MUSIC for Joint Frequency Estimation: Stability with Compressive Measurements

Wenjing Liao
Statistical and Applied Mathematical Sciences Institute
Department of Mathematics, Duke University,Durham, NC, USA
wjliao@math.duke.edu

***Abstract*—This paper studies the application of MUtiple SIgnal Classification (MUSIC) algorithm on Multiple Measurement Vector (MMV) problem for the purpose of frequency parameter estimation while s true frequencies are located in the *continuum* of a bounded domain and sensors are randomly selected from a Uniform Linear Array (ULA). The MUSIC algorithm amounts to identifying a noise subspace from measurements, forming a noise-space correlation function and searching the s smallest local minima of the noise-space correlation function. Under the assumption that the true frequencies are separated by at least one Rayleigh Length (RL), we show that with high probability the noise-space correlation function is stably perturbed by noise if the number of sensors $n \sim \mathcal{O}(s)$ up to a logarithmic factor by means of a compressive version of discrete Ingham inequalities. As the theory implies, our numerical experiments demonstrate that the reconstruction error of MUSIC with n random sensors makes little difference once n is above a point of transition.**

***Index Terms*—MUSIC, joint frequency parameter estimation, random sensing, compressive discrete Ingham inequalities**

I. Introduction

Multiple Measurement Vector (MMV) problem addresses the recovery of a set of signal vectors that share common nonzero support. Specifically, let N, M and s be positive integers. In the MMV context, s, $N+1$ and M denote signal sparsity, the number of sensor elements and snapshots, respectively. For a sensing matrix $\Phi \in \mathbb{C}^{(N+1)\times s}$, the observation matrix $\bar{Y}^{\varepsilon} \in \mathbb{C}^{(N+1)\times M}$, contaminated by noise $\bar{E} \in \mathbb{C}^{(N+1)\times M}$, is generated from signal $Z \in \mathbb{C}^{s\times M}$ via

$$\bar{Y}^{\bar{E}} = \bar{Y} + \bar{E} = \Phi Z + \bar{E}. \tag{1}$$

Here every row of Z is nonzero and the sensing matrix Φ is unknown.

In this paper we explore a special kind of MMV problem for joint frequency parameter estimation which has wide applications in spectral estimation [26], array imaging [20] and inverse scattering [8]. Let $\mathbb{T} = [0, 1)$ be the torus with wrapped distance $d(\omega_j, \omega_l) = \min_{n\in\mathbb{Z}} |\omega_j + n - \omega_l|$. For any $\omega \in \mathbb{T}$ we define the imaging vector at ω as

$$\phi(\omega) = [1 \; e^{-2\pi i\omega} \; e^{-2\pi i 2\omega} \; \ldots \; e^{-2\pi i N\omega}]^T \in \mathbb{C}^{N+1}. \tag{2}$$

Columns of Φ are composed of imaging vectors at an unknown support set $\mathcal{S} = \{\omega_1, \ldots, \omega_s\}$, i.e.,

$$\Phi = [\phi(\omega_1) \; \phi(\omega_2) \; \ldots \; \phi(\omega_s)] \in \mathbb{C}^{(N+1)\times s} \tag{3}$$

and our objective is to estimate the support set $\mathcal{S}$. We assume $\omega_j \neq \omega_l, \forall j \neq l$. In the context of signal processing and array imaging, the model formulated in (1)(2)(3) corresponds to a ULA sensing geometry where sensors are placed on a one dimensional equispaced lattice as depicted in Figure 1.

Fig. 1. Uniform Linear Array (ULA).

The field of compressive sensing [2], [9] provides us with a new technology of reconstructing a sparse signal from a small of number of random linear measurements. The idea of saving data with probabilistic sensing suggests us explore the MMV problem with a random sensing geometry, i.e., using a few sensors randomly selected from the ULA [1], [16], [23], [33], [36]. Specifically, let $\mathcal{N}$ be a subset of $\{0, 1, \ldots, N\}$ indicating the sensor indices being selected and let $P_{\mathcal{N}}$ be the operator of selecting rows indexed to $\mathcal{N}$. Denote $n = |\mathcal{N}|$, $\Phi^{\mathcal{N}} = P_{\mathcal{N}}\Phi$ and $\phi^{\mathcal{N}}(\omega) = P_{\mathcal{N}}\phi(\omega)$. Data collected in the compressive sensing regime are

$$Y^E = Y + E = \Phi^{\mathcal{N}} Z + P_{\mathcal{N}}\bar{E}, \tag{4}$$

where $Y = \Phi^{\mathcal{N}} Z$ and $E = P_{\mathcal{N}}\bar{E}$.

An important quantity in frequency parameter estimation is Rayleigh Length (RL) [7] which is mathematically defined to be the distance between the center and the first zero of the Dirichlet kernel $D_N(\omega) = \int_{-N/2}^{N/2} e^{2\pi i t\omega} dt = (\sin \pi\omega N)/(\pi\omega)$. Hence 1 RL $= 1/N$.

The MUSIC algorithm, proposed by Schmidt [24], [25], has been widely applied to the MMV problem in various settings. In this paper we focus on the stability analysis of MUSIC in the random sensing regime. Among existing works on stability analysis of MUSIC, [12], [22] are mostly related to the present work. In [12], a stability criterion is given in the random sensing regime when true frequencies are assumed to be located exactly on a grid of spacing ≥ 1 RL. However, gridding error on this coarse grid can be very large [6], [11], [13], [14]. In order to allow for frequencies distributed on a continuum, we derived a perturbation estimate on noise-space correlation function for MUSIC in [22] where the gridding barrier is overcome by discrete Ingham inequalities. Theory in [22] gives rise to a stability analysis of MUSIC for MMV problem when the sensing geometry follows a standard ULA and sparsity information is not utilized. The main contribution of this paper is to develop a stability analysis of MUSIC with sparsity of signals exploited to reduce the number of sensors.

In literature there has been a lot of work on frequency parameter estimation in the form of Single Measurement Vectors (SMV) [3], [4], [5], [10], [11], [14], [15], [22], [29]. In this paper, we focus on the MMV problem which has wide applications in array imaging [20] and inverse scattering [8].

The paper is organized as follows. We introduce the MUSIC algorithm in Section II, discuss its stability in the compressive sensing regime in Section III, perform numerical simulations in Section IV and conclude in Section V. Some notations are defined here. For $p \in (0,\infty)$ and $\mathbf{v} \in \mathbb{C}^n$, $\|\mathbf{v}\|_p = (\sum_{j=1}^n |\mathbf{v}_j|^p)^{1/p}$. For matrix $A \in \mathbb{C}^{m\times n}$, A^T and A^* is the transpose and conjugate transpose of A.We use $\sigma_{\max}(A)$, $\sigma_{\min}(A)$ and $\sigma_j(A)$ to denote the maximum, minimum nonzero and the j-th singular value of A. Frobenius norm $\|A\|_F = (\sum_{j=1}^m \sum_{l=1}^n |A_{jl}|^2)^{1/2}$ and spectral norm $\|A\|_2 = \sigma_{\max}(A)$.

II. MUSIC ALGORITHM

For the MMV problem formulated in (2)(3)(4), MUSIC is performed through identifying a noise subspace from Y^E, forming a noise-space correlation function and searching the s smallest local minima of the noise-space correlation function.

In the noiseless case, $Y^E = Y = \Phi^{\mathcal{N}} Z$. The following lemma is the key of MUSIC.

Lemma 1. *If the following three conditions are satisfied*

1) *rank*$(Z) = s$,
2) *rank*$(\Phi^{\mathcal{N}}) = s$,
3) *rank*$([\Phi^{\mathcal{N}} \ \phi^{\mathcal{N}}(\omega)]) = s+1$ *for all* $\omega \notin \mathcal{S}$,

then $\mathrm{Range}(Y)$=$\mathrm{Range}(\Phi^{\mathcal{N}})$=$\mathrm{span}\{\phi^{\mathcal{N}}(\omega_1),\ldots,\phi^{\mathcal{N}}(\omega_s)\}$ *and the support set* $\mathcal{S}$ *can be identified through the criterion:*

$$\omega \in \mathcal{S} \text{ if and only if } \phi^{\mathcal{N}}(\omega) \in \mathrm{Range}(Y).$$

In practice, MUSIC algorithm is realized through extracting peaks of the imaging function. If conditions in Lemma 1 are satisfied, let the Singular Value Decomposition (SVD) of Y be

$$Y = [\underbrace{U_1}_{n\times s} \ \ \underbrace{U_2}_{n\times(n-s)}] \underbrace{\mathrm{diag}(\sigma_1,\ldots,\sigma_s,0,\ldots)}_{n\times M} [\underbrace{V_1}_{M\times s} \ \ \underbrace{V_2}_{M\times(M-s)}]^*,$$

where $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_s > 0$. The column space of U_1 and U_2 are called signal space and noise space, respectively. Signal space is the same as $\mathrm{Range}(Y)$ and noise space is the null space of Y^*. Let $\mathcal{P}_1$ and $\mathcal{P}_2$ be the orthogonal projection onto signal space and noise space. For any $\mathbf{v} \in \mathbb{C}^n$,

$$\mathcal{P}_1\mathbf{v} = U_1(U_1^*\mathbf{v}) \ \text{ and } \ \mathcal{P}_2\mathbf{v} = U_2(U_2^*\mathbf{v}).$$

Suppose conditions in Lemma 1 are satisfied. Then $\omega \in \mathcal{S}$ if and only if $\mathcal{P}_2\phi^{\mathcal{N}}(\omega) = \mathbf{0}$. Therefore the noise-space correlation function

$$\mathcal{R}(\omega) = \frac{\|\mathcal{P}_2\phi^{\mathcal{N}}(\omega)\|_2}{\|\phi^{\mathcal{N}}(\omega)\|_2} = \frac{\|U_2^*\phi^{\mathcal{N}}(\omega)\|_2}{\|\phi^{\mathcal{N}}(\omega)\|_2},$$

vanishes and the imaging function $\mathcal{J}(\omega) = 1/\mathcal{R}(\omega)$ peaks exactly on $\mathcal{S}$.

In the presence of noise, $Y^E = Y + E = \Phi^{\mathcal{N}}Z + E$. Suppose the SVD of Y^E is

$$Y^E = [\underbrace{U_1^E}_{n\times s} \ \ \underbrace{U_2^E}_{n\times(n-s)}] \underbrace{\mathrm{diag}(\sigma_1^E,\ldots,\sigma_s^E,\ldots)}_{n\times M} [\underbrace{V_1^E}_{M\times s} \ \ \underbrace{V_2^E}_{M\times(M-s)}]^*.$$

The noise-space correlation function and imaging function become

$$\mathcal{R}^E(\omega) = \frac{\|\mathcal{P}_2^E\phi^{\mathcal{N}}(\omega)\|_2}{\|\phi^{\mathcal{N}}(\omega)\|_2} = \frac{\|U_2^{E*}\phi^{\mathcal{N}}(\omega)\|_2}{\|\phi^{\mathcal{N}}(\omega)\|_2}$$

and $\mathcal{J}^E(\omega) = 1/\mathcal{R}^E(\omega)$ where $\mathcal{P}_1^E$ and $\mathcal{P}_2^E$ stand for the orthogonal projection onto the column space of U_1^E and U_2^E respectively.

In the MUSIC algorithm, recovered support $\hat{\mathcal{S}} = \{\hat{\omega}_1,\ldots,\hat{\omega}_s\}$ is identified through the s smallest local minima of $\mathcal{R}^E(\omega)$ or the s largest local maxima of $\mathcal{J}^E(\omega)$.

III. STABILITY OF MUSIC WITH COMPRESSIVE MEASUREMENTS

Let $p \gg N$ be a prime integer. In this section we assume that true frequencies $\{\omega_j\}$ are located on a fine grid with grid spacing $\ell = 1/p$, i.e., $\mathcal{S} \subset \mathcal{L} = \left\{\frac{0}{p},\frac{1}{p},\ldots,\frac{p-1}{p}\right\}$. If true frequencies are not exactly located on $\mathcal{L}$, we can approximate them by the nearest grid points and the gridding error, to be incorporated in E, is small while $p \gg N$ [14, Figure 1].

In the noiseless case, $s+1$ sensors guarantees a successful recovery of s distinct frequencies.

Theorem 1. *Suppose* $E = \mathbf{0}$, *rank*$(Z) = s$ *and* $\mathcal{S} \subset \mathcal{L}$. *Let* $\mathcal{R}(\omega)$ *and* $\mathcal{J}(\omega)$ *be evaluated on* $\mathcal{L}$. *If* $n \geq s+1$, *then* $\omega \in \mathcal{S}$ *if and only if* $\mathcal{R}(\omega) = 0$ *and* $\mathcal{J}(\omega) = \infty$.

Proof: The proof is based on [2, Lemma 1.2], [30, Corollary 1.4], which implies that when p is a prime and $n \geq s+1$, $\mathrm{rank}(\Phi^{\mathcal{N}}) = s$ and $\forall\omega \in \mathcal{L}$, $\mathrm{rank}([\Phi^{\mathcal{N}} \ \phi^{\mathcal{N}}(\omega)]) = s+1$ if and only if $\omega \notin \mathcal{S}$. ■

Performance of MUSIC in the presence of noise relies on how much the noise-space correlation function is perturbed from $\mathcal{R}(\omega)$ to $\mathcal{R}^E(\omega)$. This question is answered by the following Lemma in [22].

Lemma 2. *Suppose* $\|E\|_2 < \sigma_s$. *Then*

$$|\mathcal{R}^E(\omega) - \mathcal{R}(\omega)| \leq \frac{4\sigma_1 + 2\|E\|_2}{(\sigma_s - \|E\|_2)^2}\|E\|_2. \quad (5)$$

Remark 1. *Lemma 2 implies that for every* $\omega_j \in \mathcal{S}$, *there exits a local minimum of* $\mathcal{J}^E(\omega)$, *denoted by* $\hat{\omega}_j$, *such that* $\hat{\omega}_j \to \omega_j$ *as* $\|E\|_2 \to 0$. *The rate of convergence depends on* σ_1, σ_s *and the function* $\mathcal{R}^2(\omega)$ *in the neighborhood of* ω_j *[22, Theorem 4].*

To make the bounds (5) explicit and meaningful, we provide an upper bound of σ_1 and a lower bound of σ_s while frequencies in $\mathcal{S}$ are separated above 1 RL and the number of sensors $n \sim \mathcal{O}(s)$ up to a logarithmic factor.

As $\sigma_1 = \sigma_{\max}(\Phi^{\mathcal{N}} Z) \leq \sigma_{\max}(\Phi^{\mathcal{N}})\sigma_{\max}(Z)$ and $\sigma_s = \sigma_{\min}(\Phi^{\mathcal{N}} Z) \geq \sigma_{\min}(\Phi^{\mathcal{N}})\sigma_{\min}(Z)$, the key is to estimate singular values of $\Phi^{\mathcal{N}}$.

We are motivated by the classical Ingham inequalities [17], [19], [35, (pp.162-164)] which studies the stability of complex exponential sums in the system $\{e^{-2\pi i \omega_j t},\ t \in [0,T],\ \omega_j \in \mathbb{R},\ j = 1,\ldots,s\}$ while $\{\omega_j\}$ satisfies a gap condition. A discrete version of Ingham inequalities corresponding to discrete samples $t = 0,1,\ldots,T$ was proved in [22, Theorem 2]. A main result here is the following compressive version of discrete Ingham inequalities whose proof is in the appendix.

Theorem 2. *Suppose $N > 10$ and $\mathcal{S}$ satisfies the gap*

$$q = \min_{j \neq l} d(\omega_j, \omega_l) > \frac{1}{N}\sqrt{\frac{2}{\pi}}\left(\frac{2}{\pi} - \frac{4}{N}\right)^{-\frac{1}{2}}. \quad (6)$$

Denote

$$A(q,N) = \frac{2}{\pi} - \frac{2}{\pi N^2 q^2} - \frac{4}{N},$$

$$B(q,N) = \begin{cases} \frac{4\sqrt{2}}{\pi} + \frac{\sqrt{2}}{\pi N^2 q^2} + \frac{3\sqrt{2}}{N} & N \text{ even} \\ \left(1 + \frac{1}{N}\right)\left(\frac{4\sqrt{2}}{\pi} + \frac{\sqrt{2}}{\pi (N+1)^2 q^2} + \frac{3\sqrt{2}}{N+1}\right) & N \text{ odd.} \end{cases}$$

Let $\tau \in (0, A(q,N))$. If the index set $\mathcal{N}$ of size n is selected uniformly at random from $\{0,\ldots,N\}$, then

$$\frac{1}{n}\sigma^2_{\min}(\Phi^{\mathcal{N}}) \geq A(q,N) - \tau, \quad (7)$$

$$\frac{1}{n}\sigma^2_{\max}(\Phi^{\mathcal{N}}) \leq B(q,N) + \tau \quad (8)$$

with probability at least $1 - \epsilon$ provided $n > \frac{8s}{\tau^2}\log\frac{4s}{\epsilon}$.

Theorem 2 leads to a perturbation estimate of the noise-space correlation function with respect to noise E in the random sensing regime while $\mathcal{S}$ satisfies the gap condition (6).

Theorem 3. *Suppose $N > 10$ and $\mathcal{S}$ satisfies the gap condition* (6). *Let $\tau \in (0, A(q,N))$. Suppose*

$$\frac{\|E\|_2}{\sqrt{n}} < \sigma_{\min}(Z)\sqrt{A(q,N) - \tau}.$$

If the index set $\mathcal{N}$ of size n is selected uniformly at random from $\{0,\ldots,N\}$, then

$$|\mathcal{R}^E(\omega) - \mathcal{R}(\omega)| \leq \frac{4\sigma_{\max}(Z)\sqrt{B(q,N)+\tau} + \frac{2\|E\|_2}{\sqrt{n}}}{\left(\sigma_{\min}(Z)\sqrt{A(q,N)-\tau} - \frac{\|E\|_2}{\sqrt{n}}\right)^2} \cdot \frac{\|E\|_2}{\sqrt{n}} \quad (9)$$

with probability at least $1 - \epsilon$ provided $n > \frac{8s}{\tau^2}\log\frac{4s}{\epsilon}$.

In the case of sensing with a complete ULA where $n = N+1$, one can expect

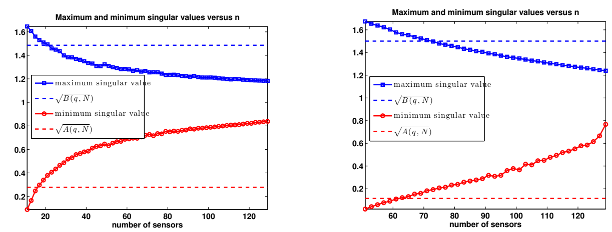
$$|\mathcal{R}^E(\omega) - \mathcal{R}(\omega)| \leq \frac{4\sigma_{\max}(Z)\sqrt{B(q,N)} + \frac{2\|E\|_2}{\sqrt{n}}}{\left(\sigma_{\min}(Z)\sqrt{A(q,N)} - \frac{\|E\|_2}{\sqrt{n}}\right)^2} \cdot \frac{\|E\|_2}{\sqrt{n}}.$$

Theorem 3 demonstrates that similar stability bound is achieved with high probability while sensors are randomly selected from a ULA with size $n \sim \mathcal{O}(s)$ up to a logarithmic factor. We conclude that for sparse signals the strategy of saving data with random sensing applies to the MUSIC algorithm.

IV. Numerical Experiments

In this section we provide some numerical experiments about $\sigma_{\max}(\Phi^{\mathcal{N}})/\sqrt{n}$, $\sigma_{\min}(\Phi^{\mathcal{N}})/\sqrt{n}$ and the performance of MUSIC as n increases.

Let $N = 128$. We fix two support sets with sparsity $s_1 = 10$, gap $q_1 > 1/N$ and $s_2 = 50$, gap $q_2 > 1/N$, respectively. Figure 2 shows the average $\sigma_{\max}(\Phi^{\mathcal{N}})/\sqrt{N}$ and $\sigma_{\min}(\Phi^{\mathcal{N}})/\sqrt{N}$ in 100 trails for two support sets as n increases from $s+1$ to $N+1$. The upper bound $\sqrt{B(q,N)}$ and lower bound $A(q,N)$ in Theorem 2 are represented by dotted lines.



(a) $s = 10$ (b) $s = 50$

Fig. 2. $\sigma_{\max}(\Phi^{\mathcal{N}})/\sqrt{n}$ and $\sigma_{\min}(\Phi^{\mathcal{N}})/\sqrt{n}$ versus n.

Next we fix $s = 50$, $M = 70$, $N = 128$ and then RL $= 1/128$. A signal $Z \in \mathbb{R}^{50\times 70}$ is generated with i.i.d. random entries $Z_{jl} \sim \text{Normal}(0, 1/M)$ and a support $\mathcal{S}$ is fixed with sparsity $s = 50$ and gap $q \in (1,2)$ RL. Noise matrix E contains i.i.d. random normal variables, i.e., $\text{Re}(E_{kj}) \sim \text{Normal}(0,\sigma^2)$ and $\text{Im}(E_{kj}) \sim \text{Normal}(0,\sigma^2)$. Noise to Signal Ratio (NSR) is defined as $\frac{\mathbb{E}\|E\|_F}{\|Y\|_F} = \frac{\sigma\sqrt{2nM}}{\|Y\|_F}$. The recovered support $\hat{\mathcal{S}}$ is identified through the s largest local maxima of $\mathcal{J}^E(\omega)$. Reconstruction error is measured by the Hausdorff distance between $\mathcal{S}$ and $\hat{\mathcal{S}}$, i.e., $d(\mathcal{S},\hat{\mathcal{S}}) = \max\left\{\max_{\hat\omega\in\hat{\mathcal{S}}}\min_{\omega\in\mathcal{S}} d(\omega,\hat\omega), \max_{\omega\in\mathcal{S}}\min_{\hat\omega\in\hat{\mathcal{S}}} d(\omega,\hat\omega)\right\}$.

Let n sensors be randomly selected from a ULA. For NSR $= 50\%, 60\%, 70\%$, we vary n from $s+1$ to $N+1$ and record the average reconstruction error $d(\mathcal{S},\hat{\mathcal{S}})$ in 100 trials. Figure 3 shows the average reconstruction error in the unit of RL versus n. It is clear that the reconstruction error with n random sensors is similar to the error with a complete ULA once n is large enough.

V. Conclusions

We have provided a stability analysis and numerical experiments of the MUSIC algorithm for joint frequency parameter estimation in the random sensing regime. In comparison with existing analysis of MUSIC [18], [21], [27], [28], [31], our theory features an explicit estimate of the singular values of the matrix $\Phi^{\mathcal{N}}$ due to compressive discrete Ingham inequalities.

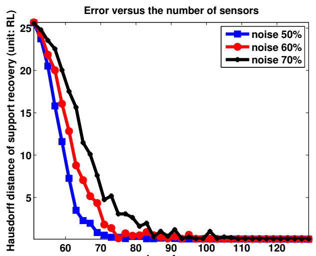


Fig. 3. Average reconstruction error $d(\mathcal{S}, \hat{\mathcal{S}})$ in 100 trials versus n.

ACKNOWLEDGEMENT

Wenjing Liao is grateful to support from NSF DMS-0847388 and SAMSI under Grant NSF DMS-1127914.

APPENDIX

Proof of Theorem 2 combines discrete Ingham inequalities[22] and matrix Bernstein inequalities[32]. The proof is composed of two parts. In Section A, we use a technique in [2] to translate results from a Bernoulli probability model into results for the uniform probability model. Let P_Δ be the operator of randomly selecting rows of Φ using the Bernoulli model. Singular values of $\frac{1}{\sqrt{n}} P_\Delta \Phi$ are estimated in Section B.

A. The Bernoulli model

We consider a subset Δ of $\{0, \ldots, N\}$ sampled using the Bernoulli model. We first create a sequence

$$\delta_k = \begin{cases} 1 & \text{with probability } p \\ 0 & \text{with probability } 1-p \end{cases}, \; p = \frac{n}{N}, \; k = 0, \ldots, N.$$

and then set $\Delta = \{k \mid \delta_k = 1, \; k = 0, \ldots, N\}$. Define Failure($\mathcal{N}$) as the event where (7)(8) in Theorem 2 fail on $\mathcal{N}$. The argument in [2, pp. 496] shows that $\mathbb{P}(\text{Failure}(\mathcal{N})) \le 2\mathbb{P}(\text{Failure}(\Delta))$ and then it suffices to prove that singular values of $\frac{1}{\sqrt{n}} P_\Delta \Phi$ satisfy (7)(8) with probability at least $1 - \epsilon/2$.

B. Singular values of $\frac{1}{\sqrt{n}} P_\Delta \Phi$

Denote $R = \frac{1}{n}(P_\Delta \Phi)^*(P_\Delta \Phi)$. We define

$$\mathbf{r}_k = [e^{-2\pi i k \omega_1} \; e^{-2\pi i k \omega_2} \; \ldots \; e^{-2\pi i k \omega_s}] \in \mathbb{C}^s,$$

$$R_k = \frac{1}{n} \delta_k \mathbf{r}_k^* \mathbf{r}_k, \; k = 0, \ldots, N.$$

$$R = \sum_{k=0}^{N} R_k, \quad \mathbb{E}R_k = \frac{1}{N} \mathbf{r}_k^* \mathbf{r}_k, \quad \mathbb{E}R = \sum_{k=0}^{N} \mathbb{E}R_k = \frac{1}{N} \Phi^* \Phi.$$

We first recall discrete Ingham inequalities [22, Theorem 2] which provide an estimate of eigenvalues of $\mathbb{E}R$ and then apply matrix Bernstein inequalities to show that R deviates little from $\mathbb{E}R$.

Lemma 3. *Suppose $\mathcal{S}$ satisfes the gap condition* (6)*. Then*

$$A(q, N)\|\mathbf{c}\|_2^2 \le \frac{1}{N}\|\Phi \mathbf{c}\|_2^2 \le B(q, N)\|\mathbf{c}\|_2^2, \quad \forall \mathbf{c} \in \mathbb{C}^s.$$

In other words,

$$\lambda_{\max}(\mathbb{E}R) \le B(q, N) \;\; \text{and} \;\; \lambda_{\min}(\mathbb{E}R) \ge A(q, N).$$

Lemma 4 (Matrix Bernstein Inequality [32, Theorem 1.6.2])**.** *Let $R_0, \ldots, R_N$ be $s \times s$ independent random matrices. Assume that each matrix has bounded deviation from its mean: $\|R_k - \mathbb{E}R_k\|_2 \le F$, for $k = 0, \ldots, N$. Form $R = \sum_{k=0}^{N} R_k$ and $\sigma^2 = \max\{\|\mathbb{E}[(R - \mathbb{E}R)(R - \mathbb{E}R)^*]\|_2, \|\mathbb{E}[(R - \mathbb{E}R)^*(R - \mathbb{E}R)]\|_2\}$. Then*

$$\mathbb{P}(\|R - \mathbb{E}R\|_2 \ge \tau) \le 2s \exp\left(\frac{-\tau^2/2}{\sigma^2 + F\tau/3}\right) \text{ for all } \tau \ge 0.$$

Applying matrix Bernstein inequality on R yields:

Lemma 5. *Suppose $N > 10$ and $\mathcal{S}$ satisfies the gap condition* (6)*. If $\tau \in (0, 3)$, we have $\|R - \mathbb{E}R\|_2 \le \tau$ with probability at least $1 - \epsilon/2$ provided $n > \frac{8s}{\tau^2} \log \frac{4s}{\epsilon}$.*

Proof: We first compute the quantities necessary for the application of Lemma 4.

$$\|R_k - \mathbb{E}R_k\|_2 = \|\frac{1}{n}(\delta_k - p)\mathbf{r}_k^* \mathbf{r}_k\|_2 \le \frac{1}{n}\|\mathbf{r}_k^* \mathbf{r}_k\|_2 = \frac{s}{n}.$$

$$\sigma^2 = \|\mathbb{E}[(R - \mathbb{E}R)(R - \mathbb{E}R)]\|_2$$
$$= \left\| \mathbb{E}\left[\sum_{k=0}^{N} \frac{1}{n}(\delta_k - p)\mathbf{r}_k^* \mathbf{r}_k\right] \left[\sum_{k'=0}^{N} \frac{1}{n}(\delta_{k'} - p)\mathbf{r}_{k'}^* \mathbf{r}_{k'}\right] \right\|_2$$
$$= \left\| \frac{1}{n^2} \sum_{k=0}^{N} \mathbb{E}(\delta_k - p)^2 \|r_k\|_2^2 \mathbf{r}_k^* \mathbf{r}_k \right\|_2 = \frac{Np(1-p)s}{n^2} \left\| \frac{1}{N} \sum_{k=0}^{N} \mathbf{r}_k^* \mathbf{r}_k \right\|_2.$$

According to Lemma 3, $\left\| \frac{1}{N} \sum_{k=0}^{N} \mathbf{r}_k^* \mathbf{r}_k \right\|_2 = \frac{1}{N} \|\Phi^* \Phi\|_2 \le B(q, N) < 3$ when $q > \frac{1}{N}$ and $N > 10$. As a result

$$\sigma^2 \le \frac{3Ns(1-p)p}{n^2} = \frac{3s(1-p)}{n} \le \frac{3s}{n}.$$

Applying Lemma 4 yields

$$\mathbb{P}(\|R - \mathbb{E}R\|_2 \ge \tau) \le 2s \exp\left(\frac{-\tau^2/2}{\frac{3s}{n} + \frac{s\tau}{3n}}\right)$$
$$\le 2s \exp\left(\frac{-\tau^2/2}{\frac{4s}{n}}\right) \text{ while } \tau < 3 = 2s \exp\left(-\frac{n\tau^2}{8s}\right) \le \frac{\epsilon}{2},$$

if $n > \frac{8s}{\tau^2} \log \frac{4s}{\epsilon}$. ∎

Combining Lemma 3 and Lemma 5 gives rise to:

Lemma 6. *Suppose $N > 10$ and the gap condition* (6) *is satisfied. Let $\tau \in (0, A(q, N))$. Then*

$$\frac{1}{n}\sigma_{\min}^2(P_\Delta \Phi) \ge A(q, N) - \tau,$$
$$\frac{1}{n}\sigma_{\max}^2(P_\Delta \Phi) \le B(q, N) + \tau$$

with probability at least $1 - \epsilon/2$ provided $n > \frac{8s}{\tau^2} \log \frac{4s}{\epsilon}$.

Proof: According to Weyl Theorem [34], $\lambda_{\min}(R) \ge \lambda_{\min}(\mathbb{E}R) - \|R - \mathbb{E}R\|_2$ and $\lambda_{\max}(R) \le \lambda_{\max}(\mathbb{E}R) + \|R - \mathbb{E}R\|_2$. Lemma 3, Lemma 5 and Weyl Theorem give rise to Lemma 6. ∎

Finally Section A and Lemma 6 in Section B give rise to Theorem 2.

REFERENCES

[1] Y. I. Abramovich, D. A. Gray, A. Y. Gorokhov and N. K. Spencer, "Comparison of DOA estimation performance for various types of sparse antenna array geometries", Proc. EUSIPCO. **96**, 1996.

[2] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory*, **52(2)** pp.489-509, 2006.

[3] E. J. Candès and C. Fernandez-Granda,"Towards a mathematical theory of super-resolution," *Communications on Pure and Applied Mathematics*, 2013.

[4] E. J. Candès and C. Fernandez-Granda,"Super-resolution from noisy data," *Journal of Fourier Analysis and Applications*, **19(6)**, pp.1229-1254, 2013.

[5] Yuxin Chen and Yuejie Chi, "Robust spectral compressed sensing via structured matrix completion," arXiv:1304.8126, 2013.

[6] Y. Chi, L. L. Scharf, A. Pezeshki and R. A. Calderbank, "Sensitivity to basis mismatch in compressed sensing", *IEEE Transactions on Signal Processing*, **59(5)**, pp.2182-2195, 2011.

[7] A. J. Den Dekker and A. van den Bos, "Resolution: a survey," *Journal of Optical Society of America A,* **14(3)**, pp. 547-557, 1997.

[8] A. J. Devaney, "Super-resolution processing of multi-static data using time reversal and MUSIC," *Journal of the Acoustical Society of America*, 2000.

[9] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, **52(4)** pp.1289-1306, 2006.

[10] D.L. Donoho, "Superresolution via sparsity constraints," *SIAM Journal on Mathematical Analysis*, **23**, pp.1309-1331, 1992.

[11] M.F. Duarte and R.G. Baraniuk, "Spectral compressive sensing," *Applied and Computational Harmonic Analysis*, **35(1)**, pp. 111-129, 2013.

[12] A. Fannjiang, "The MUSIC algorithm for sparse objects: a compressed sensing analysis," *Inverse Problems*, **27**, 035013, 2011.

[13] A. Fannjiang and W. Liao, "Mismatch and resolution in compressive imaging," *SPIE Optical Engineering+ Applications, International Society for Optics and Photonics*, 2011.

[14] A. Fannjiang, and W. Liao, "Coherence pattern-guided compressive sensing with unresolved grids," *SIAM Journal on Imaging Sciences,* **5(1)**, pp.179-202, 2012.

[15] A. Fannjiang and W. Liao, "Super-resolution by compressive sensing algorithms," *IEEE Asilomar Conference on Signals, Systems and Computers*, 2012.

[16] J. J. Fuchs, "Extension of the Pisarenko method to sparse linear arrays," *IEEE Transactions on Signal Processing* **45(10)**, pp.2413-2421, 1997.

[17] A. E. Ingham, "Some trigonometrical inequalities with applications to the theory of series," *Mathematische Zeitschrift*, **41(1)**, pp.367-379, 1936.

[18] B. A. Johnson, Y. I. Abramovich and X. Mestre, "MUSIC, G-MUSIC, and maximum-likelihood performance breakdown," *IEEE Transactions on Signal Processing*, **56(8)**, pp.3944-3958, 2008.

[19] V. Komornik and P. Loreti, *Fourier Series in Control Theory*, New York: Springer, 2005.

[20] H. Krim and M. Viberg, "Two decades of array signal processing research: the parametric approach," *IEEE Signal Processing Magazine*, **13(4)**, pp.67-94, 1996.

[21] F. Li, H. Liu, and R. J. Vaccaro, "Performance analysis for DOA estimation algorithms: unification, simplification, and observations," *IEEE Transactions on Aerospace and Electronic Systems*, **29(4)**, pp.1170-1184, 1993.

[22] W. Liao and A. Fannjiang, "MUSIC for single-snapshot spectral estimation: stability and super-resolution," submitted, 2014.

[23] P. Pakrooh, A. Pezeshki and L. L. Scharf, "Threshold effects in parameter estimation from compressed data," *GlobalSIP*, 2013.

[24] R. O. Schmidt, "A signal subspace approach to multiple emitter location and spectral estimation," Ph.D. thesis, Stanford Univ,. Stanford, CA, Nov. 1981.

[25] R. O. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Transactions on Antennas and Propagation*, **34(3)**, pp.276-280, 1986.

[26] P. Stoica and R. Moses, *Spectral Analysis of Signals*, Prentice Hall, New Jersey, 2005.

[27] P. Stoica and A. Nehorai, "MUSIC, maximum likelihood, and Cramer-Rao bound," *IEEE Transaction on Acoustics, Speech and Signal Processing*, **37(5)**, pp.720-741, 1989.

[28] A. L. Swindlehurst and T. Kailath, "A performance analysis of subspace-based methods in the presence of model errors. I. The MUSIC algorithm," *IEEE Transactions on Signal Processing*, **40(7)**, pp.1758-1774, 1992.

[29] G. Tang, B. Bhaskar, P. Shah, and B. Recht, "Compressed sensing off the grid," *IEEE Transactions on Information Theory*, **59(11)**, pp. 7465-7490, 2013.

[30] T. Tao, "An uncertainty principle for cyclic groups of prime order," *Mathematical Research Letters*, **12(1)**, pp. 121-127, 2005.

[31] J. K. Thomas, L. L. Scharf and D. W. Tufts, "The probability of a subspace swap in the SVD," *IEEE Transactions on Signal Processing*, **43(3)**, pp.730-736, 1995.

[32] J. A. Tropp, "User-friendly tail bounds for sums of random matrices," *Foundations of Computational Mathematics*, **12(4)** pp.389-434, 2012.

[33] S. A. Vorobyov, A. B. Gershman and K. M. Wong, "Maximum likelihood direction-of-arrival estimation in unknown noise fields using sparse sensor arrays," *IEEE Transactions on Signal Processing* **53(1)**, pp.34-43, 2005.

[34] H. Weyl, "Das asymptotische Verteilungsgesetz der Eigenwerte linearer partieller Differentialgleichungen (mit einer Anwendung auf die Theorie der Hohlraumstrahlung)," *Mathematische Annalen*, **71(4)**, pp.441-479, 1912.

[35] R. M. Young, *An Introduction to Nonharmonic Fourier Series*, Academic Press, New York, 1980.

[36] M. D. Zoltowski and K. T. Wong, "Closed-form eigenstructure-based direction finding using arbitrary but identical subarrays on a sparse uniform Cartesian array grid," *IEEE Transactions on Signal Processing*, **48(8)**, pp.2205-2210, 2000.