

On the smallest singular value of multivariate Vandermonde matrices with clustered nodes

Stefan Kunis* Dominik Nagel*

July 18, 2019

We prove lower bounds for the smallest singular value of rectangular, multivariate Vandermonde matrices with nodes on the complex unit circle. The nodes are “off the grid”, groups of nodes cluster, and the studied minimal singular value is bounded below by the product of inverted distances of a node to all other nodes in the specific cluster. By providing also upper bounds for the smallest singular value, this completely settles the univariate case and pairs of nodes in the multivariate case, both including reasonable sharp constants. For larger clusters, we show that the smallest singular value depends also on the geometric configuration within a cluster.

Key words and phrases: Vandermonde matrix, colliding nodes, cluster, condition number, restricted Fourier matrices, frequency analysis, super resolution.

2010 AMS Mathematics Subject Classification : 15A18, 65T40, 42A15.

1 Introduction

Vandermonde matrices appear e.g. in the stability analysis of super-resolution algorithms like Prony’s method [6, 12], the matrix pencil method [11, 19], the ESPRIT algorithm [22, 21, 16], and the MUSIC algorithm [23, 17]. We are interested in the case of nodes on the complex unit circle and a large polynomial degree, the matrices then generalize the classical discrete Fourier matrices to non-equispaced nodes and the involved polynomial degree is also called bandwidth. If all nodes are well-separated, bounds on the condition number are established for example in [5, 14, 19, 2, 8] for the univariate case and in [14, 12] at least partially for the multivariate case. For node sets with distances of which some are below the inverse bandwidth, the behavior of the smallest singular value is subject of current research. The seminal paper [9] coined the term (inverse) super-resolution factor for the product of the bandwidth and the minimal separation of the nodes. For M nodes on a grid, the results in [9, 7] imply that the smallest singular value is at most as small as the inverse super-resolution factor raised to the power of $M - 1$ if the super-resolution factor is greater than 1. More recently, the practically relevant situation of clustered nodes was studied in [20, 1, 15, 3, 13, 4, 8]. In the univariate case and for different setups, all of these refinements are able to replace the exponent $M - 1$

*Osnabrück University, Institute of Mathematics {skunis,dnagel}@uos.de

by the smaller number $m - 1$, where m denotes the number of nodes that are in the largest cluster of nodes.

Here, we refine the proof technique developed in the second version of [15] and extend it to arbitrary dimensions. In contrast to [15], we only use the information on the biggest cluster size, minimal separation between clusters and a the worst case cluster complexity (or a minimal separation between nodes) instead of taking the structure of each cluster into account. In summary, our contributions are:

- i) a refined analysis of the univariate case, cf. [15], eliminating the dependence on the total number of nodes, weakening a technical condition on the cluster separation, and improving constants, mainly by
 - a) a geometric packing argument and
 - b) an improved estimate of Dirichlet kernels and Lagrange-like basis functions;
- ii) a multidimensional generalization, including
 - a) a quantitative estimate for the well-separated case,
 - b) a sharp estimate for pair clusters in higher dimensions, and
 - c) an example on the limitations for larger clusters in higher dimensions.

The outline of this paper is as follows: Section 2 fixes notation, states the problem and gives some definitions. Furthermore, we generalize the so-called robust duality lemma from the second version of [15] to the multivariate case. In Section 3, we introduce some auxiliary functions which are used to prove our main results in Section 4. Additionally, we give examples with specified parameters, present implications of our result for special node configurations like pair clusters and well separated nodes, and compare them with existing results. In Section 5, upper bounds on the smallest singular value for the univariate case and for pair clusters in higher dimensions are presented - these match the lower bounds from our main theorem. Furthermore, an example of a triple cluster in two dimensions is given which shows that geometric properties beyond pairwise distances are needed for understanding the multivariate case. Finally, in Section 6 numerical experiments are presented that support statements and comparisons from preceding sections.

2 Preliminaries

Definition 2.1 (Setting). *We denote the component of a vector by bracketing and setting a subscript, unless its components are defined differently. Let $d \in \mathbb{N}$ be a given dimension and $\Omega := \{\mathbf{t}_1, \dots, \mathbf{t}_M\} \subset [0, 1)^d$ a set of points. The corresponding nodes are given by $\mathbf{z}_j := e^{2\pi i \mathbf{t}_j} \in \mathbb{T}^d, j = 1, \dots, M$, where $\mathbb{T} := \{z \in \mathbb{C}: |z| = 1\}$ denotes the complex unit circle. We identify the unit interval with the unit circle and therefore, we do not make a difference between the \mathbf{t}_j and \mathbf{z}_j and call them both nodes. Throughout the paper, $\|\cdot\|_2$ denotes the euclidean norm for vectors and also its induced norm for matrices, and analogously $\|\cdot\|_\infty$ the max-norm. Let $n \in \mathbb{N}$ be a degree, set $N := n+1$ and assume $M < N^d$. We are interested in the multivariate, rectangular Vandermonde matrix*

$$\mathbf{A} := \mathbf{A}(\Omega, n) := \left(\mathbf{z}_j^\alpha \right)_{\substack{j=1, \dots, M \\ \alpha \in \mathbb{N}_0^d, \|\alpha\|_\infty \leq n}} \in \mathbb{C}^{M \times N^d}, \quad (2.1)$$

and its smallest singular value

$$\sigma_{\min}(\mathbf{A}) := \min_{\substack{\mathbf{v} \in \mathbb{C}^M \\ \|\mathbf{v}\|_2=1}} \|\mathbf{A}^* \mathbf{v}\|_2.$$

The following lemma builds the core of the proof technique developed in the second version of [15] which we adapt here to the multivariate setting.

Lemma 2.2 (Robust duality, cf. [15, v2, Prop. 2]). *Let Ω and \mathbf{A} be given as in Definition 2.1. If for any unit norm vector $\mathbf{v} = (v_1, \dots, v_M)^\top \in \mathbb{C}^M$, and $\boldsymbol{\epsilon} = (\epsilon_1, \dots, \epsilon_M)^\top \in \mathbb{C}^M$ with $\|\boldsymbol{\epsilon}\|_2 \leq 1$, there exists a trigonometric polynomial of max-degree at most $n \in \mathbb{N}$, i.e.,*

$$f \in \mathcal{P}(n) := \left\{ g: \mathbb{T}^d \rightarrow \mathbb{C}: g(\mathbf{t}) = \sum_{\boldsymbol{\alpha} \in \mathbb{N}_0^d, \|\boldsymbol{\alpha}\|_\infty \leq n} \hat{g}_\alpha e^{2\pi i \boldsymbol{\alpha} \cdot \mathbf{t}}, \hat{g}_\alpha \in \mathbb{C} \right\},$$

such that $f(\mathbf{t}_j) = v_j + \epsilon_j$ for each $j = 1, \dots, M$, then

$$\|\mathbf{A}^* \mathbf{v}\|_2 \geq (1 - \|\boldsymbol{\epsilon}\|_2) \|f\|_{L^2(\mathbb{T}^d)}^{-1}.$$

Proof. Define the discrete measure $\mu := \sum_{j=1}^M v_j \delta_{\mathbf{t}_j}$. Its Fourier coefficients are given by

$$\hat{\mu}(\boldsymbol{\alpha}) = \int_{\mathbb{T}^d} e^{-2\pi i \boldsymbol{\alpha} \cdot \mathbf{t}} d\mu(\mathbf{t}) = \sum_{j=1}^M v_j z_j^{-\boldsymbol{\alpha}} = (\mathbf{A}^* \mathbf{v})_\alpha, \quad \boldsymbol{\alpha} \in \mathbb{N}_0^d, \|\boldsymbol{\alpha}\|_\infty \leq n.$$

On the one hand, using the interpolation property of f and the lower triangular inequality of the absolute value, we have

$$\left| \int_{\mathbb{T}^d} \bar{f} d\mu \right| = \left| \sum_{j=1}^M \overline{f(\mathbf{t}_j)} v_j \right| = \left| \|\mathbf{v}\|_2^2 + \sum_{j=1}^M \bar{\epsilon}_j v_j \right| \geq \|\mathbf{v}\|_2^2 - \|\mathbf{v}\|_2 \|\boldsymbol{\epsilon}\|_2 = (1 - \|\boldsymbol{\epsilon}\|_2),$$

and on the other hand, using $f \in \mathcal{P}(n)$, the Cauchy–Schwarz inequality and Parseval’s identity, we have

$$\left| \int_{\mathbb{T}^d} \bar{f} d\mu \right| = \left| \sum_{\boldsymbol{\alpha} \in \mathbb{N}_0^d, \|\boldsymbol{\alpha}\|_\infty \leq n} \overline{\hat{f}_\alpha} \hat{\mu}(\boldsymbol{\alpha}) \right| \leq \|\hat{f}\|_2 \|\mathbf{A}^* \mathbf{v}\|_2 = \|f\|_{L^2(\mathbb{T}^d)} \|\mathbf{A}^* \mathbf{v}\|_2.$$

■

The advantage of that lemma is, if $\mathbf{v} \in \mathbb{C}^M$ is a unit norm vector such that $\sigma_{\min}(\mathbf{A}) = \|\mathbf{A}^* \mathbf{v}\|_2$, it suffices to construct a function $f \in \mathcal{P}(n)$ almost interpolating the values of \mathbf{v} in order to provide a lower bound.

The following definition is similar to the ‘localized clumps’ model from the second version of [15]. We did some renaming in terms of [3] and use a normalization by N rather than n .

Definition 2.3 (Geometry of nodes). *The wrap-around distance between two nodes $\mathbf{t}, \mathbf{t}' \in [0, 1)^d$ is defined by*

$$|\mathbf{t} - \mathbf{t}'|_{\mathbb{T}^d} := \min_{\mathbf{r} \in \mathbb{Z}^d} \|\mathbf{t} - \mathbf{t}' + \mathbf{r}\|_\infty.$$

i) A subset of nodes is called cluster if it is contained in a cube of length $1/N$. For two clusters $\Lambda', \Lambda'' \subset \Omega$, we define

$$\text{dist}(\Lambda', \Lambda'') := \min\{|\mathbf{t}' - \mathbf{t}''|_{\mathbb{T}^d} : \mathbf{t}' \in \Lambda', \mathbf{t}'' \in \Lambda''\}.$$

ii) The node set Ω is called a clustered node configuration with L clusters if it can be written as

$$\Omega = \bigcup_{l=1}^L \Lambda_l,$$

where the Λ_l are clusters and the (normalized) minimal cluster separation ρ fulfills

$$\rho := N \min_{1 \leq l < l' \leq L} \text{dist}(\Lambda_l, \Lambda_{l'}) > 1.$$

We order $|\Lambda_1| \geq |\Lambda_2| \geq \dots \geq |\Lambda_L|$ and denote the cardinality of the biggest cluster by $\lambda := |\Lambda_1|$. In passing, we note that the node set Ω is called well separated with normalized separation ρ if $\lambda = 1$. Moreover, we define the partitioning of \mathbb{T}^d into shells by

$$J_m := J_m(\Omega, N, \rho) := \left\{ \mathbf{t} \in \mathbb{T}^d : m\rho \leq N |\mathbf{t}|_{\mathbb{T}^d} < (m+1)\rho \right\}, \quad m = 0, \dots, \left\lfloor \frac{N}{2\rho} \right\rfloor.$$

iii) The cluster complexity is defined by

$$\mathcal{C} := \mathcal{C}(\Omega, N) := \max_{j=1, \dots, M} \prod_{\mathbf{t}' \in \Omega : 0 < |\mathbf{t}_j - \mathbf{t}'|_{\mathbb{T}^d} \leq 1/N} \frac{1}{N |\mathbf{t}_j - \mathbf{t}'|_{\mathbb{T}^d}}$$

and finally, we define the (normalized) minimal separation

$$\tau := N \min_{1 \leq j < j' \leq M} |\mathbf{t}_j - \mathbf{t}_{j'}|_{\mathbb{T}^d}.$$

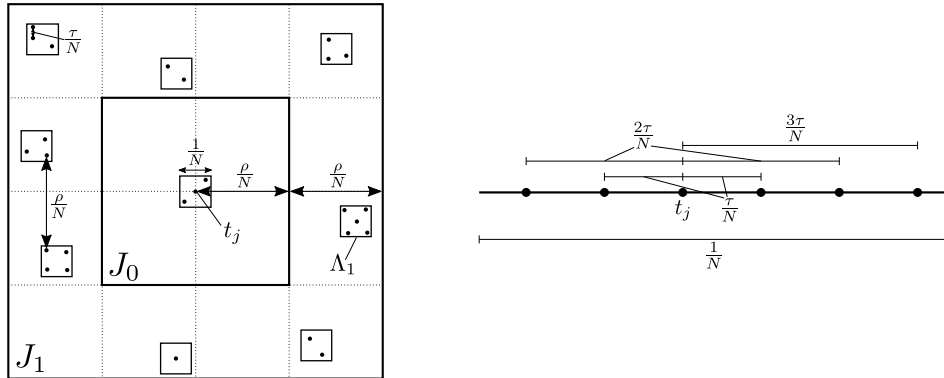


Figure 2.1: Left: Clustered node configuration and partitioning into shells for $d = 2$; right: a cluster which maximizes \mathcal{C} for $d = 1$.

Remark 2.4. (Geometry of nodes) With the notation of Definition 2.3, we note that

i) the inequality $\sin(x) \geq 2x/\pi$ for $0 \leq x \leq \pi/2$ implies

$$|z - z'| = 2 \sin(\pi |t - t'|_{\mathbb{T}}) \geq 4 |t - t'|_{\mathbb{T}}, \quad z := e^{2\pi it}, z' := e^{2\pi it'} \in \mathbb{T}. \quad (2.2)$$

A higher order approximation is given in the second version of [15],

$$|z - z'| \geq 2\pi \left(1 - \frac{\pi^2 |t - t'|_{\mathbb{T}}^2}{3} \right)^{1/2} |t - t'|_{\mathbb{T}}. \quad (2.3)$$

ii) A necessary condition on N for the existence of a clustered node configuration with L clusters is $L\rho^d \leq N^d$, with equality if and only if all nodes are equispaced. Similarly, if $N \geq L^{1/d}(\rho + 1)$, then equispaced cluster with arbitrary node configuration within each cluster exist. Moreover, the cluster separation ρ needs to scale at least linearly in the biggest cluster size λ . If on the contrary, $\rho < \lambda/4$ and $d = 1$ for simplicity of the argument, then let λ nodes form a cluster (length at most $1/N$) and place one node as far as possible away. With fixed N , we have $2\rho = N - 1$ and therefore, $\rho < \lambda/4$ is equivalent to $N \leq M/2 + 1/2$ and thus $\text{rank}(\mathbf{A}) \leq N < M$. On the other hand, $\rho > \lambda$ already implies $N \geq L(\rho + 1) \geq M$.

Finally note, that the packing argument in [14, Lemma 4.5] yields

$$|J_m \cap \Omega| \leq 2^d (2^d - 1) m^{d-1} \lambda,$$

see also Figure 2.1 (left).

iii) The cluster complexity can be upper bounded by the normalized minimal separation as follows. For $d \in \mathbb{N}$, we have $\mathcal{C} \leq \tau^{1-\lambda}$ and equality for $\lambda = 1$ and $\lambda = 2$. Refined for $d = 1$, it is easy to see that the cluster complexity is maximized by an equispaced cluster with λ nodes separated by τ/N and taking distances from the center node, see Figure 2.1 (right). By logarithmic convexity, direct calculation, and Stirling's approximation, we thus have

$$\mathcal{C} \leq \frac{1}{\tau^{\lambda-1}} \left(\left\lfloor \frac{\lambda-1}{2} \right\rfloor! \cdot \left\lceil \frac{\lambda-1}{2} \right\rceil! \right)^{-1} \leq \frac{1}{\tau^{\lambda-1} \Gamma(\frac{\lambda+1}{2})^2} \leq \frac{(2e)^{\lambda-1}}{\lambda^\lambda} \cdot \frac{1}{\tau^{\lambda-1}} \quad (2.4)$$

and similarly

$$\max_{\Omega} \mathcal{C} = \frac{1}{\tau^{\lambda-1}} \left(\left\lfloor \frac{\lambda-1}{2} \right\rfloor! \cdot \left\lceil \frac{\lambda-1}{2} \right\rceil! \right)^{-1} \geq \frac{(2e)^{\lambda-1}}{\lambda^{\lambda+1}} \cdot \frac{1}{\tau^{\lambda-1}}, \quad (2.5)$$

where the maximum is taken over all clustered node configurations with normalized minimal separation τ and the largest cluster containing λ nodes.

3 Auxiliary functions

Lemma 3.1 (Modified Dirichlet kernel). For $m, \beta \in \mathbb{N}$ the modified Dirichlet kernel is defined as $d_m: [0, 1) \rightarrow \mathbb{C}$,

$$d_m(t) := \frac{1}{m+1} \sum_{k=0}^m e^{2\pi ikt} = \begin{cases} 1, & t = 0, \\ \frac{e^{\pi imt}}{m+1} \cdot \frac{\sin(\pi(m+1)t)}{\sin(\pi t)}, & t \neq 0. \end{cases}$$

We define the powers of the multivariate modified Dirichlet kernel by

$$d_m^\beta : [0, 1]^d \rightarrow \mathbb{C}, \quad d_m^\beta(\mathbf{t}) := \left(\prod_{\ell=1}^d d_m((\mathbf{t})_\ell) \right)^\beta \in \mathcal{P}(m\beta).$$

If $m \geq \beta$ and $\mathbf{t} \in \mathbb{T}^d \setminus \{\mathbf{0}\}$, then

$$i) |d_m(\mathbf{t})| \leq d_m(\mathbf{0}) = 1,$$

$$ii) |d_m(\mathbf{t})| \leq \frac{1}{2(m+1)|\mathbf{t}|_{\mathbb{T}^d}},$$

$$iii) \|d_m^\beta\|_{L^2(\mathbb{T}^d)}^2 \leq \frac{1}{(m+1)^{d\beta^{d/2}}},$$

$$iv) \left| \left\langle d_m^\beta, d_m^\beta(\cdot - \mathbf{t}) \right\rangle_{L^2(\mathbb{T}^d)} \right| \leq \frac{1}{2(m+1)^{d\beta^{(d-1)/2}}} \cdot \frac{1}{(m+1)^\beta |\mathbf{t}|_{\mathbb{T}^d}^\beta}.$$

Proof. First, note that

$$|d_m(\mathbf{t})| \leq \left(\frac{1}{m+1} \sum_{k=0}^m |e^{2\pi i k t}| \right)^d = 1 = d_m(\mathbf{0})$$

and the point-wise bound follows in the univariate case by

$$|d_m(t)| = \frac{1}{m+1} \left| \frac{\sin(\pi(m+1)t)}{\sin(\pi t)} \right| \leq \frac{1}{(m+1)|\sin(\pi t)|} \leq \frac{1}{2(m+1)|t|_{\mathbb{T}}}.$$

Second, in the multivariate case, setting $t := |\mathbf{t}|_{\mathbb{T}^d}$, and using i) and the univariate bound yield

$$|d_m(\mathbf{t})| = \prod_{\ell=1}^d |d_m((\mathbf{t})_\ell)| \leq |d_m(t)| \leq \frac{1}{2(m+1)|t|_{\mathbb{T}}} = \frac{1}{2(m+1)|\mathbf{t}|_{\mathbb{T}^d}}.$$

Note that $\|d_m\|_{L^2(\mathbb{T}^d)}^2 = \|d_m\|_{L^2(\mathbb{T})}^{2d}$ and therefore, the third assertion is proven for the univariate case as follows. For $m \geq \beta$, Parseval's identity and direct calculation show

$$\|d_m\|_{L^2(\mathbb{T})}^2 = \frac{1}{m+1}, \quad \|d_m^2\|_{L^2(\mathbb{T})}^2 = \frac{1}{m+1} \left[\frac{2}{3} + \frac{1}{3(m+1)^2} \right] \leq \frac{1}{m+1} \cdot \frac{19}{27} \leq \frac{1}{m+1} \cdot \frac{1}{\sqrt{2}},$$

$$\|d_m^3\|_{L^2(\mathbb{T})}^2 = \frac{1}{m+1} \left[\frac{11}{20} + \frac{1}{4(m+1)^2} + \frac{1}{5(m+1)^4} \right] \leq \frac{1}{m+1} \cdot \frac{145}{256} \leq \frac{1}{m+1} \cdot \frac{1}{\sqrt{3}}.$$

For $x \in [0, 1]$ and $m \geq 4$, the estimates in [18, Proof of Lemma 2] yield

$$\frac{\sin \pi x}{(m+1) \sin \frac{\pi}{m+1} x} \leq \exp \left(-\frac{\pi^2((m+1)^2 - 1)}{6(m+1)^2} x^2 \right) \leq \exp \left(-\frac{4\pi^2 x^2}{25} \right)$$

and thus, for $m \geq \beta \geq 4$, the remaining estimate

$$\|d_m^\beta\|_{L^2(\mathbb{T})}^2 = \frac{2}{(m+1)^{2\beta}} \int_0^{1/2} \left| \frac{\sin(\pi(m+1)t)}{\sin(\pi t)} \right|^{2\beta} dt$$

$$\begin{aligned}
&= \frac{2}{m+1} \left[\frac{1}{(m+1)^{2\beta}} \left(\int_0^1 \left(\frac{\sin(\pi x)}{\sin(\frac{\pi}{m+1}x)} \right)^{2\beta} dx + \int_1^{\frac{m+1}{2}} \left| \frac{\sin(\pi x)}{\sin(\frac{\pi}{m+1}x)} \right|^{2\beta} dx \right) \right] \\
&\leq \frac{2}{m+1} \left[\int_0^\infty \exp\left(-\frac{8\beta\pi^2 x^2}{25}\right) dx + \int_1^\infty \left(\frac{1}{2x}\right)^{2\beta} dx \right] \\
&= \frac{1}{m+1} \left[\frac{5}{2\sqrt{2\pi}} \frac{1}{\sqrt{\beta}} + \frac{2^{1-2\beta}}{2\beta-1} \right] \leq \frac{1}{m+1} \cdot \frac{1}{\sqrt{\beta}}.
\end{aligned}$$

In order to prove the fourth assertion, note $|t|_{\mathbb{T}} \leq |t-t'|_{\mathbb{T}} + |t'|_{\mathbb{T}} \leq 2 \max\{|t-t'|_{\mathbb{T}}, |t'|_{\mathbb{T}}\}$ and hence, i) and ii) yield

$$|d_m(t')| |d_m(t'-t)| \leq \frac{1}{2(m+1)} \min \left\{ \frac{1}{|t'-t|_{\mathbb{T}}}, \frac{1}{|t'|_{\mathbb{T}}} \right\} \leq \frac{1}{(m+1)|t|_{\mathbb{T}}}$$

and $\|d_m d_m(\cdot - t)\|_{L^\infty(\mathbb{T})} \leq ((m+1)|t|_{\mathbb{T}})^{-1}$. Moreover, direct computation gives

$$\left| \langle d_m, d_m(\cdot - t) \rangle_{L^2(\mathbb{T})} \right| = \frac{|d_m(t)|}{m+1} \leq \frac{1}{m+1} \cdot \frac{1}{2} \cdot \frac{1}{(m+1)|t|_{\mathbb{T}}}.$$

and with $z = e^{2\pi i t}$ and Parseval's identity also

$$\begin{aligned}
\left| \langle d_m^2, d_m^2(\cdot - t) \rangle_{L^2(\mathbb{T})} \right| &= \frac{1}{(m+1)^4} \int_{\mathbb{T}} \left(\frac{\sin \pi(m+1)t'}{\sin \pi t'} \cdot \frac{\sin \pi(m+1)(t'-t)}{\sin \pi(t'-t)} \right)^2 dt' \\
&= \left| \sum_{k=-m}^m (m+1-|k|)^2 z^k \right| \\
&= \frac{|(z+1)z(z^{m+1}-z^{-m-1}) + 4(m+1)z(1-z)|}{(m+1)^4 |z-1|^3} \\
&\leq \frac{1}{m+1} \cdot \frac{1}{2} \cdot \frac{1}{(m+1)^2 |t|_{\mathbb{T}}^2}.
\end{aligned}$$

Finally, let t be the coordinate with $|t|_{\mathbb{T}} = |\mathbf{t}|_{\mathbb{T}^d}$, then the Cauchy–Schwarz inequality, iii), and the above yield (noting that $e^{-2\pi i m t'} d_m^2(t') \geq 0$ and omitting the second last line if $\beta = 1$)

$$\begin{aligned}
\left| \langle d_m^\beta, d_m^\beta(\cdot - \mathbf{t}) \rangle_{L^2(\mathbb{T}^d)} \right| &\leq \left| \int_{\mathbb{T}^d} d_m^\beta(\mathbf{t}') \overline{d_m^\beta(\mathbf{t}' - \mathbf{t})} d\mathbf{t}' \right| \\
&\leq \|d_m^\beta\|_{L^2(\mathbb{T})}^{2(d-1)} \left| \int_{\mathbb{T}} d_m^\beta(t') \overline{d_m^\beta(t' - t)} dt' \right| \\
&\leq \|d_m^\beta\|_{L^2(\mathbb{T})}^{2(d-1)} \|d_m d_m(\cdot - t)\|_{L^\infty(\mathbb{T})}^{\beta-2} \left| \langle d_m^2, d_m^2(\cdot - t) \rangle_{L^2(\mathbb{T})} \right| \\
&\leq \frac{1}{2(m+1)^{d\beta(d-1)/2}} \cdot \frac{1}{(m+1)^\beta |t|_{\mathbb{T}}^\beta}.
\end{aligned}$$

■

Lemma 3.2 (Lagrange-like basis with decay, cf. [15, v2, Lem. 3]). *Let $\beta, d, M, n \in \mathbb{N}$, β be even, $\Omega = \{\mathbf{t}_1, \dots, \mathbf{t}_M\} \subset [0, 1]^d$ be a clustered node configuration and $n \geq 2\beta^2 \lambda$. Then for each $\mathbf{t}_j \in \Omega$ with $\mathbf{t}_j \in \Lambda_l$ for some $l = l(j)$, there exists an $I_j \in \mathcal{P}(n)$, such that*

- i) $I_j(\mathbf{t}_k) = \delta_{jk}$ for all $\mathbf{t}_k \in \Lambda_l$,
- ii) $|I_j(\mathbf{t})| \leq \frac{\beta^\beta \lambda^{\beta+\lambda-1}}{(2N|\mathbf{t}-\mathbf{t}_j|_{\mathbb{T}^d})^\beta} \mathcal{C}$ for all $\mathbf{t} \neq \mathbf{t}_j$, and

$$\text{iii) } \left| \langle I_k, I_j \rangle_{L^2(\mathbb{T}^d)} \right| \leq \frac{\lambda^d \beta^{d/2}}{N^d} \lambda^{2\lambda-2} \mathcal{C}^2 \begin{cases} 1, & \mathbf{t}_k \in \Lambda_l, \\ \frac{\sqrt{\beta}}{2} \left(\frac{\lambda\beta}{N|\mathbf{t}_j-\mathbf{t}_k|_{\mathbb{T}^d}} \right)^\beta, & \text{otherwise.} \end{cases}$$

Proof. We define the functions I_j as product of a Lagrange polynomial G_j within the cluster and a fast decaying function H_j . Let $j \in \{1, \dots, M\}$ be fixed and define the j -th Lagrange polynomial within its cluster Λ_l , $l = l(j)$, as follows. If $|\Lambda_l| = 1$, we simply set $G_j \equiv 1$. Otherwise, let

$$Q := \left\lfloor \frac{n}{\lambda} \right\rfloor \geq \frac{n - \lambda + 1}{\lambda} \geq \frac{N}{2\lambda} \quad (3.1)$$

denote the 'blow-up-factor' and for $\mathbf{t}_k \in \Lambda_l \setminus \{\mathbf{t}_j\}$ let $\ell(k)$ be the index of the vector component that realizes the distance $|\mathbf{t}_j - \mathbf{t}_k|_{\mathbb{T}^d}$. We immediately have $|Q\mathbf{t}_j - Q\mathbf{t}_k|_{\mathbb{T}^d} = Q|\mathbf{t}_j - \mathbf{t}_k|_{\mathbb{T}^d} \neq 0$ and thus

$$G_j(\mathbf{t}) := \prod_{\mathbf{t}_k \in \Lambda_l \setminus \{\mathbf{t}_j\}} \frac{e^{2\pi i Q(\mathbf{t})_{\ell(k)}} - e^{2\pi i Q(\mathbf{t}_k)_{\ell(k)}}}{e^{2\pi i Q(\mathbf{t}_j)_{\ell(k)}} - e^{2\pi i Q(\mathbf{t}_k)_{\ell(k)}}}$$

fulfills $G_j(\mathbf{t}_k) = \delta_{j,k}$ and by inequality (2.2)

$$\|G_j\|_{L^\infty(\mathbb{T}^d)} \leq \prod_{\mathbf{t}_k \in \Lambda_l \setminus \{\mathbf{t}_j\}} \frac{1}{2Q \left| (\mathbf{t}_j)_{\ell(k)} - (\mathbf{t}_k)_{\ell(k)} \right|_{\mathbb{T}}} \leq \lambda^{\lambda-1} \mathcal{C}. \quad (3.2)$$

We proceed by setting

$$P := \left\lfloor \frac{n}{\lambda\beta} \right\rfloor, \quad P+1 \geq \frac{n - \lambda\beta + 1}{\lambda\beta} + 1 \geq \frac{N}{\lambda\beta},$$

and $H_j(\mathbf{t}) := d_P^\beta(\mathbf{t} - \mathbf{t}_j)$. Lemma 3.1 yields $H_j(\mathbf{t}_j) = 1$,

$$\begin{aligned} |H_j(\mathbf{t})| &\leq \left(\frac{1}{2(P+1)|\mathbf{t} - \mathbf{t}_j|_{\mathbb{T}^d}} \right)^\beta \leq \left(\frac{\lambda\beta}{2N|\mathbf{t} - \mathbf{t}_j|_{\mathbb{T}^d}} \right)^\beta, & \mathbf{t} \neq \mathbf{t}_j, \\ \left| \langle H_k, H_j \rangle_{L^2(\mathbb{T}^d)} \right| &\leq \left\| d_P^\beta \right\|_{L^2(\mathbb{T}^d)}^2 \leq \frac{1}{(P+1)^{d\beta d/2}} \leq \frac{\lambda^d \beta^{d/2}}{N^d}, & \mathbf{t}_k \in \Lambda_l, \\ \left| \langle H_k, H_j \rangle_{L^2(\mathbb{T}^d)} \right| &= \left| \left\langle d_P^\beta, d_P^\beta(\cdot - (\mathbf{t}_j - \mathbf{t}_k)) \right\rangle_{L^2(\mathbb{T}^d)} \right| \leq \frac{\lambda^d \beta^{(d+1)/2} (\lambda\beta)^\beta}{2N^d (N|\mathbf{t}_j - \mathbf{t}_k|_{\mathbb{T}^d})^\beta}, & \mathbf{t}_k \notin \Lambda_l. \end{aligned}$$

Finally, we define $I_j(\mathbf{t}) := G_j(\mathbf{t})H_j(\mathbf{t})$. This yields $I_j \in \mathcal{P}(n)$ since $G_j \in \mathcal{P}(Q(\lambda-1))$, $H_j \in \mathcal{P}(P\beta)$, and

$$P\beta + (\lambda-1)Q \leq \frac{n}{\lambda} + (\lambda-1)\frac{n}{\lambda} = n.$$

Moreover, this function has the desired property $I_j(\mathbf{t}_k) = \delta_{jk}$ for all $\mathbf{t}_k \in \Lambda_l$ and the two remaining inequalities follow by $|I_j(\mathbf{t})| \leq \|G_j\|_{L^\infty(\mathbb{T}^d)} |H_j(\mathbf{t})|$ and by using $e^{-\pi i \beta P(\mathbf{t} - \mathbf{t}_j)} H_j(\mathbf{t}) \geq 0$, also $\left| \langle I_k, I_j \rangle_{L^2(\mathbb{T}^d)} \right| \leq \|G_j\|_{L^\infty(\mathbb{T}^d)}^2 \left| \langle H_k, H_j \rangle_{L^2(\mathbb{T}^d)} \right|$. \blacksquare

Remark 3.3. Following the calculation in the second version of [15, p. 36], we can improve (3.2) to

$$\|G_j\|_{L^\infty(\mathbb{T}^d)} \leq \left(1 - \frac{\pi^2}{3\lambda^2}\right)^{\frac{1-\lambda}{2}} \left(\frac{N}{\lambda} / \left\lfloor \frac{n}{\lambda} \right\rfloor\right)^{\lambda-1} \left(\frac{\lambda}{\pi}\right)^{\lambda-1} \mathcal{C} \leq 2.4 \left(\frac{C(n)\lambda}{\pi}\right)^{\lambda-1} \mathcal{C}$$

with $C(n) \rightarrow 1$ for $n \rightarrow \infty$ and where the first two bracketed terms are due to (2.3) and (3.1), respectively.

4 A lower bound on the smallest singular value

In this chapter we work out the multivariate extension of Theorem 1 in the second version of [15]. Additionally, we do an improvement on the cluster separation condition, especially make the cluster separation independent on the number of nodes M . Furthermore, we provide an improved estimate on the smallest singular value $\sigma_{\min}(\mathbf{A})$ only depending on the biggest cluster size λ and not on the number of all nodes M .

Theorem 4.1. Let $\beta, d, N, M \in \mathbb{N}$, $\beta \geq d+1$ even, $\Omega = \{\mathbf{t}_1, \dots, \mathbf{t}_M\} \subset [0, 1)^d$ be a clustered node configuration and $N > 2\beta^2\lambda$. Moreover, assume the cluster separation

$$\rho \geq \lambda\beta \left(\beta^{1/2}2^d(2^d - 1)\lambda^\lambda \zeta(\beta - d + 1)\mathcal{C}\right)^{\frac{1}{\beta}}. \quad (4.1)$$

Then the smallest singular value of the Vandermonde matrix $\mathbf{A} \in \mathbb{C}^{M \times N^d}$ from Definition 2.1 is bounded by

$$\sigma_{\min}(\mathbf{A}) \geq \left(1.5 \cdot \beta^{d/4} \lambda^{\lambda+d/2-1/2}\right)^{-1} \frac{N^{d/2}}{\mathcal{C}}.$$

Proof. We apply the robust duality from Lemma 2.2, with $\mathbf{v} \in \mathbb{C}^M$, $\|\mathbf{v}\|_2 = 1$, such that $\sigma_{\min}(\mathbf{A}) = \|\mathbf{A}^*\mathbf{v}\|_2$, and

$$f := \sum_{k=1}^M v_k I_k,$$

where the Lagrange-like basis functions I_k are given by Lemma 3.2. The interpolation errors $\epsilon_j = f(\mathbf{t}_j) - v_j$ fulfill $\epsilon = \mathbf{K}\mathbf{v}$, where $\mathbf{K} \in \mathbb{C}^{M \times M}$ has the entries

$$K_{j,k} := \begin{cases} 0, & j = k, \\ I_k(\mathbf{t}_j), & j \neq k. \end{cases}$$

We proceed by $\|\epsilon\|_2 \leq \|\mathbf{K}\|_2 \leq \|\tilde{\mathbf{K}}\|_2$, where the second inequality follows from monotonicity of the norm [10, p. 520] (or [13, Lem. A.2]) and Lemma 3.2 i) and ii) with

$$\tilde{K}_{j,k} := \begin{cases} 0, & \mathbf{t}_j \in \Lambda_{l(k)}, \\ \frac{\beta^\beta \lambda^{\beta+\lambda-1}}{(2N|\mathbf{t}_k - \mathbf{t}_j|_{\mathbb{T}^d})^\beta} \mathcal{C}, & \text{otherwise.} \end{cases}$$

Since $\tilde{\mathbf{K}} \in \mathbb{R}^{M \times M}$ is symmetric, we bound the spectral norm by the maximum norm and apply the packing argument from Definition 2.3 ii) and Remark 2.4 ii) to get

$$\|\epsilon\|_2 \leq \max_{j=1, \dots, M} \sum_{\substack{k=1 \\ k \neq j}}^M \tilde{K}_{j,k} \leq \lambda 2^d (2^d - 1) \sum_{m=1}^{\lfloor N/2\rho \rfloor} m^{d-1} \max_{\mathbf{t} \in J_m} \frac{\beta^\beta \lambda^{\beta+\lambda-1}}{(2N|\mathbf{t}|_{\mathbb{T}^d})^\beta} \mathcal{C}$$

$$\leq 2^{d-\beta}(2^d - 1)\lambda^{\beta+\lambda}\beta^\beta\mathcal{C}\zeta(\beta - d + 1)\rho^{-\beta}.$$

Condition (4.1) and $\beta \geq 2$ imply $\|\epsilon\|_2 \leq \frac{1}{4\sqrt{2}}$. To bound the L^2 -norm of f , let $\hat{\mathbf{K}} := (|\langle I_k, I_j \rangle|)_{j,k=1,\dots,M} \in \mathbb{R}^{M \times M}$. The triangle inequality, symmetry of $\hat{\mathbf{K}}$, Lemma 3.2 iii), and the packing argument from Definition 2.3 ii) and Remark 2.4 ii) yield

$$\begin{aligned} \|f\|_{L^2(\mathbb{T}^d)}^2 &= \sum_{j=1}^M \sum_{k=1}^M v_k \bar{v}_j \langle I_k, I_j \rangle_{L^2(\mathbb{T}^d)} \leq \max_{\|\mathbf{w}\|_2=1} \mathbf{w}^* \hat{\mathbf{K}} \mathbf{w} \leq \|\hat{\mathbf{K}}\|_\infty \leq \max_j \sum_{k=1}^M |\langle I_k, I_j \rangle_{L^2(\mathbb{T}^d)}| \\ &\leq \frac{\lambda^d \beta^{d/2}}{N^d} \lambda^{2\lambda-2} \mathcal{C}^2 \left(\lambda + \lambda 2^d (2^d - 1) \sum_{m=1}^{\lfloor N/2\rho \rfloor} m^{d-1} \max_{\mathbf{t} \in J_m} \frac{\sqrt{\beta}}{2} \frac{(\lambda\beta)^\beta}{(N|\mathbf{t}|_{\mathbb{T}^d})^\beta} \right) \\ &\leq \frac{\lambda^d \beta^{d/2}}{N^d} \lambda^{2\lambda-1} \mathcal{C}^2 \left(1 + \lambda^\beta \beta^{\beta+1/2} 2^{d-1} (2^d - 1) \zeta(\beta - d + 1) \rho^{-\beta} \right). \end{aligned}$$

Condition (4.1) implies

$$\|f\|_{L^2(\mathbb{T}^d)} \leq \sqrt{\frac{3}{2}} \left(\frac{\lambda\sqrt{\beta}}{N} \right)^{d/2} \lambda^{\lambda-1/2} \mathcal{C}$$

and Lemma 2.2 finally the result. \blacksquare

For $d = 1$, Remark 2.4 iii) applied to the cluster complexity yields:

Corollary 4.2. *Under the assumptions of Theorem 4.1 with $d = 1$ and $\beta = 2$, we have*

$$\sigma_{\min}(\mathbf{A}) \geq \frac{1}{1.8(2e)^{\lambda-1}} \cdot \sqrt{N} \tau^{\lambda-1}.$$

Example 4.3 (Specific choices of β). *Specific choices of β in Theorem 4.1 yield the following:*

i) *By choosing $\beta = d + 1$ or $\beta = d + 2$ for d being odd or even, respectively, and some additional cosmetics, the condition*

$$\rho \geq 6d\lambda \left(\lambda^\lambda \mathcal{C} \right)^{\frac{1}{d+1}}$$

implies our best estimate

$$\sigma_{\min}(\mathbf{A}) \geq \left(3d^{d/4} \lambda^{\lambda+d/2-1/2} \right)^{-1} \frac{N^{d/2}}{\mathcal{C}}.$$

ii) *By choosing $\beta = 2 \lceil \frac{1}{2} \log(2^d(2^d - 1)\lambda^\lambda \zeta(2)\mathcal{C}) \rceil$ and noting that $2^{\beta/\beta} \leq 1.2$ for β even and ${}^{\log \mathcal{C}} \sqrt{\mathcal{C}} = e$, our weakest condition*

$$\rho \geq 3.3\lambda (2.5 + 1.4d + \lambda \log \lambda + \log \mathcal{C}),$$

implies

$$\sigma_{\min}(\mathbf{A}) \geq \left(1.5 \cdot (2.5 + 1.4d + \lambda \log \lambda + \log \mathcal{C})^{d/4} \lambda^{\lambda+d/2-1/2} \right)^{-1} \frac{N^{d/2}}{\mathcal{C}}.$$

Example 4.4 (Well separated nodes). For $\lambda = 1$, we have $\mathcal{C} = 1$ and the nodes are well separated. For $\rho \geq 6d$, Example 4.3 i) yields

$$\sigma_{\min}(\mathbf{A}) \geq \frac{N^{d/2}}{3d^{d/4}}.$$

Note that Theorem 4.1 always assumes $\rho \geq \beta \geq d + 1$. This compares to [12], where $\rho \geq 3 + 2 \log d$ already suffices for $\sigma_{\min}(\mathbf{A}) > 0$. Using Theorem 4.1 directly for $d = 1$ and $\beta = 2$, then $\rho \geq 4.4$ implies

$$\sigma_{\min}(\mathbf{A}) \geq \frac{\sqrt{N}}{1.8}.$$

This compares to [2, 19], which provide under the same condition on ρ , $\sigma_{\min}(\mathbf{A}) \geq \sqrt{N} \cdot \sqrt{1 - 1/\rho} \geq \sqrt{N}/1.14$.

Example 4.5 (Pair clusters). For $\lambda = 2$, we have $\mathcal{C} = 1/\tau$ and at most pairs of nodes form clusters. Example 4.3 i) with

$$\rho \geq 12d \left(\frac{4}{\tau} \right)^{\frac{1}{d+1}}$$

implies

$$\sigma_{\min}(\mathbf{A}) \geq \frac{\tau N^{d/2}}{12 \cdot 2^{d/2-1/2} \cdot d^{d/4}}.$$

Example 4.6 (Pair clusters, comparison). Let $d = 1$ and $\lambda = 2$. We apply Theorem 4.1 with $\beta = 2$, $\beta = 2 \left\lceil \frac{1}{2} \log \left(\frac{\pi^2}{3} \lambda^\lambda \mathcal{C} \right) \right\rceil$ and $\beta = 2\lambda$, respectively. These results are compared to [15, Thm. 1] (with minor corrections and where we simplified slightly $\lfloor n/\lambda \rfloor \approx n/\lambda$), to [13, Thm. 4.9] (under the additional assumption that all nodes inside the clusters have the same separation), and to [8, Cor. 4.2] (with a minor improvement for $\tau \leq 1$ and in estimating [8, Eq. (8)]). These comparisons are also presented in section 6.1 numerically.

Ref.	Thm. 4.1			[15]	[13]	[8]
$\rho \geq$	$\frac{17.3}{\sqrt{\tau}}$	$34.9 + 6.6 \log \tau $	$\frac{29}{\sqrt[3]{\tau}}$	$\frac{42.5 \sqrt[4]{M}}{\sqrt[3]{\tau}}$	$25(\log(\lfloor \frac{M}{4} \rfloor) + 1)$	3
$\sigma_{\min}(\mathbf{A}) \geq$	$\frac{\tau \sqrt{N}}{7.2}$	$\frac{\tau \sqrt{N}}{6 \sqrt[4]{5.3 + \log \tau }}$	$\frac{\tau \sqrt{N}}{8.6}$	$\frac{\tau \sqrt{N}}{4.5 \sqrt{M}}$	$\frac{\tau \sqrt{N}}{3.5}$	$\frac{\tau \sqrt{N}}{1.7}$

Example 4.7. (Comparison with [15]) Let $d = 1$ and $\beta = 2\lambda$, then $N > 2\lambda^3$ and $\rho \geq 4.4\lambda^{5/2}\mathcal{C}^{\frac{1}{2\lambda}}$ imply

$$\sigma_{\min}(\mathbf{A}) \geq \left(1.8 \cdot C_0^{\lambda-1} \cdot \lambda^{\lambda+1/4} \right)^{-1} \frac{\sqrt{N}}{\mathcal{C}},$$

where we set $C_0 = 1$ for the moment. This can be compared to [15, Thm. 1], where after minor corrections $N > 2\lambda^2$ and $\rho \geq 10\lambda^{5/2}(MC)^{\frac{1}{2\lambda}}$, imply

$$\sigma_{\min}(\mathbf{A}) \geq \left(1.5 \cdot C_0^{\lambda-1} \cdot \sqrt{M} \lambda^\lambda \right)^{-1} \frac{\sqrt{N}}{\mathcal{C}}.$$

According to Remark 3.3, $C_0 \in (\pi^{-1}, 1]$ depending on λ and n . In total, we have a stronger condition on N but our condition on ρ is always weaker and our estimate on $\sigma_{\min}(\mathbf{A})$ is sharper if $M > 2$. This comparison is also presented in Figure 6.2.

Example 4.8 (All nodes cluster). *Let $d = 1$ and $\lambda = M$. If $N > 8M$, then Corollary 4.2 implies*

$$\sigma_{\min}(\mathbf{A}) \geq \frac{1}{1.8(2e)^{M-1}} \cdot \sqrt{N}\tau^{M-1}.$$

This compares to [3], where the restriction of the nodes to an interval of length $1/(2M^2)$ and $N \geq 4M^3$ imply

$$\sigma_{\min}(\mathbf{A}) \geq \frac{1}{2^M M^{2M-1}} \cdot \sqrt{N}\tau^{M-1},$$

but, note that the definition of a clustered node configuration in [3] is in principle more flexible than ours.

5 Upper bounds and beyond distances

In this section, we show that the obtained lower bounds are sharp for $d = 1$ and for $\lambda = 2$, respectively. Moreover, we show for $d > 1$ and nodes in generic position (e.g. not all nodes on a line for $d = 2$), that the cluster complexity C is not the optimal quantity to understand the situation here. If we assume a normalized minimal separation τ between nodes, then the estimate in Theorem 4.1 is sub-optimal with respect to the order in τ we can derive from the cluster complexity. For this, we give an example with one cluster of three nodes in the bivariate case, $d = 2$.

Example 5.1 (Matching bounds for $d = 1$). *In the second version of [15, Prop. 3] an upper bound on $\sigma_{\min}(\mathbf{A})$ is given for a clustered node configuration that consists of at least one cluster of λ equispaced, τ separated nodes. After further simplifications, we can derive*

$$\min_{\Omega} \sigma_{\min}(\mathbf{A}) \leq (\pi\lambda)^{1/4} \pi^{\lambda-1} \sqrt{N} \tau^{\lambda-1} (1 + \tau C(\lambda) \sqrt{N}), \quad C(\lambda) := 2\pi \sum_{l=0}^{\lambda} \binom{\lambda-1}{l} \frac{l^\lambda}{\lambda!}.$$

Together with Remark 3.3 and Corollary 4.2 this assures that for sufficiently large $N \in \mathbb{N}$, small τ and $\lambda \geq 2$, there exist constants $c_1 \leq c_2$ such that

$$\sqrt{N} (c_1 \tau)^{\lambda-1} \leq \min_{\Omega} \sigma_{\min}(\mathbf{A}) \leq \sqrt{N} (c_2 \tau)^{\lambda-1},$$

where the minimum is taken over all clustered node configurations Ω with at least one cluster of λ nodes with normalized minimal separation τ .

This was also expected in [3, Rem. 3.5]. In particular note that the lower bound in Remark 2.4 iii) implies that the term λ^λ in Theorem 4.1 cannot be avoided.

Example 5.2 (Matching bounds for $\lambda = 2$). *Let $d \in \mathbb{N}$, $\lambda \geq 2$, and $\tau \leq 1$ be such that $|\mathbf{t}_1 - \mathbf{t}_2|_{\mathbb{T}^d} = \tau/N$, then the Cauchy interlacing theorem for eigenvalues ([10, Thm. 4.3.28]) and the binomial formula yield*

$$\begin{aligned} \sigma_{\min}(\mathbf{A})^2 &\leq N^d (1 - |d_n(\mathbf{t}_1 - \mathbf{t}_2)|) \leq N^d \left(1 - \left(1 - \frac{\pi^2 \tau^2}{6}\right)^d\right) \\ &\leq N^d \frac{\pi^2 \tau^2}{6} \sum_{k=0}^{d-1} \left(1 - \frac{\pi^2 \tau^2}{6}\right)^k \leq \frac{\pi^2 \tau^2 d N^d}{6}. \end{aligned}$$

Together with Example 4.5, there exists constants $c_1(d) \leq c_2(d)$ such that

$$c_1(d)N^{d/2}\tau \leq \min_{\Omega} \sigma_{\min}(\mathbf{A}) \leq c_2(d)N^{d/2}\tau,$$

where the minimum is taken over all clustered node configurations Ω with at least one cluster of $\lambda = 2$ nodes with normalized minimal separation τ .

Example 5.3 (Triple cluster). Let $d = 2$, $N \in \mathbb{N}$, and $\Omega = \{\mathbf{t}_1, \mathbf{t}_2, \mathbf{t}_3\} \subset [0, 1]^2$ with

$$\mathbf{t}_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \mathbf{t}_2 = \frac{\nu}{N} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \quad \mathbf{t}_3 = \frac{\nu}{N} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}, \quad \nu \in \left(0, \frac{1}{2}\right], \quad a_1^2 + a_2^2 = b_1^2 + b_2^2 = 1, \quad a_1 b_1 + a_2 b_2 \leq 0.$$

and hence, the normalized minimal separation of Ω is $\nu/\sqrt{2} \leq \tau \leq \nu$. Then the smallest singular value of the corresponding Vandermonde matrix \mathbf{A} fulfills

$$\sigma_{\min}(\mathbf{A}) = \begin{cases} \Theta(\tau^2), & \text{antipodal nodes, } a_1 b_1 + a_2 b_2 = -1, \\ \Theta(\tau), & \text{otherwise,} \end{cases}$$

and this can be seen as follows: Define the real matrix

$$\mathbf{M} := \begin{pmatrix} 1 & u & v \\ u & 1 & w \\ v & w & 1 \end{pmatrix} := \left(e^{-\pi i m(\mathbf{t}_j - \mathbf{t}_k)} d_m(\mathbf{t}_j - \mathbf{t}_k) \right)_{j,k=1,2,3},$$

note that $\sigma_{\min}(\mathbf{A})^2 = \sigma_{\min}(\mathbf{A}\mathbf{A}^*) = \sigma_{\min}(\mathbf{M}) = \|\mathbf{M}^{-1}\|_2^{-1}$, and use the explicit formula

$$\|\mathbf{M}^{-1}\|_2 = \frac{1}{|u^2 + v^2 + w^2 - 2uvw - 1|} \left\| \begin{pmatrix} w^2 - 1 & u - vw & v - uw \\ u - vw & v^2 - 1 & w - uv \\ v - uw & w - uv & u^2 - 1 \end{pmatrix} \right\|_2. \quad (5.1)$$

The univariate Taylor expansion

$$e^{-\pi i n \nu / N} d_n \left(\frac{\nu}{N} \right) = 1 - \alpha_n \nu^2 + \gamma_n \nu^4 + \mathcal{O}(\nu^6), \quad \alpha_n, \gamma_n \neq 0,$$

and $a_1^2 + a_2^2 = 1 = b_1^2 + b_2^2$ yield

$$u = 1 - \alpha_n \nu^2 + (\alpha_n^2 a_1^2 a_2^2 + \gamma_n (a_1^4 + a_2^4)) \nu^4 + \mathcal{O}(\nu^6)$$

and similar expressions for the other quantities. By direct computation, we see that the entries in the matrix on the right hand side of (5.1) are all $\mathcal{O}(\nu^2)$ and for example the diagonal entry $u^2 - 1$ is $\Theta(\nu^2)$ independent of \mathbf{a} and \mathbf{b} . Hence, the norm of that matrix is $\Theta(\nu^2)$. Similarly, the denominator of (5.1) can be computed to be

$$u^2 + v^2 + w^2 - 2uvw - 1 = \begin{cases} \mathcal{O}(\nu^6), & a_1 b_1 + a_2 b_2 = -1, \\ \Theta(\nu^4), & \text{otherwise.} \end{cases}$$

Finally, this yields

$$\sigma_{\min}(\mathbf{A}) = \begin{cases} \mathcal{O}(\nu^2), & a_1 b_1 + a_2 b_2 = -1, \\ \Theta(\nu), & \text{otherwise,} \end{cases}$$

and together with Theorem 4.1 the assertion.

6 Numerics

In this section we do four different experiments. Two of them are to compare our results with recent results from the literature ($d = 1$) and two of them underline our results from section 5. All computations were carried out using MATLAB R2017b.

6.1 Pair clusters

In order to compare our results (see Example 4.6) with the ones from the second version of [15, Thm. 2], [8] and [13], we set $d = 1$, $N = 2^{15} + 1$ ([13] requires odd N without further considerations), and take $M = 4$ and $M = 20$ nodes, respectively. The node configuration consists of uniformly placed clusters (at l/N , $l = 1, \dots, M/2$) that include two nodes each. The first cluster realizes the minimal separation τ , which is picked logarithmically uniformly at random from $[10^{-12}, 1]$, i.e. $t_1 = 0$ and $t_2 = \tau/N$. The further clusters have nodes $t_{2l} = l/N$ and $t_{2l+1} = (l + \delta)/N$ for $l = 1, \dots, (M - 1)/2$, where $\delta \in [\tau, 2\tau]$ (parameter $c = 2$ in [13, Thm. 4.7]) is picked uniformly randomly. Afterwards, we compute $\sigma_{\min}(\mathbf{A})$, where \mathbf{A} is the Vandermonde matrix defined in (2.1) corresponding to the node configuration. For each M we pick 50 instances of τ and the results are presented in Figure 6.1. This clustered node

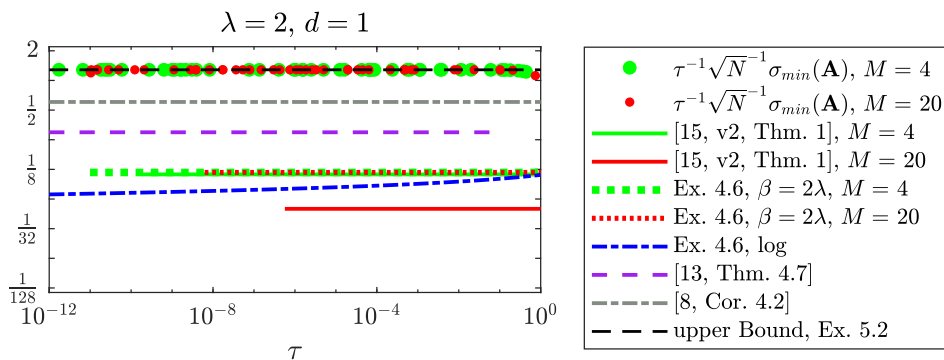


Figure 6.1: Comparison of different results for the case of pair clusters as in Example 4.6.

configuration fulfills $\rho \geq \frac{2N}{M} - 1$ independently of τ . Theorem 4.1 and the second version of [15, Thm. 1] make restrictions to τ through the condition on ρ . Therefore, choosing β logarithmically as in Example 4.3 ii) requires $\tau \geq e^{-\frac{35.9-2N/M}{6.6}}$, which is below 10^{-200} for both $M = 4$ and $M = 20$. The second version of [15] and our result with $\beta = 4$ requires respectively

$$\tau \geq \frac{43^4 M}{\rho^4} \approx \begin{cases} 1.9 \cdot 10^{-10}, & M = 4, \\ 5.9 \cdot 10^{-7}, & M = 20, \end{cases} \quad \tau \geq \frac{29^4}{\rho^4} \approx \begin{cases} 9.8 \cdot 10^{-12}, & M = 4, \\ 6.1 \cdot 10^{-9}, & M = 20. \end{cases}$$

6.2 Bigger clusters

In this numerical example, we confirm our results in the univariate case, $d = 1$, for bigger clusters of size $\lambda = 5$ and compare them with the results from the second version of [15]. The polynomial degree is set to $N = 2^{15}$. We build up clustered node configurations with $L = 2$ ($M = 10$) and $L = 10$ ($M = 50$) clusters placed equispaced at $\frac{l}{L}$ for $l = 0, \dots, L - 1$. At each cluster position the cluster nodes start to lie equispaced with separation $\frac{\tau}{N}$, where

$\tau \in [10^{-4}, 1/4]$ (the right hand interval bound is due to cluster lying in an interval of length $1/N$) is picked logarithmically uniformly at random. Afterwards the smallest singular value $\sigma_{\min}(\mathbf{A})$ is computed. This procedure is repeated 100 times for the respective choice of L and the results are presented in Figure 6.2. We use the statements from Example 4.7 with $C_0 = (1 - \frac{\pi^2}{3\lambda^2})^{-1/2} N/\lambda \lfloor n/\lambda \rfloor^{-1}$. Since $d = 1$, the worst case cluster complexity is estimated by (2.4) to $\mathcal{C} \leq \tau^{-4}/4$.

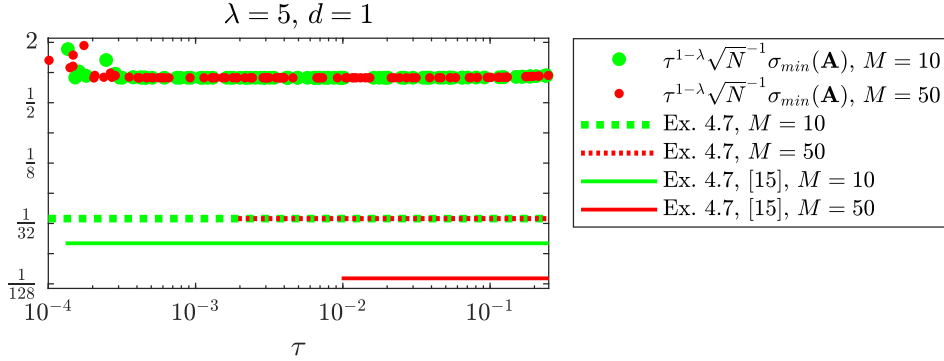


Figure 6.2: Node configurations with bigger clusters. Lower bounds on $\sigma_{\min}(\mathbf{A})$. Comparison with [15, v2, Thm. 1] as presented in Example 4.7 with estimate from Remark 3.3.

6.3 Pair clusters, bivariate

We present a numerical experiment in order to confirm our results for the higher dimensional case and set $d = 2$. Randomized clustered node configurations of $L = 2$, $L = 20$ and $L = 40$ clusters with 2 nodes each are constructed for 100 different minimal separations τ , respectively. Then the smallest singular values of the corresponding Vandermonde matrices $\sigma_{\min}(\mathbf{A})$ are computed and the upper bound from Example 5.2 and the lower bound from Example 4.5 are shown. The results are presented in Figure 6.3. The node configurations are built as follows. The minimal separation τ is picked logarithmically uniformly at random in $[10^{-3}, 1]$. We set $N = 10^3$ so that the condition on ρ in Example 4.5 together with the left interval bound for τ make $\rho \geq \rho_{\min}$ (value shown in the figure) necessary. Two clusters realize the cluster separation ρ_{\min} and for the remaining clusters, we pick a position in $[0, 1]^2$ uniformly at random. The positions are fixed for the respective choice of L and do not change for different τ . Each cluster is constructed randomly by setting one node to $(0, 0)$ and one to either $(a, 1)$ or $(1, a)$ for some $a \in [0, 1]$. Then we scale the clusters by τ and move them to their respective cluster positions.

6.4 One triple cluster, bivariate

Here we present a numerical experiment for Example 5.3. We set $N = 100$, $d = 2$ and build the triple cluster consisting of the nodes $\mathbf{t}_1 = (0, 0)^T$, $\mathbf{t}_2 = (-\sqrt{1 - a^2}\nu/N, a\nu/N)^T$ and $\mathbf{t}_3 = (\nu/N, 0)^T$ (see Figure 6.4, left), where $\tau = \nu\sqrt{1 - a^2} \in [10^{-6}, 1/2]$ is picked logarithmically uniformly at random. Then we compute the smallest singular value of the Vandermonde matrix $\sigma_{\min}(\mathbf{A})$. This is repeated 100 times for $a = 0.1$ and $a = 0$ each. The results are presented in Figure 6.4 (right). We see the asymptotic behavior with respect to τ calculated

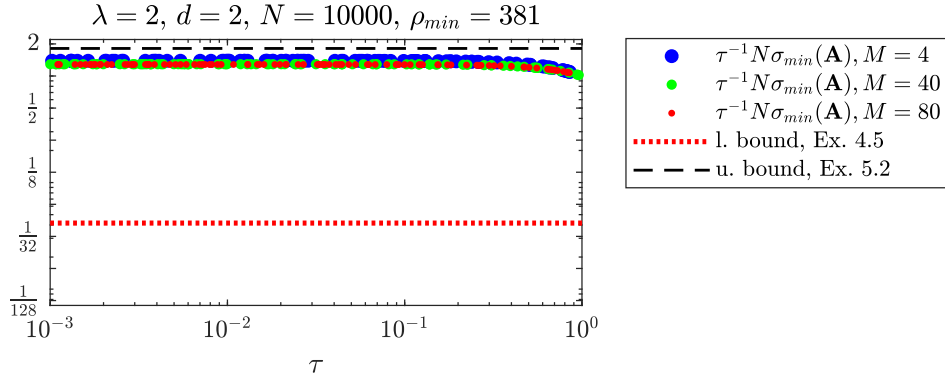


Figure 6.3: Upper and lower bounds on $\sigma_{\min}(\mathbf{A})$ for bivariate pair clusters as in Examples 4.5 and 5.2.

in Example 5.3. Furthermore, for nodes not being antipodal, we observe that the asymptotic starts when τ becomes smaller than the displacement parameter a .

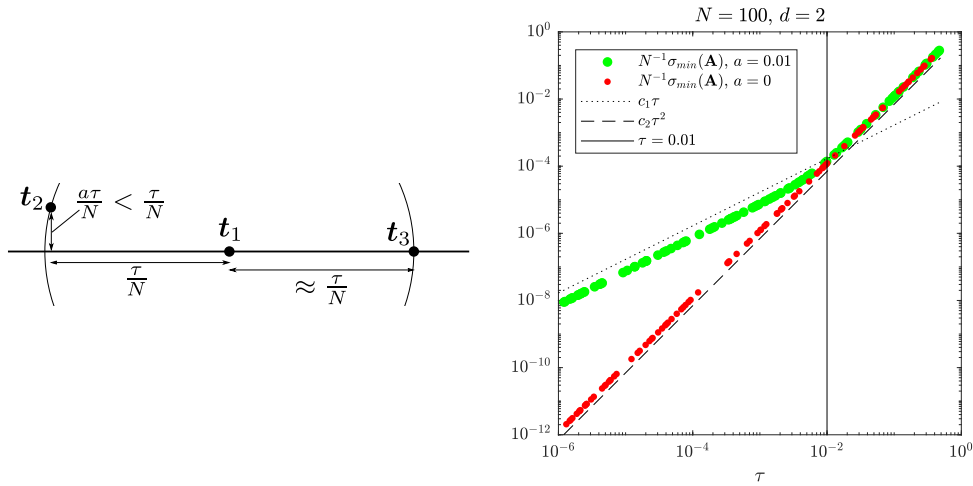


Figure 6.4: Triple cluster, almost antipodal nodes, cf. Example 5.3. Left: sketch of node positions. Right: numerical results.

Acknowledgements. The authors thank Jürgen Prestin for discussions on Lemma 3.1 and gratefully acknowledge support by the projects DFG-GK1916 and DFG-SFB944.

References

- [1] A. Akinshin, G. Goldman, and Y. Yomdin. Geometry of error amplification in solving Prony system with near-colliding nodes. *ArXiv e-prints*, Jan. 2017.
- [2] C. Aubel and H. Bölcskei. Vandermonde matrices with nodes in the unit disk and the large sieve. *Appl. Comput. Harmon. Anal.*, 47(1):53–86, jul 2019.
- [3] D. Batenkov, L. Demanet, G. Goldman, and Y. Yomdin. Stability of partial Fourier matrices with clustered nodes. *ArXiv e-prints*, Sept. 2018.

- [4] D. Batenkov, G. Goldman, and Y. Yomdin. Super-resolution of near-colliding point sources. *arXiv e-prints*, Apr. 2019.
- [5] F. S. V. Bazán. Conditioning of rectangular Vandermonde matrices with nodes in the unit disk. *SIAM J. Matrix Anal. Appl.*, 21(2):679–693, 1999.
- [6] B. G. R. de Prony. Essai expérimental et analytique: sur les lois de la dilatabilité de fluides élastique et sur celles de la force expansive de la vapeur de l’alkool, a différentes températures. *Journal de l’école polytechnique*, 1(22):24–76, 1795.
- [7] L. Demanet and N. Nguyen. The recoverability limit for superresolution via sparsity. *ArXiv e-prints*, Feb. 2015.
- [8] B. Diederichs. Well-Posedness of Sparse Frequency Estimation. *arXiv e-prints*, May 2019.
- [9] D. L. Donoho. Superresolution via sparsity constraints. *SIAM J. Math. Anal.*, 23(5):1309–1331, 1992.
- [10] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, New York, USA, 2nd edition, 2013.
- [11] Y. Hua and T. K. Sarkar. Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise. *IEEE Trans. Acoust. Speech Signal Process.*, 38(5):814–824, 1990.
- [12] S. Kunis, H. M. Möller, T. Peter, and U. von der Ohe. Prony’s method under an almost sharp multivariate Ingham inequality. *J. Fourier Anal. Appl.*, 24(5):1306–1318, 2018.
- [13] S. Kunis and D. Nagel. On the condition number of Vandermonde matrices with pairs of nearly-colliding nodes. *arXiv e-prints*, Dec. 2018.
- [14] S. Kunis and D. Potts. Stability results for scattered data interpolation by trigonometric polynomials. *SIAM J. Sci. Comput.*, 29:1403–1419, 2007.
- [15] W. Li and W. Liao. Stable super-resolution limit and smallest singular value of restricted Fourier matrices. *ArXiv e-prints*, Sept. 2017.
- [16] W. Li, W. Liao, and A. Fannjiang. Super-resolution limit of the ESPRIT algorithm. *arXiv e-prints*, May. 2019.
- [17] W. Liao and A. Fannjiang. MUSIC for single-snapshot spectral estimation: stability and super-resolution. *Appl. Comput. Harmon. Anal.*, 40(1):33–67, 2016.
- [18] L. Mattner and B. Roos. Maximal probabilities of convolution powers of discrete uniform distributions. *Statist. Probab. Lett.*, 78(17):2992 – 2996, 2008.
- [19] A. Moitra. Super-resolution, extremal functions and the condition number of Vandermonde matrices. In *STOC’15—Proceedings of the 2015 ACM Symposium on Theory of Computing*, pages 821–830. ACM, New York, 2015.
- [20] V. I. Morgenshtern and E. J. Candès. Super-resolution of positive sources: the discrete setup. *SIAM J. Imaging Sci.*, 9(1):412–444, 2016.

- [21] D. Potts and M. Tasche. Error estimates for the ESPRIT algorithm. In *Large truncated Toeplitz matrices, Toeplitz operators, and related topics*, volume 259 of *Oper. Theory Adv. Appl.*, pages 621–648. Birkhäuser/Springer, Cham, 2017.
- [22] R. Roy and T. Kailath. ESPRIT - estimation of signal parameters via rotational invariance techniques. *IEEE Trans. Acoust. Speech Signal Process.*, 37(7):984–995, 1989.
- [23] R. O. Schmidt. Multiple emitter location and signal parameter estimation. *IEEE Trans. Antennas and Propagation*, 34(3):276–280, 1986.