ASYMPTOTIC CONSISTENCY OF ESPRIT DOA ESTIMATION WITH LARGE ARRAYS

Wei Yang, Zhengyu Wang, Zenan Ling, Hafiz Tiomoko Ali, Robert C. Qiu, Zhenyu Liao†*

Huazhong University of Science and Technology, Wuhan, China

ABSTRACT

In this paper, we show that the popular ESPRIT DoA estimation method remains consistent, in the *large array and limited sample* regime where the number samples/snapshots T and the number of sensors N are both large and comparable. Albeit that the sample covariance matrix is a poor estimator of the population covariance in this setting, ESPRIT still provides consistent estimates of the DoA so long as the corresponding signal-to-noise ratio is *above* a certain threshold. Simulations are provided to elaborate the theoretical analysis.

***Index Terms*—** DoA estimation, ESPRIT, random matrix theory, sample covariance matrix, subspace method

1. INTRODUCTION

Commonly used subspace Direction-of-Arrival (DoA) methods such as MUSIC [1] and ESPRIT [2] are based on the idea to retrieve DoA information from the so-called "signal subspace" of sample covariance matrix of received signals. Consider the sample covariance matrix (SCM) computed from T snapshots on an array composed of N sensors, it is now well known that in the *large array and limited sample regime* when T is *not* much larger than N, the SCM is a poor estimator of the population covariance in a spectral sense, see, for example [3, 4] and Section 2.2 below. As a consequence, one should not expect, a priori, that the aforementioned subspace methods could provide consistent estimate of the true DoAs in this setting where N, T are both large and comparable.

With the rapid development of random matrix theory (RMT) during the past decade, many methods in statistics, signal processing, and machine learning have been revisited in the large-dimensional setting, resulting in novel insights and improved algorithms that are more adapted to large-dimensional data [5]. In the case of subspace DoA methods, it has been shown in [6] that the popular MUSIC method, despite the spectral inconsistency of the SCM in the large N, T regime, still provides DoA consistent estimates, in widely spaced DoA scenario when the DoAs $\theta_1 \ldots, \theta_K$ of interest are considered fixed as $N, T \to \infty$. In the case of closely spaced sources (that models the scenario of, e.g., DoAs placed within the order of a beam-width), on the other hand, it has been established that the classical MUSIC approach is bound to fail, and improved estimators such G-MUSIC [7] should be used instead [6].

Here, we propose to analyze the equally popular DoA subspace method of ESPRIT [2] (to be recalled in Section 2.1 below) in the *large array and limited sample* regime. While it is observed empirically that the ESPRIT method outperforms MUSIC in this regime in some cases [8] and not in some other cases [9], its theoretical analysis remains an open problem [10], due to its mathematically involved form than, e.g., the MUSIC method.

In this paper, we show that despite the overall poor performance of the SCM in the large-dimensional setting as $N, T \to \infty$ with $N/T \to c \in (0, \infty)$, the standard ESPRIT method still produces *consistent* estimates of the true DoAs, so long as (i) the DoAs are widely spaced (see Assumption 2 below and the discussion thereafter); and (ii) the sources are uncorrelated (Assumption 3); and (iii) the signal-to-noise ratio (SNR) is above a certain threshold (Assumption 4).

We further provide, in Section 4, numerical simulations on the case where one or more of the above assumptions are violated. These finite-dimensional experiments suggest that despite its asymptotic consistency, ESPRIT yields estimates with larger variance than MUSIC and the G-MUSIC approach. It would thus be of future interest to extend the RMT analysis by relaxing one or more of the above assumptions, so as to propose improved ESPRIT-type DoA methods.

Notations. We denote scalars by lowercase letters, vectors by bold lowercase, and matrices by bold uppercase. We use $\mathbb{R}$ for the set of real numbers and $\mathbb{C}$ the set of complex numbers. For a complex matrix $\mathbf{A}$, we denote $\mathbf{A}^\mathsf{T}$, $\mathbf{A}^\mathsf{H}$ the transpose and conjugate transpose of $\mathbf{A}$. We use $\|\cdot\|$ for the Euclidean norm for vectors and the spectral/operator norm for matrices. We use $\mathcal{N}(m, \sigma^2)$ for the real Gaussian distribution with mean m and variance σ^2, and say z follows a complex circular Gaussian distribution and denote $z \sim \mathcal{CN}(m, \sigma^2)$ if $z = x + \imath y$ with independent x, y such that $x \sim \mathcal{N}(\Re[m], \sigma^2/2)$ and $y \sim \mathcal{N}(\Im[m], \sigma^2/2)$.

*Work done as an independent researcher, email: tio.hafiz@gmail.com.

†This work is partially supported by the National Natural Science Foundation of China (via fund NSFC-12141107 and NSFC-62206101), the Fundamental Research Funds for the Central Universities of China (2021XXJS110), the Interdisciplinary Research Program of HUST (2023JCYJ012), the Key Research and Development Program of Hubei (2021BAA037) and of Guangxi (GuiKe-AB21196034).

2. SYSTEM MODEL AND PRELIMINARIES

In this paper, we consider a Unitary Linear Array (ULA) of N sensors that receive K narrow-band and far-field source signals with DoA $\theta_1, \ldots, \theta_K$. The received signal at time $t = 1, \ldots, T$ is thus given by

$$\mathbf{x}(t) = \sum_{k=1}^{K} \mathbf{a}(\theta_k) s_k(t) + \mathbf{n}(t) \in \mathbb{C}^N, \tag{1}$$

with complex signal $s_k(t) \in \mathbb{C}$, $\mathbf{a}(\theta_k) \in \mathbb{C}^N$ the steering vector of source $k \in \{1, \ldots, K\}$ at DoA θ_k given by[1]

$$\mathbf{a}(\theta_k) = [1,\ e^{\imath\theta_k}, \ldots, e^{\imath(N-1)\theta_k}]^{\mathsf{T}}/\sqrt{N} \in \mathbb{C}^N, \tag{2}$$

and complex circular Gaussian white noise $\mathbf{n}(t) \in \mathbb{C}^N$ having i.i.d. $\mathcal{CN}(0,1)$ entries for all t. This model can be rewritten in matrix form as $\mathbf{X} = \mathbf{AS} + \mathbf{N}$, with $\mathbf{X} = [\mathbf{x}(1), \ldots, \mathbf{x}(T)] \in \mathbb{C}^{N \times T}$ the received signal matrix, $\mathbf{A} = [\mathbf{a}(\theta_1), \ldots, \mathbf{a}(\theta_k)] \in \mathbb{C}^{N \times K}$ the matrix of steering vectors, $\mathbf{S} = [\mathbf{s}(1), \ldots, \mathbf{s}(T)] \in \mathbb{C}^{K \times T}$ the source signal matrix, and noise matrix $\mathbf{N} = [\mathbf{n}(1), \ldots, \mathbf{n}(T)] \in \mathbb{C}^{N \times T}$.

We positive ourselves under the following assumptions.

Assumption 1 (Large array and limited sample regime)**.** *As $N, T \to \infty$, we have $N/T \to c \in (0, \infty)$ with K fixed.*

Assumption 2 (Widely spaced DoAs)**.** *The DoAs $\theta_1, \ldots, \theta_K$ are fixed as $N \to \infty$.*

In practice, the scenario considered in Assumption 2 corresponds to the case where the DoAs have angular separation much larger than a beam-width $2\pi/N$. In particular, we have that $\mathbf{A}^{\mathsf{H}}\mathbf{A} \to \mathbf{I}_K$ as $N, T \to \infty$ under Assumption 2.

Assumption 3 (Uncorrelated Gaussian source)**.** *The source signal matrix satisfies $\mathbf{S} = \mathbf{P}^{1/2}\tilde{\mathbf{Z}}$ for some $\tilde{\mathbf{Z}} \in \mathbb{C}^{K \times N}$ of i.i.d. $\mathcal{CN}(0,1)$ entries and deterministic diagonal $\mathbf{P} \in \mathbb{R}^{K \times K}$ having positive diagonal entries.*

Assumption 3 is commonly considered in the literature and is also known as the *unconditional model* [10]. Under Assumption 3, the received signal $\mathbf{X}$ equals, in distribution, to

$$\mathbf{X} \overset{d}{=} (\mathbf{I}_N + \mathbf{APA}^{\mathsf{H}})^{1/2}\mathbf{Z} \equiv \mathbf{C}^{1/2}\mathbf{Z},\ \mathbf{C} \equiv \mathbf{I}_N + \mathbf{APA}^{\mathsf{H}}, \tag{3}$$

for random matrix $\mathbf{Z} \in \mathbb{C}^{N \times T}$ having i.i.d. $\mathcal{CN}(0,1)$ entries.

2.1. ESPRIT algorithm

In the following, we recall the ESPRIT estimation procedure.

1. Define two selection matrices $\mathbf{J}_1, \mathbf{J}_2 \in \mathbb{R}^{n \times N}$ that both select n among N rows of the received signal matrix $\mathbf{X} \in \mathbb{C}^{N \times T}$ with "distance" $\Delta \geq 1$ in such a way that

   $\mathbf{J}_1^{\mathsf{T}} = [\mathbf{e}_\ell,\ \ldots,\ \mathbf{e}_{n+\ell-1}]$, $\mathbf{J}_2^{\mathsf{T}} = [\mathbf{e}_{\ell+\Delta},\ \ldots,\ \mathbf{e}_{n+\ell+\Delta-1}]$,

   for $\mathbf{e}_i$ the canonical vector of $\mathbb{R}^N$ with $[\mathbf{e}_i]_j = \delta_{ij}$;
2. compute the SCM as $\hat{\mathbf{C}} = \frac{1}{T}\mathbf{X}\mathbf{X}^{\mathsf{H}}$ to retrieve $\hat{\mathbf{U}}_K = [\hat{\mathbf{u}}_1, \ldots, \hat{\mathbf{u}}_K] \in \mathbb{C}^{N \times K}$ the *signal subspace* composed of the top K eigenvectors $\hat{\mathbf{u}}_1, \ldots, \hat{\mathbf{u}}_K$ associated to the largest K eigenvalues of $\hat{\mathbf{C}}$;
3. compute $\mathbf{\Phi} = (\mathbf{J}_1\hat{\mathbf{U}}_K)^{\dagger}\mathbf{J}_2\hat{\mathbf{U}}_K \in \mathbb{C}^{K \times K}$, with $(\mathbf{A})^{\dagger}$ the Moore–Penrose pseudoinverse of $\mathbf{A}$;
4. return the estimate of the DoA $\hat{\theta}_k$ from the *angle* of $\lambda_k(\mathbf{\Phi})$, the kth (complex) eigenvalue of $\mathbf{\Phi}$.

2.2. Spectral inconsistency of large-dimensional SCM

In the following, we recall a few results from large-dimensional RMT that provide precise eigenspectral characterization of SCM in the large N, T regime.

Assumption 4 (Subspace separation)**.** *Under the settings and notations of Assumption 1–3, we assume, as $N \to \infty$, that the top eigenvalues $\lambda_k(\mathbf{APA}^{\mathsf{H}})$ of $\mathbf{APA}^{\mathsf{H}} \in \mathbb{C}^{N \times N}$ satisfy*

$$\lambda_1(\mathbf{APA}^{\mathsf{H}}) \to \rho_1 > \ldots > \lambda_K(\mathbf{APA}^{\mathsf{H}}) \to \rho_K > \sqrt{c}. \tag{4}$$

Note that in the widely spaced DoA scenario in Assumption 2, we have $\mathbf{A}^{\mathsf{H}}\mathbf{A} \to \mathbf{I}_K$ and $\lambda_k(\mathbf{APA}^{\mathsf{H}}) = \lambda_k(\mathbf{A}^{\mathsf{H}}\mathbf{AP}) \to [\mathbf{P}]_{kk}$ by Sylvester's determinant identity. As such, ρ_k is, in this case, the limit of the kth diagonal entries of $\mathbf{P}$, and refers to the signal-to-noise ratio (SNR) of the source k.

We have the following result, due to a sequence of previous efforts [11, 12, 13, 14]. See also [4, Chapter 2].

Theorem 1 (Spectral characterization of large-dimensional SCM)**.** *Under the settings and notations of Assumption 1–3, we have, for $\mathbf{X} \in \mathbb{C}^{N \times T}$ defined* (3) *and as $N, T \to \infty$ with $N/T \to c \in (0, \infty)$ that, with probability one, the eigenvalue distribution of the SCM $\hat{\mathbf{C}} = \mathbf{X}\mathbf{X}^{\mathsf{H}}/T$ converges weakly to the popular Marčenko-Pastur law*

$$\mu(dx) = (1 + c^{-1})^{+}\delta_0(x) + \frac{\sqrt{(x - E_-)^{+}(E_+ - x)^{+}}\, dx}{2\pi c x},$$

with $E_\pm = (1 \pm \sqrt{c})^2$ and $(x)^{+} = \max(x, 0)$. Moreover, let Assumption 4 hold and denote $\hat{\lambda}_1 \geq \ldots \geq \hat{\lambda}_N$ the ordered eigenvalues of $\hat{\mathbf{C}}$, we have

$$\hat{\lambda}_i \to \begin{cases} \bar{\lambda}_i = 1 + \rho_i + c\frac{1+\rho_i}{\rho_i} > E_+, & i \leq K \\ E_+ = (1 + \sqrt{c})^2, & i > K \end{cases}; \tag{5}$$

almost surely as $N \to \infty$.

Theorem 1 states that for N, T both large and comparable, the eigenvalues of the SCM $\hat{\mathbf{C}}$, instead of being close to those of its population counterpart $\mathbf{C} = \mathbf{I}_N + \mathbf{APA}^{\mathsf{H}}$ in (3), spread out on the interval $[E_-, E_+]$ of length $4\sqrt{c} \gg 0$. Moreover, under the additional Assumption 4, it is known that the largest eigenvalues of $\hat{\mathbf{C}}$ (that are due to the "signal" $\mathbf{APA}^{\mathsf{H}}$) are guaranteed to "separate" from those due to the random white

[1] The normalization by $\sqrt{N}$ is made so that $\mathbf{a}(\theta_k)$ is of unit norm. Here, we use θ_k for the DoA in the Fourier space as in [6], which is related to the "physical" angle ϕ_k of the source wave via $\theta_k = \frac{2\pi d}{\lambda_0}\sin(\phi_k)$.

noise. However, even in this case, the (empirical) eigenvalues $\hat{\lambda}_i$ of $\hat{\mathbf{C}}$ are significantly larger than the population ones $(1+\rho_i)$, by a quantity that is proportional to the dimension ratio $c = \lim N/T$.

Despite this eigenspectral inconsistency in Theorem 1, we will show, in the sequel, that the classical ESPRIT method still provides DoA consistent estimates for N, T large.

3. MAIN RESULTS

In this section, we perform asymptotic analysis of the ESPRIT method as $N, T, n \to \infty$. Assume that the matrix $\mathbf{\Phi}_1 \equiv \hat{\mathbf{U}}_K^{\mathsf{H}} \mathbf{J}_1^{\mathsf{H}} \mathbf{J}_1 \hat{\mathbf{U}}_K \in \mathbb{C}^{K \times K}$ is invertible (which can be shown to happen with probability one in the large $n, N, T \to \infty$ limit), one has, for $\mathbf{\Phi}$ defined in Section 2.1 that

$$\mathbf{\Phi} = (\hat{\mathbf{U}}_K^{\mathsf{H}} \mathbf{J}_1^{\mathsf{H}} \mathbf{J}_1 \hat{\mathbf{U}}_K)^{-1} \hat{\mathbf{U}}_K^{\mathsf{H}} \mathbf{J}_1^{\mathsf{T}} \mathbf{J}_2 \hat{\mathbf{U}}_K \equiv \mathbf{\Phi}_1^{-1} \mathbf{\Phi}_2, \tag{6}$$

which needs be evaluated to assess the asymptotic performance of ESPRIT.

Let us start with the $(1,1)$ diagonal entry of the (in general non-Hermitian) matrix $\mathbf{\Phi}_2$ by introducing the resolvent as

$$\mathbf{Q}(z) \equiv \left(\mathbf{Z}\mathbf{Z}^{\mathsf{H}}/T - z\mathbf{I}_N \right)^{-1}, \tag{7}$$

for $z \in \mathbb{C}$ not eigenvalue of $\mathbf{Z}\mathbf{Z}^{\mathsf{H}}/T$. Write

$$\begin{aligned}
[\mathbf{\Phi}_2]_{11} &= \hat{\mathbf{u}}_1^{\mathsf{H}} \mathbf{J}_1^{\mathsf{T}} \mathbf{J}_2 \hat{\mathbf{u}}_1 = \sum_{i=\ell}^{n+\ell-1} \mathbf{e}_{i+\Delta}^{\mathsf{T}} \hat{\mathbf{u}}_1 \hat{\mathbf{u}}_1^{\mathsf{H}} \mathbf{e}_i \\
&= -\frac{1}{2\pi\imath} \sum_{i=\ell}^{n+\ell-1} \oint_{\Gamma_1} \mathbf{e}_{i+\Delta}^{\mathsf{T}} (\hat{\mathbf{C}} - z\mathbf{I}_N)^{-1} \mathbf{e}_i \, dz \\
&= -\frac{1}{2\pi\imath} \sum_{i=\ell}^{n+\ell-1} \oint_{\Gamma_1} \mathbf{e}_{i+\Delta}^{\mathsf{T}} \mathbf{C}^{-\frac{1}{2}} \Big(\frac{1}{T} \mathbf{Z}\mathbf{Z}^{\mathsf{H}} - z\mathbf{I}_N \\
&\quad + z\mathbf{A}(\mathbf{P}^{-1} + \mathbf{A}^{\mathsf{H}}\mathbf{A})^{-1}\mathbf{A}^{\mathsf{H}} \Big)^{-1} \mathbf{C}^{-\frac{1}{2}} \mathbf{e}_i \, dz, \\
&= -\frac{1}{2\pi\imath} \sum_{i=\ell}^{n+\ell-1} \oint_{\Gamma_1} \mathbf{e}_{i+\Delta}^{\mathsf{T}} \mathbf{C}^{-\frac{1}{2}} \big(\mathbf{Q}(z) - z\mathbf{Q}(z)\mathbf{A} \\
&\quad \times \left(\mathbf{P}^{-1} + \mathbf{A}^{\mathsf{H}}\mathbf{A} + z\mathbf{A}^{\mathsf{H}}\mathbf{Q}(z)\mathbf{A} \right)^{-1} \mathbf{A}^{\mathsf{H}}\mathbf{Q}(z) \big) \mathbf{C}^{-\frac{1}{2}} \mathbf{e}_i \, dz, \\
&= \frac{1}{2\pi\imath} \sum_{i=\ell}^{n+\ell-1} \oint_{\Gamma_1} z \mathbf{e}_{i+\Delta}^{\mathsf{T}} \mathbf{C}^{-\frac{1}{2}} \mathbf{Q}(z)\mathbf{A} \\
&\quad \times \Big(\underbrace{\mathbf{P}^{-1} + \mathbf{A}^{\mathsf{H}}\mathbf{A} + z\mathbf{A}^{\mathsf{H}}\mathbf{Q}(z)\mathbf{A}}_{\mathbf{T}_1 \in \mathbb{C}^{K \times K}} \Big)^{-1} \mathbf{A}^{\mathsf{H}}\mathbf{Q}(z)\mathbf{C}^{-\frac{1}{2}} \mathbf{e}_i \, dz, \\
&= \frac{1}{2\pi\imath} \oint_{\Gamma_1} z \operatorname{tr}(\mathbf{T}_1^{-1} \underbrace{\mathbf{A}^{\mathsf{H}}\mathbf{Q}(z)\mathbf{C}^{-\frac{1}{2}} \mathbf{J}_1^{\mathsf{T}} \mathbf{J}_2 \mathbf{C}^{-\frac{1}{2}} \mathbf{Q}(z)\mathbf{A}}_{\mathbf{T}_2 \in \mathbb{C}^{K \times K}}) \, dz,
\end{aligned} \tag{8}$$

for Γ_1 a positively (counterclockwise) oriented path on the complex plane that circles *only* the first eigenvalue $\hat{\lambda}_1$ of $\hat{\mathbf{C}}$ associated with eigenvector $\hat{\mathbf{u}}_1$ (that exists almost surely according to Theorem 1 as $N, T \to \infty$), where we used the definition of $\mathbf{J}_1$ and $\mathbf{J}_2$ in the first line, the Cauchy's integral formula in the second, twice Woodbury matrix identity in the third and fifth line, and then the fact that $\mathbf{e}_{i+\Delta}^{\mathsf{T}} \mathbf{C}^{-\frac{1}{2}} \mathbf{Q}(z) \mathbf{C}^{-\frac{1}{2}} \mathbf{e}_i$ almost surely does *not* have a pole inside the surface circled by Γ_1 (again by Theorem 1).

Now, note that the entries of $\mathbf{T}_1, \mathbf{T}_2 \in \mathbb{C}^{K \times K}$ (i) are matrices of fixed size as $N, T \to \infty$; and (ii) contain bilinear forms of $\mathbf{Q}(z)$ and $\mathbf{Q}(z)\mathbf{B}\mathbf{Q}(z)$ for $\mathbf{B} \in \mathbb{C}^{N \times N}$ of bounded spectral norm. The following result provides asymptotic approximation of these quantities in the $N, T \to \infty$ limit.

Lemma 1 (First- and second-order deterministic equivalents for resolvent, [4, Theorem 2.4])**.** *For random matrix $\mathbf{Z} \in \mathbb{C}^{N \times T}$ having i.i.d. $\mathcal{CN}(0,1)$ entries, $z_1, z_2 \in \mathbb{C}$ not eigenvalue of $\mathbf{Z}\mathbf{Z}^{\mathsf{H}}/T$ and deterministic matrix $\mathbf{B} \in \mathbb{C}^{N \times N}$ of bounded Euclidean norm, then, the resolvent $\mathbf{Q}(z_1) = (\mathbf{Z}\mathbf{Z}^{\mathsf{H}}/T - z_1\mathbf{I}_N)^{-1}$ admits the following deterministic equivalents*

$$\begin{aligned}
\mathbf{Q}(z_1) &\leftrightarrow m(z_1) \cdot \mathbf{I}_N, \\
\mathbf{Q}(z_1)\mathbf{B}\mathbf{Q}(z_2) &\leftrightarrow m(z_1)m(z_2) \cdot \mathbf{B} + \eta(z_1, z_2) \frac{1}{T} \operatorname{tr}(\mathbf{B}) \cdot \mathbf{I}_N,
\end{aligned}$$

where $\mathbf{Q} \leftrightarrow \bar{\mathbf{Q}}$ denotes that $\mathbf{b}_1^{\mathsf{H}}(\mathbf{Q} - \bar{\mathbf{Q}})\mathbf{b}_2 \to 0$ almost surely as $N, T \to \infty$ for $\mathbf{b}_1, \mathbf{b}_2 \in \mathbb{C}^N$ of bounded norm, with $\eta(z_1, z_2) = \frac{m^2(z_1)m^2(z_2)}{(1+cm(z_1))(1+cm(z_2)) - cm(z_1)m(z_2)}$ and $m(z)$ the unique Stieltjes transform solution to the Marčenko-Pastur equation [11]

$$zcm^2(z) - (1 - c - z)m(z) + 1 = 0. \tag{9}$$

In particular, for $z_1 = z_2 = z$, we get $\eta(z, z) = \frac{m'(z)m^2(z)}{(1+cm(z))^2}$.

With the help of Lemma 1, we can approximate, in the large N, T limit, both $\mathbf{T}_1$ and $\mathbf{T}_2$ by their deterministic equivalents, and consequently the diagonal entries of $\mathbf{\Phi}_2$ as per (8). The off-diagonal entries of $\mathbf{\Phi}_2$ can be treated similarly. Taking $\mathbf{J}_2 = \mathbf{J}_1$ further allows one to retrieve the approximation of $\mathbf{\Phi}_1$. We summarize these results in the following result.

Proposition 1 (Asymptotic approximations of $\mathbf{\Phi}$)**.** *Let Assumption 1–4 hold and let n the size of selection matrices $\mathbf{J}_1, \mathbf{J}_2$ satisfy $n/N \to \tau \in (0,1)$, we have, for $\mathbf{\Phi}_1, \mathbf{\Phi}_2$ defined in* (6) *that*

$$\|\mathbf{\Phi}_1 - \bar{\mathbf{\Phi}}_1\| \to 0, \quad \|\mathbf{\Phi}_2 - \bar{\mathbf{\Phi}}_2\| \to 0, \quad \|\mathbf{\Phi} - \bar{\mathbf{\Phi}}\| \to 0, \tag{10}$$

almost surely as $n, N, T \to \infty$, with $\bar{\mathbf{\Phi}} = \bar{\mathbf{\Phi}}_1^{-1} \bar{\mathbf{\Phi}}_2$ and

$$\bar{\mathbf{\Phi}}_1 = \operatorname{diag} \left\{ \frac{n}{N} \frac{1 - c\rho_k^{-2}}{1 + c\rho_k^{-1}} + \frac{n}{N} \frac{1 + \rho_k^{-1}}{1 + \rho_k/c} \right\}_{k=1}^{K}, \tag{11}$$

$$\bar{\mathbf{\Phi}}_2 = \operatorname{diag} \left\{ \frac{n}{N} \frac{1 - c\rho_k^{-2}}{1 + c\rho_k^{-1}} \cdot e^{\imath \Delta \theta_k} \right\}_{k=1}^{K}. \tag{12}$$

It is interesting, yet not surprising, to note from Proposition 1 that due to the eigenspectral inconsistency of SCM in the large N, T regime (as, e.g., in Theorem 1), both $\mathbf{\Phi}_1$ and $\mathbf{\Phi}_2$ diverge from their initial design ($\mathbf{\Phi}_1$ by a multiplicative as well as an additive factor, and $\mathbf{\Phi}_2$ only by a multiplicative factor). These "bias" terms, as one may expect, vanish (i) in the *infinite sample* regime as $c = \lim N/T \to 0$ or (ii) in the *high SNR* regime as $\rho_k \to \infty$.

What is actually surprising and counterintuitive to us is that, despite these bias terms in the large N, T regime in $\mathbf{\Phi}_1, \mathbf{\Phi}_2$ (and consequently in $\mathbf{\Phi}$), due to the crucial fact that (i) these bias terms are *real* and (ii) $\mathbf{\Phi}_1, \mathbf{\Phi}_2$ are asymptotically diagonal, the angles of the complex eigenvalues of $\mathbf{\Phi}$ remain the same to their population counterpart, providing thus consistent estimates of the true DoAs. This is described in the following result, the proof of which follows from Proposition 1 in a straightforward manner.

Theorem 2 (High-dimensional consistency of ESPRIT)**.** *Under Assumption 1–4, let* n *the size of selection matrices* $\mathbf{J}_1, \mathbf{J}_2$ *satisfy* $n/N \to \tau \in (0, 1)$ *and let* $\hat{\theta}_k$ *denote the DoA estimate from the ESPRIT algorithm in Section 2.1, we have, for* $k \in \{1, \ldots, K\}$ *that* $\hat{\theta}_k - \theta_k \to 0$ *almost surely as* $N \to \infty$.

Theorem 2 states that for widely spaced DoAs and uncorrelated source, ESPRIT is still able to provide DoA consistent estimates in the *large array and limited sample* regime, just as the MUSIC or G-MUSIC approach. In the following section, we provide numerical simulations to testify our asymptotic analysis on *large but finite* N, T, as well as on scenarios when one or more of the above assumptions is *not* valid.

4. NUMERICAL SIMULATIONS

Figure 1 validate the asymptotic approximation in Proposition 1 for large but finite N, T, n and under Assumption 1–4. We observe the approximation $\mathbf{\Phi} \simeq \bar{\mathbf{\Phi}}$ becomes increasingly more accurate as N, T, n grow large together.

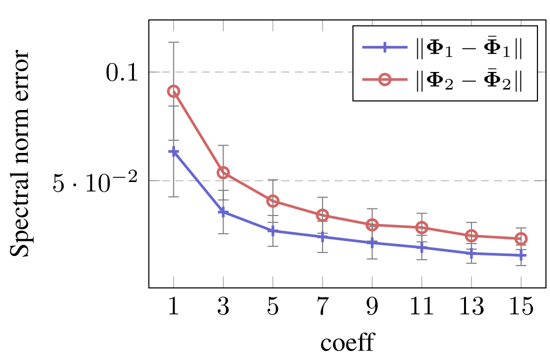

Fig. 1. Spectral norm errors of $\|\mathbf{\Phi}_1 - \bar{\mathbf{\Phi}}_1\|$ and $\|\mathbf{\Phi}_2 - \bar{\mathbf{\Phi}}_2\|$ as in Proposition 1, with $N/T = 1/4$, n/N fixed, and $N = \text{coeff} \times 128$. Results obtained over 100 independent runs.

Figure 2 compares the mean squared errors of different subspace DoA methods of MUSIC, G-MUSIC [7], and ESPRIT, to the corresponding Cramér–Rao bound [15, 16], for both uncorrelated and correlated sources (note that the latter case violates our Assumption 3). While it has been shown in our Theorem 2 (and numerically confirmed in Figure 1) that ESPRIT produces consistent DoA estimates, its variance is observed in Figure 2 to be significantly larger than MUSIC and G-MUSIC. The (asymptotic) characterization of the variance of MUSIC and G-MUSIC are provided in [6], and it would be of future interest to extends such analysis to ESPRIT. Also, note interestingly that a *phase transition* behavior is observed for all three methods: at SNR $= 0$dB for uncorrelated and SNR $= 2$dB for correlated source. This is again an empirical manifestation of the counterintuitive large-dimensional behavior of SCM eigenspectra in Theorem 1.

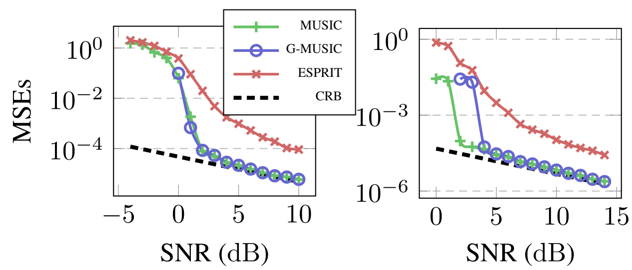

Fig. 2. Empirical mean squared errors (MSEs) of different DoA methods for widely spaced DoA (with $\theta_1 = 0$ and $\theta_2 = 5 \times 2\pi/N$) versus SNR, for (**left**) uncorrelated and (**right**) correlated source, with $N = 40$ and $T = 80$. Results obtained by averaging over 400 independent runs.

Figure 3 compares the MSEs of the three methods in the case of closely spaced DoA (that violates Assumption 2). We see that the improved G-MUSIC beats the standard MUSIC by a large margin and gets close to the Cramér–Rao bound, as in line with [6]; while ESPRIT appears to be sub-optimal in this setting, for which an improved estimate is needed.

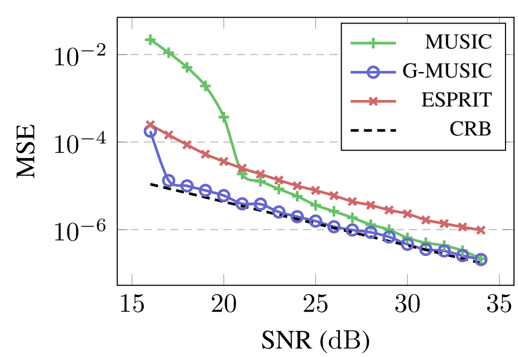

Fig. 3. Empirical MSEs for closely spaced DoA (with $\theta_1 = 0, \theta_2 = 0.25 \times 2\pi/N$) versus SNR, with $N = 40$ and $T = 80$. Results obtained by averaging over 400 independent runs.

5. REFERENCES

[1] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Transactions on Antennas and Propagation*, vol. 34, no. 3, pp. 276–280, Mar. 1986.

[2] A. Paulraj, R. Roy, and T. Kailath, "A subspace rotation approach to signal parameter estimation," *Proceedings of the IEEE*, vol. 74, no. 7, pp. 1044–1046, 1986.

[3] Romain Couillet and Mérouane Debbah, *Random Matrix Methods for Wireless Communications*, Cambridge University Press, 2011.

[4] Romain Couillet and Zhenyu Liao, *Random Matrix Methods for Machine Learning*, Cambridge University Press, 2022.

[5] Debashis Paul and Alexander Aue, "Random matrix theory in statistics: A review," *Journal of Statistical Planning and Inference*, vol. 150, pp. 1–29, July 2014.

[6] Pascal Vallet, Xavier Mestre, and Philippe Loubaton, "Performance analysis of an improved MUSIC DoA estimator," *IEEE Transactions on Signal Processing*, vol. 63, no. 23, pp. 6407–6422, 2015.

[7] Xavier Mestre and Miguel Ángel Lagunas, "Modified subspace algorithms for DoA estimation with large arrays," *IEEE Transactions on Signal Processing*, vol. 56, no. 2, pp. 598–614, 2008.

[8] R. Roy, A. Paulraj, and T. Kailath, "Comparative performance of ESPRIT and MUSIC for direction-of-arrival estimation," in *ICASSP '87. IEEE International Conference on Acoustics, Speech, and Signal Processing*, Apr. 1987, vol. 12, pp. 2344–2347.

[9] T.B. Lavate, V.K. Kokate, and A.M. Sapkal, "Performance Analysis of MUSIC and ESPRIT DOA Estimation Algorithms for Adaptive Array Smart Antenna in Mobile Communication," in *2010 Second International Conference on Computer and Network Technology*, Apr. 2010, pp. 308–311.

[10] Pascal Vallet and Philippe Loubaton, "On the Performance of MUSIC With Toeplitz Rectification in the Context of Large Arrays," *IEEE Transactions on Signal Processing*, vol. 65, no. 22, pp. 5848–5859, Nov. 2017.

[11] Vladimir A Marcenko and Leonid Andreevich Pastur, "Distribution of eigenvalues for some sets of random matrices," *Mathematics of the USSR-Sbornik*, vol. 1, no. 4, pp. 457, 1967.

[12] Jack W. Silverstein and Zhidong Bai, "On the Empirical Distribution of Eigenvalues of a Class of Large Dimensional Random Matrices," *Journal of Multivariate Analysis*, vol. 54, no. 2, pp. 175–192, 1995.

[13] Jinho Baik and Jack W. Silverstein, "Eigenvalues of large sample covariance matrices of spiked population models," *Journal of Multivariate Analysis*, vol. 97, no. 6, pp. 1382–1408, 2006.

[14] Florent Benaych-Georges and Raj Rao Nadakuditi, "The eigenvalues and eigenvectors of finite, low rank perturbations of large random matrices," *Advances in Mathematics*, vol. 227, no. 1, pp. 494–521, 2011.

[15] P. Stoica and Arye Nehorai, "MUSIC, maximum likelihood, and Cramer-Rao bound," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no. 5, pp. 720–741, May 1989.

[16] B. Ottersten, M. Viberg, and T. Kailath, "Performance analysis of the total least squares ESPRIT algorithm," *IEEE Transactions on Signal Processing*, vol. 39, no. 5, pp. 1122–1135, May 1991.