Performance Analysis of Spatial Smoothing Schemes in the Context of Large Arrays

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

***Abstract*—This paper addresses the statistical behavior 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 modeled by an asymptotic regime in which NL and M both tend to ∞ 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 modeling 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.**

***Index Terms*— Large random matrices, spatial smoothing, DoA estimation.**

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 number of available samples N of the observed signal is much larger than the number of sensors M of the array (see e.g. [15] 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 [11] using large random matrix theory. [11] was followed by various works such as [6]–[8], [18]. 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 developed to address coherent sources ([5], [12], [14]), can be used to artificially increase the number of snapshots (see e.g. [15] and the references therein, see also the recent related contributions [16], [17] 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 tends to ∞. This context is not the most relevant when M is large because L must be chosen such 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 tend to ∞ 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 entries. To establish this result, we rely on the recent result [9] addressing the behaviour of the singular values of large block-Hankel random matrices built from i.i.d. Gaussian sequences. [9] 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 neighborhood of the support of the above distribution. This, and other additional technical results derived in the present paper, allow 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

Manuscript received January 20, 2015; revised June 17, 2015 and August 10, 2015; accepted September 01, 2015. Date of publication September 17, 2015; date of current version December 08, 2015. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Marius Pesavento. This work was supported in part by project ANR-12-MONU-OOO3 DIONISOS

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

P. Vallet is with the Laboratoire de l'Intégration du Matériau au Système (CNRS, Université Bordeaux, IPB), 33405 Talence, France (e-mail: pascal.vallet@ipb.fr).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2015.2480044

[18] and [6]. We deduce from the results of [19] 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 specify 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 [6] 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 [19]. Finally, Section V presents numerical experiments sustaining our theoretical results.

Notations: For a complex matrix $\mathbf{A}$, we denote by $\mathbf{A}^T, \overline{\mathbf{A}}, \mathbf{A}^*$ its transpose, conjugate and its conjugate transpose, and by Tr $(\mathbf{A})$ and $\|\mathbf{A}\|$ its trace and spectral norm. If $\mathbf{A}$ is a $P \times P$ hermitian matrix, we denote by $\lambda_1(\mathbf{A}) \geq \lambda_2(\mathbf{A}) \geq \ldots \geq \lambda_P(\mathbf{A})$ the eigenvalues of matrix $\mathbf{A}$ arranged in the decreasing order. The identity matrix will be $\mathbf{I}$. 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 (boundlessness 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-dimensional 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 matrix $\mathbb{E}[\mathbf{v}_n\mathbf{v}_n^*] = \sigma^2\mathbf{I}$.

We note that assuming the source signals as unknown deterministic also allows to cover the case of random signals because any realization of a random signal can be considered as a deterministic signal. 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 by

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

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. [15] 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

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

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

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

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 observation 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 $\frac{\mathbf{Y}_N^{(L)}\mathbf{Y}_N^{(L)*}}{NL}$. 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 $\frac{\mathbf{Y}_N^{(L)}\mathbf{Y}_N^{(L)*}}{NL}$ has the same asymptotic behaviour than $\mathbf{A}^{(L)}\left(\frac{\mathbf{S}_N\mathbf{S}_N^*}{NL} \otimes \mathbf{I}_L\right)\mathbf{A}^{(L)*} + \sigma^2\mathbf{I}_{M-L+1}$. 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 $\frac{\mathbf{Y}_N^{(L)}\mathbf{Y}_N^{(L)*}}{NL}$ 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

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

where we recall that if $\mathbf{A}$ is a matrix, then, $\|\mathbf{A}\|$ represents the spectral norm of $\mathbf{A}$. 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 tends to ∞ 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 modeling 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 tend to ∞ (even faster than N if $\beta < 1/2$) but condition $\alpha < 2/3$ (or equivalently $\beta > 1/3$) implies that the convergence speed of L to $+\infty$ is not arbitrarily fast. As explained in Section II-B, condition $L = \mathcal{O}(M^\alpha)$ with $\alpha < 2/3$ implies that matrix $\mathbf{V}_N^{(L)}$, behaves, in some sense, as a random matrix with i.i.d. entries, and that the results of [6] and [19] obtained in the case $L = 1$ can be extended to asymptotic regime (11).

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. 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), the empirical covariance matrix $\frac{\mathbf{Y}_N^{(L)}\mathbf{Y}_N^{(L)*}}{NL}$ is not a good estimate of the true covariance matrix $\mathbf{A}^{(L)}\left(\frac{\mathbf{S}_N\mathbf{S}_N^*}{NL} \otimes \mathbf{I}_L\right)\mathbf{A}^{(L)*} + \sigma^2\mathbf{I}_{M-L+1}$ in the sense that

$$\left\|\frac{\mathbf{Y}_N^{(L)}\mathbf{Y}_N^{(L)*}}{NL} - \left(\mathbf{A}^{(L)}\left(\frac{\mathbf{S}_N\mathbf{S}_N^*}{NL} \otimes \mathbf{I}_L\right)\mathbf{A}^{(L)*} + \sigma^2\mathbf{I}_{M-L+1}\right)\right\|$$

does not converge towards 0 almost surely. Roughly speaking, this is because the true covariance matrix depends on $\frac{(M-L+1)(M-L)}{2} = \mathcal{O}(M^2)$ parameters, and that the number of independent random variables that are available for estimation is equal to MN, which, in regime (11), is of course not sufficient. Therefore, (7) is no more valid, and hence, (10) is questionable. In this paper, we show that it is possible to generalize the G-MUSIC estimators introduced in [6] and [19] in the case where $L = 1$ to the context of spatial smoothing schemes in regime (11). In order to explain this unformally, we denote by $\mathbf{X}_N$, $\mathbf{Z}_N$, and $\mathbf{B}_N$ the matrices defined by $\mathbf{X}_N = \frac{\mathbf{Y}_N^{(L)}}{\sqrt{NL}}$, $\mathbf{Z}_N = \frac{\mathbf{V}_N^{(L)}}{\sqrt{NL}}$, and $\mathbf{B}_N = \frac{1}{\sqrt{NL}}\mathbf{A}^{(L)}(\mathbf{S}_N \otimes \mathbf{I}_L)$ (we do not mention that these matrices depend on L in order to simplify the notations), and observe that

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

We denote by $(\lambda_{k,N})_{k=1,\ldots,K}$ and $(\mathbf{u}_{k,N})_{k=1,\ldots,K}$ the non zero eigenvalues and related eigenvectors of matrix $\mathbf{B}_N\mathbf{B}_N^*$, and by $(\hat{\lambda}_{k,N})_{k=1,\ldots,M-L+1}$ and $(\hat{\mathbf{u}}_{k,N})_{k=1,\ldots,M-L+1}$ the eigenvalues and eigenvectors of matrix $\mathbf{X}_N\mathbf{X}_N^*$. Matrix $\mathbf{X}_N$ coincides with the sum of rank K deterministic matrix $\mathbf{B}_N$ and block-Hankel random matrix $\mathbf{Z}_N$ due to the noise, and it is of course of fundamental interest to specify the behaviour of the K largest eigenvalues $(\hat{\lambda}_{k,N})_{k=1,\ldots,K}$ and related eigenvectors $(\hat{\mathbf{u}}_{k,N})_{k=1,\ldots,K}$ in the asymptotic regime (11). If matrix $\mathbf{Z}_N$ was i.i.d., the results of [3] and [6] would imply that, under the so-called separation condition

$$\lambda_{K,N} > \sigma^2\sqrt{c_*},\ \text{for each } N \text{ large enough} \tag{12}$$

then, for each $k = 1, \ldots, K$, it would hold that

$$\begin{aligned}&\mathbf{a}_{M-L+1}(\theta)^*\hat{\mathbf{u}}_{k,N}\hat{\mathbf{u}}_{k,N}^*\mathbf{a}_{M-L+1}(\theta)\\ &= h_*(\hat{\lambda}_{k,N})\mathbf{a}_{M-L+1}(\theta)^*\mathbf{u}_{k,N}\mathbf{u}_{k,N}^*\mathbf{a}_{M-L+1}(\theta) + o(1) \quad a.s.\end{aligned} \tag{13}$$

for each θ, where function h_* is a known function (see Section III for more details). This would immediately imply that the traditional pseudo-spectrum estimate $\hat{\eta}_N^{(t)}(\theta)$ would verify

$$\hat{\eta}_N^{(t)}(\theta)=1-\sum_{k=1}^{K}h_*(\hat{\lambda}_{k,N})\mathbf{a}_{M-L+1}(\theta)^*\mathbf{u}_{k,N}\mathbf{u}_{k,N}^*\mathbf{a}_{M-L+1}(\theta) + o(1) \quad a.s. \tag{14}$$

and that the true MUSIC pseudo-spectrum

$$\begin{aligned}\eta_N(\theta) &= \mathbf{a}_{M-L+1}(\theta)^*\mathbf{\Pi}_N^{(L)}\mathbf{a}_{M-L+1}(\theta)\\ &= 1 - \mathbf{a}_{M-L+1}(\theta)^*\sum_{k=1}^{K}\mathbf{u}_{k,N}\mathbf{u}_{k,N}^*\mathbf{a}_{M-L+1}(\theta)\end{aligned}$$

could be estimated consistently by

$$\hat{\eta}_N(\theta) = 1 - \sum_{k=1}^{K}\mathbf{a}_{M-L+1}(\theta)^*\frac{\hat{\mathbf{u}}_{k,N}\hat{\mathbf{u}}_{k,N}^*}{h_*\left(\hat{\lambda}_{k,N}\right)}\mathbf{a}_{M-L+1}(\theta) \tag{15}$$

While matrix $\mathbf{Z}_N$ is of course not i.i.d. as soon as $L > 1$, we prove in this paper that the fundamental identity (13), in principle valid when $\mathbf{Z}_N$ is i.i.d., still holds in the asymptotic regime (11). Using the approach of [3], (13) appears as a consequence of the following results (see Proposition 1):

i) The eigenvalue distribution of matrix $\mathbf{Z}_N\mathbf{Z}_N^*$ converges almost surely towards the so-called Marcenko-Pastur (see Section III-B for more details). More importantly, almost surely, for each N large enough, the eigenvalues of $\mathbf{Z}_N\mathbf{Z}_N^*$ lie in the interval $[0, \sigma^2(1+\sqrt{c_*})^2 + \epsilon]$ where ϵ can be arbitrarily small.

ii) The entries of matrices $(\mathbf{Z}_N\mathbf{Z}_N^* - z\mathbf{I})^{-1}$, $(\mathbf{Z}_N^*\mathbf{Z}_N - z\mathbf{I})^{-1}$ behave as the entries of matrices $m_*(z)\mathbf{I}_M$ and $\tilde{m}_*(z)\mathbf{I}_{NL}$, where $m_*(z)$ represents the Stieltjes transform of the Marcenko-Pastur distribution and where $\tilde{m}_*(z) = c_*m_*(z) - (1-c_*)/z$, while the entries of $(\mathbf{Z}_N\mathbf{Z}_N^* - z\mathbf{I})^{-1}\mathbf{Z}_N$ converge almost surely towards 0.

(i) follows directly from [9] where it is shown that the eigenvalue distribution of matrix $\mathbf{Z}_N\mathbf{Z}_N^*$ converges almost surely towards the Marcenko-Pastur as soon as $N \to +\infty$ holds, and that the non zero eigenvalues of $\mathbf{Z}_N\mathbf{Z}_N^*$ are arbitrarily close from interval $[\sigma^2(1-\sqrt{c_*})^2, \sigma^2(1+\sqrt{c_*})^2]$ when N is large enough, provided parameter β defined in (11) satisfies $\beta > 1/3$. If however $\beta \leq 1/3$, the latter property is not guaranteed, and the general approach of [3] fails. This explains why parameter β cannot be arbitrarily small. (ii) does not follow directly from [9], and needs therefore some extra technical efforts (see Appendix A).

(i) and (ii) not only imply (14) and the consistency of $\hat{\eta}_N(\theta)$ for each θ, but also that

$$\sup_{\theta\in[-\pi,\pi]}\left|\hat{\eta}_N^{(t)}(\theta) - \left(1-\sum_{k=1}^{K}h_*(\hat{\lambda}_{k,N})\mathbf{a}_{M-L+1}(\theta)^*\mathbf{u}_{k,N}\mathbf{u}_{k,N}^*\mathbf{a}_{M-L+1}(\theta)\right)\right| \to 0 \quad a.s. \tag{16}$$

and

$$\sup_{\theta\in[-\pi,\pi]}|\hat{\eta}_N(\theta) - \eta_N(\theta)| \to 0 \quad a.s. \tag{17}$$

These uniform consistency properties allow to study the asymptotic behaviour of the traditional MUSIC SS estimates $(\hat{\theta}_{k,N}^{(t)})_{k=1,\ldots,K}$ and of the G-MUSIC SS estimates $(\hat{\theta}_{k,N})_{k=1,\ldots,K}$ defined as the K most significant local minima of $|\hat{\eta}_N(\theta)|$. More precisely, (16) and (17) allow to generalize immediately in the asymptotic regime (11) the proof of Theorem 3 of [6] and the proof of Theorem 1 of [19] (these theorems address the case $L = 1$), and to conclude that, under the separation condition (12), it holds that:

- $(\hat{\theta}_{k,N}^{(t)})_{k=1,\ldots,K}$ and $(\hat{\theta}_{k,N})_{k=1,\ldots,K}$ are consistent and verify

$$M\left(\hat{\theta}_{k,N}^{(t)} - \theta_k\right) \to 0 \ a.s. \tag{18}$$

$$M\left(\hat{\theta}_{k,N} - \theta_k\right) \to 0 \ a.s. \tag{19}$$

(18) and (19) hold when the DoA $(\theta_k)_{k=1,\ldots,K}$ are fixed parameters that do not depend on M and N. In practice, 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

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

tends to ∞. Adapting the proof of Theorem 3 of [19], we obtain that:

- If $K = 2$, $\frac{\mathbf{S}_N\mathbf{S}_N^*}{N} \to \mathbf{I}_2$, and if the 2 DoAs scale with M, N such that $\theta_{2,N} - \theta_{1,N} = \mathcal{O}\left(\frac{1}{M}\right)$, then the G-MUSIC SS estimates still verify (19) while the traditional MUSIC SS estimates no longer verify (18)

As in the case $L = 1$, the separation condition (12) ensures that the K largest eigenvalues of the empirical covariance matrix $\frac{\mathbf{Y}_N^{(L)}\mathbf{Y}_N^{(L)*}}{NL}$ 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 tend to ∞ 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

$$\lambda_K \left(\mathbf{A}_{M-L+1}^* \mathbf{A}_{M-L+1} \mathbf{D}\right) > \frac{\sigma^2 \sqrt{d_*}}{\sqrt{L}} \tag{20}$$

for each (M,N) large enough. If $L=1$, the separation condition introduced in the context of (unsmoothed) G-MUSIC algorithms ([6]) 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 $L \ll M$, 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 $L \ll M$. 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, as well as matrix $\mathbf{Z}_N$ defined by $\mathbf{Z}_N = \frac{1}{\sqrt{NL}} \mathbf{V}_N^{(L)}$. The entries of $\mathbf{Z}_N$ have of course variance σ^2/NL. In this section, $\mathbf{B}_N$ represents a deterministic $(M+L-1) \times NL$ matrix verifying

$$\sup_N \|\mathbf{B}_N\| < +\infty, \ \mathrm{Rank}(\mathbf{B}_N) = K \tag{21}$$

for each N large enough, and not necessarily matrix $\frac{1}{\sqrt{NL}} \mathbf{A}^{(L)} \left(\mathbf{S}_N \otimes \mathbf{I}_L\right)$ as in Section II-B.

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

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

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

$$\mathbf{X}_N = \mathbf{B}_N + \mathbf{Z}_N \tag{22}$$

$\mathbf{X}_N$ can thus be interpreted as a rank K perturbation of the random block-Hankel matrix $\mathbf{Z}_N$. The purpose of this section is to formalize claim (13), and to present rigorous results characterizing 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 $(\hat{\mathbf{u}}_{k,N})_{k=1,\ldots,K}$. As shown in Section III-B, matrix $\mathbf{Z}_N$ satisfies properties that allow to follow the approach of [3]. For the reader's convenience, we provide in Section III-A a short introduction to this approach in order to highlight the importance of the results of Section III-B devoted to the asymptotic properties of the eigenvalues of $\mathbf{Z}_N \mathbf{Z}_N^*$.

A. Introduction to the Approach of [3]

The approach of [3] allows to check if some of the K largest eigenvalues $(\hat{\lambda}_{k,N})_{k=1,\ldots,K}$ of matrix $\mathbf{X}_N \mathbf{X}_N^*$ escape from the interval $[0, \sigma^2(1+\sqrt{c_*})^2]$ when $N \to +\infty$, and to evaluate the behaviour of the corresponding eigenvectors. We present the formulation of subsection 5-3 and subsection 5-6 in [4], which, while being equivalent to [3], is more direct. For N large enough, it appears that the eigenvalues of $\mathbf{Z}_N \mathbf{Z}_N^*$ cannot exceed $\sigma^2(1+\sqrt{c_*})^2 + \epsilon$ where ϵ can be chosen arbitrarily small (see statement (ii) in Proposition 1 below). In order to characterize the eigenvalues of $\mathbf{X}_N \mathbf{X}_N^*$ that exceed $\sigma^2(1+\sqrt{c_*})^2 + \epsilon$, it is sufficient to express $\mathbf{X}_N \mathbf{X}_N^* - z\mathbf{I}$ as

$$\mathbf{X}_N \mathbf{X}_N^* - z\mathbf{I} = \mathbf{Z}_N \mathbf{Z}_N^* - z\mathbf{I} + \left(\mathbf{U}_N, \mathbf{Z}_N \tilde{\mathbf{U}}_N \mathbf{\Lambda}_N^{\frac{1}{2}}\right) \times \begin{pmatrix} \mathbf{\Lambda}_N & \mathbf{I}_K \\ \mathbf{I}_K & 0 \end{pmatrix} \begin{pmatrix} \mathbf{U}_N^* \\ \mathbf{\Lambda}_N^{\frac{1}{2}} \tilde{\mathbf{U}}_N^* \mathbf{Z}_N^* \end{pmatrix}$$

If z is chosen real and greater than $\sigma^2(1+\sqrt{c_*})^2 + \epsilon$, matrix $\mathbf{Z}_N \mathbf{Z}_N^* - z\mathbf{I}$ is invertible, and if we denote by $\mathbf{Q}_N(z)$ the so-called resolvent of $\mathbf{Z}_N \mathbf{Z}_N^*$ defined by

$$\mathbf{Q}_N(z) = \left(\mathbf{Z}_N \mathbf{Z}_N^* - z\mathbf{I}\right)^{-1} \tag{23}$$

then, $\mathbf{X}_N \mathbf{X}_N^* - z\mathbf{I}$ can be written as

$$\mathbf{X}_N \mathbf{X}_N^* - z\mathbf{I} = \left(\mathbf{Z}_N \mathbf{Z}_N^* - z\mathbf{I}\right) \times \left(\mathbf{I} + \mathbf{Q}_N(z) \left(\mathbf{U}_N, \mathbf{Z}_N \tilde{\mathbf{U}}_N \mathbf{\Lambda}_N^{\frac{1}{2}}\right) \begin{pmatrix} \mathbf{\Lambda}_N & \mathbf{I}_K \\ \mathbf{I}_K & 0 \end{pmatrix} \begin{pmatrix} \mathbf{U}_N^* \\ \mathbf{\Lambda}_N^{\frac{1}{2}} \tilde{\mathbf{U}}_N^* \mathbf{Z}_N^* \end{pmatrix}\right) \tag{24}$$

Therefore, if $z > \sigma^2(1+\sqrt{c_*})^2 + \epsilon$, z is eigenvalue of $\mathbf{X}_N \mathbf{X}_N^*$ if and only if

$$\det\left(\mathbf{I} + \mathbf{Q}_N(z) (\mathbf{U}_N, \mathbf{Z}_N \tilde{\mathbf{U}}_N \mathbf{\Lambda}_N^{\frac{1}{2}}) \begin{pmatrix} \mathbf{\Lambda}_N & \mathbf{I}_K \\ \mathbf{I}_K & 0 \end{pmatrix} \times \begin{pmatrix} \mathbf{U}_N^* \\ \mathbf{\Lambda}_N^{\frac{1}{2}} \tilde{\mathbf{U}}_N^* \mathbf{Z}_N^* \end{pmatrix}\right) = 0$$

or equivalently, if and only if $\det(\mathbf{T}_N(z)) = 0$ where $\mathbf{T}_N(z)$ is the $2K \times 2K$ matrix defined by

$$\mathbf{T}_N(z) = \mathbf{I}_{2K} + \begin{pmatrix} \mathbf{U}_N^* \\ \mathbf{\Lambda}_N^{\frac{1}{2}} \tilde{\mathbf{U}}_N^* \mathbf{Z}_N^* \end{pmatrix} \mathbf{Q}_N(z) \times \left(\mathbf{U}_N, \mathbf{Z}_N \tilde{\mathbf{U}}_N \mathbf{\Lambda}_N^{\frac{1}{2}}\right) \begin{pmatrix} \mathbf{\Lambda}_N & \mathbf{I}_K \\ \mathbf{I}_K & 0 \end{pmatrix}$$

It turns out that it is possible to evaluate the behaviour of the entries of matrix $\mathbf{T}_N(z)$ when $N \to +\infty$. More precisely, the entries of $\mathbf{T}_N(z)$ depend on bilinear forms of matrices $\mathbf{Q}_N(z)$, $\mathbf{Q}_N(z)\mathbf{Z}_N$, and $\mathbf{Z}_N^*\mathbf{Q}_N(z)\mathbf{Z}_N = \mathbf{I} + z\tilde{\mathbf{Q}}_N(z)$ where $\tilde{\mathbf{Q}}_N(z)$ is the resolvent of matrix $\mathbf{Z}_N^*\mathbf{Z}_N$ defined by

$$\tilde{\mathbf{Q}}_N(z) = \left(\mathbf{Z}_N^*\mathbf{Z}_N - z\mathbf{I}\right)^{-1} \tag{25}$$

Statement (iii) of Proposition 1 shows that it is possible to characterize the asymptotic behaviour of these bilinear forms, and thus the limit form of the equation verified by the eigenvalues of $\mathbf{X}_N\mathbf{X}_N^*$ that exceed $\sigma^2(1+\sqrt{c_*})^2 + \epsilon$. This analysis allows to establish (38) and (39).

In order to evaluate the behaviour of the eigenvector $\hat{\mathbf{u}}_{k,N}$ associated to an eigenvalue $\hat{\lambda}_{k,N}$ that converges towards a value ρ_k that exceeds $\sigma^2(1+\sqrt{c_*})^2$, we use the identity

$$\hat{\mathbf{u}}_{k,N}\hat{\mathbf{u}}_{k,N}^* = \frac{1}{2i\pi}\int_{\mathcal{C}_k} \left(\mathbf{X}_N\mathbf{X}_N^* - z\mathbf{I}\right)^{-1} dz \tag{26}$$

where $\mathcal{C}_k$ is a contour enclosing the eigenvalue $\hat{\lambda}_{k,N}$, and not the other eigenvalues of $\mathbf{X}_N\mathbf{X}_N^*$. In order to obtain (40), it is sufficient to express matrix $\left(\mathbf{X}_N\mathbf{X}_N^* - z\mathbf{I}\right)^{-1}$ in terms of $\mathbf{Q}_N(z)$ using (24), to evaluate the asymptotic behaviour of the corresponding entries using statement (iii) of Proposition 1, and eventually to compute the limiting behaviour of the contour integral at the righthandside of (26).

B. 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

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

is known to satisfy the fundamental equation

$$m_{\sigma^2,c}(z) = \frac{1}{-z + \sigma^2\frac{1}{1+\sigma^2 c m_{\sigma^2,c}(z)}} \tag{27}$$

or equivalently,

$$m_{\sigma^2,c}(z) = \frac{1}{-z\left(1+\sigma^2\tilde{m}_{\sigma^2,c}(z)\right)} \tag{28}$$

$$\tilde{m}_{\sigma^2,c}(z) = \frac{1}{-z\left(1+\sigma^2 c m_{\sigma^2,c}(z)\right)} \tag{29}$$

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 (28) and (29) 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

$$w_*(z) = \frac{1}{z\, m_*(z)\, \tilde{m}_*(z)} \tag{30}$$

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

$$\phi_*(w) = \frac{(w+\sigma^2)(w+\sigma^2 c_*)}{w} \tag{31}$$

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

$$\phi_*\left(w_*(z)\right) = z \tag{32}$$

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

The main results of this paper are based on 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 $z \in \mathbb{C} - \mathbb{R}^+$,

$$\frac{1}{M-L+1}\mathrm{Tr}\left(\mathbf{Q}_N(z)\right) - m_*(z) \to 0 \ a.s. \tag{33}$$

- (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 $c_* \le 1$, 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 $z \in \mathbb{C} - \mathbb{R}^+$,

$$\mathbf{a}_N^*\left(\mathbf{Q}_N(z) - m_*(z)\mathbf{I}\right)\mathbf{b}_N \to 0 \ a.s. \tag{34}$$

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 $z \in \mathbb{C} - \mathbb{R}^+$, it holds that

$$\tilde{\mathbf{a}}_N^*\left(\tilde{\mathbf{Q}}_N(z) - \tilde{m}_*(z)\mathbf{I}\right)\tilde{\mathbf{b}}_N \to 0 \ a.s. \tag{35}$$

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

$$\mathbf{a}_N^*\left(\mathbf{Q}_N(z)\mathbf{Z}_N\right)\tilde{\mathbf{b}}_N \to 0 \ a.s. \tag{36}$$

Finally, for each $\epsilon > 0$, convergence properties (34)–(36) hold uniformly w.r.t. z on each compact subset of $\mathbb{C} - [0, x_*^+ + \epsilon]$.

(i) and (ii) follow directly from [9] while (iii) requires some additional work. See Appendix A for more details.

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, (33) was established for the first time in [10], the almost sure location of the eigenvalues of $\mathbf{Z}_N\mathbf{Z}_N^*$ can be found in [1] (see Theorem 5–11), while (34), (35) and (36) are trivial modifications of Lemma 5 of [6].

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 $0 < \beta \le 1$, or equivalently if $L = \mathcal{O}(M^\alpha)$ for each $0 \le \alpha < 1$. L can therefore tends to ∞ 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).

C. The K Largest Eigenvalues and Eigenvectors of $\mathbf{X}_N\mathbf{X}_N^*$

As mentioned in Section II-B, 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

$$\lambda_s > \sigma^2\sqrt{c_*} \tag{37}$$

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

$$\hat{\lambda}_{k,N} \xrightarrow[N\to\infty]{a.s.} \rho_k = \phi(\lambda_k) = \frac{(\lambda_k + \sigma^2)(\lambda_k + \sigma^2 c)}{\lambda_k} > x_*^+ \tag{38}$$

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

$$\hat{\lambda}_{k,N} \to x_*^+ \ a.s. \tag{39}$$

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

$$\mathbf{d}_{1,N}^*\hat{\mathbf{u}}_{k,N}\hat{\mathbf{u}}_{k,N}^*\mathbf{d}_{2,N} = h_*(\rho_k)\mathbf{d}_{1,N}^*\mathbf{u}_{k,N}\mathbf{u}_{k,N}^*\mathbf{d}_{2,N} + o(1) \quad a.s., \tag{40}$$

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

$$h_*(z) = \frac{w_*(z)^2 - \sigma^4 c_*}{w_*(z)(w_*(z) + \sigma^2 c_*)} \tag{41}$$

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 $\frac{1}{L}\mathbf{A}^{(L)}(\mathbf{S}_N\mathbf{S}_N^*/N \otimes \mathbf{I}_L)\mathbf{A}^{(L)*}$. 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 [6] and [19] 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

$$\lambda_K > \sigma^2\sqrt{c_*} \tag{42}$$

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

$$\hat{\eta}_N(\theta) = (\mathbf{a}_{M-L+1}(\theta))^* \left(\mathbf{I} - \sum_{k=1}^{K} \frac{\hat{\mathbf{u}}_{k,N}\hat{\mathbf{u}}_{k,N}^*}{h_*\left(\hat{\lambda}_{k,N}\right)}\right) \mathbf{a}_{M-L+1}(\theta) \tag{43}$$

is consistent, and verifies moreover

$$\sup_{\theta\in[-\pi,\pi]} |\hat{\eta}_N(\theta) - \eta_N(\theta)| \xrightarrow[n\to\infty]{a.s.} 0, \tag{44}$$

The consistency of estimator (43) is a direct consequence of (40) and (38). The uniform consistency property (44) can be proved as Proposition 1 in [6]. We notice that the proof of this Proposition uses extensively Lemma 5 in [6], which, in the context of the present paper has to be replaced by item (iii) of Proposition 1.

In order to obtain some insights on condition (42) and on the potential benefits of the spatial smoothing, we explicit the separation condition (42) when M and N tend to ∞ 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

$$\begin{aligned}&\frac{1}{L}\mathbf{A}^{(L)}\left(\frac{\mathbf{S}_N\mathbf{S}_N^*}{N} \otimes \mathbf{I}_L\right)\mathbf{A}^{(L)*}\\ &= \frac{M-L+1}{M}\,\mathbf{A}_{M-L+1}\left(\frac{\mathbf{S}_N\mathbf{S}_N^*}{N} \bullet \mathbf{A}_L^T\overline{\mathbf{A}}_L\right)\mathbf{A}_{M-L+1}^*\end{aligned} \tag{45}$$

where $\bullet$ represents the Hadamard (i.e. element wise) product of matrices, and where $\overline{\mathbf{A}}_L$ stands for the complex conjugation operator of the elements of matrix $\mathbf{A}_L$. 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 $\mathbf{D} \bullet \mathrm{Diag}(\mathbf{A}_L^T\overline{\mathbf{A}}_L) = \mathbf{D}$. Therefore, $\frac{\mathbf{S}_N\mathbf{S}_N^*}{N} \bullet (\mathbf{A}_L^T\overline{\mathbf{A}}_L) \simeq \mathbf{D}$ when N 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

$$\lambda_K\left(\mathbf{A}_{M-L+1}^*\mathbf{A}_{M-L+1}\mathbf{D}\right) > \frac{\sigma^2\sqrt{d_*}}{\sqrt{L}} \tag{46}$$

for each (M, N) large enough. If $L = 1$, the separation condition introduced in the context of (unsmoothed) G-MUSIC algorithms ([6]) 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 $L \ll M$, 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 $L \ll M$. 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 $(\hat{\theta}_{k,N})_{k=1,\dots,K}$ by

$$\hat{\theta}_{k,N} = \underset{\theta \in \mathcal{I}_k}{\operatorname{argmin}} \left|\hat{\eta}_N(\theta)\right|, \tag{47}$$

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 [6], the uniform consistency (44) as well as the particular structure of directional vectors $\mathbf{a}_{M-L+1}(\theta)$ imply the following result which can be proved like Theorem 3 of [6].

Theorem 3: Under condition (42), the DoA G-MUSIC SS estimates $(\hat{\theta}_{k,N})_{k=1,\dots,K}$ verify

$$M\left(\hat{\theta}_{k,N} - \theta_k\right) \to 0 \ a.s. \tag{48}$$

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

Remark 2: We remark that under the extra assumption that $\frac{\mathbf{S}_N\mathbf{S}_N^*}{N}$ converges towards a diagonal matrix, [6] (see also [20] for more general matrices $\mathbf{S}$) proved when $L = 1$ that $M^{3/2}(\hat{\theta}_{k,N} - \theta_k)$ converges in distribution towards a Gaussian distribution. It would be interesting to generalize the results of [6] and [20] 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 [19] concerning the consistency of the traditional estimates $(\hat{\theta}_{k,N}^{(t)})_{k=1,\dots,K}$ in the case $L = 1$. In particular, while the traditional estimate $\hat{\eta}_N^{(t)}(\theta)$ of the pseudo-spectrum is not consistent, it is shown in [19] (see Theorem 1) that if $L = 1$, then the arguments of its local minima $(\hat{\theta}_{k,N}^{(t)})_{k=1,\dots,K}$ are consistent and verify

$$M\left(\hat{\theta}_{k,N}^{(t)} - \theta_k\right) \to 0 \ a.s. \tag{49}$$

for each $k = 1, \dots, K$ if the separation condition is verified. The proof of Theorem 1 in [19] can be immediately adapted to the context of the present paper. For this, it is sufficient to follow the proof of [19], and to use Theorem 1, as well as the uniform consistency property

$$\sup_{\theta \in [-\pi,\pi]} \left| \hat{\eta}_N^{(t)}(\theta) - \left(1 - \sum_{k=1}^{K} h_*(\rho_k)\mathbf{a}_{M-L+1}(\theta)^*\mathbf{u}_{k,N}\mathbf{u}_{k,N}^*\mathbf{a}_{M-L+1}(\theta)\right)\right| \to 0 \ a.s. \tag{50}$$

which can be proved in the same way that (44). We note that, as $\hat{\lambda}_{k,N} \to \rho_k$, then (50) and (16) are equivalent. Therefore, the following result holds.

Theorem 4: Under condition (42), the DoA traditional MUSIC SS estimates $(\hat{\theta}_{k,N}^{(t)})_{k=1,\dots,K}$ verify

$$M\left(\hat{\theta}_{k,N}^{(t)} - \theta_k\right) \to 0 \ a.s. \tag{51}$$

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

Remark 3: It is established in [19] in the case $L = 1$ that if $\frac{\mathbf{S}_N\mathbf{S}_N^*}{N}$ converges towards a diagonal matrix, then $M^{3/2}(\hat{\theta}_{k,N}^{(t)} - \theta_k)$ has a Gaussian behaviour, and that the corresponding variance coincides with the asymptotic variance of $M^{3/2}(\hat{\theta}_{k,N} - \theta_k)$. 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 (49) are valid when the DoAs $(\theta_k)_{k=1,\dots,K}$ are fixed parameters, i.e. do not depend on M and N. Therefore, the ratio

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

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 $L = 1$, [19] also addressed the case where the DoA's $(\theta_{k,N})_{k=1,\dots,K}$ depend on N, M and verify $\theta_{k,N} - \theta_{l,N} = \mathcal{O}\left(\frac{1}{M}\right)$. 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, [19] 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 [19] that the G-MUSIC estimates still verify (48), but that, in general, $M(\hat{\theta}_{k,N}^{(t)} - \theta_k)$ does not converge towards 0. The results of [19] 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 [19], 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 [19]. 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

$$1 - \left|\frac{\operatorname{sinc} \kappa}{2}\right| > \sigma^2 c_* \tag{52}$$

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

$$\hat{\theta}_{k,N} = \underset{\theta \in \mathcal{I}_{k,N}}{\operatorname{argmin}} \left|\hat{\eta}_N(\theta)\right| \tag{53),}$$

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

$$M(\hat{\theta}_{k,N} - \theta_{k,N}) \to 0 \ a.s. \tag{54}$$

In general, the traditional MUSIC SS estimates defined by (53) when the G-MUSIC estimate $\hat{\eta}_N(\theta)$ is replaced by the traditional spectrum estimate $\hat{\eta}_N^{(t)}(\theta)$ are such that $M(\hat{\theta}_{k,N}^{(t)} - \theta_{k,N})$ 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 such that the 2 source signals have power 1. The signal to noise ratio is thus equal to $\text{SNR} = 1/\sigma^2$.

TABLE I
MINIMUM VALUE OF SNR FOR SEPARATION CONDITION

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

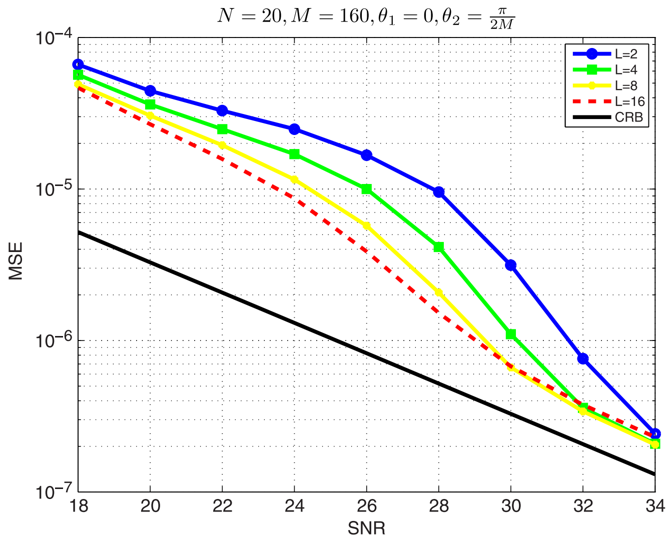

Fig. 1. Empirical MSE of G-MUSIC SS estimator $\hat{\theta}_1$ versus SNR.

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.

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

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 Fig. 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 are also represented. As expected, it is seen that the performance tends to increase with L until $L = 16$. In Fig. 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.

In Fig. 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.

We finally consider the case $L = 128$, and compare in Fig. 4 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

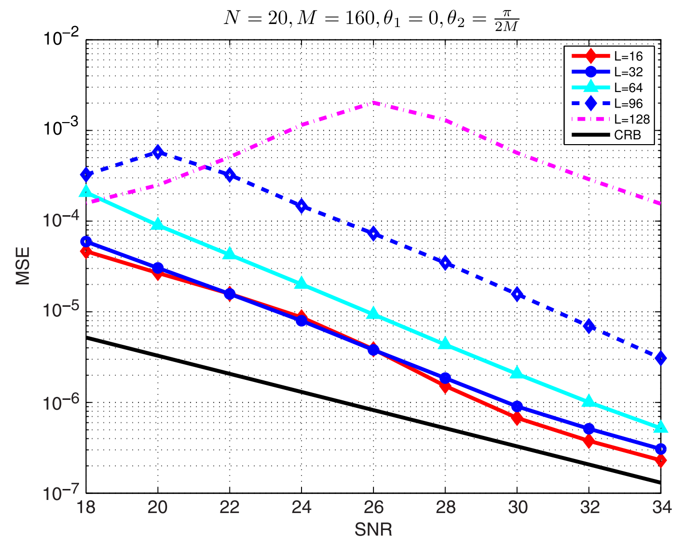

Fig. 2. Empirical MSE of G-MUSIC SS estimator $\hat{\theta}_1$ versus SNR.

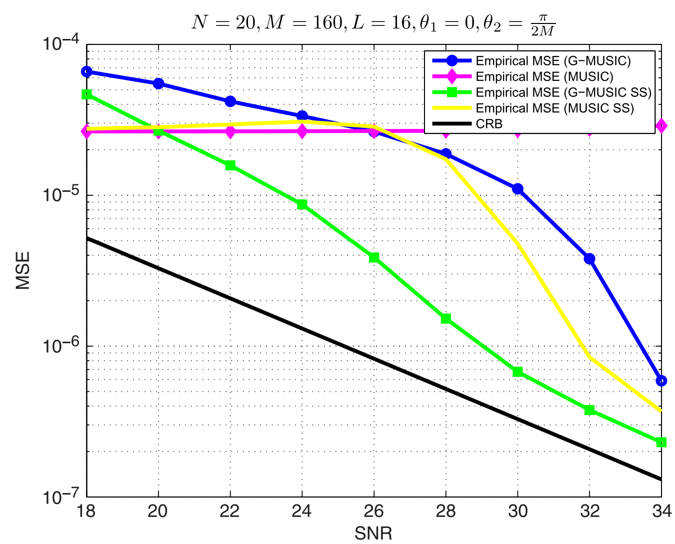

Fig. 3. Empirical MSE of different estimators of θ_1 when $L = 16$.

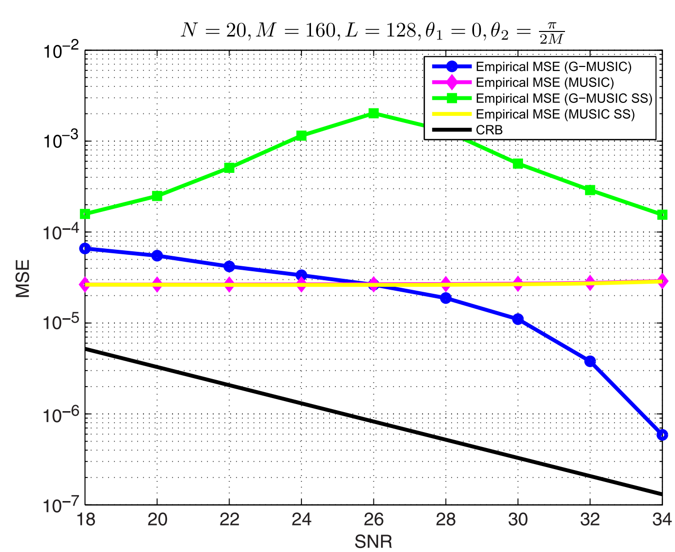

Fig. 4. Empirical MSE of different estimators of θ_1 when $L = 128$.

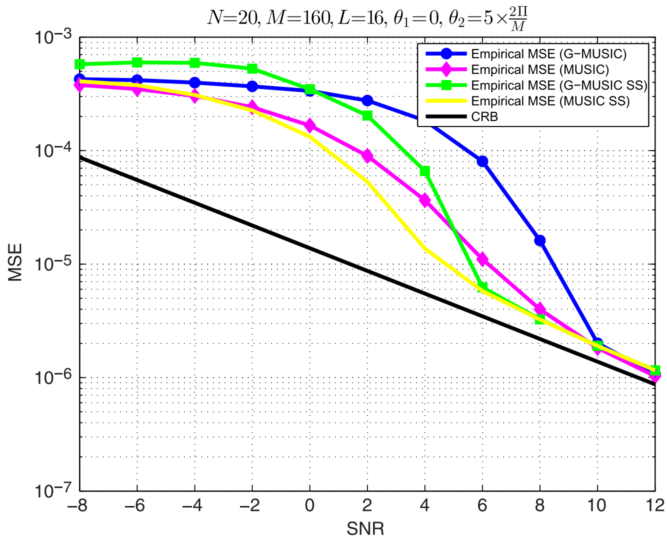

Fig. 5. Empirical MSE of different estimators of θ_1 when L = 16 and widely spaced DoAs.

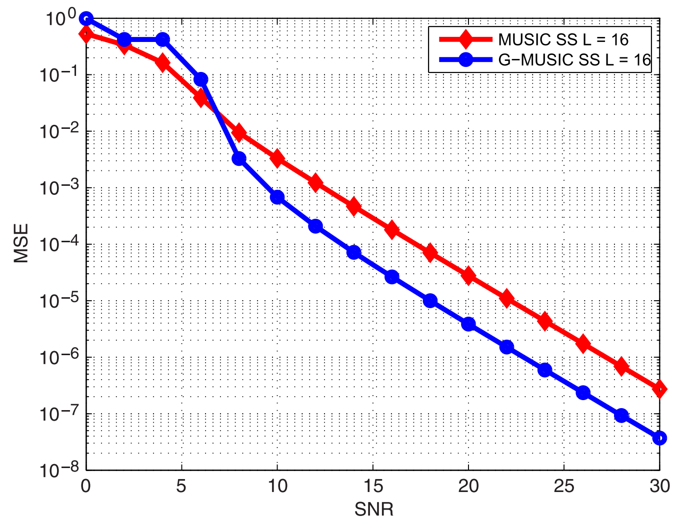

Fig. 6. Mean of the MSE(localization function), K = 5, L = 16.

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 confirms that the use of spatial smoothing schemes allows to improve the performance.

We finally consider the case of $K = 5$ sources located at $-2\pi/18, -\pi/18, 0, \pi/18, 2\pi/18$ while M and N are still equal to 160 and 20, and $L = 16$. We evaluate by Monte-Carlo simulations $\frac{1}{5}\sum_{k=1}^{5}\mathbb{E}|\hat{\eta}_N(\theta_k) - \eta_N(\theta_k)|^2$ as well as $\frac{1}{5}\sum_{k=1}^{5}\mathbb{E}|\hat{\eta}_N^{(t)}(\theta_k) - \eta_N(\theta_k)|^2$ i.e. the means of the MSE of the estimated localization functions, evaluated at the true angles. We recall that the G-MUSIC SS estimate of the localization function is consistent, but that it is not the case of the MUSIC SS estimate. This is confirmed by Fig. 6 which shows that the MSE of the G-MUSIC SS estimate is significantly smaller than the MSE of MUSIC SS estimate. It is also seen that the MSE of the MUSIC SS estimate converges towards 0 when the signal to noise ratio tends to ∞. This is because for each $z > x_*^+$, $h_*(z) \to 1$ if $\sigma^2 \to 0$. Therefore, (50) implies that $\hat{\eta}_N^{(t)}(\theta) \simeq \eta_N(\theta)$ for each θ when $\sigma^2 \simeq 0$.

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 tend to ∞ 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 [6], and have deduced from [19] their properties.

Appendix A
Proof of Proposition 1

The proof of Proposition 1 is based on the results of [9]. In order to explain this, we denote by $\mathbf{W}_N$ the $NL \times (M-L+1)$ matrix defined by

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

$\mathbf{W}_N$ can be written as $\mathbf{W}_N = \left(\mathbf{W}_N^{(1)T}, \ldots, \mathbf{W}_N^{(N)T}\right)^T$ where matrices $(\mathbf{W}_N^{(n)})_{n=1,\ldots,N}$ are independent identically distributed $L \times (M-L+1)$ Hankel matrices built from i.i.d. standard complex Gaussian sequences with variance $\frac{\sigma^2}{M-L+1}$. [9] studied the asymptotic behaviour of the empirical eigenvalue distribution of $\mathbf{W}_N\mathbf{W}_N^*$ as well the almost sure location of its eigenvalues in the asymptotic regime (11). As $\mathbf{Z}_N^*\mathbf{Z}_N$ coincides with $c_N\mathbf{W}_N\mathbf{W}_N^*$ and that, apart 0, the eigenvalues of $\mathbf{Z}_N^*\mathbf{Z}_N$ and $\mathbf{Z}_N\mathbf{Z}_N^*$ coincide, it is clear that the results of [9] can be used in order to establish items (i) and (ii) of Proposition 1. To help the reader to connect the results of [9] to the context of the present paper, we mention that the integers (M, N) in [9] should be exchanged by $(N, M-L+1)$.

We omit the proof of (i) and first briefly justify (ii). For this, we mention that Theorem 1.1 in [9] implies that almost surely, for each $\delta > 0$ and for N large enough, all the eigenvalues of $\mathbf{W}_N\mathbf{W}_N^*$ are located in $[\sigma^2(1-\sqrt{c_*^{-1}})^2 - \delta, \sigma^2(1+\sqrt{c_*^{-1}})^2 + \delta]$ if $c_*^{-1} \leq 1$, and in $[\sigma^2(1-\sqrt{c_*^{-1}})^2 - \delta, \sigma^2(1+\sqrt{c_*^{-1}})^2 + \delta] \cup \{0\}$ if $c_*^{-1} > 1$. As $\mathbf{Z}_N^*\mathbf{Z}_N = c_N\mathbf{W}_N\mathbf{W}_N^*$, and that $c_N \to c_*$, we obtain that all the eigenvalues of $\mathbf{Z}_N^*\mathbf{Z}_N$ belong to $[\sigma^2(1-\sqrt{c_*})^2 - \delta c_*, \sigma^2(1+\sqrt{c_*})^2 + \delta c_*]$ if $c_*^{-1} \leq 1$, and in $[\sigma^2(1-\sqrt{c_*})^2 - \delta c_*, \sigma^2(1+\sqrt{c_*})^2 + \delta c_*] \cup \{0\}$ if $c_*^{-1} > 1$. As the non zero eigenvalues of $\mathbf{Z}_N^*\mathbf{Z}_N$ and of $\mathbf{Z}_N\mathbf{Z}_N^*$ coincide, we deduce immediately that (ii) holds.

(iii) depends on the asymptotic properties of $\mathbf{Q}_N(z)$ and $\tilde{\mathbf{Q}}_N(z)$. If we denote by $\mathbf{Q}_{N,W}(z)$ and $\tilde{\mathbf{Q}}_{N,W}(z)$ the resolvents of matrices $\mathbf{W}_N\mathbf{W}_N^*$ and $\mathbf{W}_N^*\mathbf{W}_N$ respectively, it is easily seen that $\mathbf{Q}_N(z) = c_N^{-1}\tilde{\mathbf{Q}}_{N,W}(zc_N^{-1})$ and that $\tilde{\mathbf{Q}}_N(z) = c_N^{-1}\mathbf{Q}_{N,W}(zc_N^{-1})$. As $c_N \to c_*$, $\mathbf{Q}_N(z)$ and $\tilde{\mathbf{Q}}_N(z)$ behave as $c_*^{-1}\tilde{\mathbf{Q}}_{N,W}(zc_*^{-1})$ and $c_*^{-1}\mathbf{Q}_{N,W}(zc_*^{-1})$. As $z \in \mathbb{C} - \mathbb{R}^+$ if and only if $zc_*^{-1} \in \mathbb{C} - \mathbb{R}^+$, in order to establish (35) and (34), it is sufficient to establish the following properties of $\mathbf{Q}_{N,W}$ and $\tilde{\mathbf{Q}}_{N,W}$:

$$\tilde{\mathbf{a}}_N^* \left(\mathbf{Q}_{N,W}(z) - t_*(z)\mathbf{I}\right) \tilde{\mathbf{b}}_N \to 0 \; a.s. \tag{55}$$

(equivalent to (35)), and

$$\mathbf{a}_N^* \left(\tilde{\mathbf{Q}}_{N,W}(z) - \tilde{t}_*(z)\mathbf{I}\right) \mathbf{b}_N \to 0 \; a.s. \tag{56}$$

(equivalent to (34)) for each $z \in \mathbb{C} - \mathbb{R}^+$. Here, $t_*(z)$ and $\tilde{t}_*(z)$ are the Stieltjes transforms of the Marcenko-Pastur distributions of parameters (σ^2, c_*^{-1}) and $(\sigma^2 c_*^{-1}, c_*)$ which are related to $m_*(z)$ and $\tilde{m}_*(z)$ by the relations $m_*(z) = c_*^{-1}\, \tilde{t}_*(zc_*^{-1})$ and $\tilde{m}_*(z) = c_*^{-1}\, t_*(zc_*^{-1})$.

Proof of (55): While (55) does not appear explicitly in [9], it can be deduced rather easily from the various intermediate results proved in [9]. For this, we first remark that

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

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 [9]). More precisely, ξ can be considered as a smooth function $\xi(\mathbf{W}_N, \overline{\mathbf{W}}_N)$ of the entries of $\mathbf{W}_N$ and of matrix $\overline{\mathbf{W}}_N$ whose entries are the complex conjugates of the entries of $\mathbf{W}_N$. For each $n = 1, \ldots, N$, we denote by $\mathbf{W}_{i,j}^n$ the entry (i,j) of $L \times (M - L + 1)$ matrix $\mathbf{W}_N^{(n)}$. Then, the Poincaré-Nash inequality is a concentration inequality which states that

$$\begin{aligned}
\mathrm{Var}(\xi) \leq &\sum_{n,n'} \sum_{i,j,i',j'} \mathbb{E}\left[\frac{\partial \xi}{\partial \mathbf{W}_{i,j}^n} \mathbb{E}\left(\mathbf{W}_{i,j}^n (\mathbf{W}_{i',j'}^{n'})^*\right) \left(\frac{\partial \xi}{\partial \mathbf{W}_{i',j'}^{n'}}\right)^*\right] \\
+ &\sum_{n,n'} \sum_{i,j,i',j'} \mathbb{E}\left[\left(\frac{\partial \xi}{\partial \overline{\mathbf{W}}_{i,j}^n}\right)^* \mathbb{E}\left(\mathbf{W}_{i,j}^n (\mathbf{W}_{i',j'}^{n'})^*\right) \left(\frac{\partial \xi}{\partial \overline{\mathbf{W}}_{i',j'}^{n'}}\right)\right]
\end{aligned}$$

We notice that the structure of $\mathbf{W}_N$ implies that

$$\mathbb{E}\left(\mathbf{W}_{i,j}^n (\mathbf{W}_{i',j'}^{n'})^*\right) = \frac{\sigma^2}{M - L + 1} \delta(n - n')\delta(i + j = i' + j')$$

so that the above sums reduce to simpler terms. The above upper bound of $\mathrm{Var}(\xi)$ was evaluated in Proposition 3-1 in [9] (see Eq. (3.2)). Exchanging (M, N) by $(N, M - L + 1)$ in this proposition, we obtain immediately that $\mathrm{Var}(\xi) = \mathbb{E}|\xi|^2 = \mathcal{O}\left(\frac{L}{M-L+1}\right) = \mathcal{O}\left(\frac{L}{M}\right)$. 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 k, it holds that

$$\mathbb{E}|\xi|^{2k} = \mathcal{O}\left(\left(\frac{L}{M}\right)^k\right) \tag{57}$$

(57) can be established by induction on k. As mentioned above, (57) is verified for $k = 1$. We now assume that it holds until integer $k - 1$, and prove (57). For this, we use the obvious relation:

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

In order to manage $\mathrm{Var}(\xi^k)$, we use again the Poincaré-Nash inequality. As

$$\frac{\partial \xi^k}{\partial \mathbf{W}_{i,j}^n} = k\xi^{k-1} \frac{\partial \xi}{\partial \mathbf{W}_{i,j}^n}$$

the Poincaré-Nash inequality leads to

$$\begin{aligned}
&\mathrm{Var}(\xi^k) \\
&\leq k \sum_{n,n'} \sum_{i,j,i',j'} \\
&\mathbb{E}\left[|\xi|^{2k-2} \frac{\partial \xi}{\partial \mathbf{W}_{i,j}^n} \mathbb{E}\left(\mathbf{W}_{i,j}^n (\mathbf{W}_{i',j'}^{n'})^*\right) \left(\frac{\partial \xi}{\partial \mathbf{W}_{i',j'}^{n'}}\right)^*\right]
\end{aligned} \tag{58}$$

$$\begin{aligned}
&+ k \sum_{n,n'} \sum_{i,j,i',j'} \\
&\mathbb{E}\left[|\xi|^{2k-2} \left(\frac{\partial \xi}{\partial \overline{\mathbf{W}}_{i,j}^n}\right)^* \mathbb{E}\left(\mathbf{W}_{i,j}^n (\mathbf{W}_{i',j'}^{n'})^*\right) \left(\frac{\partial \xi}{\partial \overline{\mathbf{W}}_{i',j'}^{n'}}\right)\right]
\end{aligned} \tag{59}$$

Following the proof of Proposition 3-1 in [9], it is easy to check that the Poincaré-Nash inequality leads to

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

where C is a constant that depends on z but not on the dimensions L, M, N. As (57) is assumed to hold until integer $k - 1$, this implies that $\mathrm{Var}(\xi^k) = \mathcal{O}((L/M)^k)$. The Schwartz inequality leads immediately to

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

which is a $\mathcal{O}((L/M)^k)$ term. This establishes (57). As $L = \mathcal{O}(M^\alpha)$ with $\alpha < 2/3$, it is clear that $(L/M)^3$ verifies

$$\left(\frac{L}{M}\right)^3 = \mathcal{O}\left(\frac{1}{M^{1+2-3\alpha}}\right)$$

Therefore, (57) 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

$$\tilde{\mathbf{a}}_N^* \left(\mathbb{E}\left(\mathbf{Q}_{N,W}(z)\right) - t_*(z)\mathbf{I}\right) \tilde{\mathbf{b}}_N \to 0 \tag{60}$$

For this, we first simplify the notations and denote by $\mathbf{W}, \tilde{\mathbf{W}}, \mathbf{Q}$, the matrices $\mathbf{W}_N, \tilde{\mathbf{W}}_N, \mathbf{Q}_{N,W}(z)$. 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 $(i_1 + (n_1 - 1)L, i_2 + (n_2 - 1)L)$. We also denote

$(\mathbf{w}_j)_{j=1,\ldots,M-L+1}$ the columns of $\mathbf{W}$. Although it is not stated explicitly in [9], (60) can be deduced from various intermediate evaluations. In order to be more specific, we mention that it is proved in [9] that matrix $\mathbb{E}(\mathbf{Q})$ can be written as

$$\mathbb{E}\left(\mathbf{Q}(z)\right) = \mathbf{I}_N \otimes \mathbf{R}_N(z) + \mathbf{\Delta}_N(z) \tag{61}$$

(see Eq. (4.14) in [9]) where $\mathbf{R}_N(z)$ is a $L \times L$ matrix whose expression is omitted, and where $\mathbf{\Delta}_N(z)$ is shown to verify $\tilde{\mathbf{a}}_N^* \mathbf{\Delta}_N(z) \tilde{\mathbf{b}}_N \to 0$ using the Poincaré-Nash inequality (see Eq. (5.3) in [9]). As this will be useful to establish (56), we give some insights on the proof of (61). [9] uses the identity

$$\mathbb{E}\left[(\mathbf{Q}_{i_1,i_2}^{n_1,n_2}\right] = -\frac{1}{z}\delta(i_1-i_2)\delta(n_1-n_2)+\frac{1}{z}\mathbb{E}\left[(\mathbf{Q}\mathbf{W}\mathbf{W}^*)_{i_1,i_2}^{n_1,n_2}\right] \tag{62}$$

It turns out that the second term of the righthandside of (62) can be expressed in terms of the entries of $\mathbb{E}(\mathbf{Q})$ and of other terms that tend to 0. To obtain the corresponding expression, [9] evaluates $\mathbb{E}[(\mathbf{Q}\mathbf{w}_k\mathbf{w}_j^*)_{i_1,i_2}^{n_1,n_2}] = \mathbb{E}[(\mathbf{Q}\mathbf{w}_k)_{i_1}^{n_1}(\mathbf{w}_j^*)_{i_2}^{n_2}]$ for each k, j, i_1, i_2, n_1, n_2. For this, the identity

$$\mathbb{E}\left[(\mathbf{Q}\mathbf{w}_k)_{i_1}^{n_1}\left(\mathbf{w}_j^*\right)_{i_2}^{n_2}\right] = \sum_{i_3,n_3}\mathbb{E}\left(\mathbf{Q}_{i_1,i_3}^{n_1,n_3}\mathbf{W}_{i_3,k}^{n_3}\overline{\mathbf{W}}_{i_2,j}^{n_2}\right)$$

and the so-called the integration by parts formula (see e.g. Proposition 1.1 in [9])

$$\mathbb{E}\left(\mathbf{Q}_{i_1,i_3}^{n_1,n_3}\mathbf{W}_{i_3,k}^{n_3}\overline{\mathbf{W}}_{i_2,j}^{n_2}\right) = \sum_{i',j'}\mathbb{E}\left(\mathbf{W}_{i_3,k}^{n_3}\overline{\mathbf{W}}_{i',j'}^{n_3}\right)\mathbb{E}\left[\frac{\partial\left(\mathbf{Q}_{i_1,i_3}^{n_1,n_3}\overline{\mathbf{W}}_{i_2,j}^{n_2}\right)}{\partial\overline{\mathbf{W}}_{i',j'}^{n_3}}\right]$$

are used. After some calculations, this allows to express

$$\mathbb{E}\left[(\mathbf{Q}\mathbf{W}\mathbf{W}^*)_{i_1,i_2}^{n_1,n_2}\right] = \sum_{j=1}^{M-L+1}\left(\mathbf{Q}\mathbf{w}_j\mathbf{w}_j^*\right)_{i_1,i_2}^{n_1,n_2}$$

in terms of the entries of $\mathbb{E}(\mathbf{Q})$ and of other terms that tend to 0, and to plug the corresponding expression into (62). This, in turn, leads to (61).

In order to complete the proof of (60), it remains to justify that

$$\tilde{\mathbf{a}}_N^*\left(\mathbf{I}_N \otimes \mathbf{R}_N(z) - t_*(z)\mathbf{I}\right)\tilde{\mathbf{b}}_N \to 0$$

or equivalently that

$$\tilde{\mathbf{a}}_N^*\left(\mathbf{I}_N \otimes \mathbf{R}_N(z) - t_N(z)\mathbf{I}\right)\tilde{\mathbf{b}}_N \to 0 \tag{63}$$

where $t_N(z)$ is the Stieltjes transform of the Marcenko-Pastur distribution of parameters (σ^2, c_N^{-1}), which, of course, verifies $t_N(z) - t_*(z) \to 0$ because $c_N \to c_*$. The reader may check that (63) follows from Corollary 5.1, Theorem 7.1 and Eq. (7.3) in [9].

Sketch of Proof of (56): As above, we denote $(M-L+1)\times(M-L+1)$ matrix $\tilde{\mathbf{Q}}_{N,W}(z))$ by $\tilde{\mathbf{Q}}$ in order to simplify the notations. Using the Poincaré-Nash inequality, it can be proved like above that

$$\mathbf{a}_N^*\left(\tilde{\mathbf{Q}} - \mathbb{E}(\tilde{\mathbf{Q}})\right)\mathbf{b}_N \to 0 \; a.s.$$

and establish that

$$\mathbf{a}_N^*\left(\mathbb{E}\left(\tilde{\mathbf{Q}}(z)\right) - \tilde{t}_*(z)\mathbf{I}\right)\mathbf{b}_N \to 0 \tag{64}$$

for each $z \in \mathbb{C} - \mathbb{R}^+$. The behaviour of matrix $\mathbb{E}(\tilde{\mathbf{Q}})$ is not studied in [9]. However, it can be evaluated using the results of [9]. We first briefly justify that

$$\mathbb{E}(\tilde{\mathbf{Q}}) = \tilde{\mathbf{R}} + \tilde{\mathbf{\Delta}} \tag{65}$$

where $\tilde{\mathbf{R}}$ is a certain $(M-L+1)\times(M-L+1)$ matrix, and where $\tilde{\mathbf{\Delta}}$ verifies $\mathbf{a}_N^*\tilde{\mathbf{\Delta}}\mathbf{b}_N \to 0$. The proof of (65) uses the same ingredients than the proof of (61). We first remark that

$$\mathbf{W}^*\mathbf{Q}\mathbf{W} = \tilde{\mathbf{Q}}\mathbf{W}^*\mathbf{W} = \mathbf{I} + z\tilde{\mathbf{Q}} \tag{66}$$

The above mentioned evaluation of $\mathbb{E}[(\mathbf{Q}\mathbf{w}_k)_{i_1}^{n_1}(\mathbf{w}_j^*)_{i_2}^{n_2}]$ for each k, j, i_1, i_2, n_1, n_2 allows to calculate $\mathbb{E}((\mathbf{W}^*\mathbf{Q}\mathbf{W})_{j,k})$ in terms of the entries of $\mathbb{E}(\mathbf{Q})$ and of $\mathbb{E}(\tilde{\mathbf{Q}})$, and other terms that converge towards 0. Plugging this relation as well as (61) into (66) leads to the expression (65). As previously, $\mathbf{a}_N^*\tilde{\mathbf{\Delta}}\mathbf{b}_N \to 0$ is obtained using the Poincaré-Nash inequality.

The proof of $\mathbf{a}_N^*(\tilde{\mathbf{R}}(z) - \tilde{t}_*(z)\mathbf{I})\mathbf{b}_N \to 0$ is omitted because it needs the introduction of several notations of [9], and does not bring new insight.

Proof of (36): 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 $\mathbb{E}(\mathbf{Q}_N(z)\mathbf{Z}_N) = 0$. In order to complete the proof of (36), it is sufficient to establish that if we denote by κ_N the random variable $\kappa_N = \mathbf{a}_N^*(\mathbf{Q}_N(z)\mathbf{Z}_N)\tilde{\mathbf{b}}_N$, then, for each $p \geq 1$, it holds that

$$\mathbb{E}\left|\kappa_N - \mathbb{E}(\kappa_N)\right|^{2p} = \mathcal{O}\left(\left(\frac{L}{M}\right)^p\right) \tag{67}$$

Choosing p large enough leads to $\kappa_N - \mathbb{E}(\kappa_N) = \kappa_N \to 0 \; a.s.$ as expected. (67) can be proved like above by using the Poincaré-Nash inequality.

We finally justify that for each $\epsilon > 0$, (34)–(36) hold uniformly w.r.t. z on each compact subset of $\mathbb{C} - [0, x_*^+ + \epsilon]$. We just prove that it is the case for (36). 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 [13, 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

$$\sup_{z\in\mathcal{K}}|\kappa_N(z)| \leq \eta \tag{68}$$

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

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

Therefore, it holds that

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

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

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

The conclusion follows from the hypothesis that vectors $\mathbf{a}_N$ and $\tilde{\mathbf{b}}_N$ satisfy $\sup_N(\|\mathbf{a}_N\|, \|\tilde{\mathbf{b}}_N\|) < +\infty$. (68) implies that the sequence of analytic functions $(\kappa_N)_{N\geq 1}$ is a normal family. Therefore, Montel's theorem [13, p. 282]implies the existence of 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 (36) holds for each $z \in \mathbb{C} - \mathbb{R}^+$, 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*, 2nd ed. New York, NY, USA: Springer Ser. Statist., 2010.

[2] F. Benaych-Georges and R. R. Nadakuditi, "The eigenvalues and eigenvectors of finite, low rank perturbations of large random matrices," *Adv. Math.*, vol. 227, no. 1, pp. 494–521, 2011.

[3] F. Benaych-Georges and R. R. Nadakuditi, "The singular values and vectors of low rank perturbations of large rectangular random matrices," *J. Multivar. Anal.*, vol. 111, pp. 120–135, 2012.

[4] R. Couillet, "Robust spiked random matrices and a robust G-MUSIC estimator," *J. Multivar. Anal.*, vol. 140, pp. 139–161, 2015.

[5] 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 M.I.T. Lincoln Lab, Lexington, MA, USA, Rep. 582, 1982.

[6] W. Hachem, P. Loubaton, X. Mestre, J. Najim, and P. Vallet, "A subspace estimator for fixed rank perturbations of large random matrices," *J. Multivar. Anal.*, vol. 114, pp. 427–447, 2012.

[7] 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 Appl.*, vol. 1, no. 2, 2012.

[8] B. A. Johnson, Y. I. Abramovich, and X. Mestre, "MUSIC, G-MUSIC, maximum-likelihood performance breakdown," *IEEE Trans. Signal Process.*, vol. 56, no. 8, pp. 3944–3958, 2008.

[9] P. Loubaton, "On the almost sure location of the singular values of certain Gaussian block-Hankel large random matrices," *J. Theoret. Probabil.*, May 12, 2015 [Online]. Available: http://link.springer.com/article/10.1007/s10959-015-0614-z

[10] V. A. Marcenko and L. A. Pastur, "Distribution of eigenvalues for some sets of random matrices," *Math. USSR-Sbornik*, vol. 1, p. 457, 1967.

[11] X. Mestre, "Improved estimation of eigenvalues and eigenvectors of covariance matrices using their sample estimates," *IEEE Trans. Inf. Theory*, vol. 54, no. 11, pp. 5113–5129, 2008.

[12] S. U. Pillai and B. H. Kwon, "Performance analysis of MUSIC type high resolution estimators for direction finding in correlated and coherent scenes," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 8, pp. 1176–1189, 1989.

[13] W. Rudin", *Real and Complex Analysis*, 3rd ed. New York, NY, USA: McGraw-Hill, 1987.

[14] T. J. Shan, M. Wax, and T. Kailath, "On spatial smoothing for estimation of coherent signals," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 33, no. 8, pp. 806–811, 1985.

[15] P. Stoica and A. Nehorai, "MUSIC, maximum likelihood, Cramer-Rao bound," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 37, no. 5, pp. 720–741, 1989.

[16] A. Thakre, M. Haardt, and K. Giridhar, "Single snapshot spatial smoothing with improved effective array aperture," *IEEE Signal Process. Lett.*, vol. 16, no. 6, pp. 505–508, 2008.

[17] A. Thakre, M. Haardt, and K. Giridhar, "Tensor-based spatial smoothing (TB-SS) using multiple snapshots," *IEEE Trans. Signal Process.*, vol. 58, no. 5, pp. 2715–2728, 2010.

[18] P. Vallet, P. Loubaton, and X. Mestre, "Improved subspace estimation for multivariate observations of high dimension: The deterministic signal case," *IEEE Trans. Inf. Theory*, vol. 58, no. 2, pp. 1043–1068, Feb. 2012.

[19] P. Vallet, P. Loubaton, and X. Mestre, "A statistical comparison between MUSIC and G-MUSIC," in *Proc. ICASSP*, Brisbane, Australia, Apr. 19–24, 2015, pp. 2829–2833.

[20] P. Vallet, X. Mestre, and P. Loubaton, "A CLT for G-MUSIC DOA estimator," in *Proc. Eur. Signal Process. Conf. (EUSIPCO)*, Bucharest, Romania, 2012.

Gia-Thuy Pham (S'14) received the M.Sc. degrees in fundamental mathematics and applied mathematics, in 2011 and 2013, respectively, both from the Université Paris-Sud XI, France.

Since 2013, he has been pursuing the Ph.D. degree at the Université Paris-Est, France. His research interests include statistical signal processing and statistics of large random matrices.

Philippe Loubaton (M'88–SM'06–F'08) received the M.Sc. and Ph.D. degrees from Ecole Nationale Supérieure des Télécommunications, Paris, France, in 1981 and 1988, respectively.

From 1982 to 1986, he was a member of the technical staff of Thomson-CSF/RGS, where he worked in digital communications. From 1986 to 1988, he worked with the Institut National des Télécommunications as an Assistant Professor of Electrical Engineering. In 1988, he joined the Signal Processing Department, Ecole Nationale Supérieure des Télécommunications, Paris. Since 1995, he has been a Professor of Electrical Engineering at Paris-Est Marne la Vallé University, Champs sur Marne, France. After some work on two parameter stationary stochastic processes, his present research interests are in statistical signal processing, and digital communications with a special emphasis on blind equalization, multiuser communication systems and multicarrier modulations, performance analysis of large communication systems via large random matrices methods, and statistics of large random matrices.

Dr. Loubaton served as an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING (1999–2001, 2005–2007), IEEE COMMUNICATIONS LETTERS (2001–2003), and EURASIP *Signal Processing* (2006–2012). He was a member of the IEEE Signal Processing for Communications Technical Committee (1998–2004).

Pascal Vallet (S'08–M'11) received the M.Sc. and Ph.D. degrees from ESIEE Paris, France, in 2008 and the Université Paris-Est in 2011, respectively.

Since 2012, he has been an Associate Professor with Bordeaux INP, France. His research interests include statistical signal processing for digital communication systems.