Conditioning of restricted Fourier matrices and super-resolution of MUSIC

Weilin Li
Courant Institute of Mathematical Sciences
New York University
Email: weilinli@cims.nyu.edu

Wenjing Liao
School of Mathematics
Georgia Institute of Technology
Email: wliao60@gatech.edu

***Abstract*—This paper studies stable recovery of a collection of point sources from its noisy $M+1$ low-frequency Fourier coefficients. We focus on the super-resolution regime where the minimum separation of the point sources is below $1/M$. We propose a separated clumps model where point sources are clustered in far apart sets, and prove an accurate lower bound of the Fourier matrix with nodes restricted to the source locations. This estimate gives rise to a theoretical analysis on the super-resolution limit of the MUSIC algorithm.**

I. Introduction

In imaging and signal processing, S point sources are usually represented by a discrete measure: $\mu(\omega) = \sum_{j=1}^{S} x_j \delta_{\omega_j}(\omega)$, where $x = \{x_j\}_{j=1}^{S} \in \mathbb{C}^S$ represents the source amplitudes and $\Omega = \{\omega_j\}_{j=1}^{S} \subseteq \mathbb{T} := [0,1)$ represents the source locations. A uniform array of $M+1$ sensors collects the noisy Fourier coefficients of μ, denoted by $y \in \mathbb{C}^{M+1}$. One can write

$$y = \Phi_M x + \eta, \tag{I.1}$$

where $\Phi_M = \Phi_M(\Omega)$ is the $(M+1) \times S$ *Fourier* or *Vandermonde* matrix (with nodes on the unit circle):

$$\Phi_M(\Omega) = \begin{bmatrix} 1 & \dots & 1 \\ e^{-2\pi i\omega_1} & \dots & e^{-2\pi i\omega_S} \\ \vdots & \vdots & \vdots \\ e^{-2\pi i M\omega_1} & \dots & e^{-2\pi i M\omega_S} \end{bmatrix},$$

and $\eta \in \mathbb{C}^{M+1}$ represents noise.

Our goal is to accurately recover μ, especially the support Ω, from y. The measurements y contains information about μ at a coarse resolution of approximately $1/M$, whereas we would like to estimate μ with a higher resolution. In the noiseless setting where $\eta = 0$, the measure μ can be exactly recovered by many methods. With noise, the stability of this inverse problem depends on Ω. A crucial quantity is the *minimum separation* between the two closest points in Ω, defined as

$$\Delta = \Delta(\Omega) = \min_{1\le j<k\le S} |\omega_j - \omega_k|_{\mathbb{T}},$$

where $|\cdot|_{\mathbb{T}}$ is the metric on the torus $\mathbb{T}$. In imaging, $1/M$ is regarded as the standard resolution. As a manifestation of the Heisenberg uncertainty principle, recovery is sensitive to noise whenever $\Delta < 1/M$, which case is referred as super-resolution. The *super-resolution factor* (SRF) is M/Δ, standing for the maximum number of points in Ω that is contained in an interval of length $1/M$.

Prior mathematical work on super-resolution can be placed in three main categories: (a) the min-max error of super-resolution was studied in [1], [2] when point sources are on a fine grid of $\mathbb{R}$; (b) when Ω is *well-separated* such that $\Delta \ge C/M$ for some constant $C > 1$, some representative methods include total variation minimization (TV-min) [3], [4], [5], greedy algorithms [6], and subspace methods [7], [8]. These results address the issue of discretization error [6] arising in sparse recovery, but they do not always succeed when $\Delta < 1/M$; (c) when $\Delta < 1/M$, certain assumptions on the signs of μ are required by many optimization-based methods [9], [10], [11]. Alternatively, *subspace methods* exploit a low-rank factorization of the data and can recover *complex* measures, but there are many unanswered questions related to its stability that we would like to address.

This paper focuses on a highly celebrated subspace method, called MUltiple SIgnal Classification (MUSIC) [12]. An important open problem is to understand the *super-resolution limit* of MUSIC: characterize the support sets Ω and noise level for which MUSIC can stably recover all measures μ supported in Ω within a prescribed accuracy. Prior numerical experiments in [7] showed that MUSIC can succeed even when $\Delta < 1/M$, but a rigorous justification was not provided. This is one of our main motivations for the theory presented in this paper and in our more detailed preprint [13].

As a result of Wedin's theorem [17], [18], the stability of MUSIC obeys, in an informal manner,

$$\text{Sensitivity} \leq \underbrace{\frac{\text{Constant}}{x_{\min}\sigma_{\min}^2(\Phi_M)}}_{\text{Noise amplification factor}} \cdot \underbrace{Q(\eta)}_{\text{Noise term}},$$

where $x_{\min} = \min_j |x_j|$, $\sigma_{\min}(\Phi_M)$ is the smallest non-zero singular value of Φ_M, and $Q(\eta)$ is a quantity depending on noise. Therefore, MUSIC can accurately estimate μ provided that the noise term is sufficiently small compared to the noise amplification factor which depends crucially on $\sigma_{\min}(\Phi_M)$.

In the separated case $\Delta > 1/M$, accurate estimates for $\sigma_{\min}(\Phi_M)$ and $\sigma_{\max}(\Phi_M)$ are known [14], [15], [8], [7]. In the super-resolution regime $\Delta < 1/M$, the value of $\sigma_{\min}(\Phi_M)$ is extremely sensitive to the "geometry" or configuration of Ω, and a more sophisticated description of the "geometry" of Ω other than the minimum separation is required. Based on this observation, we define a *separated clumps* model where Ω consists of well-separated subsets, where each subset contains several closely spaced points. This situation occurs naturally in applications where point sources clustered in far apart sets.

Under this separated clumps model, we provide a lower bound of $\sigma_{\min}(\Phi_M)$ with the dominant term scaling like $\text{SRF}^{-\lambda+1}$, where λ is the cardinality of the largest clump. This is a significant improvement on existing lower bounds with continuous measurements where the exponents depend on the total sparsity S [1], [2]. We use this estimate to rigorously establish the resolution limit of MUSIC and explain numerical results. More comprehensive explanations, comparisons, simulations, and proofs can be found in [13].

II. Minimum singular value of Vandermonde matrices

We first define a geometric model of Ω where the point sources are clustered into far apart clumps.

Assumption 1 (Separated clumps model)**.** Let M and A be a positive integers and $\Omega \subseteq \mathbb{T}$ have cardinality S. We say that Ω consists of A *separated clumps* with parameters (M, S, α, β) if the following hold.

1) Ω can be written as the union of A disjoint sets $\{\Lambda_a\}_{a=1}^A$, where each *clump* Λ_a is contained in an interval of length $1/M$.
2) $\Delta \geq \alpha/M$ with $\max_{1\leq a\leq A}(\lambda_a - 1) < 1/\alpha$ where λ_a is the cardinality of Λ_a.
3) If $A > 1$, then the distance between any two clumps is at least β/M.

There are many types of discrete sets that consist of separated clumps. Extreme examples include when Ω is a single clump containing all S points, and when Ω consists of S clumps containing a single point. While our theory applies to both extremes, the in-between case where Ω consists of several clumps each of modest size is the most interesting, and developing a theory of super-resolution for this case has turned out to be quite challenging.

Under this separated clumps model, we expect $\sigma_{\min}(\Phi_M)$ to be an ℓ^2 aggregate of A terms, where each term only depends on the "geometry" of each clump.

Theorem 1. *Let $M \geq S^2$. Assume Ω satisfies Assumption 1 with parameters (M, S, α, β) for some $\alpha > 0$ and*

$$\beta \geq \max_{1\leq a\leq A} \frac{20S^{1/2}\lambda_a^{5/2}}{\alpha^{1/2}}. \qquad \text{(II.1)}$$

Then there exist explicit constants $C_a > 0$ such that

$$\sigma_{\min}(\Phi_M) \geq \sqrt{M}\Big(\sum_{a=1}^{A}\big(C_a\alpha^{-\lambda_a+1}\big)^2\Big)^{-\frac{1}{2}}. \qquad \text{(II.2)}$$

The main feature of this theorem is the exponent on $\text{SRF} = 1/\alpha$, which depends on the cardinality of each clump as opposed to the total number of points. Let λ be the cardinality of the largest clump: $\lambda = \max_{a=1}^A \lambda_a$. Theorem 1 implies

$$\sigma_{\min}(\Phi_M) \geq C\sqrt{M}\ \text{SRF}^{-\lambda+1}. \qquad \text{(II.3)}$$

Previous results [1], [2] strongly suggest (we avoid using "imply" because they studied a similar inverse problem but with continuous, rather than discrete measurements like the ones considered here) that

$$\sigma_{\min}(\Phi_M) \geq C\sqrt{M}\ \text{SRF}^{-S+1}. \qquad \text{(II.4)}$$

By comparing the inequalities (II.3) and (II.4), we see that our lower bound is dramatically better when all of the point sources are not located within a single clump. These results are also consistent with our intuition that $\sigma_{\min}(\Phi_M)$ is smallest when Ω consists of S closely spaced points; more details about this can be found in [13]. In [16], a lower bound of $\sigma_{\min}(\Phi_M)$ is derived for a model called clustered nodes; a detail comparison between Theorem 1 and results in [16] can be found in [13].

The following theorem provides an upper bound on $\sigma_{\min}(\Phi_M)$ when Ω contains λ consecutive points spaced by α/M, and this shows that the dependence on SRF in inequality (II.3) is optimal.

Theorem 2. *Let $\lambda \le S \le M-1$, and there exists a constant $c>0$ depending only on λ such that the following hold: for any $0<\alpha\le c(M+1)^{-1/2}$, $\omega\in\mathbb{T}$ and $\Omega\subseteq\mathbb{T}$ of cardinality S that contains the set $\omega+\{0,\alpha/M,\ldots,(\lambda-1)\alpha/M\}$, we have $\sigma_{\min}(\Phi_M)\le C_\lambda \alpha^{\lambda-1}$.*

III. MUSIC AND ITS SUPER-RESOLUTION LIMIT

In signal processing, the MUSIC algorithm [12], has been widely used due to its superior numerical performance among subspace methods. MUSIC relies upon the Vandermonde decomposition of a Hankel matrix, and its stability to noise can be formulated as a matrix perturbation problem.

Throughout the following exposition, we assume that L is an integer satisfying the inequalities $S\le L\le M+1-S$. The Hankel matrix of y is

$$\mathcal{H}(y)=\begin{bmatrix} y_0 & y_1 & \cdots & y_{M-L}\\ \vdots & \vdots & \ddots & \vdots\\ y_L & y_{L+1} & \cdots & y_M\end{bmatrix}.$$

If we denote the noiseless measurement vector by $y^0=\Phi_M(\Omega)x$, then it is straightforward to verify that we have the following Vandermonde decomposition

$$\mathcal{H}(y^0)=\Phi_L \mathrm{diag}(x_1,\ldots,x_S)\Phi_{M-L}^T.$$

Observe that both Φ_L and Φ_{M-L} have full column rank when $S\le L\le M+1-S$ and that $H(y^0)$ has rank S. The Singular Value Decomposition (SVD) of $\mathcal{H}(y^0)$ is of the form

$$\mathcal{H}(y^0)=[U\ W]\ \mathrm{diag}(\sigma_1,\ldots,\sigma_S,0,\ldots,0)V^*,$$

where $\sigma_1\ge\ldots\ge\sigma_S$ are the non-zero singular values of $\mathcal{H}(y^0)$. The columns of $U\in\mathbb{C}^{(L+1)\times S}$ and $W\in\mathbb{C}^{(L+1)\times(L+1-S)}$ span $\mathrm{Range}(\mathcal{H}(y^0))$ and $\mathrm{Range}(\mathcal{H}(y^0))^\perp$ respectively, which are called the signal space and the noise space.

For any $\omega\in\mathbb{T}$ and positive integer L, we define the *steering vector* of length $L+1$ to be

$$\phi_L(\omega)=[1\ e^{-2\pi i\omega}\ e^{-2\pi i2\omega}\ \ldots\ e^{-2\pi iL\omega}]^T.$$

Algorithm 1 MUltiple SIgnal Classification

Input: $y\in\mathbb{C}^{M+1}$, sparsity S, L.

1: Form Hankel matrix $\mathcal{H}(y)\in\mathbb{C}^{(L+1)\times(M-L+1)}$

2: Compute the SVD of $\mathcal{H}(y)$:

$$\mathcal{H}(y)=[\widehat{U}\ \widehat{W}]\mathrm{diag}(\widehat{\sigma}_1,\ldots,\widehat{\sigma}_S,\widehat{\sigma}_{S+1},\ldots)\widehat{V}^*,$$

where $\widehat{U}\in\mathbb{C}^{(L+1)\times S}$, $\widehat{W}\in\mathbb{C}^{(L+1)\times(L+1-S)}$.

3: Compute the imaging function $\widehat{\mathcal{J}}(\omega)=\|\phi_L(\omega)\|_2/\|\widehat{W}^*\phi_L(\omega)\|_2,\ \omega\in[0,1)$.

Output: $\widehat{\Omega}=\{\widehat{\omega}_j\}_{j=1}^S$ corresponding to the S largest local maxima of $\widehat{\mathcal{J}}$.

MUSIC is based on the following observation that

$$\omega\in\Omega \text{ iff } \phi_L(\omega)\in\mathrm{Range}(\mathcal{H}(y^0))=\mathrm{Range}(U).$$

Table I: Functions in the MUSIC algorithm

	Noise-space correlation function	Imaging function
Noiseless	$\mathcal{R}(\omega)=\frac{\|W^*\phi_L(\omega)\|_2}{\|\phi_L(\omega)\|_2}$	$\mathcal{J}(\omega)=\frac{1}{\mathcal{R}(\omega)}$
Noisy	$\widehat{\mathcal{R}}(\omega)=\frac{\|\widehat{W}^*\phi_L(\omega)\|_2}{\|\phi_L(\omega)\|_2}$	$\widehat{\mathcal{J}}(\omega)=\frac{1}{\widehat{\mathcal{R}}(\omega)}$

This observation can be reformulated in terms of the noise-space correlation function $\mathcal{R}(\omega)$ and the imaging function $\mathcal{J}(\omega)$ (see Table I for their definitions), as summarized in the following lemma.

Lemma 1. *Let $S\le L\le M+1-S$. Then*

$$\omega\in\{\omega_j\}_{j=1}^S \Longleftrightarrow \mathcal{R}(\omega)=0 \Longleftrightarrow \mathcal{J}(\omega)=\infty.$$

To summarize this discussion: in the noiseless case where we have access to y^0, the source locations can be exactly identified through the zeros of the noise-space correlation function $\mathcal{R}(\omega)$ or the peaks of the imaging function $\mathcal{J}(\omega)$.

In the presence of noise, we only have access to $\mathcal{H}(y)$, which is a perturbation of $H(y^0)$:

$$\mathcal{H}(y)=\mathcal{H}(y^0)+\mathcal{H}(\eta).$$

The noise-space correlation and imaging functions are perturbed to $\widehat{\mathcal{R}}(\omega)$ and $\widehat{\mathcal{J}}(\omega)$ respectively. Stability of MUSIC depends on the perturbation of the noise-space correlation function from $\mathcal{R}(\omega)$ to $\widehat{\mathcal{R}}(\omega)$ which we measure by

$$\|\widehat{\mathcal{R}}-\mathcal{R}\|_\infty := \max_{\omega\in[0,1)}|\widehat{\mathcal{R}}(\omega)-\mathcal{R}(\omega)|.$$

By using Wedin's theorem [17], [18, Theorem 3.4], we can prove the following perturbation bound.

Proposition 1. *Let $S \leq L \leq M + 1 - S$. Suppose $2\|\mathcal{H}(\eta)\|_2 < x_{\min}\sigma_{\min}(\Phi_L)\sigma_{\min}(\Phi_{M-L})$. Then*

$$\|\widehat{\mathcal{R}} - \mathcal{R}\|_\infty \leq \frac{2\|\mathcal{H}(\eta)\|_2}{x_{\min}\sigma_{\min}(\Phi_L)\sigma_{\min}(\Phi_{M-L})}.$$

If η is independent Gaussian noise, i.e., $\eta \sim \mathcal{N}(0, \sigma^2 I)$, the spectral norm of $\mathcal{H}(\eta)$ satisfies the following concentration inequality [19, Theorem 4]:

Lemma 2. *If $\eta \sim \mathcal{N}(0, \sigma^2 I)$, then*

$$\mathbb{E}\|\mathcal{H}(\eta)\|_2 \leq \sigma\sqrt{2C(M,L)\log(M+2)},$$

$$\mathbb{P}\left\{\|\mathcal{H}(\eta)\|_2 \geq t\right\} \leq (M+2)\exp\left(-\frac{t^2}{2\sigma^2 C(M,L)}\right),$$

for $t > 0$, and $C(M,L) = \max(L+1, M-L+1)$.

Combining Proposition 1, Lemma 2 and Theorem 1 gives rise to a stability analysis of MUSIC:

Theorem 3. *Let M be an even integer satisfying $M \geq 2S^2$ and set $L = M/2$. Fix parameters $\varepsilon > 0$, $\nu > 1$, and let $\eta \sim \mathcal{N}(0, \sigma^2 I)$. Assume Ω satisfies Assumption 1 with parameters (L, S, α, β) for some $\alpha > 0$ and β satisfying* (II.1). *There exist explicit constants $c_a > 0$ such that if*

$$\frac{\sigma}{x_{\min}} < C(M,\nu)\Big(\sum_{a=1}^{A} c_a^2 \alpha^{-2(\lambda_a - 1)}\Big)^{-1}\varepsilon,$$

$$C(M,\nu) = \frac{M}{32\sqrt{\nu(M+2)\log(M+2)}},$$

then with probability no less than $1-(M+2)^{-(\nu-1)}$,

$$\|\widehat{\mathcal{R}} - \mathcal{R}\|_\infty \leq \varepsilon.$$

In order to guarantee an ε-perturbation of the noise-space correlation function, the noise-to-signal ratio should follow the scaling law

$$\frac{\sigma}{x_{\min}} \propto \sqrt{\frac{M}{\log M}}\left(\sum_{a=1}^{A} c_a^2 \alpha^{-2(\lambda_a - 1)}\right)^{-1}\varepsilon.$$

Let λ be the cardinality of the largest clump. By (II.3), this scaling law reduces to

$$\frac{\sigma}{x_{\min}} \propto \sqrt{\frac{M}{\log M}}\alpha^{2\lambda - 2}\varepsilon = \sqrt{\frac{M}{\log M}}\,\mathrm{SRF}^{-(2\lambda-2)}\varepsilon.$$

The resolution limit of MUSIC is exponential in SRF, but the exponent only depends on the cardinality of the separated clumps instead of the total sparsity S. These estimates are verified by numerical experiments in [13].

Acknowledgment

Wenjing Liao is supported by NSF-DMS-1818751.

References

[1] D. L. Donoho, "Superresolution via sparsity constraints," *SIAM Journal on Mathematical Analysis*, vol. 23, no. 5, pp. 1309–1331, 1992.

[2] L. Demanet and N. Nguyen, "The recoverability limit for superresolution via sparsity," *arXiv preprint arXiv:1502.01385*, 2015.

[3] E. J. Candès and C. Fernandez-Granda, "Super-resolution from noisy data," *Journal of Fourier Analysis and Applications*, vol. 19, no. 6, pp. 1229–1254, 2013.

[4] V. Duval and G. Peyré, "Exact support recovery for sparse spikes deconvolution," *Foundations of Computational Mathematics*, vol. 15, no. 5, pp. 1315–1355, 2015.

[5] W. Li, "Elementary L^∞ error estimates for super-resolution de-noising," *arXiv preprint arXiv:1702.03021*, 2017.

[6] A. C. Fannjiang and W. Liao, "Coherence-pattern guided compressive sensing with unresolved grids," *SIAM Journal on Imaging Sciences*, vol. 5, no. 1, pp. 179–202, 2012.

[7] W. Liao and A. Fannjiang, "MUSIC for single-snapshot spectral estimation: Stability and super-resolution," *Applied and Computational Harmonic Analysis*, vol. 40, no. 1, pp. 33–67, 2016.

[8] A. Moitra, "Super-resolution, extremal functions and the condition number of Vandermonde matrices," *Proceedings of the Forty-Seventh Annual ACM Symposium on Theory of Computing*, 2015.

[9] V. I. Morgenshtern and E. J. Candes, "Super-resolution of positive sources: The discrete setup," *SIAM Journal on Imaging Sciences*, vol. 9, no. 1, pp. 412–444, 2016.

[10] Q. Denoyelle, V. Duval, and G. Peyré, "Support recovery for sparse super-resolution of positive measures," *Journal of Fourier Analysis and Applications*, vol. 23, no. 5, pp. 1153–1194, 2017.

[11] J. J. Benedetto and W. Li, "Super-resolution by means of Beurling minimal extrapolation," *Applied and Computational Harmonic Analysis*, 2018.

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

[13] W. Li and W. Liao, "Stable super-resolution limit and smallest singular value of restricted fourier matrices," *arXiv preprint arXiv:1709.03146*, 2017.

[14] H. L. Montgomery and R. C. Vaughan, "Hilbert's inequality," *Journal of the London Mathematical Society*, vol. 2, no. 1, pp. 73–82, 1974.

[15] J. D. Vaaler, "Some extremal functions in Fourier analysis," *Bulletin of the American Mathematical Society*, vol. 12, no. 2, pp. 183–216, 1985.

[16] D. Batenkov, L. Demanet, G. Goldman, and Y. Yomdin, "Stability of partial Fourier matrices with clustered nodes," *arXiv preprint arXiv:1809.00658*, 2018.

[17] P.-Å. Wedin, "Perturbation bounds in connection with singular value decomposition," *BIT Numerical Mathematics*, vol. 12, no. 1, pp. 99–111, 1972.

[18] R.-C. Li, "Relative perturbation theory: II. Eigenspace and singular subspace variations," *SIAM Journal on Matrix Analysis and Applications*, vol. 20, no. 2, pp. 471–492, 1998.

[19] W. Liao, "Music for multidimensional spectral estimation: stability and super-resolution," *IEEE Transactions on Signal Processing*, vol. 63, no. 23, pp. 6395–6406, 2015.