AN IMPROVED SUBSPACE-BASED ALGORITHM IN THE SMALL SAMPLE SIZE REGIME

Xavier Mestre, Francisco Rubio

Centre Tecnològic de Telecomunicacions de Catalunya (CTTC)
Parc Mediterrani de la Tecnologia, Av/ del Canal Olímpic s/n, Castelldefels, 08860 Barcelona, Spain
URL: http://www.cttc.es/, e-Mail:xavier.mestre@cttc.es

ABSTRACT

A new method for subspace identification in array signal processing applications is proposed. The method is based on random matrix theory and provides consistent estimates even when the observation dimension increases without bound at the same rate as the number of observations. This guarantees a good behavior in finite sample size situations, where the number of sensors and the number of samples have the same order of magnitude. Consistency of the algorithm holds in situations where the signal and noise subspaces are asymptotically separable in the sense that, in the asymptotic sample eigenvalue distribution, signal and noise eigenvalues generate different spectral clusters.

1. INTRODUCTION

Subspace-based algorithms exploit the orthogonality between signal and noise subspaces in order to derive estimators of quantities that parametrize one of these subspaces. One of the main drawbacks of the subspace-based methods is the fact that their performance drops dramatically when either the signal to noise ratio or the number of samples fall below a certain threshold [1][2]. This behavior has usually been attributed to the presence of a "subspace swap", meaning that some of the noise eigenvectors of the sample covariance matrix are mistakenly associated with the signal subspace and vice-versa [3][4]. In this paper, we take a different approach to the analysis of the performance breakdown effect in subspace methods. We intentionally focus on the small sample size regime in array signal processing applications, whereby the number of samples (N) and the number of sensors (M) are of the same order of magnitude. Since the finite case is essentially intractable, we concentrate on the asymptotic behavior, i.e. we assume that both quantities are large but have the same order of magnitude ($M \to \infty$, $N \to \infty$, $M/N \to c$, $0 < c < \infty$). This type of limit is usually very useful in order to characterize the finite sample size behavior of array processing architectures.

Let us consider a collection of N complex valued array observations, $\mathbf{y}(n) \in \mathbb{C}^{M\times 1}$, $n = 1 \dots N$ obtained from an array of $M > 1$ sensors. Let $\mathbf{R}$ represent the true $M \times M$ covariance matrix of the observation, with eigenvectors $\{\mathbf{e}_i,\ i = 1 \dots M\}$ and associated eigenvalues $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_M$. Each eigenvalue is repeated according to its multiplicity, and the eigenvectors associated with eigenvalues of multiplicity higher than one are taken to be an arbitrary orthonormal basis of the associated subspace. Let $\mathbf{E}_N \in \mathbb{C}^{M\times K}$ denote a matrix containing an orthonormal basis of eigenvectors spanning the noise subspace, i.e. those associated with the smallest eigenvalue of $\mathbf{R}$ (which has multiplicity K, so that $\lambda_1 = \ldots = \lambda_K$). We will assume throughout the paper the the dimension of the noise subspace, namely K, is known.

Subspace identification algorithms are based on the property that any vector $\mathbf{s}$ lying on the signal subspace is orthogonal to the columns of $\mathbf{E}_N$, that is $\mathbf{E}_N^H\mathbf{s} = \mathbf{0}$. In most array signal processing applications, one has access to a parametrized version of the vector $\mathbf{s} = \mathbf{s}(\boldsymbol{\theta})$ as a function of some parameter vector $\boldsymbol{\theta}$. This parameter vector can be found by inspecting those values of $\boldsymbol{\theta}$ that have the property $\mathbf{E}_N^H\mathbf{s}(\boldsymbol{\theta}) = \mathbf{0}$, i.e. values of $\boldsymbol{\theta}$ for which the column $\mathbf{s}(\boldsymbol{\theta})$ lies on the signal subspace. In practice, though, the eigenvectors $\{\mathbf{e}_i\}$ are known, and must be estimated from the received data. Let $\hat{\mathbf{R}}$ represent the sample covariance matrix drawn from N consecutive samples. We will denote by $\{\hat{\mathbf{e}}_i,\ i = 1 \dots M\}$ and $\hat{\lambda}_1 \leq \hat{\lambda}_2 \leq \ldots \leq \hat{\lambda}_M$ the eigenvectors and associated eigenvalues of $\hat{\mathbf{R}}$. In particular, $\hat{\mathbf{E}}_N$ will be a matrix containing the noise eigenvectors of the sample covariance matrix. Since in general $\hat{\mathbf{E}}_N \neq \mathbf{E}_N$, the columns of $\hat{\mathbf{E}}_N$ are not completely orthogonal to any vector $\mathbf{s}$ lying on the signal subspace. However, assuming that $\hat{\mathbf{E}}_N$ is a close approximation of $\mathbf{E}_N$, one can expect that the projection of a vector in the signal subspace onto the columns of $\hat{\mathbf{E}}_N$ will in general have a very small norm. Based on that principle, subspace identification techniques estimate a certain set of parameters $\boldsymbol{\theta}$ by minimizing the norm of such a projection, i.e. the minimizers of the cost function

$$\tilde{\eta}(\boldsymbol{\theta}) = \mathbf{s}^H(\boldsymbol{\theta})\,\hat{\mathbf{E}}_N\hat{\mathbf{E}}_N^H\mathbf{s}(\boldsymbol{\theta}). \tag{1}$$

This paper will only be concerned with subspace-based algorithms that conform to this quadratic structure. One can imagine $\tilde{\eta}(\boldsymbol{\theta})$ as an estimator of the true cost function $\eta(\boldsymbol{\theta}) = \mathbf{s}^H(\boldsymbol{\theta})\,\mathbf{E}_N\mathbf{E}_N^H\mathbf{s}(\boldsymbol{\theta})$, which is in practice unknown. In spectral analysis and DoA detection applications, this estimating procedure is usually referred to as the MUSIC algorithm [5, 6], and the cost function in (1) as the MUSIC pseudospectrum. Under the appropriate statistical assumptions, and when the number of samples increases without bound ($N \to \infty$) and the observation dimension is taken to be a fixed quantity ($M < \infty$), $\tilde{\eta}(\boldsymbol{\theta})$ tends almost surely to the deterministic original cost function $\eta(\boldsymbol{\theta})$. If this convergence is uniform in $\boldsymbol{\theta}$, one can ensure that the estimates obtained are consistent as $N \to \infty$, $M < \infty$. In this situation, we will state that the estimator is N-consistent.

In practice, however, both the number of samples (N) and the number of sensors (M) are finite quantities. Therefore, it seems

This work was partially funded by the Spanish Ministry of Science and Technology (MCYT) under projects FIT-330210-2005-23, FIT-330220-2005-108 and the European Comission IST-2002-507525, IST-2002-508009.

quite natural to investigate the performance of subspace algorithms whenever M, N have the same order of magnitude. This issue has been less studied in the literature, mainly due to the high complexity involved. However, a lot of insight can be gained by analyzing the asymptotic behavior of the cost function in (1) when $M, N \to \infty$ for $M/N = c$, $0 < c < \infty$. Note that this is a good approximation of reality, in the sense that M, N have the same order of magnitude as in a practical situation. In order to carry out such bidimensional limits, one must resort to random matrix theory techniques.

2. ASYMPTOTIC DESCRIPTION OF $\hat{\mathbf{R}}$ AND M, N-INCONSISTENCY OF TRADITIONAL APPROACHES

Random matrix theory is a branch of statistics that is devoted to the study of the asymptotic behavior of the eigenvalues and eigenvectors when their dimensions tend to infinity. A popular result of random matrix theory states that, under some appropriate technical assumptions, the empirical eigenvalue distribution of $\hat{\mathbf{R}}$ tends (as $M, N \to \infty$ at the same rate) to a compactly supported non-random density. Furthermore, one can characterize the asymptotic distribution of the eigenvalues of $\hat{\mathbf{R}}$ in terms of the asymptotic distribution of the eigenvalues of $\mathbf{R}$. That characterization is usually carried out in terms of the Stieltjes transforms of the corresponding densities of eigenvalues, which are complex functions of the form

$$\hat{m}(z) = \mathbf{s}^H \left(\hat{\mathbf{R}} - z\mathbf{I}_M\right)^{-1} \mathbf{s}, \quad \hat{b}(z) = \frac{1}{M} \operatorname{tr}\left[\left(\hat{\mathbf{R}} - z\mathbf{I}_M\right)^{-1}\right] \tag{2}$$

when associated with $\hat{\mathbf{R}}$ or

$$m(z) = \mathbf{s}^H \left(\mathbf{R} - z\mathbf{I}_M\right)^{-1} \mathbf{s}, \quad b(z) = \frac{1}{M} \operatorname{tr}\left[\left(\mathbf{R} - z\mathbf{I}_M\right)^{-1}\right] \tag{3}$$

when associated with $\mathbf{R}$.

The convergence of the functions in (2) as $M, N \to \infty$ at the same rate can be studied under the following additional statistical assumptions:

(As1) The observation vectors $\mathbf{y}(n)$ can be modeled as $\mathbf{y}(n) = \mathbf{R}^{1/2}\mathbf{u}(n)$, where $\mathbf{R}^{1/2}$ is a positive definite $M \times M$ Hermitian deterministic matrix, and $\mathbf{u}(n)$, $n = 1 \ldots N$, is a collection of independent and identically distributed complex $M \times 1$ random vectors. The real and imaginary parts of the entries of $\mathbf{u}(n)$ are all i.i.d. absolutely continuous random variables with zero mean, variance $1/2$ and finite absolute eighth moments.

(As2) the spatial correlation matrix $\mathbf{R}$ has uniformly bounded spectral radius for all M and the functions $m(z)$ and $b(z)$ have a limit as $M \to \infty$ for all $z \in \mathbb{C}^+ = \{z \in \mathbb{C} : \operatorname{Im}[z] < 0\}$. Furthermore, $\sup_M \left(\sqrt{M} \max_{1 \le i \le M} \left|\{\mathbf{s}\}_i\right|\right) < +\infty$, where $\{\mathbf{s}\}_i$ denotes the ith entry of the column vector $\mathbf{s}$.

For any $z \in \mathbb{C}^+$, and under **(As1-As2)** the functions $\hat{m}(z)$ and $\hat{b}(z)$ are asymptotically close to two deterministic counterparts $\bar{m}(z)$ and $\bar{b}(z)$, in the sense that [7, 8]

$$|\hat{m}(z) - \bar{m}(z)| \to 0, \quad \left|\hat{b}(z) - \bar{b}(z)\right| \to 0$$

almost surely for all $z \in \mathbb{C}^+$ as $M, N \to \infty$ at the same rate, where $\bar{b}(z) = b$ is the unique solution to the following equation in the set $\left\{b \in \mathbb{C} : -(1-c)/z + cb \in \mathbb{C}^+\right\}$:

$$b = \frac{1}{M} \sum_{m=1}^{M} \frac{1}{\lambda_m \left(1 - c - czb\right) - z}, \tag{4}$$

and

$$\bar{m}(z) = \sum_{m=1}^{M} \frac{\mathbf{s}^H \mathbf{e}_m \mathbf{e}_m^H \mathbf{s}}{\lambda_m \left(1 - c - cz\bar{b}(z)\right) - z}. \tag{5}$$

With this result, we observe that $\hat{b}(z)$ and $\hat{m}(z)$ are asymptotically close to the deterministic counterparts $\bar{b}(z)$ and $\bar{m}(z)$, which are much easier to analyze. To simplify the analysis, and since we are only interested in the asymptotic behavior as an approximation of the finite reality, we will assume that the empirical eigenvalue distribution of $\mathbf{R}$ is not altered as $M \to \infty$. In this case, $\bar{b}(z)$ is the Stieltjes transform associated with the asymptotic sample eigenvalue distribution, the density of which will be denoted by $q(x)$. This density can easily be retrieved from $\bar{b}(z)$ using the Stieltjes inversion formula, namely $q(x) = \lim_{y \to 0^+} \frac{1}{\pi} \operatorname{Im}\left[\bar{b}(x + \mathrm{j}\, y)\right]$, valid for all $x \in \mathbb{R}$ such that $q(x)$ exists. In practice, one observes that the density $q(x)$ presents a series of eigenvalue clusters, each one associated to one or more eigenvalues of the true covariance matrix (see further [9]). As the number of samples per sensor increases ($c \to 0$) or the true eigenvalues become more separated from one another, each true eigenvalue tends to generate a single eigenvalue cluster separated from the rest in the asymptotic sample eigenvalue distribution. This will be further exploited in the next section.

Let us now use these results to study the behavior of the traditional subspace estimation function in (1) as $M, N \to \infty$ at the same rate. Note, first, that we can express that cost function in terms of $\hat{m}(z)$ as follows (from now on, we obviate the dependence on the parameters θ):

$$\tilde{\eta} = \lim_{\epsilon \to 0} \lim_{y \to 0} \frac{1}{\pi} \int_0^{\hat{\lambda}_K + \epsilon} \operatorname{Im}\left[\hat{m}(x + \mathrm{j}\, y)\right] dx \tag{6}$$

where $\epsilon > 0$. This can be readily seen by replacing $\hat{m}(x + \mathrm{j}\, y)$ above with the expression given in (2).

Using (6) and noting from (5) that $|\hat{m}(z) - \bar{m}(z)| \to 0$ as $M, N \to \infty$ at the same rate, we can ensure that $|\tilde{\eta} - \bar{\eta}| \to 0$ almost surely, where

$$\bar{\eta} = \sum_{m=1}^{M} w_m \mathbf{s}^H \mathbf{e}_m \mathbf{e}_m^H \mathbf{s}$$

$$w_m = \int_0^{\theta} \frac{c\lambda_m x q(x) dx}{\left(\lambda_m \left(1 - c - cx p(x)\right) - x\right)^2 + \left(c\lambda_m x \pi q(x)\right)^2}$$

and with $p(x)$ and $q(x)$ denoting the real part and the scaled imaginary part of $\bar{b}(z)$ when restricted to the real axis, i.e. $\lim_{y \to 0^+} \bar{b}(x + \mathrm{j}\, y) = p(x) + \pi\, \mathrm{j}\, q(x)$, and $\theta > 0$ being the unique positive value such that

$$\int_0^{\theta} q(x) dx = \frac{K}{M}.$$

It is important to note that, when $M, N \to \infty$ at the same rate, the traditional subspace-based estimator in (1) does not concentrate all the energy on the true noise eigenvectors. Indeed, we see from the expression of the coefficients w_m that these values can be non-zero even when $m > K$, implying that part of the energy

concentrated on $\hat{\mathbf{E}}_N\hat{\mathbf{E}}_N^H$ is spilled onto the true signal subspace instead of remaining concentrated on the noise subspace. This energy spilling onto the signal subspace is the cause for not having complete orthogonality between the sample noise subspace and the true signal subspace when M and N have the same dimension (as it occurs in a finite sample size situation). This is the cause for the performance breakdown effect of subspace-based algorithms.

3. AN M, N-CONSISTENT SUBSPACE ESTIMATOR

From all the above, we can conclude that traditional subspace-approaches are N-consistent but not M, N-consistent. In this section, we propose a new subspace-based estimator that is M, N-consistent, i.e. consistent not only when $N \to \infty$ for a fixed M, but also when $M, N \to \infty$ at the same rate. This will guarantee a better behavior of the algorithm whenever M and N have the same magnitude.

Note, first of all, that we can give an analytical expression for η as a function of $m(z)$ in (3), simply by making use of the Cauchy integral formula, i.e.

$$\eta = \frac{1}{2\pi\,\mathrm{j}}\oint_{C^-}\frac{\mathbf{s}^H\mathbf{E}_N\mathbf{E}_N^H\mathbf{s}}{\lambda_1 - z}dz = \frac{1}{2\pi\,\mathrm{j}}\oint_{C^-} m(z)dz \qquad (7)$$

where C^- is a negatively[1] oriented contour taking values on $\mathbb{C}\backslash\{\lambda_1,\ldots,\lambda_M\}$ and enclosing only the noise eigenvalue, namely $\lambda_1 = \ldots = \lambda_K$. In particular, we can choose a contour generated as follows. Consider the following complex function, defined on the positive real axis $f : \mathbb{R}^+ \longrightarrow \mathbb{C}$,

$$f(x) = \frac{x}{1 - c - cx\bar{b}(x)} \qquad (8)$$

where, with some abuse of notation, $\bar{b}(x) = \lim_{y\to 0^+}\bar{b}(x + \mathrm{j}\,y)$, where $\bar{b}(z)$ is as defined in the last section. It can be shown [10] that $f(x)$ in (8) is continuous on $\mathbb{R}^+$ and differentiable on all the domain except for a finite number of points, corresponding to the boundary of the support of the limiting sample eigenvalue distribution. On the other hand, the real part of $f(x)$ is monotonically increasing with x, whereas the imaginary part of $f(x)$ is positive for x in the support of the limiting sample eigenvalue distribution and zero elsewhere.

In order to derive the estimators, we need to assume that there exists separation between noise and signal eigenvalue clusters in the asymptotic distribution function of the eigenvalues of $\hat{\mathbf{R}}$. This can be technically expressed as follows (see further [10]).

(**As3**) **Splitting condition**. We will assume that the number of available samples per sensor is higher than a predefined value,

$$\frac{N}{M} > \frac{1}{M}\sum_{m=1}^{M}\left(\frac{\lambda_m}{\lambda_m - \eta}\right)^2$$

where η is the smallest real valued solution to the following equation:

$$\frac{1}{M}\sum_{m=1}^{M}\frac{\lambda_m^2}{(\lambda_m - \eta)^3} = 0. \qquad (9)$$

This is the minimum number of samples per sensor that is needed in order to guarantee separation of the noise and first signal eigenvalue cluster of the asymptotic eigenvalue distribution of $\hat{\mathbf{R}}$.

[1] We follow the usual convention of orienting positive contours counterclockwise.

From the properties of $f(x)$ described above, and under (**As3**), one can always choose two real-valued constants, σ_1 and σ_2, $\sigma_1 < \sigma_2$, such that both $f(\sigma_1)$ and $f(\sigma_2)$ are real-valued, and the interval $[f(\sigma_1), f(\sigma_2)]$ contains the noise eigenvalue only. Now, it can be seen that, if we move x from σ_1 to σ_2, the function $f(x)$ concatenated with its conjugate (denoted by $f^*(x)$) as x goes back from σ_2 to σ_1 describes a curve with the same properties as the contour C^- above. Consequently, we can express (7) as

$$\eta = \frac{1}{2\pi\,\mathrm{j}}\int_{\sigma_1}^{\sigma_2} m(f(x))f'(x)dx - \frac{1}{2\pi\,\mathrm{j}}\int_{\sigma_1}^{\sigma_2} m(f^*(x))\,(f^*(x))'\,dx \qquad (10)$$

where $f'(x)$ and $(f^*(x))'$ denote the derivative of $f(x)$ and $f^*(x)$ respectively.

Observe now that, using (5) restricted to $z = x \in \mathbb{R}^+$, we can write

$$\bar{m}(x)\left(1 - c - cx\bar{b}(x)\right) = \sum_{m=1}^{M}\frac{\mathbf{s}^H\mathbf{e}_m\mathbf{e}_m^H\mathbf{s}}{\lambda_m - f(x)} = m\left(f(x)\right). \qquad (11)$$

Hence, multiplying both sides of the last equations and its conjugate by $f'(x)$ and $(f^*(x))'$ respectively, and using (10), we obtain

$$\eta = \frac{1}{\pi}\int_{\sigma_1}^{\sigma_2}\mathrm{Im}\left[\bar{m}(x)\left(1 - c - cx\bar{b}(x)\right)f'(x)\right]dx$$

or, from the definition of $f(x)$ in (8),

$$\eta = \frac{1}{\pi}\int_{\sigma_1}^{\sigma_2}\mathrm{Im}\left[\bar{m}(x)\frac{1 - c + cx^2\bar{b}'(x)}{1 - c - cx\bar{b}(x)}\right]dx.$$

It can be shown that the integrand is bounded over $\mathbb{C}$, and by virtue of the Dominated Convergence Theorem,

$$\eta = \lim_{y\to 0^+}\frac{1}{2\pi\,\mathrm{j}}\int_{\sigma_1}^{\sigma_2}\bar{m}(z)\frac{1 - c + cz^2\bar{b}'(z)}{1 - c - cz\bar{b}(z)}dx - \lim_{y\to 0^-}\frac{1}{2\pi\,\mathrm{j}}\int_{\sigma_1}^{\sigma_2}\bar{m}(z)\frac{1 - c + cz^2\bar{b}'(z)}{1 - c - cz\bar{b}(z)}dx$$

with $z = x + \mathrm{j}\,y$ and where we have *defined* $\bar{m}(z) = \bar{m}^*(z^*)$ and $\bar{b}(z) = \bar{b}^*(z^*)$ when $z \in \mathbb{C}^- = \{z \in \mathbb{C} : \mathrm{Im}\,[z] < 0\}$. Using again the fact that the integrands in the above equation are bounded, we can finally express η as

$$\eta = \lim_{y\to 0^+}\frac{1}{2\pi\,\mathrm{j}}\oint_{\partial R_y^-}\bar{m}(z)\frac{1 - c + cz^2\bar{b}'(z)}{1 - c - cz\bar{b}(z)}dz \qquad (12)$$

where ∂R_y^- is a negatively oriented rectangular contour with vertices $\{\sigma_1 - \mathrm{j}\,y, \sigma_2 - \mathrm{j}\,y, \sigma_2 + \mathrm{j}\,y, \sigma_1 + \mathrm{j}\,y\}$.

The advantage of expressing this quantity in terms of the complex variable z is the fact that, when $z \in \mathbb{C}\backslash\mathbb{R}$, we have a trivial way of M, N-consistently estimating the functions $\bar{m}(z)$ and $\bar{b}(z)$. Indeed, by definition, $\bar{m}(z)$ and $\bar{b}(z)$ are (pointwise) M, N-consistently estimated by $\hat{m}(z)$ and $\hat{b}(z)$ for any $z \in \mathbb{C}\backslash\mathbb{R}$ as $M, N \to \infty$. Therefore, it seems reasonable to obtain a potentially consistent estimator by replacing $\bar{m}(z)$ and $\bar{b}(z)$ with $\hat{m}(z)$ and $\hat{b}(z)$ in the equations above. Inserting $\hat{m}(z)$ and $\hat{b}(z)$ into (12) and using the asymptotic eigenvalue separation arguments in [11], one can show that the proposed estimator takes the form [10]

$$\hat{\eta} = \sum_{m=1}^{M}\varphi_m\mathbf{s}^H\hat{\mathbf{e}}_m\hat{\mathbf{e}}_m^H\mathbf{s},$$

with φ_m, $m = 1 \dots M$, defined as follows:

$$\varphi_m = \begin{cases} 1 & m \leq [M-N]^+ \\ 1 - \phi_m + \psi_m & [M-N]^+ + 1 \leq m \leq K \\ -\phi_m & m > K \end{cases}$$

$$\phi_m = \sum_{k=[M-N]^+ +1}^{K} \frac{\hat{\lambda}_m \left(\hat{\lambda}_k - \mu_k\right)}{\left(\hat{\lambda}_m - \hat{\lambda}_k\right)\left(\hat{\lambda}_m - \mu_k\right)}$$

$$\psi_m = \sum_{\substack{k=[M-N]^+ +1 \\ k \neq m}}^{M} \frac{\hat{\lambda}_k \left(\hat{\lambda}_m - \mu_m\right)}{\left(\hat{\lambda}_k - \hat{\lambda}_m\right)\left(\hat{\lambda}_k - \mu_m\right)}$$

and where $\mu_1 \leq \dots \leq \mu_M$, are the real-valued solutions to the following equation in μ

$$\frac{1}{M} \sum_{k=1}^{M} \frac{\hat{\lambda}_k}{\hat{\lambda}_k - \mu} = \frac{1}{c}. \tag{13}$$

It is worth pointing out that the final proposed estimator takes the form of a weighted subspace identification algorithm that uses both signal and noise subspaces. Our approach gives the (asymptotically) optimum way of combining both subspaces in order to maximize the separation in the small sample size regime. We also stress that $\hat{\eta}$ is not guaranteed to be positive for every value of $\mathbf{s}$. In practice, this has no further implications in most subspace-based identification methods such as MUSIC.

4. NUMERICAL EVALUATION

We compare the performance of the MUSIC algorithm using our subspace estimator with the traditional version of the algorithm. The scenario consisted of two sources located at 35° and 40° of azimuth impinging on a uniform linear array of 10 sensors separated half a wavelength apart and received in ominidirectional background noise. All the signals were mutually independent, circularly symmetric, Gaussian distributed, with zero mean. Figure 1 compares the mean squared error obtained by averaging 500 realizations of the proposed and traditional MUSIC algorithms when $N = 7$ samples and $N = 70$ samples were available at the array. Results are shown for different value of the signal-to-noise (SNR), assumed equal for the two sources. The unconditional Cramér-Rao Bound (CRB) is also shown for informative purposes. Observe that the proposed MUSIC algorithm presents an SNR threshold much lower than the traditional MUSIC algorithm. Furthermore, the proposed approach shows also a better performance in the high SNR region.

5. CONCLUSIONS

We have presented a new algorithm for subspace-based estimation in array processing applications that, under certain conditions, provides consistent estimates when both the number of samples and the number of sensors increase without bound at the same rate. This property guarantees a good performance in the small sample size regime, i.e. in situations were the number of samples and the number of sensors have the same order of magnitude. Simulation results indicate that the threshold at which breakdown of the MUSIC algorithm occurs is significantly reduced with our approach. Interestingly enough, the algorithm also improves the traditional MUSIC performance at high SNR.

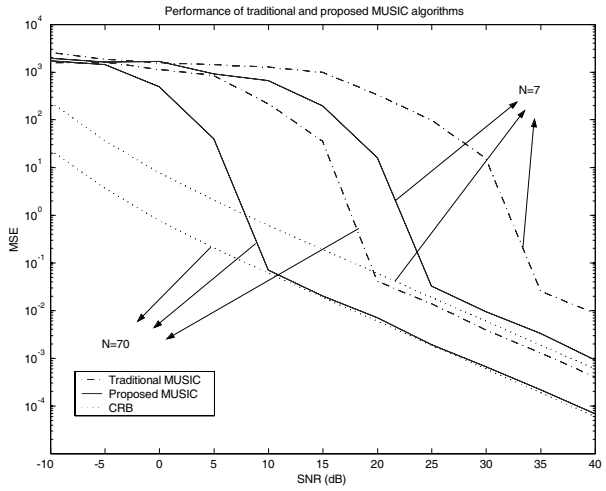

Fig. 1. Comparative evaluation of the MSE achieved by the proposed and traditional MUSIC algorithms as a function of SNR.

6. REFERENCES

[1] D. Tufts, A. Kot, and R. Vaccaro, The Threshold Effect in Signal Processing Algorithms Which Use an Estimated Subspace, in SVD and signal processing, II: Algorithms, analysis and applications, New York, Elsevier, 1991.

[2] P. Stoica, V. Šimonytė, and T. Söderström, "On the resolution performance of spectral analysis," Signal processing, vol. 44, pp. 153–161, June 1995.

[3] J. Thomas, L. Scharf, and D. Tufts, "The probability of a subspace swap in the SVD," IEEE Trans. on Signal Processing, vol. 43, pp. 730–736, Mar. 1995.

[4] M. Hawkes, A. Nehorai, and P. Stoica, "Performance breakdown of subspace-based methods: Prediction and cure," in Proceedings of the IEEE ICASSP 2001, pp. 4005–4008.

[5] R. Schmidt, "Multiple emitter localization and signal parameter estimation," in Proc. of the RADC, Spectral Estimation Workshop,, pp. 243–258, 1979.

[6] G. Bienvenu, "Influence of spatial coherence of the background noise on high resolution passive methods," Proc. IEEE ICASSP, 1979, pp. 306–309.

[7] J. Silverstein, "Strong convergence of the empirical distribution of eigenvalues of large dimensional random matrices," Journal of Multivariate Analysis, vol. 5, pp. 331–339, 1995.

[8] V. Girko, An Introduction to Statistical Analysis of Random Arrays. The Netherlands: VSP, 1998.

[9] J. Silverstein and P. Combettes, "Signal detection via spectral theory of large dimensional random matrices," IEEE Trans. on Signal Processing, vol. 40, pp. 2100–2105, Aug. 1992.

[10] X. Mestre, "Improved subspace-based estimation in the small sample size regime," 2005. Preprint. Available at http://www.cttc.es/drafts/cttc-rc-2005-005.pdf.

[11] Z. Bai and J. Silverstein, "No eigenvalues outside the support of the limiting spectral distribution of large dimensional sample covariance matrices," Annals of probability, vol. 26, pp. 316–345, 1998.