On the Performance of MUSIC with Toeplitz Rectification in the Context of Large Arrays

Pascal Vallet, *Member, IEEE*, and Philippe Loubaton, *Fellow, IEEE*

Abstract—When using subspace methods for DoA estimation such as MUSIC, it is well known that a performance loss occurs when the number of available samples N is not large compared to the number of sensors M. This degradation is mainly due to the use of the Sample Correlation Matrix (SCM), which is a poor estimator of the true correlation matrix of the observations in this situation. When the latter exhibits a Toeplitz structure, a standard trick consists in correcting the structure of the SCM by averaging its entries along the subdiagonals. This procedure, known as Toeplitz rectification, is widely known to improve the estimation of the true correlation matrix, hence the performance of the corresponding subspace methods. In this paper, we propose a statistical analysis of the MUSIC method using Toeplitz rectified SCM (refered to as R-MUSIC), in the context where M, N are of the same order of magnitude. More precisely, considering the asymptotic regime in which M, N converge to infinity at the same rate, we prove the consistency and asymptotic normality of the R-MUSIC DoA estimates. Numerical simulations show the accurate prediction provided by the proposed theoretical analysis.

Index Terms—Subspace DoA estimation, Large sensors array, Toeplitz rectification

I. INTRODUCTION

CONSIDER K narrowband and far-field source signals impinging on an array of M sensors with Direction of Arrival (DoA) $\theta_1, \ldots, \theta_K$. The received signal is usually modeled as the multivariate time series $(\mathbf{y}_n)_{n \in \mathbb{Z}}$ with

$$\mathbf{y}_n = \sum_{k=1}^{K} \mathbf{a}(\theta_k) s_{k,n} + \mathbf{v}_n, \tag{1}$$

where $\mathbf{a}(\theta_1), \ldots, \mathbf{a}(\theta_K)$ are the M-dimensional steering vectors associated with the DoA and the array geometry, $s_{k,n}$ represents the k-th transmit source signal and $\mathbf{v}_n$ is an additive complex circular Gaussian noise, assumed spatially and temporally white.

DoA estimation, that is the estimation of $\theta_1, \ldots, \theta_K$ from N samples $\mathbf{y}_1, \ldots, \mathbf{y}_N$, is a classical problem in array processing, and several algorithms such as maximum likelihood, subspace methods or spatial filtering based estimators have been proposed and extensively studied during the last four decades (see e.g. [2] for a review). One of the most prominent criterion for assessing the performance of such algorithms is the capacity to resolve closely spaced DoA, which is in general improved when the number M of sensors in the array increases. The recent technological advances in the development of large dimensional arrays, e.g. in massive multi-antenna systems for wireless communications [3], have renewed the interest in studying high resolution methods.

In general, DoA estimation methods involve intricate statistics of the received signal, and in particular most of them can be written as functions of the Sample Correlation Matrix (SCM) of the received signal $\hat{\mathbf{R}}$, defined as

$$\hat{\mathbf{R}} = \frac{1}{N} \sum_{n=1}^{N} \mathbf{y}_n \mathbf{y}_n^*.$$

Due to the inherent complexity of the underlying estimators, the statistical analysis of these methods, in terms of consistency, Mean Square Error (MSE), resolution probability etc., is usually performed in asymptotic regimes, the most standard one being the *large sample size regime* where the number N of available samples for estimation goes to infinity, while all the other parameters $(M, K, \theta_1, \ldots, \theta_K$, etc.) remain fixed. In this regime, standard statistical results on the asymptotic behaviour of the SCM can be exploited to predict the performance of DoA estimators (see e.g. [4] in the context of subspace methods). In practice, the predictions are quite reliable for a finite N, as long as $M \ll N$ and the DoA are sufficiently spaced.

Nevertheless, in situations where $M \approx N$, the large sample size predictions are in general quite poor, and one needs to adapt the asymptotic regime to model this non standard context, which occurs when the number of sensors is large, or when the sample size is small comparatively (e.g. for short time stationary signals). In the seminal paper [5], a non standard asymptotic regime was proposed, in which M, N goes to infinity at the same rate, that is $\frac{M}{N} \to c > 0$. Using tools from Random Matrix Theory predicting the asymptotic behaviour of the largest eigenvalues and associated eigenvectors of the SCM, asymptotics of various subspace methods cost functions (MUSIC, SS-MUSIC) were derived, in the *unconditional model* (i.e. $(\mathbf{s}_n)$ modeled as zero mean Gaussian and temporally i.i.d.), and an improved MUSIC estimator (termed as G-MUSIC) was proposed. Using similar tools, the *conditional model* (i.e. $(\mathbf{s}_n)$ modeled as deterministic and unknown) was explored in [6]–[9], and a G-MUSIC estimator was also derived in this context, and its consistency, MSE and asymptotic normality was studied in the double asymptotic regime mentioned above. The special case of closely spaced DoA (i.e. with a spacing of the order $\mathcal{O}(\frac{1}{M})$) was also studied;

in this case, G-MUSIC is still able to separate asymptotically the sources while it is no longer the case for MUSIC (cf [9] for the precise meaning of this result). Other works in the large M, N regime include the study of MUSIC with time-correlated noise [10], Beamspace-MUSIC [11], Robust MUSIC/G-MUSIC [12] or MUSIC with Spatial Smoothing [13].

When dealing with such high dimensional observations, the usual key point is to develop estimators of certain functionals of the true correlation matrix of the received signal (defined as $\mathbf{R} = \mathbb{E}[\hat{\mathbf{R}}]$ for both the conditional and unconditional model), which are consistent in the doubly asymptotic regime $M, N \to \infty$ at the same rate. In particular, the consistent estimation of the correlation matrix itself in the spectral norm sense is a challenging problem in the statistics community (see e.g. [14]–[17]) and requires restrictive assumptions such as the vanishing of the off-diagonal entries of $\mathbf{R}$ at a sufficiently fast rate, as $M \to \infty$.

In the special case of a Toeplitz correlation matrix $\mathbf{R} = (r_{i-j})$, $i, j \in \{1, \ldots, M\}$, which occurs in our context e.g. for the unconditional model with spatially uncorrelated sources and a Uniform Linear Array (ULA), it is usual to force the SCM $\hat{\mathbf{R}}$ to have a Toeplitz structure by averaging its coefficients along the diagonals, which leads to a new correlation matrix estimator. This procedure, known as *Toeplitz rectification*, was first proposed in the signal processing community by [18] and later more specifically studied by [19] in the context of array processing for a large class of array geometries.

For any M, N, [16] obtained precise upper bounds on the quadratic risk (in the spectral norm sense) of the Toeplitz rectification scheme associated with a tapering/banding, for estimating correlation matrices belonging to certain classes of Toeplitz matrices. In particular, consistency of these estimates holds when M or $N \to \infty$, as long as the entries of $\mathbf{R}$ outside the main diagonal satisfy a certain rate of decay. Later, it was shown in [17] that the consistency of the rectified SCM holds without banding nor tapering, under the specific regime $\frac{M}{N} \to c > 0$ as $M, N \to \infty$. Nevertheless, in the context of model (1), it turns out that the off-diagonal entries of $\mathbf{R}$ all have the same order of magnitude and do not satisfy the rate of decay required in [16], [17], hence these results cannot be used in the array processing specific context, for which a different analysis is needed.

In this paper, we develop a statistical analysis of the Toeplitz rectification scheme in the high dimensional asymptotic regime $M, N \to \infty$ such that $\frac{M}{N} \to c > 0$, assuming a ULA geometry for the array (for which the works of [16], [17] cannot be used). The analysis is provided under both the conditional and unconditional models, which are summarized in Section II along with the assumptions used throughout the paper. Surprisingly, we prove in Section III that the Rectified SCM, without any resort to tapering nor banding, provides a consistent estimator of the true correlation matrix $\mathbf{R}$. This result is then exploited in Section IV to study the statistical performance, in terms of consistency and asymptotic normality, of the MUSIC method for DoA estimation, when using the Rectified SCM (the method will be referred to as

R-MUSIC). In particular, an accurate approximation of the MSE is derived, and a stagnation phenomenon is shown to occur for high Signal to Noise Ratio (SNR). Moreover, it is also proved that R-MUSIC has the ability to separate closely spaced DoA, as for the G-MUSIC method. Finally, Section V presents some numerical simulations illustrating the previous results, and Section VI concludes the paper.

Notations: For a complex matrix $\mathbf{A}$, we denote by $\mathbf{A}^T, \mathbf{A}^*$ its transpose and its conjugate transpose, and by $\text{Tr}(\mathbf{A})$, $\|\mathbf{A}\|_2$ and $\|\mathbf{A}\|_F$ its trace, spectral norm and Frobenius norm respectively. The Hadamard product (entrywise product of matrices) is denoted $\odot$. 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 real normal distribution with mean m and variance σ^2 is denoted $\mathcal{N}_{\mathbb{R}}(\alpha, \sigma^2)$. A random variable Z follows the complex Gaussian distribution $\mathcal{N}_{\mathbb{C}}(\alpha, \sigma^2)$ if $Z = X + iY$ with X, Y independent such that $X \sim \mathcal{N}_{\mathbb{R}}\left(\text{Re}(\alpha), \frac{\sigma^2}{2}\right)$ and $Y \sim \mathcal{N}_{\mathbb{R}}\left(\text{Im}(\alpha), \frac{\sigma^2}{2}\right)$. The expectation and variance of a complex random variable Z will be denoted $\mathbb{E}[Z]$ and $\mathbb{V}[Z]$. A random vector $\mathbf{X}$ follows the distribution $\mathcal{N}_{\mathbb{C}^k}(\boldsymbol{\alpha}, \mathbf{R})$ if $\mathbf{b}^*\mathbf{X} \sim \mathcal{N}_{\mathbb{C}}\left(\mathbf{b}^*\boldsymbol{\alpha}, \|\mathbf{R}^{1/2}\mathbf{b}\|_2^2\right)$ for all deterministic (column) vector $\mathbf{b}, \boldsymbol{\alpha} \in \mathbb{C}^k$ and $k \times k$ positive definite matrix $\mathbf{R}$. For a sequence of random variables $(X_n)_{n \in \mathbb{N}}$ and a random variable X, we write

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

when X_n converges respectively almost surely (a.s.) 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. MODELS AND ASSUMPTIONS

From now on, we consider a ULA of M sensors receiving K narrowband and far-field source signals with DoA $\theta_1, \ldots, \theta_K$, and collecting N samples in the $M \times N$ matrix $\mathbf{Y}$ modeled as

$$\mathbf{Y} = \mathbf{A}\mathbf{S} + \mathbf{V},$$

where

- $\mathbf{A} = [\mathbf{a}(\theta_1), \ldots, \mathbf{a}(\theta_K)]$ is the $M \times K$ steering vectors matrix, with [1]

$$\theta \mapsto \mathbf{a}(\theta) = \frac{1}{\sqrt{M}}\left[1, e^{i\theta}, \ldots, e^{i(M-1)\theta}\right]^T,$$

and $\theta_1, \ldots, \theta_K \in (-\pi, \pi)$. [2]
- $\mathbf{S}$ is the $K \times N$ matrix of source signals,
- $\mathbf{V}$ is an additive noise matrix with i.i.d. entries $\mathcal{N}_{\mathbb{C}}(0, \sigma^2)$.

---

[1] In this paper, we assume unit-norm steering vectors by introducing the normalization factor $\frac{1}{\sqrt{M}}$, in order to keep the matrix $\mathbf{A}$ (hence the SNR) bounded as $M \to \infty$.

[2] In this paper, θ_k refers to the Direction of Arrival in the Fourier domain, and is related to the "physical" elevation angle $\phi_k \in (0, \pi)$ of the source wave by the parametrization

$$\theta_k = \frac{2\pi d \cos(\phi_k)}{\lambda},$$

where d, λ are respectively the sensors interspacing and wavelength.

For our analysis, we work with both the conditional and unconditional models, as well as uncorrelated sources, which is summarized by the following two assumptions on the signal matrix $\mathbf{S}$.

Assumption 1 (Conditional model)**.** *Signal matrix $\mathbf{S}$ is deterministic and*

$$\frac{1}{N}\mathbf{S}\mathbf{S}^* = \mathbf{\Gamma} + o\left(\frac{1}{N^{1/4}}\right)$$

as $N \to \infty$, where $\mathbf{\Gamma}$ is a positive diagonal matrix.

We underline the fact that the rate of convergence of the SCM $\frac{\mathbf{S}\mathbf{S}^*}{N}$ to the true correlation matrix $\mathbf{\Gamma}$ imposed in Assumption 1 is merely technical and not restrictive in practice, since it is satisfied for a wide range of stationary linear processes (see e.g. [20]).

Assumption 2 (Unconditional model)**.** *Signal matrix $\mathbf{S}$ is modeled as $\mathbf{S} = \mathbf{\Gamma}^{1/2}\mathbf{X}$, with $\mathbf{X}$ a $K \times N$ random matrix having i.i.d $\mathcal{N}_{\mathbb{C}}(0,1)$ entries, and where $\mathbf{\Gamma}$ is a positive diagonal matrix.*

Note that under Assumption 1, $\mathbf{Y}$ is a non centered Gaussian matrix, while under Assumption 2, $\mathbf{Y} \overset{\mathcal{D}}{=} \left(\mathbf{A}\mathbf{\Gamma}\mathbf{A}^* + \sigma^2\mathbf{I}\right)^{1/2}\mathbf{W}$, with $\mathbf{W}$ having $\mathcal{N}_{\mathbb{C}}(0,1)$ entries.

Moreover, for both assumptions, we denote by $\mathbf{R}$ the covariance matrix of the received signal defined as

$$\mathbf{R} = \mathbb{E}[\hat{\mathbf{R}}] = \begin{cases} \mathbf{A}\frac{\mathbf{S}\mathbf{S}^*}{N}\mathbf{A}^* + \sigma^2\mathbf{I} & \text{under Assumption 1,} \\ \mathbf{A}\mathbf{\Gamma}\mathbf{A}^* + \sigma^2\mathbf{I} & \text{under Assumption 2} \end{cases}$$

For our asymptotic analysis, we consider that $N = N(M)$ is a function of M such that $\frac{M}{N} \to c > 0$, while K is independent of M. The DoA $\theta_1, \ldots, \theta_K$ may depend on M for the results in Section III, while we distinguish two situations for the results in Section IV (see remark below). For the sake of readability, we may add a subscript M to avoid ambiguity for quantities depending on M.

Remark 1. *The assumption that $\theta_1, \ldots, \theta_K$ are fixed with respect to M is relevant for scenarios where the source DoA are widely spaced compared to the array beamwidth $\frac{2\pi}{M}$ (see [9]); this assumption is used to obtain the results of Section IV-A, IV-B and IV-C. The scenario of closely spaced DoA, having an angular spacing of the order of a beamwidth requires to consider DoA depending on M. This assumption and its impact on the results of this paper is discussed in Section IV-D.*

Under the asymptotic framework described previously, we notice that for both Assumptions 1 and 2

$$\left\|\mathbf{R} - \left(\mathbf{A}\mathbf{\Gamma}\mathbf{A}^* + \sigma^2\mathbf{I}\right)\right\|_2 \xrightarrow[M\to\infty]{} 0,$$

where matrix $\mathbf{A}\mathbf{\Gamma}\mathbf{A}^* + \sigma^2\mathbf{I}$ has a Toeplitz structure, a fact which will be fully exploited in the next section to study the rectification scheme.

Remark 2. *The signal model used in this paper differs substantially from the one in [17], which involves a sequence of correlation coefficients $(r_n)_{n\in\mathbb{Z}} \in \ell^1(\mathbb{Z})$ and the sequence $(\mathbf{R}_M)_{M\geq 1}$ of Toeplitz correlations matrices with*

$\mathbf{R}_M = (r_{i-j})_{i,j\in\{1,\ldots,M\}}$. *In our model, we are considering a triangular array of correlation coefficients since we deal with a sequence $(\mathbf{R}_M)_{M\geq 1}$ of correlation matrices with M-dependent entries. Moreover, the absolute summability assumption is not satisfied here.*

III. THE RECTIFIED SCM

Let us denote by $\hat{\mathbf{R}}$ the SCM of the received signal defined as the matrix

$$\hat{\mathbf{R}} = \frac{\mathbf{Y}\mathbf{Y}^*}{N}.$$

Since $\hat{\mathbf{R}}$ fails to preserve a Toeplitz structure, it was originally proposed in [18] to replace it by its closest neighbor in the subspace of Toeplitz matrices $\mathcal{T}$. This subspace being endowed with the usual Hilbert-Schmidt inner product, let us consider the orthonormal basis $\mathbf{E}_{-M+1}, \ldots, \mathbf{E}_{M-1}$ where

$$\mathbf{E}_m = \sqrt{\frac{1}{M - |m|}}\mathbf{J}^m,$$

with $\mathbf{J}$ the upper $M \times M$ shift matrix [3] and denote by $\pi_{\mathcal{T}}$ the linear map representing the orthogonal projection onto $\mathcal{T}$, given by

$$\pi_{\mathcal{T}}(\mathbf{X}) = \sum_{m=-M+1}^{M-1} \text{Tr}\left(\mathbf{E}_m^*\mathbf{X}\right)\mathbf{E}_m. \quad (2)$$

The Toeplitz rectification (or simply Toeplitzification) of $\hat{\mathbf{R}}$ is then defined as the projected matrix

$$\tilde{\mathbf{R}} = \pi_{\mathcal{T}}(\hat{\mathbf{R}}).$$

Rewriting the individual elements of $\tilde{\mathbf{R}}$ leads to

$$\left[\tilde{\mathbf{R}}\right]_{i,j} = \frac{1}{M - |i-j|} \sum_{k-l=i-j} \left[\hat{\mathbf{R}}\right]_{k,l},$$

which means that the Toeplitzification can be also interpreted as an averaging along the diagonals of $\hat{\mathbf{R}}$.

It turns out that the rectified SCM is a consistent estimate of $\mathbf{R}$, as stated by the following result, whose proof is deferred to Appendix B.

Theorem 1. *Under Assumptions 1 or 2, we have*

$$\left\|\tilde{\mathbf{R}} - \mathbf{R}\right\|_2 \xrightarrow[M\to\infty]{a.s.} 0.$$

Theorem 1 implies in particular

$$\sup_{\|\mathbf{b}_1\|_2, \|\mathbf{b}_2\|_2=1} \left|\mathbf{b}_1^*\left(\tilde{\mathbf{R}} - \mathbf{R}\right)\mathbf{b}_2\right| \xrightarrow[M\to\infty]{a.s.} 0, \quad (3)$$

where the supremum is taken over all unit norm vectors $\mathbf{b}_1, \mathbf{b}_2$. Another consequence of Theorem 1, which immediately follows from Weyl's inequality, is given by the convergence

$$\max_{k=1,\ldots,M} \left|\tilde{\lambda}_k - \lambda_k\right| \xrightarrow[M\to\infty]{a.s.} 0, \quad (4)$$

where $\tilde{\lambda}_1 \geq \ldots \geq \tilde{\lambda}_M$ denote the eigenvalues of the rectified S.C.M $\tilde{\mathbf{R}}$ and $\lambda_1 \geq \ldots \geq \lambda_K > \lambda_{K+1} = \ldots = \lambda_M = \sigma^2$ denote the eigenvalues of $\mathbf{R}$.

---

[3] By convention, $\mathbf{J}^0 = \mathbf{I}$ and for $m \geq 0$, $\mathbf{J}^{-m} = (\mathbf{J}^T)^m$.

Note that the result of Theorem 1 is in stark contrast with the behaviour of the non-rectified SCM $\hat{\mathbf{R}}$, which is not a consistent estimator. Indeed, although the weaker convergence

$$\mathbf{b}_1^* \left(\hat{\mathbf{R}} - \mathbf{R} \right) \mathbf{b}_2 \xrightarrow[M \to \infty]{a.s.} 0$$

is still valid, the convergence in spectral norm does not hold. Indeed, even in the pure noise case, that is $\mathbf{Y} = \mathbf{V}$ ($\mathbf{R} = \sigma^2 \mathbf{I}$), it is well-known, (see e.g [21]) that $\left\| \hat{\mathbf{R}} \right\|_2 \to \sigma^2(1 + \sqrt{c})^2 > \sigma^2$ a.s as $M \to \infty$.

We conclude this section by the following corollary which will be used to study the behaviour of the rectified MUSIC cost function and whose proof is deferred to Appendix A.

Corollary 1. *Let $\tilde{\mathbf{\Pi}}, \mathbf{\Pi}$ be the orthogonal projection matrices onto the eigenspace associated with the $M - K$ smallest eigenvalues of $\tilde{\mathbf{R}}$ and $\mathbf{R}$ respectively. Then under Assumption 1 or 2,*

$$\left\| \tilde{\mathbf{\Pi}} - \mathbf{\Pi} \right\|_2 \xrightarrow[M \to \infty]{a.s.} 0. \tag{5}$$

IV. A STATISTICAL ANALYSIS OF R-MUSIC

Having characterized the first order behaviour of the rectified S.C.M $\tilde{\mathbf{R}}$ in terms of spectral norm, we are now able to study the statistical performance of the MUSIC algorithm based on $\tilde{\mathbf{R}}$, in terms of consistency and asymptotic normality.

A. Consistency of R-MUSIC

We recall that the MUSIC method relies on the property that the DoA $\theta_1, \ldots, \theta_K$ are the unique zeros of the cost function

$$\eta(\theta) = \|\mathbf{\Pi} \mathbf{a}(\theta)\|_2^2, \tag{6}$$

where $\mathbf{\Pi}$ is the orthogonal projection matrix onto the kernel of $\mathbf{R} - \sigma^2 \mathbf{I}$. The standard estimate of the MUSIC cost function consists in replacing $\mathbf{\Pi}$ by its corresponding estimate $\hat{\mathbf{\Pi}}$ from the S.C.M $\hat{\mathbf{R}}$, leading to

$$\hat{\eta}(\theta) = \left\| \hat{\mathbf{\Pi}} \mathbf{a}(\theta) \right\|_2^2.$$

The DoA are then estimated as the K deepest local minima of $\theta \mapsto \hat{\eta}(\theta)$. In the same way, the Rectified-MUSIC (R-MUSIC) method is based on the rectified S.C.M, and the DoA estimate are taken as the K deepest local minima of

$$\tilde{\eta}(\theta) = \left\| \tilde{\mathbf{\Pi}} \mathbf{a}(\theta) \right\|_2^2, \tag{7}$$

where $\tilde{\mathbf{\Pi}}$ is the orthogonal projection matrix associated with the $M - K$ smallest eigenvalues of $\tilde{\mathbf{R}}$. The cost function estimate (7) of (6) is consistent; indeed, from Corollary 1, we immediately obtain

$$\sup_{\theta \in (-\pi, \pi]} |\tilde{\eta}(\theta) - \eta(\theta)| \xrightarrow[M \to \infty]{a.s.} 0. \tag{8}$$

In order to study the behaviour of the R-MUSIC DoA estimates, we need to define the latter properly. Consider K compact intervals $\mathcal{I}_1, \ldots, \mathcal{I}_K \subset (-\pi, \pi]$ such that $\theta_k \in \mathcal{I}_k$ for all k. The R-MUSIC DoA estimate of θ_k is then defined as

$$\tilde{\theta}_k = \operatorname*{argmin}_{\theta \in \mathcal{I}_k} \tilde{\eta}(\theta).$$

It turns out that the property (8) automatically implies the following consistency result (see [7, Prop. 4.1.] for a proof).

Theorem 2. *Under both Assumptions 1 and 2, and assuming $\theta_1, \ldots, \theta_K$ are fixed with respect to M, the R-MUSIC DoA estimators satisfy for all $k = 1, \ldots, K$,*

$$M \left(\tilde{\theta}_k - \theta_k \right) \xrightarrow[M \to \infty]{a.s.} 0.$$

The R-MUSIC method thus provides consistent DoA estimates, with an error decreasing faster than $\frac{1}{M}$; this property is usually refered to as M-*consistency*.

The result of Theorem 2 is similar to the M-consistency result obtained for the methods MUSIC/G-MUSIC in [9]. Nevertheless, the latter also require the additional condition

$$\frac{\gamma_K}{\sigma^2} > \sqrt{c},$$

where γ_K is the smallest positive eigenvalue of $\mathbf{\Gamma}$. This condition, refered to as *subspace separation*, is needed to ensure clear signal/noise subspace separation in the spectrum of $\hat{\mathbf{R}}$, and can be interpreted as a minimal SNR constraint for the source with the lowest energy. Since the R-MUSIC method does not require such an assumption, a performance gain over MUSIC/G-MUSIC is expected, especially in the low SNR regions, as it will be shown on numerical examples in Section V.

B. Asymptotic normality of R-MUSIC: the conditional model

We now study the 2nd order behaviour of the R-MUSIC DoA estimates, in terms of asymptotic normality, and consider first the settings of Assumption 1 (conditional model). For that purpose, we define

$$\rho_k^2 = \frac{\sigma^2 c \left(\sigma^2 \|\mathbf{T}_k\|_F^2 + 2 \left\| \mathbf{T}_k \mathbf{A} \frac{\mathbf{S}}{\sqrt{N}} \right\|_F^2 \right)}{\left\| \mathbf{\Pi} \frac{\mathbf{a}'(\theta_k)}{M} \right\|_2^4}, \tag{9}$$

where $\mathbf{T}_k$ is the Hermitian Toeplitz matrix defined by

$$\mathbf{T}_k = \frac{1}{2} \pi_{\mathcal{T}} \left((\mathbf{A}\mathbf{\Gamma}\mathbf{A}^*)^\sharp \mathbf{a}(\theta_k) \frac{\mathbf{a}'(\theta_k)^*}{M} \mathbf{\Pi} \right)$$
$$+ \frac{1}{2} \pi_{\mathcal{T}} \left(\mathbf{\Pi} \frac{\mathbf{a}'(\theta_k)}{M} \mathbf{a}(\theta_k)^* (\mathbf{A}\mathbf{\Gamma}\mathbf{A}^*)^\sharp \right), \tag{10}$$

where $(.)^\sharp$ stands for the pseudo-inverse. Note that the term ρ_k^2 will serve as an asymptotic variance for the k-th R-MUSIC DoA estimate. We also define the following quantity, which will describe an asymptotic bias:

$$\Delta_k = \frac{1}{M} \frac{\operatorname{Tr} \left(\mathbf{A} \left(\frac{\mathbf{S}\mathbf{S}^*}{N} - \mathbf{\Gamma} \right) \mathbf{A}^* \mathbf{T}_k \right)}{\left\| \mathbf{\Pi} \frac{\mathbf{a}'(\theta_k)}{M} \right\|_2^2}. \tag{11}$$

Theorem 3. *Under Assumption 1, and assuming $\theta_1, \ldots, \theta_K$ are fixed with respect to M, the R-MUSIC DoA estimators satisfy, for all $k = 1, \ldots, K$,*

$$\frac{M^{3/2} \left(\tilde{\theta}_k - \theta_k - \Delta_k \right)}{\rho_k} \xrightarrow[M \to \infty]{\mathcal{D}} \mathcal{N}(0, 1). \tag{12}$$

The proof of Theorem 3 is deferred to Appendix C.

Similarly to Theorem 2, the result of Theorem 3 holds whatever the noise power σ^2 and or the source power matrix $\boldsymbol{\Gamma}$ are, which is again in contrast to the asymptotic normality results for G-MUSIC/MUSIC obtained in [9].

Looking at the statement of Theorem 3, we notice the presence of an asymptotic bias term Δ_k, which is independent of σ^2, as well as a variance term ρ_k^2 vanishing as $\sigma^2 \to 0$. This sheds light on a major drawback of R-MUSIC. Indeed, from the proof of Theorem 3, one can deduce that

$$\mathbb{E}\left|\tilde{\theta}_k - \theta_k\right|^2 \xrightarrow[\sigma^2 \to 0]{} |\Delta_k|^2, \tag{13}$$

showing that the MSE of R-MUSIC saturates in high SNR regions at the bias term $|\Delta_k|^2$. Nevertheless, this bias is mitigated for large dimensions (and fixed SNR) since we also have

$$\Delta_k = o\left(\frac{1}{M^{3/2}}\right),$$

as $M \to \infty$. Modifying the proof of Theorem 3 with a uniform integrability argument, we can also deduce that

$$\mathbb{E}\left|\tilde{\theta}_k - \theta_k\right|^2 = \frac{\rho_k^2}{M^3} + o\left(\frac{1}{M^3}\right), \tag{14}$$

We conclude from this discussion that in practice, the bias contribution in the MSE is negligible at low SNR and large M, N, while it becomes preponderant for reasonable dimensions M, N and high SNR.

Finally, we mention that ρ_k^2 can be approximated by a more explicit expression, using Lemma 3 in Appendix C-A:

$$\mathbb{E}\left|\tilde{\theta}_k - \theta_k\right|^2 = \frac{1}{M^3} \frac{6\sigma^2 c(\gamma_k + \sigma^2)}{\gamma_k^2} + o\left(\frac{1}{M^3}\right). \tag{15}$$

Note also that the asymptotic MSE (15) coincides with the one of G-MUSIC [9] when $\frac{\gamma_k}{\sigma^2} \gg 1$ (large SNR), as well as the CRB [4].

C. Asymptotic normality of R-MUSIC: the unconditional model

In this section, we present the asymptotic normality result for the unconditional model (Assumption 2). For ease of reading, we keep the same notation for the variance term (9), which is now expressed as

$$\rho_k^2 = \frac{c\left\|\mathbf{R}^{1/2}\mathbf{T}_k\mathbf{R}^{1/2}\right\|_F^2}{\left\|\boldsymbol{\Pi}\frac{\mathbf{a}'(\theta_k)}{M}\right\|_2^4}, \tag{16}$$

where $\mathbf{T}_k$ is defined in (10) (recall that $\mathbf{R} = \mathbf{A}\boldsymbol{\Gamma}\mathbf{A}^* + \sigma^2\mathbf{I}$ in the unconditional model).

The equivalent of Theorem 3 is given by the following result, whose proof is also sketched in Appendix C.

Theorem 4. *Under Assumption 2, and assuming $\theta_1, \dots, \theta_K$ are fixed with respect to M, the R-MUSIC DoA estimators satisfy, for all $k = 1, \dots, K$,*

$$\frac{M^{3/2}\left(\tilde{\theta}_k - \theta_k\right)}{\rho_k} \xrightarrow[M \to \infty]{\mathcal{D}} \mathcal{N}(0, 1). \tag{17}$$

A noticeable fact about Theorem 4 is the absence of an asymptotic bias term, compared to Theorem 3, which is due to the specific unconditional model, where the source signal are assumed random. In this unconditional case, we still observed a stagnation of the MSE for high SNR, which is this time directly contained in the variance term (16) since

$$\rho_k^2 \xrightarrow[\sigma^2 \to 0]{} \frac{c\left\|(\mathbf{A}\boldsymbol{\Gamma}\mathbf{A}^*)^{1/2}\mathbf{T}_k(\mathbf{A}\boldsymbol{\Gamma}\mathbf{A}^*)^{1/2}\right\|_F^2}{\left\|\boldsymbol{\Pi}\frac{\mathbf{a}'(\theta_k)}{M}\right\|_2^4}.$$

Nevertheless, this phenomenon is negligible for large M, N since we also have (see Lemma 3)

$$\rho_k^2 \xrightarrow[M \to \infty]{} \frac{6\sigma^2 c(\gamma_k + \sigma^2)}{\gamma_k^2}. \tag{18}$$

D. Some comments on the closely spaced DoA scenario

When dealing with large array of sensors and widely spaced source DoA (with a spacing much larger than a beamwidth $\frac{2\pi}{M}$), it is well-known that most of the DoA estimation methods perform similarly in terms of MSE (from the asymptotic regime $M \to \infty$ point of view). In particular, it is the case for MUSIC and its improved version G-MUSIC (see e.g. [9]), and also for R-MUSIC from the comments formulated at the end of Section IV-B.

In that case, a more relevant scenario consists in considering closely spaced DoA, with a spacing of magnitude $\mathcal{O}\left(\frac{1}{M}\right)$ [22], [9], which is summarized in the following assumption.

Assumption 3. *We consider $K = 2$, $\boldsymbol{\Gamma} = \mathbf{I}$ and DoA depending on M in such a way that*

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

where $\alpha > 0$ is independent of M.

In that case, the R-MUSIC DoA estimates are now defined as

$$\tilde{\theta}_{k,M} = \underset{\theta \in \mathcal{I}_{k,M}}{\arg\min}\, \tilde{\eta}(\theta) \tag{19}$$

where $\mathcal{I}_{k,M}$ is the interval

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

Under Assumption 3, the classical MUSIC method (when defined similarly to (19)) looses the M-consistency property (see Section IV-A), while the improved G-MUSIC method succeeds in separating the DoA provided the separation condition

$$1 - \left|\frac{\sin \alpha/2}{\alpha/2}\right| > \sigma^2\sqrt{c}$$

holds. Since (7) is a uniformly consistent estimate of the cost function (6), the M-consistency for the R-MUSIC DoA estimates is in fact automatically verified, by using the proof of [9, Th.6]. Therefore, under Assumption 1 or 2, and under the additional Assumption 3, the R-MUSIC DoA estimates defined by (19) satisfy, for $k \in \{1, 2\}$,

$$M\left(\tilde{\theta}_{k,M} - \theta_{k,M}\right) \xrightarrow[M \to \infty]{a.s.} 0.$$

Concerning the asymptotic normality results, one can see that the proof of Theorems 3 and 4 remains unchanged in the context of Assumption 3, and the convergence (12) and (17) are still valid as long as

$$0 < \liminf_{M\to\infty} \rho_k \le \limsup_{M\to\infty} \rho_k < \infty,$$

for $k \in \{1,2\}$, where ρ_k is defined by (9) or (16). This is ensured by Lemma 4 in Appendix C-A, and thus we also obtain under Assumption 3

$$\frac{M^{3/2}\left(\tilde{\theta}_k - \theta_k\right)}{\rho_k} \xrightarrow[M\to\infty]{\mathcal{D}} \mathcal{N}(0,1),$$

where ρ_k is given by (9) or (16).

V. NUMERICAL RESULTS

In this section, we provide numerical results on the R-MUSIC DoA estimates, illustrating the theoretical predictions of the previous sections.

Simulations settings. For all the simulations, we consider $K = 2$ sources, with $M = 40$ sensors and $N = 80$ samples, and a total of 500 Monte-Carlo runs to evaluate the DoA estimates (except for Figures 5 and 4 where $N = 30$). The signal matrix $\mathbf{S}$ is of the form

$$\mathbf{S} = \mathbf{\Gamma}^{1/2}\mathbf{X}\tilde{\mathbf{\Gamma}}^{1/2}, \tag{20}$$

where

- $\mathbf{X}$ is a matrix with standard i.i.d. complex circular Gaussian entries. In the conditional case, only one realization of $\mathbf{X}$ is generated and kept fixed for all runs while in the unconditional case, a new $\mathbf{X}$ is generated at each run;
- $\mathbf{\Gamma}$ represents the spatial correlation between sources which will be fixed to $\mathbf{I}$;
- $\tilde{\mathbf{\Gamma}}$ represents the temporal correlation of the source signals; it will be either fixed to $\tilde{\mathbf{\Gamma}} = \mathbf{I}$ or

$$\tilde{\mathbf{\Gamma}} = \left(\xi^{|i-j|}\right)_{i,j\in\{1,\dots,N\}},$$

with $\xi = 0.7$, depending whether we consider temporally white signals or AR(1) signals.

For all the simulations in the conditional model provided below, we have generated a unique $\mathbf{X}$ matrix which is reused for the different Figures (except again for Figures 5 and 4). The SNR is defined as $-10\log(\sigma^2)$ and the estimated DoA are taken as the K largest local maxima of the cost function under study.

Widely spaced DoA. We first deal with the widely spaced DoA scenario, by considering DoA $\theta_1 = 0$ and $\theta_2 = 5 \times \frac{2\pi}{M}$. In Figures 1, 2 and 3, we plot, as a function of the SNR, the empirical MSE of $\hat{\theta}_1$ as well as the theoretical MSE predicted in Theorems 3 and 4, that is

$$\text{MSE}(\hat{\theta}_1) = \begin{cases} \frac{\rho_1^2}{M^3} + \Delta_1^2 & \text{for the conditional case,} \\ \frac{\rho_1^2}{M^3} & \text{for the unconditional case} \end{cases}$$

where Δ_1 is defined in (11) and ρ_1 in (9) or (16). In Figures 1 and 2, the conditional case is considered for temporally white and AR(1) signals respectively (recall that we reuse the same matrix $\mathbf{X}$ in (20) for both plots), together with the associated Cramer-Rao bounds. In both cases, a good match is observed between the empirical and predicted MSE. In
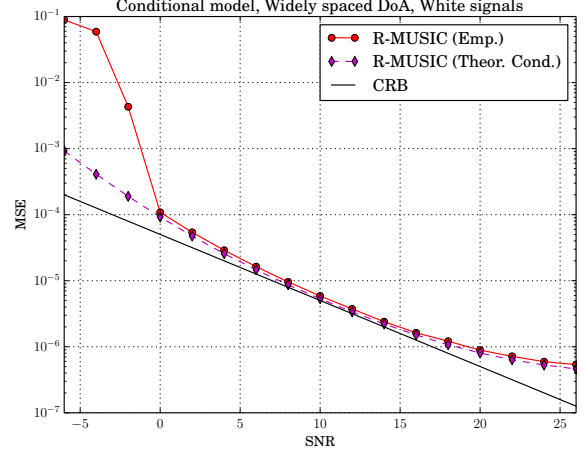

Fig. 1. Empirical and Theoretical MSE vs SNR (conditional model, temporally white signals, widely spaced DoA)
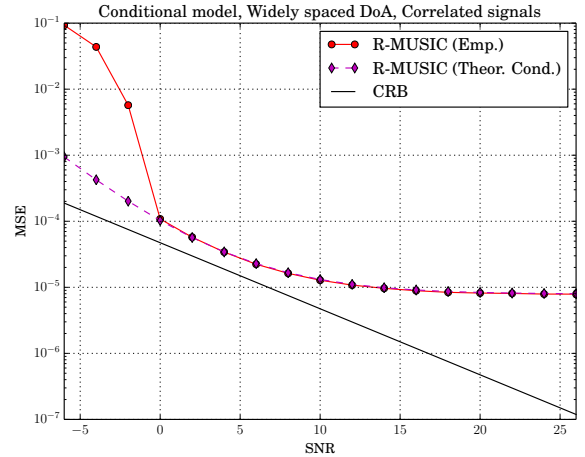

Fig. 2. Empirical and Theoretical MSE vs SNR (conditional model, temporally correlated signals, widely spaced DoA)

Figure 3, the unconditional case is considered for both temporally white and AR(1) signals. In this case, we observe a less accurate prediction when temporal correlation is present, a situation which is not covered by the assumptions of the unconditional scenario (cf. Assumption 2).

Closely spaced DoA. Figures 4 and 5 provide simulations for the closely spaced DoA scenario for the conditional and unconditional models respectively, where $\theta_1 = 0$ and $\theta_2 = 0.25 \times \frac{2\pi}{M}$, and where the source signals are temporally white Gaussian. The number of samples is fixed to $N = 30$ to illustrate an undersampled situation, for which the analysis we provide is still valid. In that case, the performance of R-MUSIC severely degrades, and the
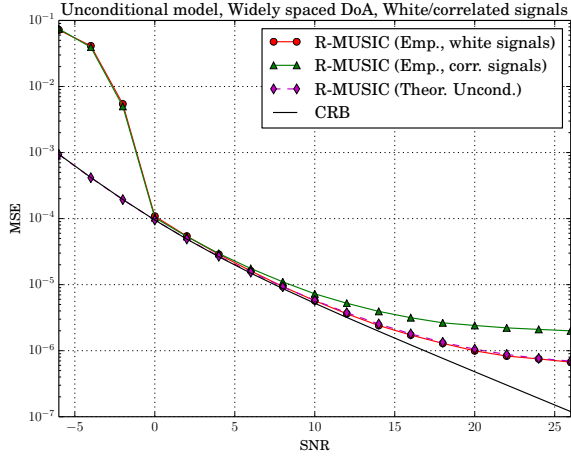
Fig. 3. Empirical and Theoretical MSE vs SNR (unconditional model, temporally white and AR(1) signals, widely spaced DoA)

Fig. 5. Empirical and Theoretical MSE vs SNR (unconditional model, temporally white signals, closely spaced DoA, undersampled)

predicted MSE obtained in Theorem 3 may present some discrepancy with the empirical one, see the discussion on the limitations of the asymptotic analysis below.
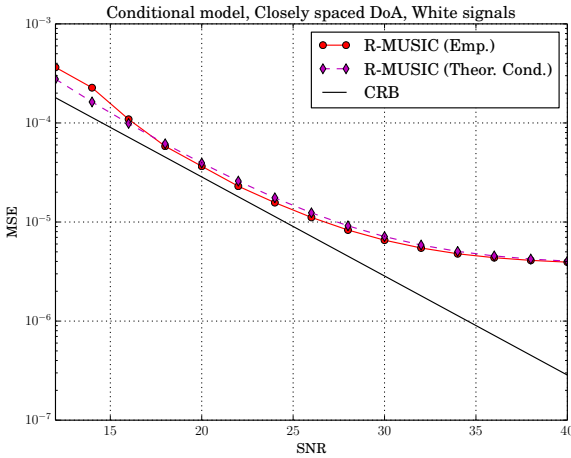

Fig. 4. Empirical and Theoretical MSE vs SNR (conditional model, temporally white signals, closely spaced DoA, undersampled)

Comparisons with other methods. Finally, we also provide in Figures 7 a comparison with the G-MUSIC method mentioned in Section IV-A, for closely spaced DoA ($\theta_1 = 0$ and $\theta_2 = 0.25 \times \frac{2\pi}{M}$) and temporally white source signals, in the context of the unconditional model. Although the theoretical performance of these methods in the large dimensional regime is still an open problem, the ESPRIT and Root-MUSIC methods, which are known to outperform MUSIC for moderate sample size [23], are also simulated. To compare similar algorithms, the "Root" version of R-MUSIC is implemented. We notice, as expected, a gain of performance for R-MUSIC in the low SNR region in terms of threshold points, that is the SNR under which the MSE significantly deteriorates and departs from the one predicted by the asymptotic analysis (these points can be seen around 16 dB and 24 dB for the R-MUSIC and G-MUSIC respectively). This difference between the two methods can
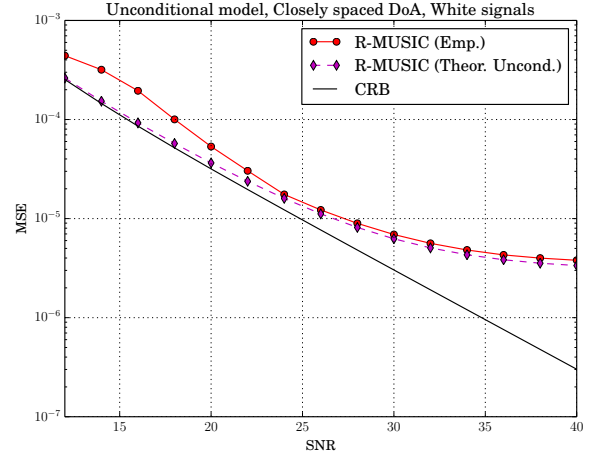

Fig. 6. Empirical and Theoretical MSE vs SNR of Root R-MUSIC, Root MUSIC, ESPRIT and G-MUSIC (unconditional model, temporally white signals, widely spaced DoA)

be explained by the fact that G-MUSIC requires a sufficiently large SNR to work properly (cf. discussion in Section IV-A). However, G-MUSIC, which is efficient when the SNR goes to infinity, clearly outperforms R-MUSIC for the high SNR region, which suffers from the SNR stagnation.

Potential limitations of the asymptotic analysis. As we have seen previously in Figure 3, the unconditional model is not robust to temporal correlation of signals, and the theoretical analysis is no more accurate in that situation. Comparatively, the conditional model is able to handle temporal correlations in its very assumption, and the theoretical analysis we have provided is valid in that case, as we have noticed e.g. in Figure 2. Nevertheless, as stated in Assumption 1, the spatial sample correlation matrix $\frac{\mathbf{SS}^*}{N}$ of the source signals must be "close" in practice to a diagonal matrix; this situation may not be fully verified for small sample size N and in that case, we may observe some discrepancy between the empirical and asymptotic MSE.
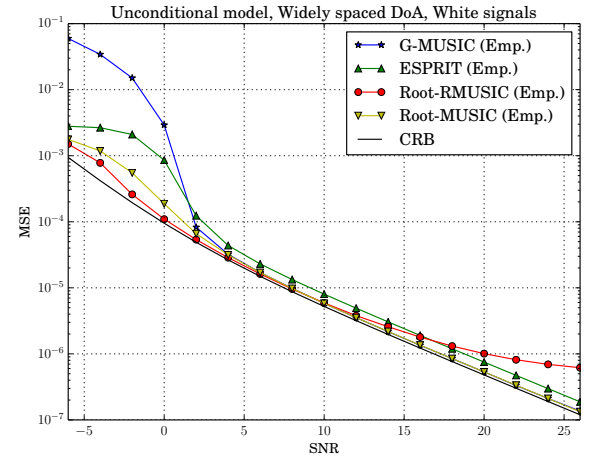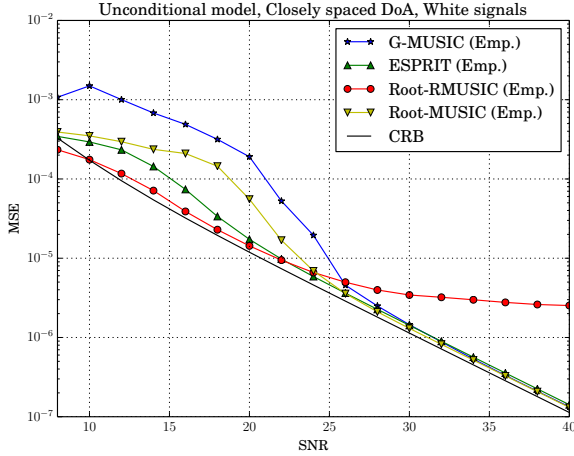
Fig. 7. Empirical and Theoretical MSE vs SNR of Root R-MUSIC, Root MUSIC, ESPRIT and G-MUSIC (unconditional model, temporally white signals, closely spaced DoA)

VI. Conclusion

In this paper, we have provided a statistical analysis of the R-MUSIC method, in terms of consistency and asymptotic normality, in the regime where both the number of sensors and the number of snapshots go to infinity at the same rate. The study has been conducted under both the conditional and the unconditional signal model, which are commonly used in array processing, as well as for two scenarios involving widely and closely spaced DoA, where in the latter case the source DoA have a spacing of the order of a beamwidth (i.e. inversely proportional to the number of sensors). In particular, it has been established that R-MUSIC is able to asymptotically separate closely spaced DoA. Compared to G-MUSIC, another method sharing this latter property, R-MUSIC does not require a minimal SNR to be valid. The asymptotic analysis of the MSE has also enlighted the well-known threshold effect in the large SNR region, which demonstrates that the use of R-MUSIC method is essentially interesting for the low SNR region. Finally, numerical results have demonstrated the accuracy of the asymptotic results, and we have underlined the potential limitations of the theoretical analysis, especially in situations involving small sample size or temporal correlations.

Appendix A
Proof of Corollary 1

Let $\mathcal{C}$ be a circle, counterclockwise oriented, enclosing σ^2 and leaving outside $\gamma_1 + \sigma^2, \ldots, \gamma_K + \sigma^2$, where $\gamma_1, \ldots, \gamma_K$ are the diagonal entries of $\mathbf{\Gamma}$. Since the K largest eigenvalues of $\mathbf{R}$ have $\gamma_1 + \sigma^2, \ldots, \gamma_K + \sigma^2$ as limits when $M \to \infty$ whereas the $M - K$ smallest ones coincide with σ^2, Theorem 1 through convergence of the eigenvalues (4) thus implies that the K largest eigenvalues of $\tilde{\mathbf{R}}$ converge outside $\mathcal{C}$ and the $M - K$ smallest eigenvalues converge inside $\mathcal{C}$. Using residue theorem, we consequently obtain the following

contour integral representation, a.s. for all large M,

$$\tilde{\mathbf{\Pi}} - \mathbf{\Pi} = \frac{1}{2\pi i} \oint_{\mathcal{C}} \left(\left(\tilde{\mathbf{R}} - z\mathbf{I} \right)^{-1} - (\mathbf{R} - z\mathbf{I})^{-1} \right) \mathrm{d}z$$
$$= \frac{1}{2\pi i} \oint_{\mathcal{C}} \left(\tilde{\mathbf{R}} - z\mathbf{I} \right)^{-1} \left(\mathbf{R} - \tilde{\mathbf{R}} \right) (\mathbf{R} - z\mathbf{I})^{-1} \mathrm{d}z.$$

Finally, since a.s.

$$\limsup_{M \to \infty} \sup_{z \in \mathcal{C}} \left\| \left(\tilde{\mathbf{R}} - z\mathbf{I} \right)^{-1} \right\|_2 < \infty, \tag{21}$$

as well as

$$\limsup_{M \to \infty} \sup_{z \in \mathcal{C}} \left\| (\mathbf{R} - z\mathbf{I})^{-1} \right\|_2 < \infty, \tag{22}$$

we easily deduce (5).

Appendix B
Proof of Theorem 1

A. Preliminary results

The following lemma gives upper bounds on the spectral norm of any Toeplitz rectified matrix. Note that the second bound, although far from being optimal, is sufficient for the technicalities of this paper.

Lemma 1. *There exists a constant $C > 0$ such that for any $M \in \mathbb{N}^*$ and $M \times M$ complex matrix $\mathbf{C}$, we have*

$$\|\pi_{\mathcal{T}}(\mathbf{C})\|_2 \leq \sup_{\nu \in (-\pi, \pi]} |\mathrm{Tr}\left(\mathbf{C}\mathbf{L}(\nu)\right)| \tag{23}$$

where

$$\mathbf{L}(\nu) = \sum_{m=-M+1}^{M-1} \frac{\exp(im\nu)}{M - |m|} \mathbf{J}^m.$$

Moreover, $\|\mathbf{L}(\nu)\|_F = \mathcal{O}\left(\sqrt{\log(M)} \right)$.

Proof. Denote by $\mathbf{\Omega} = (\omega_{i-j})_{i,j=1,\ldots,M}$ the Toeplitz matrix $\pi_{\mathcal{T}}(\mathbf{C})$, where by definition

$$\omega_m = \frac{1}{M - |m|} \mathrm{Tr}\left(\mathbf{C}\mathbf{J}^m\right).$$

Using the standard bound on spectral norm of Toeplitz matrices based on Parseval's Theorem, we have

$$\|\pi_{\mathcal{T}}(\mathbf{C})\|_2 \leq \sup_{(-\pi, \pi]} \left| \sum_{m=-M+1}^{M-1} \omega_m \exp(im\nu) \right|$$
$$\leq \sup_{(-\pi, \pi]} |\mathrm{Tr}\left(\mathbf{C}\mathbf{L}(\nu)\right)|,$$

which proves (23). Finally, using classical results on harmonic series, we have

$$\|\mathbf{L}(\nu)\|_F^2 = \|\mathbf{L}(0)\|_F^2$$
$$\leq \frac{1}{M} + 2 \sum_{m=1}^{M-1} \frac{1}{M - |m|}$$
$$\leq C \log(M).$$

$\square$

The next property is a concentration bound which follows directly from mixing two Chernoff bounds for sub-Gaussian and sub-exponential distributions (see e.g. [24]).

Lemma 2. *Let* $\mathbf{B} = [\mathbf{b}_1, \ldots, \mathbf{b}_N]$ *be a deterministic* $M \times N$ *matrix,* $\mathbf{x}_1, \ldots, \mathbf{x}_N$ *be independent random vectors such that* $\mathbf{x}_n \sim \mathcal{N}_{\mathbb{C}^M}(\mathbf{b}_n, \mathbf{I})$, $\mathbf{\Xi}$ *a* $M \times M$ *Hermitian matrix and define*

$$Z = \frac{1}{N} \sum_{n=1}^{N} \mathbf{x}_n^* \mathbf{\Xi} \mathbf{x}_n.$$

Then there exists constant $\kappa_1, \kappa_2 > 0$ *and independent of* $M, N, \mathbf{B}, \mathbf{\Xi}$ *such that for all* $\epsilon \geq 0$,

$$\mathbb{P}\left(|Z - \mathbb{E}[Z]| \geq \epsilon\right) \leq$$
$$\kappa_1 \exp\left(-\kappa_2 \min\left\{\frac{N\epsilon^2}{\|\mathbf{\Xi}\|_F^2}, \frac{N\epsilon^2}{\operatorname{Tr}\left(\mathbf{\Xi}^2 \frac{\mathbf{B}\mathbf{B}^*}{N}\right)}, \frac{N\epsilon}{\|\mathbf{\Xi}\|_2}\right\}\right).$$

B. The proof

Conditional model. We start the proof by considering Assumption 1. Using triangular inequality, we obtain

$$\left\|\tilde{\mathbf{R}} - \mathbf{R}\right\|_2 \leq \chi_{1,M} + \chi_{2,M} + \chi_{3,M},$$

where

$$\chi_{1,M} = \left\|\tilde{\mathbf{R}} - \mathbb{E}[\tilde{\mathbf{R}}]\right\|_2,$$
$$\chi_{2,M} = \left\|\mathbb{E}[\tilde{\mathbf{R}}] - (\mathbf{A}\mathbf{\Gamma}\mathbf{A}^* + \sigma^2\mathbf{I})\right\|_2,$$
$$\chi_{3,M} = \left\|\mathbf{A}\left(\frac{\mathbf{S}\mathbf{S}^*}{N} - \mathbf{\Gamma}\right)\mathbf{A}^*\right\|_2.$$

By hypothesis, we obviously have $\chi_{3,M} \to 0$ as $M \to \infty$. Using (23) in Lemma 1 below (see Appendix B-A), we have

$$\chi_{2,M} \leq \sup_{\nu \in (-\pi, \pi]} |g_M(\nu)|, \tag{24}$$

where

$$g_M(\nu) = \operatorname{Tr}\left(\mathbf{A}\left(\frac{\mathbf{S}\mathbf{S}^*}{N} - \mathbf{\Gamma}\right)\mathbf{A}^*\mathbf{L}(\nu)\right), \tag{25}$$

with

$$\mathbf{L}(\nu) = \sum_{m=-M+1}^{M-1} \frac{\exp(\mathrm{i}m\nu)}{M - |m|}\mathbf{J}^m. \tag{26}$$

A straightforward bound gives

$$|g_M(\nu)| \leq \left\|\frac{\mathbf{S}\mathbf{S}^*}{N} - \mathbf{\Gamma}\right\|_2 \sum_{k,\ell=1}^{K} |\mathbf{a}(\theta_k)^*\mathbf{L}(\nu)\mathbf{a}(\theta_\ell)|.$$

Since for any θ, θ', $|\mathbf{a}(\theta)^*\mathbf{J}^m\mathbf{a}(\theta')| \leq \frac{M-|m|}{M}$, we deduce using (26) that

$$|\mathbf{a}(\theta)^*\mathbf{L}(\nu)\mathbf{a}(\theta')| \leq 2,$$

for all $\nu \in (-\pi, \pi]$, and since K is independent of M, we easily deduce that $\chi_{2,M} \to 0$ as $M \to \infty$.

It now remains to prove that $\chi_{1,M} \to 0$ a.s. as $M \to \infty$. As above, using Lemma 1, we get

$$\chi_{1,M} \leq \sup_{\nu \in (-\pi, \pi]} |h_M(\nu)|, \tag{27}$$

where

$$h_M(\nu) = \operatorname{Tr}\left(\left(\hat{\mathbf{R}} - \mathbb{E}[\hat{\mathbf{R}}]\right)\mathbf{L}(\nu)\right). \tag{28}$$

This can be rewritten as

$$h_M(\nu) = \sum_{k,\ell=0}^{M-1} \left[\hat{\mathbf{R}} - \mathbb{E}[\hat{\mathbf{R}}]\right]_{k,\ell} \frac{\exp\left(\mathrm{i}(k-\ell)\nu\right)}{M - |k-\ell|}$$
$$= M\,\mathbf{a}(\nu)^* \left(\left(\hat{\mathbf{R}} - \mathbb{E}[\hat{\mathbf{R}}]\right) \odot \mathbf{L}(0)\right)^T \mathbf{a}(\nu)$$

Obviously, for all $\nu_1, \nu_2 \in (-\pi, \pi]$, we have $\|\mathbf{a}(\nu_1) - \mathbf{a}(\nu_2)\|_2 \leq CM|\nu_1 - \nu_2|$, where $C > 0$ is independent of M, ν_1, ν_2. To simplify the notations in what follows, the quantity C will be reused as a generic constant which may change from one line to another, but remains independent of M, ν_1, ν_2.

Using the fact that the spectral norm is submultiplicative with respect to the Hadamard product (see e.g. [25]), we obtain

$$|h_M(\nu_1) - h_M(\nu_2)| \leq$$
$$CM^2|\nu_1 - \nu_2| \left\|\hat{\mathbf{R}} - \mathbb{E}[\hat{\mathbf{R}}]\right\|_2 \|\mathbf{L}(0)\|_2.$$

Using the bound $\|\mathbf{L}(0)\|_2 \leq \|\mathbf{L}(0)\|_F = \mathcal{O}\left(\sqrt{\log(M)}\right)$ (Lemma 1) and the fact that (see [26], [27]) a.s.,

$$\left\|\hat{\mathbf{R}} - \mathbb{E}[\hat{\mathbf{R}}]\right\|_2 = \mathcal{O}(1),$$

as $M \to \infty$, we deduce that

$$|h_M(\nu_1) - h_M(\nu_2)| \leq CM^2\sqrt{\log(M)}|\nu_1 - \nu_2|. \tag{29}$$

Consider now the set of points

$$\mathcal{V}_M = \left\{\frac{2\pi m}{M^3} - \pi : m = 0, \ldots, M^3 - 1\right\}.$$

For each $\nu \in (-\pi, \pi]$, let ν_M be the closest point to ν in $\mathcal{V}_M$; thus $|\nu - \nu_M| = \mathcal{O}(M^{-3})$. As a consequence, using (29), we obtain

$$\sup_{\nu \in (-\pi, \pi]} |h_M(\nu)| \leq$$
$$\sup_{\nu \in (-\pi, \pi]} |h_M(\nu) - h_M(\nu_M)| + \sup_{\nu \in (-\pi, \pi]} |h_M(\nu_M)|,$$

which implies that a.s.,

$$\sup_{\nu \in (-\pi, \pi]} |h_M(\nu)| \leq \max_{\nu \in \mathcal{V}_M} |h_M(\nu)| + \mathcal{O}\left(\frac{\sqrt{\log(M)}}{M}\right). \tag{30}$$

Now, replacing $\hat{\mathbf{R}} = \frac{1}{N}\sum_{n=1}^{N} \mathbf{y}_n\mathbf{y}_n^*$ in (28), we notice that $h_M(\nu)$ can be rewritten more explicitely as

$$h_M(\nu) = \frac{1}{N}\sum_{n=1}^{N}\left(\mathbf{y}_n^*\mathbf{L}(\nu)\mathbf{y}_n - \mathbb{E}[\mathbf{y}_n^*\mathbf{L}(\nu)\mathbf{y}_n]\right),$$

where we recall that $\mathbf{y}_1, \ldots, \mathbf{y}_N$ denote the N statistically independent columns of $\mathbf{Y}$. Using Markov's inequality and

Lemma 2 below (with $\mathbf{B} = \frac{\mathbf{AS}}{\sigma}$, $\mathbf{\Xi} = \sigma^2\mathbf{L}(\nu)$), we obtain for all $\epsilon > 0$

$$\mathbb{P}\left(\max_{\nu \in \mathcal{V}_M} |h_M(\nu)| \geq \epsilon\right) \leq \sum_{\nu \in \mathcal{V}_M} \mathbb{P}\left(|h_M(\nu)| \geq \epsilon\right)$$
$$= \mathcal{O}\left(\frac{1}{M^\ell}\right),$$

for all integer ℓ, and therefore $\chi_{1,M} \to 0$ a.s.

Unconditional model. In the unconditional model, we directly need to control the term

$$\chi_{1,M} = \left\|\tilde{\mathbf{R}} - \mathbb{E}[\tilde{\mathbf{R}}]\right\|_2 = \left\|\pi_{\mathcal{T}}\left(\hat{\mathbf{R}} - \mathbb{E}[\hat{\mathbf{R}}]\right)\right\|_2,$$

since $\mathbf{R} = \mathbb{E}[\hat{\mathbf{R}}] = \mathbb{E}[\tilde{\mathbf{R}}]$. As in the conditional case, we have $\chi_{1,M} \leq \sup_{(-\pi,\pi]} |h_M(\nu)|$ where h_M is defined by (28) and the discretization trick (30) is still verified. Finally, invoking again Lemma 2 with $\mathbf{\Xi} = \mathbf{R}^{1/2}\mathbf{L}(\nu)\mathbf{R}^{1/2}$, $\mathbf{B} = \mathbf{0}$, we also end up with $\mathbb{P}\left(\max_{\nu \in \mathcal{V}_M} |h_M(\nu)| \geq \epsilon\right) = \mathcal{O}\left(M^{-\ell}\right)$ for all $\ell \in \mathbb{N}^*$, and thus $\chi_{1,M} \to 0$ a.s. as $M \to \infty$.

<div align="center">

APPENDIX C

PROOF OF THEOREM 3

</div>

A. Preliminary results

The proof of the following lemma, which is straightforward and relies on discrete approximations of the Beta function, is omitted.

Lemma 3. *Let $\mathbf{T}_k$ be the matrix defined in* (10) *and assume $\theta_1, \ldots, \theta_K$ are fixed with respect to M. Then, for all $p, q, k \in \{1, \ldots, K\}$, as $M \to \infty$*

1) $\|\mathbf{T}_k\|_F^2 = \frac{1}{24\gamma_k^2} + \mathcal{O}\left(\frac{1}{\sqrt{M}}\right)$,
2) $\mathbf{a}(\theta_p)^*\mathbf{T}_k^2\mathbf{a}(\theta_q) = \frac{\delta_{k-p}\delta_{k-q}}{48\gamma_k} + \mathcal{O}\left(\frac{1}{\sqrt{M}}\right)$,
3) $\mathbf{a}(\theta_p)^*\mathbf{T}_k\mathbf{a}(\theta_q) = \mathcal{O}\left(\frac{1}{\sqrt{M}}\right)$.

Moreover,

$$\left\|\mathbf{\Pi}\frac{\mathbf{a}'(\theta_k)}{M}\right\|_2^2 = \frac{1}{12} + \mathcal{O}\left(\frac{1}{M}\right).$$

The next Lemma allows to study $\|\mathbf{T}_k\|_F^2$ and $\left\|\mathbf{\Pi}\frac{\mathbf{a}'(\theta_k)}{M}\right\|_2^2$ in the context of closely spaced DoA.

Lemma 4. *Let $\mathbf{T}_k$ be the matrix defined in* (10). *Then, under Assumption 3, for $k \in \{1, 2\}$,*

$$0 < \liminf_{M \to \infty} \|\mathbf{T}_k\|_F^2 \leq \limsup_{M \to \infty} \|\mathbf{T}_k\|_F^2 < \infty. \quad (31)$$

Moreover,

$$0 < \liminf_{M \to \infty} \left\|\mathbf{\Pi}\frac{\mathbf{a}'(\theta_k)}{M}\right\|_2^2 \leq \limsup_{M \to \infty} \left\|\mathbf{\Pi}\frac{\mathbf{a}'(\theta_k)}{M}\right\|_2^2 < \infty. \quad (32)$$

Proof. The proof of (32) is given in [9, App. D]. We prove (31) for $k = 1$, the case $k = 2$ being similar. Let

$$\mathbf{d}_1 = \mathbf{A}(\mathbf{A}^*\mathbf{A})^{-2}\mathbf{A}^*\mathbf{a}(\theta_1) \quad \text{and} \quad \mathbf{d}_2 = \mathbf{\Pi}\frac{\mathbf{a}'(\theta_1)}{M}.$$

Then

$$\|\mathbf{T}_1\|_F^2 =$$
$$\frac{1}{4}\sum_{m=-M+1}^{M-1}\frac{1}{M-|m|}\left|\sum_{n-n'=m}\left(d_{1,n}\overline{d_{2,n'}} + d_{2,n}\overline{d_{1,n'}}\right)\right|^2.$$

Denote $\lambda = \int_0^1 \mathrm{e}^{\mathrm{i}\alpha t}\mathrm{d}t$ and $\gamma = \int_0^1 t\mathrm{e}^{\mathrm{i}\alpha t}\mathrm{d}t$ for short. Using discrete approximations of integrals, we obtain after some computations that

$$\|\mathbf{T}_1\|_F^2 \xrightarrow[M \to \infty]{} \frac{1}{2\left(1 - |\lambda|^2\right)^4}\int_0^1 \frac{|h(t)|^2}{1-t}\mathrm{d}t.$$

where

$$h(t) = \int_0^{1-t}\left(g\left(s+t, s\right) - \overline{g\left(s, s+t\right)}\right)\mathrm{d}s,$$

and

$$g(u, v) = \left(1 - \overline{\lambda}\mathrm{e}^{\mathrm{i}\alpha u}\right)\left(\left(1 - |\lambda|^2\right)v\right.$$
$$\left. + \left(\frac{\lambda}{2} + \gamma\right)\mathrm{e}^{-\mathrm{i}\alpha v} - \left(\frac{1}{2} - \overline{\lambda}\gamma\right)\right).$$

From this, it is straightforward to see that

$$\limsup_{M \to \infty}\|\mathbf{T}_1\|_F^2 < \infty.$$

To handle the limit inferior, we notice that $h(t)$ is a linear combination of the functions $x_0(t) = 1$,

$$x_1(t) = t, \quad x_2(t) = t^2,$$
$$x_3(t) = \mathrm{e}^{\mathrm{i}\alpha t}, \quad x_4(t) = t\mathrm{e}^{\mathrm{i}\alpha t},$$

as well as $\overline{x_3}$ and $\overline{x_4}$, which are linearly independent in the space $\mathrm{L}_{\mathbb{C}}^2([0,1])$. One can check that the coefficients in this linear combination are all zeros only if $\alpha = 0$. Since by Assumption 3, $\alpha > 0$, we deduce that the continuous function h is not identically zero, and thus

$$\liminf_{M \to \infty}\|\mathbf{T}_1\|_F^2 > 0.$$

□

B. The proof

In this proof, the notation C is used to denote a generic positive constant, independent of M, which may change from one line to another.

Following the standard approach for performance analysis of M-estimators, we start with the Taylor expansion of the R-MUSIC cost function around θ_k. From Theorem 2, $\tilde{\eta}$ admits on $\mathcal{I}_k$ a local minima at $\tilde{\theta}_k$ a.s. for all large M, and therefore we have

$$0 = \tilde{\eta}'(\tilde{\theta}_k)$$
$$= \tilde{\eta}'(\theta_k) + \left(\tilde{\theta}_k - \theta_k\right)\tilde{\eta}^{(2)}(\theta_k) + \left(\tilde{\theta}_k - \theta_k\right)^2\frac{\tilde{\eta}^{(3)}(\overline{\theta}_k)}{2},$$

where $\overline{\theta}_k \in \left(\min(\theta_k, \tilde{\theta}_k), \max(\theta_k, \tilde{\theta}_k)\right)$. By noticing that the n-th derivative of $\theta \mapsto \mathbf{a}(\theta)$ satisfies

$$\sup_{\theta \in (-\pi, \pi]} \left\| \mathbf{a}^{(n)}(\theta) \right\|_2 \sim M^n$$

as $M \to \infty$, we readily have

$$\frac{\tilde{\eta}^{(3)}(\overline{\theta}_k)}{M^3} \leq$$
$$C \frac{\left\| \mathbf{a}^{(2)}(\overline{\theta}_k) \right\|_2 \left\| \mathbf{a}'(\overline{\theta}_k) \right\|_2 + \left\| \mathbf{a}^{(3)}(\overline{\theta}_k) \right\|_2 \left\| \mathbf{a}(\overline{\theta}_k) \right\|_2}{M^3} \left\| \tilde{\mathbf{\Pi}} \right\|_2,$$

for some constant $C > 0$, and thus $M^{-3} \tilde{\eta}^{(3)}(\overline{\theta}_k) = O(1)$ a.s. as $M \to \infty$. Moreover, from (3), we have a.s. as $M \to \infty$,

$$\frac{\tilde{\eta}^{(2)}(\theta_k)}{M^2} = 2 \frac{\mathbf{a}'(\theta_k)^* \mathbf{\Pi} \mathbf{a}'(\theta_k)}{M^2} + o(1).$$

This implies that

$$M^{3/2} \left(\tilde{\theta}_k - \theta_k \right) = -\sqrt{M} \frac{\mathrm{Re}\left(\frac{\mathbf{a}'(\theta_k)^*}{M} \tilde{\mathbf{\Pi}} \mathbf{a}(\theta_k) \right)}{\frac{\mathbf{a}'(\theta_k)^*}{M} \mathbf{\Pi} \frac{\mathbf{a}'(\theta_k)}{M}} + o_{\mathbb{P}}(1). \tag{33}$$

The problem thus reduces to compute the asymptotic distribution of $\sqrt{M} \mathrm{Re}\left(\frac{\mathbf{a}'(\theta_k)^*}{M} \tilde{\mathbf{\Pi}} \mathbf{a}(\theta_k) \right)$. For that purpose, we express, as in the proof of Corollary 1, the projector $\tilde{\mathbf{\Pi}}$ in terms of the rectified S.C.M. $\tilde{\mathbf{R}}$ by means of a contour integral which is easier to study. At this point, we separate the conditional and unconditional models.

Conditional model. If $\mathcal{C}$ is a circle counterclockwise oriented, enclosing σ^2 and leaving outside $\gamma_1 + \sigma^2, \ldots, \gamma_K + \sigma^2$ (recall that $\gamma_1, \ldots, \gamma_K$ are the diagonal entries of $\mathbf{\Gamma}$), then we have a.s. for sufficient large M

$$\tilde{\mathbf{\Pi}} - \mathbf{\Pi} = \frac{1}{2\pi} \oint_{\mathcal{C}} \left(\left(\tilde{\mathbf{R}} - z\mathbf{I} \right)^{-1} - \left(\overline{\mathbf{R}} - z\mathbf{I} \right)^{-1} \right) \mathrm{d}z,$$
$$= \frac{1}{2\pi \mathrm{i}} \oint_{\mathcal{C}} \left(\tilde{\mathbf{R}} - z\mathbf{I} \right)^{-1} \tilde{\mathbf{\Delta}} \left(\overline{\mathbf{R}} - z\mathbf{I} \right)^{-1} \mathrm{d}z,$$

where we have denoted $\tilde{\mathbf{\Delta}} = \tilde{\mathbf{R}} - \overline{\mathbf{R}}$, with

$$\overline{\mathbf{R}} = \mathbf{A} \mathbf{\Gamma} \mathbf{A}^* + \sigma^2 \mathbf{I}$$

being the Toeplitz matrix obtained by replacing $\frac{\mathbf{SS}^*}{N}$ with $\mathbf{\Gamma}$ in the true covariance matrix $\mathbf{R} = \mathbf{A} \frac{\mathbf{SS}^*}{N} \mathbf{A}^* + \sigma^2 \mathbf{I}$. Similarly, developing further to the second order, we get

$$\tilde{\mathbf{\Pi}} - \mathbf{\Pi} = \mathbf{\Psi}_1 + \mathbf{\Psi}_2, \tag{34}$$

where

$$\mathbf{\Psi}_1 = \frac{1}{2\pi \mathrm{i}} \oint_{\mathcal{C}} \left(\overline{\mathbf{R}} - z\mathbf{I} \right)^{-1} \tilde{\mathbf{\Delta}} \left(\overline{\mathbf{R}} - z\mathbf{I} \right)^{-1} \mathrm{d}z$$

and

$$\mathbf{\Psi}_2 =$$
$$- \frac{1}{2\pi \mathrm{i}} \oint_{\mathcal{C}} \left(\tilde{\mathbf{R}} - z\mathbf{I} \right)^{-1} \tilde{\mathbf{\Delta}} \left(\overline{\mathbf{R}} - z\mathbf{I} \right)^{-1} \tilde{\mathbf{\Delta}} \left(\overline{\mathbf{R}} - z\mathbf{I} \right)^{-1} \mathrm{d}z.$$

Rewriting the first term of (34) thanks to residue theorem leads to

$$\mathbf{\Psi}_1 = -\mathbf{\Pi} \tilde{\mathbf{\Delta}} (\mathbf{A} \mathbf{\Gamma} \mathbf{A}^*)^\sharp - (\mathbf{A} \mathbf{\Gamma} \mathbf{A}^*)^\sharp \tilde{\mathbf{\Delta}} \mathbf{\Pi}.$$

Concerning the second term, we have the upper bound

$$\| \mathbf{\Psi}_2 \|_2 \leq$$
$$C \left\| \tilde{\mathbf{\Delta}} \right\|_2^2 \sup_{z \in \mathcal{C}} \left\| \left(\tilde{\mathbf{R}} - z\mathbf{I} \right)^{-1} \right\|_2 \sup_{z \in \mathcal{C}} \left\| \left(\mathbf{R} - z\mathbf{I} \right)^{-1} \right\|_2^2,$$

which indicates that $\mathbf{\Psi}_2$ will be negligible compared to $\mathbf{\Psi}_1$ in the remainder of the analysis. Indeed, from the proof of Theorem 1, using (24), (25) and (27), (28), we obtain

$$\left\| \tilde{\mathbf{\Delta}} \right\|_2^2 \leq C \left(\sup_{(-\pi, \pi]} |h_M(\nu)|^2 + \sup_{(-\pi, \pi]} |g_M(\nu)|^2 \right).$$

From the rate of convergence specified in Assumption 1, it is obvious that

$$\sup_{(-\pi, \pi]} |g_M(\nu)|^2 = o\left(\frac{1}{\sqrt{M}} \right).$$

Recalling (30), we have a.s.

$$\sup_{\nu \in (-\pi, \pi]} |h_M(\nu)|^2 \leq C \max_{\nu \in \mathcal{V}_M} |h_M(\nu)|^2 + o\left(\frac{1}{\sqrt{M}} \right).$$

Using Lemma 2, we obtain that

$$\mathbb{P}\left(\sqrt{N} \max_{\nu \in \mathcal{V}_M} |h_M(\nu)|^2 \geq \epsilon \right) = \mathcal{O}\left(\frac{1}{M^\ell} \right)$$

for all $\ell \in \mathbb{N}^*$, and consequently we deduce that a.s.

$$\sup_{\nu \in (-\pi, \pi]} |h_M(\nu)|^2 = o\left(\frac{1}{\sqrt{M}} \right).$$

In view of (21) and (22), we have proved that a.s.,

$$\| \mathbf{\Psi}_2 \|_2 = o\left(\frac{1}{\sqrt{M}} \right),$$

and therefore

$$\sqrt{M} \mathrm{Re}\left(\frac{\mathbf{a}'(\theta_k)^*}{M} \tilde{\mathbf{\Pi}} \mathbf{a}(\theta_k) \right)$$
$$= \sqrt{M} \mathrm{Re}\left(\mathbf{d}_1^* \tilde{\mathbf{\Delta}} \mathbf{d}_2 \right) + o_{\mathbb{P}}(1)$$
$$= \sqrt{M} \mathrm{Re}\left(\mathbf{d}_1^* \left(\tilde{\mathbf{R}} - \mathbb{E}[\tilde{\mathbf{R}}] \right) \mathbf{d}_2 \right)$$
$$+ \sqrt{M} \mathrm{Re}\left(\mathbf{d}_1^* \left(\mathbb{E}[\tilde{\mathbf{R}}] - \overline{\mathbf{R}} \right) \mathbf{d}_2 \right) + o_{\mathbb{P}}(1) \tag{35}$$

where $\mathbf{d}_1 = \mathbf{\Pi} \frac{\mathbf{a}'(\theta_k)}{M}$ and $\mathbf{d}_2 = (\mathbf{A} \mathbf{\Gamma} \mathbf{A}^*)^\sharp \mathbf{a}(\theta_k)$. Some straightforward algebra leads to

$$\mathrm{Re}\left(\mathbf{d}_1^* \left(\mathbb{E}[\tilde{\mathbf{R}}] - \overline{\mathbf{R}} \right) \mathbf{d}_2 \right) = \mathrm{Tr}\left(\mathbf{A} \left(\frac{\mathbf{SS}^*}{N} - \mathbf{\Gamma} \right) \mathbf{A}^* \mathbf{T}_k \right)$$

and

$$\mathrm{Re}\left(\mathbf{d}_1^* \left(\tilde{\mathbf{R}} - \mathbb{E}[\tilde{\mathbf{R}}] \right) \mathbf{d}_2 \right)$$
$$= \mathrm{Tr}\left(\left(\hat{\mathbf{R}} - \mathbf{R} \right) \mathbf{T}_k \right)$$
$$= \frac{1}{N} \sum_{n=1}^{N} \left(\mathbf{y}_n^* \mathbf{T}_k \mathbf{y}_n - \mathbb{E}\left[\mathbf{y}_n^* \mathbf{T}_k \mathbf{y}_n \right] \right),$$

where $\mathbf{T}_k$ is defined in (10). We remark that this last expression coincides with the mean of independent random

variables, whose asymptotic distribution is standard to study. Let $Z_{n,M} = \mathbf{y}_n^* \mathbf{T}_k \mathbf{y}_n - \mathbb{E}\left[\mathbf{y}_n^* \mathbf{T}_k \mathbf{y}_n\right]$ and define

$$s_M^2 = \sum_{n=1}^N \mathbb{V}[Z_{n,M}]$$

$$= N\left(\sigma^4 \|\mathbf{T}_k\|_F^2 + 2\sigma^2 \left\|\mathbf{T}_k \mathbf{A} \frac{\mathbf{S}}{\sqrt{N}}\right\|_F^2\right).$$

In view of Lemma 3 in Appendix C-A below, $s_M^2 \sim CM$ as $M \to \infty$. Moreover, $\mathbb{E}\left|Z_{n,M}\right|^4 \le CM^4$ and using Lemma 2, we easily obtain, for all $\epsilon > 0$,

$$\frac{1}{s_M^2} \sum_{n=1}^N \mathbb{E}\left[Z_{n,M}^2 \mathbb{1}_{|Z_{n,M}| \ge s_M \epsilon}\right]$$

$$\le \frac{C}{M} \sum_{n=1}^N \left(\mathbb{E}|Z_{n,M}|^4\right)^{1/2} \mathbb{P}\left(|Z_{n,M}| \ge s_M \epsilon\right)^{1/2}$$

$$\le CM \sum_{n=1}^N \mathbb{P}\left(|Z_{n,M}| \ge s_M \epsilon\right)^{1/2}$$

$$\xrightarrow[M\to\infty]{} 0.$$

Consequently, the Lindeberg condition to apply the Central Limit Theorem for triangular arrays [28] is satisfied and it follows that

$$\frac{\sqrt{M}\mathrm{Re}\left(\mathbf{d}_1^* \left(\tilde{\mathbf{R}} - \mathbb{E}[\tilde{\mathbf{R}}]\right) \mathbf{d}_2\right)}{\sqrt{\sigma^2 c \left(\sigma^2 \|\mathbf{T}_k\|_F^2 + 2\left\|\mathbf{T}_k \mathbf{A} \frac{\mathbf{S}}{\sqrt{N}}\right\|_F^2\right)}} \xrightarrow[M\to\infty]{\mathcal{D}} \mathcal{N}_{\mathbb{R}}(0,1).$$

(36)

Going back to (35) and (33), and taking into account the fact that the variance (9) has a positive limit as $M \to \infty$ (Lemma 3), we eventually obtain (12).

Unconditional model. Under Assumption 2, the proof of Theorem 4 follows very closely the steps for the conditional case. Indeed, similar to (34), we can show that

$$\tilde{\mathbf{\Pi}} - \mathbf{\Pi} = -\mathbf{\Pi}\tilde{\mathbf{\Delta}}\left(\mathbf{A}\mathbf{\Gamma}\mathbf{A}^*\right)^\sharp - \left(\mathbf{A}\mathbf{\Gamma}\mathbf{A}^*\right)^\sharp \tilde{\mathbf{\Delta}}\mathbf{\Pi} + \mathbf{\Psi}_2,$$

where $\tilde{\mathbf{\Delta}} = \tilde{\mathbf{R}} - \mathbf{R}$ and $\|\mathbf{\Psi}_2\|_2 = o\left(M^{-1/2}\right)$ a.s. as $M \to \infty$. Consequently,

$$\sqrt{M}\mathrm{Re}\left(\frac{\mathbf{a}'(\theta_k)^*}{M}\tilde{\mathbf{\Pi}}\mathbf{a}(\theta_k)\right) = -\frac{\sqrt{M}}{N}\sum_{n=1}^N Z_{n,N}$$

where $Z_{n,N} = \mathbf{y}_n^* \mathbf{T}_k \mathbf{y}_n - \mathbb{E}[\mathbf{y}_n^* \mathbf{T}_k \mathbf{y}_n]$ with $\mathbf{T}_k$ defined in (10) (recall that in this case, $\mathbf{y}_1, \ldots, \mathbf{y}_N$ are the i.i.d. $\mathcal{N}_{\mathbb{C}^M}(\mathbf{0}, \mathbf{R})$ columns of $\mathbf{Y}$). We immediately have $\mathbb{V}[Z_{n,M}] = \left\|\mathbf{R}^{1/2}\mathbf{T}_k\mathbf{R}^{1/2}\right\|_F^2$, and the Lindeberg condition to apply the CLT is verified as previously, using Lemma 2, and the result of Theorem 4 follows.

REFERENCES

[1] P. Vallet and P. Loubaton, "Toeplitz rectification and DoA estimation with MUSIC," in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2014, pp. 2237–2241.
[2] H. Krim and M. Viberg, "Two decades of array signal processing research: the parametric approach," *IEEE Signal Process. Mag.*, vol. 13, no. 4, pp. 67–94, 1996.
[3] E. Larsson, O. Edfors, F. Tufvesson, and T. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag*, vol. 52, no. 2, pp. 186–195, 2014.
[4] P. Stoica and A. Nehorai, "MUSIC, Maximum Likelihood, and Cramer-Rao bound," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, no. 5, pp. 720–741, 1989.
[5] X. Mestre and M. Lagunas, "Modified subspace algorithms for DoA estimation with large arrays," *IEEE Trans. Signal Process.*, vol. 56, no. 2, pp. 598–614, 2008.
[6] 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, Feb. 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] ——, "A subspace estimator for fixed rank perturbations of large random matrices," *J. Multivariate Anal.*, vol. 114, pp. 427–447, 2013.
[9] P. Vallet, X. Mestre, and P. Loubaton, "Performance Analysis of an Improved MUSIC DoA Estimator," *IEEE Trans. Signal Process.*, vol. 63, no. 23, pp. 6407–6422, Dec 2015.
[10] J. Vinogradova, R. Couillet, and W. Hachem, "Statistical inference in large antenna arrays under unknown noise pattern," *IEEE Trans. Signal Process.*, vol. 61, no. 22, pp. 5633–5645, 2013.
[11] P. Vallet, X. Mestre, P. Loubaton, and R. Couillet, "Asymptotic analysis of Beamspace-MUSIC in the context of large arrays," in *Proc. SAM'14*, 2014, pp. 469–472.
[12] R. Couillet, "Robust spiked random matrices and a robust G-MUSIC estimator," *J. Multivariate Anal.*, vol. 140, pp. 139–161, 2015.
[13] G. T. Pham, P. Loubaton, and P. Vallet, "Performance Analysis of Spatial Smoothing Schemes in the Context of Large Arrays," *IEEE Trans. Signal Process.*, vol. 64, no. 1, pp. 160–172, 2016.
[14] O. Ledoit and M. Wolf, "A well-conditioned estimator for large-dimensional covariance matrices," *J. Multivariate Anal.*, vol. 88, no. 2, pp. 365–411, 2004.
[15] P. J. Bickel and E. Levina, "Regularized estimation of large covariance matrices," *Ann. Statist.*, vol. 36, no. 1, pp. 199–227, 2008.
[16] T. T. Cai, Z. Ren, and H. H. Zhou, "Optimal rates of convergence for estimating Toeplitz covariance matrices," *Probab. Theory Related Fields*, vol. 156, no. 1-2, pp. 101–143, 2013.
[17] J. Vinogradova, R. Couillet, and W. Hachem, "Estimation of Toeplitz covariance matrices in large dimensional regime with application to source detection," *IEEE Trans. Signal Process.*, vol. 63, no. 18, pp. 4903–4913, 2015.
[18] J. Cadzow, "Signal enhancement using canonical projection operators," in *Proc. IEEE ICASSP'87*, vol. 12, 1987, pp. 673–676.
[19] P. Forster, "Generalized rectification of cross spectral matrices for arrays of arbitrary geometry," *IEEE Trans. Signal Process.*, vol. 49, no. 5, pp. 972–978, 2001.
[20] C. Heyde, "An iterated logarithm result for autocorrelations of a stationary linear process," *Ann. Prob.*, pp. 328–332, 1974.
[21] S. Geman, "A limit theorem for the norm of random matrices," *Ann. Prob.*, pp. 252–261, 1980.
[22] F. Vincent, O. Besson, and E. Chaumette, "Approximate maximum likelihood estimation of two closely spaced sources," *Signal Process.*, vol. 97, pp. 83–90, 2014.
[23] R. Roy, A. Paulraj, and T. Kailath, "Comparative performance of ESPRIT and MUSIC for direction-of-arrival estimation," in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 12. IEEE, 1987, pp. 2344–2347.
[24] R. Vershynin, "Introduction to the non-asymptotic analysis of random matrices," in *Compressed Sensing*, Y. C. Eldar and G. Kutyniok, Eds. Cambridge University Press, 2012, pp. 210–268.
[25] R. H. Y. C. Johnson and R. A. Horn, "Topics in matrix analysis," *Cambridge University*, 1991.
[26] F. Benaych-Georges and R. R. Nadakuditi, "The singular values and vectors of low rank perturbations of large rectangular random matrices," *J. Multivariate Anal.*, vol. 111, no. 0, pp. 120–135, 2012.
[27] P. Loubaton and P. Vallet, "Almost Sure Localization of the Eigenvalues in a Gaussian Information Plus Noise Model. Application to the Spiked Models." *Electron. J. Probab.*, vol. 16, pp. 1934–1959, 2011.
[28] P. Billingsley, *Probability and measure*, 3rd ed. John Wiley & Sons Ltd., 1995.