

SINGLE-EXPONENTIAL BOUNDS FOR THE SMALLEST SINGULAR VALUE OF VANDERMONDE MATRICES IN THE SUB-RAYLEIGH REGIME

DMITRY BATENKOV AND GIL GOLDMAN

ABSTRACT. Following recent interest by the community, the scaling of the minimal singular value of a Vandermonde matrix with nodes forming clusters on the length scale of Rayleigh distance on the complex unit circle is studied. Using approximation theoretic properties of exponential sums, we show that the decay is only single exponential in the size of the largest cluster, and the bound holds for arbitrary small minimal separation distance. We also obtain a generalization of well-known bounds on the smallest eigenvalue of the generalized prolate matrix in the multi-cluster geometry. Finally, the results are extended to the entire spectrum.

1. INTRODUCTION

For an ordered set of distinct nodes $\mathcal{X} = \{x_1, \dots, x_s\}$ with $x_j \in (-\pi, \pi]$, and $N \geq s - 1$, consider the $(N + 1) \times s$ Vandermonde matrix

$$\mathbf{V}_N(\mathcal{X}) := \left[e^{ikx_j} \right]_{\substack{1 \leq j \leq s \\ 0 \leq k \leq N}}. \quad (1.1)$$

This class of matrices is the subject of numerous recent investigations in the applied harmonic analysis community, e.g. [1, 4, 5, 14, 17, 18, 19, 20, 21, 22]. While interesting in their own right, the spectral properties of \mathbf{V}_N are closely related to the problem of super-resolution (SR) under sparsity constraints, which also received a lot of attention in recent years [6, 8, 9, 11]. In the SR context, the smallest singular value $\sigma_{\min}(\mathbf{V}_N) := \min_{\mathbf{c} \in \mathbb{C}^s, \|\mathbf{c}\|_2=1} \|\mathbf{V}_N \mathbf{c}\|_2$ controls the limit of stable recovery of a superposition of Dirac masses supported on \mathcal{X} from its first $N + 1$ Fourier coefficients, while the singular subspaces play a major role in various SR algorithms (e.g. MUSIC and ESPRIT) [13, 19, 20, 21].

Let Δ denote the minimal separation (in the wrap-around sense) between the elements of \mathcal{X} . With s fixed, two distinct asymptotic regimes are known:

- (1) When $N\Delta \gtrsim O(1)$, the matrix \mathbf{V}_N is well-conditioned, and $\sigma_{\min}(\mathbf{V}_N) = O(\sqrt{N})$.
- (2) When $N\Delta \ll 1$, $\sigma_{\min}(\mathbf{V}_N)$ can be as small as $O(\sqrt{N}(N\Delta)^{s-1})$.

The well-conditioned case 1) has been studied in [1, 16, 22, 23, 25], by various tools from harmonic analysis and analytic number theory. The separation condition $\Delta \gtrsim \frac{1}{N}$ plays a major role in the analysis of the convex relaxations of the SR problem [7, 8].

Case 2) corresponds to the so-called ‘‘sub-Rayleigh’’ regime, where $N = 2\pi\Delta^{-1}$ is precisely the Rayleigh resolution limit. The possibility to resolve closely spaced point sources from low-frequency measurements with arbitrary precision was already established by G. de Prony in 1795 [26]¹, providing the symbolic-algebraic basis for many other reconstruction algorithms that followed. However,

2010 *Mathematics Subject Classification*. Primary 15A18, 65T40, 65F20.

Key words and phrases. Vandermonde matrices with nodes on the unit circle, nonuniform Fourier matrices, sub-Rayleigh resolution, singular values, super-resolution, condition number.

D.B. was supported by the ISRAEL SCIENCE FOUNDATION (grant No. 1793/20).

¹English translation of the original Prony’s paper can be found in [2].

without additional prior information regarding the geometry of \mathcal{X} , the sensitivity to noise (“condition number”) of the SR problem in the sub-Rayleigh regime may be as large as SRF^{2s-1} , where $\text{SRF} := (N\Delta)^{-1}$ is the “super-resolution factor”. This quickly becomes prohibitive already for moderate values of s . The exponent $2s - 1$ corresponds to the worst-case scenario where all the nodes of \mathcal{X} are clustered together and approximately equispaced, e.g. $x_{j+1} = x_j + \Delta_j$ with $\Delta_j \approx \Delta$, $j = 1, \dots, s - 1$.

When $\Delta_j = \Delta$ for all $1 \leq j \leq s - 1$, \mathbf{V}_N is a contiguous submatrix of the DFT matrix (also known as the “prolate matrix” in the literature [29]), and the scaling of $\sigma_{\min}(\mathbf{V}_N)$, in the asymptotic regime $N \rightarrow \infty$, $N\Delta < c$ directly follows from the “Bell Labs theory” of the spectral concentration problem [27] (see also [3]). See Section 3 for further discussion.

Now suppose that an a-priori information is available, according to which only a small number of nodes can be clustered, with the different clusters separated by $\theta \gtrsim \frac{1}{N}$ (see Definition 2.3 below). It has been recently shown by several groups that in this case, $\sigma_{\min}(\mathbf{V}_N) \asymp \sqrt{N}(N\Delta)^{\ell-1}$ where ℓ is the largest multiplicity of any such cluster. Accordingly, the SR condition number will scale as $\text{SRF}^{2\ell-1}$ [4, 5, 6, 10, 17, 18, 19].

The scaling of the proportionality constants in the order estimates above, in particular, their dependence on ℓ and s , are a subject of ongoing research. This question is of importance in the regime where $\ell \ll s$, so that the factor $\text{SRF}^{2\ell-1}$ is significantly smaller than SRF^{2s-1} .

In this paper we prove a single-exponential in ℓ and linear in s lower bound for $\sigma_{\min}(\mathbf{V}_N(\mathcal{X}))$ in the multi-cluster geometry (Theorem 2.2), of the form

$$\sigma_{\min}(\mathbf{V}_N(\mathcal{X})) \geq c_1 \sqrt{N}(N\Delta/c_2)^{\ell-1}, \quad (1.2)$$

where c_1, c_2 are absolute constants, independent of ℓ, s, N, Δ , (and in fact, $c_2 = 32\pi e$), holding whenever

$$c_3(\ell)s/\theta \leq N \leq c_4(\ell)/(\tau\Delta s). \quad (1.3)$$

Relative to prior works on the subject, in particular [18, 19] (see Section 3 below), our single-exponential in ℓ bound (1.2) holds for a fixed N and *all* sufficiently small Δ . Applying a simple limiting argument, in Theorem 2.3 we also generalize Slepian’s bound for the smallest eigenvalue of the prolate matrix (see above) in the non-equispaced multi-cluster case.

The main technical contribution of this paper is a new method of proof of the bound (1.2) for a single cluster (Theorem 2.1), which was previously shown in this setting in [18, Example 4.8] (again, see details in Section 3). The proof is based on applying the classical Turan’s inequality for exponential sums and Salem’s inequality to the analysis of stability of Vandermonde matrices with nodes on the unit circle. The extension of this result to Vandermonde matrices with multiple sets of clustered nodes separated by $\theta \gtrsim \frac{1}{N}$, is done by invoking our recent result [5, Theorem 2.2], which, in turn, shows that the column subspaces in \mathbb{C}^{N+1} corresponding to each cluster are nearly orthogonal.

Our results can easily be extended to show single-exponential scaling for *all the singular values* of \mathbf{V}_N (resp. eigenvalues of the prolate matrix). We present some details of these extensions in Section 7, however for the sake of brevity we do not provide the full derivations.

2. MAIN RESULTS

Definition 2.1 (Wrap-around distance). *For $x, y \in \mathbb{R}$, we define the wrap-around distance*

$$d(x, y) := |\text{Arg exp } \iota(x - y)| = |x - y \pmod{-\pi, \pi}| \in [0, \pi],$$

where for $z \in \mathbb{C} \setminus \{0\}$, $\text{Arg}(z)$ is the principal value of the argument of z , taking values in $(-\pi, \pi]$.

We denote by $\mathbb{T} = \mathbb{R} \pmod{-\pi, \pi}$ the periodic interval of length 2π .

Definition 2.2 (Single cluster configuration). *The node set $\mathcal{X} = \{x_1, \dots, x_\ell\} \subset (-\pi, \pi]$ is said to form a $(\Delta, \ell, \tau)_{\mathbb{T}}$ -cluster, for some $\ell - 1 \leq \tau \leq \frac{\pi}{\Delta}$, if*

$$\forall x, y \in \mathcal{X}, x \neq y : \quad \Delta \leq d(x, y) \leq \tau\Delta.$$

Below we write $C_k(\ell)$, for some indexes $k = 1, \dots$, to indicate constants that depend only on ℓ . Our first main result is the following.

Theorem 2.1. *Let \mathcal{X} form a $(\Delta, \ell, \tau)_{\mathbb{T}}$ -clustered configuration. Then there exist a constant $C_1(\ell)$ and absolute constants $C_2 = 32\pi e, C_3$, such that for any N satisfying $C_1(\ell) \leq N \leq \frac{2\pi}{\tau\Delta}$,*

$$\sigma_{\min}(\mathbf{V}_N(\mathcal{X})) \geq C_3 \sqrt{N} \left(\frac{N\Delta}{C_2} \right)^{\ell-1}. \quad (2.1)$$

Definition 2.3 (Multi-cluster configuration, periodic interval). *The node set $\mathcal{X} = \{x_1, \dots, x_s\} \subset (-\pi, \pi]$ is said to form a $(\Delta, \theta, s, \ell, \tau)_{\mathbb{T}}$ -clustered configuration for some $\Delta > 0, 1 \leq \ell \leq s, \ell - 1 \leq \tau \leq \frac{\pi}{\Delta}$ and $\theta > 0$, if for each x_j , there exist at most ℓ distinct nodes*

$$\mathcal{X}^{(j)} = \{x_{j,k}\}_{k=1, \dots, r_j} \subset \mathcal{X}, \quad 1 \leq r_j \leq \ell, \quad x_{j,1} \equiv x_j,$$

such that the following conditions are satisfied:

(1) *For any $y \in \mathcal{X}^{(j)} \setminus \{x_j\}$, we have*

$$\Delta \leq d(y, x_j) \leq \tau\Delta.$$

(2) *For any $y \in \mathcal{X} \setminus \mathcal{X}^{(j)}$, we have*

$$d(y, x_j) \geq \theta.$$

Theorem 2.2. *There exist constants $C_4(\ell), C_5(\ell)$ and absolute constants $C_6 = 32\pi e, C_7$, such that for any \mathcal{X} forming a $(\Delta, \theta, s, \ell, \tau)_{\mathbb{T}}$ -clustered configuration and N satisfying $\frac{C_4 s}{\theta} \leq N \leq \frac{C_5}{s\tau\Delta}$,*

$$\sigma_{\min}(\mathbf{V}_N(\mathcal{X})) \geq C_7 \sqrt{N} \left(\frac{N\Delta}{C_6} \right)^{\ell-1}. \quad (2.2)$$

Definition 2.4 (Generalized prolate matrix). *Let $\mathcal{X} = \{x_1, \dots, x_s\} \subset \mathbb{R}$ be a collection of s pairwise distinct points on the real line. We define the generalized prolate matrix as follows:*

$$\mathbf{G}(\mathcal{X}) = \left[\frac{1}{2} \int_{-1}^1 e^{i\omega(x_j - x_k)} \right]_{j,k=1}^s \in \mathbb{R}^{s \times s}.$$

Note that $\mathbf{G}(\mathcal{X})$ is symmetric and positive definite (see e.g. [4, Proposition 2.6]).

In a manner completely analogous to Definition 2.2, we define the notion of a clustered configuration appropriate for this setting.

Definition 2.5 (Multi-cluster configuration, real line). *The node set $\mathcal{X} = \{x_1, \dots, x_s\} \subset \mathbb{R}$ is said to form a $(\Delta, \theta, s, \ell, \tau)_{\mathbb{R}}$ -clustered configuration for some $\Delta > 0, 1 \leq \ell \leq s, \ell - 1 \leq \tau$ and $\theta > 0$, if for each x_j , there exist at most ℓ distinct nodes*

$$\mathcal{X}^{(j)} = \{x_{j,k}\}_{k=1, \dots, r_j} \subset \mathcal{X}, \quad 1 \leq r_j \leq \ell, \quad x_{j,1} \equiv x_j,$$

such that the following conditions are satisfied:

(1) *For any $y \in \mathcal{X}^{(j)} \setminus \{x_j\}$, we have*

$$\Delta \leq |y - x_j| \leq \tau\Delta.$$

(2) *For any $y \in \mathcal{X} \setminus \mathcal{X}^{(j)}$, we have*

$$|y - x_j| \geq \theta.$$

The next theorem is a direct corollary of Theorem 2.2, and it should be compared to (3.1).

Theorem 2.3. *There exist absolute constants C_8 , $C_9 = 16\pi e$ and constants $C_{10}(\ell)$, $C_{11}(\ell)$, such that for any \mathcal{X} forming a $(\Delta, \theta, s, \ell, \tau)_{\mathbb{R}}$ -clustered configuration with $\theta \geq C_{10}s$ and $s\tau\Delta \leq C_{11}$,*

$$\lambda_{\min}(\mathbf{G}(\mathcal{X})) \geq C_8 \left(\frac{\Delta}{C_9} \right)^{2(\ell-1)}. \quad (3.2)$$

Remark 2.1. *Frequently the definition of the prolate matrix contains an additional bandwidth parameter $\Omega > 0$, so that the inner products are considered in an interval $[-\Omega, \Omega]$ [4]. In the present paper we do not lose any generality by rescaling Ω to 1.*

3. PRIOR ART

In this section only, $c, c_1, \dots, c', \dots, C, \dots$ denote generic constants which might be different in different formulas, and which do not depend on N, Δ .

Let us start with the setting $\ell = s$. Recalling Definition 2.4, we have, as $N \rightarrow \infty$, that

$$(2N)^{-1} \sigma_{\min}^2(\mathbf{V}_N(\mathcal{X}/N)) \rightarrow \lambda_{\min}(\mathbf{G}(\mathcal{X})).$$

In the equispaced setting $x_{j+1} = x_j + \Delta$, $j = 1, \dots, s-1$, the matrix \mathbf{G} is precisely the ‘‘prolate matrix’’ [27, 29], and it holds that

$$\begin{aligned} \lambda_{\min}(\mathbf{G}(\mathcal{X})) &= C_{EQ}(s) \Delta^{2s-2} \{1 + O(\Delta)\}, \quad \Delta \ll 1; \\ C_{EQ}(s) &:= \frac{2^{2s-2}}{(2s-1) \binom{2s-2}{s-1}^3} \asymp_s \left(\frac{1}{4} \right)^{2s-2}. \end{aligned} \quad (3.1)$$

Here \asymp_s means ‘‘up to polynomial in s and $1/s$ factors’’. This gives

$$\sigma_{\min}(\mathbf{V}_N) \asymp_s \sqrt{N} \left(\frac{N\Delta}{4} \right)^{s-1} \left\{ 1 + O(\sqrt{N\Delta}) \right\}.$$

Non-asymptotic bounds for the case of node configurations \mathcal{X} with minimal separation of at least Δ are available as well. An explicit construction in [19, Proposition 3] (following [11]) gives

$$\min_{\mathcal{X}} \sigma_{\min}(\mathbf{V}_N(\mathcal{X})) \lesssim_s \sqrt{N} \left(\frac{N\Delta}{2} \right)^{s-1}, \quad N\Delta < \frac{2\pi}{C_{LL}(s)\sqrt{N}}, \quad C_{LL}(s) = 2\pi \sum_{j=0}^{s-1} \binom{s-1}{j} \frac{j^s}{s!}.$$

For a single cluster setting $\ell = s$, Theorem 2.1 has been proven in [18, Example 4.8] with the better constant $C_2 = 4\pi e$ (in the earlier work [19] this constant was not explicit). The reduction in the tightness of constant in our work might be related to the fact that our constant is also valid for all the singular values, see Theorem 7.1 below.

Turning to the more general case $\ell \leq s$ and cluster separation θ (as in Definition 2.3), in [19] and later [18, Corollary 4.2] it was shown that

$$\sigma_{\min}(\mathbf{V}_N(\mathcal{X})) \geq \frac{5}{9} \sqrt{N} \left(\frac{N\Delta}{4\pi e} \right)^{\ell-1}, \quad N\Delta < 1.$$

However, the above holds under the condition $N\theta > c_\gamma (N\Delta)^{-\gamma}$ with any $\gamma > 0$ and $c_\gamma \rightarrow \infty$ as $\gamma \rightarrow 0$. For fixed N, θ , this prevents $\Delta \rightarrow 0$ in order for the bound to continue to hold. In contrast, assuming $N\theta > c'$ and $N\Delta < c''$ with c', c'' depending only on s, ℓ , it was shown in [5, Theorem 2.3, Corollary 2.1] that for all \mathcal{X} satisfying the clustering geometric assumptions, we have

$$C' \sqrt{N} (N\Delta)^{\ell-1} \leq \sigma_{\min}(\mathbf{V}_N(\mathcal{X})) \leq \frac{1}{2} \sqrt{N\ell e} (\tau N\Delta)^{\ell-1}. \quad (3.2)$$

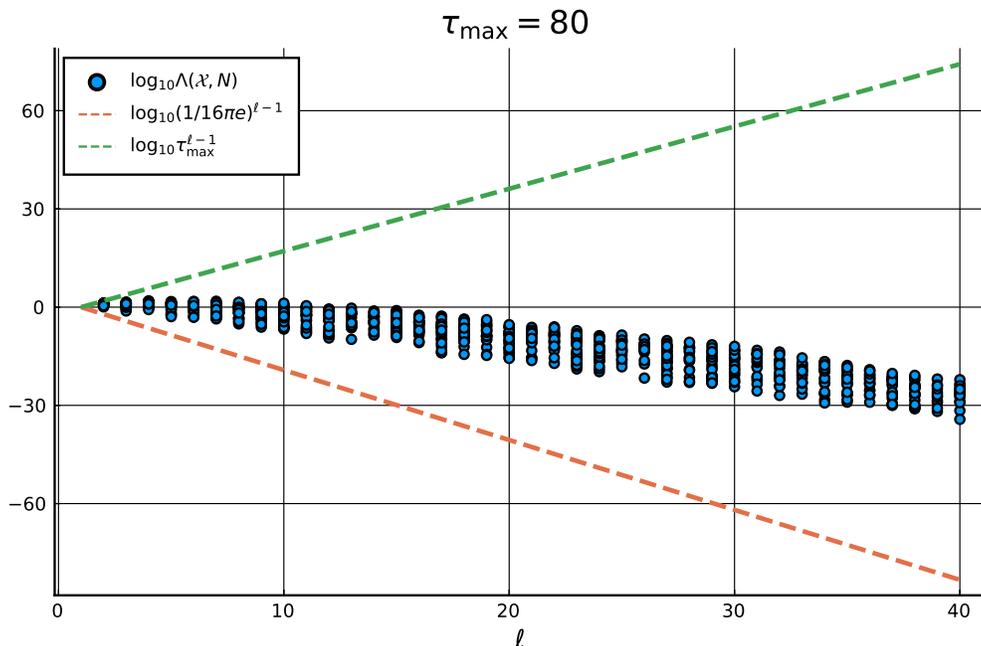


FIGURE 1. Single cluster - dependence of σ_{\min} on ℓ . We plot ℓ vs $\log_{10} \Lambda(\mathcal{X}, N)$. N is varying between 60 and 300, Δ is varying between 10^{-25} and 10^{-4} , while ℓ varies from 2 to 40 and τ varies from $\ell - 1$ to $\tau_{\max} = 80$. The lower and upper bounds are shown as dashed lines.

Here $\tau \geq (\ell - 1)$ is a uniformity parameter, controlling the overall extent of any cluster (see Definition 2.3). However, the constant C' was not explicit. In [4] the same scaling for the lower bound was shown with $C' = \frac{\sqrt{\pi/2}}{(s\sqrt{2\pi})^{2s-1}}$, albeit under the additional assumption that $x_j \in \frac{\pi}{2s^2}(-1, 1]$ for all $j = 1, \dots, s$.

More details on the above developments are available in [5, Section 1.4], [18, Examples 4.7,4.8], [4, Remarks 3.5,3.7] and [19, Remark 4].

Our single-exponential in ℓ bound of Theorem 2.2, which holds for a fixed N and *all* sufficiently small Δ , thus provides an improvement upon the above mentioned results (except, possibly, for the value of the absolute constant C_6 , see Section 4).

4. NUMERICAL EXPERIMENTS

In this section we estimate the exponential dependence of $\sigma_{\min}(\mathbf{V}_N(\mathcal{X}))$ on ℓ numerically, by computing

$$\Lambda(\mathcal{X}, N) := \sigma_{\min}(\mathbf{V}_N(\mathcal{X})) / (\sqrt{N}(N\Delta)^{\ell-1}).$$

Varying ℓ, τ and N, Δ fixed, we expect that

$$(1/c_2)^{\ell-1} \lesssim \Lambda(\mathcal{X}, N) \lesssim (1/c'_2)^{\ell-1}.$$

Our theoretical results indicate that the above holds with $c_2 \leq 16\pi e$ and $c'_2 \geq 1/\tau$ (see resp. (6.10) and (3.2)).

As can be seen from Fig. 1, both the upper and lower bounds are correct, although the corresponding constants $16\pi e$ and $1/\tau$ are not tight.

All numerical tests were performed in arbitrary precision arithmetic.

5. DISCUSSION

It is an interesting open question whether a bound of the type (1.2) should hold in the multi-cluster geometry, for $N\theta > c'$, $N\Delta < c''$ where c', c'' do not depend on s and with no essential further restrictions on \mathcal{X} . If this is the case, then it is plausible that the super-resolution problem for a practically infinite spike train ($s \gg 1$) with small sub-Rayleigh clusters (a model analogous to Donoho's Rayleigh regular measures, [11]) can be essentially decoupled into treating each cluster separately.

There is a room for further refinement regarding the bounds themselves, as there is a relatively large gap in the constants between the upper bound in (3.2) and (1.2).

6. PROOFS

6.1. Preliminaries on exponential sums. We review some preliminary results about exponential sums and their implications to the problem at hand.

Definition 6.1. *Given a vector $\mathbf{c} \in \mathbb{C}^\ell$ and $\mathcal{X} = \{x_1, \dots, x_\ell\} \subset \mathbb{R}$, we define the exponential sum*

$$P(t) = P_{\mathbf{c}, \mathcal{X}}(t) := \sum_{j=1}^{\ell} c_j e^{itx_j}.$$

The number of nonzero c_j 's is called the degree of P . The set of all exponential sums of degree at most ℓ is denoted by \mathfrak{I}_ℓ .

Remark 6.1. *Our definition of exponential sums covers the case of purely imaginary exponents only to simplify the presentation. More general results for arbitrary complex exponents are available in e.g. [24, 12].*

We denote by μ the Lebesgue measure on \mathbb{R} .

Given an interval I and a (complex valued) continuous function $f \in C(I)$, $1 \leq p \leq \infty$, we denote

$$\|f\|_{L^p(I)} := \begin{cases} \left(\frac{1}{\mu(I)} \int_I |f(t)|^p dt \right)^{1/p}, & 1 \leq p < \infty; \\ \sup_{t \in I} |f(t)|, & p = \infty. \end{cases}$$

Exponential sums satisfy many classical inequalities from approximation theory. In particular, we have the following estimates.

Proposition 6.1 (Turan's inequality). *Let $P \in \mathfrak{I}_\ell$, and let $\Omega \subset I$ be intervals with positive Lebesgue measure. Then*

$$\|P\|_{L^\infty(I)} \leq \left(\frac{4e\mu(I)}{\mu(\Omega)} \right)^{\ell-1} \|P\|_{L^\infty(\Omega)}.$$

Proof. See a review on Turan's lemma on p.7 of [24] (and Turan's original result [28], in German). □

Proposition 6.2 (Nikolskii-type inequality). *Let $P \in \mathfrak{I}_\ell$, then*

$$\|P\|_{L^p[0,1]} \leq \left(\frac{\pi\ell}{2} \right)^{2/q-2/p} \|P\|_{L^q[0,1]}, \quad 0 < q < p \leq \infty, \quad q \leq 2.$$

Proof. This is Theorem 2.5 in [12]. □

Applying the above with $I = [0, \frac{4\pi}{\Delta}]$, $\Omega = [0, N]$, $p = \infty$ and $q = 2$ yields the following.

Corollary 6.1. *Let \mathcal{X} form an $(\Delta, \ell, \tau)_{\mathbb{T}}$ -clustered configuration and let $N \leq \frac{4\pi}{\Delta}$, then for any $\mathbf{c} \in \mathbb{C}^\ell$*

$$\|P_{\mathbf{c}, \mathcal{X}}\|_{L^2([0, N])} \geq \frac{2}{\pi\ell} \left(\frac{N\Delta}{16\pi e} \right)^{\ell-1} \|P_{\mathbf{c}, \mathcal{X}}\|_{L^2([0, \frac{4\pi}{\Delta}])}. \quad (6.1)$$

Proof. Indeed, we have

$$\|P_{\mathbf{c}, \mathcal{X}}\|_{L^2([0, \frac{4\pi}{\Delta}])} \leq \|P_{\mathbf{c}, \mathcal{X}}\|_{L^\infty([0, \frac{4\pi}{\Delta}])} \leq \left(\frac{4e \cdot 4\pi}{N\Delta} \right)^{\ell-1} \|P_{\mathbf{c}, \mathcal{X}}\|_{L^\infty[0, N]} \leq \frac{\pi\ell}{2} \left(\frac{16\pi e}{N\Delta} \right)^{\ell-1} \|P_{\mathbf{c}, \mathcal{X}}\|_{L^2[0, N]}.$$

□

Now consider an exponential sum $P_{\mathbf{c}, \mathcal{X}}$ where Δ is the minimal separation of the nodes in \mathcal{X} . The next result states that for intervals I with length of the order of $\frac{1}{\Delta}$ or more, the coefficients norm $\|\mathbf{c}\|_2$ and $\|P_{\mathbf{c}, \mathcal{X}}\|_{L^2(I)}$ are related by an absolute constant. This is contrary to the case when the length of I is smaller than $\frac{1}{\Delta}$, in which case the constant will depend on $\Delta \cdot \mu(I)$ and ℓ , as we will show below.

Proposition 6.3. *Let \mathcal{X} form an $(\Delta, \ell, \tau)_{\mathbb{T}}$ -clustered configuration and let $\mathbf{c} \in \mathbb{C}^\ell$. Then, there exists an absolute constant C_{12} such that*

$$\|P_{\mathbf{c}, \mathcal{X}}\|_{L^2([0, \frac{4\pi}{\Delta}])}^2 \geq C_{12} \|\mathbf{c}\|_2^2.$$

Proof. It directly follows from [30, Vol.I, Chapter V, Th. 9.1]² stating that for an interval I such that $\mu(I) = \frac{2\pi(1+\delta)}{\Delta}$, $\delta > 0$, there exists an absolute constant C such that

$$\|\mathbf{c}\|_2^2 \leq C(1 - \delta^{-1}) \|P_{\mathbf{c}, \mathcal{X}}\|_{L^2(I)}^2.$$

□

Finally we require the following Bernstein type inequality bounding the maximum absolute value of the derivative of an exponential sum on $[0, 1]$, by its maximum absolute value on $[0, 1]$. See proof in [12, Theorem 2.20]

Proposition 6.4. *Let $P_{\mathbf{c}, \mathcal{X}}$ be an exponential sum, $\mathbf{c} \in \mathbb{C}^\ell$ and $\mathcal{X} = \{x_1, \dots, x_\ell\} \subset \mathbb{R}$, then*

$$\|P'_{\mathbf{c}, \mathcal{X}}\|_{L^\infty([0, 1])} \leq C_{13} \left(108\ell^5 + \sum_{j=1}^{\ell} x_j^2 \right)^{\frac{1}{2}} \|P_{\mathbf{c}, \mathcal{X}}\|_{L^\infty([0, 1])}.$$

6.2. Proof of Theorem 2.1. Let \mathcal{X} form $(\Delta, \ell, \tau)_{\mathbb{T}}$ -clustered configuration as in Theorem 2.1 and assume, without loss of generality, that \mathcal{X} is centered around the origin, i.e. $\min_j x_j + \max_j x_j = 0$, which implies that $\mathcal{X} \subset [-\tau\Delta/2, \tau\Delta/2]$.

Let

$$\|P\|_{2, N} := \left(\sum_{k=0}^N |P(k)|^2 \right)^{1/2}.$$

Then

$$\sigma_{\min}(\mathbf{V}_N(\mathcal{X})) = \min_{\|\mathbf{c}\|_2=1} \|P_{\mathbf{c}, \mathcal{X}}\|_{2, N}. \quad (6.2)$$

²Nazarov calls this type of inequality Salem's Inequality, see [24, page 8].

Fix some $\mathbf{c} \in \mathbb{C}^\ell$ such that $\|\mathbf{c}\|_2 = 1$ and $N \leq \frac{2\pi}{7\Delta} \leq \frac{4\pi}{\Delta}$. Combining Corollary 6.1 and Proposition 6.3 we obtain

$$\|P_{\mathbf{c},\mathcal{X}}\|_{L^2([0,N])} \geq \frac{2}{\pi\ell} \left(\frac{N\Delta}{16\pi e}\right)^{\ell-1} \|P_{\mathbf{c},\mathcal{X}}\|_{L^2([0,\frac{4\pi}{\Delta}])} \geq \frac{2}{\pi\ell} \left(\frac{N\Delta}{16\pi e}\right)^{\ell-1} \sqrt{C_{12}}\|\mathbf{c}\|_2 = \frac{C_{14}}{\ell} \left(\frac{N\Delta}{16\pi e}\right)^{\ell-1}. \quad (6.3)$$

At this point, we ‘‘almost’’ have the required result, what is left is to relate $\|P\|_{2,N}$ and the norm $\|P_{\mathbf{c},\mathcal{X}}\|_{L^2([0,N])}$, as follows.

Define

$$Q_{\mathbf{c},\mathcal{X},N}(u) := \sum_{j=1}^{\ell} c_j e^{2Nu x_j}.$$

Then $Q_{\mathbf{c},\mathcal{X},N}(u) = P_{\mathbf{c},\mathcal{X}}(Nu)$ and

$$\|P_{\mathbf{c},\mathcal{X}}\|_{L^2([0,N])} = \|Q_{\mathbf{c},\mathcal{X},N}\|_{L^2([0,1])}.$$

Put

$$T_{\mathbf{c},\mathcal{X},N}(u) := Q_{\mathbf{c},\mathcal{X},N}(u)\bar{Q}_{\mathbf{c},\mathcal{X},N}(u) = |Q_{\mathbf{c},\mathcal{X},N}(u)|^2.$$

We have that

$$\begin{aligned} \|P_{\mathbf{c},\mathcal{X}}\|_{L^2([0,N])}^2 &= \|T_{\mathbf{c},\mathcal{X},N}\|_{L^1([0,1])} = \int_0^1 T_{\mathbf{c},\mathcal{X},N}(u) du, \\ \|P_{\mathbf{c},\mathcal{X}}\|_{2,N}^2 &= \sum_{k=0}^N T_{\mathbf{c},\mathcal{X},N}\left(\frac{k}{N}\right). \end{aligned} \quad (6.4)$$

$T_{\mathbf{c},\mathcal{X},N} = \sum_{j=1}^w b_j e^{iu\lambda_j}$ is an exponential sum of maximal degree $w := \ell^2 - \ell + 1$, with the frequencies satisfying $|\lambda_j| \leq \tau N\Delta$. Consequently by Proposition 6.4

$$\|T'_{\mathbf{c},\mathcal{X},N}\|_{L^\infty([0,1])} \leq C_{13}(108w^5 + w(\tau N\Delta)^2)^{1/2} \|T_{\mathbf{c},\mathcal{X},N}\|_{L^\infty([0,1])}. \quad (6.5)$$

Approximating the integral by a Riemann sum and using equation (6.5) we have

$$\begin{aligned} \left| \int_0^1 T_{\mathbf{c},\mathcal{X},N}(u) du - \frac{1}{N} \sum_{k=0}^N T_{\mathbf{c},\mathcal{X},N}\left(\frac{k}{N}\right) \right| &\leq \frac{1}{2N} \|T'_{\mathbf{c},\mathcal{X},N}\|_{L^\infty([0,1])} + \frac{1}{N} T_{\mathbf{c},\mathcal{X},N}(0) \\ &\leq \frac{1}{2N} C_{13}(108w^5 + w(\tau N\Delta)^2)^{1/2} \|T_{\mathbf{c},\mathcal{X},N}\|_{L^\infty([0,1])} \\ &\quad + \frac{1}{N} T_{\mathbf{c},\mathcal{X},N}(0) \\ &\leq \frac{1}{N} \left(\frac{1}{2} C_{13}(108w^5 + w(\tau N\Delta)^2)^{1/2} + 1 \right) \|T_{\mathbf{c},\mathcal{X},N}\|_{L^\infty([0,1])}. \end{aligned}$$

By assumption $N \leq \frac{2\pi}{7\Delta}$, therefore for an absolute constant C_{15}

$$\left| \int_0^1 T_{\mathbf{c},\mathcal{X},N}(u) du - \frac{1}{N} \sum_{k=0}^N T_{\mathbf{c},\mathcal{X},N}\left(\frac{k}{N}\right) \right| \leq \frac{C_{15}\ell^5}{N} \|T_{\mathbf{c},\mathcal{X},N}\|_{L^\infty([0,1])}. \quad (6.6)$$

In addition, using Proposition 6.2 with $p = \infty, q = 1$ and $n = w$, we have

$$\|T_{\mathbf{c},\mathcal{X},N}\|_{L^\infty([0,1])} \leq \left(\frac{\pi w}{2}\right)^2 \|T_{\mathbf{c},\mathcal{X},N}\|_{L^1([0,1])}. \quad (6.7)$$

Therefore, substituting (6.7) into (6.6), we obtain

$$\left| \|T_{\mathbf{c},\mathcal{X},N}\|_{L^1([0,1])} - \frac{1}{N} \sum_{k=0}^N T_{\mathbf{c},\mathcal{X},N} \left(\frac{k}{N} \right) \right| \leq \frac{C_{15}\ell^5}{N} \left(\frac{\pi w}{2} \right)^2 \|T_{\mathbf{c},\mathcal{X},N}\|_{L^1([0,1])} \leq C_{16} \frac{\ell^9}{N} \|T_{\mathbf{c},\mathcal{X},N}\|_{L^1([0,1])} \quad (6.8)$$

For $N \geq 2C_{16}\ell^9$, we get from (6.8) that

$$\frac{1}{N} \sum_{k=0}^N T_{\mathbf{c},\mathcal{X},N} \left(\frac{k}{N} \right) \geq \frac{1}{2} \|T_{\mathbf{c},\mathcal{X},N}\|_{L^1([0,1])}.$$

By (6.4) we conclude that

$$\|P_{\mathbf{c},\mathcal{X}}\|_{2,N}^2 \geq \frac{N}{2} \|P_{\mathbf{c},\mathcal{X}}\|_{L^2([0,N])}^2. \quad (6.9)$$

Finally substituting (6.9) into (6.3) we get that

$$\|P_{\mathbf{c},\mathcal{X}}\|_{2,N} \geq \frac{C_{14}}{\ell} \sqrt{\frac{N}{2}} \left(\frac{N\Delta}{16\pi e} \right)^{\ell-1}. \quad (6.10)$$

Note that for all $\ell \geq 1$ we have $2^{\ell-1} \geq \ell$. Since \mathbf{c} was arbitrary, using the relation (6.2) completes the proof of Theorem 2.1 with $C_1(\ell) = 2C_{16}\ell^9$, $C_2 = 32\pi e$ and $C_3 = C_{14}/\sqrt{2}$. \square

6.3. Proof of Theorem 2.2. Let \mathcal{X} form a $(\Delta, \theta, s, \ell, \tau)_{\mathbb{T}}$ -clustered configuration. Then there exists an M -partition $\mathcal{X} = \bigsqcup_{j=1}^M \mathcal{C}^{(j)}$ such that for each $j \in \{1, \dots, M\}$:

- (1) $\mathcal{C}^{(j)}$ form a $(\Delta, \ell^{(j)}, \tau)_{\mathbb{T}}$ -cluster according to Definition 2.2, where $\ell^{(j)} \leq \ell$;
- (2) $d(x, y) \geq \theta$, $\forall x \in \mathcal{C}^{(j)}, \forall y \in \mathcal{X} \setminus \mathcal{C}^{(j)}$.

By (2.1), we have that for each $j = 1, \dots, M$

$$\sigma_{\min}(\mathbf{V}_N(\mathcal{C}^{(j)})) \geq C_3 \sqrt{N} \left(\frac{N\Delta}{C_2} \right)^{\ell-1}, \quad C_1(\ell) \leq N \leq \frac{2\pi}{\tau\Delta}. \quad (6.11)$$

We now apply Theorem 2.2 in [5], whose reduced version reads as follows.

Proposition 6.5. *Let \mathcal{X} form a $(\Delta, \theta, s, \ell, \tau)_{\mathbb{T}}$ -clustered configuration. Then there exist constants $C_{17}(\ell), C_{18}(\ell)$, depending only on ℓ , such that whenever*

$$\frac{C_{17}s}{\theta} \leq N \leq \frac{C_{18}}{s\tau\Delta}, \quad (6.12)$$

we have

$$\sigma_{\min}(\mathbf{V}_N(\mathcal{X})) \geq \frac{1}{2} \min_{j=1, \dots, M} \sigma_{\min}(\mathbf{V}_N(\mathcal{C}^{(j)})).$$

Since $\theta \leq \pi$ and $s \geq 1$, both lower bounds on N in (6.11) and (6.12) are satisfied whenever $N\theta \geq s \max(C_{17}(\ell)\pi, C_{17})$. On the other hand, $N\tau\Delta < \min(2\pi, C_{18})$ implies the corresponding upper bounds on N . This completes the proof of Theorem 2.2 with $C_7 = C_3/2$, $C_4(\ell) = \max(C_{17}, C_{17}(\ell)\pi)$, $C_5(\ell) = \min(2\pi, C_{18}(\ell))$ and $C_6 = C_2 = 32\pi e$. \square

6.4. Proof of Theorem 2.3.

Proof. Let $\theta \geq \frac{C_4 s}{2}$ and $s\tau\Delta \leq \frac{C_5}{2}$, where the constants $C_4 = C_4(\ell)$, $C_5 = C_5(\ell)$, are the same as in Theorem 2.2. Now let \mathcal{X} form a $(\Delta, \theta, s, \ell, \tau)_{\mathbb{R}}$ -clustered configuration.

For any $N \in \mathbb{N}$ and $j \in \{1, \dots, s\}$ put $\xi_j = \xi_{j,N} := \frac{x_j}{N}$ and $\boldsymbol{\xi} = \boldsymbol{\xi}^{(N)} := \{\xi_1, \dots, \xi_s\}$, and define the following shifted in frequency and normalized Vandermonde like matrix

$$\tilde{\mathbf{V}}_N(\boldsymbol{\xi}) := \frac{1}{\sqrt{2N}} [\exp(ik\xi_j)]_{k=-N, \dots, N}^{j=1, \dots, s}.$$

We have

$$\frac{1}{2} \int_{-1}^1 \exp(i\omega t) d\omega = \lim_{N \rightarrow \infty} \frac{1}{2N} \sum_{k=-N}^N \exp\left(i \frac{k}{N} t\right).$$

Consequently $\mathbf{G}(\mathcal{X}) = \lim_{N \rightarrow \infty} \tilde{\mathbf{V}}_N(\boldsymbol{\xi})^H \tilde{\mathbf{V}}_N(\boldsymbol{\xi})$, and so by continuity of eigenvalues [15, Section 2.4.9] we have that

$$\lambda_{\min}(\mathbf{G}(\mathcal{X})) = \lim_{N \rightarrow \infty} \lambda_{\min}(\tilde{\mathbf{V}}_N(\boldsymbol{\xi})^H \tilde{\mathbf{V}}_N(\boldsymbol{\xi})) = \lim_{N \rightarrow \infty} \sigma_{\min}^2(\tilde{\mathbf{V}}_N(\boldsymbol{\xi})). \quad (6.13)$$

For N large enough we have $\{\xi_{1,N}, \dots, \xi_{s,N}\} \subset (-\pi, \pi]$ and we can write $\tilde{\mathbf{V}}_N(\boldsymbol{\xi})$ as

$$\tilde{\mathbf{V}}_N(\boldsymbol{\xi}) = \frac{1}{\sqrt{2N}} \mathbf{V}_{2N}(\boldsymbol{\xi}) \cdot \text{diag}\left(e^{-iN\xi_1}, \dots, e^{-iN\xi_s}\right), \quad (6.14)$$

where $\text{diag}(e^{-iN\xi_1}, \dots, e^{-iN\xi_s})$ is the $s \times s$ diagonal matrix with $(e^{-iN\xi_1}, \dots, e^{-iN\xi_s})$ as its main diagonal. By (6.14) clearly

$$\sigma_{\min}(\tilde{\mathbf{V}}_N(\boldsymbol{\xi})) = \frac{1}{\sqrt{2N}} \sigma_{\min}(\mathbf{V}_{2N}(\boldsymbol{\xi})). \quad (6.15)$$

One can validate that for each N , $\{\xi_{j,N}\}_{j=1}^s$ form a $(\frac{\Delta}{N}, \frac{\theta}{N}, s, \ell, \tau)_{\mathbb{T}}$ -clustered configuration and, on the other hand, the assumptions $\theta \geq \frac{C_4 s}{2}$ and $s\tau\Delta \leq \frac{C_5}{2}$ imply that $\frac{C_4 s}{(\frac{\theta}{N})} \leq 2N \leq \frac{C_5}{s\tau(\frac{\Delta}{N})}$. Now

we apply Theorem 2.2 and obtain $\sigma_{\min}(\mathbf{V}_{2N}(\boldsymbol{\xi})) \geq C_7 \sqrt{2N} \left(\frac{2\Delta}{C_6}\right)^{\ell-1}$ and therefore using (6.15) we have

$$\sigma_{\min}(\tilde{\mathbf{V}}_N(\boldsymbol{\xi})) \geq C_7 \left(\frac{2\Delta}{C_6}\right)^{\ell-1}.$$

Finally using (6.13) we get that $\lambda_{\min}(\mathbf{G}(\mathcal{X})) \geq C_7^2 \left(\frac{2\Delta}{C_6}\right)^{2\ell-2}$. This proves Theorem 2.3 with $C_8 = C_7^2$, $C_9 = \frac{C_6}{2} = 16\pi e$, $C_{10} = \frac{C_4}{2}$ and $C_{11} = \frac{C_5}{2}$. \square

7. ENTIRE SPECTRUM

As mentioned in the Introduction, our proofs can be extended to provide scaling for *all the singular values* of \mathbf{V}_N (resp. eigenvalues of \mathbf{G} .)

For a single cluster, we have the following more general result from which Theorem 2.1 immediately follows as a corollary.

Theorem 7.1. *Let \mathcal{X} form a $(\Delta, \ell, \tau)_{\mathbb{T}}$ -clustered configuration. Denote the singular values of $\mathbf{V}_N(\mathcal{X})$ by*

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_\ell \equiv \sigma_{\min}.$$

Then for any N satisfying $C_1(\ell) \leq N \leq \frac{2\pi}{\tau\Delta}$, there holds

$$\sigma_m(\mathbf{V}_N(\mathcal{X})) \geq C_3 \sqrt{N} \left(\frac{N\Delta}{C_2} \right)^{m-1}, \quad m = 1, 2, \dots, \ell. \quad (7.1)$$

All the constants are the same as in Theorem 2.1.

Proof outline. Fix $m = 1, 2, \dots, \ell$, and let $\mathbf{c} \in \mathbb{C}^m$ with $\|\mathbf{c}\|_2 = 1$ be arbitrary. Furthermore, denote by \mathcal{X}_m the ordered set $\{x_1, \dots, x_m\} \subseteq \mathcal{X}$. By the Courant-Fischer minmax principle we have, extending (6.2), that

$$\sigma_m(\mathbf{V}_N(\mathcal{X})) \geq \min_{\mathbf{c} \in \mathbb{C}^m, \|\mathbf{c}\|_2=1} \|P_{\mathbf{c}, \mathcal{X}_m}\|_{2,N}, \quad m = 1, 2, \dots, \ell. \quad (7.2)$$

Now we can repeat the computation from Section 6.2, replacing ℓ with m and \mathcal{X} with \mathcal{X}_m . \square

In order to provide appropriate extensions of Theorem 2.2 and Theorem 2.3, recall the construction of the M -partition of \mathcal{X} from Section 6.3. Now for each $m = 1, 2, \dots, \ell$ let q_m be the number of clusters among the $\mathcal{C}^{(j)}$ of multiplicity at least m :

$$q_m := \#\{m : m \leq \ell^{(j)}\}. \quad (7.3)$$

The extension of Theorem 2.2 to include all the singular values is the following.

Theorem 7.2. *Let \mathcal{X} form a $(\Delta, \theta, s, \ell, \tau)_{\mathbb{T}}$ as in Theorem 2.2. Then for N as in Theorem 2.2, for each $m = 1, 2, \dots, \ell$ there are precisely q_m singular values of $\mathbf{V}_N(\mathcal{X})$ bounded from below by*

$$C_7 \sqrt{N} \left(\frac{N\Delta}{C_6} \right)^{m-1}. \quad (7.4)$$

To prove this result, we repeat the proof from Section 6.3, replacing Proposition 6.5 with its “full” version from [5] which reads as follows.

Proposition 7.1 (Theorem 2.2 in [5]). *Let \mathcal{X} form a $(\Delta, \theta, s, \ell, \tau)_{\mathbb{T}}$ -clustered configuration. Let $\tilde{\sigma}_1 \geq \tilde{\sigma}_2 \geq \dots \geq \tilde{\sigma}_s$ denote the union of all the singular values of the matrices $\mathbf{V}_N(\mathcal{C}^{(j)})$ in non-increasing order, and $\sigma_1 \geq \dots \geq \sigma_s$ denote the singular values of $\mathbf{V}_N(\mathcal{X})$. Then whenever (6.12) holds, we have*

$$\sigma_j \geq \frac{1}{2} \tilde{\sigma}_j, \quad j = 1, \dots, s.$$

As for Theorem 2.3, we can define the numbers q_m in a similar manner with respect to the clustered configurations on \mathbb{R} , and then we have the following.

Theorem 7.3. *For any \mathcal{X} forming a $(\Delta, \theta, s, \ell, \tau)_{\mathbb{R}}$ -clustered configuration as in Theorem 2.3, for each $m = 1, 2, \dots, \ell$ there are precisely q_m eigenvalues of $\mathbf{G}(\mathcal{X})$ bounded from below by*

$$C_8 \left(\frac{\Delta}{C_9} \right)^{2(m-1)}.$$

The proof is identical to that of Theorem 2.3, noting that (6.13) and (6.15) hold for all the singular values, and using Theorem 7.2 in place of Theorem 2.2.

REFERENCES

- [1] Céline Aubel and Helmut Bölcskei. Vandermonde matrices with nodes in the unit disk and the large sieve. *Applied and Computational Harmonic Analysis*, August 2017. doi:10.1016/j.acha.2017.07.006.
- [2] J.R. Auton. Investigation of Procedures for Automatic Resonance Extraction from Noisy Transient Electromagnetics Data. Volume III. Translation of Prony's Original Paper and Bibliography of Prony's Method. Technical report, Effects Technology Inc., Santa Barbara, CA, 1981.
- [3] Alex H. Barnett. How exponentially ill-conditioned are contiguous submatrices of the Fourier matrix? *arXiv:2004.09643 [cs, math]*, April 2020. arXiv:2004.09643.
- [4] Dmitry Batenkov, Laurent Demanet, Gil Goldman, and Yosef Yomdin. Conditioning of Partial Nonuniform Fourier Matrices with Clustered Nodes. *SIAM Journal on Matrix Analysis and Applications*, 44(1):199–220, January 2020. doi:10/ggjwzb.
- [5] Dmitry Batenkov, Benedikt Diederichs, Gil Goldman, and Yosef Yomdin. The spectral properties of Vandermonde matrices with clustered nodes. *Linear Algebra and its Applications*, August 2020. doi:10.1016/j.laa.2020.08.034.
- [6] Dmitry Batenkov, Gil Goldman, and Yosef Yomdin. Super-resolution of near-colliding point sources. *Information and Inference: A Journal of the IMA*, 10(2):515–572, June 2021. doi:10.1093/imaiai/iaaa005.
- [7] Emmanuel J. Candès and Carlos Fernandez-Granda. Super-Resolution from Noisy Data. *Journal of Fourier Analysis and Applications*, 19(6):1229–1254, December 2013. doi:10.1007/s00041-013-9292-3.
- [8] Emmanuel J. Candès and Carlos Fernandez-Granda. Towards a Mathematical Theory of Super-resolution. *Communications on Pure and Applied Mathematics*, 67(6):906–956, June 2014. doi:10.1002/cpa.21455.
- [9] Laurent Demanet and Nam Nguyen. The recoverability limit for superresolution via sparsity. 2014.
- [10] Benedikt Diederichs. Well-Posedness of Sparse Frequency Estimation. *arXiv:1905.08005 [math]*, May 2019. arXiv:1905.08005.
- [11] D.L. Donoho. Superresolution via sparsity constraints. *SIAM Journal on Mathematical Analysis*, 23(5):1309–1331, 1992.
- [12] T Erdélyi. Inequalities for exponential sums. *Sbornik: Mathematics*, 208(3):433–464, March 2017. doi:10.1070/SM8670.
- [13] Albert Fannjiang. Compressive Spectral Estimation with Single-Snapshot ESPRIT: Stability and Resolution. *arXiv:1607.01827 [cs, math]*, July 2016. arXiv:1607.01827.
- [14] PJSG Ferreira. Super-resolution, the recovery of missing samples and vandermonde matrices on the unit circle. In *Proceedings of the Workshop on Sampling Theory and Applications, Loen, Norway*, 1999.
- [15] Roger A. Horn and Charles R. Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge ; New York, 2nd ed edition, 2012.
- [16] A. E. Ingham. Some trigonometrical inequalities with applications to the theory of series. *Mathematische Zeitschrift*, 41(1):367–379, December 1936. doi:10.1007/BF01180426.
- [17] Stefan Kunis and Dominik Nagel. On the condition number of Vandermonde matrices with pairs of nearly-colliding nodes. *Numerical Algorithms*, July 2020. doi:10.1007/s11075-020-00974-x.
- [18] Stefan Kunis and Dominik Nagel. On the smallest singular value of multivariate Vandermonde matrices with clustered nodes. *Linear Algebra and its Applications*, 604:1–20, November 2020. doi:10.1016/j.laa.2020.06.003.
- [19] Weilin Li and Wenjing Liao. Stable super-resolution limit and smallest singular value of restricted Fourier matrices. *Applied and Computational Harmonic Analysis*, 51:118–156, March 2021. doi:10.1016/j.acha.2020.10.004.
- [20] Weilin Li, Wenjing Liao, and Albert Fannjiang. Super-resolution limit of the ESPRIT algorithm. *IEEE Transactions on Information Theory*, pages 1–1, 2020. doi:10/ggrnpw.
- [21] Wenjing Liao and Albert Fannjiang. MUSIC for single-snapshot spectral estimation: Stability and super-resolution. *Applied and Computational Harmonic Analysis*, 40(1):33–67, January 2016. doi:10.1016/j.acha.2014.12.003.
- [22] Ankur Moitra. Super-resolution, Extremal Functions and the Condition Number of Vandermonde Matrices. In *Proceedings of the Forty-Seventh Annual ACM on Symposium on Theory of Computing*, STOC '15, pages 821–830, New York, NY, USA, 2015. ACM. doi:10.1145/2746539.2746561.
- [23] H. L. Montgomery and R. C. Vaughan. Hilbert's Inequality. *Journal of the London Mathematical Society*, s2-8(1):73–82, May 1974. doi:10.1112/jlms/s2-8.1.73.
- [24] F.L. Nazarov. Local estimates of exponential polynomials and their applications to inequalities of uncertainty principle type. *St Petersburg Mathematical Journal*, 5(4):663–718, 1994.
- [25] M. Negreanu and E. Zuazua. Discrete Ingham Inequalities and Applications. *SIAM Journal on Numerical Analysis*, 44(1):412–448, January 2006. doi:10.1137/050630015.
- [26] R. Prony. Essai experimental et analytique. *J. Ec. Polytech.(Paris)*, 2:24–76, 1795.

- [27] D. Slepian. Prolate spheroidal wave functions, fourier analysis, and uncertainty – V: The discrete case. *Bell System Technical Journal*, The, 57(5):1371–1430, May 1978. doi:10.1002/j.1538-7305.1978.tb02104.x.
- [28] Paul Turán. *Eine neue Methode in der Analysis und deren Anwendungen*. Akadémiai Kiadó, 1953.
- [29] J.M. Varah. The prolate matrix. *Linear Algebra and its Applications*, 187:269–278, July 1993. doi:10.1016/0024-3795(93)90142-B.
- [30] A. Zygmund. *Trigonometric Series. Vols. I, II*. Cambridge University Press, New York, 1959.

DEPARTMENT OF APPLIED MATHEMATICS, SCHOOL OF MATHEMATICAL SCIENCES, TEL AVIV UNIVERSITY, P.O. BOX 39040, TEL AVIV 6997801, ISRAEL

Email address: `dbatenkov@tauex.tau.ac.il`

DEPARTMENT OF MATHEMATICS, THE WEIZMANN INSTITUTE OF SCIENCE, REHOVOT 76100, ISRAEL

Email address: `gil.goldman@weizmann.ac.il`