside the unit disk from those outside. In contrast, the inner–outer factorization always involves the outer factor, along with the Blaschke product corresponding to the roots of $\tilde{C}(z)$ appearing in its maximum-phase factor.

The fact that (13) has multiple solutions leads to the widely known fact that the second-order statistics are "phase-blind." This means that the autocorrelation sequence alone only allows the outer factor $\tilde{C}_0(z)$ to be computed. To get $\tilde{C}(z)$, one has to tell which roots of $\tilde{C}_0(z)$ have to be reflected with respect to the unit circle, or, in other words, one has to specify the suitable Blaschke multiplier $\tilde{K}(z)$. This has to be done by using other information than the autocorrelation sequence of the output of $\tilde{C}(z)$ on the unknown driving process $w(n)$. For various classes of driving processes that are not Gaussian, higher-order cumulants such as in [26] or [27] can be used. This essentially works for driving processes that have a non-symmetric probability distribution, yielding nonzero odd-order cumulants. Unfortunately, in the Gaussian case, not only such techniques are useless but, worse, there is actually no way to solve this problem. More precisely, let us consider the factorization $\tilde{C}(z)=\tilde{K}(z)\tilde{C}_0(z)$. Then, the output b of the filter $\tilde{C}(z)$, whose Gaussian driving noise is $w(n)$, is the same as the output of $\tilde{C}_0(z)$, whose driving noise $w'(n)$ is given by:

$$w'(n) = \sum_{l \geq 0} k(l) w(n-l) \quad \forall n \in \mathbf{Z}. \tag{24}$$

Filtering $w(n)$ with $\tilde{K}(z)$ does not change its power spectrum, since it is all-pass. However, the question is whether $w'(n)$ still remains a Gaussian white noise or not. If it does, there is no way to distinguish between filtering the unknown Gaussian white noise $w(n)$ with $\tilde{C}(z)$, or filtering the also unknown Gaussian white noise $w'(n)$ with $\tilde{C}_0(z)$. Unfortunately the answer is true, not only for a Blaschke product but for any inner function (and only for inner functions). Even if it seems to be well known in practice and is generally illustrated in various papers with numerical examples, we give here a rigorous formalization of this phenomenon and its proof:

Theorem *According to (24) and if* $w(n) \sim \mathcal{N}(0, \sigma^2)$, $w'(n) \sim \mathcal{N}(0, \sigma'^2)$, *where* $\sigma'^2 = \sigma^2(\sum_n |k(n)|^2) = \sigma^2 \|\tilde{K}\|_2^2$ *according to the Parseval identity, this implies that:*

$$\sigma = \sigma' \quad \text{if and only if} \quad \|\tilde{K}\|_2 = 1 \tag{25}$$

Proof Let us recall Lukacs and King extension of Bernstein's theorem [23]: if $(w(n))_{n \in \mathbb{Z}}$ is a family of independent variables (not necessarily identically distributed) with variance σ_n^2 and if $X = \sum_n \alpha_n w(n)$ and $Y = \sum_n \beta_n w(n)$ are linear combinations of $w(n)$, then X and Y are independent if and only if all $w(n)$ are normal and $\sum_n \alpha_n \beta_n \sigma_n^2 = 0$. Therefore, if $X = w'(n)$ and $Y = w'(m)$ where $n \neq m$, the above independence condition leads to:

$$\sigma^2 \sum_l k(l) k(l + (n-m)) = 0 \tag{26}$$

Using the Parseval identity, (26) can be expressed by means of the scalar product of $\tilde{K}(z)$ and $z^{n-m}\tilde{K}(z)$ in L^2 of the unit circle:

$$0 = \sum_l k(l) k(l + (n-m)) = \left\langle \tilde{K}(z), z^{n-m}\tilde{K}(z) \right\rangle_{L^2}$$

$$= \frac{1}{2\pi} \int_0^{2\pi} \tilde{K}\left(e^{j\theta_1}\right) \overline{\tilde{K}\left(e^{j\theta_1}\right)} e^{-(n-m)j\theta_1} d\theta_1 \tag{27}$$

$$= \frac{1}{2\pi} \int_0^{2\pi} \left|\tilde{K}\left(e^{j\theta_1}\right)\right|^2 e^{-(n-m)j\theta_1} d\theta_1. \tag{28}$$

Therefore $w'(n)$ and $w'(m)$ are independent for any $n \neq m$ if and only if all the Fourier coefficients of $|\tilde{K}|^2$ are zero, except for the central coefficient. It happens if and only if $|\tilde{K}|$ is constant on the unit circle. Given (25), $(w'(n))_n$ are i.i.d. $\mathcal{N}(0, \sigma^2)$ if and only if $|\tilde{K}| = 1$ on the unit circle, hence if and only if $\tilde{K}$ is an inner function, and the proof is complete. □

The straightforward consequence is that, besides its outer factor, there is no way to recover any other information on a MA process driven by a Gaussian process, by solely knowing its Gaussian nature. To do this, one would need to know more information on the driving noise.

3 Simulation results and comparative study

In this section, the relevance of our approach is evaluated in terms of accuracy and computational cost. Note that we compare our approach with ILS-CN [8], YWILS [9] and PEM based [11] methods. For this purpose, two sets of simulations are presented.

3.1 First simulations: protocol, results and comments

Let us consider the following simulation protocol: The AR process order p is set to 2. The AR parameters are 1, −1.0463, 0.7921 corresponding to the two poles $p_1 = 0.89e^{j2\pi/3}$ and $p_2 = \overline{p_1}$. Three types of additive MA noise are now considered and are defined by their zeros:

1. $z_1 = -3$. The detailed results are provided in Table1 for different numbers of samples.
2. $z_1 = 2.853 + j0.927$ and $z_2 = \overline{z_1}$. The zeros are outside the unit disk in the z-plane. The estimated zeros should hence be outside the unit disk in the z-plane, as explained in the discussion in the above Sect. 2.3. The results are depicted in Fig. 1 where the true and estimated poles are given in the z-plane. Moreover, Fig. 2a–c gives the estimated zeros.
3. $z_1 = 0.247 + j0.076$ and $z_2 = \overline{z_1}$. The zeros are inside the unit disk in the z-plane. The estimated zeros should hence correspond to estimations of their inverse conjugates. Fig. 2b–d gives the estimated zeros.

In the following and without loss of generality, let the signal-to-noise ratio (SNR), which is the ratio between the powers of the AR signal and the MA noise, be equal to 15 dB. Concerning our method, s and r are set to 1 and 5, respectively.

For the three types of zeros and according to Table 1, Figs. 1 and 2, the approach we propose provides reliable results. According to various tests we did, using <500 samples (e.g., 200 samples) does not make it possible to obtain accurate estimates of MA parameters for every method. In addition, the higher the number of available samples is, the more accurate the estimations of the MA parameters are. According to Table 1, the comparative study with the PEM

Table 1 Comparative study between our method, PEM [11], YWILS [9] and ILS-CN [8] for the first simulation protocol where $z_1 = -3$

Number of samples		Our method				PEM [11]			
$p = 2, q = 1$	True	500	1000	2000	4000	500	1000	2000	4000
a_1	−1.046	−1.0462	−1.0465	−1.045	−1.0458	−1.0344	−1.036	−1.0365	−1.0361
a_2	0.7921	0.7937	0.793	0.7921	0.7908	0.7782	0.7804	0.7804	0.7805
p_1	0.89	0.8909	0.8905	0.8893	0.8897	0.8821	0.8834	0.8834	0.8835
θ_1	0.9425	0.9433	0.9427	0.9434	0.9422	0.9443	0.9443	0.9439	0.9442
p_2	0.89	0.8909	0.8905	0.8893	0.8897	0.8821	0.8834	0.8834	0.8835
θ_2	−0.9425	−0.9433	−0.9427	−0.9434	−0.9422	−0.9443	−0.9443	−0.9439	−0.9442
b_1“mean”	3	3.5424	3.2143	3.1111	3.0216	2.9864	2.9836	2.9861	2.9738
b_1“median”	3	3.6406	3.3784	3.1027	3.0568	–			
σ_u^2	1	0.9736	0.9932	0.9942	1.0008	1.0228	1.0174	0.9587	0.9517

Number of samples		YWILS [9]				ILS-CN [8]			
$p = 2, q = 1$	True	500	1000	2000	4000	500	1000	2000	4000
a_1	−1.046	−1.0945	−1.0449	−1.077	−1.037	−1.0965	−1.0461	−1.070	−1.041
a_2	0.7921	0.8670	0.7946	0.8367	0.8123	0.8760	0.7469	0.8637	0.8532
p_1	0.89	0.93	0.8911	0.9144	0.9024	0.893	0.9101	0.9234	1.024
θ_1	0.9425	0.9416	0.9443	0.9411	0.9421	0.9141	0.9943	0.8911	0.9331
p_2	0.89	0.93	0.8911	0.9144	0.9024	0.893	0.9101	0.9234	1.024
θ_2	−0.9425	−0.9416	−0.9443	−0.9411	−0.9421	−0.9141	−0.9943	−0.8911	−0.9331
σ_u^2	1	91.0336	28.7616	19.4237	6.152	–			

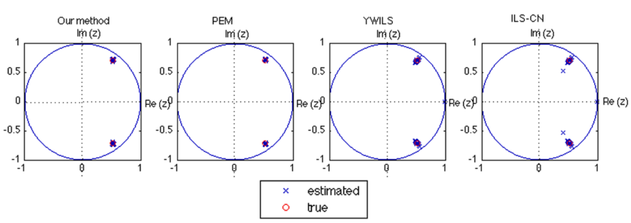
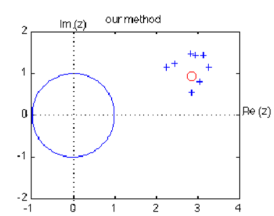
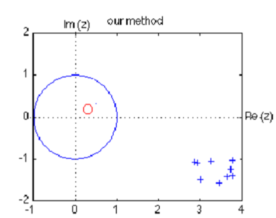


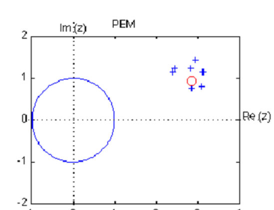
Fig. 1 True and estimated AR poles in the z-plane for the first simulation protocol to evaluate estimation accuracy for SNR = 15 dB, the number of samples = 2000, $z_1 = 2.853 + j0.927$ and $z_2 = \overline{z_1}$



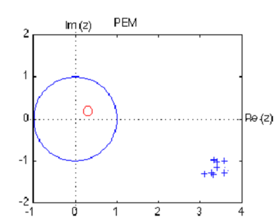
(a) zero outside the unit disc “our method”.



(b) zero inside the unit disc “our method”.



(c) zero outside the unit disc “PEM”.



(d) zero inside the unit disc “PEM”.

Fig. 2 Our method versus PEM [11] to estimate the zeros of the MA process when $z_1 = 2.853 + j0.927$ and $z_2 = \overline{z_1}$ **a–c** and $z_1 = 0.247 + j0.076$ and $z_2 = \overline{z_1}$ **b–d** (in order to make it clearer, z_1 is only presented)

also shows that there is “almost” no differences between them, especially when the number of samples is higher or equal than 1000 samples. The PEM is, however, less sensitive to the number of samples than our method. Even if the PEM approach is asymptotically unbiased, it is sensitive to the initial estimate. Therefore, when the starting estimates are poor, the final results may be poor as well. Although initial estimates of the AR parameters can be obtained by using the HOYW equations, this is not necessarily the case for the MA parameters. In addition, it should be noted that the latter is an iterative method that requires the computation of the innovation of a Kalman filter at each step. Its computational load is hence much more higher than the one of our new method. The PEM we suggested using in [11] takes three times longer than our approach. Our new method is hence a good compromise between estimation performance and computational cost.

3.2 Second simulations: protocol, results and comments

Now, let us look at the second simulation protocol: p and q are, respectively, set to 4 and 1. The AR parameters are cho-

Table 2 Comparative study between our method, PEM [11], YWILS [9] and ILS-CN [8], based on 5000 realizations and 1000 samples available, $p = 4$ and $q = 1$ (PPEV means process parameter error vector and NT the type of the norm)

	Methods	PEM [11]		Our method		YWILS [9]		ILS-CN [8]	
SNR	PPEV	NT							
		$\|\cdot\|_2^2$	$\|\cdot\|_\infty$	$\|\cdot\|_2^2$	$\|\cdot\|_\infty$	$\|\cdot\|_2^2$	$\|\cdot\|_\infty$	$\|\cdot\|_2^2$	$\|\cdot\|_\infty$
5 dB	AR	1.55	1.87	1.73	1.96	3.09	2.20	3.55	2.22
	MA	0.30	0.92	0.45	1.29	Not considered by the model			
10 dB	AR	1.23	1.74	1.43	1.80	2.58	2.11	2.74	2.20
	MA	1.50	1.73	1.79	2.07	Not considered by the model			
15 dB	AR	1.06	1.71	1.29	1.78	2.20	1.98	2.36	2.06
	MA	2.19	3.60	2.26	4.56	Not considered by the model			
20 dB	AR	0.77	1.51	0.92	1.72	1.12	1.93	1.34	2.01
	MA	3.19	4.81	3.99	5.25	Not considered by the model			
25 dB	AR	0.40	1.50	0.49	1.58	0.59	1.92	0.67	1.97
	MA	4.92	8.35	5.25	9.78	Not considered by the model			

sen randomly by selecting the corresponding AR poles, $p_1 = r_1 e^{j\phi_1}$, $p_2 = \overline{p_1}$, $p_3 = r_2 e^{j\phi_2}$, $p_4 = \overline{p_3}$ where $\{r_i\}_{i=1,2}$ and $\{\phi_i\}_{i=1,2}$ are independent random variables uniformly distributed in the range $[0, 1[$ and $[0, \pi[$, respectively. The MA parameter is chosen randomly, so that the zero z_1 is a random variable $\in [-1, -5]$. As the zero is outside the unit disk in the z-plane, the identification is here possible according to the discussion of Sect. 2.3. We look at the AR parameter estimation error vector[2] defined by: $\tilde{\mathbf{a}} = \left[(a_1 - \hat{a}_1) \cdots (a_p - \hat{a}_p)\right]$ and the MA parameters estimation error vector defined by: $\tilde{\mathbf{c}} = \left[(c_0 - \hat{c}_0) \cdots (c_q - \hat{c}_q)\right]$.

As shown in Table 2, the PEM we presented in [11] and our new method provide significant results compared with YWILS and ILS-CN methods. Globally, the estimations of the AR parameters and the MA parameter are comparatively accurate for SNR higher than 10 dB. In addition, if the SNR increases, the MA parameter estimation error increases. This is due to the fact that the MA process is the noise and not the signal. Hence, the power driven by the AR process makes it "difficult" to have a good estimation of the properties of the disturbance i.e., the MA noise.

3.3 Conclusions

In the above simulation sets, we can see that the PEM we suggested using in [11] is the most accurate, but its main drawback is its computational cost. In some applications, the accuracy is not necessarily the only priority. Thus, in mobile communication systems, the identification of the system can be required, but very accurate estimates of the AR and MA parameters are not useful to deduce the transmitted symbols when using Kalman filtering. A happy medium between estimation precision and computational cost is usually rather preferred. In the next section, we hence compare our new method and the PEM [11].

4 Application in channel estimation

In mobile communication systems, the transmitted signal arrives at the receiver from different propagation paths. Indeed, multipath causes distortions on the transmitted signal. The estimation of the channel is essential to achieve coherent symbol detection at the receiver. In this section, an orthogonal frequency division multiplexing (OFDM) system is considered with a flat fading channel on each subcarrier. Thus, the received signal on the mth subcarrier $y_m(n)$ corresponds to the product of the training sequence symbols $s_m(n)$ and the channel $h_m(n)$. The symbols are assumed to belong to a binary phase shift keying (BPSK) constellation, i.e., $s(n) \in \{-1, 1\}$. In addition, we assume that the received signal is disturbed by a colored noise $b_m(n)$. Therefore, it can be modeled as follows:

$$y_m(n) = s_m(n)h_m(n) + b_m(n) \qquad (29)$$

In the following $h_m(n)$ and $b_m(n)$ are assumed to be, respectively, modeled by AR and MA models. Our purpose is hence to estimate the AR and MA parameters during the training period, i.e., by means of a training sequence known both at the transmitter and the receiver. A specific choice must be done on the training sequence.

4.1 State space representation of the system

In order to get the corresponding state space representation of the system, let us introduce the following quantities:

$$\xi_j(n) = \sum_{i=1}^{q+1-j} c_{i+j-1} w(n-i) \quad \forall j \in 1, \ldots, q \qquad (30)$$

[2] More particularly, the square of its 2-norm defined by: $\|\,.\,\|_2^2 = \tilde{\mathbf{a}}\tilde{\mathbf{a}}^T$ and its ∞-norm defined by $\|\,.\,\|_\infty = max\{(a_i - \hat{a}_i)\}_{i=1\ldots p}$.

This implies that:

$$\xi_1(n) = \sum_{i=1}^{q} c_i w(n-i) \tag{31}$$

and,

$$\xi_j(n) = \xi_{j+1}(n-1) + c_j w(n-1) \quad \forall q \geqslant j \geqslant 2 \tag{32}$$

Given $\xi_1(n)$, (29) can be written as follows:

$$y_m(n) = s_m(n)h_m(n) + \xi_1(n) + c_0 w(n) \tag{33}$$

In this case, let us define the state vector as follows:

$$\mathbf{x}(n) = \left[h_m(n)\cdots h_m(n-p+1)\xi_1(n)\cdots\xi_q(n)\right]^T \tag{34}$$

The corresponding state space representation of the system satisfies:

$$\begin{aligned} \mathbf{x}(n) &= \mathbf{F}\mathbf{x}(n-1) + \mathbf{G}\mathbf{u}(n) \\ y_m(n) &= \mathbf{s}(n)\mathbf{x}(n) + c_0 w(n) \end{aligned} \tag{35}$$

where $\mathbf{s}(n) = [s_m(n)0\cdots010\cdots0]$,

$$\mathbf{F} = \begin{bmatrix} -a_1 & \cdots & -a_p & & & \\ 1 & 0\cdots & 0 & & \mathbf{O}_{p\times q} & \\ \vdots & \ddots & \vdots & & & \\ 0 & \cdots 1 & 0 & & & \\ & & & 0 & 1\cdots & 0 \\ & \mathbf{O}_{q\times p} & & \vdots & \ddots & \vdots \\ & & & & & 1 \\ & & & 0 & \cdots & 0 \end{bmatrix} \tag{36}$$

$$\mathbf{G} = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & c_1 & \cdots & c_q \end{bmatrix}^T \tag{37}$$

and

$$\mathbf{u}(n) = \left[\, u(n) \;\; w(n-1) \,\right]^T \tag{38}$$

For this case, let us define $z(n)$ as follows:

$$\begin{aligned} z(n) &= s_m(n)y_m(n) = s_m(n)^2 h_m(n) + s_m(n)b_m(n) \\ &= h_m(n) + s_m(n)b_m(n) \end{aligned} \tag{39}$$

Taking the correlation of (39) we get,

$$R_{zz}(\tau) = R_{hh}(\tau) + s_m(n)s_m(n-\tau)R_{bb}(\tau) \tag{40}$$

One can notice that the process $z(n)$ is not necessarily wide-sense stationary because its correlation function expressed in (40) depends on the symbol $s_m(n)$ or another previous symbol $s_m(n-\tau)$. To avoid the above phenomenon, the symbols used in the training sequence must be the same. Indeed, in this case, Eq. (40) becomes:

$$R_{zz}(\tau) = R_{hh}(\tau) + R_{bb}(\tau) \tag{41}$$

Then, the overdetermined HOYW equations in this system can be considered.

Remark Using subspace methods for identification such as N4SID could have been *a priori* of interest. This method was, for instance, used for speech enhancement in [28] where the speech signal was modeled by an AR process whereas the additive noise was white. In [29], channel estimation was seen as a realization issue where the additive noise was also white. The above methods operate with the following steps:

1. Estimating the quadruple defined by the transition matrix, the observation vector and the covariance matrices of the model noise and the observation noise of the state space representation of the system.
2. Using Kalman filtering with this quadruple to estimate the speech signal or the channel. In [28] and [29], the authors suggested using this method because the observation equation in the state space representation of the system had two distinct parts: One corresponding to the product between the observation vector and the state vector was related to the channel alone, whereas the second one defined the additive noise. In our paper where the additive noise is colored, if we look at (35), one can see that this property is not satisfied as both channel and noise are involved in the definition of the state vector.

4.2 Simulations and results for channel estimation

We consider an OFDM system with BPSK modulation and 64 subcarriers. The fading channels are approximated by an AR process with an order[3] p set to 2; the AR parameters are set to $a_1 = -1.0463$ and $a_2 = 0.7921$. The received signal is disturbed by an additive colored noise modeled by a MA process with an order q set to 1; the MA parameters are set to $c_0 = 1$ and $c_1 = -3$. The training sequence symbols are all set to 1 with a length $N = 512$.

It is well known that the received signal after the fast Fourier transform (FFT) can be described as in (29) [30]. By using the proposed method, and assuming the AR and MA orders are known, the AR parameters, the MA parameters and the variance of the driving process can be estimated during the training sequence period. For the true data, using Kalman

[3] It should be noted that the channel is approximated with high-order AR model. Here, we assume a second-order AR model to approximate the channel as an example.

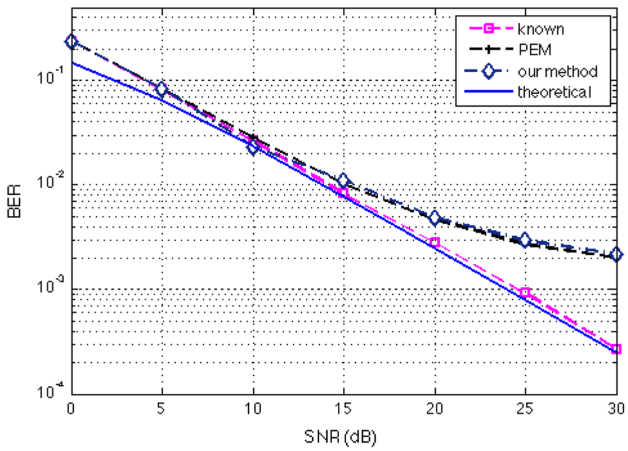


Fig. 3 BER versus SNR of the OFDM using the estimated channel based on our method, the PEM [11], when the AR and MA parameters are known

filter with the estimated parameters makes it possible to predict the channel and then to retrieve the transmitted data. The effect of the channel estimation error on the system performance is shown in Fig. 3, which shows that the higher the SNR is, the smaller the bit error rate (BER) is. By considering a number of samples available equal to 2000, the performance of the PEM-based approach and our new method are approximately the same. However, the computational cost of our new method based on inner–outer factorization is much lower. Therefore, this application confirms the relevance of our method.

5 Conclusions

The identification method proposed in this paper consists in estimating the AR parameters by means of an overdetermined set of HOYW equations and in estimating the MA parameters by using the inner–outer factorization. A comparative study with other methods such as the PEM we suggested in a previous paper confirms the efficiency of the proposed scheme. Our approach corresponds to a good compromise between computational cost and estimation accuracy. Combined with Kalman filtering, it can be of interest for channel estimation in mobile communication system, as pointed out by our simulation results.

References

1. Nehorai, A., Stoica, P.: Adaptive algorithms for constrained ARMA signals in the presence of noise. IEEE Trans. Acoust. Speech Sign. Process. **36**(8), 1282–1291 (1988)
2. Petitjean, J., Diversi, R., Grivel, E., Guidorzi, R., Roussilhe, P.: Recursive errors-in-variables approach for AR parameter estimation from noisy observations. Application to radar sea clutter rejection. In: Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), pp. 3401–3404 (2009)
3. Wu, W., Chen, P.: Adaptive AR modeling in white Gaussian noise. IEEE Trans. Sign. Process. **45**(5), 1184–1191 (1997)
4. Chung, W., Un, C.: Iterative autoregressive parameter estimation in presence of additive white noise. IEEE Electron. Lett. **27**(20), 1800–1802 (1991)
5. Davila, C.: A subspace approach to estimation of autoregressive parameters from noisy measurements. IEEE Trans. Sign. Process. **46**(2), 530–534 (1998)
6. Petitjean, J., Grivel, E., Bobillet, W., Roussilhe, P.: Multichannel AR parameter estimation from noisy observations as an errors-in-variables issue. Signal, Image Video Process. **4**(2), 209–220 (2010)
7. Bobillet, W., Diversi, R., Grivel, E., Guidorzi, R., Najim, M., Soverini, U.: Speech enhancement combining optimal smoothing and Errors-In-Variables identification of noisy AR processes. IEEE Trans. Sign. Process. **55**(12), 5564–5578 (2007)
8. Zheng, W.: Estimation of the parameters of autoregressive signals from colored noise-corrupted measurements. IEEE Sign. Process. Lett. **7**(7), 201–204 (2000)
9. Mahmoudi, A., Karimi, M.: Parameter estimation of autoregressive signals from observations corrupted with colored noise. Sign. Process. **90**(1), 157–164 (2010)
10. Mahmoudi, A., Karimi, M., Amindavar, H.: Parameter estimation of autoregressive signals in presence of colored AR(1) noise as a quadratic eigenvalue problem. Sign. Process. **92**(4), 1151–1156 (2012)
11. Diversi, R., Ijima, H., Grivel, E.: Prediction error method to estimate the AR parameters when the process is disturbed by a colored noise. In: Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), pp. 6143–6147 (2013)
12. Gibson, J., Boneung, K., Gray, S.: Filtering of colored noise for speech enhancement and coding. IEEE Trans. Sign. Process. **39**(8), 1732–1742 (1991)
13. Durbin, J.: Efficient estimation of parameters in moving average models. Biometrika **46**(3/4), 306–316 (1959)
14. Stoica, P., McKelvey, T., Mari, J.: MA estimation in polynomial time. IEEE Trans. Sign. Process. **48**(7), 1999–2012 (2000)
15. Dumitrescu, B., Tabus, I., Stoica, P.: On the parameterization of positive real sequences and MA parameter estimation. IEEE Trans. Sign. Process. **49**(11), 2630–2639 (2001)
16. Moses, R., Liu, D.: Optimal nonnegative definite approximations of estimated moving average covariance sequences. IEEE Trans. Sign. Process. **39**(9), 2007–2015 (1991)
17. Stoica, A., Moses, R., Stoica, P.: Enforcing positiveness on estimated spectral densities. Electron. Lett. **29**(23), 2009–2011 (1993)
18. Bhansali, R.: A simulation study of autoregressive and window estimators of the inverse correlation function. J R Stat Soc Appl Stat **32**(2), 141–149 (1983)
19. Rusek, F., Anderson, J.: Constrained capacities for faster-than-Nyquist signaling. IEEE Trans. Inf. Theory **55**(2), 764–775 (2009)
20. Box, G., Jenkins, D.: Time Series Analysis, Forecasting and Control. Hoden-Day, San Francisco (1970)
21. Ljung, L.: System Identification—Theory for the User. Prentice Hall, Englewood Cliffs, NJ (1999)
22. Merchan, F., Turcu, F., Grivel, E., Najim, M.: Rayleigh fading channel simulator based on inner-outer factorization. Sign. Process. **90**(1), 24–33 (2010)
23. Lukacs, E., King, E.: A property of the normal distribution. Ann. Math. Stat. **25**(2), 389–394 (1954)
24. Najim, M.: Modeling Estimation and Optimal Filtering in Signal Processing. Chapter 3. ISTE Ltd and Wiley, New York (2008)
25. Hoffman, K.: Banach Spaces of Analytic Functions. Prentice Hall Inc., Englewood Cliffs NJ (1962)

26. Nikias, C., Mendel, J.: Signal processing with higher order spectra. IEEE Sign. Process. Mag. **10**(3), 10–37 (1993)
27. Lii, K., Rosenblatt, M.: Deconvolution and estimation of transfer function phase and coefficients for nonGaussian linear processes. Ann. Stat. **10**(4), 1195–1208 (1982)
28. Grivel, E., Gabrea, M., Najim, M.: Speech enhancement as a realization issue. Sign. Process., Elsevier **82**(12), 1963–1978 (2002)
29. Grolleau, J., Grivel, E., Najim, M.: Subspace identification method for Rayleigh channel estimation. In: Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), vol. 3, pp. 449–452 (2007)
30. Jamoos, A., Abdou, A., Abdel-Nour, H., Grivel, E.: Two cross-coupled $H\infty$ filters for fading channel estimation in OFDM systems. In: Proceedings of the Computer, Information, Systems Sciences and Engineering (CISSE), pp. 349–353. Springer, Heidelberg (2010)