IMPROVED SUBSPACE DOA ESTIMATION METHODS WITH LARGE ARRAYS: THE DETERMINISTIC SIGNALS CASE

P. Vallet, P. Loubaton

Université Paris-Est
IGM LabInfo, UMR-CNRS 8049
5 Bd. Descartes, Champs sur Marne
77454 Marne la Vallée Cedex 2, France
e-mail:{vallet,loubaton}@univ-mlv.fr

X. Mestre

Centre Tecnològic de Telecomunicacions de Catalunya
Av. del Canal Olímpic, s/n, PMT
08860 Castelldefels, Barcelona, Spain
e-mail: xavier.mestre@cttc.cat

ABSTRACT

This paper is devoted to the subspace DoA estimation using a large antennas array when the number of available snapshots is of the same order of magnitude than the number of sensors. In this context, the traditional subspace methods fail because the empirical covariance matrix of the observations is a poor estimate of the true covariance matrix. Mestre et al. proposed recently to study the behaviour of the traditional estimators when the number of antennas M and the number of snapshots N converge to $+\infty$ at the same rate. Using large random matrix theory results, they showed that the traditional subspace estimate is not consistent in the above asymptotic regime and they proposed a new consistent subspace estimate which outperforms the standard subspace method for realistic values of M and N. However, the work of Mestre et al. assumes that the source signals are independent and identically distributed in the time domain. The goal of the present paper is to propose new consistent estimators of the DoAs in the case where the source signals are modelled as unknown deterministic signals. This, in practice, allows to use the proposed approach whatever the statistical properties of the source signals are.

***Index Terms*—** DoA, Large Random Matrix Theory, MUSIC

1. INTRODUCTION

Subspace DoA estimation methods using antenna arrays (such as MUSIC) have been extensively studied in the past because they offer good complexity versus performance trade-off. Their statistical performance have been mainly characterized in the case where the number of snapshots N converge to $+\infty$ while the number of antennas M remains fixed. In practice, the corresponding conclusions are valid in finite sample size if N is much greater than M. However, this assumption is often not realistic if the number of antennas is ”large” because in practice, the number of available snapshots is limited. In order to study the statistical performance of the subspace estimates in this context, Mestre et al [4] proposed to consider the asymptotic regime in which M and N converge to $+\infty$ at the same rate, i.e. $M, N \to +\infty, c = \frac{M}{N}$ converges towards a strictly positive constant. Using Large Random Matrix Theory (LRMT) results, [4] proved that the traditional DoAs subspace estimators are asymptotically biased, and proposed consistent estimators which outperform the standard ones, for realistic values of M and N. The work [4] however assumes that the various source signals may only be correlated in the spatial domain, so that in the time domain they are assumed to be independent identically distributed (i.i.d.) sequences. Hence, when the source signals are correlated in the time domain, the various equations used to predict the behaviour of uncorrelated random matrices are not valid. Therefore, the approach developed in [4] do not provide a consistent estimator. The purpose of this paper is to propose consistent subspace estimators when the source signals are modelled as non observable deterministic sequences. In practice, this context is relevant whatever the properties of the source signals because realizations of any kind of stochastic processes can be seen as deterministic sequences. The present approach is again based on LRMT results, but in contrast with [4], the observation is modelled as a noisy non zero mean random matrix, a model recently introduced in [2] and referred to as the ”Information plus Noise model”.

This paper is structured as follows. In section 2, we present the signal model and the addressed problem. In section 3, we provide some background material on the asymptotic eigenvalue distribution of the empirical covariance matrix. In section 4, we evaluate the asymptotic behaviour of the standard subspace estimate, and propose a new consistent estimate. In section 5, numerical results illustrate the performance of our new approach.

2. THE ADDRESSED PROBLEM

We assume that K narrow band deterministic source signals $(s_k)_{k=1,\ldots,K}$ are received by an antenna array of M elements, $K < M$. We assume for simplicity that the array is linear with equispaced antennas. The corresponding M dimensional observation signal $\mathbf{y}_n$ (at discrete time n) is supposed to be given by

$$\mathbf{y}_n = \mathbf{A}\mathbf{s}_n + \mathbf{v}_n \tag{1}$$

where $\mathbf{A} = (\mathbf{a}(\theta_1), \ldots, \mathbf{a}(\theta_K))$ is the matrix that contains the steering vectors of the K sources and where $\mathbf{v}_n$ is an additive white noise with covariance matrix $\mathbb{E}(\mathbf{v}_n\mathbf{v}_n^H) = \sigma^2\mathbf{I}_M$. $\mathbf{s}_n$ is defined by $\mathbf{s}_n = (s_{1,n}, \ldots, s_{K,n})^T$. We assume that $\mathbf{y}_n$ is available from $n = 1$ to $n = N$, and that $M < N$, or equivalently that $c = \frac{M}{N}$ is strictly less than 1. We note that it is possible to generalize our results in case where $c > 1$; the presentation of the corresponding results would however complicate the present paper. We denote by $\mathbf{Y} = (\mathbf{y}_1, \ldots, \mathbf{y}_N)$ the observed matrix which can be written as

$$\mathbf{Y} = \mathbf{AS} + \mathbf{V} \tag{2}$$

where $\mathbf{S}$ and $\mathbf{V}$ are defined as $\mathbf{Y}$. We denote by $\mathbf{\Pi}$ the orthogonal projection matrix on the ”noise subspace”, which in our context is defined as the orthogonal complement of the column space of matrix $\mathbf{A}$. In the following, we assume that the empirical covariance matrix of $\mathbf{S}$ defined by $\frac{1}{N}\mathbf{S}\mathbf{S}^H$ is full rank. Therefore, the noise subspace

978-1-4244-2354-5/09/$25.00 ©2009 IEEE

coincides with the kernel of the covariance matrix $\mathbf{R}$ defined as

$$\mathbf{R} = \frac{1}{N}\mathbf{A}\mathbf{S}\mathbf{S}^H\mathbf{A}^H \tag{3}$$

We denote by $(\lambda_k)_{k=1,\ldots,M}$ the eigenvalues of matrix $\mathbf{R}$ arranged in increasing order and by $(\mathbf{e}_k)_{k=1,\ldots,M}$ the corresponding unit norm eigenvectors. We note in particular that $\lambda_1 = \ldots = \lambda_{M-K} = 0$ while the remaining eigenvalues are strictly positive and that $\mathbf{\Pi} = \sum_{k=1}^{M-K} \mathbf{e}_k\mathbf{e}_k^H$. The subspace method is based on the observation that the angles $(\theta_k)_{k=1,\ldots,K}$ coincide with the K solutions of the equation $\mathbf{a}(\theta)^H\mathbf{\Pi}\mathbf{a}(\theta) = 0$. In order to be able to use this last observation, it is in practice necessary to estimate the function $\mathbf{a}(\theta)^H\mathbf{\Pi}\mathbf{a}(\theta)$ (called "localization function") for each $\theta \in [-\pi, \pi]$, or more generically to estimate

$$\eta = \mathbf{b}^H\mathbf{\Pi}\mathbf{b} \tag{4}$$

for each deterministic M-dimensional vector $\mathbf{b}$. If $N \to +\infty$ while M is fixed, the empirical covariance matrix of the observations $\hat{\mathbf{R}}$ of $\mathbf{Y}$ defined by

$$\hat{\mathbf{R}} = \frac{1}{N}\mathbf{Y}\mathbf{Y}^H \tag{5}$$

converges towards the matrix $\mathbf{R} + \sigma^2\mathbf{I}_M$ in the sense that

$$\|\hat{\mathbf{R}} - (\mathbf{R} + \sigma^2\mathbf{I})\|_{s.p} \to 0 \quad a.s \tag{6}$$

where $\|\|_{s.p}$ represents the spectral norm and a.s the almost sure convergence. We denote by $(\hat{\lambda}_k)_{k=1,\ldots,M}$ the eigenvalues of $\hat{\mathbf{R}}$ arranged in increasing order and by $(\hat{\mathbf{e}}_k)_{k=1,\ldots,M}$ the corresponding eigenvectors. (6) implies that $\hat{\eta}_{trad} - \eta \to 0 \ \ a.s$ where $\hat{\eta}_{trad}$ is the traditional estimator of the localization function defined by

$$\hat{\eta}_{trad} = \sum_{k=1}^{M-K} \mathbf{b}^H\hat{\mathbf{e}}_k\hat{\mathbf{e}}_k^H\mathbf{b} \tag{7}$$

However, relation (6) does not hold in the asymptotic regime $M, N \to +\infty$ in such a way that $c = \frac{M}{N}$ converges towards a non zero constant. In particular, it is shown in section 4 that $\hat{\eta}_{trad} - \eta$ does not converge to 0.

3. THE ASYMPTOTIC EIGENVALUE DISTRIBUTION OF MATRIX $\hat{\mathbf{R}}$

In this section, we review certain results related to the behaviour of the eigenvalue distribution of matrix $\hat{\mathbf{R}}$ when $M, N \to +\infty$ in such a way that $c = \frac{M}{N}$ converges towards a non zero constant.

The eigenvalue distribution of $\hat{\mathbf{R}}$ is characterized by its distribution function $\hat{F}(\lambda) = \frac{1}{M}\mathrm{card}\{\hat{\lambda}_k : \hat{\lambda}_k \leq \lambda,\ k = 1,\ldots,M\}$ where card denotes the cardinality of a set. $\hat{F}(\lambda)$ represents the proportion of the eigenvalues of $\hat{\mathbf{R}}$ which are lower than or equal to λ and its associated probability measure $d\hat{F}(\lambda)$ is $\frac{1}{M}\sum_{k=1}^{M}\delta(\lambda - \hat{\lambda}_k)$.

We first review some results which follow immediately from [2] and [3]. The distribution function $\hat{F}(\lambda)$ is clearly random and under additional technical assumptions, there exists a deterministic distribution function $F(\lambda)$ such that, $\hat{F}(\lambda) - F(\lambda) \to 0$ a.s $\forall\, \lambda$ (see [2]). The probability measure associated to F, denoted $dF(\lambda)$, is called the asymptotic eigenvalue distribution of the matrix $\hat{\mathbf{R}}$, and its characterization allows to obtain useful informations on the behaviour of the $(\hat{\lambda}_k)_{k=1,\ldots,M}$. The support of $dF(\lambda)$ is a compact subset $\mathcal{S}$ of $\mathbb{R}^+$, $dF(\lambda)$ is absolutely continuous, and its density is continuous on $\mathcal{S}$ and differentiable on the interior $\mathring{\mathcal{S}}$ of $\mathcal{S}$. The measure $dF(\lambda)$ is characterized by its Stieltjes transform $m(z)$ defined for each $z \in \mathbb{C} - \mathcal{S}$ by

$$m(z) = \int_{\mathcal{S}} \frac{1}{\lambda - z}\, dF(\lambda) \tag{8}$$

We note that $m(z)$ is holomorphic on $\mathbb{C} - \mathcal{S}$. We denote by $f(w)$ the function defined by $f(w) = \frac{1}{N}\mathrm{Trace}(\mathbf{R} - w\mathbf{I})^{-1}$, and we consider the following equation w.r.t. m :

$$\frac{cm}{1 + \sigma^2 cm} = f\left(z(1 + \sigma^2 cm)^2 - \sigma^2(1 - c)(1 + \sigma^2 cm)\right) \tag{9}$$

Consider the set $\mathbb{C}^+ = \{z \in \mathbb{C} : \mathrm{Im}(z) > 0\}$. For each $z \in \mathbb{C}^+$, $m(z)$ is the unique solution of equation (9) for which $\mathrm{Im}(m(z)) > 0$ and $\mathrm{Im}(zm(z)) > 0$. $m(z)$ also satisfies (9) on $z \in \mathbb{C} - \mathcal{S}$, and is real valued on $\mathbb{R} - \mathcal{S}$. Moreover, for each $x \in \mathcal{S}$ the limit $\lim_{y \to 0^+} m(x + iy)$ exists, and is still denoted $m(x)$ to simplify the notations. Finally, $m(x)$ satisfies the equation (9) on $\mathring{\mathcal{S}}$, $\mathrm{Im}(m(x)) > 0$ on $\mathring{\mathcal{S}}$, and the function $x \to \frac{1}{\pi}\mathrm{Im}(m(x))$ coincides with the density of measure $dF(x)$.

We now present a characterization of $\mathcal{S}$ which is more explicit than the analysis provided in [3]. The proof is omitted due to the lack of space.

Theorem 1 *We recall that we are in the case where $c < 1$. Let $\phi(w)$ be the function defined on $\mathbb{R} - \{\lambda_1, \ldots, \lambda_M\}$ by*

$$\phi(w) = w(1 - \sigma^2 f(w))^2 + (1 - c)\sigma^2(1 - \sigma^2 f(w)) \tag{10}$$

The number of local maxima of ϕ satisfying

$$1 - \sigma^2 f(w) > 0 \ \text{ and } \ \phi(w) > 0 \tag{11}$$

is an even number $2Q$.
These local maxima are denoted by $\{w_q^-, w_q^+\}_{q=1,\ldots,Q}$ and they satisfy

$$w_1^- < 0 < w_1^+ \leq w_2^- < w_2^+ \leq \ldots \leq w_Q^- < w_Q^+ \tag{12}$$

If we denote by $x_q^- = \phi(w_q^-)$ and $x_q^+ = \phi(w_q^+)$ the values taken by ϕ at the local maxima, then,

$$0 < x_1^- < x_1^+ \leq x_2^- < x_2^+ \leq \ldots \leq x_Q^- < x_Q^+ \tag{13}$$

Moreover, $\mathcal{S}$ is the reunion of Q compact intervals called "clusters"

$$\mathcal{S} = \bigcup_{q=1}^{Q} [x_q^-, x_q^+] \tag{14}$$

Finally, each eigenvalue of $\mathbf{R}$ belongs to one of the intervals (w_q^-, w_q^+).

In order to have a better understanding of this result, we consider the case where matrix $\mathbf{R}$ has a finite number m of distinct eigenvalues denoted $\overline{\lambda}_1, \ldots, \overline{\lambda}_m$ which remain fixed when M and N increase. In other words, the eigenvalue distribution of matrix $\mathbf{R}$ converges towards a Dirac measure concentrated at points $(\overline{\lambda}_j)_{j=1,\ldots,m}$. We note that $\overline{\lambda}_1 = 0$. If c is close to 0, matrix $\hat{\mathbf{R}}$ tends to be very close to $\mathbf{R} + \sigma^2\mathbf{I}$. Therefore, the asymptotic eigenvalue distribution dF is itself close from a Dirac measure at points $(\overline{\lambda}_j + \sigma^2)_{j=1,\ldots,m}$. For each $q = 1, \ldots, m$, $[x_q^-, x_q^+]$ and $[w_q^-, w_q^+]$ are small width intervals containing $\overline{\lambda}_q + \sigma^2$ and $\overline{\lambda}_q$ respectively. If c increases, the width of the various intervals tend to increase, so that some of the intervals may merge.

We now introduce a useful definition.

Definition 1 *We say that the eigenvalue λ_k $(k = 1, \ldots, M)$ of the matrix $\mathbf{R}$ is associated with the cluster $\left[x_q^-, x_q^+\right]$ if $\lambda_k \in \left[w_q^-, w_q^+\right]$. Moreover, we say that an eigenvalue λ_k is separated from another eigenvalue λ_n if they are not associated with the same cluster.*

The first cluster $[x_1^-, x_1^+]$ plays a special role because we see from Theorem 1 that it is associated with the eigenvalue 0 of matrix $\mathbf{R}$. Under a certain condition we omit, the eigenvalue 0 is the only one to be associated with the first cluster (i.e 0 is separated from the other eigenvalues of $\mathbf{R}$). We now state a conjecture which is necessary for the validity of the next section's results.

Conjecture 1 *Assume that 0 is separated from the other eigenvalues of $\mathbf{R}$ for each large M, N. Consider an interval (a, b) containing $[x_1^-, x_1^+]$ and for which $b < x_2^-$. Then, if M, N are large enough, almost surely, the first $M - K$ eigenvalues of $\hat{\mathbf{R}}$ belong to (a, b) and the other ones do not belong to (a, b).*

A quite similar result was established in [1] in the context of zero mean large random matrices. In our context, it has not yet been established formally. However, numerical simulations clearly indicate that this conjecture is correct. See e.g. the discussion [3]. From now on, we assume that the following statement is verified.

Assumption 1 *The eigenvalue 0 is separated from the other eigenvalues and Conjecture 1 is correct.*

We now state a technical, but very important lemma.

Lemma 1 *Let $w(z)$ be the function defined on $\mathbb{C}$ by*

$$w(z) = z(1 + \sigma^2 cm(z))^2 - \sigma^2(1 - c)(1 + \sigma^2 cm(z)) \quad (15)$$

Then, $w(z)$ is holomorphic on $\mathbb{C} - \mathcal{S}$, and $\mathrm{Im}(w(z)) \neq 0$ if $z \in \mathbb{C} - \mathbb{R}$. Moreover, $w(x)$ is real if $x \in \mathbb{R} - \mathcal{S}$.
If we denote by $\mathcal{C}$ the set defined by

$$\mathcal{C} = \{w(x) : x \in [x_1^-, x_1^+]\} \cup \{w(x)^* : x \in [x_1^-, x_1^+]\} \quad (16)$$

then $\mathcal{C}$ is a closed contour enclosing 0 but no other eigenvalue of matrix $\mathbf{R}$. Finally, the winding number of $\mathcal{C}$ around 0 is equal to 1.

We finish this section by a useful convergence result.

Proposition 1 *As $M, N \to \infty$ at the same rate, $\frac{1}{M}\mathrm{Trace}(\hat{\mathbf{R}} - z\mathbf{I})^{-1}$ converges a.s for $z \in \mathbb{C} - \mathcal{S}$ towards $m(z)$. Moreover, the entries of matrix $(\hat{\mathbf{R}} - z\mathbf{I})^{-1}$ converge a.s for $z \in \mathbb{C} - \mathcal{S}$ towards the entries of the matrix $\mathbf{T}(z)$ defined by*

$$\mathbf{T}(z) = (1 + \sigma^2 cm(z)) \left[\mathbf{R} - w(z)\mathbf{I}\right]^{-1} \quad (17)$$

Remark: In order to connect Proposition 1 with the results used in [4], we recall the asymptotic behaviour of the entries of $(\hat{\mathbf{R}} - z\mathbf{I})^{-1}$ in the case where the source signal is temporally uncorrelated. Let $m_{iid}(z)$ be the solution to the equation

$$m_{iid}(z) = \frac{1}{M}\mathrm{Tr}\, \mathbf{T}_{iid}(z)$$
$$\mathbf{T}_{iid}(z) = \left[\left(\mathbf{A}\mathbf{A}^\dagger + \sigma^2\mathbf{I}\right)(1 - c - czm_{iid}(z)) - z\mathbf{I}\right]^{-1}$$

Then the entries of $(\hat{\mathbf{R}} - z\mathbf{I})^{-1}$ have the same asymptotic behaviour that the entries of $\mathbf{T}_{iid}(z)$. One can verify that the entries of $\mathbf{T}(z)$, which depend on $\mathbf{S}$, converge to the entries of $\mathbf{T}_{iid}(z)$. Therefore, in this context, our estimator is essentially equivalent to the proposal of [4]. However, if $\mathbf{S}$ is not temporally iid, then the entries of $(\hat{\mathbf{R}} - z\mathbf{I})^{-1}$ do not have the same asymptotic behaviour that the entries of $\mathbf{T}_{iid}(z)$. In this context, the various equations that allow to derive the estimator of [4] differ from what we propose in the present paper and do not lead in principle to a consistent estimator.

4. DERIVATION OF A CONSISTENT SUBSPACE ESTIMATE

We first give a description of the asymptotic behaviour of the conventional MUSIC estimate $\hat{\eta}_{trad}$ of η defined by (7). One can show that when $M, N \to +\infty$, $c = \frac{M}{N} \to 0$, $\hat{\eta}_{trad}$ consistently estimates $\overline{\eta}$. However, if $M, N \to +\infty$, $c = \frac{M}{N}$ converges towards a non zero constant, $\hat{\eta}_{trad}$ becomes inconsistent. We now present our new consistent estimator of η.

Theorem 2 *As $M, N \to \infty$ at the same rate, $\eta = \mathbf{b}^H \mathbf{\Pi} \mathbf{b}$ is consistently estimated by $\hat{\eta}_{new}$ defined by*

$$\hat{\eta}_{new} = \sum_{k=1}^{M} \hat{\beta}_k\, \mathbf{b}^H \hat{\mathbf{e}}_k \hat{\mathbf{e}}_k^H \mathbf{b} \quad (18)$$

where $(\hat{\beta}_k)_{k=1,\ldots,M-K}$ are defined by

$$\hat{\beta}_k = 1 + \frac{\sigma^2}{N} \sum_{l=M-K+1}^{M} \frac{1}{\hat{\lambda}_l - \hat{\lambda}_k} + \frac{2\sigma^2}{N} \sum_{l=M-K+1}^{M} \frac{\hat{\lambda}_k}{(\hat{\lambda}_k - \hat{\lambda}_l)^2}$$
$$- \sigma^2(1 - c)\left(\sum_{l=M-K+1}^{M} \frac{1}{\hat{\lambda}_l - \hat{\lambda}_k} - \sum_{l=M-K+1}^{M} \frac{1}{\hat{\mu}_l - \hat{\lambda}_k}\right) \quad (19)$$

and where $(\hat{\beta}_k)_{k=M-K+1,\ldots,M}$ are defined by

$$\hat{\beta}_k = \frac{\sigma^2}{N} \sum_{l=1}^{M-K} \frac{1}{\hat{\lambda}_k - \hat{\lambda}_l} - \frac{2\sigma^2}{N} \sum_{l=1}^{M-K} \frac{\hat{\lambda}_k}{(\hat{\lambda}_k - \hat{\lambda}_l)^2}$$
$$+ \sigma^2(1 - c)\left(\sum_{l=1}^{M-K} \frac{1}{\hat{\lambda}_k - \hat{\mu}_l} - \sum_{l=1}^{M-K} \frac{1}{\hat{\lambda}_k - \hat{\lambda}_l}\right) \quad (20)$$

The $(\hat{\mu}_l)_{l=1,\ldots,M}$ are the solutions (arranged in increasing order) of the equation $1 + \frac{\sigma^2}{N}\mathrm{Trace}\left(\hat{\mathbf{R}} - x\mathbf{I}\right)^{-1} = 0$.

Sketch of the proof : The starting point is based on Lemma 1 which allows to express $\eta = \mathbf{b}^H \mathbf{\Pi} \mathbf{b}$ with the Cauchy Integral Formula as

$$\eta = \frac{1}{2i\pi} \int_{\mathcal{C}^-} \mathbf{b}^H (\mathbf{R} - \lambda\mathbf{I})^{-1} \mathbf{b}\, \mathrm{d}\lambda \quad (21)$$

where the notation $\mathcal{C}^-$ means that the contour $\mathcal{C}$ defined by (16) is negatively oriented. Using the parametrization defined in (16), we immediately get that

$$\eta = \frac{1}{\pi}\mathrm{Im}\left[\int_{x_1^-}^{x_1^+} \mathbf{b}^H (\mathbf{R} - w(x)\mathbf{I})^{-1} \mathbf{b}\, w'(x)\mathrm{d}x\right] \quad (22)$$

where $w'(x)$ represents the derivative of $w(x)$. As $w(x)$ is real if $x \in \mathbb{R} - \mathcal{S}$, (22) is equal to

$$\eta = \frac{1}{\pi}\mathrm{Im}\left[\int_a^b \mathbf{b}^H (\mathbf{R} - w(x)\mathbf{I})^{-1} \mathbf{b}\, w'(x)\mathrm{d}x\right] \quad (23)$$

where $a < x_1^- < x_1^+ < b < x_2^-$. We notice that $\mathrm{Im}(w(z)) \neq 0$ if $z \in \mathbb{C} - \mathbb{R}$, and that the integrand on the right hand side of (22) is holomorphic on $\mathbb{C} - \mathbb{R}$. Therefore, (23) can be written as

$$\eta = \frac{1}{2i\pi} \int_{\partial\mathcal{R}_y^-} \mathbf{b}^H (\mathbf{R} - w(z)\mathbf{I})^{-1} \mathbf{b}\, w'(z)\mathrm{d}z \quad (24)$$

where $\partial\mathcal{R}_y^-$ represents the negatively oriented boundary of the rectangle $\mathcal{R}_y = \{u + iv : u \in [a, b], v \in [-y, y]\}$. Proposition 1 implies that η is also given by

$$\eta = \frac{1}{2i\pi} \int_{\partial\mathcal{R}_y^-} \frac{\mathbf{b}^H \mathbf{T}(z)\mathbf{b}}{1 + \sigma^2 cm(z)}\, w'(z)\mathrm{d}z \qquad (25)$$

The integrand on the right hand side of (25) can be consistently estimated using the results in Proposition 1. Indeed, Proposition 1 implies that $m(z)$ and $\mathbf{b}^H\mathbf{T}(z)\mathbf{b}$ can be consistently estimated by $\frac{1}{M}\mathrm{Trace}(\hat{\mathbf{R}} - z\mathbf{I})^{-1}$ and $\mathbf{b}^H(\hat{\mathbf{R}} - z\mathbf{I})^{-1}\mathbf{b}$ respectively for each $z \in \partial\mathcal{R}_y^-$. In order to estimate $w'(z)$, we evaluate $w'(z)$ in terms of $m(z)$ and $m'(z)$, and replace these two functions by their corresponding consistent estimates. Hence, the final consistent estimate $\hat{\eta}_{new}$ is obtained by replacing the integrand in (25) by its consistent estimate, denoted by $\hat{g}(z)$. In other words,

$$\hat{\eta}_{new} = \frac{1}{2i\pi} \int_{\partial\mathcal{R}_y^-} \hat{g}(z)\, \mathrm{d}z \qquad (26)$$

where $\hat{g}(z)$ is a rational function of z. Therefore, the integral can be evaluated using the residue theorem. The poles of $\hat{g}(z)$ are the eigenvalues $(\hat{\lambda}_k)_{k=1,\ldots,M}$ as well as the zeros of $1 + c\frac{\sigma^2}{M}\mathrm{Trace}(\hat{\mathbf{R}} - z\mathbf{I})^{-1}$. Conjecture 1 implies that if M and N are large enough, the first $M - K$ eigenvalues $(\hat{\lambda}_k)_{k=1,\ldots,M-K}$ belong to the rectangle $\mathcal{R}_y$ while the last ones $(\hat{\lambda}_k)_{k=M-K+1,\ldots,M}$ are located outside $\mathcal{R}_y$. Moreover, it can be shown that the $(\hat{\mu}_k)_{k=1,\ldots,M-K}$ belong to the rectangle $\mathcal{R}_y$ while the $(\hat{\mu}_k)_{k=M-K+1,\ldots,M}$ are located outside $\mathcal{R}_y$. These remarks allow to establish that $\hat{\eta}_{new}$ is given by (18).

5. NUMERICAL RESULTS

We compare the results provided by the traditional subspace estimate, the new estimate (18) (referred to in the figure as "General Case"), and the improved estimate of [4] derived under the assumption that the source signals are i.i.d. sequences (referred to as "Unconditional Case").

We consider two closely spaced sources with equal power impinging on a uniform linear array from DoAs of 16° and 18° w.r.t the broadside of the antenna array. The emitted symbols come from a 16-QAM constellation and are filtered by a raised-cosine with a roll-of of 0.5. The oversampling rate of the two sources are respectively 2 and 4. The number of antennas is $M = 20$ and the number of snapshots is $N = 40$. The distance between two consecutive antennas is half a wavelength. The estimates are obtained by evaluating the three localization functions for $\mathbf{b} = \mathbf{a}(\theta)$ for different values of θ. In each case, the two estimated angles are defined as the two deepest local minima of the estimated localization function.

We compare in Figure 1 the outlier probability of the three approaches versus the SNR. An outlier is declared when one of the two estimated angles is separated from the true one by more than half of the separation between the two true sources. To achieve a probability of 0.5, we notice that we have a gain of 10 dB by using the new estimator instead of the traditional one.

In Figure 2, we plot the square root of the relative mean square error (referred to as "Relative Standard Deviation") of the estimates w.r.t. the SNR, and the Cramer-Rao Bound (CRB) for correlated sources. Concerning the achievement of the CRB, this figure shows an improvement of 4 dB.

Moreover, for the parameters chosen in the simulations, the eigenvalue 0 separation condition is verified near 8 dB, and we clearly see in the two figures that the performances begin to improve near this SNR value. However, it is quite surprising (in regards to the remark at the end of Section 3) that the improved estimator in the unconditional case have the same performance than the new one, even with time-correlated sources. The link between these two estimators could be an interesting topic.

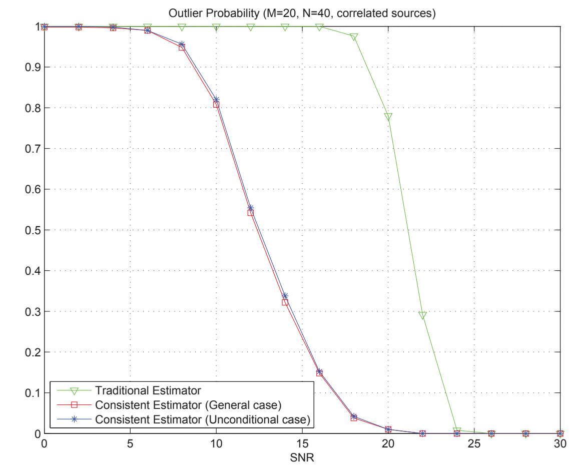

Fig. 1. Outlier Probability

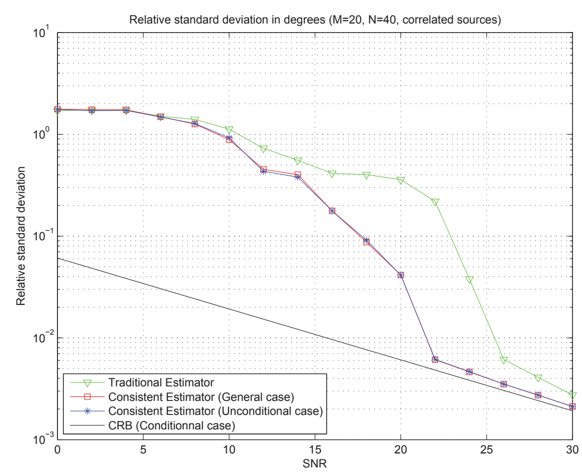

Fig. 2. Relative Standard Deviation

6. REFERENCES

[1] Z.D. Bai, J.W. Silverstein, "Exact separation of eigenvalues of large dimensional sample covariance matrices", *Ann. Probab.*, vol. 27, no. 3, pp. 1536-1555, 1999.

[2] R.B. Dozier, J.W. Silverstein, "On the empirical distribution of eigenvalues of large dimension information-plus-noise type matrices", *Journal of Multivariate Analysis 98* (2007), pp. 678-694.

[3] R.B. Dozier, J.W. Silverstein, "Analysis of the Limiting Spectral Distribution of Large Dimensional Information-Plus-Noise Type Matrices", *Journal of Multivariate Analysis 98* (2007), pp. 1099-1122.

[4] X. Mestre, M.A. Lagunas, "Modified subspace algorithms for DoA estimation with large arrays", *IEEE Trans. on Signal Processing*, vol. 56, no. 2, February 2008.