

# Performance Analysis of an Improved MUSIC DoA Estimator

Pascal Vallet, Xavier Mestre, Philippe Loubaton

► **To cite this version:**

Pascal Vallet, Xavier Mestre, Philippe Loubaton. Performance Analysis of an Improved MUSIC DoA Estimator. IEEE Transactions on Signal Processing, Institute of Electrical and Electronics Engineers, 2015, 63 (23), pp.6407 - 6422. 10.1109/TSP.2015.2465302 . hal-01579060

**HAL Id: hal-01579060**

**<https://hal-upec-upem.archives-ouvertes.fr/hal-01579060>**

Submitted on 30 Aug 2017

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Performance analysis of an improved MUSIC DoA estimator

Pascal Vallet, *Member, IEEE*, Xavier Mestre, *Senior Member, IEEE*, and Philippe Loubaton, *Fellow, IEEE*

**Abstract**—This paper addresses the statistical performance of subspace DoA estimation using a sensor array, in the asymptotic regime where the number of samples and sensors both converge to infinity at the same rate. Improved subspace DoA estimators were derived (termed as G-MUSIC) in previous works, and were shown to be consistent and asymptotically Gaussian distributed in the case where the number of sources and their DoA remain fixed. In this case, which models widely spaced DoA scenarios, it is proved in the present paper that the traditional MUSIC method also provides DoA consistent estimates having the same asymptotic variances as the G-MUSIC estimates. 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 G-MUSIC estimates are still able to consistently separate the sources, while this is no longer the case for the MUSIC ones. The asymptotic variances of G-MUSIC estimates are also evaluated.

**Index Terms**—Subspace DoA estimation, large sensor arrays, random matrix theory

## I. INTRODUCTION

THE problem of estimating the directions of arrival (DoA) of source signals with an array of sensors is fundamental in statistical signal processing, and several methods have been developed and characterized in terms of performance, during the past 40 years. Among the most popular high resolution methods, subspace algorithms such as MUSIC [17] are widely used. It is well known (see e.g. [19]) that subspace methods suffer the so-called “threshold effect”, which involves a severe degradation when either the Signal to Noise Ratio (SNR) and/or the sample size are not large enough. In contrast, the threshold breakdown is less significant for Maximum Likelihood (ML) techniques, and occurs for a much lower SNR and/or sample size. However, due to their reduced complexity since they involve a one-dimensional search over the set of possible DoA, subspace methods are usually preferred over ML which requires a multi-dimensional search.

The study of the statistical performance of MUSIC algorithm has received a lot of attention, see e.g. [18], and

P. Vallet is with Laboratoire de l’Intégration du Matériau au Système (CNRS, Univ. Bordeaux, Bordeaux INP), 351, Cours de la Libération 33405 Talence (France), pascal.vallet@bordeaux-inp.fr

X. Mestre is with Centre Tecnològic de Telecomunicacions de Catalunya (CTTC), Av. Carl Friedrich Gauss 08860 Castelldefels, Barcelona (Spain), xavier.mestre@cttc.cat

P. Loubaton is with Laboratoire d’Informatique Gaspard Monge (CNRS, Université Paris-Est/MLV), 5 Bd. Descartes 77454 Marne-la-Vallée (France), loubaton@univ-mlv.fr

This work was partially supported by the Catalan Government under grant 2014 SGR 1567, the French Projects ANR-12-MONU-0003 DIONISOS and GDR ISIS Jeunes Chercheurs. The material of this paper was partly presented in the conference papers [14] [21].

its behaviour has been mainly characterized in the situation where the number of available samples  $N$  of the observed signal is much larger than the number of sensors  $M$  of the array. However, there may exist some situations where this hypothesis is not realistic, for example when the number of sensors  $M$  is large and the signals have short-time duration or short time stationarity. In this case,  $M$  and  $N$  are of the same order of magnitude, and the standard statistical analysis of MUSIC is irrelevant. This is mainly because the sample correlation matrix of the observations, on which MUSIC mainly relies, does not properly estimate the true covariance matrix. In this context, the standard estimate of the MUSIC angular “pseudo-spectrum” does not appear to be consistent. To model this more stringent scenario, it was proposed in [13] to consider a new asymptotic regime in which both  $M, N$  converges to infinity at the same rate, that is

$$M, N \rightarrow \infty \text{ such that } \frac{M}{N} \rightarrow c > 0.$$

Based on results from random matrix theory, giving a precise description of the behaviour of the eigenvalues and eigenvectors of large random matrices, an improved MUSIC DoA technique, termed as “G-MUSIC”, was derived in [13] in the unconditional model case, that is, by assuming that the source signals are Gaussian and temporally white. This method was based on a novel estimator of the “pseudo-spectrum” function. Other related works concerning the unconditional case include [9] as well as [10] where the source number detection is addressed. Later, [20] addressed the more general conditional model case, i.e. the source signals are modelled as non observable deterministic signals. Using an approach similar to [13], a different estimator of the pseudo-spectrum was proposed. More recently, the work of [23] extends the improved subspace estimation of [20] to the situation where the noise may be correlated in time. We also mention the recent series of works [4] [3] [5] on robust subspace estimation, in the context of impulsive noise.

Experimentally, it can be observed that in certain scenarios, MUSIC and G-MUSIC show quite similar performance, while in other contexts G-MUSIC outperforms MUSIC. In this paper which is focused on the conditional case, we explain this behaviour and provide a complete description of the statistical performance of MUSIC and G-MUSIC. Roughly speaking, we prove that if the DoAs are widely spaced compared to  $\frac{1}{M}$ , MUSIC and G-MUSIC have a similar behaviour, while MUSIC fails when the DoAs are closely spaced. More precisely, we establish the following results.

- When the number of sources  $K$  and the corresponding DoA remain fixed as  $M, N \rightarrow \infty$  (a regime which

models widely spaced sources), we show that, while the pseudo-spectrum estimate of MUSIC is inconsistent, its minimization w.r.t. the DoA provides  $N$ -consistent<sup>1</sup> estimates. Moreover, in the case of asymptotically uncorrelated source signals, the MUSIC DoA estimates share the same asymptotic MSE as G-MUSIC.

- For two sources with an angular spacing of the order of a beamwidth, that is  $\mathcal{O}(M^{-1})$  as  $M, N \rightarrow \infty$ , we show that G-MUSIC remains  $N$ -consistent while MUSIC is not  $N$ -consistent anymore, which means that MUSIC is no longer able to asymptotically separate the DoA.

#### A. Problem formulation and previous works

Let us consider the situation where  $K$  narrow-band and far-field source signals are impinging on a uniform linear array of  $M$  sensors, with  $K < M$ . The received signal at the output of the array is usually modeled as a complex  $M$ -variate time series  $(\mathbf{y}_n)_{n \geq 1}$  given by

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

where

- $\mathbf{A} = [\mathbf{a}(\theta_1), \dots, \mathbf{a}(\theta_K)]$  is the  $M \times K$  matrix of steering vectors  $\mathbf{a}(\theta_1), \dots, \mathbf{a}(\theta_K)$ , with  $\theta_1, \dots, \theta_K$  the source signals DoA, and  $\mathbf{a}(\theta) = \frac{1}{\sqrt{M}}[1, \dots, e^{i(M-1)\theta}]^T$ ;
- $\mathbf{s}_n \in \mathbb{C}^K$  contains the source signals received at time  $n$ , considered as unknown deterministic;
- $(\mathbf{v}_n)_{n \geq 1}$  is a temporally and spatially white circularly symmetric complex Gaussian noise with spatial covariance  $\mathbb{E}[\mathbf{v}_n \mathbf{v}_n^*] = \sigma^2 \mathbf{I}$ .

By assuming that  $N$  observations  $\mathbf{y}_1, \dots, \mathbf{y}_N$  are collected in the  $M \times N$  matrix

$$\mathbf{Y}_N = [\mathbf{y}_1, \dots, \mathbf{y}_N] = \mathbf{A}\mathbf{S}_N + \mathbf{V}_N, \quad (1)$$

with  $\mathbf{S}_N = [\mathbf{s}_1, \dots, \mathbf{s}_N]$  and  $\mathbf{V}_N = [\mathbf{v}_1, \dots, \mathbf{v}_N]$ , the DoA estimation problem thus consists in estimating the  $K$  DoA  $\theta_1, \dots, \theta_K$  from the matrix of samples  $\mathbf{Y}_N$ .

Subspace methods are based on the observation that the source contributions  $\mathbf{A}\mathbf{s}_1, \dots, \mathbf{A}\mathbf{s}_N$  are confined in the so-called signal subspace of dimension  $K$ , defined as  $\text{span}\{\mathbf{a}(\theta_1), \dots, \mathbf{a}(\theta_K)\}$ . By assuming that the signal sample covariance  $N^{-1}\mathbf{S}_N\mathbf{S}_N^*$  is full rank,  $\theta_1, \dots, \theta_K$  are the unique zeros of the pseudo-spectrum

$$\eta(\theta) = \mathbf{a}(\theta)^* \mathbf{\Pi} \mathbf{a}(\theta), \quad (2)$$

where  $\mathbf{\Pi}$  is the orthogonal projection matrix onto the noise subspace, defined as the orthogonal complement of the signal subspace, and which coincides in that case with the kernel of  $N^{-1}\mathbf{A}\mathbf{S}_N\mathbf{S}_N^*\mathbf{A}^*$  of dimension  $M - K$ .

Since  $\mathbf{\Pi}$  is not available in practice, it must be estimated from the observation matrix  $\mathbf{Y}_N$ . This estimation is traditionally performed by using the so-called sample correlation matrix of the observations (SCM)

$$\frac{\mathbf{Y}_N \mathbf{Y}_N^*}{N} = \frac{1}{N} \sum_{n=1}^N \mathbf{y}_n \mathbf{y}_n^*,$$

<sup>1</sup>An estimator  $\hat{\theta}_N$  of a (possibly depending on  $N, M$ ) DoA  $\theta_N$  is defined as  $N$ -consistent if almost surely,  $N(\hat{\theta}_N - \theta_N) \rightarrow 0$  as  $M, N \rightarrow \infty$ .

and  $\mathbf{\Pi}$  is directly estimated by considering its sample estimate  $\hat{\mathbf{\Pi}}_N$ , i.e. the corresponding orthogonal projection matrix onto the eigenspace associated with the  $M - K$  smallest eigenvalues of  $\frac{\mathbf{Y}_N \mathbf{Y}_N^*}{N}$ . The MUSIC method thus consists in estimating the DoA  $\theta_1, \dots, \theta_K$  as the  $K$  most significant minima of the estimated pseudo-spectrum

$$\hat{\eta}_N^{(t)}(\theta) = \mathbf{a}(\theta)^* \hat{\mathbf{\Pi}}_N \mathbf{a}(\theta),$$

where the superscript  $(t)$  refers to “traditional estimate”.

The SCM is known to be an accurate estimator of the true covariance matrix when the number of available samples  $N$  is much larger than the observation dimension  $M$ . Indeed, in the asymptotic regime where  $M$  is constant and  $N$  converges to infinity, under some technical conditions, the law of large numbers ensures that

$$\left\| \frac{\mathbf{Y}_N \mathbf{Y}_N^*}{N} - \left( \mathbf{A} \frac{\mathbf{S}_N \mathbf{S}_N^*}{N} \mathbf{A}^* + \sigma^2 \mathbf{I} \right) \right\| \rightarrow 0, \quad (3)$$

almost surely (a.s.) as  $N \rightarrow \infty$ , where  $\|\cdot\|$  stands for the spectral norm. This implies that

$$\left\| \hat{\mathbf{\Pi}}_N - \mathbf{\Pi} \right\| \xrightarrow[N \rightarrow \infty]{a.s.} 0 \quad (4)$$

i.e. the sample projection matrix  $\hat{\mathbf{\Pi}}_N$  is a consistent estimator of  $\mathbf{\Pi}$ . Moreover, (4) directly implies the uniform consistency of the traditional pseudo-spectrum estimate

$$\sup_{\theta \in [-\pi, \pi]} \left| \hat{\eta}_N^{(t)}(\theta) - \eta(\theta) \right| \xrightarrow[N \rightarrow \infty]{a.s.} 0. \quad (5)$$

The  $K$  MUSIC DoA estimates, defined formally, for  $k = 1, \dots, K$ , by

$$\hat{\theta}_{k,N}^{(t)} = \underset{\theta \in \mathcal{I}_k}{\text{argmin}} \hat{\eta}_N^{(t)}(\theta),$$

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

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

Several accurate approximations of the MSE on the MUSIC DoA estimates have been obtained (see e.g. [18] and the references therein).

In the situation where  $M, N$  are of the same order of magnitude, (3), and therefore (4) as well as (5), are no longer true. To analyze this situation, [13] proposed to consider the non standard asymptotic regime in which

$$M, N \rightarrow \infty \text{ such that } \frac{M}{N} \rightarrow c > 0. \quad (6)$$

In [20], an estimator  $\hat{\eta}_N(\theta)$  of the pseudo-spectrum  $\eta(\theta)$  was derived. Under an extra assumption, called the separation condition, it was proved to be consistent in the new asymptotic regime (6), that is

$$\hat{\eta}_N(\theta) - \eta(\theta) \rightarrow 0,$$

almost surely, when <sup>2</sup>  $M, N \rightarrow \infty$  such that  $\frac{M}{N} \rightarrow c > 0$ . In the case where the number of sources  $K$  remains fixed when

<sup>2</sup>Note that in that case  $\eta(\theta)$  depends on  $M$  (and thus implicitly on  $N$ ). In the next sections, a subscript  $N$  will be added to make clear this dependence.

$M$  and  $N$  increase, the separation condition was shown to hold if the eigenvalues of  $\mathbf{A} \frac{\mathbf{S}_N \mathbf{S}_N^*}{N} \mathbf{A}^*$  are above the threshold  $\sigma^2 \sqrt{c}$  [20, Section III-C]. Note that a similar estimator was previously derived in [13] in the unconditional source signal case. A stronger result of uniform convergence over  $\theta$  was proved in [6], that is

$$\sup_{\theta \in [-\pi, \pi]} |\hat{\eta}_N(\theta) - \eta(\theta)| \rightarrow 0,$$

almost surely. When  $K$  and the DoA  $(\theta_k)_{k=1, \dots, K}$  remain fixed, the G-MUSIC DoA estimates, defined for  $k = 1, \dots, K$  by  $\hat{\theta}_{k,N} = \operatorname{argmin}_{\theta \in \mathcal{I}_k} \hat{\eta}_N(\theta)$ , were also shown to be  $N$ -consistent, that is

$$N \left( \hat{\theta}_{k,N} - \theta_k \right) \rightarrow 0$$

almost surely, when  $M, N \rightarrow \infty$  such that  $\frac{M}{N} \rightarrow c > 0$ . More recently, [7] also proposed a second-order analysis of the G-MUSIC DoA estimates (in the conditional case), in terms of a Central Limit Theorem (CLT) in the latter asymptotic regime.

The work in [7] assumes that the source signals are spatially uncorrelated asymptotically, that is  $N^{-1} \mathbf{S}_N \mathbf{S}_N^*$  converges to a positive diagonal matrix as  $N \rightarrow \infty$ , and both [6] and [7] that the source DoA  $\theta_1, \dots, \theta_K$  are fixed with respect to  $M, N$ . This latter assumption is suitable for practical scenarios in which the source DoA are widely spaced. However, for scenarios in which the source DoA are closely spaced, e.g. with an angular separation of the order  $\mathcal{O}(M^{-1})$ , the analysis of G-MUSIC provided in [6] and [7] are not relevant anymore.

In this paper, we address a theoretical comparison between the performance of MUSIC and G-MUSIC in the two following scenarios.

In a first scenario, in which the number of sources  $K$  and the corresponding DoA  $\theta_1, \dots, \theta_K$  are considered fixed with respect to  $M, N$  (referred to as “widely spaced DoA”) and where it is known that G-MUSIC is  $N$ -consistent, we prove that, while the traditional MUSIC pseudo-spectrum estimate  $\hat{\eta}_N^{(t)}(\theta)$  is inconsistent, the MUSIC algorithm is  $N$ -consistent and that the two methods exhibit the same asymptotic Gaussian distributions. We remark that the analysis provided for this scenario allows spatial correlation between the different source signals.

In a second scenario, we consider  $K = 2$  spatially uncorrelated source signals with DoA  $\theta_1$  and  $\theta_2$  depending on  $M, N$  such that their angular separation  $\theta_1 - \theta_2 = \mathcal{O}(M^{-1})$ , when  $M, N$  converge to infinity at the rate. We show in this context that the G-MUSIC DoA estimates remain  $N$ -consistent while MUSIC loses its  $N$ -consistency. We also provide in this scenario the asymptotic distribution for the G-MUSIC DoA estimates.

To obtain the asymptotic distribution of G-MUSIC under the two previous scenarios, we rely on a Central Limit Theorem (CLT) which extends the results obtained in [7] using a different approach, and which allows situations involving spatial correlations between sources and closely spaced DoA. A CLT for the traditional MUSIC DoA estimates is also given in the first scenario using the same technique. The proofs of these results need the use of large random matrix theory technics, and appear to be quite long and technical. Therefore,

we choose to not include them in the present paper. However, the derivations are available on-line at [22].

## B. Organization and notations

*Organization of the paper:* In section II, we review some basic random matrix theory results, concerning the asymptotic behaviour of the eigenvalues of the SCM in the case where the number of sources  $K$  remains fixed when  $M$  and  $N$  increase. We then make use of these results to introduce the estimator of any bilinear form of the noise subspace projector  $\mathbf{\Pi}$ , derived in [20]. We also give a Central Limit Theorem (CLT) for this estimator, which will be used in the subsequent sections to derive the asymptotic distribution of the G-MUSIC DoA estimates. In section III, we prove that MUSIC and G-MUSIC are both  $N$ -consistent in the scenario where the source DoA are widely spaced. However, in a closely spaced DoA scenario, we prove that MUSIC is not  $N$ -consistent, while G-MUSIC is still  $N$ -consistent. Finally, we provide in section IV an analysis of G-MUSIC and MUSIC DoA estimates in terms of Asymptotic Gaussianity. In particular, it is shown that MUSIC and G-MUSIC exhibit exactly the same asymptotic MSE in the widely spaced DoA scenario and for asymptotically uncorrelated source signals. Some numerical experiments are provided which confirm the accuracy of the predicted performance of both methods.

*Notations:* For a complex matrix  $\mathbf{A}$ , we denote by  $\mathbf{A}^T, \mathbf{A}^*$  its transpose and its conjugate transpose, and by  $\operatorname{tr}(\mathbf{A})$  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. The notation  $\operatorname{span}\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$  will refer to the vector space generated by  $\mathbf{x}_1, \dots, \mathbf{x}_n$ . The real normal distribution with mean  $m$  and variance  $\sigma^2$  is denoted  $\mathcal{N}_{\mathbb{R}}(\alpha, \sigma^2)$  and the multivariate normal distribution in  $\mathbb{R}^k$ , with mean  $\mathbf{m}$  and covariance  $\mathbf{\Gamma}$  is denoted in the same way  $\mathcal{N}_{\mathbb{R}^k}(\mathbf{m}, \mathbf{\Gamma})$ . A complex random variable  $Z = X + iY$  follows the distribution  $\mathcal{N}_{\mathbb{C}}(\alpha + i\beta, \sigma^2)$  if  $X$  and  $Y$  are independent with respective distributions  $\mathcal{N}_{\mathbb{R}}(\alpha, \frac{\sigma^2}{2})$  and  $\mathcal{N}_{\mathbb{R}}(\beta, \frac{\sigma^2}{2})$ . The expectation and variance of a complex random variable  $Z$  will be denoted  $\mathbb{E}[Z]$  and  $\mathbb{V}[Z]$ . For a sequence of random variables  $(X_n)_{n \in \mathbb{N}}$  and a random variable  $X$ , we write

$$X_n \xrightarrow[n \rightarrow \infty]{a.s.} X \text{ and } X_n \xrightarrow[n \rightarrow \infty]{\mathcal{D}} X$$

when  $X_n$  converges respectively with probability one and in distribution to  $X$ . 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. ASYMPTOTIC BEHAVIOUR OF THE SAMPLE EIGENVALUES AND EIGENVECTORS

In this section, we present some basic results from random matrix theory describing the behaviour of the eigenvalues of the SCM  $\frac{\mathbf{Y}_N \mathbf{Y}_N^*}{N}$ , in the asymptotic regime where  $M, N$  converge to infinity such that  $\frac{M}{N} \rightarrow c > 0$ . These results are required to properly introduce the improved subspace estimator of [20]. To that end, we will work with the following more general model, referred to as “Information plus Noise” in the literature.

We consider  $M, N, K \in \mathbb{N}^*$  such that  $K < M$  and  $M = M(N)$ , is a function of  $N$  satisfying <sup>3</sup>

$$c_N = \frac{M}{N} = c + o\left(\frac{1}{\sqrt{N}}\right) \quad (7)$$

as  $N \rightarrow \infty$ . Thus, in the remainder, the notation  $N \rightarrow \infty$  will refer to the double asymptotic regime  $M, N \rightarrow \infty$ ,  $M/N \rightarrow c > 0$ . We also assume that  $K$  is fixed with respect to  $N$  (for the general case where  $K$  may possibly go to infinity with  $N$ , see [20]). We consider the sequence of random matrices  $(\Sigma_N)_{N \geq 1}$  of size  $M \times N$  where <sup>4</sup>

$$\Sigma_N = \mathbf{B}_N + \mathbf{W}_N, \quad (8)$$

with

- $\mathbf{B}_N$  a rank  $K$  deterministic matrix satisfying  $\sup_N \|\mathbf{B}_N\| < \infty$ ,
- $\mathbf{W}_N$  having i.i.d.  $\mathcal{N}_{\mathbb{C}}\left(0, \frac{\sigma^2}{N}\right)$  entries .

We denote by  $\lambda_{1,N} \geq \dots \geq \lambda_{K,N}$  the non zero eigenvalues of  $\mathbf{B}_N \mathbf{B}_N^*$  and by  $\mathbf{u}_{1,N}, \dots, \mathbf{u}_{K,N}$  the respective unit norm eigenvectors.  $(\mathbf{u}_{k,N})_{k=K+1, \dots, M}$  are unit norm mutually orthogonal vectors of the kernel of  $\mathbf{B}_N \mathbf{B}_N^*$ . Equivalently,  $\hat{\lambda}_{1,N} \geq \dots \geq \hat{\lambda}_{M,N}$  are the eigenvalues of the matrix  $\Sigma_N \Sigma_N^*$  and  $\hat{\mathbf{u}}_{1,N}, \dots, \hat{\mathbf{u}}_{M,N}$  the respective unit norm eigenvectors.

#### A. The asymptotic spectral distribution of the SCM

Let  $\hat{\mu}_N$  be the empirical spectral measure of the matrix  $\Sigma_N \Sigma_N^*$ , defined as the random probability measure

$$\hat{\mu}_N = \frac{1}{M} \sum_{k=1}^M \delta_{\hat{\lambda}_{k,N}},$$

with  $\delta_x$  the Dirac measure at point  $x$ . The distribution  $\hat{\mu}_N$  can be alternatively characterized through its Stieltjes transform defined as

$$\hat{m}_N(z) = \int_{\mathbb{R}} \frac{d\hat{\mu}_N(\lambda)}{\lambda - z} = \frac{1}{M} \text{tr} (\Sigma_N \Sigma_N^* - z\mathbf{I})^{-1}$$

where  $(\Sigma_N \Sigma_N^* - z\mathbf{I})^{-1}$  is the resolvent of the matrix  $\Sigma_N \Sigma_N^*$ .

It is well-known from [12] that for all  $z \in \mathbb{C} \setminus \mathbb{R}$ ,

$$\hat{m}_N(z) \xrightarrow[N \rightarrow \infty]{a.s.} m(z), \quad (9)$$

where

$$m(z) = \int_{\mathbb{R}} \frac{d\mu(\lambda)}{\lambda - z}$$

is the Stieltjes of a deterministic probability measure called the Marchenko-Pastur distribution, whose support coincides with the compact interval  $[\sigma^2(1 - \sqrt{c})^2, \sigma^2(1 + \sqrt{c})^2]$ , and which is defined by

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

<sup>3</sup>The condition  $\sqrt{N}(c_N - c) \rightarrow 0$  is purely technical and is in fact only needed for the validity of Theorems 3, 8 and 7 below.

<sup>4</sup>Of course, we retrieve the usual array processing model (1) by setting  $\Sigma_N = N^{-1/2} \mathbf{Y}_N$ ,  $\mathbf{B}_N = N^{-1/2} \mathbf{A} \mathbf{S}_N$  and  $\mathbf{W}_N = N^{-1/2} \mathbf{V}_N$ .

with  $x^- = \sigma^2(1 - \sqrt{c})^2$  and  $x^+ = \sigma^2(1 + \sqrt{c})^2$ .

Moreover,  $m(z)$  satisfies the following fundamental equation

$$m(z) = \frac{1}{-z(1 + \sigma^2 c m(z)) + \sigma^2(1 - c)}. \quad (10)$$

An equivalent statement of (9) is given with the following convergence in distribution

$$\hat{\mu}_N \xrightarrow[N \rightarrow \infty]{\mathcal{D}} \mu$$

which holds almost surely, that is, the empirical eigenvalue distribution of  $\Sigma_N \Sigma_N^*$  has the same asymptotic behaviour as the Marchenko-Pastur distribution. Practically, the eigenvalue histogram of  $\Sigma_N \Sigma_N^*$  matches the density of the Marchenko-Pastur distribution, for  $M, N$  large enough, as shown in Figure 1, where we have chosen  $M = 1000$ ,  $N = 2000$ ,  $\sigma^2 = 1$  and  $K = 2$  with  $\lambda_{1,N} = 5$  and  $\lambda_{2,N} = 10$ .

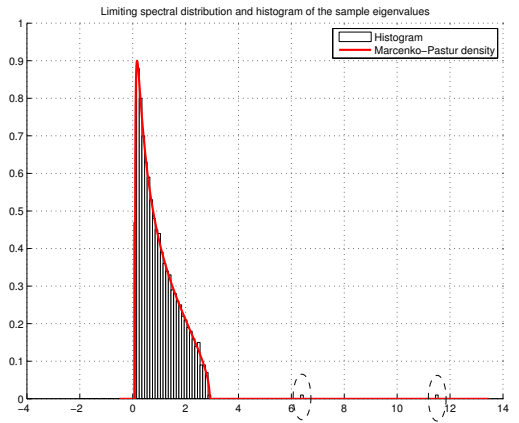


Fig. 1. Marchenko-Pastur distribution and eigenvalue histogram of  $\Sigma_N \Sigma_N^*$

**Remark 1.** The Marchenko-Pastur distribution was originally obtained as the limit distribution of the empirical eigenvalue distribution of the noise part  $\mathbf{W}_N \mathbf{W}_N^*$ . Nevertheless, the assumption that the rank  $K$  of the deterministic perturbation  $\mathbf{B}_N$  is independent of  $N$  implies that the Marchenko-Pastur limit still holds for  $\Sigma_N \Sigma_N^*$ . This fact is well known, and can be easily seen by expressing  $\hat{m}_N(z)$  in terms of the Stieltjes transform of the spectral distribution of  $\mathbf{W}_N \mathbf{W}_N^*$ . Finite rank perturbations of  $\mathbf{W}_N$  are often referred to as “spiked models” in the random matrix literature [1].

#### B. Asymptotic behaviour of the sample eigenvalues

As also noticed in Figure 1, the non zero eigenvalues  $\lambda_{1,N}$  and  $\lambda_{2,N}$  of  $\mathbf{B}_N \mathbf{B}_N^*$  generate two outliers  $\hat{\lambda}_{1,N}$ ,  $\hat{\lambda}_{2,N}$  in the spectrum of  $\Sigma_N \Sigma_N^*$ , in the sense that  $\hat{\lambda}_{1,N}$ ,  $\hat{\lambda}_{2,N}$  are outside the support  $[x^-, x^+]$  of the Marchenko-Pastur distributions, while all the remaining eigenvalues  $\hat{\lambda}_{3,N}, \dots, \hat{\lambda}_{M,N}$  concentrate around  $[x^-, x^+]$ .

In fact, under an additional condition on the non zero eigenvalues  $\lambda_{1,N}, \dots, \lambda_{K,N}$ , it is possible to characterize the behaviour of the  $K$  largest sample eigenvalues  $\hat{\lambda}_{1,N}, \dots, \hat{\lambda}_{K,N}$ . The following assumption, usually referred to as *subspace*

separation condition, ensures that the  $K$  non zero eigenvalues of  $\mathbf{B}_N \mathbf{B}_N^*$  are sufficiently separated from the  $M - K$  zero eigenvalues.

**Assumption 1.** For  $k = 1, \dots, K$ ,  $\lambda_{k,N} \rightarrow \lambda_k$  as  $N \rightarrow \infty$ , where

$$\lambda_1 > \dots > \lambda_K > \sigma^2 \sqrt{c}.$$

We note that forthcoming results remain valid if some  $(\lambda_k)_{k=1, \dots, K}$  coincide. We assume that  $\lambda_k \neq \lambda_l$  for  $k \neq l$  in order to simplify the presentation. Under the previous assumption, an accurate description of the behaviour of the eigenvalues of  $\Sigma_N \Sigma_N^*$  can be obtained.

**Theorem 1.** Under Assumption 1, for  $k = 1, \dots, K$ ,

$$\hat{\lambda}_{k,N} \xrightarrow[N \rightarrow \infty]{a.s.} \phi(\lambda_k) = \frac{(\lambda_k + \sigma^2)(\lambda_k + \sigma^2 c)}{\lambda_k}.$$

with  $\phi(\lambda_k) > x^+$ . Moreover, for all  $\epsilon > 0$ ,

$$\hat{\lambda}_{K+1,N}, \dots, \hat{\lambda}_{M,N} \in (x^- - \epsilon, x^+ + \epsilon),$$

almost surely for  $N$  large enough.

Theorem 1 is a consequence of the general results proved in [1] (see also [11] for a different, but less general, proof). Rephrased in another way, under the separation condition, the  $K$  largest eigenvalues of  $\Sigma_N \Sigma_N^*$  escape from the support of the Marchenko-Pastur distribution while the smallest  $M - K$  eigenvalues are concentrated in a neighborhood of  $[x^-, x^+]$ .

**Remark 2.** Theorem 1 in conjunction with Assumption 1 have a nice interpretation (see e.g. [15] and [2] in the conditional case). Indeed, we notice that the separation condition can be interpreted as a detectability threshold on the SNR condition, if we define the SNR to be the ratio  $\frac{\lambda_k}{\sigma^2}$ . Therefore, Theorem 1 ensures that the  $K$  “signal sample eigenvalues”  $\hat{\lambda}_{1,N}, \dots, \hat{\lambda}_{K,N}$  will be detectable in the sense that they will split from the  $M - K$  “noise sample eigenvalues” as  $N \rightarrow \infty$ , as long as the SNR is above  $\sqrt{c}$ .

### C. Estimation of the signal subspace

In this section, we introduce a consistent estimator of any bilinear form of the noise subspace orthogonal projection matrix, which was derived in [7] (see also [20]). Let us introduce the function

$$w(z) = z(1 + \sigma^2 c m(z))^2 - \sigma^2(1 - c)(1 + \sigma^2 c m(z)).$$

From the fixed point equation (10), straightforward algebra leads to the new equation

$$\phi(w(z)) = z, \quad (11)$$

and one can see easily that the function  $\lambda \mapsto \phi(\lambda)$  is a one to one correspondence from  $(\sigma^2 \sqrt{c}, +\infty)$  onto  $(\sigma^2(1 + \sqrt{c})^2, +\infty)$  with inverse function  $x \mapsto w(x)$  defined on the interval  $(\sigma^2(1 + \sqrt{c})^2, +\infty)$  (see [22]).

The following fundamental result was proved in [1] (see also [7]).

**Theorem 2.** Under Assumption 1, for all deterministic sequences of unit norm vectors  $(\mathbf{d}_{1,N})$ ,  $(\mathbf{d}_{2,N})$ , we have for  $k = 1, \dots, K$

$$\mathbf{d}_{1,N}^* \hat{\mathbf{u}}_{k,N} \hat{\mathbf{u}}_{k,N}^* \mathbf{d}_{2,N} = h(\phi(\lambda_k)) \mathbf{d}_{1,N}^* \mathbf{u}_{k,N} \mathbf{u}_{k,N}^* \mathbf{d}_{2,N} + o(1) \quad a.s.,$$

where

$$h(z) = \frac{w(z)^2 - \sigma^4 c}{w(z)(w(z) + \sigma^2 c)}.$$

Since the function  $\phi$  is the inverse of the function  $w$ , we obtain an explicit expression for  $h(\phi(\lambda_k))$ :

$$h(\phi(\lambda_k)) = \frac{\lambda_k^2 - \sigma^4 c}{\lambda_k(\lambda_k + \sigma^2 c)}.$$

Define the following bilinear form of the noise subspace orthogonal projection matrix:

$$\eta_N = \mathbf{d}_{1,N}^* \mathbf{\Pi}_N \mathbf{d}_{2,N}, \quad (12)$$

as well as its traditional estimate

$$\hat{\eta}_N^{(t)} = \mathbf{d}_{1,N}^* \hat{\mathbf{\Pi}}_N \mathbf{d}_{2,N}. \quad (13)$$

Then Theorem 2 shows in particular that

$$\hat{\eta}_N^{(t)} = \mathbf{d}_{1,N}^* \left( \mathbf{I} - \sum_{k=1}^K h(\phi(\lambda_k)) \mathbf{u}_{k,N} \mathbf{u}_{k,N}^* \right) \mathbf{d}_{2,N} + o(1), \quad (14)$$

a.s., which implies that the traditional subspace estimate is not consistent.

Moreover, Theorem 1 in conjunction with Theorem 2 directly provides a consistent estimator of (12). Indeed, under Assumption 1,

$$\hat{\eta}_N - \eta_N \xrightarrow[N \rightarrow \infty]{a.s.} 0, \quad (15)$$

where

$$\hat{\eta}_N = \mathbf{d}_{1,N}^* \left( \mathbf{I} - \sum_{k=1}^K \frac{1}{h(\hat{\lambda}_{k,N})} \hat{\mathbf{u}}_{k,N} \hat{\mathbf{u}}_{k,N}^* \right) \mathbf{d}_{2,N}. \quad (16)$$

**Remark 3.** It should be noticed that the estimator given in (16) provides in particular a consistent estimator of any  $(i, j)$ -th entry of  $\mathbf{\Pi}_N$ , by choosing  $\mathbf{d}_{1,N} = \mathbf{e}_i$  and  $\mathbf{d}_{2,N} = \mathbf{e}_j$ . However, (15) does not imply that we have a norm-consistent estimator of  $\mathbf{\Pi}_N$ , in the sense that

$$\left\| \mathbf{\Pi}_N - \left( \mathbf{I} - \sum_{k=1}^K \frac{1}{h(\hat{\lambda}_{k,N})} \hat{\mathbf{u}}_{k,N} \hat{\mathbf{u}}_{k,N}^* \right) \right\|$$

does not necessarily converge to 0 as  $N \rightarrow \infty$ .

A result concerning the asymptotic Gaussianity of the estimator  $\hat{\eta}_N$  can be also derived. Let  $\vartheta_{k,\ell}$  be defined under Assumption 1 by

$$\vartheta_{k,\ell} = \frac{\sigma^4 c (\lambda_k \lambda_\ell + (\lambda_k + \lambda_\ell) \sigma^2 + \sigma^4) (\lambda_k \lambda_\ell + \sigma^4 c)}{4 (\lambda_k^2 - \sigma^4 c) (\lambda_\ell^2 - \sigma^4 c) (\lambda_k \lambda_\ell - \sigma^4 c)}$$

$$\begin{aligned} \chi_{k,\ell}^{(t)} = & \lambda_k \lambda_\ell (\lambda_k \lambda_\ell + \sigma^2 (\lambda_k + \lambda_\ell) + \sigma^4) \left( (1+c)(\lambda_k \lambda_\ell + \sigma^4 c) + 2\sigma^2 c (\lambda_k + \lambda_\ell) \right) \\ & - c (\lambda_k \lambda_\ell - \sigma^4 c) (\lambda_k \lambda_\ell + \sigma^2 (\lambda_k + \lambda_\ell) + \sigma^4 c)^2. \end{aligned}$$

for  $1 \leq k, \ell \leq K$ , and by

$$\vartheta_{k,\ell} = \frac{\sigma^2 (\lambda_k + \sigma^2)}{4(\lambda_k^2 - \sigma^4 c)}$$

for  $k \leq K, \ell \geq K+1$ , with  $\vartheta_{k,\ell} = \vartheta_{\ell,k}$ , and set  $\vartheta_{k,\ell} = 0$  for  $k, \ell \geq K+1$ . Define finally

$$\gamma_N = \sum_{k,\ell=1}^M \vartheta_{k,\ell} |\mathbf{u}_{k,N}^* (\mathbf{d}_{1,N} \mathbf{d}_{2,N}^* + \mathbf{d}_{2,N} \mathbf{d}_{1,N}^*) \mathbf{u}_{\ell,N}|^2. \quad (17)$$

We then have the following result.

**Theorem 3.** *Under Assumption 1, if  $\liminf_N \gamma_N > 0$ , then*

$$\sqrt{N} \frac{\text{Re}(\hat{\eta}_N - \eta_N)}{\sqrt{\gamma_N}} \xrightarrow[N \rightarrow \infty]{\mathcal{D}} \mathcal{N}_{\mathbb{R}}(0, 1). \quad (18)$$

The proof of Theorem 3, which requires the use of technical tools from random matrix theory, is not included in the paper and is available in [22].

To conclude this section, we also provide a result on the asymptotic Gaussianity of the classical subspace estimator (13), which will prove to be useful to study the behaviour of MUSIC in the next section. In the same way as (17), we define

$$\vartheta_{k,\ell}^{(t)} = \frac{\sigma^4 c}{4} \frac{\chi_{k,\ell}^{(t)}}{\lambda_k \lambda_\ell (\lambda_k + \sigma^2 c)^2 (\lambda_\ell + \sigma^2 c)^2 (\lambda_k \lambda_\ell - \sigma^4 c)}$$

for  $1 \leq k, \ell \leq K$ , where  $\chi^{(t)}(k, \ell)$  is given at the top of the page (note that  $\chi_{k,\ell}^{(t)} > 0$ ), and by

$$\vartheta_{k,\ell}^{(t)} = \frac{\sigma^2 (\lambda_k + \sigma^2) (\lambda_k^2 - \sigma^4 c)}{4\lambda_k^2 (\lambda_k + \sigma^2 c)^2}.$$

for  $k \leq K, \ell \geq K+1$ , with  $\vartheta_{k,\ell}^{(t)} = \vartheta_{\ell,k}^{(t)}$ , and set  $\vartheta_{k,\ell}^{(t)} = 0$  for  $k, \ell \geq K+1$ . Define finally

$$\gamma_N^{(t)} = \sum_{k,\ell=1}^M \vartheta_{k,\ell}^{(t)} |\mathbf{u}_{k,N}^* (\mathbf{d}_{1,N} \mathbf{d}_{2,N}^* + \mathbf{d}_{2,N} \mathbf{d}_{1,N}^*) \mathbf{u}_{\ell,N}|^2. \quad (19)$$

Then the following result holds.

**Theorem 4.** *Under Assumption 1, if  $\liminf_N \gamma_N^{(t)} > 0$ , then*

$$\sqrt{N} \frac{\text{Re}(\hat{\eta}_N^{(t)} - \eta_N^{(t)})}{\sqrt{\gamma_N^{(t)}}} \xrightarrow[N \rightarrow \infty]{\mathcal{D}} \mathcal{N}_{\mathbb{R}}(0, 1), \quad (20)$$

where

$$\eta_N^{(t)} = \mathbf{d}_{1,N}^* \left( \mathbf{I} - \sum_{k=1}^K \frac{\lambda_{k,N}^2 - \sigma^2 c_N}{\lambda_{k,N} (\lambda_{k,N} + \sigma^2 c_N)} \mathbf{u}_{k,N} \mathbf{u}_{k,N}^* \right) \mathbf{d}_{2,N}.$$

The proof of Theorem 4 is given in [22].

#### D. Connections with others improved subspace estimators

Estimator (16) is valid under the hypothesis that the number of sources  $K$  remains fixed when  $N \rightarrow +\infty$ . We recall that, under the hypothesis that the source signals are deterministic, or equivalently in the conditional case, [20] proposed a consistent estimator of  $\eta_N$ , say  $\hat{\eta}_{N,c}$ , valid whatever  $K$  is, and that it was proved in [20] that

$$\hat{\eta}_{N,c} - \hat{\eta}_N \rightarrow 0 \text{ a.s.} \quad (21)$$

It is even established in [22, Remark 3.] that

$$\hat{\eta}_{N,c} - \hat{\eta}_N = o_{\mathbb{P}} \left( \frac{1}{\sqrt{N}} \right) \quad (22)$$

Therefore, if  $K$  is fixed, the original subspace estimator derived in [20] appears to be equivalent to the estimator (16).

If the  $K$  dimensional source signal  $(\mathbf{s}_n)$  is assumed to be i.i.d. complex Gaussian, or equivalently in the unconditional case, [13] proposed another consistent estimator, denoted  $\hat{\eta}_{N,u}$ , also valid whatever  $K$  is in the unconditional case. When  $K$  is fixed, and when  $(\mathbf{s}_n)$  is deterministic, that is, in the conditional case, it is shown in the Appendix A that

$$\hat{\eta}_{N,u} - \hat{\eta}_N = o_{\mathbb{P}} \left( \frac{1}{\sqrt{N}} \right) \quad (23)$$

Therefore, if  $K$  is fixed, the subspace estimator of [13], in principle valid in the unconditional case, behaves as  $\hat{\eta}_N$ , or equivalently as the estimator  $\hat{\eta}_{N,c}$  derived in [20] in the conditional case. In conclusion, if  $K$  is fixed, in the conditional case, the estimators  $\hat{\eta}_{N,u}$ ,  $\hat{\eta}_{N,c}$  and  $\hat{\eta}_N$  are all equivalent. In section IV-B, simulations are provided to illustrate that  $\hat{\eta}_N$  and  $\hat{\eta}_{N,u}(\theta)$  present the same performance, in the context of DoA estimation.

### III. ANALYSIS OF THE CONSISTENCY OF G-MUSIC AND MUSIC

From now on, we use the results of section II for  $\boldsymbol{\Sigma}_N = N^{-1/2} \mathbf{Y}_N$ ,  $\mathbf{B}_N = N^{-1/2} \mathbf{A}(\boldsymbol{\theta}) \mathbf{S}_N$ ,  $\mathbf{W}_N = N^{-1/2} \mathbf{V}_N$ ,  $\mathbf{d}_{1,N} = \mathbf{d}_{2,N} = \mathbf{a}(\boldsymbol{\theta})$  and assume that Assumption 1 holds. Based on the subspace estimator (16), [7] proposed the improved pseudo-spectrum estimator

$$\hat{\eta}_N(\boldsymbol{\theta}) = 1 - \sum_{k=1}^K \frac{1}{h(\hat{\lambda}_{k,N})} |\mathbf{a}(\boldsymbol{\theta})^* \hat{\mathbf{u}}_{k,N}|^2, \quad (24)$$

**Remark 4.** *The pseudo-spectrum estimator (24) can be viewed as a weighted version of the traditional pseudo-spectrum estimator*

$$\hat{\eta}_N^{(t)}(\boldsymbol{\theta}) = 1 - \sum_{k=1}^K |\mathbf{a}(\boldsymbol{\theta})^* \hat{\mathbf{u}}_{k,N}|^2.$$

Therefore, there is no additional computational cost by using this improved pseudo-spectrum estimator (which gives the G-MUSIC method described below), since it also relies on an eigenvalues/eigenvectors decomposition of the SCM  $\frac{1}{N}\mathbf{Y}_N\mathbf{Y}_N^*$ . Moreover, in the traditional asymptotic regime where  $\frac{M}{N} \rightarrow 0$ , by setting  $c = 0$ , we remark that  $h(z) = 1$  and thus the improved pseudo-spectrum estimator reduces to the traditional one.

From (15), we have directly that  $\hat{\eta}_N(\theta) - \eta_N(\theta) \rightarrow 0$  a.s. as  $N \rightarrow \infty$ , for all  $\theta$ . In Hachem et al. [6], this convergence was also proved to be uniform, that is

$$\sup_{\theta \in [-\pi, \pi]} |\hat{\eta}_N(\theta) - \eta_N(\theta)| \xrightarrow[N \rightarrow \infty]{a.s.} 0, \quad (25)$$

The resulting DoA estimation method, termed as G-MUSIC, consists in estimating  $\theta_1, \dots, \theta_K$  as the  $K$  most significant minima of  $\theta \mapsto \hat{\eta}_N(\theta)$ .

Concerning the traditional pseudo-spectrum estimator  $\hat{\eta}_N^{(t)}(\theta)$ , Theorem 2 directly implies that for all  $\theta$ ,

$$\hat{\eta}_N^{(t)}(\theta) - \eta_N^{(t)}(\theta) \xrightarrow[N \rightarrow \infty]{a.s.} 0,$$

where

$$\eta_N^{(t)}(\theta) = 1 - \sum_{k=1}^K \frac{\lambda_k^2 - \sigma^4 c}{\lambda_k (\lambda_k + \sigma^2 c)} |\mathbf{a}(\theta)^* \mathbf{u}_{k,N}|^2. \quad (26)$$

#### A. $N$ -consistency for widely spaced DoA

In this section, we consider a widely spaced DoA scenario. In practice, such a situation occurs e.g. when the DoA have an angular separation much larger than a beamwidth  $\frac{2\pi}{M}$ . Mathematically speaking, we will therefore consider that the DoA  $\theta_1, \dots, \theta_K$  are fixed with respect to  $N$ . In that case,  $\mathbf{A}^* \mathbf{A} \rightarrow \mathbf{I}$  and the separation condition (Assumption 1) holds if and only if the eigenvalues of  $\frac{\mathbf{S}_N \mathbf{S}_N^*}{N}$  converge to  $\lambda_1 > \dots > \lambda_K > \sigma^2 \sqrt{c}$ . To summarize, we make the following assumption.

**Assumption 2.**  $K, \theta_1, \dots, \theta_K$  are independent of  $N$ , and the eigenvalues of  $N^{-1} \mathbf{S}_N \mathbf{S}_N^*$  converge to

$$\lambda_1 > \dots > \lambda_K > \sigma^2 \sqrt{c}.$$

Note that Assumption 2 allows in particular spatial correlation between sources, since  $\frac{1}{N} \mathbf{S}_N \mathbf{S}_N^*$  may converge to a positive definite matrix, which is not necessarily constrained to be diagonal.

To study the consistency of G-MUSIC and MUSIC, we need to define ‘‘properly’’ the corresponding estimators, to avoid identifiability issues. As it is usually done in the theory of M-estimation, we consider  $\mathcal{I}_1, \dots, \mathcal{I}_K \subset [-\pi, \pi]$   $K$  compact disjoint intervals such that  $\theta_k \in \text{Int}(\mathcal{I}_k)$  (Int denotes the interior of a set), and formally define the G-MUSIC and MUSIC DoA estimators as <sup>5</sup>

$$\hat{\theta}_{k,N} = \underset{\theta \in \mathcal{I}_k}{\text{argmin}} \hat{\eta}_N(\theta) \quad \text{and} \quad \hat{\theta}_{k,N}^{(t)} = \underset{\theta \in \mathcal{I}_k}{\text{argmin}} \hat{\eta}_N^{(t)}(\theta). \quad (27)$$

<sup>5</sup>Note that the G-MUSIC cost function can be negative due to the presence of the weighting factor  $h(\hat{\lambda}_{k,N})^{-1}$  in (24).

We have the following result, whose proof is deferred to Appendix B.

**Theorem 5.** Under Assumption 2, for  $k = 1, \dots, K$ ,

$$\hat{\theta}_{k,N} = \theta_k + o\left(\frac{1}{N}\right) \quad \text{and} \quad \hat{\theta}_{k,N}^{(t)} = \theta_k + o\left(\frac{1}{N}\right),$$

with probability one.

The results of Theorem 5 show that both the G-MUSIC and MUSIC methods have the same first order behaviour, i.e. are  $N$ -consistent, when the angles  $\theta_1, \dots, \theta_K$  are fixed with respect to  $N$ . In section IV, it will be further shown that the MUSIC method also has the same asymptotic MSE as the G-MUSIC method as  $N \rightarrow \infty$ .

#### B. $N$ -consistency for closely spaced DoA

In this section, we study the consistency of G-MUSIC and MUSIC in a closely spaced DoA scenario, where we let the DoA  $\theta_{1,N}, \dots, \theta_{K,N}$  depends on  $N$  and converge to the same value at rate  $\mathcal{O}\left(\frac{1}{M}\right)$ . To simplify the presentation, we only consider  $K = 2$  sources with DoA  $\theta_{1,N}$  and  $\theta_{2,N} = \theta_{1,N} + \frac{\alpha}{N}$ , where  $\alpha > 0$ , and assume asymptotic uncorrelated sources with equal powers, that is  $N^{-1} \mathbf{S}_N \mathbf{S}_N^* \rightarrow \mathbf{I}$ . In this case, it is easily seen that the two non null signal eigenvalues of  $\frac{\mathbf{A} \mathbf{S}_N \mathbf{S}_N^* \mathbf{A}^*}{N}$  converge to

$$\lambda_1(\alpha) = 1 + \left| \text{sinc}\left(\frac{\alpha c}{2}\right) \right| \quad \text{and} \quad \lambda_2(\alpha) = 1 - \left| \text{sinc}\left(\frac{\alpha c}{2}\right) \right|.$$

where  $\text{sinc}(x) = \sin(x)/x$  if  $x \neq 0$  and  $\text{sinc}(0) = 1$ . Therefore, the subspace separation condition (Assumption 1) holds if and only if  $\lambda_2(\alpha) > \sigma^2 \sqrt{c}$ . To summarize, we consider the following assumption.

**Assumption 3.** We assume that  $K = 2$ ,

$$\frac{\mathbf{S}_N \mathbf{S}_N^*}{N} \xrightarrow[N \rightarrow \infty]{} \mathbf{I},$$

and that the DoA  $\theta_{1,N}, \theta_{2,N}$  depend on  $N$  in such a way that

$$\theta_{2,N} = \theta_{1,N} + \frac{\alpha}{N},$$

where  $\alpha > 0$  satisfies

$$\left| \text{sinc}\left(\frac{\alpha c}{2}\right) \right| < 1 - \sigma^2 \sqrt{c}.$$

Since the DoA are not fixed with respect to  $N$ , we define, in the same way as (27), the G-MUSIC and MUSIC DoA estimates as

$$\hat{\theta}_{k,N} = \underset{\theta \in \mathcal{I}_{k,N}}{\text{argmin}} \hat{\eta}_N(\theta) \quad \text{and} \quad \hat{\theta}_{k,N}^{(t)} = \underset{\theta \in \mathcal{I}_{k,N}}{\text{argmin}} \hat{\eta}_N^{(t)}(\theta) \quad (28)$$

where  $\mathcal{I}_{k,N}$  is defined as the compact interval

$$\mathcal{I}_{k,N} = \left[ \theta_{k,N} - \frac{\alpha - \epsilon}{2N}, \theta_{k,N} + \frac{\alpha - \epsilon}{2N} \right],$$

with  $0 < \epsilon < \alpha$ . The  $N$ -consistency results for G-MUSIC and MUSIC in the closely spaced DoA scenario can be summarized as follows.



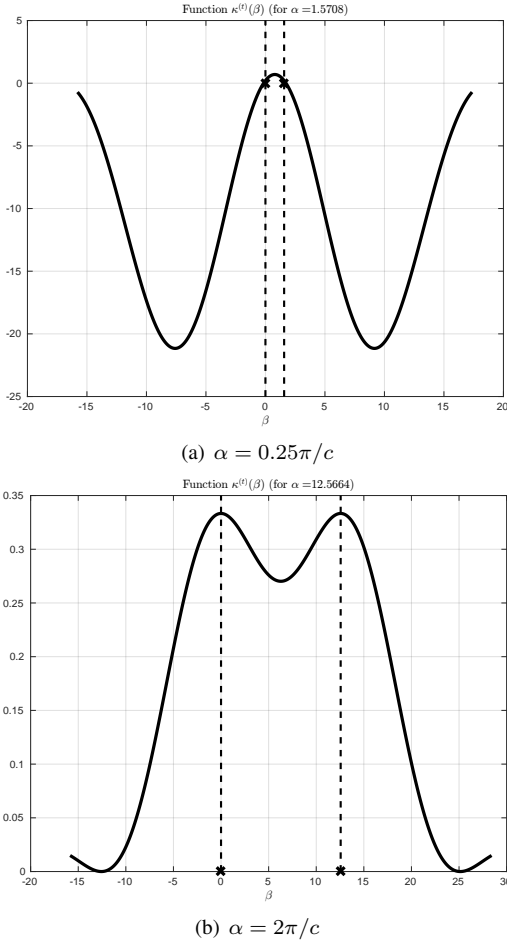


Fig. 2. Function  $\beta \mapsto \kappa^{(t)}(\beta)$  for  $\sigma = 1$ ,  $c = 0.5$  and different values of  $\alpha$  (the dashed lines represent the location of 0 and  $\alpha$ )

**Theorem 6.** Under Assumption 3, for  $k \in \{1, 2\}$ ,

$$\hat{\theta}_{k,N} = \theta_{k,N} + o\left(\frac{1}{N}\right), \quad (29)$$

with probability one. Moreover, if 0 and  $\alpha$  are not local maxima of the function  $\beta \mapsto \kappa^{(t)}(\beta)$  defined by

$$\begin{aligned} \kappa^{(t)}(\beta) = & \frac{(\lambda_1(\alpha)^2 - \sigma^4 c) (\text{sinc}(\beta c/2) + \text{sinc}((\beta - \alpha)c/2))^2}{2\lambda_1(\alpha)^2 (\lambda_1(\alpha) + \sigma^2 c)} \\ & + \frac{(\lambda_2(\alpha)^2 - \sigma^4 c) (\text{sinc}(\beta c/2) - \text{sinc}((\beta - \alpha)c/2))^2}{2\lambda_2(\alpha)^2 (\lambda_2(\alpha) + \sigma^2 c)}, \quad (30) \end{aligned}$$

then  $N \left( \hat{\theta}_{k,N}^{(t)} - \theta_{k,N} \right)$  does not converge to 0.

The proof of Theorem 6 is deferred to Appendix C.

Theorem 6 shows that the G-MUSIC method remains  $N$ -consistent when two sources have DoA with a spacing of the order  $\mathcal{O}(M^{-1})$  while MUSIC may not be able to consistently separate the two DoA if the spacing parameter  $\alpha$  is not a local maximum of the function defined in (30) (numerical examples are given in Figure 2). This confirms the superiority of G-MUSIC over MUSIC in closely spaced DoA situations and low sample size situations.

### C. Remarks on the spatial periodogram

Regarding the previous results on the consistency of the MUSIC estimator for widely spaced and closely spaced scenarios, it is natural to ask how traditional “low resolution” techniques for DoA estimation behave.

Considering the classical spatial periodogram cost function, that is

$$\hat{\eta}_N^{(p)}(\theta) = \mathbf{a}(\theta)^* \frac{\mathbf{Y}_N \mathbf{Y}_N^*}{N} \mathbf{a}(\theta),$$

we can prove, such as in [6, Sec. 3.3], that

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

where  $\eta_N^{(p)}(\theta) = \mathbf{a}(\theta)^* \left( \mathbf{A} \frac{\mathbf{S}_N \mathbf{S}_N^*}{N} \mathbf{A}^* + \sigma^2 \mathbf{I} \right) \mathbf{a}(\theta)$ . Moreover, following the steps of the proof of Theorem 5 for the MUSIC estimates, we end up as well with

$$\hat{\theta}_{k,N}^{(p)} = \theta_k + o\left(\frac{1}{N}\right),$$

with probability one, where  $\hat{\theta}_{k,N}^{(p)} = \arg\max_{\theta \in \mathcal{I}_k} \hat{\eta}_N^{(p)}(\theta)$ . Therefore, the spatial periodogram also provides consistent estimate in the widely-spaced DoA scenario, without any requirements on the sources power (i.e. without need of the separation condition  $\lambda_k > \sigma^2 \sqrt{c}$ ,  $k = 1, \dots, K$ ). This confirms the well-known fact that the use of subspace methods, especially MUSIC, is not necessarily a relevant choice for estimating the DoA of widely spaced sources. Nevertheless, in certain scenarios involving correlated source signals and widely spaced DoA and for which the spatial periodogram may exhibit a non negligible bias at high SNR, the use of subspace methods may still be interesting (see numerical illustrations in Section IV-B).

However, in the scenario of closely spaced DoA, one can also prove, following the steps of Theorem 6, that the spatial periodogram suffers the same drawback as MUSIC, and is not capable of consistently separating two DoA with an angular spacing of the order  $\mathcal{O}\left(\frac{1}{M}\right)$ . Simulations are provided in the next section to illustrate these facts.

## IV. ASYMPTOTIC GAUSSIANTY OF G-MUSIC AND MUSIC

We now apply the results of Theorems 3 and 4 to obtain a Central Limit Theorem for the G-MUSIC and MUSIC DoA estimates. The results for G-MUSIC will be valid for both the widely spaced and closely spaced DoA scenarios introduced in the previous section, while the CLT for MUSIC will be only valid for the widely spaced DoA scenario, since it is not  $N$ -consistent for the other situation.

### A. CLT for G-MUSIC and MUSIC

The following Theorem, whose proof is given in Appendix D, provides the asymptotic Gaussianity of the G-MUSIC DoA estimates, under Assumption 2 or Assumption 3. We denote by  $\mathbf{a}'(\theta)$  and  $\mathbf{a}''(\theta)$  respectively the first and second order derivatives w.r.t.  $\theta$  of the function  $\theta \rightarrow \mathbf{a}(\theta)$ .

**Theorem 7.** Under Assumption 2 or Assumption 3,

$$N^{3/2} \sqrt{\frac{(\mathbf{d}_{1,N}^* \mathbf{\Pi}_N \mathbf{d}_{1,N})^2}{\gamma_N}} (\hat{\theta}_{k,N} - \theta_{k,N}) \xrightarrow[N \rightarrow \infty]{\mathcal{D}} \mathcal{N}_{\mathbb{R}}(0, 1). \quad (31)$$

for  $k = 1, \dots, K$ , where  $\gamma_N$  is defined by (17), with  $\mathbf{d}_{1,N} = N^{-1} \mathbf{a}'(\theta_{k,N})$ ,  $\mathbf{d}_{2,N} = \mathbf{a}(\theta_{k,N})$ .

In particular, by considering the settings of Assumption 2 and adding the following spatial uncorrelation condition

$$\frac{\mathbf{S}_N \mathbf{S}_N^*}{N} \xrightarrow[N \rightarrow \infty]{} \text{diag}(\lambda_1, \dots, \lambda_K), \quad (32)$$

we obtain, using the usual asymptotic orthogonality between  $\mathbf{a}(\theta_k)$  and  $\mathbf{u}_{k',N}$  for  $k \neq k'$  (see e.g. [7, Lem. 8]),

$$\mathbf{d}_{1,N}^* \mathbf{\Pi}_N \mathbf{d}_{1,N} \xrightarrow[N \rightarrow \infty]{} \frac{c^2}{12} \quad \text{and} \quad \gamma_N \xrightarrow[N \rightarrow \infty]{} \frac{c^2 \sigma^2 (\lambda_k + \sigma^2)}{24 (\lambda_k^2 - \sigma^4 c)}.$$

Thus, we retrieve the results of [7] under this particular assumption:

$$N^{3/2} (\hat{\theta}_{k,N} - \theta_k) \xrightarrow[N \rightarrow \infty]{\mathcal{D}} \mathcal{N}_{\mathbb{R}} \left( 0, \frac{6 \sigma^2 (\lambda_k + \sigma^2)}{c^2 (\lambda_k^2 - \sigma^4 c)} \right), \quad (33)$$

Therefore, Theorem 7 extends the results of [7] to more general scenarios of correlated sources and not necessarily widely distributed sources.

Concerning the MUSIC method, we obtain the following result in the widely spaced DoA scenario.

**Theorem 8.** Let  $\eta_N^{(t)(2)}(\theta)$  be the second order derivative of  $\theta \mapsto \eta_N^{(t)}(\theta)$  defined in (26). Under Assumption 2, and if

$$\liminf_{N \rightarrow \infty} |\eta_N^{(t)(2)}(\theta_k)| > 0,$$

it holds that

$$N^{3/2} \sqrt{\frac{\eta_N^{(t)(2)}(\theta_k)^2}{4\gamma_N^{(t)}}} (\hat{\theta}_{k,N}^{(t)} - \theta_k) \xrightarrow[N \rightarrow \infty]{\mathcal{D}} \mathcal{N}_{\mathbb{R}}(0, 1), \quad (34)$$

for  $k = 1, \dots, K$ , where  $\gamma_N^{(t)}$  is defined by (19) by setting  $\mathbf{d}_{1,N} = N^{-1} \mathbf{a}'(\theta_k)$  and  $\mathbf{d}_{2,N} = \mathbf{a}(\theta_k)$ ,

The proof of Theorem 8, which is based on the CLT of Theorem 4, is similar to the one of Theorem 7 and is therefore omitted.

Theorem 8, having been derived under Assumption 2, allows in particular correlation between source signals. Moreover, by assuming asymptotic uncorrelation between sources, i.e. that (32) holds, we obtain

$$\frac{1}{N^2} \eta_N^{(t)(2)}(\theta_k) \xrightarrow[N \rightarrow \infty]{} \frac{c^2 (\lambda_k^2 - \sigma^4 c)}{6 \lambda_k (\lambda_k + \sigma^2 c)}, \quad (35)$$

and

$$\gamma_N^{(t)} \xrightarrow[N \rightarrow \infty]{} \frac{c^2 (\lambda_k + \sigma^2) (\lambda_k^2 - \sigma^4 c)}{24 \lambda_k^2 (\lambda_k + \sigma^2 c)^2},$$

which implies

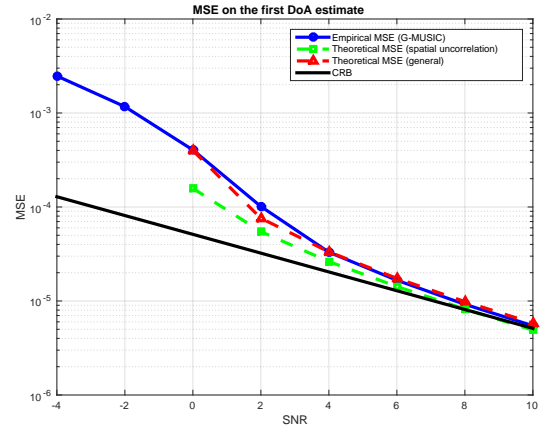
$$N^{3/2} (\hat{\theta}_{k,N}^{(t)} - \theta_k) \xrightarrow[N \rightarrow \infty]{\mathcal{D}} \mathcal{N}_{\mathbb{R}} \left( 0, \frac{6 \sigma^2 (\lambda_k + \sigma^2)}{c^2 (\lambda_k^2 - \sigma^4 c)} \right). \quad (36)$$

The striking fact about Theorem 8 is that, in the widely spaced scenario, the variance of the MUSIC estimates obtained in (36) coincides with the variance of the G-MUSIC estimates (33) previously derived in [7]. This shows that MUSIC and G-MUSIC present exactly the same asymptotic performance for widely spaced DoA and uncorrelated sources, which reinforces the conclusions given in Section III-A.

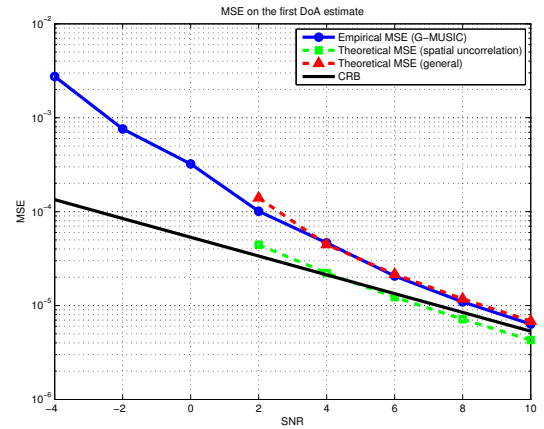
### B. Numerical examples

In this section, we provide numerical simulations illustrating the results given in the previous sections.

To illustrate the similarity between the theoretical MSE (formula of Theorem 7) and its approximation for uncorrelated source signal and wide-space DoA (specific formula of (33)), we plot these two formulas in Figure 3(a) and Figure 3(b), together with the empirical MSE of the G-MUSIC estimate  $\hat{\theta}_{1,N}$  and the Cramer-Rao bound (CRB). The parameters are  $K = 2$ ,  $M = 40$ ,  $N = 80$ ,  $\text{SNR} = -10 \log(\sigma^2)$ . In Figure 3(a), we consider the context of wide-space DoA with uncorrelated source signal, by choosing a signal matrix  $\mathbf{S}_N$  with standard i.i.d  $\mathcal{N}_{\mathbb{C}}(0, 1)$  entries, and setting  $\theta_1 = 0$ ,  $\theta_2 = 5 \times \frac{2\pi}{M}$ . The separation condition  $\lambda_K > \sigma^2 \sqrt{c}$  occurs around  $\text{SNR} = 0$  dB. In this situation, we notice that the two MSE formulas match, as discussed in Section IV-A. In



(a) Uncorrelated source signals



(b) Correlated source signals

Fig. 3. Empirical MSE of  $\hat{\theta}_{1,N}$  for widely spaced DoA versus SNR (dB)

Figure 3(b), we consider the context of wide-space DoA with

significant correlation between source signals, by choosing a matrix  $\mathbf{S}_N = \mathbf{R}^{1/2}\mathbf{X}_N$  with  $\mathbf{R} = [1, 0.4; 0.4, 1]$  and  $\mathbf{X}_N$  having standard i.i.d  $\mathcal{N}_{\mathbb{C}}(0, 1)$  entries. The separation condition occurs around  $\text{SNR} = 2$  dB. We notice that the MSE formula of Theorem 7 is relatively accurate while a discrepancy may occur for the formula (33), since the spatial uncorrelation is not fulfilled in that case.

In Figure 4, we consider the context of wide-space DoA and uncorrelated source signals, and compare the performance of G-MUSIC, MUSIC and DoA estimation with spatial periodogram, in terms of MSE on the first DoA estimate. The empirical MSE of  $\hat{\theta}_{1,N}$  together with its theoretical MSE given in Theorem 7, as well as the empirical MSE of  $\hat{\theta}_{1,N}^{(t)}$  and  $\hat{\theta}_{1,N}^{(p)}$  are plotted. The parameters are  $M = 40$ ,  $N = 80$ , and  $\theta_1 = 0$ ,  $\theta_2 = 5 \times \frac{2\pi}{M}$ . The signal matrix  $\mathbf{S}_N$  has standard i.i.d  $\mathcal{N}_{\mathbb{C}}(0, 1)$  entries, and the separation condition occurs around  $\text{SNR} = 0$  dB. We notice in Figure 4 that the performance

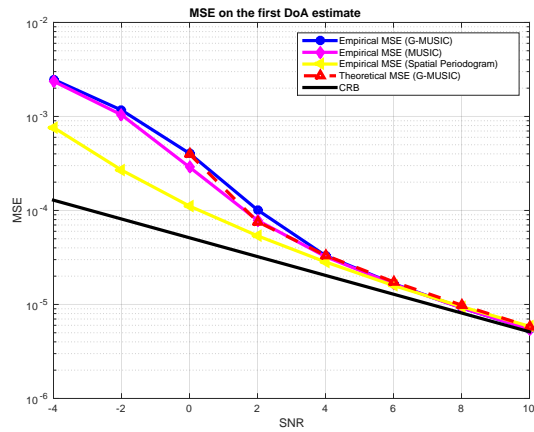


Fig. 4. Empirical MSE of  $\hat{\theta}_{1,N}$ ,  $\hat{\theta}_{1,N}^{(t)}$  and  $\hat{\theta}_{1,N}^{(p)}$  for widely spaced DoA and uncorrelated source signals, versus SNR (dB)

of G-MUSIC, MUSIC as well as the DoA estimate from the spatial periodogram coincide, since the source DoA are widely spaced (five times the beamwidth  $\frac{2\pi}{M}$ ). We also notice that the threshold effect of the spatial periodogram is less significant, since it is not constrained by the subspace separation condition (see Section III-C).

In Figure 5, we consider the same simulation as for Figure 4, except that we add significant correlation between sources, by taking  $\mathbf{S}_N = \mathbf{R}^{1/2}\mathbf{X}_N$  with  $\mathbf{R} = [1, 0.4; 0.4, 1]$  and  $\mathbf{X}_N$  having standard i.i.d  $\mathcal{N}_{\mathbb{C}}(0, 1)$  entries. Again, we notice that both G-MUSIC and MUSIC perform well, since the DoA are widely spaced. Concerning the spatial periodogram method, we notice that a strong bias occurs at high SNR, which corresponds to the well-known effect of source correlation on spatial beamforming techniques (see [16]).

To illustrate the asymptotic Gaussianity of the G-MUSIC and MUSIC estimates predicted in Theorems 7 and 8, we plot in Figure 6 the histograms of  $\hat{\theta}_{2,N}$  and  $\hat{\theta}_{2,N}^{(t)}$  (5000 draws), with the parameters used in Figure 5 (widely spaced DoA and correlated source signals,  $\text{SNR}=6$  dB).

Figure 7 illustrates the closely spaced DoA scenario, and the parameters are the same as in Figure 4, except for the DoA fixed to  $\theta_1 = 0$ ,  $\theta_2 = 0.25 \times \frac{2\pi}{M}$ . The separation

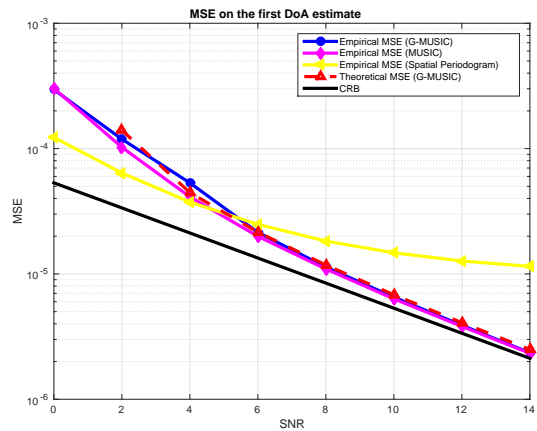
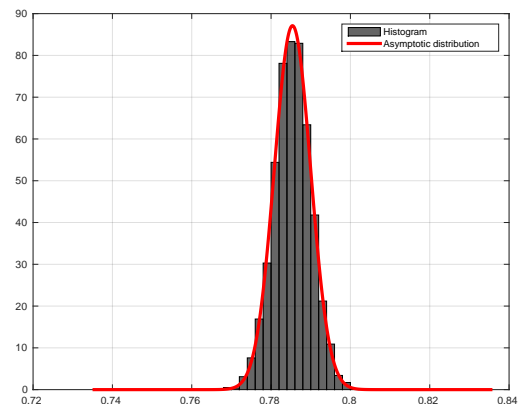
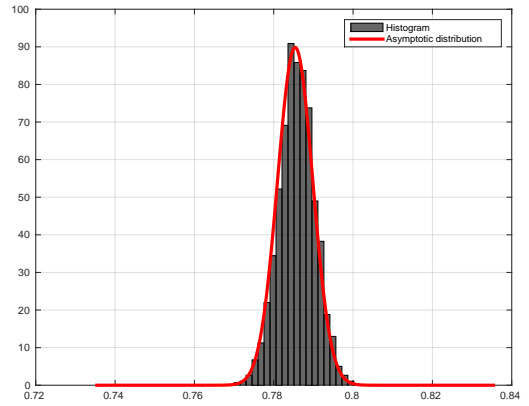


Fig. 5. Empirical MSE of  $\hat{\theta}_{1,N}$ ,  $\hat{\theta}_{1,N}^{(t)}$  and  $\hat{\theta}_{1,N}^{(p)}$  for widely spaced DoA and correlated source signals, versus SNR (dB)



(a) G-MUSIC



(b) MUSIC

Fig. 6. Histograms of  $\hat{\theta}_{2,N}$  and  $\hat{\theta}_{2,N}^{(t)}$  compared to their asymptotic Gaussian distribution

condition is fulfilled for all SNR. One can observe that a strong difference occurs between the performances of the G-MUSIC and MUSIC methods, e.g. a difference of 4 dB between the threshold points of G-MUSIC and MUSIC can be measured, which illustrates the result of Theorem 6. Moreover, we notice the poor performance of the spatial periodogram DoA estimate, which suffers from the well-known resolution loss, since the DoA spacing is lower than a beamwidth.

Similarly, in Figure 8, we keep the same parameters as for

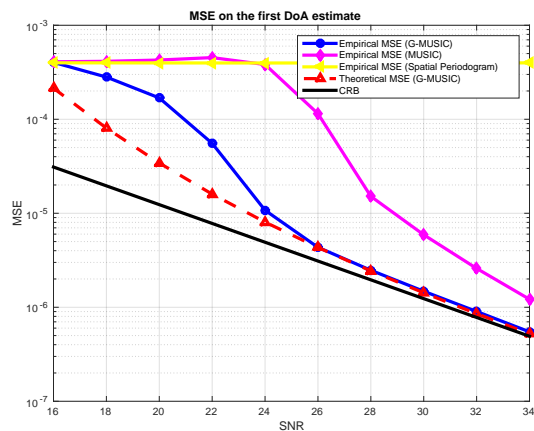


Fig. 7. Empirical MSE of  $\hat{\theta}_{1,N}$ ,  $\hat{\theta}_{1,N}^{(t)}$  and  $\hat{\theta}_{1,N}^{(p)}$  for closely spaced DoA versus SNR (dB)

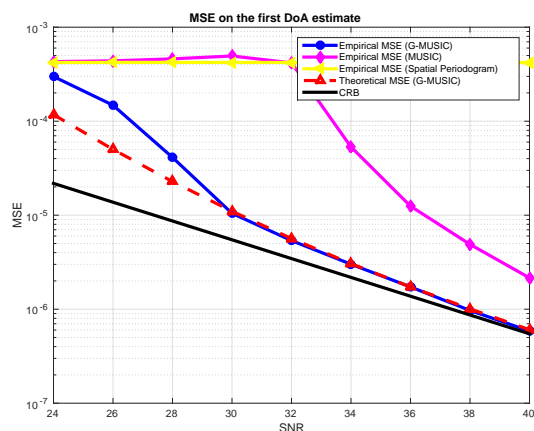


Fig. 8. Empirical MSE of  $\hat{\theta}_{1,N}$ ,  $\hat{\theta}_{1,N}^{(t)}$  and  $\hat{\theta}_{1,N}^{(p)}$  for closely spaced DoA (undersampled), versus SNR (dB)

Figure 7 except that  $M = 40$  and  $N = 20$ . Thus, we consider an “undersampled” scenario in which  $N > M$ . In that case, G-MUSIC still outperforms the MUSIC estimates, with about 6 dB between the threshold points.

In Figure 9, we provide the empirical MSE of MUSIC together with the theoretical MSE given in Theorem 8. The parameters are  $M = 40$ ,  $N = 80$ ,  $\theta_1 = 0$ ,  $\theta_2 = 5 \times \frac{2\pi}{M}$ , and correlated source signals with correlation matrix  $\mathbf{R} = [1, 0.4; 0.4, 1]$  and the separation condition occurs near 2 dB. One can observe the accuracy of the theoretical MSE predicted in Theorem 8.

Finally, we provide in Figures 10 and 11 a comparison between the conditional and unconditional G-MUSIC estimates, using respectively the same scenarios as for Figure 4 and 7. The unconditional G-MUSIC estimator is computed with the formula of [13]. We observe that the two estimators exhibit the same empirical MSE as soon as the separation condition is fulfilled (around SNR = 2 dB for Figure 10 and verified for all SNR in Figure 11), which illustrates the remarks in Section II-D on the connections between both estimators.

## V. CONCLUSION

In this paper, we have addressed a statistical comparison of the performance of the G-MUSIC and MUSIC method for

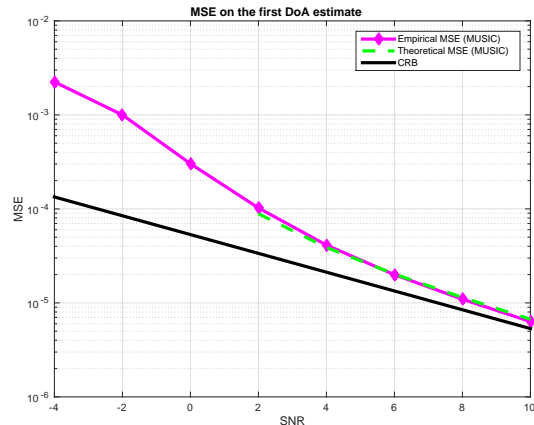


Fig. 9. Empirical MSE of  $\hat{\theta}_{1,N}^{(t)}$  for widely spaced DoA and correlated sources, versus SNR (dB)

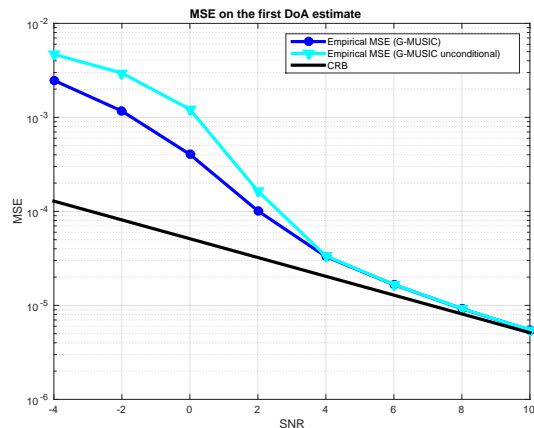


Fig. 10. Empirical MSE of  $\hat{\theta}_{1,N}$  (conditional and unconditional G-MUSIC), for widely spaced DoA and uncorrelated source, versus SNR (dB)

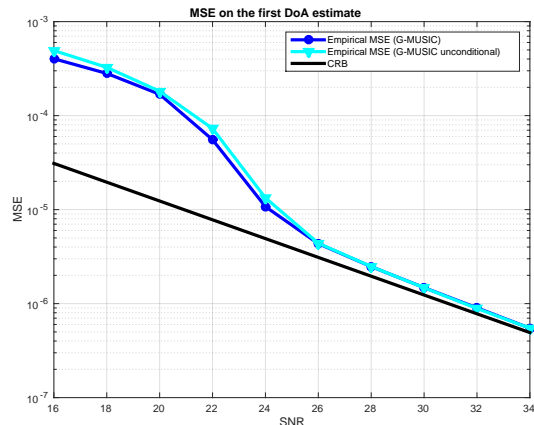


Fig. 11. Empirical MSE of  $\hat{\theta}_{1,N}$  (conditional and unconditional G-MUSIC), for closely spaced DoA and uncorrelated sources, versus SNR (dB)

DoA estimation, in an asymptotic regime where the number of sensors  $M$  and the number of samples  $N$  both converge to infinity at the same rate. Two scenarios were considered. In a first scenario where the source DoA are widely spaced (i.e. fixed with respect to  $M, N$ ), we have proved that both MUSIC and G-MUSIC exhibit the same asymptotic performance in terms of consistency and asymptotic Gaussianity. In a second scenario where the source DoA are closely spaced (i.e. with an angular separation of the order of a beamwidth  $\mathcal{O}(M^{-1})$ ), we have proved that G-MUSIC is still  $N$ -consistent, while MUSIC is no more able to separate the DoA. The asymptotic Gaussianity of G-MUSIC and the identification of its asymptotic MSE provided in this paper hold under general conditions, including correlation between sources, and extend previous existing results which were only valid for asymptotically uncorrelated source signals.

#### APPENDIX A

##### COMPARISON BETWEEN THE UNCONDITIONAL SUBSPACE ESTIMATOR OF [13] WITH THE ESTIMATOR (16).

In this section, we establish (23) when the source signals are deterministic signals satisfying Assumption 1. For this, we first recall that the unconditional estimator  $\hat{\eta}_{N,u}$  proposed in (16) can be written as

$$\hat{\eta}_{N,u} = \frac{1}{2i\pi} \int_{\partial\mathcal{R}} \mathbf{d}_{1,N} (\boldsymbol{\Sigma}_N \boldsymbol{\Sigma}_N^* - z\mathbf{I})^{-1} \mathbf{d}_{2,N} \hat{g}_N(z) dz, \quad (37)$$

where  $\hat{g}_N(z)$  is defined by

$$\hat{g}_N(z) = \frac{(1 - c_N) + cz^2 \hat{m}'_N(z)}{(1 - c_N) - c_N z \hat{m}_N(z)},$$

and where  $\partial\mathcal{R}$  is a contour enclosing the interval  $[\sigma^2(1 - \sqrt{c})^2 - \epsilon, \sigma^2(1 + \sqrt{c})^2 + \epsilon]$ ,  $\epsilon$  being chosen in such a way that  $\sigma^2(1 + \sqrt{c})^2 + \epsilon < \lambda_K$ , and where we recall that  $\hat{m}_N(z) = \frac{1}{M} \text{tr}(\boldsymbol{\Sigma}_N \boldsymbol{\Sigma}_N^* - z\mathbf{I})^{-1} (\hat{m}'_N(z))$  represents the derivative of  $\hat{m}_N(z)$  w.r.t.  $z$ . Using condition (7), it is easily seen that  $\hat{m}_N(z) = m(z) + o_{\mathbb{P}}\left(\frac{1}{\sqrt{N}}\right)$ , that  $\hat{m}'_N(z) = m'(z) + o_{\mathbb{P}}\left(\frac{1}{\sqrt{N}}\right)$ , and using an additional argument such as in [22, Sec. 4.1], one can show

$$\hat{\eta}_{N,c} = \frac{1}{2i\pi} \int_{\partial\mathcal{R}} \mathbf{d}_{1,N}^* (\boldsymbol{\Sigma}_N \boldsymbol{\Sigma}_N^* - z\mathbf{I})^{-1} \mathbf{d}_{2,N} g(z) dz + o_{\mathbb{P}}\left(\frac{1}{\sqrt{N}}\right),$$

where

$$g(z) = \frac{(1 - c) + cz^2 m'(z)}{(1 - c) - czm(z)}.$$

It is established in [22] that

$$\begin{aligned} \hat{\eta}_N = & \frac{1}{2i\pi} \int_{\partial\mathcal{R}} \mathbf{d}_{1,N}^* (\boldsymbol{\Sigma}_N \boldsymbol{\Sigma}_N^* - z\mathbf{I})^{-1} \mathbf{d}_{2,N} \frac{w'(z)}{1 + \sigma^2 cm(z)} dz \\ & + o_{\mathbb{P}}\left(\frac{1}{\sqrt{N}}\right). \end{aligned}$$

The conclusion follows from the identity

$$g(z) = \frac{w'(z)}{1 + \sigma^2 cm(z)},$$

which can be checked easily.

#### APPENDIX B

##### PROOF OF THEOREM 5

The consistency of G-MUSIC is already established in [6], and we prove hereafter the consistency of MUSIC.

From Theorem 2, we have for all  $\theta \in [-\pi, \pi]$ ,

$$\hat{\eta}_N^{(t)}(\theta) = \eta_N^{(t)}(\theta) + o(1), \quad (38)$$

with probability one, where

$$\eta_N^{(t)}(\theta) = 1 - \mathbf{a}(\theta)^* \mathbf{U}_N \mathbf{D} \mathbf{U}_N^* \mathbf{a}(\theta),$$

with  $\mathbf{U}_N = [\mathbf{u}_{1,N}, \dots, \mathbf{u}_{K,N}]$  and  $\mathbf{D} = \text{diag}(d_1, \dots, d_K)$  with

$$d_k = \frac{\lambda_k^2 - \sigma^4 c}{\lambda_k (\lambda_k + \sigma^2 c)}.$$

It is easily seen that  $d_1 > d_2 > \dots > d_K$ . Applying verbatim the steps of [6, Sec. 3.3], (38) can be strengthened to

$$\sup_{\theta \in [-\pi, \pi]} \left| \hat{\eta}_N^{(t)}(\theta) - \eta_N^{(t)}(\theta) \right| \xrightarrow[N \rightarrow \infty]{a.s.} 0. \quad (39)$$

Using the fact that  $\mathbf{U}_N$  and  $\mathbf{A}$  share the same image, we have

$$\mathbf{U}_N = \mathbf{A} (\mathbf{A}^* \mathbf{A})^{-1/2} \mathbf{V}_N,$$

where  $\mathbf{V}_N$  is a  $K \times K$  unitary matrix given by  $\mathbf{V}_N = (\mathbf{A}^* \mathbf{A})^{-1/2} \mathbf{A}^* \mathbf{U}_N$ . Since  $\theta_1, \dots, \theta_K$  are fixed with respect to  $N$ , we also have  $\mathbf{A}^* \mathbf{A} \rightarrow \mathbf{I}_K$  as  $N \rightarrow \infty$ . It is clear that if  $l \neq k$ , then it holds that

$$\sup_{\theta \in \mathcal{I}_k} |\mathbf{a}(\theta)^* \mathbf{a}(\theta_l)| \rightarrow 0$$

From this, we obtain immediately that

$$\sup_{\theta \in \mathcal{I}_k} \|\mathbf{a}(\theta)^* \mathbf{U}_N - \mathbf{a}(\theta)^* \mathbf{a}(\theta_k) \mathbf{e}_k^* \mathbf{V}_N\| \rightarrow 0$$

and that, for all  $k = 1, \dots, K$ ,

$$\sup_{\theta \in \mathcal{I}_k} \left| \eta_N^{(t)}(\theta) - \left(1 - |\mathbf{a}(\theta)^* \mathbf{a}(\theta_k)|^2 \mathbf{e}_k^* \mathbf{V}_N \mathbf{D} \mathbf{V}_N^* \mathbf{e}_k\right) \right| \xrightarrow[N \rightarrow \infty]{} 0. \quad (40)$$

Moreover, it holds that

$$\sup_{\theta \notin \bigcup_k \mathcal{I}_k} \eta_N^{(t)}(\theta) \xrightarrow[N \rightarrow \infty]{} 1. \quad (41)$$

We claim that

$$\hat{\theta}_{k,N}^{(t)} \xrightarrow[N \rightarrow \infty]{a.s.} \theta_k. \quad (42)$$

To verify this, we first remark that (39) and (40) used at point  $\theta = \hat{\theta}_{k,N}^{(t)}$  lead to

$$\hat{\eta}_N^{(t)}(\hat{\theta}_{k,N}^{(t)}) - \left(1 - \left|\mathbf{a}(\hat{\theta}_{k,N}^{(t)})^* \mathbf{a}(\theta_k)\right|^2 \mathbf{e}_k^* \mathbf{V}_N \mathbf{D} \mathbf{V}_N^* \mathbf{e}_k\right) \rightarrow 0$$

almost surely. As function of  $\theta$ ,  $|\mathbf{a}(\theta)^* \mathbf{a}(\theta_k)|^2$  has a unique global maximum at  $\theta_k$  and that  $\mathbf{e}_k^* \mathbf{V}_N \mathbf{D} \mathbf{V}_N^* \mathbf{e}_k$  is lower

bounded by  $d_K > 0$  so we deduce that (42) holds. Otherwise, one could extract from sequence  $(\hat{\theta}_{k,N}^{(t)})_{N \geq 1}$  a subsequence  $\hat{\theta}_{k,\phi(N)}^{(t)}$  converging towards a point  $\theta_* \neq \theta_k$  almost surely. This would imply that

$$\hat{\eta}_{\phi(N)}^{(t)}(\hat{\theta}_{k,\phi(N)}^{(t)}) - \hat{\eta}_{\phi(N)}^{(t)}(\theta_*) \rightarrow 0$$

and that

$$\hat{\eta}_{\phi(N)}^{(t)}(\hat{\theta}_{k,\phi(N)}^{(t)}) - \left(1 - |\mathbf{a}(\theta_*)^* \mathbf{a}(\theta_k)|\right)^2 \mathbf{e}_k^* \mathbf{V}_N \mathbf{D} \mathbf{V}_N^* \mathbf{e}_k \rightarrow 0$$

However, (39) and (40) used at point  $\theta = \theta_k$  imply that

$$\hat{\eta}_N^{(t)}(\theta_k) - (1 - \mathbf{e}_k^* \mathbf{V}_N \mathbf{D} \mathbf{V}_N^* \mathbf{e}_k) \rightarrow 0.$$

Therefore, for  $\epsilon > 0$  small enough, it holds that

$$\hat{\eta}_{\phi(N)}^{(t)}(\theta_k) < \hat{\eta}_{\phi(N)}^{(t)}(\hat{\theta}_{k,\phi(N)}^{(t)}) - \epsilon$$

for each  $N$  large enough, a contradiction.

We now improve (42) by showing that

$$N \left( \hat{\theta}_{k,N}^{(t)} - \theta_k \right) \xrightarrow[N \rightarrow \infty]{a.s.} 0, \quad (43)$$

and for that purpose we follow the approach of [8] (also used in [6, Sec. 4]). By definition, we have

$$\begin{aligned} \eta_N^{(t)}(\hat{\theta}_{k,N}^{(t)}) &\leq \left| \eta_N^{(t)}(\hat{\theta}_{k,N}^{(t)}) - \hat{\eta}_N^{(t)}(\hat{\theta}_{k,N}^{(t)}) \right| + \hat{\eta}_N^{(t)}(\hat{\theta}_{k,N}^{(t)}) \\ &\leq \sup_{\theta \in [-\pi, \pi]} \left| \eta_N^{(t)}(\theta) - \hat{\eta}_N^{(t)}(\theta) \right| + \hat{\eta}_N^{(t)}(\theta_k), \end{aligned}$$

and from (39) and (40) used at point  $\theta = \theta_k$ , we obtain

$$\begin{aligned} \limsup_{N \rightarrow \infty} \eta_N^{(t)}(\hat{\theta}_{k,N}^{(t)}) &\leq \limsup_{N \rightarrow \infty} \hat{\eta}_N^{(t)}(\theta_k) \\ &= 1 - \liminf_{N \rightarrow \infty} \mathbf{e}_k^* \mathbf{V}_N \mathbf{D} \mathbf{V}_N^* \mathbf{e}_k \\ &< 1, \end{aligned} \quad (44)$$

where the last inequality comes from the fact that  $\mathbf{e}_k^* \mathbf{V}_N \mathbf{D} \mathbf{V}_N^* \mathbf{e}_k \geq d_K > 0$ . Assume that the sequence  $N \left( \hat{\theta}_{k,N}^{(t)} - \theta_k \right)$  is not bounded. Then we can extract a subsequence  $\varphi(N) \left( \hat{\theta}_{k,\varphi(N)}^{(t)} - \theta_k \right)$  such that

$$\varphi(N) \left| \hat{\theta}_{k,\varphi(N)}^{(t)} - \theta_k \right| \xrightarrow[N \rightarrow \infty]{} \infty.$$

This implies that  $\mathbf{a}(\hat{\theta}_{k,\varphi(N)}^{(t)})^* \mathbf{a}(\theta_k) \rightarrow 0$  and that, by (40),  $\eta_{\varphi(N)}^{(t)}(\hat{\theta}_{k,\varphi(N)}^{(t)}) \rightarrow 1$ , a contradiction with (44). Since  $N \left( \hat{\theta}_{k,N}^{(t)} - \theta_k \right)$  is bounded, we can extract a subsequence such that

$$\varphi(N) \left| \hat{\theta}_{k,\varphi(N)}^{(t)} - \theta_k \right| \xrightarrow[N \rightarrow \infty]{} \beta,$$

with  $\beta$  assumed to lie in  $[-\pi, \pi]$  without loss of generality. If  $\beta \neq 0$ , then (40) gives

$$\eta_{\varphi(N)}^{(t)}(\hat{\theta}_{k,\varphi(N)}^{(t)}) = 1 - \mathbf{e}_k^* \mathbf{V}_N \mathbf{D} \mathbf{V}_N^* \mathbf{e}_k \operatorname{sinc}(\beta c/2) + o(1)$$

with probability one. Since, in that case,

$$\limsup_{N \rightarrow \infty} \eta_N^{(t)}(\hat{\theta}_{k,N}^{(t)}) > 1 - \liminf_{N \rightarrow \infty} \mathbf{e}_k^* \mathbf{V}_N \mathbf{D} \mathbf{V}_N^* \mathbf{e}_k,$$

this contradicts (44) again.

Therefore all converging subsequences of the bounded sequence  $N \left( \hat{\theta}_{k,N}^{(t)} - \theta_k \right)$  have the same limit, which is 0, and thus the whole sequence converges itself to 0, which finally shows (43).

## APPENDIX C PROOF OF THEOREM 6

Recall from (25) that we have  $\sup_{\theta} |\hat{\eta}_N(\theta) - \eta_N(\theta)| \rightarrow_N 0$  with probability one, with

$$\eta_N(\theta) = \mathbf{a}(\theta)^* \mathbf{\Pi}_N \mathbf{a}(\theta) = 1 - \mathbf{a}(\theta)^* \mathbf{A} (\mathbf{A}^* \mathbf{A})^{-1} \mathbf{A}^* \mathbf{a}(\theta).$$

From Assumption 3, we have  $\sup_{\theta} |\eta_N(\theta) - \tilde{\eta}_N(\theta)| \rightarrow_N 0$  where

$$\tilde{\eta}_N(\theta) = 1 - \frac{1}{1 - \operatorname{sinc}\left(\frac{\alpha c}{2}\right)^2} \mathbf{a}(\theta)^* \mathbf{A} \mathbf{T} \mathbf{A}^* \mathbf{a}(\theta), \quad (45)$$

where

$$\mathbf{T} = \begin{bmatrix} 1 & -e^{i\alpha c/2} \operatorname{sinc}\left(\frac{\alpha c}{2}\right) \\ -e^{-i\alpha c/2} \operatorname{sinc}\left(\frac{\alpha c}{2}\right) & 1 \end{bmatrix}.$$

Note that

$$\limsup_{N \rightarrow \infty} \left| \hat{\eta}_N(\hat{\theta}_{1,N}) \right| \leq \limsup_{N \rightarrow \infty} |\hat{\eta}_N(\theta_{1,N})| = 0. \quad (46)$$

We next rely on the following lemma.

**Lemma 1.** *If  $(\psi_N)$  is a sequence of  $[-\pi, \pi]$  such that  $N |\psi_N - \theta_{1,N}| \rightarrow \infty$ , then*

$$\eta_N(\psi_N) \xrightarrow[N \rightarrow \infty]{} 1.$$

Moreover, for any compact  $\mathcal{K} \subset \mathbb{R}$ ,

$$\sup_{\beta \in \mathcal{K}} \left| \eta_N \left( \theta_{1,N} + \frac{\beta}{N} \right) - (1 - \kappa(\beta)) \right| \xrightarrow[N \rightarrow \infty]{} 0,$$

where

$$\begin{aligned} \kappa(\beta) &= \frac{1}{1 - \operatorname{sinc}(\alpha c/2)^2} \left( \operatorname{sinc}(\beta c/2)^2 + \operatorname{sinc}((\beta - \alpha)c/2)^2 \right. \\ &\quad \left. - 2 \operatorname{sinc}(\alpha c/2) \operatorname{sinc}(\beta c/2) \operatorname{sinc}((\beta - \alpha)c/2) \right) \end{aligned}$$

is such that  $\kappa(\beta) \leq 1$  with equality if and only if  $\beta = 0$  or  $\beta = \alpha$ .

*Proof.* The two convergences can be easily obtained from (45). It thus remains to establish that  $\kappa(\beta) \leq 1$  with equality if and only if  $\beta = 0$  or  $\beta = \alpha$ . Consider the Hilbert space  $\mathcal{L}_{\mathbb{C}}^2([0, 1])$  endowed with the usual scalar product  $\langle z_1, z_2 \rangle = \int_0^1 z_1(t) z_2(t)^* dt$ , and let  $x_1, x_2, y \in \mathcal{L}_{\mathbb{C}}^2([0, 1])$  defined by

$$x_1(t) = 1, \quad x_2(t) = e^{i\alpha ct} \quad \text{and} \quad y(t) = e^{i\beta ct}.$$

Straightforward computations show that  $\kappa(\beta)$  coincides with the squared norm of the orthogonal projection of  $y$  onto  $\operatorname{span}\{x_1, x_2\}$ . Since  $y$  is unit-norm, it is clear that  $\kappa(\beta) \leq 1$ , and the equality holds if and only if  $y \in \operatorname{span}\{x_1, x_2\}$ , which is obviously the case if and only if  $\beta = 0$  or  $\beta = \alpha$ .  $\square$

From Lemma 1, the function  $\kappa$  admits a global maximum, equal to 1, at the unique points 0 and  $\alpha$  and

$$\sup_{\beta \in [-\frac{\alpha}{2}, \frac{3\alpha}{2}]} \left| \hat{\eta}_N \left( \theta_{1,N} + \frac{\beta}{N} \right) - (1 - \kappa(\beta)) \right| \xrightarrow[N \rightarrow \infty]{} 0.$$

Thus,

$$\hat{\eta}_N \left( \hat{\theta}_{1,N} \right) = 1 - \kappa \left( N \left( \hat{\theta}_{1,N} - \theta_{1,N} \right) \right) + o(1),$$

and since  $\hat{\eta}_N(\hat{\theta}_{1,N}) \rightarrow 0$  a.s., we deduce that  $N(\hat{\theta}_{1,N} - \theta_{1,N}) \rightarrow 0$  a.s. We obtain similarly the same results for  $\hat{\theta}_{2,N}$ .

We now consider the consistency of the traditional MUSIC estimates. From (39),

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

where  $\eta_N^{(t)}(\theta) = 1 - \mathbf{a}(\theta)^* \mathbf{U}_N \mathbf{D} \mathbf{U}_N^* \mathbf{a}(\theta)$ . From Assumption 3,  $N^{-1} \mathbf{S}_N \mathbf{S}_N^* \rightarrow \mathbf{I}$ , and using the fact that

$$\mathbf{A}^* \mathbf{A} \xrightarrow[N \rightarrow \infty]{} \begin{bmatrix} 1 & e^{i\alpha c/2} \text{sinc}(\alpha c/2) \\ e^{-i\alpha c/2} \text{sinc}(\alpha c/2) & 1 \end{bmatrix},$$

together with a singular value decomposition of  $\mathbf{A}$ , straightforward computations yield

$$\sup_{\theta} \left| \eta_N^{(t)}(\theta) - \tilde{\eta}_N^{(t)}(\theta) \right| \xrightarrow[N \rightarrow \infty]{} 0,$$

where  $\tilde{\eta}_N^{(t)}(\theta) = 1 - \mathbf{a}(\theta)^* \mathbf{A} \tilde{\mathbf{V}} \tilde{\mathbf{D}} \tilde{\mathbf{V}}^* \mathbf{A}^* \mathbf{a}(\theta)$  and where  $\tilde{\mathbf{V}}$  is  $2 \times 2$  unitary matrix given by

$$\tilde{\mathbf{V}} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & e^{i\alpha c/2} \\ -e^{-i\alpha c/2} & 1 \end{bmatrix},$$

and  $\tilde{\mathbf{D}}$  is a  $2 \times 2$  diagonal matrix defined by

$$\tilde{\mathbf{D}} = \begin{bmatrix} \frac{d_1(\alpha)}{1 - \text{sinc}(\alpha c/2)} & 0 \\ 0 & \frac{d_2(\alpha)}{1 + \text{sinc}(\alpha c/2)} \end{bmatrix}$$

with

$$d_1(\alpha) = \frac{(1 - |\text{sinc}(\alpha c/2)|)^2 - \sigma^4 c}{(1 - |\text{sinc}(\alpha c/2)|)(1 - |\text{sinc}(\alpha c/2)| + \sigma^2 c)}$$

$$d_2(\alpha) = \frac{(1 + |\text{sinc}(\alpha c/2)|)^2 - \sigma^4 c}{(1 + |\text{sinc}(\alpha c/2)|)(1 + |\text{sinc}(\alpha c/2)| + \sigma^2 c)}.$$

We now use the following result, whose proof is similar to the one of Lemma 1.

**Lemma 2.** *If  $(\psi_N)$  is a sequence of  $[-\pi, \pi]$  such that  $N|\psi_N - \theta_{1,N}| \rightarrow \infty$ , then*

$$\eta_N^{(t)}(\psi_N) \xrightarrow[N \rightarrow \infty]{} 1.$$

Moreover, for any compact  $\mathcal{K} \subset \mathbb{R}$ ,

$$\sup_{\beta \in \mathcal{K}} \left| \eta_N^{(t)}\left(\theta_1 + \frac{\beta}{N}\right) - \left(1 - \kappa^{(t)}(\beta)\right) \right| \xrightarrow[N \rightarrow \infty]{} 0,$$

where

$$\kappa^{(t)}(\beta) =$$

$$(\text{sinc}(\beta c/2) - \text{sinc}((\beta - \alpha)c/2))^2 \frac{d_1(\alpha)}{2(1 - |\text{sinc}(\alpha c/2)|)}$$

$$+ (\text{sinc}(\beta c/2) + \text{sinc}((\beta - \alpha)c/2))^2 \frac{d_2(\alpha)}{2(1 + |\text{sinc}(\alpha c/2)|)}.$$

Function  $\kappa^{(t)}$  does not admit in general a local maximum at 0 or  $\alpha$ . In effect, it is easy to find values of  $\alpha$  for which  $\kappa^{(t)}(0)$  and  $\kappa^{(t)}(\alpha)$  are not local maxima of function  $\kappa^{(t)}$ . For example, if  $\alpha = \frac{\pi}{c}$ , we easily check that  $\kappa^{(t)'}(\beta) \neq 0$  for  $\beta = 0$  and  $\beta = \alpha$ .

From Lemma 2, we have with probability one,

$$\hat{\eta}_N^{(t)}(\hat{\theta}_{1,N}^{(t)}) = 1 - \kappa^{(t)}\left(N(\hat{\theta}_{1,N}^{(t)} - \theta_{1,N})\right) + o(1).$$

Assume  $N(\hat{\theta}_{1,N}^{(t)} - \theta_{1,N}) \rightarrow 0$ . Then  $\hat{\eta}_N^{(t)}(\hat{\theta}_{1,N}^{(t)}) \rightarrow 1 - \kappa^{(t)}(0)$ . If 0 and  $\alpha$  are not local maxima of  $\kappa^{(t)}$ , let  $\beta \in [-\frac{\alpha - \epsilon}{2}, \frac{\alpha - \epsilon}{2}]$  such that  $\kappa^{(t)}(0) < \kappa^{(t)}(\beta)$ , and  $(\psi_N)$  a sequence such that  $N(\psi_N - \theta_{1,N}) \rightarrow \beta$ . Then

$$\limsup_{N \rightarrow \infty} \hat{\eta}_N^{(t)}(\hat{\theta}_{1,N}^{(t)}) \leq \limsup_{N \rightarrow \infty} \hat{\eta}_N^{(t)}(\psi_N)$$

$$= 1 - \kappa^{(t)}(\beta)$$

$$< 1 - \kappa^{(t)}(0),$$

which is a contradiction.

## APPENDIX D

### PROOF OF THEOREM 7

To prove Theorem 7, we will use the classical  $\Delta$ -method, as in e.g. Hachem et al. [7].

We consider the settings of Assumption 2 or Assumption 3, and make appear the dependence in  $N$  for the DoA in both scenarios, which we denote by  $\theta_{1,N}, \dots, \theta_{K,N}$ . Let  $k = 1, \dots, K$ . Using Theorem 5 under Assumption 2 (respectively Theorem 6 under Assumption 3), as well as a Taylor expansion around  $\theta_{k,N}$ , we obtain

$$\hat{\eta}'_N(\hat{\theta}_{k,N}) = \hat{\eta}'_N(\theta_{k,N}) + (\hat{\theta}_{k,N} - \theta_{k,N}) \hat{\eta}_N^{(2)}(\theta_{k,N})$$

$$+ \frac{(\hat{\theta}_{k,N} - \theta_{k,N})^2}{2} \hat{\eta}_N^{(3)}(\tilde{\theta}_{k,N}),$$

where  $\tilde{\theta}_{k,N} \in (\min\{\hat{\theta}_{k,N}, \theta_{k,N}\}, \max\{\hat{\theta}_{k,N}, \theta_{k,N}\})$ . Since by definition,  $\hat{\eta}'_N(\hat{\theta}_{k,N}) = 0$ , we obtain

$$\hat{\theta}_{k,N} - \theta_{k,N} = - \frac{\hat{\eta}'_N(\theta_{k,N})}{\hat{\eta}_N^{(2)}(\theta_{k,N}) + \frac{\hat{\theta}_{k,N} - \theta_{k,N}}{2} \hat{\eta}_N^{(3)}(\tilde{\theta}_{k,N})}.$$

As the  $j$ -th derivative  $\mathbf{a}^{(j)}(\theta)$  satisfies  $\sup_{\theta} \|\mathbf{a}^{(j)}(\theta)\| \sim M^j$ , we deduce from [6] that <sup>6</sup>

$$\frac{1}{N^3} \hat{\eta}_N^{(3)}(\tilde{\theta}_{k,N}) = \mathcal{O}(1) \quad (47)$$

with probability one. Theorem 5 implies

$$(\hat{\theta}_{k,N} - \theta_{k,N}) \frac{\hat{\eta}_N^{(3)}(\tilde{\theta}_{k,N})}{N^2} \xrightarrow[N \rightarrow \infty]{a.s.} 0,$$

and we obtain

$$N^{3/2} (\hat{\theta}_{k,N} - \theta_{k,N}) = - \frac{\frac{1}{\sqrt{N}} \hat{\eta}'_N(\theta_{k,N})}{\frac{1}{N^2} \hat{\eta}_N^{(2)}(\theta_{k,N}) + o_{\mathbb{P}}(1)}. \quad (48)$$

By using (15) and the fact that  $\mathbf{\Pi}_N \mathbf{a}(\theta_k) = \mathbf{0}$ , we can write

$$\frac{1}{N^2} \hat{\eta}_N^{(2)}(\theta_{k,N}) = 2 \frac{\mathbf{a}'(\theta_{k,N})^* \mathbf{\Pi}_N \mathbf{a}'(\theta_{k,N})}{N} + o_{\mathbb{P}}(1)$$

<sup>6</sup>The boundedness (47) can be obtained using the techniques developed in the proof of [6, Th. 3.1] (see also equation (1.3) in the introduction part of this reference).

Under Assumption 2, the basic convergences  $\mathbf{A}^* \mathbf{A} \rightarrow \mathbf{I}$ , as  $N \rightarrow \infty$ , as well as

$$\left\| \frac{1}{N} \mathbf{a}'(\theta_{k,N}) \right\|^2 \xrightarrow{N \rightarrow \infty} \frac{c^2}{3}$$

and

$$\left| \frac{1}{N} \mathbf{a}'(\theta_{k,N})^* \mathbf{a}(\theta_{\ell,N}) \right| \xrightarrow{N \rightarrow \infty} \frac{c^2}{4} \delta_{k,\ell}$$

prove that

$$\frac{\mathbf{a}'(\theta_{k,N})^* \mathbf{\Pi}_N \mathbf{a}'(\theta_{k,N})}{N} \xrightarrow{N \rightarrow \infty} \frac{c^2}{12} > 0.$$

Under Assumption 3, we use the arguments in the proof of Lemma 1. Indeed, let  $x_1, x_2, y \in \mathcal{L}_{\mathbb{C}}^2([0, 1])$  defined by

$$x_1(t) = 1, x_2(t) = e^{i\alpha ct} \text{ and } y(t) = ict.$$

Then, we observe that  $\frac{\mathbf{a}'(\theta_{k,N})^* \mathbf{\Pi}_N \mathbf{a}'(\theta_{k,N})}{N}$  converges to the squared norm of the orthogonal projection of  $y$  onto  $\text{span}\{x_1, x_2\}^\perp$ . Thus, we deduce again that

$$\liminf_{N \rightarrow \infty} \frac{\mathbf{a}'(\theta_{k,N})^* \mathbf{\Pi}_N \mathbf{a}'(\theta_{k,N})}{N} > 0.$$

Consider now the quantity  $\gamma_N$  introduced in (17), where we set

$$\mathbf{d}_{1,N} = \frac{\mathbf{a}'(\theta_{k,N})}{N} \text{ and } \mathbf{d}_{2,N} = \mathbf{a}(\theta_{k,N}).$$

Obviously,

$$\begin{aligned} \gamma_N &\geq \frac{\mathbf{a}'(\theta_{k,N})^* \mathbf{\Pi}_N \mathbf{a}'(\theta_{k,N})}{N} \sum_{\ell=1}^K \frac{\sigma^2(\lambda_\ell + \sigma^2)}{4(\lambda_\ell^2 - \sigma^4 c)} |\mathbf{a}(\theta_{k,N})^* \mathbf{u}_{\ell,N}|^2 \\ &\geq D \frac{\mathbf{a}'(\theta_{k,N})^* \mathbf{\Pi}_N \mathbf{a}'(\theta_{k,N})}{N}, \end{aligned}$$

where  $D = \min \left\{ \frac{\sigma^2(\lambda_\ell + \sigma^2)}{4(\lambda_\ell^2 - \sigma^4 c)} : \ell = 1, \dots, K \right\} > 0$ . Therefore, under Assumption 2 or Assumption 3, we obtain

$$\liminf_{N \rightarrow \infty} \gamma_N > 0.$$

Since

$$\begin{aligned} \hat{\eta}'_N(\theta_{k,N}) &= \\ 2N \text{Re} \left( \frac{\mathbf{a}'(\theta_{k,N})^*}{N} \left( \mathbf{I} - \sum_{k=1}^K \frac{1}{h(\hat{\lambda}_{k,N})} \hat{\mathbf{u}}_{k,N} \hat{\mathbf{u}}_{k,N}^* \right) \mathbf{a}(\theta_{k,N}) \right), \end{aligned}$$

Theorem 3 applied with  $\mathbf{d}_{1,N} = \frac{\mathbf{a}'(\theta_{k,N})}{N}$ ,  $\mathbf{d}_{2,N} = \mathbf{a}(\theta_{k,N})$  gives

$$\frac{\hat{\eta}'_N(\theta_{k,N})}{2\sqrt{N}\sqrt{\gamma_N}} \xrightarrow{N \rightarrow \infty} \mathcal{N}_{\mathbb{R}}(0, 1),$$

Gathering this convergence with (48), we eventually obtain (31).

## REFERENCES

- [1] 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.
- [2] P. Bianchi, M. Debbah, M. Mada, and M. Najim. Performance of statistical tests for single source detection using random matrix theory. *IEEE Transactions on Information Theory*, 57(4):2400–2419, 2011.
- [3] Romain Couillet. Robust spiked random matrices and a robust G-MUSIC estimator. *To appear in Journal of Multivariate Analysis*, 2015.
- [4] Romain Couillet and Abba Kammoun. Robust G-MUSIC. In *Signal Processing Conference (EUSIPCO), 2014 Proceedings of the 22nd European*, pages 2155–2159. IEEE, 2014.
- [5] Couillet, Romain and Pascal, Frédéric and Silverstein, Jack W. The random matrix regime of Maronna’s M-estimator with elliptically distributed samples. *Journal of Multivariate Analysis*, 139:56–78, 2015.
- [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] 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.
- [8] E.J. Hannan. The estimation of frequency. *Journal of Applied probability*, 10(3):510–519, 1973.
- [9] 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.
- [10] S. Kritchman and B. Nadler. Non-parametric detection of the number of signals: hypothesis testing and random matrix theory. *IEEE Transactions on Signal Processing*, 57(10):3930–3941, 2009.
- [11] Philippe Loubaton and Pascal Vallet. Almost sure localization of the eigenvalues in a gaussian information plus noise model. application to the spiked models. *Electron. J. Probab.*, 16:1934–1959, 2011.
- [12] V.A. Marchenko and L.A. Pastur. Distribution of eigenvalues for some sets of random matrices. *Mathematics of the USSR-Sbornik*, 1:457, 1967.
- [13] X. Mestre and M.Á. Lagunas. Modified subspace algorithms for doa estimation with large arrays. *IEEE Transactions on Signal Processing*, 56(2):598–614, 2008.
- [14] X. Mestre, P. Vallet, P. Loubaton, and W. Hachem. Asymptotic analysis of a consistent subspace estimator for observations of increasing dimension. In *IEEE Statistical Signal Processing Workshop (SSP)*, pages 677–680. IEEE, 2011.
- [15] R.R. Nadakuditi and A. Edelman. Sample eigenvalue based detection of high-dimensional signals in white noise using relatively few samples. *IEEE Transactions on Signal Processing*, 56(7):2625–2637, 2008.
- [16] V Umapathi Reddy, Arogyaswami Paulraj, and Thomas Kailath. Performance analysis of the optimum beamformer in the presence of correlated sources and its behavior under spatial smoothing. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 35(7):927–936, 1987.
- [17] R. Schmidt. Multiple emitter location and signal parameter estimation. *IEEE Transactions on Antennas and Propagation*, 34(3):276–280, 1986.
- [18] P. Stoica and A. Nehorai. Music, maximum likelihood, and cramer-rao bound. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 37(5):720–741, 1989.
- [19] John K Thomas, Louis L Scharf, and Donald W Tufts. The probability of a subspace swap in the svd. *IEEE Transactions on Signal Processing*, 43(3):730–736, 1995.
- [20] 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.
- [21] P. Vallet, X. Mestre, and P. Loubaton. A CLT for the G-MUSIC DoA estimator. In *EUSIPCO 2012*, pages 2298–2302, 2012.
- [22] P. Vallet, X. Mestre, and P. Loubaton. A CLT for an improved subspace estimator with observations of increasing dimensions. 2015. arXiv:1502.02501.
- [23] Julia Vinogradova, Romain Couillet, and Walid Hachem. Statistical inference in large antenna arrays under unknown noise pattern. *IEEE Transactions on Signal Processing*, 61(22):5633–5645, 2013.