Available at
www.ElsevierComputerScience.com
POWERED BY SCIENCE DIRECT

Signal Processing 86 (2006) 2289–2295

SIGNAL PROCESSING

www.elsevier.com/locate/sigpro

Asymptotically minimum variance second-order estimation for complex circular processes

Jean-Pierre Delmas*, Yann Meurisse

Département CITI, Institut National des Télécommunications, 9 rue Charles Fourier, 91011 Evry Cedex, France

Received 14 February 2003; received in revised form 13 June 2005
Available online 29 November 2005

Abstract

This paper addresses asymptotically minimum variance (AMV) of parameter estimators within the class of algorithms based on second-order statistics for estimating parameter of strict-sense stationary complex circular processes. As an application, the estimation of the frequencies of cisoids for mixed spectra time series containing a sum of cisoids and an MA process is considered.

Keywords: Asymptotically minimum variance algorithms; Covariance matching estimation techniques; Frequency estimation of mixed complex circular processes; Second-order statistics-based algorithms

1. Introduction

There is considerable literature about second-order statistics-based algorithms. To provide a benchmark for the efficiency of such existing algorithms, it has been proposed to consider a general lower bound for the variance of consistent estimators based on second-order moments which is asymptotically tight (in the number of measurements). Stoica et al. with their asymptotically best consistent (ABC) estimators [1] and Porat and Friedlander [2] with their asymptotically minimum variance (AMV) estimator were the first to derive such estimators for estimating the ARMA parameters of real Gaussian processes from second-order statistics. Then, this approach was extended to high-order statistics [3] for real-valued processes and was used in many applications (see e.g., the work by Giannakis and Halford [4] for blind real-valued channel estimation). We propose to consider in this paper the case of second-order statistics derived from stationary complex circular processes.

The paper is organized as follows. Section 2 presents the AMV second-order estimator for stationary complex circular processes with a special attention to the statistics involved. It is proved that this AMV estimator is not a direct extension of the real-valued associated AMV estimator using the conjugate transpose instead of transpose. As an application, the estimation of the frequencies of cisoids for mixed spectra time series containing a sum of cisoids and an MA process is considered in Section 3. Finally, illustrative examples with comparisons with the modified Pisarenko decomposition (MPD) estimator [5] which is devoted to MA noise are given in Section 4.

*Corresponding author. Tel.: +33 1 60 76 46 32;
fax: +33 1 60 76 44 33.
E-mail address: jean-pierre.delmas@int-evry.fr (J.-P. Delmas).

0165-1684/$ - see front matter
doi:10.1016/j.sigpro.2005.10.011

The following notations are used throughout the paper. Matrices and vectors are represented by bold upper case and bold lower case characters, respectively. Vectors are by default in column orientation, while T, H, $*$ stand for transpose, conjugate transpose, conjugate, respectively. Vec($\cdot$) is the "vectorization" operator that turns a matrix into a vector by stacking the columns of the matrix one below another. arg min [resp. arg max] refers to the minimizing [resp. maximizing] argument of the proceeding expression.

2. Asymptotic minimum variance second-order estimator

We consider a strict-sense stationary complex circular process x_t whose $M \times M$ Hermitian Toeplitz structured covariance matrix $\mathbf{R}(\Theta) = \mathrm{E}(\mathbf{x}_t\mathbf{x}_t^{\mathrm{H}})$ with $\mathbf{x}_t \stackrel{\text{def}}{=} (x_t, \ldots, x_{t-M+1})^{\mathrm{T}}$, is parameterized by the real parameter $\Theta \in \mathbb{R}^L$. This parameter is supposed identifiable from $\mathbf{R}(\Theta)$ in the following sense:

$$\mathbf{R}(\Theta) = \mathbf{R}(\Theta') \Leftrightarrow \Theta = \Theta'.$$

The covariance matrix $\mathbf{R}(\Theta)$ is traditionally estimated by $\mathbf{R}_{\mathrm{T}} = (1/\mathrm{T})\sum_{t=1}^{\mathrm{T}} \mathbf{x}_t\mathbf{x}_t^{\mathrm{H}}$ or by the Hermitian Toeplitz matrix $\mathbf{R}_{\mathrm{T}}^{\mathrm{to}}$ built by averaging along the diagonals of $\mathbf{R}_{\mathrm{T}}$.

In the sequel, we restrict our study to the following mixed-spectrum processes:

$$x_t = s_t + n_t \quad \text{with } s_t \stackrel{\text{def}}{=} \sum_{k=1}^{K} a_k \mathrm{e}^{\mathrm{i}2\pi f_k t}\mathrm{e}^{\mathrm{i}\phi_k}$$

$$\text{and } n_t \stackrel{\text{def}}{=} \sum_{q=-\infty}^{+\infty} b_q u_{t-q},$$

where $(u_t)_{t=0,\pm1,\pm2,\ldots}$ is a sequence of circular complex zero-mean i.i.d. random variables where $\mathrm{E}|u_t^4| < \infty$, with $\kappa_u \stackrel{\text{def}}{=} \mathrm{Cum}(u_t, u_t^*, u_t, u_t^*)$ and $\mathrm{E}|u_t^2| = \sigma_u^2$. $(a_k)_{k=1,\ldots,K}$ and $(b_q)_{q=-\infty,\ldots,+\infty}$ are fixed positive real and complex numbers, respectively, with $\sum_{q=-\infty}^{+\infty} |b_q| < \infty$, $(f_k)_{k=1,\ldots,K}$ are fixed distinct real numbers in $(-1/2, +1/2)$, ϕ_k are random variables uniformly distributed on $[0, 2\pi)$ and $(\phi_k)_{k=1,\ldots,K}$ and u_t are mutually independent.

To extend the notion of AMV estimators [2] (also called asymptotically best consistent (ABC) estimators in [1]) to complex circular processes, two solutions can be considered. First, stacking the real and imaginary parts of the data, existing real asymptotic results can be applied. However this real-valued procedure is often more tedious and/or lacking of engineering insight. Consequently, the second approach that consists in adapting the real-valued procedure to the complex-valued data is considered. We note that because the asymptotic distribution of the second-order statistics is not circular, simply replacing the transpose operator by the conjugate transpose in the existing results has no sense.

To adapt the existing results, two conditions must be satisfied. First, the covariance $\mathbf{C}_R$ of the asymptotic distribution of $\mathbf{R}_{\mathrm{T}}$ must be nonsingular. Second, an arbitrary second-order algorithm considered as a mapping which associates to $\mathbf{R}_{\mathrm{T}}$, the estimate Θ_{T}

$$\mathbf{R}_{\mathrm{T}} \longmapsto \Theta_{\mathrm{T}} = \mathrm{alg}(\mathbf{R}_{\mathrm{T}})$$

must be complex differentiable w.r.t. $\mathbf{R}_{\mathrm{T}}$ at the point $\mathbf{R}(\Theta)$. While, the second condition is satisfied for any mapping alg(.) differentiable w.r.t. the real and imaginary part of the entries of $\mathbf{R}_{\mathrm{T}}$ because $\mathbf{R}_{\mathrm{T}}$ is Hermitian, the first one is not for the following reason. The covariance $\mathbf{C}_R^{\mathrm{to}}$ of the asymptotic distribution of $\mathbf{R}_{\mathrm{T}}^{\mathrm{to}}$ is singular because the set of the M^2 entries of $\mathbf{R}_{\mathrm{T}}^{\mathrm{to}}$ considered as random variables are linearly dependent and consequently $\mathbf{C}_R$ is singular as well because $\mathbf{C}_R = \mathbf{C}_R^{\mathrm{to}}$ as it is proved in [6].

To solve this difficulty, we could work only with the first column $\mathbf{r}_{\mathrm{T}}$ of $\mathbf{R}_{\mathrm{T}}$ because the first column $\mathbf{r}(\Theta)$ of $\mathbf{R}(\Theta)$ is one to one related to $\mathbf{R}(\Theta)$. But this choice leads to an algorithm $\mathbf{r}_{\mathrm{T}} \longmapsto \Theta_{\mathrm{T}} = \mathrm{alg}(\mathbf{r}_{\mathrm{T}})$ that is not differentiable w.r.t. $\mathbf{r}_{\mathrm{T}}$ at the point $\mathbf{r}(\Theta)$. To make this algorithm differentiable, we consider in the following the statistics $\mathbf{s}_{\mathrm{T}}$ equivalent to $\mathbf{r}_{\mathrm{T}}$ constituted by $\mathbf{r}_{\mathrm{T}}$ and $\mathbf{r}_{\mathrm{T}}^*$ where the first common real term $r_{0,\mathrm{T}}$ appears only once, i.e.

$$\mathbf{s}_{\mathrm{T}} \stackrel{\text{def}}{=} \begin{pmatrix} \mathbf{J}\mathbf{r}_{\mathrm{T}}'^* \\ r_{0,\mathrm{T}} \\ \mathbf{r}_{\mathrm{T}}' \end{pmatrix} \quad \text{with } \mathbf{r}_{\mathrm{T}} \stackrel{\text{def}}{=} \begin{pmatrix} r_{0,\mathrm{T}} \\ \mathbf{r}_{\mathrm{T}}' \end{pmatrix},$$

where $\mathbf{J}$ is the reversal permutation matrix of appropriate order (1 in the antidiagonal and 0 elsewhere). So $\mathbf{s}^*(\Theta) = \mathbf{J}\mathbf{s}(\Theta)$ where $\mathbf{s}(\Theta)$ is deduced from $\mathbf{r}(\Theta)$ as $\mathbf{s}_{\mathrm{T}}$ is deduced from $\mathbf{r}_{\mathrm{T}}$. Consequently, all algorithm differentiable w.r.t. $(\Re(\mathbf{s}), \Im(\mathbf{s}))$ or equivalently w.r.t. $(\mathbf{s}, \mathbf{s}^*)$ becomes differentiable w.r.t. $\mathbf{s}$ alone if $\delta\mathbf{s}$ is structured as

$$\delta\mathbf{s} = \begin{pmatrix} \mathbf{J}\delta\mathbf{r}'^* \\ \delta r_0 \\ \delta\mathbf{r}' \end{pmatrix}$$

and

$$\text{alg}[\mathbf{s}(\Theta)+\delta\mathbf{s}] = \text{alg}[\mathbf{s}(\Theta)] + [\mathbf{D}_s, \mathbf{D}_s^*]\begin{bmatrix}\delta\mathbf{s}\\ \delta\mathbf{s}^*\end{bmatrix} + \text{o}(\delta\mathbf{s})$$
$$= \Theta + \mathbf{D}_s^{\text{alg}}\delta\mathbf{s} + \text{o}(\delta\mathbf{s})$$

with $\mathbf{D}_s^{\text{alg}} \stackrel{\text{def}}{=} \mathbf{D}_s + \mathbf{D}_s^*\mathbf{J}$ where $\mathbf{D}_s$ and $\mathbf{D}_s^*$ are the Jacobian matrices w.r.t. $\mathbf{s}$ and $\mathbf{s}^*$ associated with the mapping alg(.) at the point $\mathbf{s}(\Theta)$. And because $\text{alg}[\mathbf{s}(\Theta)] = \Theta$ for all Θ:

$$\text{alg}[\mathbf{s}(\Theta+\delta\Theta)] = \text{alg}[\mathbf{s}(\Theta) + \mathbf{S}\delta\Theta + \text{o}(\delta\Theta)]$$
$$= \Theta + \mathbf{D}_s^{\text{alg}}\mathbf{S}\delta\Theta + \text{o}(\delta\Theta)$$
$$= \Theta + \delta\Theta.$$

Therefore $\mathbf{D}_s^{\text{alg}}$ is a left inverse of $\mathbf{S} \stackrel{\text{def}}{=} \text{d}\mathbf{s}(\Theta)/\text{d}\Theta$:

$$\mathbf{D}_s^{\text{alg}}\mathbf{S} = \mathbf{I}_{\text{L}}, \tag{2.1}$$

and this time, the covariance matrix $\mathbf{C}_s$ of the asymptotic distribution of $\mathbf{s}_{\text{T}}$ is an Hermitian positive definite matrix thanks to its algebraic structure given by the following lemma directly deduced from [6] and [7].

Lemma 1. *The statistics* $\mathbf{s}_{\text{T}}$ *converge in distribution to the complex noncircular Gaussian distribution of covariances* $\mathbf{C}_s$ *and* $\mathbf{C}'_s$:

$$\sqrt{T}(\mathbf{s}_{\text{T}} - \mathbf{s}(\Theta)) \stackrel{\mathscr{L}}{\to} \mathscr{N}_c(\mathbf{0}; \mathbf{C}_s, \mathbf{C}'_s),$$

with[1]

$$\mathbf{C}_s = \int_{-1/2}^{+1/2} S_n^2(f)\mathbf{e}(f)\mathbf{e}^{\text{H}}(f)\,\text{d}f + 2\sum_{k=1}^{K} a_k^2 S_n(f_k)\mathbf{e}(f_k)\mathbf{e}^{\text{H}}(f_k) + \kappa_u\gamma\gamma^{\text{H}} \quad \textit{and} \quad \mathbf{C}'_s = \mathbf{C}_s\mathbf{J}, \tag{2.2}$$

where $\mathbf{e}(f) \stackrel{\text{def}}{=} (\text{e}^{-\text{i}2\pi(M-1)f}, \ldots, \text{e}^{-\text{i}2\pi f}, 1, \text{e}^{\text{i}2\pi f}, \ldots, \text{e}^{\text{i}2\pi(M-1)f})^{\text{H}}$, $\mathbf{e}(f_k) \stackrel{\text{def}}{=} (\text{e}^{-\text{i}2\pi(M-1)f_k}, \ldots, \text{e}^{-\text{i}2\pi f_k}, 1, \text{e}^{\text{i}2\pi f_k}, \ldots, \text{e}^{\text{i}2\pi(M-1)f_k})^{\text{H}}$, $\gamma \stackrel{\text{def}}{=} (\gamma_{M-1}, \ldots, \gamma_1, \gamma_0, \gamma_1^*, \ldots, \gamma_{M-1}^*)^{\text{H}}$ *with* $\gamma_k \stackrel{\text{def}}{=} \text{E}(\mathbf{n}_t\mathbf{n}_{t-k}^*) = \sigma_u^2\sum_{q=-\infty}^{+\infty} b_q b_{q-k}^*$ *and* $S_n(f) \stackrel{\text{def}}{=} \sum_{k=-\infty}^{+\infty}\gamma_k \text{e}^{-\text{i}2\pi kf}$.

$\mathbf{S}$ is a full column matrix from (2.1) and we prove in Appendix A by application of Theorem 2 of [3], extended to the complex circular case:

[1]We note that the noncircular complex Gaussian asymptotic distribution of $\mathbf{s}_{\text{T}}$ is characterized by $\mathbf{C}_s$ only.

Theorem 1. *The asymptotic covariance of an estimator of* Θ *given by an arbitrary consistent second-order algorithm is bounded below by the real symmetric positive definite matrix* $(\mathbf{S}^{\text{H}}\mathbf{C}_s^{-1}\mathbf{S})^{-1}$:

$$\mathbf{C}_\Theta = \mathbf{D}_s^{\text{alg}}\mathbf{C}_s(\mathbf{D}_s^{\text{alg}})^{\text{H}} \geqslant (\mathbf{S}^{\text{H}}\mathbf{C}_s^{-1}\mathbf{S})^{-1}. \tag{2.3}$$

Furthermore, we prove in Appendix A that this lowest bound is asymptotically tight, i.e., there exists an algorithm alg(.) whose covariance of the asymptotic distribution of Θ_{T} satisfies (2.3) with equality. Therefore, Theorem 3 of [3] extends to the complex case.

Theorem 2. *The following nonlinear least square algorithm is an AMV second order algorithm.*

$$\Theta_{\text{T}} = \text{argmin}_\alpha[\mathbf{s}_{\text{T}} - \mathbf{s}(\alpha)]^{\text{H}}\mathbf{C}_s^{-1}[\mathbf{s}_{\text{T}} - \mathbf{s}(\alpha)]. \tag{2.4}$$

In practice, it is difficult to optimize the nonlinear function (2.4) which it involves the computation of $\mathbf{C}_s^{-1}$ that depends on α. Porat and Friedlander proved in [8], in the real case that the lowest bound (2.3) is also obtained if an arbitrary consistent estimate $\mathbf{C}_{s,\text{T}}$ of $\mathbf{C}_s$ is used in (2.4). This property extends to the complex case and to any Hermitian positive definite weighting matrix and we prove in Appendix A.

Theorem 3. *The covariance of the asymptotic distribution of* Θ_{T} *given by an arbitrary nonlinear least square algorithm defined by*

$$\Theta_{\text{T}} = \text{argmin}_\alpha[\mathbf{s}_{\text{T}} - \mathbf{s}(\alpha)]^{\text{H}}\mathbf{W}(\alpha)[\mathbf{s}_{\text{T}} - \mathbf{s}(\alpha)] \tag{2.5}$$

is preserved if the Hermitian positive definite weighting matrix $\mathbf{W}(\alpha)$ *is replaced by an arbitrary consistent estimate* $\mathbf{W}_{\text{T}}$ *that satisfies* $\mathbf{W}_{\text{T}} = \mathbf{W}(\Theta) + \text{O}(\mathbf{s}_{\text{T}} - \mathbf{s}(\Theta))$.

So the minimization (2.4) can be preferably replaced by the following:

$$\Theta_{\text{T}} = \text{argmin}_\alpha[\mathbf{s}_{\text{T}} - \mathbf{s}(\alpha)]^{\text{H}}\mathbf{C}_{s,\text{T}}^{-1}[\mathbf{s}_{\text{T}} - \mathbf{s}(\alpha)]. \tag{2.6}$$

Remark. Naturally, thanks to the one to one mapping $\mathbf{s}_{\text{T}} \leftrightarrow \mathbf{r}_{\text{T}}$, the following ad-hoc nonlinear least square algorithm:

$$\Theta_{\text{T}} = \text{argmin}_\alpha[\mathbf{r}_{\text{T}} - \mathbf{r}(\alpha)]^{\text{H}}\mathbf{C}_r^{-1}(\alpha)[\mathbf{r}_{\text{T}} - \mathbf{r}(\alpha)], \tag{2.7}$$

where $\mathbf{C}_r(\alpha)$ denotes the covariance of the asymptotic distribution of $\mathbf{r}_{\text{T}}$, can be considered as a second-order algorithm that is complex differentiable w.r.t. $\mathbf{s}$ alone. Consequently, its asymptotic covariance matrix satisfies $\mathbf{C}_\Theta \geqslant (\mathbf{S}^{\text{H}}\mathbf{C}_s^{-1}\mathbf{S})^{-1}$. But we will see in Section 4 that $\mathbf{C}_\Theta$ does not attain this lower bound

$(\mathbf{S}^{\mathrm{H}}\mathbf{C}_s^{-1}\mathbf{S})^{-1}$. So this ad-hoc algorithm is no longer an AMV algorithm.

3. Application to the estimation of frequencies for mixed spectra time series

3.1. AMV estimator

In the following, to satisfy the identifiability condition, we consider an MA of order Q process as linear process. In this case, the $M \times M$ covariance matrix of x_t is given by

$$\mathbf{R}(\Theta) = \sum_{k=1}^{K} a_k^2 \tilde{\mathbf{e}}(f_k)\tilde{\mathbf{e}}^{\mathrm{H}}(f_k) + \begin{pmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_Q & 0 & \cdots & 0 \\ \gamma_1^* & \gamma_0 & \cdots & \gamma_{Q-1} & \gamma_Q & \ddots & \cdots \\ \vdots & \ddots & \ddots & \gamma_{Q-2} & \gamma_{Q-1} & \ddots & 0 \\ \gamma_Q^* & \gamma_{Q-1}^* & \ddots & \ddots & \ddots & \ddots & \gamma_Q \\ 0 & \gamma_Q^* & \gamma_{Q-1}^* & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \gamma_1^* & \gamma_0 & \gamma_1 \\ 0 & \cdots & 0 & \gamma_Q^* & \cdots & \gamma_1^* & \gamma_0 \end{pmatrix},$$

where $\tilde{\mathbf{e}}(f_k) \stackrel{\mathrm{def}}{=} (1, \mathrm{e}^{\mathrm{i}2\pi f_k}, \ldots, \mathrm{e}^{\mathrm{i}2(M-1)\pi f_k})^{\mathrm{H}}$. $\mathbf{R}(\Theta)$ is parametrized by the $L = 2(Q+K)+1$ real parameters $\Theta = (\Theta_1^{\mathrm{T}}, \Theta_2^{\mathrm{T}})^{\mathrm{T}}$ with $\Theta_1 \stackrel{\mathrm{def}}{=} (f_1, \ldots, f_K)^{\mathrm{T}}$ and $\Theta_2 \stackrel{\mathrm{def}}{=} [a_1^2, \ldots, a_K^2, \gamma_0, \Re(\gamma_1), \ldots, \Re(\gamma_Q), \Im(\gamma_1), \ldots, \Im(\gamma_Q)]^{\mathrm{T}}$. We note that the first column $\mathbf{r}(\Theta)$ of $\mathbf{R}(\Theta)$ is linear with respect to Θ_2 and consequently $\mathbf{s}(\Theta)$ as well:

$$\mathbf{s}(\Theta) = \mathbf{\Psi}(\Theta_1)\Theta_2, \tag{3.1}$$

where $\mathbf{\Psi}(\Theta_1)$ is the following $(2M-1) \times (K + 2Q + 1)$ matrix:

$$\mathbf{\Psi}(\Theta_1) = \begin{pmatrix} & \mathbf{0}_{M-Q-1} & \mathbf{O}_{M-Q-1,Q} & \mathbf{O}_{M-Q-1,Q} \\ & \mathbf{0}_Q & \mathbf{J}_Q & i\mathbf{J}_Q \\ \mathbf{e}(f_1), \ldots, \mathbf{e}(f_K) & 1 & \mathbf{0}_Q^{\mathrm{T}} & \mathbf{0}_Q^{\mathrm{T}} \\ & \mathbf{0}_Q & \mathbf{I}_Q & -i\mathbf{I}_Q \\ & \mathbf{0}_{M-Q-1} & \mathbf{O}_{M-Q-1,Q} & \mathbf{O}_{M-Q-1,Q} \end{pmatrix}.$$

Therefore $\mathbf{\Psi}(\Theta_1)$ has column full rank (over the field $\mathscr{R}$) if $M - Q - 1 \geqslant K$. This condition is equivalent to having the number of unknown real parameters no larger than the number of estimating equations available, i.e., $2(Q+K)+1 \leqslant 1 + 2(M-1)$. This necessary condition is also sufficient to ensure identifiability because in this case, the vector $\mathbf{r}''(\Theta) \stackrel{\mathrm{def}}{=} (r_{Q+1}, \ldots, r_{M-1})^{\mathrm{T}}$ issued from $\mathbf{r}(\Theta)$ satisfies:

$$\mathbf{r}''(\Theta) = \mathbf{\Psi}'(\Theta_1)\Theta_2',$$

where $\mathbf{\Psi}'(\Theta_1) \stackrel{\mathrm{def}}{=} (\bar{\mathbf{e}}(f_1), \ldots, \bar{\mathbf{e}}(f_K))$ with $\bar{\mathbf{e}}(f_k) \stackrel{\mathrm{def}}{=} (\mathrm{e}^{\mathrm{i}2\pi(Q+1)f_k}, \ldots, \mathrm{e}^{\mathrm{i}2\pi(M-1)f_k})^{\mathrm{H}}$ and $\Theta_2' \stackrel{\mathrm{def}}{=} (a_1^2, \ldots, a_K^2)^{\mathrm{T}}$ and this linear Vandermonde system has an unique solution if its number of columns is less or equal than its number of lines, i.e., $M - Q - 1 \geqslant K$. We suppose in this paper, that this condition is satisfied. The minimization (2.6) with respect to Θ_2 is immediate thanks to (3.1) if Θ_2 is not restricted to be real. With a geometric procedure, we obtain:

$$\widehat{\Theta}_2 = [\mathbf{\Psi}^{\mathrm{H}}(\Theta_1)\mathbf{W}\mathbf{\Psi}(\Theta_1)]^{-1}\mathbf{\Psi}^{\mathrm{H}}(\Theta_1)\mathbf{W}\mathbf{s}_{\mathrm{T}} \tag{3.2}$$

with $\mathbf{W} \stackrel{\mathrm{def}}{=} \mathbf{C}_{s,\mathrm{T}}^{-1}$. With arguments similar to that of COMET [9], we prove in Appendix A that $\widehat{\Theta}_2$ is real-valued for all consistent estimate $\mathbf{C}_{s,\mathrm{T}}$ of $\mathbf{C}_s$ structured as $\mathbf{s}_{\mathrm{T}}\mathbf{s}_{\mathrm{T}}^{\mathrm{H}}$. Thus, $\widehat{\Theta}_2$ given by (3.2) is the real value that minimizes (2.6). $\Theta_{1,\mathrm{T}}$ is obtained by substituting $\widehat{\Theta}_2$ in (2.5):

$$\Theta_{1,\mathrm{T}} = \operatorname{argmax}_{\alpha_1} V(\alpha_1) \tag{3.3}$$

with

$$V(\alpha_1) \stackrel{\mathrm{def}}{=} \mathbf{s}_{\mathrm{T}}^{\mathrm{H}}\mathbf{W}\mathbf{\Psi}(\alpha_1)[\mathbf{\Psi}^{\mathrm{H}}(\alpha_1)\mathbf{W}\mathbf{\Psi}(\alpha_1)]^{-1}\mathbf{\Psi}^{\mathrm{H}}(\alpha_1)\mathbf{W}\mathbf{s}_{\mathrm{T}}.$$

3.2. Performance analysis

By application of Theorem 1, the covariance of the asymptotic distribution of the minimum variance second-order frequency estimator (3.3) is given by the top left $K \times K$ "frequency corner" of $(\mathbf{S}^{\mathrm{H}}\mathbf{C}_s^{-1}\mathbf{S})^{-1}$ where $\mathbf{C}_s$ is given by (2.2). If we note that $\mathbf{S} = [\mathbf{S}_1, \mathbf{\Psi}]$ with $\mathbf{S}_1 \stackrel{\mathrm{def}}{=} \partial\mathbf{s}/\partial\Theta_1$, the matrix inversion lemma gives

$$\begin{aligned}\mathbf{C}_{\Theta_1} &= (\mathbf{S}_1^{\mathrm{H}}\mathbf{C}_s^{-1}\mathbf{S}_1 - \mathbf{S}_1^{\mathrm{H}}\mathbf{C}_s^{-1}\mathbf{\Psi}[\mathbf{\Psi}^{\mathrm{H}}\mathbf{C}_s^{-1}\mathbf{\Psi}]^{-1}\mathbf{\Psi}^{\mathrm{H}}\mathbf{C}_s^{-1}\mathbf{S}_1)^{-1} \\ &= (\mathbf{S}_1^{\mathrm{H}}\mathbf{C}_s^{-1/2}\mathbf{P}^{\perp}_{\mathbf{C}_s^{-1/2}\mathbf{\Psi}}\mathbf{C}_s^{-1/2}\mathbf{S}_1)^{-1},\end{aligned} \tag{3.4}$$

where $\mathbf{P}^{\perp}_{\mathbf{C}_s^{-1/2}\mathbf{\Psi}}$ denotes the projector onto the orthocomplement of the columns of $\mathbf{C}_s^{-1/2}\mathbf{\Psi}$.

Remark. In the case of non-Gaussian additive noise, an AMV second-order algorithm devised under the Gaussian assumption (i.e. with a weighting matrix associated with the Gaussian case) is no longer an AMV second-order estimator. But in this case, the asymptotic covariance of such an algorithm is insensitive to the distribution of the additive noise n_t thanks to a result proved in [6] and is given by (3.4) where $\mathbf{C}_s$ is associated with the Gaussian case.

4. Illustrative examples

In this section we illustrate the loss of performance of several suboptimal least square algorithms compared to the AMV algorithm and to the MPD estimator of [5] which is to the best of our knowledge the only second-order algorithm specifically devoted to MA noise. More precisely, we consider the following algorithms:

- The first one is the ad-hoc algorithm (denoted AMV_r) (2.7) obtained by considering the statistics $\mathbf{r}_\text{T}$ only.
- The second ones (denoted AMV_s^g and AMV_r^g) are deduced from the AMV (2.6) and AMV_r (2.7) estimator when the weighting matrix $\mathbf{C}_{s,\text{T}}^{-1}$ and $\mathbf{C}_{r,\text{T}}^{-1}$ are, respectively, replaced by consistent estimates $\mathscr{C}_{s,\text{T}}^{-1}$ and $\mathscr{C}_{r,\text{T}}^{-1}$ of the inverse of the asymptotic covariance matrices $\mathbf{C}_s$ and $\mathbf{C}_r$ associated with the asymptotic covariance matrix $\mathscr{C}_r = r_0(\Theta)\mathbf{R}(\Theta)$ given by the erroneous signal model of independent Gaussian complex circular observations x_t for which:

$$\mathscr{C}_{s,\text{T}} = \begin{pmatrix} r_{0,\text{T}}\mathbf{J}\mathbf{R}'^{*}_{\text{T}}\mathbf{J} & r_{0,\text{T}}\mathbf{J}\mathbf{r}'^{*}_{\text{T}} & \mathbf{J}\mathbf{r}'^{*}_{\text{T}}\mathbf{r}'^{\text{H}}_{\text{T}} \\ r_{0,\text{T}}\mathbf{r}'^{\text{T}}_{\text{T}}\mathbf{J} & r^2_{0,\text{T}} & r_{0,\text{T}}\mathbf{r}'^{\text{H}}_{\text{T}} \\ \mathbf{r}'_{\text{T}}\mathbf{r}'^{\text{T}}_{\text{T}}\mathbf{J} & r_{0,\text{T}}\mathbf{r}'_{\text{T}} & r_{0,\text{T}}\mathbf{R}'_{\text{T}} \end{pmatrix},$$

where

$$\mathbf{R}_\text{T} \stackrel{\text{def}}{=} \begin{pmatrix} r_{0,\text{T}} & \mathbf{r}'^{\text{H}}_{\text{T}} \\ \mathbf{r}'_{\text{T}} & \mathbf{R}'_{\text{T}} \end{pmatrix}.$$

- The third ones (denoted LS_s and LS_r) are the unweighted least square algorithms $\Theta_\text{T} = \arg\min_\alpha \|\mathbf{s}_\text{T} - \mathbf{s}(\alpha)\|^2$ and $\Theta_\text{T} = \arg\min_\alpha \|\mathbf{r}_\text{T} - \mathbf{r}(\alpha)\|^2$.
- The fourth ones (denoted $\text{AMV}_{s'}$ and $\text{AMV}_{r'}$) are deduced from the AMV and AMV_r algorithms by eliminating the r_0 term. For these algorithms, Θ_2 does not contain the γ_0 term.
- The last ones (denoted MDP)[2] is the MPD algorithm for which the asymptotic variance is given by $\text{Var}[\theta_{1,\text{T}}] = (1/\text{T})\mathbf{D}_r^{\text{MPD}}\mathbf{C}_r(\mathbf{D}_r^{\text{MPD}})^{\text{H}}_{(1,1)}$, where the Jacobian $\mathbf{D}_r^{\text{MPD}}$ is given in Appendix B.

[2]We note, that contrary to the AMV approach for which $M \geqslant Q + K + 1$, the MPD estimator, based on an Hankel matrix built from the "zero triangles" of $\mathbf{R}(\Theta)$, requires the fixed order $M = Q + 2K + 2$.

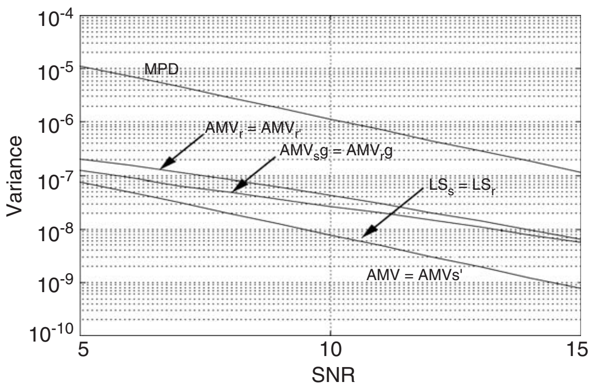

Fig. 1. Theoretical variance $\text{Var}[\theta_{1,\text{T}}]$ given by the AMV and suboptimal estimators versus SNR.

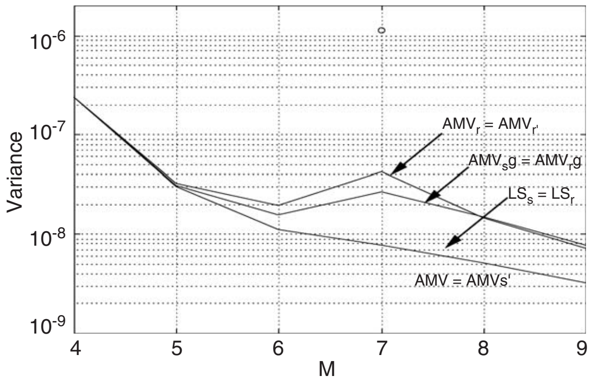

Fig. 2. Theoretical variance $\text{Var}[\theta_{1,\text{T}}]$ given by the AMV and suboptimal estimators versus M.

We show the case $Q = 1$ ($b_0 = b_1 = 0.707$), $K = 2$ (equipowered complex cisoids with $f_1 = -0.1$, $f_1 = 0.25$) and $T = 2000$, $M = 7$ for different SNRs in Fig. 1 and for SNR = 15 dB and different values of M in Fig. 2. We see that these algorithms are actually suboptimal except for $M = Q + K + 1$[3] for which they have the asymptotic performances of the AMV algorithm and the ad-hoc algorithm (denoted AMV_r) (2.7) obtained by using the conjugate transpose instead of transpose in the real-valued associated AMV estimator is outperformed by the AMV estimator. We see that the asymptotic variance $\text{Var}[\theta_{1,\text{T}}]$ given by the algorithms LS_s and LS_r, AMV_r and $\text{AMV}_{r'}$, AMV and $\text{AMV}_{s'}$ coincide, respectively. This later result generalizes a result proved in [10] for a white noise.

[3]We note that this property has been confirmed for several values of Q and K, but we have not succeeded in proving it analytically.

5. Conclusion

This paper has extended the notion of AMV second-order estimation devoted to parameters of real stationary processes to complex circular stationary processes. It has been shown that a special attention to the statistics involved is required. In particular it is proved that the AMV estimator is not a direct extension of the real-valued associated AMV estimator using the conjugate transpose instead of transpose. As an application, the estimation of the frequencies of cisoids for mixed spectra time series containing a sum of cisoids and an MA process is considered.

Appendix A. Proof of theorems

Proof of Theorem 1. Using the proof of Theorem 2 of [3] (replacing the superscript T by H), it is sufficient to prove that the Hermitian matrix $(\mathbf{S}^{\mathrm{H}}\mathbf{C}_s^{-1}\mathbf{S})^{-1}$ is real symmetric. Because $\mathbf{s}^* = \mathbf{J}\mathbf{s}$ implies $\mathbf{S}^* = \mathbf{J}\mathbf{S}$ and $\mathbf{C}_s^{\mathrm{T}} = \mathbf{J}\mathbf{C}_s\mathbf{J}$, we have $(\mathbf{S}^{\mathrm{H}}\mathbf{C}_s^{-1}\mathbf{S})^{\mathrm{T}} = \mathbf{S}^{\mathrm{T}}(\mathbf{C}_s^{-1})^{\mathrm{T}}\mathbf{S}^* = \mathbf{S}^{\mathrm{H}}\mathbf{J}(\mathbf{C}_s^{\mathrm{T}})^{-1}\mathbf{J}\mathbf{S} = \mathbf{S}^{\mathrm{H}}\mathbf{C}_s^{-1}\mathbf{S}$. □

Proof of Theorem 2. By a perturbation analysis, $\Theta_{\mathrm{T}} = \Theta + \delta\Theta$ is associated with $\mathbf{s}_{\mathrm{T}} = \mathbf{s}(\Theta) + \delta\mathbf{s}$ (with $\delta\mathbf{s}$ structured). Because $V_{\mathrm{T}}(\alpha) \stackrel{\mathrm{def}}{=} [\mathbf{s}_{\mathrm{T}} - \mathbf{s}(\alpha)]^{\mathrm{H}}\mathbf{C}_s^{-1}(\alpha)[\mathbf{s}_{\mathrm{T}} - \mathbf{s}(\alpha)]$ is minimum for $\alpha = \Theta_{\mathrm{T}}$, we have: $(V_{\mathrm{T}}(\alpha)/\mathrm{d}\alpha)_{|\alpha=\Theta+\delta\Theta} = \mathbf{0}$. Expanding this derivative, we straightforwardly obtain: $(\mathbf{S}^{\mathrm{H}}\mathbf{C}_s^{-1}\mathbf{S} + \mathbf{S}^{\mathrm{T}}\mathbf{C}_s^{-1*}\mathbf{S}^*)\delta\Theta + \mathrm{o}(\delta\Theta) = \mathbf{S}^{\mathrm{H}}\mathbf{C}_s^{-1}\delta\mathbf{s} + \mathbf{S}^{\mathrm{T}}\mathbf{C}_s^{-1*}\delta\mathbf{s}^* + \mathrm{o}(\delta\mathbf{s})$. Consequently, the algorithm (2.4) satisfies:

$$\begin{aligned}
&\mathrm{alg}[\mathbf{s}(\Theta) + \delta\mathbf{s}] \\
&\quad = \Theta + (\mathbf{S}^{\mathrm{H}}\mathbf{C}_s^{-1}\mathbf{S} + \mathbf{S}^{\mathrm{T}}\mathbf{C}_s^{-1*}\mathbf{S}^*)^{-1} \\
&\qquad \times (\mathbf{S}^{\mathrm{H}}\mathbf{C}_s^{-1}, \mathbf{S}^{\mathrm{T}}\mathbf{C}_s^{-1*}) \begin{pmatrix} \delta\mathbf{s} \\ \delta\mathbf{s}^* \end{pmatrix} + \mathrm{o}(\delta\mathbf{s}) \\
&\quad = \Theta + (\mathbf{S}^{\mathrm{H}}\mathbf{C}_s^{-1}\mathbf{S})^{-1}\mathbf{S}^{\mathrm{H}}\mathbf{C}_s^{-1}\delta\mathbf{s} + \mathrm{o}(\delta\mathbf{s}), \qquad \text{(A.1)}
\end{aligned}$$

by using $\mathbf{S}^* = \mathbf{J}\mathbf{S}$ and $\mathbf{C}_s^{\mathrm{T}}(\Theta) = \mathbf{C}_s^* = \mathbf{J}\mathbf{C}_s\mathbf{J}$ in the second equality. □

Proof of Theorem 3. Following a perturbation analysis similar to that of the proof of Theorem 2, it is straightforward to show that the differential $\mathbf{D}_s^{\mathrm{alg}} = (\mathbf{S}^{\mathrm{H}}\mathbf{W}\mathbf{S})^{-1}\mathbf{S}^{\mathrm{H}}\mathbf{W}$ of this algorithm is preserved. □

Proof of the real value of $\widehat{\Theta}_2$. If

$$\mathbf{K} \stackrel{\mathrm{def}}{=} \frac{1}{2}\begin{pmatrix} \mathbf{J}_{M-1} & \mathbf{0} & \mathbf{I}_{M-1} \\ \mathbf{0}^{\mathrm{T}} & 2 & \mathbf{0}^{\mathrm{T}} \\ i\mathbf{J}_{M-1} & \mathbf{0} & -i\mathbf{I}_{M-1} \end{pmatrix}$$

denotes the linear invertible transformation that associates to $\mathbf{s}_{\mathrm{T}}$, the real-valued vector $\boldsymbol{\eta}_{\mathrm{T}}$ comprised of the real and imaginary parts of $\mathbf{s}_{\mathrm{T}}$, $\boldsymbol{\eta}_{\mathrm{T}} = \mathbf{K}\mathbf{s}_{\mathrm{T}}$ and $\widehat{\Theta}_2$ given by (3.2) assumes the form: $[(\mathbf{K}\boldsymbol{\Psi})^{\mathrm{H}}(\mathbf{K}\mathbf{W}^{-1}\mathbf{K}^{\mathrm{H}})^{-1}(\mathbf{K}\boldsymbol{\Psi})]^{-1}(\mathbf{K}\boldsymbol{\Psi})^{\mathrm{H}}(\mathbf{K}\mathbf{W}^{-1}\mathbf{K}^{\mathrm{H}})^{-1}\mathbf{K}\mathbf{s}_{\mathrm{T}}$, where $\mathbf{K}\mathbf{s}_{\mathrm{T}}$ is real and so is $\mathbf{K}\boldsymbol{\Psi}$. It remains to examine $\mathbf{K}\mathbf{W}^{-1}\mathbf{K}^{\mathrm{H}}$. Because $\mathbf{K}\mathbf{s}_{\mathrm{T}}\mathbf{s}_{\mathrm{T}}^{\mathrm{H}}\mathbf{K}^{\mathrm{H}} = \boldsymbol{\eta}_{\mathrm{T}}\boldsymbol{\eta}_{\mathrm{T}}^{\mathrm{H}}$ is real-valued, the matrix $\mathbf{K}\mathbf{W}^{-1}\mathbf{K}^{\mathrm{H}} = \mathbf{K}\mathbf{C}_{s,\mathrm{T}}\mathbf{K}^{\mathrm{H}}$ is real-valued. □

Appendix B. MPD algorithm

MPD principles: The MPD algorithm is based on the Hankel matrix of order $K+1$

$$\mathbf{H}(\Theta) = \begin{pmatrix} r_{Q+1} & r_{Q+2} & \cdots & r_{Q+K+1} \\ r_{Q+2} & r_{Q+3} & \cdots & r_{Q+K+2} \\ \vdots & & & \vdots \\ r_{Q+K+1} & r_{Q+K+2} & \cdots & r_{Q+2K+1} \end{pmatrix}$$

built from the "zero triangles" (i.e., the part excluding the terms $\gamma_0, \ldots, \gamma_Q$) of $\mathbf{R}(\Theta)$ of order $M = Q + 2K + 2$ and it relies on the following lemma proved in [5]:

Lemma 2. $\{\mathrm{e}^{\mathrm{i}2\pi f_1}, \ldots, \mathrm{e}^{\mathrm{i}2\pi f_K}\}$ *are the roots of the polynomial* $\sum_{l=0}^{K} c_l z^l$ *if and only if* $\mathbf{H}(\Theta)\mathbf{c} = \mathbf{0}$ *with* $\mathbf{c} \stackrel{\mathrm{def}}{=} (c_0, \ldots, c_K)^{\mathrm{T}}$.

And the MPD algorithm is an extension of the mapping:

$$\mathrm{Vec}(\mathbf{R}(\Theta)) \stackrel{\mathrm{g}_1}{\longmapsto} \mathbf{H}(\Theta) \stackrel{\mathrm{g}_2}{\longmapsto} (c_0, \ldots, c_K)^{\mathrm{T}} \stackrel{\mathrm{g}_3}{\longmapsto} (\mathrm{e}^{\mathrm{i}2\pi f_1}, \ldots, \mathrm{e}^{\mathrm{i}2\pi f_K})^{\mathrm{T}} \stackrel{\mathrm{g}_4}{\longmapsto} (f_1, \ldots, f_K)^{\mathrm{T}}$$

generated by the unstructured matrix $\mathbf{R}_{\mathrm{T}}$.

The Jacobian of the MPD algorithm: An extension of g_1 is naturally obtained thanks to the Toeplization of $\mathbf{R}_{\mathrm{T}}$. So g_1 is a linear operator whose associated matrix is $\mathbf{D}_1$. An extension of g_2 can be obtained, thanks to the right singular vector of $\mathbf{H}_{\mathrm{T}}$ associated with its smallest singular value. The derivative $\mathbf{D}_2$ of this mapping is derived from [11, Theorem 8, rel. 4, p. 162]. Finally, derivatives $\mathbf{D}_3$ and $\mathbf{D}_4$ of the mappings g_3 (rooting of polynomial $C(z) = \sum_{l=0}^{K} c_l z^l$) and g_4

$((z_1, \ldots, z_K)^T \longmapsto ((1/2\pi)\arg(z_1), \ldots, (1/2\pi)\arg(z_K))^T)$ are classically derived from standard perturbation calculus. Applying the chain differential rule, the Jacobian of the MPD algorithm is $\mathbf{D}_r^{\mathrm{MPD}} = \mathbf{D}_4\mathbf{D}_3\mathbf{D}_2\mathbf{D}_1$.

References

[1] P. Stoica, B. Friedlander, T. Söderström, An approximate maximum approach to ARMA spectral estimation, in: Proceedings of the Decision and Control, Fort Lauderdale, 1985.

[2] B. Porat, B. Friedlander, Asymptotic accuracy of ARMA parameter estimation methods based on sample covariances, in: Proceedings of the Seventh IFAC/IFORS Symposium on Identification and System Parameter Estimation, New York, 1985.

[3] B. Porat, B. Friedlander, Performance analysis of parameter estimation algorithms based on high-order moments, Internat. J. Adaptive Control and Signal Process. 3 (1989) 191–229.

[4] G.B. Giannakis, S.D. Halford, Performance analysis of blind equalizers based on cyclostationarity statistics, in: Proceedings of the 26th Conference on Information Science System, Princeton, March 1994, pp. 711–716.

[5] P.J. Sherman, A.E. Frazho, High resolution spectral estimation of sinusoids in colored noise using a modified Pisarenko decomposition, in: Proceedings of the ICASSP, Tokyo, April 1986, 181–184.

[6] J.P. Delmas, Asymptotic normality of sample covariance matrix for mixed spectra time series: application to sinusoidal frequencies estimation, IEEE Trans. Inform. Theory 47 (4) (May 2001) 1681–1687.

[7] T.-H. Li, B. Kedem, S. Yakowitz, Asymptotic normality of sample autocovariances with an application in frequency estimation, Stochast. Process. Appl. (52) (1994) 329–349.

[8] B. Friedlander, B. Porat, Asymptotically optimal estimation of MA and ARMA parameters of non-Gaussian processes from high-order moments, IEEE Trans. Automat. Control 35 (1) (January 1990) 27–35.

[9] B. Ottersten, P. Stoica, R. Roy, Covariance matching estimation techniques for array signal processing, Digital Signal Process. 8 (1998) 185–210.

[10] P. Stoica, P. Händel, T. Söderström, Approximate maximum likelihood frequency estimation, Automatica 30 (1) (1994) (131-145-1847).

[11] J.R. Magnus, H. Neudecker, Matrix differential calculus with applications in statistics and econometrics, Wiley Series in Probability and Statistics, 1999 (revised edition).