Performance analysis of spatial smoothing schemes in the context of large arrays

Gia-Thuy Pham, Philippe Loubaton, Pascal Vallet

Performance analysis of spatial smoothing schemes in the context of large arrays.

Gia-Thuy Pham, Philippe Loubaton, *Fellow, IEEE* and Pascal Vallet, *Member, IEEE,*

Abstract

This paper adresses the statistical behaviour of spatial smoothing subspace DoA estimation schemes using a sensor array in the case where the number of observations N is significantly smaller than the number of sensors M, and that the smoothing parameter L is such that M and NL are of the same order of magnitude. This context is modelled by an asymptotic regime in which NL and M both converge towards ∞ at the same rate. As in recent works devoted to the study of (unsmoothed) subspace methods in the case where M and N are of the same order of magnitude, it is shown that it is still possible to derive improved DoA estimators termed as Generalized-MUSIC with spatial smoothing (G-MUSIC SS). The key ingredient of this work is a technical result showing that the largest singular values and corresponding singular vectors of low rank deterministic perturbation of certain Gaussian block-Hankel large random matrices behave as if the entries of the latter random matrices were independent identically distributed. This allows to conclude that when the number of sources and their DoA do not scale with M, N, L, a situation modelling widely spaced DoA scenarios, then both traditional and Generalized spatial smoothing subspace methods provide consistent DoA estimators whose convergence speed is faster than $\frac{1}{M}$. The case of DoA that are spaced of the order of a beamwidth, which models closely spaced sources, is also considered. It is shown that the convergence speed of G-MUSIC SS estimates is unchanged, but that it is no longer the case for MUSIC SS ones.

I. Introduction

The statistical analysis of subspace DoA estimation methods using an array of sensors is a topic that has received a lot of attention since the seventies. Most of the works were devoted to the case where the

P. Vallet is with Laboratoire de l'Intégration du Matériau au Système (CNRS, Univ. Bordeaux, IPB), 351, Cours de la Libération 33405 Talence (France), pascal.vallet@ipb.fr

G.T. Pham and P. Loubaton are with Laboratoire d'Informatique Gaspard Monge (CNRS, Université Paris-Est/MLV), 5 Bd. Descartes 77454 Marne-la-Vallée (France), loubaton@univ-mlv.fr

number of available samples N of the observed signal is much larger than the number of sensors M of the array (see e.g. [14] and the references therein). More recently, the case where M and N are large and of the same order of magnitude was addressed for the first time in [10] using large random matrix theory. [10] was followed by various works such as [7], [17], [6], [5]. The number of observations may also be much smaller than the number of sensors. In this context, it is well established that spatial smoothing schemes, originally developped to address coherent sources ([4], [13], [11]), can be used to artificially increase the number of snapshots (see e.g. [14] and the references therein, see also the recent related contributions [15], [16] devoted to the case where $N = 1$). Spatial smoothing consists in considering $L < M$ overlapping arrays with $M - L + 1$ sensors, and allows to generate artificially NL snapshots observed on a virtual array of $M - L + 1$ sensors. The corresponding $(M - L + 1) \times NL$ matrix, denoted $\mathbf{Y}_N^{(L)}$, collecting the observations is the sum of a low rank component generated by $(M - L + 1)$-dimensional steering vectors with a noise matrix having a block-Hankel structure. Subspace methods can still be developed, but the statistical analysis of the corresponding DoA estimators was addressed in the standard regime where $M - L + 1$ remains fixed while NL converges towards ∞. This context is not the most relevant when M is large because L must be chosen in such a way that the number of virtual sensors $M - L + 1$ be small enough w.r.t. NL, thus limiting the statistical performance of the estimates. In this paper, we study the statistical performance of spatial smoothing subspace DoA estimators in asymptotic regimes where $M - L + 1$ and NL both converge towards ∞ at the same rate, where $\frac{L}{M} \to 0$ in order to not affect the aperture of the virtual array, and where the number of sources K does not scale with M, N, L. For this, it is necessary to evaluate the behaviour of the K largest eigenvalues and corresponding eigenvectors of the empirical covariance matrix $\frac{\mathbf{Y}_N^{(L)}\mathbf{Y}_N^{(L)*}}{NL}$. To address this issue, we prove that the above eigenvalues and eigenvectors have the same asymptotic behaviour as if the noise contribution $\mathbf{V}_N^{(L)}$ to matrix $\mathbf{Y}_N^{(L)}$, a block-Hankel random matrix, was a Gaussian random matrix with independent identically distributed. To establish this result, we rely on the recent result [8] addressing the behaviour of the singular values of large block-Hankel random matrices built from i.i.d. Gaussian sequences. [8] implies that the empirical eigenvalue distribution of matrix $\frac{\mathbf{V}_N^{(L)}\mathbf{V}_N^{(L)*}}{NL}$ converges towards the Marcenko-Pastur distribution, and that its eigenvalues are almost surely located in the neigborhood of the support of the above distribution. This allows to generalize the results of [3] to our random matrix model, and to characterize the behaviour of the largest eigenvalues and eigenvectors of $\frac{\mathbf{Y}_N^{(L)}\mathbf{Y}_N^{(L)*}}{NL}$. We deduce from this improved subspace estimators, called DoA G-MUSIC SS (spatial smoothing) estimators, which are similar to those of [17] and [5]. We deduce from the results of [18] that when the DoAs do not scale with M, N, L, i.e. if the DoAs are widely spaced compared to aperture array, then both G-MUSIC SS and traditional MUSIC SS estimators

are consistent and converge at a rate faster than $\frac{1}{M}$. Moreover, when the DoAs are spaced of the order of $\frac{1}{M}$, the behaviour of G-MUSIC SS estimates remains unchanged, but the convergence rate of traditional subspace estimates is lower.

This paper is organized as follows. In section II, we precise the signal models, the underlying assumptions, and formulate our main results. In section III, we prove that the largest singular values and corresponding singular vectors of low rank deterministic perturbation of certain Gaussian block-Hankel large random matrices behave as if the entries of the latter random matrices were independent identically distributed. In section IV, we apply the results of section III to matrix $\mathbf{Y}_N^{(L)}$, and follow [5] in order to propose a G-MUSIC algorithm to the spatial smoothing context of this paper. The consistency and the convergence speed of the G-MUSIC SS estimates and of the traditional MUSIC SS estimates are then deduced from the results of [18]. Finally, section V present numerical experiments sustaining our theoretical results.

Notations : For a complex matrix $\mathbf{A}$, we denote by $\mathbf{A}^T, \mathbf{A}^*$ its transpose and its conjugate transpose, and by $\mathrm{Tr}\,(\mathbf{A})$ and $\|\mathbf{A}\|$ its trace and spectral norm. The identity matrix will be $\mathbf{I}$ and $\mathbf{e}_n$ will refer to a vector having all its components equal to 0 except the n-th equals to 1. For a sequence of random variables $(X_n)_{n\in\mathbb{N}}$ and a random variable X, we write

$$X_n \xrightarrow[n\to\infty]{a.s.} X$$

when X_n converges almost surely towards X. Finally, $X_n = o_{\mathbb{P}}(1)$ will stand for the convergence of X_n to 0 in probability, and $X_n = \mathcal{O}_{\mathbb{P}}(1)$ will stand for tightness (boundedness in probability).

II. Problem formulation and main results.

A. Problem formulation.

We assume that K narrow-band and far-field source signals are impinging on a uniform linear array of M sensors, with $K < M$. In this context, the M–dimensional received signal $(\mathbf{y}_n)_{n\geq 1}$ can be written as

$$\mathbf{y}_n = \mathbf{A}_M \mathbf{s}_n + \mathbf{v}_n,$$

where

- $\mathbf{A}_M = [\mathbf{a}_M(\theta_1), \ldots, \mathbf{a}_M(\theta_K)]$ is the $M \times K$ matrix of M–dimensionals steering vectors $\mathbf{a}_M(\theta_1), \ldots, \mathbf{a}_M(\theta_K)$, with $\theta_1, \ldots, \theta_K$ the source signals DoA, and $\mathbf{a}_M(\theta) = \frac{1}{\sqrt{M}}[1, \ldots, \mathrm{e}^{\mathrm{i}(M-1)\theta}]^T$;
- $\mathbf{s}_n \in \mathbb{C}^K$ contains the source signals received at time n, considered as unknown deterministic ;

- $(\mathbf{v}_n)_{n\geq 1}$ is a temporally and spatially white complex Gaussian noise with spatial covariance $\mathbb{E}[\mathbf{v}_n\mathbf{v}_n^*] = \sigma^2\mathbf{I}$.

The received signal is observed between time 1 and time N, and we collect the available observations in the $M \times N$ matrix $\mathbf{Y}_N$ defined

$$\mathbf{Y}_N = [\mathbf{y}_1, \ldots, \mathbf{y}_N] = \mathbf{A}_M\mathbf{S}_N + \mathbf{V}_N, \tag{1}$$

with $\mathbf{S}_N = [\mathbf{s}_1, \ldots, \mathbf{s}_N]$ and $\mathbf{V}_N = [\mathbf{v}_1, \ldots, \mathbf{v}_N]$. We assume that $\mathrm{Rank}(\mathbf{S}_N) = K$ for each M, N greater than K. The DoA estimation problem consists in estimating the K DoA $\theta_1, \ldots, \theta_K$ from the matrix of samples $\mathbf{Y}_N$.

When the number of observations N is much less than the number of sensors M, the standard subspace method fails. In this case, it is standard to use spatial smoothing schemes in order to artificially increase the number of observations. In particular, it is well established that spatial smoothing schemes allow to use subspace methods even in the single snapshot case, i.e. when $N = 1$ (see e.g. [14] and the references therein). If $L < M$, spatial smoothing consists in considering L overlapping subarrays of dimension $M - L + 1$. At each time n, L snapshots of dimension $M - L + 1$ are thus available, and the scheme provides NL observations of dimension $M - L + 1$. In order to be more specific, we introduce the following notations. If L is an integer less than M, we denote by $\mathcal{Y}_n^{(L)}$ the $(M - L + 1) \times L$ Hankel matrix defined by

$$\mathcal{Y}_n^{(L)} = \begin{pmatrix} \mathbf{y}_{1,n} & \mathbf{y}_{2,n} & \cdots & \cdots & \mathbf{y}_{L,n} \\ \mathbf{y}_{2,n} & \mathbf{y}_{3,n} & \cdots & \cdots & \mathbf{y}_{L+1,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{y}_{M-L+1,n} & \mathbf{y}_{M-L+2,n} & \cdots & \cdots & \mathbf{y}_{M,n} \end{pmatrix} \tag{2}$$

Column l of matrix $\mathcal{Y}_n^{(L)}$ corresponds to the observation on subarray l at time n. Collecting all the observations on the various subarrays allows to obtain NL snapshots, thus increasing artificially the number of observations. We define $\mathbf{Y}_N^{(L)}$ as the $(M - L + 1) \times NL$ block-Hankel matrix given by

$$\mathbf{Y}_N^{(L)} = \left(\mathcal{Y}_1^{(L)}, \ldots, \mathcal{Y}_N^{(L)}\right) \tag{3}$$

In order to express $\mathbf{Y}_N^{(L)}$, we consider the $(M - L + 1) \times L$ Hankel matrix $\mathcal{A}^{(L)}(\theta)$ defined from vector $\mathbf{a}_M(\theta)$ in the same way than $\mathcal{Y}_n^{(L)}$. We remark that $\mathcal{A}^{(L)}(\theta)$ is rank 1, and can be written as

$$\mathcal{A}^{(L)}(\theta) = \sqrt{L(M - L + 1)/M}\, \mathbf{a}_{M-L+1}(\theta) \left(\mathbf{a}_L(\theta)\right)^T \tag{4}$$

We consider the $(M-L+1)\times KL$ matrix $\mathbf{A}^{(L)}$

$$\mathbf{A}^{(L)} = \Big(\mathcal{A}^{(L)}(\theta_1), \mathcal{A}^{(L)}(\theta_2), \ldots, \mathcal{A}^{(L)}(\theta_K)\Big) \tag{5}$$

which, of course, is a rank K matrix whose range coincides with the subspace generated by the $(M-L+1)$-dimensional vectors $\mathbf{a}_{M-L+1}(\theta_1), \ldots, \mathbf{a}_{M-L+1}(\theta_K)$. $\mathbf{Y}_N^{(L)}$ can be written as

$$\mathbf{Y}_N^{(L)} = \mathbf{A}^{(L)}\ (\mathbf{S}_N \otimes \mathbf{I}_L) + \mathbf{V}_N^{(L)} \tag{6}$$

where matrix $\mathbf{V}_N^{(L)}$ is the block Hankel matrix corresponding to the additive noise. As matrix $\mathbf{S}_N \otimes \mathbf{I}_L$ is full rank, the extended obervation matrix $\mathbf{Y}_N^{(L)}$ appears as a noisy version of a low rank component whose range is the K–dimensional subspace generated by vectors $\mathbf{a}_{M-L+1}(\theta_1), \ldots, \mathbf{a}_{M-L+1}(\theta_K)$. Moreover, it is easy to check that

$$\mathbb{E}\left(\frac{\mathbf{V}_N^{(L)}\mathbf{V}_N^{(L)*}}{NL}\right) = \sigma^2\mathbf{I}_{M-L+1}$$

Therefore, it is potentially possible to estimate the DoAs $(\theta_k)_{k=1,\ldots,K}$ using a subspace approach based on the eigenvalues / eigenvectors decomposition of matrix $\mathbf{Y}_N^{(L)}\mathbf{Y}_N^{(L)*}/NL$. The asymptotic behaviour of spatial smoothing subspace methods is standard in the regimes where $M-L+1$ remains fixed while NL converges towards ∞. This is due to the law of large numbers which implies that the empirical covariance matrix $\mathbf{Y}_N^{(L)}\mathbf{Y}_N^{(L)*}/NL$ has the same asymptotic behaviour than $\mathbf{A}^{(L)}\ (\mathbf{S}_N\mathbf{S}_N^* \otimes \mathbf{I}_L/NL)\ \mathbf{A}^{(L)*} + \sigma^2\mathbf{I}_{M-L+1}$, In this context, the orthogonal projection matrix $\hat{\mathbf{\Pi}}_N^{(L)}$ onto the eigenspace associated to the $M-L+1-K$ smallest eigenvalues of $\mathbf{Y}_N^{(L)}\mathbf{Y}_N^{(L)*}/NL$ is a consistent estimate of the orthogonal projection matrix $\mathbf{\Pi}^{(L)}$ on the noise subspace, i.e. the orthogonal complement of $\mathrm{sp}\{\mathbf{a}_{M-L+1}(\theta_1), \ldots, \mathbf{a}_{M-L+1}(\theta_K)\}$. In other words, it holds that

$$\left\|\hat{\mathbf{\Pi}}_N^{(L)} - \mathbf{\Pi}^{(L)}\right\| \to 0\ a.s. \tag{7}$$

The traditional pseudo-spectrum estimate $\hat{\eta}_N^{(t)}(\theta)$ defined by

$$\hat{\eta}_N^{(t)}(\theta) = \mathbf{a}_{M-L+1}(\theta)^*\hat{\mathbf{\Pi}}_N^{(L)}\mathbf{a}_{M-L+1}(\theta)$$

thus verifies

$$\sup_{\theta\in[-\pi,\pi]} \left|\hat{\eta}_N^{(t)}(\theta) - \eta(\theta)\right| \xrightarrow[N\to\infty]{a.s.} 0. \tag{8}$$

where $\eta(\theta) = \mathbf{a}_{M-L+1}(\theta)^*\mathbf{\Pi}^{(L)}\mathbf{a}_{M-L+1}(\theta)$ is the MUSIC pseudo-spectrum. Moreover, the K MUSIC traditional DoA estimates, defined formally, for $k = 1, \ldots, K$, by

$$\hat{\theta}_{k,N}^{(t)} = \operatorname*{argmin}_{\theta\in\mathcal{I}_k} \hat{\eta}_N^{(t)}(\theta), \tag{9}$$

where $\mathcal{I}_k$ is a compact interval containing θ_k and such that $\mathcal{I}_k \cap \mathcal{I}_l = \emptyset$ for $k \neq l$, are consistent, i.e.

$$\hat{\theta}_{k,N}^{(t)} \xrightarrow[N\to\infty]{a.s.} \theta_k. \tag{10}$$

However, the regime where $M - L + 1$ remains fixed while NL converges towards ∞ is not very interesting in practice because the size $M-L+1$ of the subarrays may be much smaller that the number of antennas M, thus reducing the resolution of the method. We therefore study spatial smoothing schemes in regimes where the dimensions $M-L+1$ and NL of matrix $\mathbf{Y}_N^{(L)}$ are of the same order of magnitude and where $\frac{L}{M} \to 0$ in order to keep unchanged the aperture of the array. More precisely, we assume that integers N and L depend on M and that

$$M \to +\infty, N = \mathcal{O}(M^\beta),\ \frac{1}{3} < \beta \leq 1,\ c_N = \frac{M-L+1}{NL} \to c_* \tag{11}$$

In regime (11), N thus converges towards ∞ but at a rate that may be much lower than M thus modelling contexts in which N is much smaller than M. As $N \to +\infty$, it also holds that $\frac{M}{NL} \to c_*$. Therefore, it is clear that $L = \mathcal{O}(M^\alpha)$ where $\alpha = 1-\beta$ verifies with $0 \leq \alpha < 2/3$. L may thus converge towards ∞ (even faster than N if $\beta < 1/2$) but in such a way that $\frac{L}{M} \to 0$. As in regime (11) N depends on M, it could be appropriate to index the various matrices and DoA estimators by integer M rather than by integer N as in definitions (5) and (9). However, we prefer to use the index N in the following in order to keep the notations unchanged. We also denote projection matrix $\mathbf{\Pi}^{(L)}$ and pseudo-spectrum $\eta(\theta)$ by $\mathbf{\Pi}_N^{(L)}$ and $\eta_N(\theta)$ because they depend on M. Moreover, in the following, the notation $N \to +\infty$ should be understood as regime (11) for some $\beta \in (1/3, 1]$.

B. Main results.

In regime (11), (7) is no more valid. Hence, (10) is questionable. In this paper, we show that it is possible to generalize the G-MUSIC estimators introduced in [5] in the case where $L = 1$ to the context of spatial smoothing schemes in regime (11) and establish the following results. Under the separation condition that the K non zero eigenvalues of matrix $\frac{1}{L}\mathbf{A}^{(L)}\left(\frac{\mathbf{S}_N\mathbf{S}_N^*}{N} \otimes \mathbf{I}_L\right)\mathbf{A}^{(L)*}$ are above the threshold $\sigma^2\sqrt{c_*}$ for each N large enough, we deduce from [18] that:

- the spatial smoothing traditional MUSIC estimates $(\hat{\theta}_{k,N}^{(t)})_{k=1,\ldots,K}$ and the G-MUSIC SS estimates, denoted $(\hat{\theta}_{k,N})_{k=1,\ldots,K}$ are consistent and verify

$$M(\hat{\theta}_{k,N}^{(t)} - \theta_k) \ \to\ 0\ a.s., \tag{12}$$

$$M(\hat{\theta}_{k,N} - \theta_k) \ \to\ 0\ a.s. \tag{13}$$

(12) and (13) hold when the DoA $(\theta_k)_{k=1,\ldots,K}$ do not scale with M, N. In pratice, this assumption corresponds to practical situations where the DoA are widely spaced because when the DoA $(\theta_k)_{k=1,\ldots,K}$ are fixed, the ratio

$$\frac{\min_{k\neq l} |\theta_k - \theta_l|}{(2\pi)/M}$$

converges towards ∞. We deduce from [18] that:

- If $K = 2$ and that the 2 DoAs scale with M, N is such a way that $\theta_{2,N} - \theta_{1,N} = \mathcal{O}(\frac{1}{M})$, then the G-MUSIC SS estimates still verify (13) while the traditional MUSIC SS estimates no longer verify (12)

As in the case $L = 1$, the above mentioned separation condition ensures that the K largest eigenvalues of the empirical covariance matrix $(\mathbf{Y}_N^{(L)}\mathbf{Y}_N^{(L)*})/NL$ correspond to the K sources, and the signal and noise subspaces can be separated. In order to obtain some insights on this condition, and on the potential benefit of the spatial smoothing, we study the separation condition when M and N converge towards ∞ at the same rate, i.e. when $\frac{M}{N} \to d_*$, or equivalently that $\beta = 1$ and that L does not scale with N. In this case, it is clear that c_* coincides with $c_* = d_*/L$. Under the assumption that $\frac{\mathbf{S}_N\mathbf{S}_N^*}{N}$ converges towards a diagonal matrix $\mathbf{D}$ when N increases, then we establish that the separation condition holds if

$$\lambda_K \left(\mathbf{A}_{M-L+1}^*\mathbf{A}_{M-L+1}\mathbf{D}\right) > \frac{\sigma^2\sqrt{d_*}}{\sqrt{L}} \tag{14}$$

for each (M, N) large enough. If $L = 1$, the separation condition introduced in the context of (unsmoothed) G-MUSIC algorithms ([5]) is of course recovered, i.e.

$$\lambda_K \left(\mathbf{A}_M^*\mathbf{A}_M\mathbf{D}\right) > \sigma^2\sqrt{d_*}$$

If M is large and that $L << M$, matrix $\mathbf{A}_{M-L+1}^*\mathbf{A}_{M-L+1}$ is close from $\mathbf{A}_M^*\mathbf{A}_M$ and the separation condition is nearly equivalent to

$$\lambda_K \left(\mathbf{A}_M^*\mathbf{A}_M\mathbf{D}\right) > \frac{\sigma^2\sqrt{d_*}}{\sqrt{L}}$$

Therefore, it is seen that the use of the spatial smoothing scheme allows to reduce the threshold $\sigma^2\sqrt{d_*}$ corresponding to G-MUSIC method without spatial smoothing by the factor $\sqrt{L}$. Therefore, if M and N are the same order of magnitude, our asymptotic analysis allows to predict an improvement of the performance of the G-MUSIC SS methods when L increases provided $L << M$. If L becomes too large, the above rough analysis is no more justified and the impact of the diminution of the number of antennas becomes dominant, and the performance tends to decrease.

III. Asymptotic behaviour of the largest singular values and corresponding singular vectors of finite rank perturbations of certain large random block-Hankel matrices.

In this section, N, M, L still satisfy (11) while K is a fixed integer that does not scale with N. We consider the $(M+L-1)\times NL$ block-Hankel random matrix $\mathbf{V}_N^{(L)}$ defined previously, and introduce matrix $\mathbf{Z}_N$ defined

$$\mathbf{Z}_N = \frac{1}{\sqrt{NL}}\mathbf{V}_N^{(L)}$$

in order to simplify the notations. The entries of $\mathbf{Z}_N$ have of course variance σ^2/NL. In the following, $\mathbf{B}_N$ represents a deterministic $(M+L-1)\times NL$ matrix verifying

$$\sup_N \|\mathbf{B}_N\| < +\infty,\ \mathrm{Rank}(\mathbf{B}_N) = K, \tag{15}$$

for each N large enough. We denote by $\lambda_{1,N} > \lambda_{2,N} \ldots > \lambda_{K,N}$ the non zero eigenvalues of matrix $\mathbf{B}_N\mathbf{B}_N^*$ arranged in decreasing order, and by $(\mathbf{u}_{k,N})_{k=1,\ldots,K}$ and $(\tilde{\mathbf{u}}_{k,N})_{k=1,\ldots,K}$ the associated left and right singular vectors of $\mathbf{B}_N$. The singular value decomposition of $\mathbf{B}_N$ is thus given by

$$\mathbf{B}_N = \sum_{k=1}^{K} \lambda_{k,N}^{1/2}\mathbf{u}_{k,N}\tilde{\mathbf{u}}_{k,N}^* = \mathbf{U}_N\mathbf{\Lambda}_N^{1/2}\tilde{\mathbf{U}}_N^*$$

Moreover, we assume that:

Assumption 1. *The K non zero eigenvalues $(\lambda_{k,N})_{k=1,\ldots,K}$ of matrix $\mathbf{B}_N\mathbf{B}_N^*$ converge towards $\lambda_1 > \lambda_2 > \ldots > \lambda_K$ when $N \to +\infty$.*

Here, for ease of exposition, we assume that the eigenvalues $(\lambda_{k,N})_{k=1,\ldots,K}$ have multiplicity 1 and that $\lambda_k \neq \lambda_l$ for $k \neq l$. However, the forthcoming results can be easily adapted if some λ_k coincide.

We define matrix $\mathbf{X}_N$ as

$$\mathbf{X}_N = \mathbf{B}_N + \mathbf{Z}_N \tag{16}$$

$\mathbf{X}_N$ can thus be interpreted as a rank K perturbation of the block-Hankel matrix $\mathbf{Z}_N$. The purpose of this section is to study the behaviour of the K largest eigenvalues $(\hat{\lambda}_{k,N})_{k=1,\ldots,K}$ of matrix $\mathbf{X}_N\mathbf{X}_N^*$ as well as of their corresponding eigenvectors $(\hat{\mathbf{u}}_{k,N})_{k=1,\ldots,K}$. It turns out that $(\hat{\lambda}_{k,N})_{k=1,\ldots,K}$ and $(\hat{\mathbf{u}}_{k,N})_{k=1,\ldots,K}$ behave as if the entries of matrix $\mathbf{Z}_N$ where i.i.d. To see this, we have first to precise the behaviour of the eigenvalues of matrix $\mathbf{Z}_N\mathbf{Z}_N^*$ in the asymptotic regime (11).

A. Behaviour of the eigenvalues of matrix $\mathbf{Z}_N\mathbf{Z}_N^*$.

We first recall the definition of the Marcenko-Pastur distribution $\mu_{\sigma^2,c}$ of parameters σ^2 and c (see e.g. [1]). $\mu_{\sigma^2,c}$ is the probability distribution defined by

$$d\mu_{\sigma^2,c}(x) = \delta_0[1-c^{-1}]_+ + \frac{\sqrt{(x-x^-)\,(x^+-x)}}{2\sigma^2 c\pi x}\mathbb{1}_{[x^-,x^+]}(x)\,dx$$

with $x^- = \sigma^2(1-\sqrt{c})^2$ and $x^+ = \sigma^2(1+\sqrt{c})^2$. Its Stieltjes transform $m_{\sigma^2,c}(z)$ defined by

$$m_{\sigma^2,c}(z) = \int_{\mathrm{R}} \frac{\mathrm{d}\mu_{\sigma^2,c}(\lambda)}{\lambda - z}$$

is known to satisfy the fundamental equation

$$m_{\sigma^2,c}(z) = \frac{1}{-z + \sigma^2 \frac{1}{1+\sigma^2 c m_{\sigma^2,c}(z)}} \tag{17}$$

or equivalently,

$$m_{\sigma^2,c}(z) = \frac{1}{-z(1+\sigma^2\tilde{m}_{\sigma^2,c}(z))} \tag{18}$$

$$\tilde{m}_{\sigma^2,c}(z) = \frac{1}{-z(1+\sigma^2 c m_{\sigma^2,c}(z))} \tag{19}$$

where $\tilde{m}_{\sigma^2,c}(z)$ is known to coincide with Stieltjes transform of the Marcenko-Pastur distribution $\mu_{\sigma^2 c,c^{-1}} = c\mu_{\sigma^2,c} + (1-c)\delta_0$.

In order to simplify the notations, we denote by $m_*(z)$ and $\tilde{m}_*(z)$ the Stieltjes transforms of Marcenko-Pastur distributions μ_{σ^2,c_*} and $\mu_{\sigma^2 c_*,c_*^{-1}}$. $m_*(z)$ and $\tilde{m}_*(z)$ verify Equations (18) and (19) for $c = c_*$. We also denote by x_*^- and x_*^+ the terms $x_*^- = \sigma^2(1-\sqrt{c_*})^2$ and $x_*^+ = \sigma^2(1+\sqrt{c_*})^2$. We recall that function $w_*(z)$ defined by

$$w_*(z) = \frac{1}{z\,m_*(z)\,\tilde{m}_*(z)} \tag{20}$$

is analytic on $\mathbb{C} - [x_*^-, x_*^+]$, verifies $w_*(x_*^+) = \sigma^2\sqrt{c_*}$, and increases from $\sigma^2\sqrt{c_*}$ to $+\infty$ when x increases from x_*^+ to $+\infty$ (see [3], section 3.1). Moreover, if $\phi_*(w)$ denotes function defined by

$$\phi_*(w) = \frac{(w+\sigma^2)(w+\sigma^2 c_*)}{w} \tag{21}$$

then, ϕ_* increases from x_*^+ to $+\infty$ when w increases from $\sigma^2\sqrt{c_*}$ to $+\infty$. Finally, it holds that

$$\phi_*(w_*(z)) = z \tag{22}$$

for each $z \in \mathbb{C} - [x_*^-, x_*^+]$.

We denote by $\mathbf{Q}_N(z)$ and $\tilde{\mathbf{Q}}_N(z)$ the so-called resolvent of matrices $\mathbf{Z}_N\mathbf{Z}_N^*$ and $\mathbf{Z}_N^*\mathbf{Z}_N$ defined by

$$\mathbf{Q}_N(z) = (\mathbf{Z}_N\mathbf{Z}_N^* - z\mathbf{I}_{M-L+1})^{-1},\ \tilde{\mathbf{Q}}_N(z) = (\mathbf{Z}_N^*\mathbf{Z}_N - z\mathbf{I}_{NL})^{-1}$$

Then, the results of [8] imply the following proposition.

Proposition 1.

- *(i) The eigenvalue distribution of matrix $\mathbf{Z}_N\mathbf{Z}_N^*$ converges almost surely towards the Marcenko-Pastur distribution μ_{σ^2,c_*}, or equivalently, for each $z \in \mathbb{C} - \mathbb{R}^+$,*

$$\frac{1}{M-L+1}\mathrm{Tr}(\mathbf{Q}_N(z)) - m_*(z) \to 0\, a.s. \tag{23}$$

- *(ii) For each $\epsilon > 0$, almost surely, for N large enough, all the eigenvalues of $\mathbf{Z}_N\mathbf{Z}_N^*$ belong to $[\sigma^2(1-\sqrt{c_*})^2 - \epsilon, \sigma^2(1+\sqrt{c_*})^2 + \epsilon]$ if $c_* \leq 1$, and to $[\sigma^2(1-\sqrt{c_*})^2 - \epsilon, \sigma^2(1+\sqrt{c_*})^2 + \epsilon] \cup \{0\}$ if $c_* > 1$.*
- *(iii) Moreover, if $\mathbf{a}_N, \mathbf{b}_N$ are $(M-L+1)$–dimensional deterministic vectors satisfying $\sup_N(\|\mathbf{a}_N\|, \|\mathbf{b}_N\|) < +\infty$, then it holds that for each $z \in \mathbb{C}^+$*

$$\mathbf{a}_N^* \left(\mathbf{Q}_N(z) - m_*(z)\mathbf{I}\right) \mathbf{b}_N \to 0\ a.s. \tag{24}$$

Similarly, if $\tilde{\mathbf{a}}_N$ and $\tilde{\mathbf{b}}_N$ are NL–dimensional deterministic vectors verifying $\sup_N(\|\tilde{\mathbf{a}}_N\|, \|\tilde{\mathbf{b}}_N\|) < +\infty$, then for each $z \in \mathbb{C}^+$, it holds that

$$\tilde{\mathbf{a}}_N^* \left(\tilde{\mathbf{Q}}_N(z) - \tilde{m}_*(z)\mathbf{I}\right) \tilde{\mathbf{b}}_N \to 0\ a.s. \tag{25}$$

Moreover, for each $z \in \mathbb{C}^+$, it holds that

$$\mathbf{a}_N^* \left(\mathbf{Q}_N(z)\mathbf{Z}_N\right) \mathbf{b}_N \to 0\ a.s. \tag{26}$$

Finally, for each $\epsilon > 0$, convergence properties (24, 25, 26) hold uniformly w.r.t. z on each compact subset of $\mathbb{C} - [0, x_*^+ + \epsilon]$.

The proof is given in the Appendix.

Remark 1. *Proposition 1 implies that in a certain sense, matrix $\mathbf{Z}_N\mathbf{Z}_N^*$ behaves as if the entries of $\mathbf{Z}_N$ were i.i.d because Proposition 1 is known to hold for i.i.d. matrices. In the i.i.d. case, (23) was established for the first time in [9], the almost sure location of the eigenvalues of $\mathbf{Z}_N\mathbf{Z}_N^*$ can be found in [1] (see Theorem 5-11), while (24), (25) and (26) are trivial modifications of Lemma 5 of [5].*

We notice that the convergence towards the Marcenko-Pastur distribution holds as soon as $N \to +\infty$ and $\frac{M-L+1}{NL} \to c_*$. In particular, the convergence is still valid if $N = \mathcal{O}(M^\beta)$ for each $0 < \beta \leq 1$, or equivalently if $L = \mathcal{O}(M^\alpha)$ for each $0 \leq \alpha < 1$. L can therefore converge towards ∞ much faster than N. However, the hypothesis that $\beta > 1/3$, which is also equivalent to $L = \mathcal{O}(M^\alpha)$ with $\alpha < 2/3$, is necessary to establish item (ii).

B. The K largest eigenvalues and eigenvectors of $\mathbf{X}_N\mathbf{X}_N^*$.

While matrix $\mathbf{Z}_N$ does not meet the conditions formulated in [3], Proposition 1 allows to use the approach used in [3], and to prove that the K largest eigenvalues and corresponding eigenvectors of $\mathbf{X}_N\mathbf{X}_N^*$. behave as if the entries of $\mathbf{Z}_N$ were i.i.d. In particular, the following result holds.

Theorem 1. *We denote by s, $0 \leq s \leq K$, the largest integer for which*

$$\lambda_s > \sigma^2\sqrt{c_*} \tag{27}$$

Then, for $k = 1, \ldots, s$, it holds that

$$\hat{\lambda}_{k,N} \xrightarrow[N\to\infty]{a.s.} \rho_k = \phi(\lambda_k) = \frac{(\lambda_k + \sigma^2)(\lambda_k + \sigma^2 c)}{\lambda_k} > x_*^+. \tag{28}$$

Moreover, for $k = s+1, \ldots, K$, it holds that

$$\hat{\lambda}_{k,N} \to \sigma^2(1 + \sqrt{c_*})^2 \, a.s. \tag{29}$$

Finally, for all deterministic sequences of unit norm vectors $(\mathbf{d}_{1,N})$, $(\mathbf{d}_{2,N})$, we have for $k = 1, \ldots, s$

$$\mathbf{d}_{1,N}^*\hat{\mathbf{u}}_{k,N}\hat{\mathbf{u}}_{k,N}^*\mathbf{d}_{2,N} = h_*(\rho_k)\mathbf{d}_{1,N}^*\mathbf{u}_{k,N}\mathbf{u}_{k,N}^*\mathbf{d}_{2,N} + o(1) \quad a.s., \tag{30}$$

where function $h_*(z)$ is defined by

$$h_*(z) = \frac{w_*(z)^2 - \sigma^4 c_*}{w_*(z)(w_*(z) + \sigma^2 c_*)} \tag{31}$$

For the reader's convenience, we provide in the appendix some insights on the approach developed in [3] to prove (28) and (29). For more details on (30), see the proof of Theorem 2 in [5] as well as the identity

$$h_*(z) = \frac{z m_*(z)^2 \tilde{m}_*(z)}{(z m_*(z)\tilde{m}_*(z))'}$$

where $'$ represents the derivation w.r.t. z.

IV. Derivation of a consistent G-MUSIC method.

We now use the results of section III for matrix $\mathbf{X}_N = \mathbf{Y}_N^{(L)}/\sqrt{NL}$ and $\mathbf{B}_N = \frac{1}{\sqrt{NL}}\mathbf{A}^{(L)}(\mathbf{S}_N \otimes \mathbf{I}_L)$. We recall that $(\hat{\lambda}_{k,N})_{k=1,\ldots,M-L+1}$ and $(\hat{\mathbf{u}}_{k,N})_{k=1,\ldots,M-L+1}$ represent the eigenvalues and eigenvectors of the empirical covariance matrix $\mathbf{Y}_N^{(L)}\mathbf{Y}_N^{(L)*}/NL$, and that $(\lambda_{k,N})_{k=1,\ldots,K}$ and $(\mathbf{u}_{k,N})_{k=1,\ldots,K}$ are the non zero eigenvalues and corresponding eigenvectors of $\frac{1}{L}\mathbf{A}^{(L)}\left(\mathbf{S}_N\mathbf{S}_N^*/N \otimes \mathbf{I}_L\right)\mathbf{A}^{(L)*}$. We recall that $\mathbf{\Pi}_N^{(L)}$ represents the orthogonal projection matrix onto the noise subspace, i.e. the orthogonal complement

of the space generated by vectors $(\mathbf{a}_{M-L+1}(\theta_k))_{k=1,\ldots,K}$ and that $\eta_N(\theta)$ is the corresponding MUSIC pseudo-spectrum

$$\eta_N(\theta) = \mathbf{a}_{M-L+1}(\theta)^* \, \mathbf{\Pi}_N^{(L)} \, \mathbf{a}_{M-L+1}(\theta)$$

Theorem 1 allows to generalize immediately the results of [5] and [18] concerning the consistency of G-MUSIC and MUSIC DoA estimators in the case $L = 1$. More precisely:

Theorem 2. *Assume that the K non zero eigenvalues $(\lambda_{k,N})_{k=1,\ldots,K}$ converge towards deterministic terms $\lambda_1 > \lambda_2 > \ldots > \lambda_K$ and that*

$$\lambda_K > \sigma^2 \sqrt{c_*} \tag{32}$$

Then, the estimator $\hat{\eta}_N(\theta)$ of the pseudo-spectrum $\eta_N(\theta)$ defined by

$$\hat{\eta}_N(\theta) = (\mathbf{a}_{M-L+1}(\theta))^* \left(\mathbf{I} - \sum_{k=1}^{K} \frac{1}{h\left(\hat{\lambda}_{k,N}\right)} \hat{\mathbf{u}}_{k,N} \hat{\mathbf{u}}_{k,N}^* \right) \mathbf{a}_{M-L+1}(\theta) \tag{33}$$

verifies

$$\sup_{\theta \in [-\pi,\pi]} |\hat{\eta}_N(\theta) - \eta_N(\theta)| \xrightarrow[N\to\infty]{a.s.} 0, \tag{34}$$

This result can be proved as Proposition 1 in [5].

In order to obtain some insights on condition (32) and on the potential benefits of the spatial smoothing, we explicit the separation condition (32) when M and N converge towards ∞ at the same rate, i.e. when $\frac{M}{N} \to d_*$, or equivalently that $\beta = 1$ and that L does not scale with N. In this case, it is clear that c_* coincides with $c_* = d_*/L$. It is easily seen that

$$\frac{1}{L} \mathbf{A}^{(L)} \left(\frac{\mathbf{S}_N \mathbf{S}_N^*}{N} \otimes \mathbf{I}_L \right) \mathbf{A}^{(L)*} = (M - L + 1/M) \; \mathbf{A}_{M-L+1} \left(\frac{\mathbf{S}_N \mathbf{S}_N^*}{N} \bullet \mathbf{A}_L^T \overline{\mathbf{A}}_L \right) \mathbf{A}_{M-L+1}^* \tag{35}$$

where $\bullet$ represents the Hadamard (i.e. element wise) product of matrices, and where $\overline{\mathbf{B}}$ stands for the complex conjugation operator of the elements of matrix $\mathbf{B}$. If we assume that $\frac{\mathbf{S}_N \mathbf{S}_N^*}{N}$ converges towards a diagonal matrix $\mathbf{D}$ when N increases, then $\frac{\mathbf{S}_N \mathbf{S}_N^*}{N} \bullet (\mathbf{A}_L^T \overline{\mathbf{A}}_L)$ converges towards the diagonal matrix $\mathbf{D} \bullet \mathrm{Diag}\left(\mathbf{A}_L^T \overline{\mathbf{A}}_L\right) = \mathbf{D}$. Therefore, $\frac{\mathbf{S}_N \mathbf{S}_N^*}{N} \bullet (\mathbf{A}_L^T \overline{\mathbf{A}}_L) \simeq \mathbf{D}$ when is large enough. Using that $\frac{L}{M} \to 0$, we obtain that the separation condition is nearly equivalent to

$$\lambda_K \left(\mathbf{A}_{M-L+1} \mathbf{D} \; \mathbf{A}_{M-L+1}^* \right) > \frac{\sigma^2 \sqrt{d_*}}{\sqrt{L}}$$

or to

$$\lambda_K \left(\mathbf{A}_{M-L+1}^* \mathbf{A}_{M-L+1} \mathbf{D} \right) > \frac{\sigma^2 \sqrt{d_*}}{\sqrt{L}} \tag{36}$$

for each (M,N) large enough. If $L=1$, the separation condition introduced in the context of (unsmoothed) G-MUSIC algorithms ([5]) is of course recovered, i.e.

$$\lambda_K\left(\mathbf{A}_M^*\mathbf{A}_M\mathbf{D}\right) > \sigma^2\sqrt{d_*}$$

for each (M,N) large enough. If M is large and that $L << M$, matrix $\mathbf{A}_{M-L+1}^*\mathbf{A}_{M-L+1}$ is close from $\mathbf{A}_M^*\mathbf{A}_M$ and the separation condition is nearly equivalent to

$$\lambda_K\left(\mathbf{A}_M^*\mathbf{A}_M\mathbf{D}\right) > \frac{\sigma^2\sqrt{d_*}}{\sqrt{L}}$$

Therefore, it is seen that the use of the spatial smoothing scheme allows to reduce the threshold $\sigma^2\sqrt{d_*}$ corresponding to G-MUSIC method without spatial smoothing by the factor $\sqrt{L}$. Hence, if M and N are the same order of magnitude, our asymptotic analysis allows to predict an improvement of the performance of the G-MUSIC methods based on spatial smoothing when L increases provided $L << M$. If L becomes too large, the above rough analysis is no more justified and the impact of the diminution of the number of antennas becomes dominant, and the performance tends to decrease. This analysis is sustained by the numerical simulations presented in section V.

We define the DoA G-MUSIC SS estimates $(\hat{\theta}_{k,N})_{k=1,\ldots,K}$ by

$$\hat{\theta}_{k,N} = \underset{\theta\in\mathcal{I}_k}{\operatorname{argmin}} \left|\hat{\eta}_N(\theta)\right|, \tag{37}$$

where $\mathcal{I}_k$ is a compact interval containing θ_k and such that $\mathcal{I}_k \cap \mathcal{I}_l = \emptyset$ for $k \neq l$. As in [5], (34) as well as the particular structure of directional vectors $\mathbf{a}_{M-L+1}(\theta)$ imply the following result which can be proved as Theorem 3 of [5]

Theorem 3. *Under condition (32), the DoA G-MUSIC SS estimates $(\hat{\theta}_{k,N})_{k=1,\ldots,K}$ verify*

$$M\left(\hat{\theta}_{k,N}-\theta_k\right) \to 0 \ a.s. \tag{38}$$

for each $k=1,\ldots,K$.

Remark 2. *We remark that under the extra assumption that $\frac{\mathbf{S}_N\mathbf{S}_N^*}{N}$ converges towards a diagonal matrix,[5] (see also [19] for more general matrices $\mathbf{S}$) proved when $L=1$ that $M^{3/2}\left(\hat{\theta}_{k,N}-\theta_k\right)$ converges in distribution towards a Gaussian distribution. It would be interesting to generalize the results of [5] and [19] to the G-MUSIC estimators with spatial smoothing in the asymptotic regime (11). This is a difficult task that is not within the scope of the present paper.*

Theorem 1 also allows to generalize immediately the results of [18] concerning the consistency of the traditional estimates $(\hat{\theta}_{k,N}^{(t)})_{k=1,\ldots,K}$ in the case $L=1$. In particular, while the traditional estimate $\hat{\eta}_N^{(t)}$ of the pseudo-spectrum is not consistent, it is shown in [18] that if $L=1$, then the arguments of its local minima $(\hat{\theta}_{k,N}^{(t)})_{k=1,\ldots,K}$ are consistent and verify

$$M\left(\hat{\theta}_{k,N}^{(t)}-\theta_k\right)\to 0\ a.s. \tag{39}$$

for each $k=1,\ldots,K$ if the separation condition is verified. The reader can check that Theorem 1 allows to generalize immediately this behaviour to the traditional DoA MUSIC estimates with spatial smoothing in regime (11). More precisely, the following result holds.

Theorem 4. *Under condition (32), the DoA traditional MUSIC SS estimates $(\hat{\theta}_{k,N}^{(t)})_{k=1,\ldots,K}$ verify*

$$M\left(\hat{\theta}_{k,N}^{(t)}-\theta_k\right)\to 0\ a.s. \tag{40}$$

for each $k=1,\ldots,K$.

Remark 3. *It is established in [18] in the case $L=1$ that if $\frac{\mathbf{S}_N\mathbf{S}_N^*}{N}$ converges towards a diagonal matrix, then $M^{3/2}\left(\hat{\theta}_{k,N}^{(t)}-\theta_k\right)$ has a Gaussian behaviour, and that the corresponding variance coincides with the asymptotic variance of $M^{3/2}\left(\hat{\theta}_{k,N}-\theta_k\right)$. In particular, if $L=1$, the asymptotic performance of MUSIC and G-MUSIC estimators coincide. It would be interesting to check whether this result still holds true for the MUSIC and G-MUSIC estimators with spatial smoothing.*

Theorems 2 and 3 as well as (39) assume that the DoAs $(\theta_k)_{k=1,\ldots,K}$ are fixed parameters, i.e. do not scale with M. Therefore, the ratio

$$\frac{\min_{k\neq l}|\theta_k-\theta_l|}{(2\pi)/M}$$

converges towards $+\infty$. In practice, this context is thus able to model practical situations in which $\sup_{k\neq l}|\theta_k-\theta_l|$ is significantly larger than the aperture of the array. In the case $L=1$, [18] also addressed the case where the DoA's $(\theta_{k,N})_{k=1,\ldots,K}$ depend on N,M and verify $\theta_{k,N}-\theta_{l,N}=\mathcal{O}(\frac{1}{M})$. This context allows to capture practical situations in which the DoA's are spaced of the order of a beamwidth. In order to simplify the calculations, [18] considered the case $K=2$, $\theta_{2,N}=\theta_{1,N}+\frac{\alpha}{N}$ and where matrix $\frac{\mathbf{S}_N\mathbf{S}_N^*}{N}\to\mathbf{I}_2$. However, the results can be generalized easily to more general situations. It is shown in [18] that the G-MUSIC estimates still verifiy (40), but that, in general, $M\left(\hat{\theta}_{k,N}^{(t)}-\theta_k\right)$ does not converge towards 0. The results of [18] can be generalized immediately to the context of G-MUSIC estimators with spatial smoothing in regime (11). For this, we have to assume that $\theta_{2,N}=\theta_{1,N}+\frac{\kappa}{M}$

(in [18], M and N are of the same order of magnitude so that the assumptions $\theta_{2,N} = \theta_{1,N} + \frac{\alpha}{N}$ and $\theta_{2,N} = \theta_{1,N} + \frac{\kappa}{M}$ are equivalent), and to follow the arguments of section 4 in [18]. The conclusion of this discussion is the following Theorem.

Theorem 5. *Assume $K = 2$, $\theta_{2,N} = \theta_{1,N} + \frac{\kappa}{M}$, and that $\frac{\mathbf{S}_N \mathbf{S}_N^*}{N} \to \mathbf{I}_2$. If the separation condition*

$$1 - |\operatorname{sinc} \kappa/2| > \sigma^2 c_* \tag{41}$$

holds, then the G-MUSIC SS estimates $(\hat{\theta}_{k,N})_{k=1,2}$ defined by

$$\hat{\theta}_{k,N} = \underset{\theta \in \mathcal{I}_{k,N}}{\operatorname{argmin}} \left| \hat{\eta}_N(\theta) \right|, \tag{42}$$

where $\mathcal{I}_{k,N} = [\theta_{k,N} - \frac{\kappa-\epsilon}{2N}, \theta_{k,N} + \frac{\kappa-\epsilon}{2N}]$ for ϵ small enough, verify

$$M \left(\hat{\theta}_{k,N} - \theta_{k,N} \right) \to 0 \ a.s. \tag{43}$$

In general, the traditional MUSIC SS estimates defined by (42) when the G-MUSIC estimate $\hat{\eta}_N(\theta)$ is replaced by the traditional spectrum estimate $\hat{\eta}_N^{(t)}(\theta)$ are such that $M \left(\hat{\theta}_{k,N}^{(t)} - \theta_{k,N} \right)$ does not converge towards 0.

V. Numerical examples

In this section, we provide numerical simulations illustrating the results given in the previous sections. We first consider 2 closely spaced sources with DoAs $\theta_1 = 0$ and $\theta_2 = \frac{\pi}{2M}$, and we assume that $M = 160$ and $N = 20$. The $2 \times N$ signal matrix is obtained by normalizing a realization of a random matrix with $\mathcal{N}_{\mathbb{C}}(0, 1)$ i.i.d. entries in such a way that the 2 source signals have power 1. The signal to noise ratio is thus equal to $\mathrm{SNR} = 1/\sigma^2$. Table I provides the minimum value of SNR for which the separation condition, in its finite length version (i.e. when the limits $(\lambda_k)_{k=1,\ldots,K}$ and c_* are replaced by $(\lambda_{k,N})_{k=1,\ldots,K}$ and c_N respectively) holds, i.e.

$$(\sigma^2)^{-1} = \frac{1}{\lambda_{K,N}} \sqrt{(M - L + 1)/NL}$$

It is seen that the minimal SNR first decreases but that it increases if L is large enough. This confirms the discussion of the previous section on the effect of L on the separation condition.

In figure 1, we represent the mean-square errors of the G-MUSIC SS estimator $\hat{\theta}_1$ for $L = 2, 4, 8, 16$ versus SNR. The corresponding Cramer-Rao bounds is also represented. As expected, it is seen that the performance tends to increase with L until $L = 16$. In figure 2, L is equal to 16, 32, 64, 96, 128.

For $L = 32$, it is seen that the MSE tends to degrade at high SNR w.r.t. $L = 16$, while the performance severely degrades for larger values of L.

L	2	4	8	16	32	64	96	128
SNR	33.46	30.30	27.46	25.31	24.70	28.25	36.11	51.52

TABLE I: Minimum value of SNR for separation condition

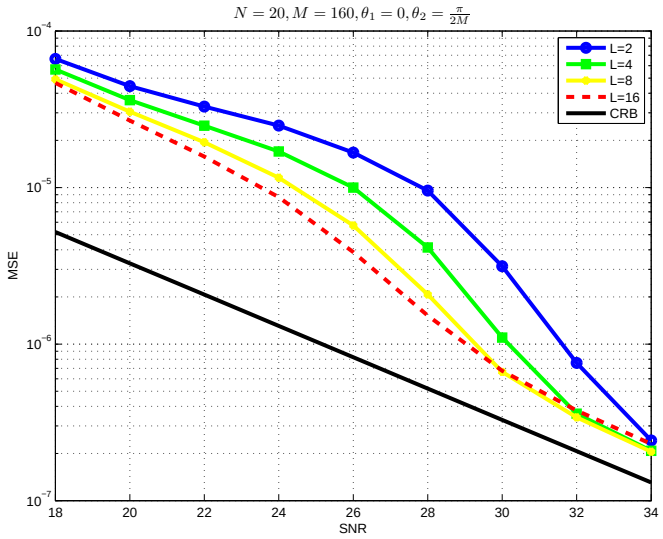

Fig. 1: Empirical MSE of G-MUSIC SS estimator $\hat{\theta}_1$ versus SNR

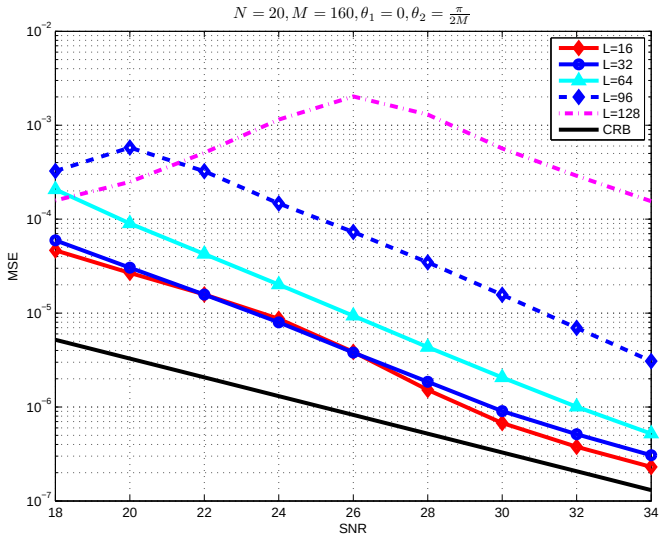

Fig. 2: Empirical MSE of G-MUSIC SS estimator $\hat{\theta}_1$ versus SNR

In Figure 3, parameter L is equal to 16. We compare the performance of G-MUSIC SS with the standard MUSIC method with spatial smoothing. We also represent the MSE provided by G-MUSIC and MUSIC for $L = 1$. The standard unsmoothed MUSIC method of course completely fails, while the use of the G-MUSIC SS provides a clear improvement of the performance w.r.t. MUSIC SS and unsmoothed G-MUSIC.

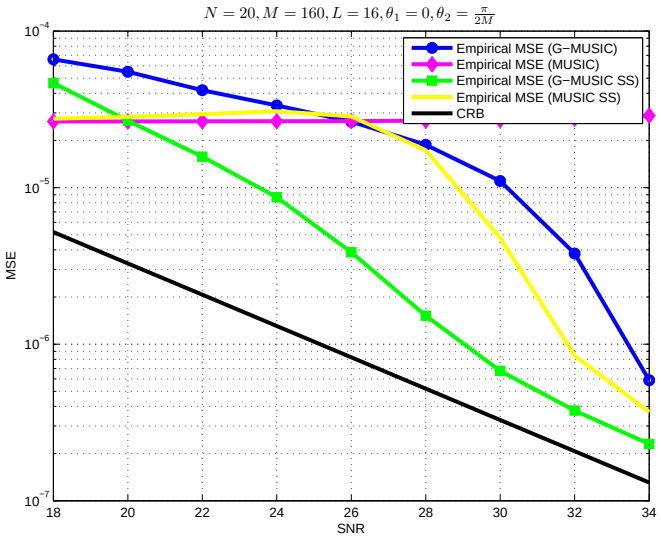

Fig. 3: Empirical MSE of different estimators of θ_1 when L=16

We finally consider the case $L = 128$, and compare as above G-MUSIC SS, MUSIC SS, unsmoothed G-MUSIC and unsmoothed MUSIC. G-MUSIC SS completely fails because L and M are of the same order of magnitude. Theorem 2 is thus no more valid, and the pseudo-spectrum estimate is not consistent.

We now consider 2 widely spaced sources with DoAs $\theta_1 = 0$ and $\theta_2 = 5\frac{2\pi}{M}$, and keep the same parameters as above. We consider the case $L = 16$, and represent in Fig. 5 the performance of MUSIC, G-MUSIC, MUSIC-SS, and G-MUSIC-SS. It is first observed that, in contrast with the case of closely spaced DoAs, MUSIC-SS and G-MUSIC-SS have the same performance when the SNR is above the threshold 6 dB. This is in accordance with Theorem 4, and tends to indicate that, as in the case $L = 1$, if $\frac{\mathbf{S}_N \mathbf{S}_N^*}{N}$ converges towards a diagonal matrix, then the asymptotic performance of G-MUSIC-SS and MUSIC-SS coincide (see Remark 3). The comparison between the methods with and without spatial smoothing also confirm that the use of spatial smoothing schemes allow to improve the performance.

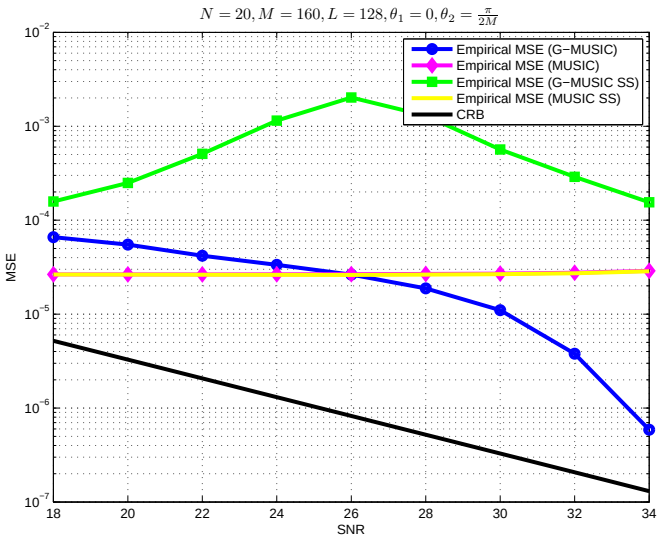

Fig. 4: Empirical MSE of different estimators of θ_1 when L=128

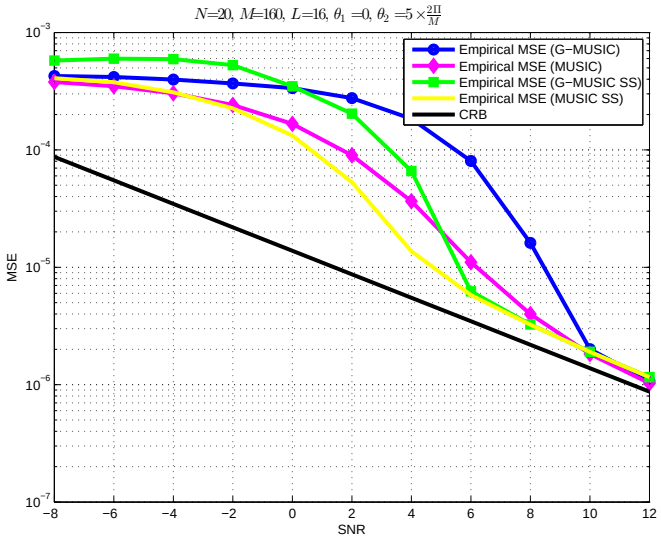

Fig. 5: Empirical MSE of different estimators of θ_1 when L=16 and widely spaced DoAs

VI. Conclusion

In this paper, we have addressed the behaviour of subspace DoA estimators based on spatial smoothing in asymptotic regimes where M and NL converge towards ∞ at the same rate. For this, we have evaluated the behaviour of the largest singular values and corresponding singular vectors of large random matrices defined as additive low rank perturbations of certain random block-Hankel matrices, and established that

they behave as if the entries of the block-Hankel matrices were i.i.d. Starting from this result, we have shown that it is possible to generalize the G-estimators introduced in [5], and have deduced from [18] their properties.

APPENDIX A
INSIGHTS ON THE PROOF OF (28) AND (29).

We first recall that [3] established that for $1 \leq k \leq K$, if $\hat{\lambda}_{k,N}$ does not converge towards a limit strictly greater than x_*^+, then $\hat{\lambda}_{k,N}$ converges towards x_*^+. We have therefore to evaluate the behaviour of the eigenvalues of $\mathbf{X}_N\mathbf{X}_N^*$ that are greater than $x_*^+ + \epsilon$ for some $\epsilon > 0$.

If $\mathbf{C}$ represents a $P \times Q$ matrix, we denote by $\underline{\mathbf{C}}$ the $(P+Q) \times (P+Q)$ hermitian matrix defined by

$$\underline{\mathbf{C}} = \begin{bmatrix} 0 & \mathbf{C} \\ \mathbf{C}^* & 0 \end{bmatrix}. \tag{44}$$

Then, the non zero eigenvalues of $\underline{\mathbf{C}}$ coincide with the (positive and negative) square roots of the non zero eigenvalues of matrix $\mathbf{C}\mathbf{C}^*$, and the corresponding eigenvectors are the $(P+Q)$–dimensional vectors $(\mathbf{a}_k^T, \pm\mathbf{b}_k^T)^T$ where $(\mathbf{a}_k, \mathbf{b}_k)$ represent the pairs of left and right singular vectors of $\mathbf{C}$. Therefore, $\lambda > x_*^+ + \epsilon$ is eigenvalue of $\mathbf{X}_N\mathbf{X}_N^*$ if and only if $\sqrt{\lambda} > (x_*^+ + \epsilon)^{1/2}$ is eigenvalue of matrix $\underline{\mathbf{X}}_N$. We consider the singular value decomposition

$$\mathbf{B}_N = \mathbf{U}_N \mathbf{\Lambda}_N^{1/2} \tilde{\mathbf{U}}_N^*$$

of matrix $\mathbf{B}_N$ and express $\underline{\mathbf{X}}_N$ as

$$\underline{\mathbf{X}}_N = \begin{bmatrix} 0 & \mathbf{Z}_N \\ \mathbf{Z}_N^* & 0 \end{bmatrix} + \begin{bmatrix} 0 & \mathbf{U}_N \mathbf{\Lambda}_N^{1/2} \tilde{\mathbf{U}}_N^* \\ \tilde{\mathbf{U}}_N \mathbf{\Lambda}_N^{1/2} \mathbf{U}_N^* & 0 \end{bmatrix}$$

which can be written as

$$\begin{bmatrix} 0 & \mathbf{Z}_N \\ \mathbf{Z}_N * & 0 \end{bmatrix} + \underbrace{\begin{bmatrix} \mathbf{U}_N & 0 \\ 0 & \tilde{\mathbf{U}}_N \mathbf{\Lambda}_N^{1/2} \end{bmatrix}}_{\mathbf{D}} \mathbf{J} \underbrace{\begin{bmatrix} \mathbf{U}_N^* & 0 \\ 0 & \mathbf{\Lambda}_N^{1/2} \tilde{\mathbf{U}}_N^* \end{bmatrix}}_{\mathbf{D}^*}$$

where $\mathbf{J}$ is defined by

$$\mathbf{J} = \begin{pmatrix} 0 & \mathbf{I}_K \\ \mathbf{I}_K & 0 \end{pmatrix}$$

Consider $x > (x_*^+ + \epsilon)^{1/2}$. Then, x is not a singular value of $\mathbf{Z}_N\mathbf{Z}_N^*$, and therefore, not an eigenvalue of $\underline{\mathbf{Z}}_N$. Therefore, it holds that

$$\begin{aligned}\det(\underline{\mathbf{X}}_N - x\mathbf{I}) &= \det(\underline{\mathbf{Z}}_N - x\mathbf{I} + \mathbf{DJD}^*) \\ &= \det(\underline{\mathbf{Z}}_N - x\mathbf{I})\det(\mathbf{I} + (\underline{\mathbf{Z}}_N - x\mathbf{I})^{-1}\mathbf{DJD}^*) \\ &= \det(\underline{\mathbf{Z}}_N - x\mathbf{I})\det(\mathbf{I}_{2K} + \mathbf{JD}^*(\underline{\mathbf{Z}}_N - x\mathbf{I})^{-1}\mathbf{D})\end{aligned}$$

after noticing $\mathbf{J} = \mathbf{J}^{-1}$. For $w \in \mathbb{C} - [-(x_*^+ + \epsilon)^{1/2}, (x_*^+ + \epsilon)^{1/2}]$, we denote by $\mathbf{S}_N(w)$ the $2K \times 2K$ matrix defined by

$$\mathbf{S}_N(w) = \mathbf{I}_{2K} + \mathbf{JD}^*(\underline{\mathbf{Z}}_N - w\mathbf{I})^{-1}\mathbf{D}$$

Using the identity

$$(\underline{\mathbf{Z}}_N - w\mathbf{I})^{-1} = \begin{bmatrix} w\mathbf{Q}_N(w^2) & \mathbf{Q}_N(w^2)\mathbf{Z}_N \\ \mathbf{Z}_N^*\mathbf{Q}_N(w^2) & w\tilde{\mathbf{Q}}_N(w^2) \end{bmatrix}$$

we obtain immediately that

$$(\mathbf{S}_N(w))_{1,1} = \mathbf{I}_K + \mathbf{\Lambda}_N^{1/2}\tilde{\mathbf{U}}_N^*\mathbf{Z}_N^*\mathbf{Q}_N(w^2)\mathbf{U}_N$$

$$(\mathbf{S}_N(w))_{1,2} = w\mathbf{\Lambda}_N^{1/2}\mathbf{U}_N^*\widetilde{\mathbf{Q}}_N(w^2)\tilde{\mathbf{U}}_N\mathbf{\Lambda}_N^{1/2}$$

$$(\mathbf{S}_N(w))_{2,1} = w\mathbf{U}_N^*\mathbf{Q}_N(w^2)\mathbf{U}_N$$

$$(\mathbf{S}_N(w))_{2,2} = \mathbf{I}_K + \mathbf{U}_N^*\mathbf{Q}_N(w^2)\mathbf{Z}_N\tilde{\mathbf{U}}_N\mathbf{\Lambda}_N^{1/2}$$

Item (iii) of Proposition 1 implies that the elements of $\mathbf{S}_N(w)$ converge almost surely, uniformly on the compact subsets of $\mathbb{C} - [-(x_*^+ + \epsilon)^{1/2}, (x_*^+ + \epsilon)^{1/2}]$ towards the elements of matrix $\mathbf{S}_*(w)$ defined by

$$\mathbf{S}_*(w) = \begin{bmatrix} \mathbf{I}_K & w\widetilde{m}_*(w^2)\mathbf{\Lambda} \\ wm(w^2)\mathbf{I}_K & \mathbf{I}_K \end{bmatrix}$$

It is easy to check that $\det(\mathbf{S}_N(w))$ and $\det(\mathbf{S}_*(w))$ are functions of w^2. We define functions s_N and s_* on $\mathbb{C} - [0, x_*^+ + \epsilon]$ by $s_N(w^2) = \det(\mathbf{S}_N(w))$ and $s_*(w^2) = \det(\mathbf{S}_*(w))$. It is clear that almost surely, $s_N(z) \to s_*(z)$ uniformly on the compact subsets of $\mathbb{C} - [0, (x_*^+ + \epsilon)]$. Therefore, in order to precise the behaviour of the eigenvalues of $\mathbf{X}_N\mathbf{X}_N^*$ that are greater than $x_*^+ + \epsilon$ (i.e. the solutions of the equation $s_N(x) = 0$ greater than $x_*^+ + \epsilon$), it is first useful to characterize the solutions of the equation $s_*(x) = 0$. The equation $s_*(x) = 0$ is equivalent to

$$\Pi_{k=1}^K (1 - \lambda_k x m_*(x)\tilde{m}_*(x)) = 0$$

or equivalently to

$$w_*(x) = \lambda_k$$

for $k = 1, \ldots, K$. Using the properties of function w_*, we obtain immediately that if $\epsilon < \rho_s - x_*^+ = \phi_*(\lambda_s) - x_*^+$, then the solutions of $s_*(x) = 0$ that are greater than $x_*^+ + \epsilon$ coincide with the $(\rho_k)_{k=1,\ldots,s}$ defined by $\rho_k = \phi_*(\lambda_k)$ for $k = 1, \ldots, s$. Using this, it can be proved using appropriate arguments that, almost surely, for N large enough, then the s greatest eigenvalues $(\lambda_{k,N})_{k=1,\ldots,s}$ of $\mathbf{X}_N\mathbf{X}_N^*$ are greater than $x_*^+ + \epsilon$, and that $\lambda_{k,N} \to \rho_k$ for $k = 1, \ldots, s$ [1].

APPENDIX B
PROOF OF PROPOSITION 1.

The proof of Proposition 1 is based on the results of [8]. In order to explain this, we denote by $\mathbf{W}_N$ the $NL \times (M - L + 1)$ matrix defined by

$$\mathbf{W}_N = \frac{1}{\sqrt{c_N}} \mathbf{Z}_N^*$$

The variance of the entries of $\mathbf{W}_N$ is equal to $\frac{\sigma^2}{M-L+1}$. Therefore, matrix $\mathbf{W}_N$ is similar to the matrices studied in [8] except that the integers (M, N) in [8] should be exchanged by $(N, M-L+1)$. In particular, after this replacement, it is clear that the asymptotic regime (11) coincides with the regime in [8]. In order to recall the results of [8], we denote by $t_N(z)$, $\tilde{t}_N(z)$, $t_*(z)$ and $\tilde{t}_*(z)$ the Stieltjes transforms of the Marcenko-Pastur distributions of parameters (σ^2, c_N^{-1}), $(\sigma^2 c_N^{-1}, c_N)$, (σ^2, c_*^{-1}) and $(\sigma^2 c_*^{-1}, c_*)$. Moreover, $\mathbf{Q}_{N,W}(z)$ and $\tilde{\mathbf{Q}}_{N,W}(z)$ represent the resolvents of matrices $\mathbf{W}_N\mathbf{W}_N^*$ and $\mathbf{W}_N^*\mathbf{W}_N$ respectively. It is shown in [8] (see Section 6) that the eigenvalue distribution of $\mathbf{W}_N\mathbf{W}_N^*$ converges almost surely towards $\mu_{\sigma^2, c_*^{-1}}$, a statement equivalent to

$$\frac{1}{NL}\mathrm{Tr}(\mathbf{Q}_{N,W}(z)) - t_*(z) \to 0 \; a.s.$$

or to

$$\frac{1}{M-L+1}\mathrm{Tr}(\tilde{\mathbf{Q}}_{N,W}(z)) - \tilde{t}_*(z) \to 0 \; a.s.$$

for each $z \in \mathbb{C}^+$. As we have

$$\mathbf{Z}_N^*\mathbf{Z}_N = c_N\mathbf{W}_N\mathbf{W}_N^* \tag{45}$$

[1]The arguments used in [3] require the uniform convergence of s_N towards s_* on the set $\mathrm{Re}(z) > \mathbf{x}_*^+ + \epsilon$, a property that is not established in Proposition 1. However, the proof of the contuinity lemma 2.1 in [2] can be simplified, and only needs the uniform convergence on compact sets.

it holds that the resolvent $\tilde{\mathbf{Q}}_N(z)$ of $\mathbf{Z}_N^*\mathbf{Z}_N$ is equal to

$$\tilde{\mathbf{Q}}_N(z) = c_N^{-1}\mathbf{Q}_{N,W}(zc_N^{-1})$$

As $c_N \to c_*$, it is clear $\frac{1}{NL}\mathrm{Tr}(\tilde{\mathbf{Q}}_N(z))$ behaves as $\frac{1}{c_*}\frac{1}{NL}\mathrm{Tr}(\mathbf{Q}_{N,W}(zc_*^{-1})$. Similarly, $\frac{1}{M-L+1}\mathrm{Tr}(\mathbf{Q}_N(z))$ behaves as $\frac{1}{c_*}\frac{1}{M-L+1}\mathrm{Tr}(\tilde{\mathbf{Q}}_{N,W}(zc_*^{-1}))$. Therefore, for each $z \in \mathbb{C}^+$, it holds that

$$\frac{1}{NL}\mathrm{Tr}(\tilde{\mathbf{Q}}_N(z)) - c_*^{-1}\, t_*(zc_*^{-1}) \to 0 \; a.s.$$

and that

$$\frac{1}{M-L+1}\mathrm{Tr}(\mathbf{Q}_N(z)) - c_*^{-1}\, \tilde{t}_*(zc_*^{-1}) \to 0 \; a.s.$$

Using Equations (18, 19), it is easy to verify that $m_*(z) = c_*^{-1}\,\tilde{t}_*(zc_*^{-1})$ and $\tilde{m}_*(z) = c_*^{-1}\,t_*(zc_*^{-1})$. This establishes (23) and the convergence of the eigenvalue distribution of $\mathbf{Z}_N\mathbf{Z}_N^*$ towards μ_{σ^2,c_*}.

Asymptotic regime (11) implies that $L = \mathcal{O}(M^\alpha) = \mathcal{O}((M-L+1)^\alpha)$ where $\alpha < 2/3$. Therefore, [8] implies that for each $\epsilon > 0$, almost surely, for N large enough, the eigenvalues of $\mathbf{W}_N\mathbf{W}_N^*$ are located in $[\sigma^2(1-\sqrt{c_*^{-1}})^2 - \epsilon, \sigma^2(1+\sqrt{c_*^{-1}})^2] \cup \{0\}\mathbb{1}(c_*^{-1} > 1)$. (45) and the convergence of c_N towards c_* lead immediately to item (ii) of Proposition 1.

Using the same arguments as above, (25) appears as a consequence of

$$\tilde{\mathbf{a}}_N^* \left(\mathbf{Q}_{N,W}(z) - t_*(z)\mathbf{I}\right)\tilde{\mathbf{b}}_N \to 0 \; a.s. \tag{46}$$

While (46) does not appear explicitely in [8], it can be deduced rather easily from the various intermediate results proved in [8]. For this, we first remark that

$$\tilde{\mathbf{a}}_N^* \left(\mathbf{Q}_{N,W}(z) - t_*(z)\mathbf{I}\right)\tilde{\mathbf{b}}_N = \tilde{\mathbf{a}}_N^* \left(\mathbf{Q}_{N,W}(z) - \mathbb{E}(\mathbf{Q}_{N,W}(z))\right)\tilde{\mathbf{b}}_N + \tilde{\mathbf{a}}_N^* \left(\mathbb{E}(\mathbf{Q}_{N,W}(z)) - t_*(z)\mathbf{I}\right)\tilde{\mathbf{b}}_N$$

and establish that the 2 terms at the right hand side of the above equation converge towards 0. In order to simplify the notations, we denote by ξ the first term. The almost sure convergence of ξ towards 0 follows from the Poincaré-Nash inequality (see e.g. Proposition 2 of [8]). Exchanging (M,N) by $(N, M-L+1)$ in Proposition 6 of [8], we obtain immediately that $\mathbb{E}|\xi|^2 = \mathcal{O}(\frac{L}{M-L+1}) = \mathcal{O}(\frac{L}{M})$. As $L/M \to 0$, this implies that ξ converges in probability towards 0. In order to prove the almost sure convergence, we briefly justify that for each n, it holds that

$$\mathbb{E}|\xi|^{2n} = \mathcal{O}\left((L/M)^n\right) \tag{47}$$

(47) can be established by induction on n. As mentioned above, (47) is verified for $n = 1$. We now assume that it holds until integer $n-1$, and prove (47). For this, we use the obvious relation:

$$\mathbb{E}|\xi|^{2n} = (\mathbb{E}|\xi|^n)^2 + \mathrm{Var}(\xi^n)$$

Using the Poincaré-Nash inequality as in the proof of Proposition 6 of [8], we obtain easily that

$$\mathrm{Var}(\xi^n) \leq C\,\frac{L}{M}\mathbb{E}(|\xi|^{2n-2})$$

where C is a constant that depends on z but not on the dimensions L, M, N. As (47) is assumed to hold until integer $n-1$, this implies that $\mathrm{Var}(\xi^n) = \mathcal{O}\left((L/M)^n\right)$. The Schwartz inequality leads immediately to

$$\left(\mathbb{E}|\xi|^n\right)^2 \leq \mathbb{E}(|\xi|^2)\,\mathbb{E}(|\xi|^{2n-2})$$

which is a $\mathcal{O}\left((L/M)^n\right)$ term. This establishes (47). As $L = \mathcal{O}(M^\alpha)$ with $\alpha < 2/3$, it is clear that $(L/M)^3$ verifies

$$(L/M)^3 = \frac{1}{M^{1+2-3\alpha}}$$

Therefore, (47) for $n = 3$ leads to

$$\mathbb{E}\left(|\xi|^6\right) = \mathcal{O}\left(\frac{1}{M^{1+2-3\alpha}}\right)$$

As $2-3\alpha > 0$, the use of the Markov inequality and of the Borel-Cantelli lemma imply that ξ converges towards 0 almost surely as expected.

It remains to justify that $\tilde{\mathbf{a}}_N^*\left(\mathbb{E}(\mathbf{Q}_{N,W}(z)) - t_*(z)\mathbf{I}\right)\tilde{\mathbf{b}}_N$ converges towards 0. Although it is not stated explicitely in [8], it can immediately deduced from Eq. (5.3) in Proposition 8, as well as on Corollary 1, Theorem 2, and formula (7.3).

(24) is equivalent to

$$\mathbf{a}_N^*\left(\tilde{\mathbf{Q}}_{N,W}(z) - \tilde{t}_*(z)\mathbf{I}\right)\mathbf{b}_N \to 0\ a.s. \tag{48}$$

It can be proved as above that

$$\mathbf{a}_N^*\left(\tilde{\mathbf{Q}}_{N,W}(z) - \mathbb{E}(\tilde{\mathbf{Q}}_{N,W}(z))\right)\mathbf{b}_N \to 0\ a.s.$$

and establish that

$$\mathbf{a}_N^*\left(\mathbb{E}(\tilde{\mathbf{Q}}_{N,W}(z)) - \tilde{t}_*(z)\mathbf{I}\right)\mathbf{b}_N \to 0 \tag{49}$$

for each $z \in \mathbb{C}^+$. The behaviour of matrix $\mathbb{E}(\tilde{\mathbf{Q}}_{N,W}(z))$ is not studied in [8]. However, it can be evaluated using the results of [8]. For this, we first simplify the notations and denote by $\mathbf{W}, \tilde{\mathbf{W}}, \mathbf{Q}, \tilde{\mathbf{Q}}$ the matrices $\mathbf{W}_N, \tilde{\mathbf{W}}_N, \mathbf{Q}_{N,W}(z)$, and $\tilde{\mathbf{Q}}_{N,W}(z)$. Moreover, $\mathbf{Q}$ is a $NL \times NL$ block matrix, so that we denote by $\mathbf{Q}_{i_1,i_2}^{n_1,n_2}$ its entry $(i_1 + (n_1-1)L, i_2 + (n_2-1)L)$.

As in [8], we denote by $\tau^{(N)}(.)$ and $\mathcal{T}^{(N)}_{M-L+1,L}(.)$ the operators defined by

$$\tau^{(N)}(\mathbf{Q})(i) = \frac{1}{NL}\mathrm{Tr}\left(\mathbf{Q}(\mathbf{I}\otimes\mathbf{J}_L^i)\right)$$

$$\mathcal{T}^{(N)}_{M-L+1,L}(\mathbf{Q}) = \sum_{i=-(L-1)}^{L-1} \tau^{(N)}(\mathbf{Q})(i)\mathbf{J}^{*i}_{M-L+1}$$

where $\mathbf{J}_L$ is the $L\times L$ "shift" matrix defined by $(\mathbf{J}_L)_{i,j} = \delta(j-i=1)$.

Replacing integers (M,N) by integers $(N, M-L+1)$ in Equation (4.6) of [8], we obtain that

$$\mathbb{E}\left[(\mathbf{QW})^{n_1}_{i_1,k}(\mathbf{W}^*)^{n_2}_{j,i_2}\right] = \frac{\sigma^2}{M-L+1}\mathbb{E}\left(\mathbf{Q}^{n_1,n_2}_{i_1,i_2-(k-j)}\right)\mathbb{1}_{1\le i_2-(k-j)\le L}$$
$$-\frac{\sigma^2}{c_N}\sum_{i=-(L-1)}^{L-1}\mathbb{1}_{1\le k-i\le M-L+1}\mathbb{E}\left[\tau^{(N)}(\mathbf{Q})(i)(\mathbf{QW})^{n_1}_{i_1,k-i}(\mathbf{W}^*)^{n_2}_{j,n_2}\right]$$

with $1\le j,k\le M-L+1$, $1\le n_1,n_2\le N$, $1\le i_1,i_2\le L$.

Setting $u=k-i$, the second term of the righthandside of the above equation can also be written as

$$\frac{\sigma^2}{c_N}\sum_{u=1}^{M-L+1}\mathbb{E}\left[\tau^{(N)}(\mathbf{Q})(k-u)\mathbb{1}_{-(L-1)\le k-u\le L-1}(\mathbf{QW})^{n_1}_{i_1,u}(\mathbf{W}^*)^{n_2}_{j,n_2}\right]$$

Now setting $n=n_1=n_2, i=i_1=i_2$, and summing over n and i, we obtain

$$\mathbb{E}\left(\mathbf{W}^*\mathbf{QW}\right)_{j,k} = \frac{\sigma^2}{c_N}\tau^{(N)}(\mathbb{E}(\mathbf{Q}))(k-j)\mathbb{1}_{-(L-1)\le k-j\le L-1}$$
$$-\frac{\sigma^2}{c_N}\mathbb{E}\left(\sum_{u=1}^{M-L+1}\tau^{(N)}(\mathbf{Q})(k-u)\mathbb{1}_{-(L-1)\le k-u\le L-1}(\mathbf{W}^*\mathbf{QW})_{j,u}\right)$$

and using that $\tau^{(N)}(\mathbf{Q})(k-u)\mathbb{1}_{-(L-1)\le k-u\le L-1} = \left(\mathcal{T}^{(N)}_{M-L+1,L}(\mathbf{Q})\right)_{k,u}$, we get that

$$\mathbb{E}\left(\mathbf{W}^*\mathbf{QW}\right)_{j,k} = \frac{\sigma^2}{c_N}\left(\mathcal{T}^{(N)}_{M-L+1,L}(\mathbb{E}(\mathbf{Q}))\right)_{k,j} - \frac{\sigma^2}{c_N}\mathbb{E}\left(\mathcal{T}^{(N)}_{M-L+1,L}(\mathbf{Q})\mathbf{W}^T\mathbf{Q}^T\overline{\mathbf{W}}\right)_{k,j}$$

We express matrix $\mathbf{Q} = \mathbb{E}(\mathbf{Q}) + \mathring{\mathbf{Q}}$, and obtain that

$$\mathbb{E}\left(\mathbf{W}^*\mathbf{QW}\right)_{j,k} = \frac{\sigma^2}{c_N}\left(\mathcal{T}^{(N)}_{M-L+1,L}(\mathbb{E}(\mathbf{Q}))\right)_{k,j}$$
$$-\frac{\sigma^2}{c_N}\left(\mathcal{T}^{(N)}_{M-L+1,L}(\mathbb{E}(\mathbf{Q}))\mathbb{E}(\mathbf{W}^T\mathbf{Q}^T\overline{\mathbf{W}})\right)_{k,j}$$
$$-\frac{\sigma^2}{c_N}\mathbb{E}\left(\mathcal{T}^{(N)}_{M-L+1,L}(\mathring{\mathbf{Q}})\mathbf{W}^T\mathbf{Q}^T\overline{\mathbf{W}}\right)_{k,j}$$

Noticing the equation,

$$\mathbf{W}^T\mathbf{Q}^T\overline{\mathbf{W}} = \tilde{\mathbf{Q}}^T\mathbf{W}^T\overline{\mathbf{W}}$$

we obtain that

$$\mathbb{E}\left(\mathbf{W}^*\mathbf{QW}\right) = \frac{\sigma^2}{c_N}\mathcal{T}^{(N)}_{M-L+1,L}\left(\mathbb{E}(\mathbf{Q}^T)\right) - \frac{\sigma^2}{c_N}\mathbb{E}\left(\mathbf{W}^*\mathbf{W}\tilde{\mathbf{Q}}\right)\mathcal{T}^{(N)}_{M-L+1,L}\left(\mathbb{E}(\mathbf{Q}^T)\right) - \frac{\sigma^2}{c_N}\mathbb{E}\left(\mathbf{W}^*\mathbf{W}\tilde{\mathbf{Q}}\mathcal{T}^{(N)}_{M-L+1,L}(\mathring{\mathbf{Q}})\right)$$

Moreover we notice that

$$\mathbf{W}^*\mathbf{Q}\mathbf{W} = \tilde{\mathbf{Q}}\mathbf{W}^*\mathbf{W} = \mathbf{W}^*\mathbf{W}\tilde{\mathbf{Q}} = \mathbf{I} + z\tilde{\mathbf{Q}} \tag{50}$$

Therefore, it holds that

$$\mathbf{I} + z\mathbb{E}(\tilde{\mathbf{Q}}) = \frac{\sigma^2}{c_N}\mathcal{T}^{(N)}_{M-L+1,L}(\mathbb{E}(\mathbf{Q}^T)) - \frac{\sigma^2}{c_N}(\mathbf{I} + z\mathbb{E}(\tilde{\mathbf{Q}}))\mathcal{T}^{(N)}_{M-L+1,L}\left(\mathbb{E}(\mathbf{Q}^T)\right) + \tilde{\Delta}$$

where

$$\tilde{\Delta} = -\frac{\sigma^2}{c_N}\mathbb{E}\left(\mathbf{W}^*\mathbf{W}\tilde{\mathbf{Q}}\mathcal{T}^{(N)}_{M-L+1,L}(\overset{\circ}{\mathbf{Q}})\right) \tag{51}$$

This leads to the equation

$$z\mathbb{E}(\tilde{\mathbf{Q}})\left(\mathbf{I} + \frac{\sigma^2}{c_N}\mathcal{T}^{(N)}_{M-L+1,L}(\mathbb{E}(\mathbf{Q}^T))\right) = -\mathbf{I} + \tilde{\Delta} \tag{52}$$

Lemma 2 of [8] (used when (M, N) is replaced by $(M - L + 1, N)$) implies that matrix

$$\mathbf{I} + \frac{\sigma^2}{c_N}\mathcal{T}^{(N)}_{M-L+1,L}(\mathbb{E}(\mathbf{Q}))$$

is invertible for $z \in \mathbb{C}^+$, and that its inverse, denoted $\mathbf{H}$, verifies

$$\|\mathbf{H}\| \leq \frac{|z|}{\mathrm{Im}(z)} \tag{53}$$

for $z \in \mathbb{C}^+$. (52) implies that

$$\mathbb{E}(\tilde{\mathbf{Q}}) = -\frac{\mathbf{H}^T}{z} + \tilde{\Delta}\mathbf{H}^T$$

Therefore, (49) is equivalent to

$$\mathbf{a}_N^*\left(-\frac{\mathbf{H}^T}{z} - \tilde{t}_*(z)\mathbf{I} + \tilde{\Delta}\mathbf{H}^T\right)\mathbf{b}_N \to 0$$

Using the same technics as in Proposition 8 (see Eq. 5.3) of [8] as well as (53), we obtain immediately that

$$\mathbf{a}_N^*\tilde{\Delta}\mathbf{H}^T\mathbf{b}_N \to 0$$

It thus remains to establish that

$$\mathbf{a}_N^*\left(-\frac{\mathbf{H}^T}{z} - \tilde{t}_*(z)\mathbf{I}\right)\mathbf{b}_N \to 0 \tag{54}$$

For this, we use the identity

$$-\frac{\mathbf{H}^T}{z} - \tilde{t}_*(z)\mathbf{I} = -\mathbf{H}^T\left(\frac{\mathbf{I}}{z\tilde{t}_*(z)} + (\mathbf{H}^T)^{-1}\right)\tilde{t}_*(z)$$

$t_*(z)$ and $\tilde{t}_*(z)$ satisfy the relation $\frac{-1}{z\tilde{t}_*(z)} = 1 + \frac{\sigma^2}{c_N}t_*(z)$. Hence, the right hand side of the above equation can be written as

$$-\mathbf{H}^T\left((-1 - \frac{\sigma^2}{c_N}t_*(z))\mathbf{I} + \mathbf{I} + \frac{\sigma^2}{c_N}\mathcal{T}^{(N)}_{M-L+1,L}\left(\mathbb{E}(\mathbf{Q}^T)\right)\right)\tilde{t}_*(z)$$

Corollary 1 and Theorem 2 of [8] imply that

$$\left\| \mathcal{T}^{(N)}_{M-L+1,L} \left(\mathbb{E}(\mathbf{Q}^T) - t_*(z)\mathbf{I} \right) \right\| \to 0$$

if $z \in \mathbb{C}^+$. This and (53) leads to

$$\mathbf{a}_N^* \mathbf{H}^T \mathcal{T}^{(N)}_{M-L+1,L} \left(\mathbb{E}(\mathbf{Q}^T(z)) - t_*(z)\mathbf{I} \right) \mathbf{b}_N \to 0$$

and to (54). This completes the proof of (49).

We now establish (26). For this, we first remark that for each $\theta \in \mathbb{R}$, the distribution of matrix $\mathbf{Z}_N e^{i\theta}$ coincides with the distribution of $\mathbf{Z}_N$. Therefore, it holds that

$$\mathbb{E}\left(\mathbf{Q}_N(z) \mathbf{Z}_N e^{i\theta} \right) = \mathbb{E}\left(\mathbf{Q}_N(z) \mathbf{Z}_N \right)$$

which implies that $\mathbb{E}\left(\mathbf{Q}_N(z) \mathbf{Z}_N \right) = 0$. In order to complete the proof of (26), it is sufficient to establish that if we denote by κ_N the random variable $\kappa_N = \mathbf{a}_N^* \left(\mathbf{Q}_N(z) \mathbf{Z}_N \right) \mathbf{b}_N$, then, for each $p \geq 1$, it holds that

$$\mathbb{E}\left| \kappa_N - \mathbb{E}(\kappa_N) \right|^{2p} = \mathcal{O}\left(\left(\frac{L}{M} \right)^p \right) \tag{55}$$

Choosing p large enough leads to $\kappa_N - \mathbb{E}(\kappa_N) = \kappa_N \to 0$ $a.s.$ as expected. (55) can be proved as above by using the Poincaré-Nash inequality.

We finally justify that for each $\epsilon > 0$, (24, 25, 26) hold uniformly w.r.t. z on each compact subset of $\mathbb{C} - [0, x_*^+ + \epsilon]$. We just prove that it the case for (26). By item (ii), almost surely, function $z \to \kappa_N(z)$ is analytic on $\mathbb{C} - [0, x_*^+ + \epsilon]$. We use a standard argument based on Montel's theorem ([12], p.282). We first justify that for each compact subset $\mathcal{K} \subset \mathbb{C} - [0, x_*^+ + \epsilon]$, then it exists a constant η such that

$$\sup_{z \in \mathcal{K}} |\kappa_N(z)| \leq \eta \tag{56}$$

for each N large enough. We consider the singular value decomposition of matrix $\mathbf{Z}_N$:

$$\mathbf{Z}_N = \mathbf{\Gamma}_N \mathbf{\Delta}_N \mathbf{\Theta}_N^*$$

where $\mathbf{\Delta}_N$ represents the diagonal matrix of non zero singular values of $\mathbf{Z}_N$. $\kappa_N(z)$ can be written as

$$\kappa_N(z) = \mathbf{a}_N^* \mathbf{\Gamma}_N \left(\mathbf{\Delta}_N^2 - z\mathbf{I} \right)^{-1} \mathbf{\Delta}_N \mathbf{\Theta}_N^* \mathbf{b}_N$$

Therefore, it holds that

$$|\kappa_N(z)| \leq \left\| \left(\mathbf{\Delta}_N^2 - z\mathbf{I} \right)^{-1} \mathbf{\Delta}_N \right\| \|\mathbf{a}_N\| \|\mathbf{b}_N\|$$

Item (ii) implies that the entries of $\mathbf{\Delta}_N^2$ are located into $[0, x_*^+ + \epsilon]$. for each N large enough. Therefore, for each $z \in \mathcal{K}$, it holds that

$$\left\| \left(\mathbf{\Delta}_N^2 - z\mathbf{I} \right)^{-1} \mathbf{\Delta}_N \right\| \leq \frac{1}{\mathrm{dist}([0, x_*^+ + \epsilon], \mathcal{K})}$$

The conclusion follows from the hypothesis that vectors $\mathbf{a}_N$ and $\mathbf{b}_N$ satisfy $\sup_N(\|\mathbf{a}_N\|, \|\mathbf{b}_N\|) < +\infty$. (56) implies that the sequence of analytic functions $(\kappa_N)_{N\geq 1}$ is a normal family . Therefore, it exists a subsequence extracted from $(\kappa_N)_{N\geq 1}$ that converges uniformly on each compact subset of $\mathbb{C} - [0, x_*^+ + \epsilon]$ towards a certain analytic function κ_*. As (26) holds for each $z \in \mathbb{C}^+$, function κ_* is identically zero. We have thus shown that each converging subsequence extracted from $(\kappa_N)_{N\geq 1}$ converges uniformly towards 0 on each compact subset of $\mathbb{C} - [0, x_*^+ + \epsilon]$. This, in turn, shows that the whole sequence converges uniformly on each compact subset of $\mathbb{C} - [0, x_*^+ + \epsilon]$ as expected.

REFERENCES

[1] Z.D. Bai and J.W. Silverstein. *Spectral analysis of large dimensional random matrices*. Springer Series in Statistics, 2nd edition, 2010.

[2] Florent Benaych-Georges and Raj Rao Nadakuditi. The eigenvalues and eigenvectors of finite, low rank perturbations of large random matrices. *Adv. in Math.*, 227(1):494–521, 2011.

[3] Florent Benaych-Georges and Raj Rao Nadakuditi. The singular values and vectors of low rank perturbations of large rectangular random matrices. *Journal of Multivariate Analysis*, 111:120–135, 2012.

[4] J.E Evans, D.F Sun, and J.R Johnson. Application of advanced signal processing techniques to angle estimation in ATC navigation and surveillance systems. 1982. M.I.T. Lincoln LAB, Lexington, M.A., Rep. 582.

[5] W. Hachem, P. Loubaton, X. Mestre, J. Najim, and P. Vallet. A Subspace Estimator for Fixed Rank Perturbations of Large Random Matrices. *Journal of Multivariate Analysis*, 114:427–447, 2012. arXiv:1106.1497.

[6] W. Hachem, P. Loubaton, X. Mestre, J. Najim, and P. Vallet. Large information plus noise random matrix models and consistent subspace estimation in large sensor networks. *Random Matrices: Theory and Applications*, 1(2), 2012.

[7] B.A. Johnson, Y.I. Abramovich, and X. Mestre. MUSIC, G-MUSIC, and Maximum-Likelihood performance breakdown . *IEEE Transactions on Signal Processing*, 56(8):3944 – 3958, 2008.

[8] P. Loubaton. On the almost sure location of the singular values of certain Gaussian block-Hankel large random matrices. 2014. arXiv:1405.2006.

[9] V.A. Marcenko and L.A. Pastur. Distribution of eigenvalues for some sets of random matrices. *Mathematics of the USSR-Sbornik*, 1:457, 1967.

[10] X. Mestre. Improved estimation of eigenvalues and eigenvectors of covariance matrices using their sample estimates. *Information Theory, IEEE Transactions on*, 54(11):5113–5129, 2008.

[11] S.U. Pillai and B.H. Kwon. Performance analysis of MUSIC type high resolution estimators for direction finding in correlated and coherent scenes . *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 37(8):1176–1189, 1989.

[12] W. Rudin. *Real and Complex Analysis*. McGraw-Hill, 3nd edition, 1987.

[13] T.J. Shan, M. Wax, and T. Kailath. On spatial smoothing for estimation of coherent signals. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 33(8):806–811, 1985.

[14] P. Stoica and A. Nehorai. MUSIC, maximum likelihood, and Cramer-Rao bound. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 37(5):720–741, 1989.

[15] A. Thakre, M. Haardt, and K. Giridhar. Single snapshot spatial smoothing with improved effective array aperture . *IEEE Signal Processing Letters*, 16(6):505–508, 2008.

[16] A. Thakre, M. Haardt, and K. Giridhar. Tensor-based spatial smoothing (TB-SS) using multiple snapshots . *IEEE Transactions on Signal Processing*, 58(5):2715–2728, 2010.

[17] P. Vallet, P. Loubaton, and X. Mestre. Improved Subspace Estimation for Multivariate Observations of High Dimension: The Deterministic Signal Case. *IEEE Transactions on Information Theory*, 58(2), Feb. 2012. arXiv: 1002.3234.

[18] P. Vallet, X. Mestre, and P. Loubaton. Performance analysis of an improved MUSIC DoA estimator. Accepted to ICASSP 2015, can be downloaded at http://www-syscom.univ-mlv.fr/~loubaton/dionisos/vallet-loubaton-mestre-icassp2015.pdf.

[19] P. Vallet, X. Mestre, and P. Loubaton. A CLT for G-MUSIC DOA estimator. In *European Signal Processing Conf (EUSIPCO)*, 2012.