

Super-resolution of near-colliding point sources

DMITRY BATENKOV[†]

Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
Department of Applied Mathematics, School of Mathematical Sciences, Tel-Aviv University, P.O. Box 39040, Tel-Aviv 6997801, Israel

[†]Corresponding author. Email: dbatenkov@tauex.tau.ac.il

GIL GOLDMAN AND YOSEF YOMDIN

Department of Mathematics, The Weizmann Institute of Science, Rehovot 76100, Israel
Email: gil.goldman@weizmann.ac.il yosef.yomdin@weizmann.ac.il

[Received on 14 May 2019; revised on 20 January 2020; accepted on 18 February 2020]

We consider the problem of stable recovery of sparse signals of the form

$$F(x) = \sum_{j=1}^d a_j \delta(x - x_j), \quad x_j \in \mathbb{R}, \quad a_j \in \mathbb{C},$$

from their spectral measurements, known in a bandwidth Ω with absolute error not exceeding $\epsilon > 0$. We consider the case when at most $p \leq d$ nodes $\{x_j\}$ of F form a cluster whose extent is smaller than the Rayleigh limit $\frac{1}{\Omega}$, while the rest of the nodes is well separated. Provided that $\epsilon \lesssim \text{SRF}^{-2p+1}$, where $\text{SRF} = (\Omega \Delta)^{-1}$ and Δ is the minimal separation between the nodes, we show that the minimax error rate for reconstruction of the cluster nodes is of order $\frac{1}{\Omega} \text{SRF}^{2p-1} \epsilon$, while for recovering the corresponding amplitudes $\{a_j\}$ the rate is of the order $\text{SRF}^{2p-1} \epsilon$. Moreover, the corresponding minimax rates for the recovery of the non-clustered nodes and amplitudes are $\frac{\epsilon}{\Omega}$ and ϵ , respectively. These results suggest that stable super-resolution is possible in much more general situations than previously thought. Our numerical experiments show that the well-known matrix pencil method achieves the above accuracy bounds.

Keywords: signal reconstruction; spike-trains; Fourier transform; Prony systems; sparsity; super-resolution.

1. Introduction

1.1 Super-resolution of sparse signals

The problem of mathematical super-resolution (SR) is to extract the fine details of a signal from band-limited and noisy measurements of its Fourier transform [41]. It is an inverse problem of great theoretical and practical interest.

The specifics of SR highly depend on the type of prior information assumed about the signal structure. Many theoretical and practical studies assume *signals of compact support*, in which case the SR problem is equivalent to analytic continuation (equivalently, extrapolation) of the Fourier transform. However, it can be shown that the spectrum of a compactly supported function can be extrapolated from samples of accuracy ϵ by a factor that scales at most logarithmically with the signal-to-noise ratio $\frac{1}{\epsilon}$, see e.g. [9, 18, 34, 41] and the references therein. On the other hand, in recent years, considerable progress

has been made in studying SR for *sparse signals*, which are frequently modeled as idealized *spike-trains*

$$F(x) = \sum_{j=1}^d a_j \delta(x - x_j), \quad x_j \in \mathbb{R}, \quad (1.1)$$

where δ is the ubiquitous Dirac's δ -distribution. This particular type of signals is widely used in the literature, as it is believed to capture the essential difficulty of SR with sparse priors, see e.g. [21, 26].

Let $\mathcal{F}(F)$ denote the Fourier transform of F :

$$\mathcal{F}(F)(s) = \int_{-\infty}^{\infty} F(x) e^{-2\pi i s x} dx. \quad (1.2)$$

Further, suppose that the spectral data are given as a function Φ satisfying, for some $\epsilon > 0$ and $\Omega > 0$,

$$|\Phi(s) - \mathcal{F}(F)(s)| \leq \epsilon, \quad s \in [-\Omega, \Omega]. \quad (1.3)$$

The sparse SR problem reads as follows: given Φ as above, estimate the unknown parameters of F , namely the *amplitudes* $\{a_j\}$ and the *nodes* $\{x_j\}$.

If $\epsilon = 0$, the problem can be solved exactly by a variety of parametric methods (Prony's method etc., see e.g. [51, 54] and Subsection 1.2 below). For $\epsilon > 0$, if $\mathfrak{f}$ is any reconstruction algorithm receiving Φ as an input and producing an estimate $F' = \mathfrak{f}(\Phi)$ of the signal that satisfies (1.3), then, under an appropriate definition of the distance $\|F - F'\|$, it is of great interest to have a good estimate of the noise amplification factor (or the *problem condition number*) $\mathcal{K}$ such that

$$\|F - F'\| \approx \mathcal{K} \epsilon. \quad (1.4)$$

1.2 Rayleigh limit and minimal separation

It has been well established that the difficulty of sparse SR is directly related to the *minimal separation* $\Delta = \min_{1 \leq i < j \leq d} |x_i - x_j|$ or, more precisely, to the relationship between Δ and Ω .

Without any *a priori* information, the best attainable resolution from spectral data of bandwidth Ω is of the order $\frac{1}{\Omega}$, which is also known as the *Rayleigh limit*. Both classical methods of non-parametric spectral estimation [54], as well as modern convex optimization-based methods, solve the problem under some sort of a separation condition of the form $\Delta \geq \frac{c}{\Omega}$ [5, 17, 19–21, 27, 28, 35, 53, 55], and moreover, these methods are generally considered to be stable.

On the other hand, the case $\Delta \ll \frac{1}{\Omega}$ (and arbitrary signed/complex amplitudes $\{a_j\}$) is much more difficult (Fig. 1).

The sparse SR problem has appeared already in the work by R. Prony [51], where he devised an algebraic scheme to recover the parameters $\{x_j, a_j\}$ from $2d$ equispaced measurements of $\mathcal{F}(F)$, assuming F is given by (1.1), and for arbitrary $\Delta > 0$ and $|a_j| > 0$ (see Proposition A.2 below). Since then, Prony's method and its various extensions and generalizations have been used extensively in applied and pure mathematics and engineering ([4, 48–50, 54, 57] and the references therein). While these methods provide exact recovery for $\epsilon = 0$, the question of their stability (the magnitude of $\mathcal{K}$ in (1.4)) becomes of essential interest. For instance, if it so happens that an estimate $F' = \sum_{j=1}^d a'_j \delta(x - x'_j)$

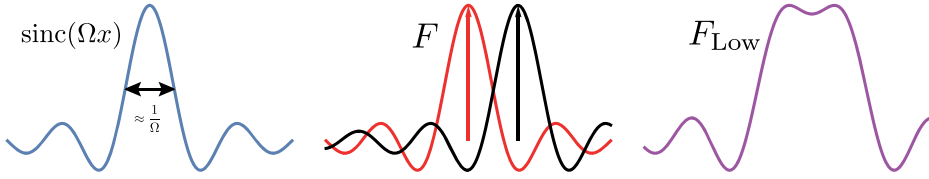


FIG. 1. The Rayleigh limit. For a signal $F(x) = \sum_j a_j \delta(x - x_j)$, its low resolution version is given by $F_{\text{Low}}(x) = \mathcal{F}^{-1}(\mathcal{F}(F) \cdot \chi_{[-\Omega, \Omega]}) \propto \sum_j a_j \text{sinc}(\Omega(x - x_j))$. $F_{\text{Low}}(x)$ will have peaks of width $\approx \frac{1}{\Omega}$, and therefore it will be increasingly difficult to recover signals for which the minimal separation between the $\{x_j\}$'s is much smaller than $\frac{1}{\Omega}$.

satisfies¹ $\min_{1 \leq j \leq d} |x'_j - x_j| \gtrsim \Delta$, then such F' may be of little practical use in many applications (because the inner structure of the sparse signal will be determined incorrectly).

The first work which examined the stability of SR in the sub-Rayleigh regime was by D. Donoho [26]. The signal F was assumed to have an infinite number of spikes $\{x_j\}$, constrained to a grid of step size Δ , with less than one spike per unit interval on average, but whose local complexity was constrained to have no more than d spikes per any interval of length d (such d is called the *Rayleigh index*). It was shown that the worst-case ℓ_2 error of such F (i.e. the ℓ_2 norm of the coefficient sequence of the difference) from continuous measurements with a band-limit Ω and perturbation of size ϵ (in L_2 sense) scales like $\text{SRF}^\alpha \epsilon$, where $\text{SRF} = \frac{1}{\Omega \Delta} > 1$ is the so-called *super-resolution factor*, and α satisfies $2d - 1 \leq \alpha \leq 2d + 1$. In [24], the authors considered the case of d -sparse signals supported on a grid and showed that the correct exponent should be $\alpha = 2d - 1$ in this case. In another recent work [39], the same scaling was shown to hold in the case of d -sparse signals and discrete Fourier measurements.

In the papers mentioned above, the error rate $\text{SRF}^{2d-1} \epsilon$ is *minimax*, meaning that on one hand, it is attained by a certain algorithm for all signals of interest, and on the other hand, there exist worst-case examples for which no algorithm can achieve an essentially smaller error. It turns out that these worst-case signals all have the structure of a cluster, where all the d nodes $\{x_j\}$ appear consecutively, i.e. $x_j = x_1 + (j - 1)\Delta$, $j = 1, \dots, d$. A natural question which arises is *if it is a priori known that only a subset of the d spikes can become clustered, can we have better reconstruction accuracy?* In this paper, we shall provide a positive answer to this question.

1.3 Main contributions

In this paper, we consider the case where the nodes $\{x_j\}$ can take arbitrary real values (the so-called *off-grid* setting), while the amplitudes $\{a_j\}$ can be arbitrary complex scalars. We further assume that exactly p nodes, $x_\kappa, \dots, x_{\kappa+p-1}$, form a small cluster of extent $h \ll \frac{1}{\Omega}$ and are approximately uniformly distributed inside the cluster, while the rest of the nodes are well separated from the cluster and from each other (see Definition 2.5 below). The approximate uniformity is expressed by the assumption that the minimal separation between any two cluster nodes is bounded from below by $\Delta = \tau h$ for some fixed $0 < \tau \leq 1$. Under these *p-clustered* assumptions, we show in Theorem 2.8 that for small enough ϵ – and, in particular, for $\epsilon \lesssim (\Omega \Delta)^{2p-1}$, the worst-case error rates of a minimax reconstruction algorithm (see

¹ We use the symbols $\lesssim, \gtrsim$ and $\asymp$ to denote order equivalence, up to constants: $A(t) \lesssim B(t)$ (or $B(t) \gtrsim A(t)$), if and only if there exists a positive constant c (depending on the specified parameters) such that $A(t) \leq cB(t)$ for all specified values of t . $A(t) \asymp B(t)$ if $A(t) \lesssim B(t)$ and also $B(t) \lesssim A(t)$.

Definition 2.2 below), receiving Φ satisfying (1.3) as an input, and returning an estimate $x'_j = x'_j(\Phi)$, $a'_j = a'_j(\Phi)$, satisfy

1. Non-cluster nodes:

$$\max_{j \notin \{\kappa, \dots, \kappa+p-1\}} |x_j - x'_j| \asymp \frac{\epsilon}{\Omega},$$

$$\max_{j \notin \{\kappa, \dots, \kappa+p-1\}} |a_j - a'_j| \asymp \epsilon.$$

2. Cluster nodes:

$$\max_{j \in \{\kappa, \dots, \kappa+p-1\}} |x_j - x'_j| \asymp \frac{\epsilon}{\Omega} (\Omega \Delta)^{-2p+2},$$

$$\max_{j \in \{\kappa, \dots, \kappa+p-1\}} |a_j - a'_j| \asymp \epsilon (\Omega \Delta)^{-2p+1}.$$

The constants appearing in our bounds depend on p, d , *a priori* bounds on the magnitudes $|a_j|$, and additional geometric parameters, but neither on Δ nor on Ω .

Our results indicate, in particular, that the non-clustered nodes $\{x_j\}_{j \notin \{\kappa, \dots, \kappa+p-1\}}$ can be recovered with much better accuracy than the cluster nodes. Let the SR factor be defined, as before, by $\text{SRF} = (\Omega \Delta)^{-1}$, then the condition number of the cluster nodes scales like SRF^{2p-1} in the SR regime $\text{SRF} \gg 1$, while the condition number of the non-cluster nodes does not depend on the SRF at all.

Our approach is to reduce the continuous measurements problem to a certain ‘Prony-type’ system of $2d$ nonlinear equations, given by equispaced measurements of $\Phi(s)$ with a carefully chosen spacing $\lambda \approx \Omega$, and analyze the sensitivity of this system to perturbations. The proofs involve techniques from quantitative singularity theory and numerical analysis. Some of the tools, in particular the ‘decimation-and-blowup’ technique, were previously developed in [1, 2, 6–8, 11–13]. The single-cluster case $p = d$ has been first analyzed in [7], while the lower bound (in a slightly less general formulation) has been essentially shown in [1]. One of the main technical results, Lemma 5.7, has been first proven in [8].

Our numerical experiments in Section 3 show that the above bounds are attained by matrix pencil (MP), a well-known, high-resolution algorithm [36, 37].

1.4 Related work and discussion

Our main results generalize several previously available bounds for both on-grid and off-grid SR [7, 24, 39], replacing the overall sparsity d with the ‘local’ sparsity p .² Compared with previous works, we also have an explicit control of the perturbation ϵ for which the stability bounds hold $\epsilon \leq C \cdot (\Omega \Delta)^{2p-1}$. So, given F satisfying the clustering assumptions and Ω , we can choose $\epsilon = c (\Omega \Delta)^{2p-1}$ such that F can be accurately resolved and c does not depend on Ω, Δ . But this also means that given $\epsilon > 0$, we can choose Δ_0 and Ω_0 such that $(\Omega_0 \Delta_0)^{2p-1} \geq \frac{\epsilon}{c}$, and for any F satisfying the clustering assumptions with $\Delta = \Delta_0$ and $\Omega = \Omega_0$, the SR problem can be accurately solved. Therefore, fixing ϵ , our results show that accurate recovery is possible for all SRF values up to $\left(\frac{1}{\epsilon}\right)^{\frac{1}{2p-1}}$ (but possibly also for higher

² Our clustering model is distinct from Donoho’s model of *sparse clumps* on a grid [26], and so the two results cannot be compared directly.

values of SRF). On the other hand, a similar argument using the lower bounds for the minimax error shows that with perturbation of magnitude ϵ , no algorithm can resolve signals having a cluster of size p and separation $\Delta_\epsilon \lesssim \frac{1}{\Omega} \epsilon^{\frac{1}{2p-1}}$, giving an upper bound for the attainable SRF values exactly matching the lower bound above. To summarize, we obtain *the best possible scaling of the attainable resolution with clustered sparsity p and absolute perturbation ϵ :*

$$\text{SRF} \asymp \sqrt[2p-1]{\frac{1}{\epsilon}}. \quad (1.5)$$

This Hölder-type scaling is much more favorable compared with SR by analytic continuation under the prior of compact signal support, where the bandwidth extrapolation factor scales only as a fractional power of $\log \frac{1}{\epsilon}$, see e.g. [9] and the references therein. Also, note that the sparse SR problem enjoys linear stability in ϵ (1.4), whereas analytic continuation exhibits stability of the form $\text{Error} \approx \epsilon^\gamma$, where $\gamma < 1$ [9, 18].

Stable SR in the on-grid setting of [24, 26, 39] is closely related to the smallest singular value of a certain class of Fourier-type matrices. Using the decimation technique (see also [22, 23]), in a recent paper [8], we have derived novel estimates³ for this quantity under the partial clustering setting (compare with [3, 15, 29, 38, 45]), and using these results, we have shown in the same paper that the asymptotic scaling of the condition number for on-grid SR in this regime is SRF^{2p-1} , matching the off-grid setting of the present paper.

The question of providing rigorous performance guarantees for high-resolution algorithms such as MP, MUSIC, ESPRIT and others, in the SR regime $\text{SRF} > 1$, is of current interest. In two very recent works, [39, 40], the authors derive stability estimates for MUSIC and ESPRIT algorithms under similar clustering assumptions, finite sampling and white Gaussian perturbation model. Their results suggest that the corresponding noise amplification factors $\mathcal{K}$ for the nodes are of the order SRF^{2p-2} with high probability. During the review of the present paper, the authors of [40] established near-optimality of ESPRIT in the bounded noise model. In particular, they showed that ESPRIT is optimal up to a factor of $1/\Omega$, i.e. $|x_j - \tilde{x}_j| \lesssim (\Omega \Delta)^{-2p-2} \epsilon$ with discrete Fourier measurements, however, requiring $\epsilon \lesssim (\Omega \Delta)^{4p-3}/\Omega$. We also mention [16, 32], where the connection between perturbation of (square) MP eigenvalues and the *a priori* distribution of these eigenvalues was established via potential theory. It will be interesting to investigate the possibility to applying these methods to the analysis of MP in the clustered setting.

Turning to other techniques, the special case of a single cluster can be solved with optimal accuracy by polynomial homotopy methods, as described in [6]; however, in order to generalize this algorithm to configurations with non-cluster nodes, we need to know the optimal decimation parameter λ . Nonlinear least squares and related methods (e.g. Variable Projections [33, 47]) apparently provide an optimal recovery rate; however, they generally require very accurate initialization. We hope that our methods may help in analyzing these techniques as well and plan to pursue this line of research in the future. For the case of positive point sources, stability rate SRF^{2p} has been established for convex optimization techniques in [46], see also a related preprint [25].

³ Estimates for the smallest singular value were independently obtained in [39] giving same asymptotic order but better absolute constants. In [10], we have obtained optimal scalings of all the singular values by different techniques.

1.5 Organization of the paper

In Section 2, we provide the necessary definitions and formulate the main results. In Section 3, we present several numerical experiments confirming the optimality of the MP algorithm. The proof of Theorem 2.6 (upper bound) is presented in Section 5. The proof of Theorem 2.7 (lower bound) is given in Section 6.

2. Minimax bounds for clustered SR

2.1 Notation and preliminaries

We shall denote by $\mathcal{P}_d$ the parameter space of signals F with complex amplitudes and real, pairwise, distinct and ordered nodes,

$$\mathcal{P}_d = \left\{ (\mathbf{a}, \mathbf{x}) : \mathbf{a} = (a_1, \dots, a_d) \in \mathbb{C}^d, \mathbf{x} = (x_1, \dots, x_d) \in \mathbb{R}^d, x_1 < x_2 < \dots < x_d \right\},$$

and identify signals F with their parameters $(\mathbf{a}, \mathbf{x}) \in \mathcal{P}_d$. In particular, this induces a structure of a linear space on $\mathcal{P}_d$. Throughout this text, we will always use the maximum norm $\|\cdot\| = \|\cdot\|_\infty$ on $\mathbb{C}^d$, $\mathbb{R}^d$ and $\mathcal{P}_d$, where for $F = (\mathbf{a}, \mathbf{x}) \in \mathcal{P}_d$

$$\|F\| = \max(\|\mathbf{a}\|_\infty, \|\mathbf{x}\|_\infty).$$

We shall denote the orthogonal coordinate projections of a signal F to the j -th node and j -th amplitude, respectively, by $P_{\mathbf{x}j} : \mathcal{P}_d \rightarrow \mathbb{R}$ and $P_{\mathbf{a}j} : \mathcal{P}_d \rightarrow \mathbb{C}$. We shall also denote the j -th component of a vector $\mathbf{v}$ by $\mathbf{v}_j$.

Let $L_\infty[-\Omega, \Omega]$ denote the space of bounded complex-valued functions defined on $[-\Omega, \Omega]$ with the norm $\|e\| = \max_{|s| \leq \Omega} |e(s)|$.

DEFINITION 2.1 Given $\Omega > 0$ and $U \subseteq \mathcal{P}_d$, we denote by $\mathfrak{F}(\Omega, U)$ the class of all admissible reconstruction algorithms, i.e.

$$\mathfrak{F}(\Omega, U) = \left\{ \mathfrak{f} : L_\infty[-\Omega, \Omega] \rightarrow U \right\}.$$

DEFINITION 2.2 Let $U \subset \mathcal{P}_d$. We consider the minimax error rate in estimating a signal $F \in U$ ⁴ from Ω -bandlimited data as in (1.3), with measurement error $\epsilon > 0$:

$$\mathcal{E}(\epsilon, U, \Omega) = \inf_{\mathfrak{f} \in \mathfrak{F}(\Omega, U)} \sup_{F \in U} \sup_{\|e\| \leq \epsilon} \|F - \mathfrak{f}(\mathcal{F}(F) + e)\|.$$

⁴ To ensure the minimax error rate is finite, depending on the noise level, we impose constraints on $U \subset \mathcal{P}_d$, namely lower and upper bounds on the magnitude of the amplitudes and the separation of the nodes. We will specify these constraints exactly in the statements of the accuracy bounds.

Similarly, the minimax errors of estimating the individual nodes, respectively, the amplitudes of $F \in U$ are defined by

$$\begin{aligned}\mathcal{E}^{\mathbf{x}j}(\epsilon, U, \Omega) &= \inf_{f \in \mathfrak{F}(\Omega, U)} \sup_{F \in U} \sup_{\|e\| \leq \epsilon} \left| P_{\mathbf{x}j}(F) - P_{\mathbf{x}j}(f(\mathcal{F}(F) + e)) \right|, \\ \mathcal{E}^{\mathbf{a}j}(\epsilon, U, \Omega) &= \inf_{f \in \mathfrak{F}(\Omega, U)} \sup_{F \in U} \sup_{\|e\| \leq \epsilon} \left| P_{\mathbf{a}j}(F) - P_{\mathbf{a}j}(f(\mathcal{F}(F) + e)) \right|.\end{aligned}$$

Let a signal $F \in \mathcal{P}_d$ be fixed. We define the ϵ -error set $E_{\epsilon, \Omega}(F)$ as the following pre-image.

DEFINITION 2.3 The error set $E_{\epsilon, \Omega}(F) \subset \mathcal{P}_d$ is the set consisting of all the signals $F' \in \mathcal{P}_d$ with

$$\left| \mathcal{F}(F')(s) - \mathcal{F}(F)(s) \right| \leq \epsilon, \quad s \in [-\Omega, \Omega].$$

We will denote by $E_{\epsilon}^{\mathbf{x}j}(F) = E_{\epsilon, \Omega}^{\mathbf{x}j}(F)$ and $E_{\epsilon}^{\mathbf{a}j}(F) = E_{\epsilon, \Omega}^{\mathbf{a}j}(F)$ the projections of the error set onto the individual nodes and the amplitudes components, respectively:

$$\begin{aligned}E_{\epsilon, \Omega}^{\mathbf{x}j}(F) &= \left\{ \mathbf{x}'_j \in \mathbb{R} : (\mathbf{a}', \mathbf{x}') \in E_{\epsilon, \Omega}(F) \right\} \equiv P_{\mathbf{x}j}E_{\epsilon, \Omega}(F), \\ E_{\epsilon, \Omega}^{\mathbf{a}j}(F) &= \left\{ \mathbf{a}'_j \in \mathbb{C} : (\mathbf{a}', \mathbf{x}') \in E_{\epsilon, \Omega}(F) \right\} \equiv P_{\mathbf{a}j}E_{\epsilon, \Omega}(F).\end{aligned}\tag{2.1}$$

For any subset V of a normed vector space with norm $\|\cdot\|$, the diameter of V is

$$\text{diam}(V) = \sup_{\mathbf{v}', \mathbf{v}'' \in V} \|\mathbf{v}' - \mathbf{v}''\|.$$

The minimax errors are directly linked to the diameter of the corresponding projections of the error set by the following easy computation, which is standard in the theory of optimal recovery [42–44] (see also [24, 26, 39]).

PROPOSITION 2.4 For $U \subset \mathcal{P}_d$, $\Omega > 0$, $1 \leq j \leq d$ and $\epsilon > 0$ we have

$$\frac{1}{2} \sup_{F: E_{\frac{1}{2}\epsilon, \Omega}(F) \subseteq U} \text{diam}(E_{\frac{1}{2}\epsilon, \Omega}^{\mathbf{x}j}(F)) \leq \mathcal{E}^{\mathbf{x}j}(\epsilon, U, \Omega) \leq \sup_{F \in U} \text{diam}(E_{2\epsilon, \Omega}(F))\tag{2.2}$$

$$\frac{1}{2} \sup_{F: E_{\frac{1}{2}\epsilon, \Omega}(F) \subseteq U} \text{diam}(E_{\frac{1}{2}\epsilon, \Omega}^{\mathbf{x}j}(F)) \leq \mathcal{E}^{\mathbf{x}j}(\epsilon, U, \Omega) \leq \sup_{F \in U} \text{diam}(E_{2\epsilon, \Omega}^{\mathbf{x}j}(F))\tag{2.3}$$

$$\frac{1}{2} \sup_{F: E_{\frac{1}{2}\epsilon, \Omega}(F) \subseteq U} \text{diam}(E_{\frac{1}{2}\epsilon, \Omega}^{\mathbf{a}j}(F)) \leq \mathcal{E}^{\mathbf{a}j}(\epsilon, U, \Omega) \leq \sup_{F \in U} \text{diam}(E_{2\epsilon, \Omega}^{\mathbf{a}j}(F))\tag{2.4}$$

Proof. We shall prove (2.2), the proof in the other cases is identical. We omit Ω from the following to reduce clutter.

Upper bound Let $\epsilon > 0$. For any $\Phi \in L_\infty[-\Omega, \Omega]$, let

$$\mathcal{B}(\epsilon, \Phi) = \{F \in U : \|\mathcal{F}(F) - \Phi\| \leq \epsilon\}.$$

Consider an oracle estimator $\mathfrak{f}_\epsilon \in \mathfrak{F}(\Omega, U)$ defined as

$$\mathfrak{f}_\epsilon(\Phi) = \begin{cases} \text{any element of } \mathcal{B}(\epsilon, \Phi) & \text{if } \mathcal{B}(\epsilon, \Phi) \neq \emptyset, \\ F_0 & \text{else,} \end{cases}$$

where F_0 is an arbitrary element of U . Now let $F \in U$ and $\Phi = \mathcal{F}(F) + e$ where $\|e\| \leq \epsilon$. Then, by definition $F \in \mathcal{B}(\epsilon, \Phi)$. Put $F' = \mathfrak{f}_\epsilon(\Phi)$, thus $\|\mathcal{F}(F') - \Phi\| \leq \epsilon$, and therefore

$$\|\mathcal{F}(F') - \mathcal{F}(F)\| \leq \|\mathcal{F}(F') - \Phi\| + \|\Phi - \mathcal{F}(F)\| = 2\epsilon.$$

We conclude that $F' \in E_{2\epsilon}(F)$ and consequently $\mathcal{E}(\epsilon, U, \Omega) \leq \|F - F'\| \leq \text{diam}(E_{2\epsilon}(F))$.

Lower bound For the lower bound, let $F \in U$ such that $E_{\frac{1}{2}\epsilon}(F) \subseteq U$. Let $\xi > 0$ small enough be fixed. There exist $F^1, F^2 \in E_{\frac{1}{2}\epsilon}(F)$ with $\|F^1 - F^2\| = \text{diam}(E_{\frac{1}{2}\epsilon}(F)) - \xi$. Let $\Phi = \mathcal{F}(F)$, and let $F' = \mathfrak{f}(\Phi)$ be the output of a certain estimator $\mathfrak{f}$ corresponding to the input Φ . We have $\|\Phi - \mathcal{F}(F^1)\|, \|\Phi - \mathcal{F}(F^2)\| \leq \epsilon$. Consequently, there exist perturbation functions e_1, e_2 satisfying $\|e_1\|, \|e_2\| \leq \epsilon$, while also

$$\mathcal{F}(F') = \Phi = \mathcal{F}(F^1) + e_1 = \mathcal{F}(F^2) + e_2.$$

By definition of the minimax error, we therefore have

$$\begin{aligned} \mathcal{E}(\epsilon, U, \Omega) &= \inf_{\mathfrak{f}} \sup_{\|e\| < \epsilon, F \in U} \|F - \mathfrak{f}(\mathcal{F}(F) + e)\| \\ &\geq \inf_{\mathfrak{f}} \max \left(\|F^1 - F'\|, \|F^2 - F'\| \right) \\ &\geq \inf_{\mathfrak{f}} \frac{1}{2} \left\{ \|F^1 - F'\| + \|F^2 - F'\| \right\} \\ &\geq \frac{1}{2} \|F^1 - F^2\| \\ &= \frac{1}{2} \text{diam}(E_{\frac{1}{2}\epsilon}(F)) - \frac{\xi}{2}. \end{aligned}$$

The lower bound follows by letting $\xi \rightarrow 0$. □

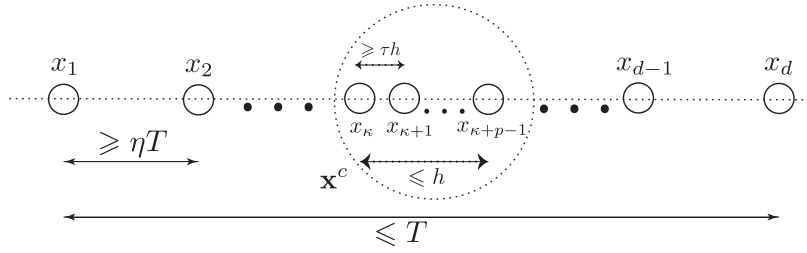


FIG. 2. A sketch of a uniform (p, h, T, τ, η) -clustered configuration $\mathbf{x} = (x_1, \dots, x_d)$ as in Definition 2.5.

2.2 Uniform estimates of minimax error for clustered configurations

The main goal of this paper is to estimate $\mathcal{E}(\epsilon, U, \Omega)$ (in fact its component-wise analogues $\mathcal{E}^{\mathbf{x},j}(\epsilon, U, \Omega)$ and $\mathcal{E}^{\mathbf{a},j}(\epsilon, U, \Omega)$) where $U \subset \mathcal{P}_d$ are certain compact subsets of $\mathcal{P}_d$ containing signals with $p \leq d$ nodes forming a small, approximately uniform, cluster. In order to have explicit bounds, we describe such sets U by additional parameters T, h, τ, η, m, M as follows.

DEFINITION 2.5 (Uniform cluster configuration, Fig. 2). Given $0 < \tau, \eta \leq 1$ and $0 < h \leq T$, a node vector $\mathbf{x} = (x_1, \dots, x_d) \in \mathbb{R}^d$ is said to form a (p, h, T, τ, η) -clustered configuration, if there exists a subset of p nodes $\mathbf{x}^c = \{x_{\kappa}, \dots, x_{\kappa+p-1}\} \subset \mathbf{x}$, $p \geq 2$, which satisfies the following conditions:

1. for each $x_j, x_k \in \mathbf{x}^c, j \neq k$,

$$\tau h \leq |x_j - x_k| \leq h;$$

2. for $x_\ell \in \mathbf{x} \setminus \mathbf{x}^c$ and $x_j \in \mathbf{x}, \ell \neq j$,

$$\eta T \leq |x_\ell - x_j| \leq T.$$

Our first main result provides an upper bound on $\text{diam}(E_{\epsilon, \Omega}(F))$, and its coordinate projections, for any signal F forming a clustered configuration as above.

THEOREM 2.6 (Upper bound) Let $F = (\mathbf{a}, \mathbf{x}) \in \mathcal{P}_d$, such that $\mathbf{x}$ forms a (p, h, T, τ, η) -clustered configuration and $0 < m \leq \|\mathbf{a}\|$. Then, there exist positive constants $C_1, \dots, C_5$, depending only on d, p, m , such that for each $\frac{C_4}{\eta T} \leq \Omega \leq \frac{C_5}{h}$ and $\epsilon \leq C_3(\Omega \tau h)^{2p-1}$, it holds that:

$$\begin{aligned} \text{diam}(E_{\epsilon, \Omega}^{\mathbf{x},j}(F)) &\leq \frac{C_1}{\Omega} \epsilon \times \begin{cases} (\Omega \tau h)^{-2p+2}, & x_j \in \mathbf{x}^c, \\ 1, & x_j \in \mathbf{x} \setminus \mathbf{x}^c; \end{cases} \\ \text{diam}(E_{\epsilon, \Omega}^{\mathbf{a},j}(F)) &\leq C_2 \epsilon \times \begin{cases} (\Omega \tau h)^{-2p+1}, & x_j \in \mathbf{x}^c, \\ 1, & x_j \in \mathbf{x} \setminus \mathbf{x}^c. \end{cases} \end{aligned}$$

REMARK 2.1 Our main focus is to investigate the error rates of the SR problem as the cluster size becomes small. Fixing the parameters p, d, m , the range of admissible Ω in Theorem 2.6, $\frac{C_4}{\eta T} \leq \Omega \leq \frac{C_5}{h}$, is non-empty for a sufficiently small cluster size h . Furthermore, we comment here that the constants C_4, C_5 actually only depend on d .

The above estimates are order optimal, as our next main theorem shows. For simplicity and without loss of generality, in the results below, we assume that the index κ is fixed.

THEOREM 2.7 (Lower bound) Let $m \leq M, 2 \leq p \leq d, \tau \leq \frac{1}{p-1}, \eta < \frac{1}{d}, T > 0$ be fixed. There exist positive constants $C'_1, \dots, C'_5$, depending only on d, p, m, M , such that for every Ω, h satisfying $h \leq C'_4 T$ and $\Omega h \leq C'_5$ there exists $F = (\mathbf{a}, \mathbf{x}) \in \mathcal{P}_d$, with $\mathbf{x}$ forming a (p, h, T, τ, η) -clustered configuration, and with $0 < m \leq \|\mathbf{a}\| \leq M < \infty$, such that for certain indices $j_1, j_2 \in \{\kappa, \dots, \kappa + p - 1\}$ and every $\epsilon \leq C'_3(\Omega \tau h)^{2p-1}$, it holds that:

$$\begin{aligned} \text{diam}(E_{\epsilon, \Omega}^{\mathbf{x}j}(F)) &\geq \frac{C'_1}{\Omega} \epsilon \times \begin{cases} (\Omega \tau h)^{-2p+2}, & \text{if } j = j_1, \\ 1, & \forall j \notin \{\kappa, \dots, \kappa + p - 1\}; \end{cases} \\ \text{diam}(E_{\epsilon, \Omega}^{\mathbf{a}j}(F)) &\geq C'_2 \epsilon \times \begin{cases} (\Omega \tau h)^{-2p+1}, & \text{if } j = j_2, \\ 1, & \forall j \notin \{\kappa, \dots, \kappa + p - 1\}. \end{cases} \end{aligned}$$

REMARK 2.2 The lower bounds for the quantities $\text{diam}(E_{\epsilon, \Omega}^{\mathbf{x}j}(F))$ were shown in (1) to hold for *any signal F with real amplitudes*; however, at the expense of the implicit dependence of the constants on the separation parameter τ . While bounding $\text{diam}(E_{\epsilon, \Omega}(F))$ (and its projections) for all signals F is an interesting question in its own right, in this paper, we use these to bound the minimax error rate, and therefore, it is sufficient to show that there exist certain signals with large enough $E_{\epsilon, \Omega}(F)$. As it turns out, it is possible to obtain a more accurate geometric description of these sets, which in turn can be used for reducing reconstruction error if additional *a priori* information is available. Work in this direction was started in (2) and we intend to provide further details of these developments in a future work.

Combining Theorems 2.6 and 2.7 with Proposition 2.4, we obtain optimal rates for the minimax error $\mathcal{E}$ and its projections as follows.

THEOREM 2.8 Let $m < M, 2 \leq p \leq d, \tau < \frac{1}{2(p-1)}, \eta < \frac{1}{2d}, T > 0$ be fixed. There exist constants c_1, c_2, c_3 , depending only on d, p, m, M such that for all $\frac{c_1}{\eta T} \leq \Omega \leq \frac{c_2}{h}$ and $\epsilon \leq c_3(\Omega \tau h)^{2p-1}$, the minimax error rates for the set

$$\begin{aligned} U &:= U(p, d, h, \tau, \eta, T, m, M) \\ &= \{(\mathbf{a}, \mathbf{x}) \in \mathcal{P}_d : 0 < m \leq \|\mathbf{a}\| \leq M < \infty, \mathbf{x} \text{ forms a } (p, h, T, \tau, \eta)\text{-clustered configuration}\}, \end{aligned}$$

satisfy the following:

1. For the non-cluster nodes:

$$\forall j \notin \{\kappa, \dots, \kappa + p - 1\} : \begin{cases} \mathcal{E}^{\mathbf{x}j}(\epsilon, U, \Omega) \asymp \frac{\epsilon}{\Omega}, \\ \mathcal{E}^{\mathbf{a}j}(\epsilon, U, \Omega) \asymp \epsilon. \end{cases}$$

2. For the cluster nodes:

$$\begin{aligned} \max_{j=\kappa, \dots, \kappa+p-1} \mathcal{E}^{\mathbf{x}j}(\epsilon, U, \Omega) &\asymp \frac{\epsilon}{\Omega} (\Omega \tau h)^{-2p+2}, \\ \max_{j=\kappa, \dots, \kappa+p-1} \mathcal{E}^{\mathbf{a}j}(\epsilon, U, \Omega) &\asymp \epsilon (\Omega \tau h)^{-2p+1}. \end{aligned}$$

The proportionality constants in the above statements depend only on d, p, m, M .

Proof. Let $C_3, C'_3, C_4, C'_4, C_5, C'_5$ be the constants from Theorems 2.6 and 2.7. Put $c_1 = C_4$ and $c_2 = \min(C_5, C'_5, C_4 C'_4)$. Let $\frac{c_1}{\eta T} \leq \Omega \leq \frac{c_2}{h}$, and $\epsilon \leq c_3(\Omega \tau h)^{2p-1}$, where $c_3 \leq \min(C_3, C'_3)$ will be determined below. It is immediately verified that Ω, h and ϵ as above satisfy the conditions of both Theorems 2.6 and 2.7.

Upper bound Directly follows from the upper bounds in Theorem 2.6 and Proposition 2.4.

Lower bound Denote $U_\epsilon = \left\{ F \in U : E_{\frac{1}{2}\epsilon, \Omega}(F) \subseteq U \right\}$. To prove the lower bounds on $\mathcal{E}$, it is sufficient to show that there exists an $F \in U_\epsilon \neq \emptyset$ such that the conclusions of Theorem 2.7 are satisfied for this F .

It is not difficult to see that for any choice of the parameters as above, the set U has a non-empty interior, and furthermore, that one can choose m', M' satisfying $m < m' < M' < M$, and also $T' = 0.99T$, $\tau' = 2\tau$ and $\eta' = 2\eta$, such that

$$U' = U(p, d, h, \tau', \eta', T', m', M') \subset U, \quad \partial U' \cap \partial U = \emptyset.$$

By construction, there exist positive constants $\tilde{C}_1, \tilde{C}_2$, independent of Ω, h and τ, η , such that

$$\begin{aligned} \inf_{u \in \partial U, u' \in \partial U'} |P_{\mathbf{x}_j}(u) - P_{\mathbf{x}_j}(u')| &\geq \tilde{C}_1 \times \begin{cases} \tau h, & x_j \in \mathbf{x}^c, \\ \eta T, & x_j \in \mathbf{x} \setminus \mathbf{x}^c; \end{cases} \\ \inf_{u \in \partial U, u' \in \partial U'} |P_{\mathbf{a}_j}(u) - P_{\mathbf{a}_j}(u')| &\geq \tilde{C}_2. \end{aligned} \quad (2.5)$$

Now, we use the fact that $\epsilon < c_3(\Omega \tau h)^{2p-1}$. Applying Theorem 2.6 to an arbitrary signal $F' \in U'$ and using the conditions $\frac{1}{\Omega} \leq \frac{\eta T}{c_1}$ and $\Omega \tau h \leq \Omega h \leq c_2$, we obtain that

$$\begin{aligned} \text{diam} \left(E_{\frac{1}{2}\epsilon}^{\mathbf{x}_j}(F') \right) &\leq \begin{cases} \frac{C_1 c_3}{2} \tau h, & x_j \in \mathbf{x}^c, \\ \frac{C_1 c_3}{2\Omega} (\Omega \tau h)^{2p-1} \leq \frac{C_1 c_3}{2c_1} c_2^{2p-1} \eta T, & x_j \in \mathbf{x} \setminus \mathbf{x}^c; \end{cases} \\ \text{diam} \left(E_{\frac{1}{2}\epsilon}^{\mathbf{a}_j}(F') \right) &\leq \begin{cases} \frac{C_2 c_3}{2}, & x_j \in \mathbf{x}^c, \\ \frac{C_2 c_3}{2} c_2^{2p-1}, & x_j \in \mathbf{x} \setminus \mathbf{x}^c. \end{cases} \end{aligned} \quad (2.6)$$

Now we set $c_3 = \min(C_3, C'_3, C''_3)$ where

$$C''_3 = \min(1, c_1) \times \min(1, c_2^{-2p+1}) \times \min\left(\frac{2\tilde{C}_1}{C_1}, \frac{2\tilde{C}_2}{C_2}\right).$$

Combining (2.5) and (2.6), we obtain that $F' \in U_\epsilon$. Since $F' \in U'$ was arbitrary, we conclude that $U' \subseteq U_\epsilon$. Since clearly $U' \neq \emptyset$, applying Proposition 2.4 and Theorem 2.7 finishes the proof. $\square$

3. Numerical optimality of MP algorithm

The main theoretical result of this paper, Theorem 2.8, establishes the best possible scalings for the SR problem with clustered nodes. In this section, we provide some numerical evidence⁵ that a certain SR algorithm, the MP method [36, 37], attains these performance bounds.

Our choice of MP is fairly arbitrary, as we believe that many high-resolution algorithms have similar behavior in the regime $\text{SRF} \gg 1$.

Throughout this section, we replace Ω by N , so that the spectral data is sampled with unit spacing.

3.1 The MP method

ALGORITHM 3.1 The Matrix Pencil algorithm

Input: Model order d

Input: Sequence $\{\tilde{m}_k\}$, $k = 0, 1, \dots, N-1$ where $N > 2d$, of the form (3.1)

Input: pencil parameter $d+1 \leq L \leq N-d$

Output: Estimates for the nodes $\{x_j\}$ and amplitudes $\{a_j\}$ as in (3.1)

1. Compute the matrices $A = \tilde{H}^\uparrow, B = \tilde{H}_\downarrow$;
2. Compute the truncated Singular Value Decomposition (SVD) of A, B of order d :

$$A = U_1 \Sigma_1 V_1^H, \quad B = U_2 \Sigma_2 V_2^H,$$

where U_1, U_2, V_1, V_2 are $L \times d$ and Σ_1, Σ_2 are $d \times d$;

3. Generate the reduced pencil

$$A' = U_2^H U_1 \Sigma_1 V_1^H V_2, \quad B' = \Sigma_2$$

where A', B' are $d \times d$;

4. Compute the generalized eigenvalues $\tilde{z}_j$ of the reduced pencil (A', B') , and put $\{\tilde{x}_j\} = \frac{1}{2\pi} \{\angle \tilde{z}_j\}$, $j = 1, \dots, d$;
5. Compute $\tilde{a}_j$ by solving the linear least squares problem

$$\tilde{\mathbf{a}} = \arg \min_{\mathbf{a} \in \mathbb{C}^d} \|\tilde{\mathbf{m}} - \tilde{V} \mathbf{a}\|_2,$$

where $\tilde{V} = \tilde{V}(\tilde{\mathbf{x}})$ is the Vandermonde matrix $\tilde{V} = \left[\exp \left(2\pi i \tilde{x}_j k \right) \right]_{k=0, \dots, N-1}^{j=1, \dots, d}$;

6. **return** the estimated $\tilde{x}_j$ and $\tilde{a}_j$.
-

⁵ The source code for reproducing the figures in this section is publicly available at <https://github.com/batenkov-group/iai-code>.

ALGORITHM 3.2 The worst-case perturbation signal

Input: Signal $F = (\mathbf{a}, \mathbf{x}) \in \mathcal{P}_d$ with $\mathbf{a} = \mathbf{a}^{(2)}$ and cluster nodes $\mathbf{x}^c = (x_1, \dots, x_p)$

Input: Noise level ϵ

Output: The perturbed signal F_ϵ

1. Compute the cluster center $\mu = \frac{x_1 + x_p}{2}$ and put $\tilde{\mathbf{x}}^c = \mathbf{x}^c - \mu$;
2. Construct the moment vector of the centered cluster: $\mathbf{g} = \left(\sum_{j=1}^p \mathbf{a}_j \tilde{\mathbf{x}}_j^k \right)_{k=0,1,\dots,2p-1} \in \mathbb{R}^{2p}$;
3. Construct the vector $\mathbf{g}'$ to be equal to $\mathbf{g}$ except the last entry: $\mathbf{g}'_k = \mathbf{g}_k$ for $k = 0, 1, \dots, 2p-2$ and $\mathbf{g}'_{2p-1} = \mathbf{g}_{2p-1} + \epsilon$;
4. Solve the Prony problem of order p with the data $\mathbf{g}'$ (for ϵ small enough, a unique solution always exists – see Proposition A.3 and (13)), obtaining a signal $F' = (\mathbf{a}', \mathbf{x}') \in \mathcal{P}_p$;
5. Move the cluster nodes back and put

$$F_\epsilon(x) = \sum_{j=p+1}^d \mathbf{a}_j \delta(x - \mathbf{x}_j) + \sum_{j=1}^p \mathbf{a}'_j \delta(x - (\mathbf{x}'_j + \mu));$$

return the signal F_ϵ .

Let $F = (\mathbf{a}, \mathbf{x}) \in \mathcal{P}_d$ as in (1.1) with $x_j \in \left[-\frac{1}{2}, \frac{1}{2}\right]$. Given the noisy Fourier measurements

$$\tilde{m}_k = \underbrace{\mathcal{F}(F)(-k)}_{=m_k} + n_k \quad (3.1)$$

$$= \sum_{j=1}^d a_j \exp(2\pi i x_j k) + n_k, \quad k = 0, 1, \dots, N-1, \quad N > 2d, \quad (3.2)$$

the MP method estimates $\tilde{F} = (\tilde{\mathbf{a}}, \tilde{\mathbf{x}})$ as follows. Consider the Hankel matrix

$$H = \begin{bmatrix} m_0 & m_1 & \dots & m_{N-L-1} \\ m_1 & m_2 & \dots & m_{N-L} \\ \vdots & \ddots & \ddots & \vdots \\ m_L & m_{L+1} & \dots & m_{N-1} \end{bmatrix} \in \mathbb{C}^{(L+1) \times (N-L)}, \quad (3.3)$$

and further let $H^\uparrow = H[0 : L-1, :]$ and $H_\downarrow = H[1 : L, :]$ be the $L \times (N-L)$ matrix obtained from H by deleting the last (respectively, the first) row. Then, it turns out that the numbers $z_j = \exp(2\pi i x_j)$ are the d non-zero generalized eigenvalues (i.e. rank-reducing numbers) of the pencil $H_\downarrow - zH^\uparrow$. If we now construct the noisy matrices $A = \tilde{H}^\uparrow, B = \tilde{H}_\downarrow$ from the available data $\{\tilde{m}_k\}_{k=0,\dots,N-1}$, we could apparently just solve the generalized eigenvalue problem with A, B . However, if $L > d$, then the pencil

Algorithm 3.3: A single experiment**Input** : p, d, h, N, ϵ **Input** : Testing scheme (either **S1** or **S2**)

- 1 Construct the signal F and the sequence $\tilde{m}_k, k = 0, \dots, N-1$ according to Subsection 3.2;
- 2 Compute the actual perturbation magnitude

$$\epsilon_0 = \max_{k=0, \dots, N-1} |\mathcal{F}(F)(-k) - \tilde{m}_k|;$$

- 3 Execute the MP method (Algorithm 3.1) with $L = \lceil \frac{N}{2} \rceil$ and obtain $F_{MP} = (\mathbf{a}^{MP}, \mathbf{x}^{MP})$;

4 **for each** j **do**

- 5 compute the error for node j :

$$e_j = \min_{\ell} |\mathbf{x}_j^{MP} - \mathbf{x}_{\ell}|;$$

- 6 The success for node j is defined as

$$Succ_j = \left(e_j < \frac{\min_{\ell \neq j} |\mathbf{x}_{\ell} - \mathbf{x}_j|}{3} \right).$$

- 7 **if** $Succ_j == \text{true}$ **then**

- 8 let $\ell(j) = \arg \min_{\ell} |\mathbf{x}_j^{MP} - \mathbf{x}_{\ell}|$;

- 9 compute normalized node error amplification factor

$$\mathcal{K}_{\mathbf{x},j} = \frac{|\mathbf{x}_j - \mathbf{x}_{\ell(j)}^{MP}| \cdot N}{\epsilon_0};$$

- 10 compute normalized amplitude error amplification factor

$$\mathcal{K}_{\mathbf{a},j} = \frac{|\mathbf{a}_j - \mathbf{a}_{\ell(j)}^{MP}|}{\epsilon_0};$$

- 11 **return** ϵ_0 , and $(\mathcal{K}_{\mathbf{x},j}, \mathcal{K}_{\mathbf{a},j}, Succ_j)$ for each node $j = 1, \dots, d$.

$B - zA$ is close to being singular, and so an additional step of low-rank approximation is required. We summarize the MP method in Algorithm 1, and the interested reader is referred to the widely available literature on the subject (e.g. [36, 37, 45, 54], and the references therein) for further details. Note that there exist numerous variants of MP, but, again, we believe the particular details to be immaterial for our discussion.

3.2 Experimental setup

3.2.1 Clustered node configurations In our experiments presented below, we constructed (p, h, T, τ, η) -clustered configurations with

$$\tau = \frac{1}{p-1}, T = \pi, \eta = \frac{\pi - h}{\pi(d-p+1)}$$

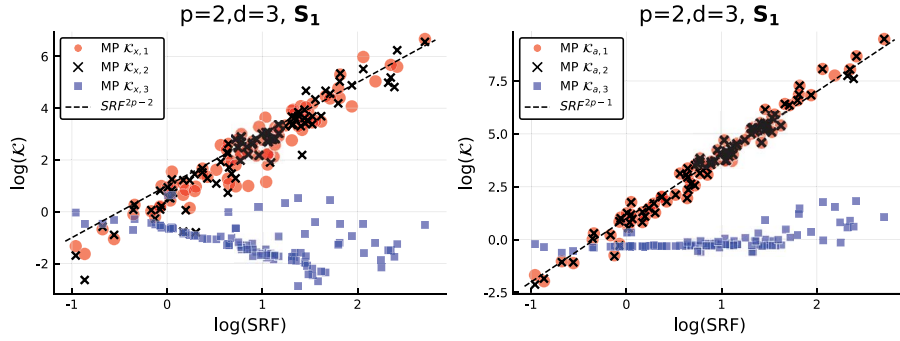


FIG. 3. The error amplification factors. Algorithm 3 was executed 100 times with $p = 2, d = 3$, scheme **S1** and varying h, N, ϵ . For cluster nodes $j = 1, 2$, the node error amplification factors $\mathcal{K}_{x,j}$ (left panel) scale like SRF^{2p-2} , while the amplitude error amplification factors $\mathcal{K}_{a,j}$ (right panel) scale like SRF^{2p-1} . For the non-cluster node $j = 3$, both error amplification factors are bounded by a constant.

as follows:

1. The cluster nodes $\mathbf{x}^c = (x_1, \dots, x_p)$ where $x_j = (j-1) \cdot \Delta$ and $\Delta = \frac{h}{p-1}$ for $j = 1, \dots, p$.
2. The non-cluster nodes were chosen to be

$$x_{p+j} = (p-1)\Delta + j \cdot \frac{\pi - (p-1)\Delta}{d-p+1}, \quad j = 1, \dots, d-p.$$

3.2.2 Choice of signal and perturbation Two different schemes were tested:

- S1** A generic signal with complex amplitude vector $\mathbf{a}^{(1)} = (t^0, t^1, t^2, \dots) \in \mathbb{C}^d$ and a bounded random perturbation sequence $\{n_k\}$, uniformly distributed in $[-\epsilon, \epsilon]$.
- S2** Worst-case scenario in accordance with the construction of Section 6 (and in particular of Theorem 6.2): a real amplitude vector $\mathbf{a}^{(2)} = (1, -1, 1, \dots) \in \mathbb{R}^d$ and the perturbed Fourier coefficient sequence $\{\tilde{m}_k\}$ of the particular signal $F_\epsilon = (\mathbf{a}', \mathbf{x}') \in \mathcal{P}_d$ constructed according to Algorithm 2:

$$\tilde{m}_k = \mathcal{F}(F_\epsilon)(-k) = \sum_{j=1}^d \mathbf{a}'_j \exp(2\pi i \mathbf{x}'_j k), \quad k = 0, \dots, N-1.$$

3.3 Results

3.3.1 Error amplification factors In the first set of experiments, we measured the actual error amplification factors $\mathcal{K}_{x,j}, \mathcal{K}_{a,j}$ as in Algorithm 3 (recall also (1.4)), choosing ϵ, N, h randomly from a pre-defined numerical range. The results are presented in Figs 3 and 4 for the testing schemes **S1** and **S2**, accordingly. The scalings of Theorem 2.8, in particular the dependence on SRF, are confirmed.

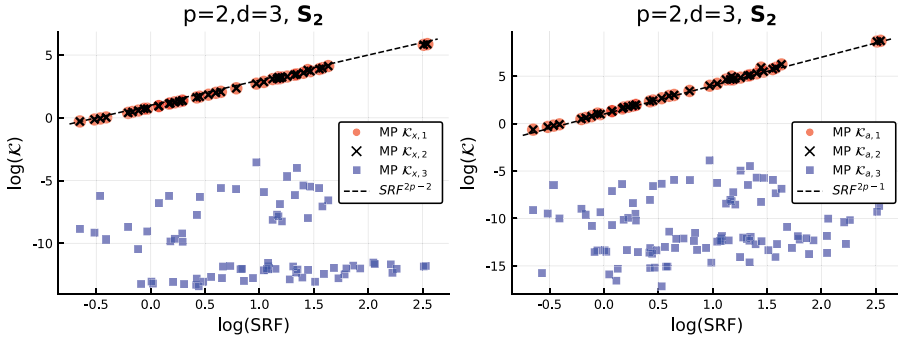


FIG. 4. Same setup as in Fig. 3, scheme **S2**. Comparing with Fig. 3, the variance of the factors corresponding to the cluster nodes is much smaller than for the case of random perturbations, indicating that the construction is indeed worst-case.

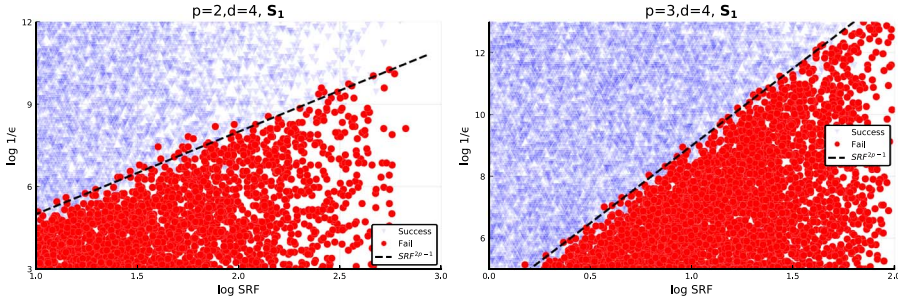


FIG. 5. Phase transition for successful recovery, random bounded perturbations (scheme **S1**) with $d = 4$ and $p = 2, 3$. Each experiment is represented by either a blue triangle (if the recovery was successful, i.e. $\text{Succ}_j = \text{True}$, $\forall j = 1, \dots, d$ as returned by Algorithm 3) or a red circle otherwise. The relationship $\epsilon_{\text{crit}} \approx \text{SRF}^{1-2p}$ for the critical value of ϵ is confirmed.

3.3.2 Noise threshold for successful recovery In the second set of experiments, we investigated the noise threshold $\epsilon \lesssim \text{SRF}^{1-2p}$ for successful recovery, as predicted by the theory. We have performed 15000 random experiments with scheme **S1** (the randomness was in the choice of h, N, ϵ and the noise sequence $\{n_k\}$) according to Algorithm 3, recording the success/failure result of each such experiment. The results for $d = 4$ and $p = 2, 3$ are presented in Fig. 5, and the theoretical scaling above is confirmed for the MP method.

Although not covered by our current theory, it is of interest to establish the recovery threshold for every node separately. In Fig. 6, we can see that for a non-cluster node, the threshold is approximately constant (i.e. does not depend on the SRF) – even though Theorem 2.6 requires $\epsilon \lesssim \text{SRF}^{1-2p}$.

4. Normalization

In the intermediate claims, instead of considering a general signal $F = (\mathbf{a}, \mathbf{x}) \in \mathcal{P}_d$, we shall usually assume that the node vector $\mathbf{x} = (x_1, \dots, x_d)$ is normalized to the interval $[-\frac{1}{2}, \frac{1}{2}]$ and centered around the origin, i.e. $x_d = -x_1$. Let us briefly argue how to obtain the general result from this special case.

Let us define the scale and shift transformations on $\mathcal{P}_d$.

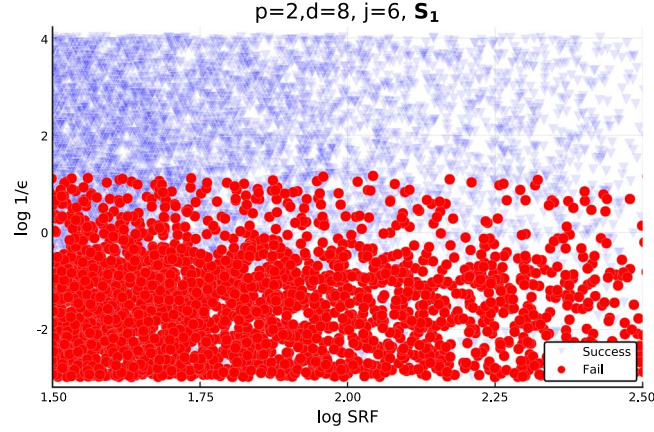


FIG. 6. Phase transition for successful recovery of a non-cluster node. Comparing with Fig. 5, the threshold is approximately constant $\epsilon_{crit} \approx \text{const}$. Here, $p = 2, d = 8$, scheme **S1**, plotted is the successful recovery of the node at index $j = 6$.

DEFINITION 4.1 For $F = \sum_{j=1}^d a_j \delta(x - x_j) \in \mathcal{P}_d$ and $\alpha \in \mathbb{R}$, we define $SH_\alpha : \mathcal{P}_d \rightarrow \mathcal{P}_d$ as follows:

$$SH_\alpha(F)(x) = \sum_{j=1}^d a_j \delta(x - (x_j - \alpha)).$$

DEFINITION 4.2 For $F = \sum_{j=1}^d a_j \delta(x - x_j) \in \mathcal{P}_d$ and $T > 0$, we define $SC_T : \mathcal{P}_d \rightarrow \mathcal{P}_d$ as follows:

$$SC_T(F)(x) = \sum_{j=1}^d a_j \delta\left(x - \frac{x_j}{T}\right).$$

By the shift property of the Fourier transform, for any $\epsilon, \Omega > 0$, we have that

$$SH_\alpha(E_{\epsilon, \Omega}(F)) = E_{\epsilon, \Omega}(SH_\alpha(F)). \quad (4.1)$$

By the scale property of the Fourier transform, we have that for any $\epsilon > 0$,

$$SC_T(E_{\epsilon, \Omega}(F)) = E_{\epsilon, \Omega T}(SC_T(F)). \quad (4.2)$$

Thus, we have the following.

PROPOSITION 4.3 Let $F = (\mathbf{a}, \mathbf{x}) \in \mathcal{P}_d$, $\alpha \in \mathbb{R}$ and $T > 0$. Then, for any $\epsilon > 0$ and $1 \leq j \leq d$ we have

$$\text{diam}(E_{\epsilon, \Omega}^{\mathbf{x}j}(F)) = T \text{diam}\left(E_{\epsilon, \Omega T}^{\mathbf{x}j}(SC_T(SH_\alpha(F)))\right) \quad (4.3)$$

$$\text{diam}(E_{\epsilon, \Omega}^{\mathbf{a}j}(F)) = \text{diam}\left(E_{\epsilon, \Omega T}^{\mathbf{a}j}(SC_T(SH_\alpha(F)))\right) \quad (4.4)$$

5. Upper bounds

5.1 Overview of the proof

The proof of Theorem 2.6, presented in the next subsections and some of the appendices, is somewhat technical. In order to help the reader, we provide an overview of the essential ideas and steps.

The main object of the study, the error set $E_{\epsilon, \Omega}(F) \subset \mathcal{P}_d$, is the pre-image of an (infinite dimensional) ϵ -cube in the data space, under the Fourier transform mapping $\mathcal{F}$ (recall (1.2) and Definition 2.3). However, it is not obvious how to obtain quantitative estimates on $\mathcal{F}^{-1}$ directly. Thus, we replace $\mathcal{F}$ with certain finite-dimensional sampled versions of it, denoted $FM_\lambda : \mathcal{P}_d \rightarrow \mathbb{C}^{2d}$, where the sampling parameter λ defines the rate at which $2d$ equispaced samples of $\mathcal{F}(F)$ are taken. The pre-images of ϵ -cubes under FM_λ define the corresponding λ -error sets $E_{\epsilon, (\lambda)} \subset \mathcal{P}_d$, and in fact, the original $E_{\epsilon, \Omega}(F)$ is contained in the intersection of all the $E_{\epsilon, (\lambda)}$. Thus, it is sufficient to bound the diameter of a single such $E_{\epsilon, (\lambda^*)}$ (see remark in the next paragraph) with a carefully chosen λ^* so that the result will be as small as possible. Such quantitative estimates are obtained by careful analysis of the row-wise norms of the Jacobian matrix of $FM_{\lambda^*}^{-1}$ and applying the so-called *quantitative inverse function theorem* (Theorem B.1). Using these estimates, the optimal λ^* is shown to be on the order of Ω , from which the upper bounds of Theorem 2.6 follow.

An additional technical complication arises from the fact that FM_λ^{-1} defines a multivalued mapping, and the full pre-image $E_{\epsilon, (\lambda)}$ contains multiple copies of a certain ‘basic’ set $A = A_{\epsilon, \lambda}$. However, when considering the intersection of all $E_{\epsilon, (\lambda)}$ s, the non-zero shifts for certain different λ s do not intersect, and therefore, eventually only the diameter of the basic set A needs to be estimated.

Below is a brief description of the different intermediate results and the organization of the remainder of Section 5.

1. In Subsection 5.2, we formally define the λ -decimated maps FM_λ , the corresponding error sets $E_{\epsilon, (\lambda)}$, and provide quantitative estimates on the Jacobian of FM_λ^{-1} in Proposition 5.3 (proved in Appendix C). These bounds essentially depend on the ‘effective separation’ of each node in $\mathbf{x}$ from its neighbors, after a blowup by a factor of λ .
2. In Subsection 5.3, we show that for a signal $F = (\mathbf{a}, \mathbf{x})$, there exists a certain range of admissible λ s, denoted by $\Lambda(\mathbf{x})$, for which the effective separation (see previous item) between the nodes in $\mathbf{x}^c$ is of the order of Ωh , while for the rest of the nodes, it is bounded from below by a constant independent of Ω, h . These estimates are proved in Proposition 5.8.
3. In Subsection 5.4, we study in detail the geometry of the error sets $E_{\epsilon, (\lambda)}$ for $\lambda \in \Lambda(\mathbf{x})$. First, we consider (in Subsection 5.4.1) the local inverses FM_λ^{-1} . For each $\lambda \in \Lambda(\mathbf{x})$, we show that the local inverse exists in a neighborhood V of radius $R \approx (\Omega h)^{2p-1}$ around $FM_\lambda(F)$ and provide estimates on the Lipschitz constants of FM_λ^{-1} on V and the diameter of $FM_\lambda^{-1}(V)$. The main bounds to that effect are proved in Proposition 5.14, using the previously established general estimates from Proposition 5.3 and the quantitative inverse function theorem (Theorem B.1).
4. Next, denoting $A = A_{R, \lambda} = FM_\lambda^{-1}(V)$, we show in Proposition 5.16 that the set $E_{\epsilon, (\lambda)}$ is a union of certain copies of A , where each such copy is obtained by shifting the nodes in A by an integer multiple of λ^{-1} and/or by permuting them.
5. In Subsection 5.5, we complete the proof. At this point, we consider the entire set $\Lambda(\mathbf{x})$. The main technical step, Proposition 5.17 (proved in Appendix F), establishes that for a certain $\lambda^* \in \Lambda(\mathbf{x})$ and all possible permutations π and shifts $\ell \in \mathbb{Z} \setminus \{0\}$, there exists a particular $\bar{\lambda} = \bar{\lambda}(\pi, \ell) \in \Lambda(\mathbf{x})$

such that the intersection between π -permutation and ℓ -shift of A_{R,λ^*} and the *entire error set* $E_{R,(\tilde{\lambda})}$ is empty. From this fact, it immediately follows that the original error set $E_{\epsilon,\Omega}(F)$ with $\epsilon = R$ is contained in A_{R,λ^*} (Proposition 5.18). The proof is finished by invoking the previously established estimates on the diameter of A_{R,λ^*} and its projections.

REMARK 5.1 We expect that the tools developed throughout the proof will also be useful to calculate the minimal finite sampling rate required to achieve the minimax error rate stated in Theorem 2.6.

5.2 λ -decimation maps

For the purpose of the following analysis, we extend the space of signals $\mathcal{P}_d$ to include signals with complex nodes and denote the extended space by $\bar{\mathcal{P}}_d$,

$$\bar{\mathcal{P}}_d = \left\{ (\mathbf{a}, \mathbf{x}) : \mathbf{a} = (a_1, \dots, a_d) \in \mathbb{C}^d, \mathbf{x} = (x_1, \dots, x_d) \in \mathbb{C}^d \right\}.$$

We will be considering specific sets of exactly $2d$ samples of the Fourier transform, made at constant rate λ as follows.

DEFINITION 5.1 For $\lambda > 0$, we define the map $FM_\lambda : \bar{\mathcal{P}}_d \cong \mathbb{C}^{2d} \rightarrow \mathbb{C}^{2d}$ by

$$FM_\lambda((\mathbf{a}, \mathbf{x})) = \mu = (\mu_0, \dots, \mu_{2d-1}), \quad \mu_k = \sum_{j=1}^d a_j e^{2\pi i x_j \lambda k}, \quad k = 0, \dots, 2d-1.$$

We call such map a λ -decimation map.

For $\lambda > 0$ and $\epsilon > 0$, we define the corresponding error set $E_{\epsilon,(\lambda)}$ as follows.

DEFINITION 5.2 The error set $E_{\epsilon,(\lambda)}(F) \subset \mathcal{P}_d$ is the set consisting of all the signals $F' \in \mathcal{P}_d$ with

$$\|FM_\lambda(F') - FM_\lambda(F)\| \leq \epsilon.$$

Similarly, we denote by $E_{\epsilon,(\lambda)}^{\mathbf{a}j}(F)$, $E_{\epsilon,(\lambda)}^{\mathbf{x}j}(F)$ the projection of the error set $E_{\epsilon,(\lambda)}(F)$ onto the corresponding amplitudes and the nodes components (compare (2.1)).

Now, consider the given spectrum $\mathcal{F}(F)(s)$, $s \in [-\Omega, \Omega]$. Clearly for each $\lambda \leq \frac{\Omega}{2d-1}$, we have that $E_{\epsilon,\Omega}(F) \subseteq E_{\epsilon,(\lambda)}(F)$ giving

$$E_{\epsilon,\Omega}(F) \subseteq \bigcap_{\lambda \in (0, \frac{\Omega}{2d-1}]} E_{\epsilon,(\lambda)}(F). \quad (5.1)$$

Hence, to prove the upper bounds in Theorem 2.6, we shall show that there exists a *certain* subset $S \subseteq \left(0, \frac{\Omega}{2d-1}\right]$ such that for each $\lambda \in S$, $\text{diam}(E_{\epsilon,(\lambda)}(F))$ can be effectively controlled.

In the next proposition, we derive a uniform bound on the norms of the inverse Jacobian of FM_λ near a signal with clustered nodes. The bounds explicitly depend on the distances between the so-called mapped nodes $z_j(\lambda) = e^{2\pi i \lambda x_j}$.

PROPOSITION 5.3 (Uniform Jacobian bounds). Let $F = (\mathbf{a}, \mathbf{x}) \in \tilde{\mathcal{P}}_d$, $\mathbf{a} = (a_1, \dots, a_d)$, $\mathbf{x} = (x_1, \dots, x_d)$ and for $\lambda > 0$ let $z_1 = e^{2\pi i \lambda x_1}, \dots, z_d = e^{2\pi i \lambda x_d}$. Suppose that for each $j = 1, \dots, d$, we have $0 < \frac{m}{2} \leq |a_j|$ and $\frac{1}{2} \leq |z_j| \leq 2$ for some $m > 0$.

Further assume that for $\tilde{\eta}, \tilde{h}$ with $1 \geq \tilde{\eta} \geq \tilde{h}$, and $\mathbf{x}^c = \{x_{\kappa}, \dots, x_{\kappa+p-1}\} \subset \mathbf{x}$, $p \geq 2$, the nodes $z_1, \dots, z_d$ satisfy:

1. For each $x_j, x_k \in \mathbf{x}^c$, $j \neq k$, we have that $|z_j - z_k| \geq \tilde{h}$.
2. For each $x_\ell \in \mathbf{x} \setminus \mathbf{x}^c$ and $x_j \in \mathbf{x}$, $\ell \neq j$, we have that $|z_\ell - z_j| \geq \tilde{\eta}$.

Then the Jacobian matrix of FM_λ at F , denoted by $J_\lambda(F)$, is non-degenerate. Furthermore, write the inverse Jacobian matrix $J_\lambda^{-1}(F)$ in the following block form $J_\lambda^{-1}(F) = \begin{bmatrix} A \\ \tilde{B} \end{bmatrix}$, where $A, \tilde{B}$ are $d \times 2d$. Then, the ℓ_1 norms of the rows of the blocks $A, \tilde{B}$ are bounded as follows:

$$\sum_{k=1}^{2d} |A_{j,k}| \leq K_1(\tilde{\eta}, d, p), \quad x_j \in \mathbf{x} \setminus \mathbf{x}^c, \quad (5.2)$$

$$\sum_{k=1}^{2d} |\tilde{B}_{j,k}| \leq K_2(m, \tilde{\eta}, d, p) \frac{1}{\lambda}, \quad x_j \in \mathbf{x} \setminus \mathbf{x}^c, \quad (5.3)$$

$$\sum_{k=1}^{2d} |A_{j,k}| \leq K_3(\tilde{\eta}, d, p) \tilde{h}^{-2p+1}, \quad x_j \in \mathbf{x}^c, \quad (5.4)$$

$$\sum_{k=1}^{2d} |\tilde{B}_{j,k}| \leq K_4(m, \tilde{\eta}, d, p) \frac{1}{\lambda} \tilde{h}^{-2p+2}, \quad x_j \in \mathbf{x}^c, \quad (5.5)$$

where $K_1(\cdot, \dots, \cdot), K_2(\cdot, \dots, \cdot), K_3(\cdot, \dots, \cdot), K_4(\cdot, \dots, \cdot)$ are constants depending only on the parameters inside the brackets.

The proof of Proposition 5.3 is given in Appendix C.

5.3 The existence of an admissible decimation

In this section, we shall prove the existence of a certain blowup factors λ , such that the mapped nodes $\{e^{2\pi i \lambda x_j}\}$ (see Proposition 5.3 above) attain ‘good’ separation properties. This result will later be used to show that for any such λ , the corresponding inverse λ -decimation map FM_λ^{-1} will have the smallest possible coordinatewise Lipschitz constants with respect to Ω, h (up to constants) (see Proposition 5.3).

DEFINITION 5.4 For each $x \in \mathbb{R}$ and $a > 0$, consider the operation $\bmod \left(-\frac{a}{2}, \frac{a}{2}\right]$ defined as

$$x \bmod \left(-\frac{a}{2}, \frac{a}{2}\right] = x - ka,$$

where k is the unique integer such that $x - ka \in \left(-\frac{a}{2}, \frac{a}{2}\right]$. Using this notation, the principal value of the complex argument function is defined as

$$\text{Arg}(re^{i\theta}) = \theta \bmod (-\pi, \pi],$$

for each $\theta \in \mathbb{R}$ and $r > 0$.

DEFINITION 5.5 For $\alpha, \beta \in \mathbb{C} \setminus \{0\}$, we define the angular distance between α, β as

$$\angle(\alpha, \beta) = \left| \operatorname{Arg} \left(\frac{\alpha}{\beta} \right) \right| = \left| (\operatorname{Arg}(\alpha) - \operatorname{Arg}(\beta)) \bmod (-\pi, \pi] \right|,$$

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

LEMMA 5.6 For $|x| = |y| = 1$, we have

$$\frac{2}{\pi} \angle(x, y) \leq |x - y| \leq \angle(x, y). \quad (5.6)$$

Proof. First,

$$|x - y| = \left| 1 - \frac{x}{y} \right| = 2 \sin \left| \frac{1}{2} \operatorname{Arg} \frac{x}{y} \right| = 2 \sin \left| \frac{\angle(x, y)}{2} \right|.$$

Then, use the fact that for any $|\theta| \leq \frac{\pi}{2}$ we have

$$\frac{2}{\pi} |\theta| \leq \sin |\theta| \leq |\theta|.$$

□

Let $F = (\mathbf{a}, \mathbf{x}) \in \mathcal{P}_d$ such that the node vector $\mathbf{x} = (x_1, \dots, x_d)$ forms a (p, h, T, τ, η) -clustered configuration, with $\mathbf{x}^c = \{x_\kappa, x_{\kappa+1}, \dots, x_{\kappa+p-1}\}$. According to Proposition 5.3, the norms of the rows of the inverse Jacobian $J_\lambda^{-1}(F)$ essentially depend on the minimal distance between the mapped nodes $z_j(\lambda) = e^{2\pi i \lambda x_j}$. After a blowup by a factor of $\lambda \leq \frac{1}{2h}$, the pairwise angular distances $\angle(\cdot, \cdot)$ (and hence the euclidean distances) between the mapped cluster nodes $z_\kappa, \dots, z_{\kappa+p-1}$ are now of order λh .

On the other hand, the non-cluster nodes are at distance larger than $\eta T \gg h$. Therefore, after the blowup by λ , the non-cluster nodes $z_1, \dots, z_{\kappa-1}, z_{\kappa+p}, \dots, z_d$ may in principle be located anywhere on the unit circle. For example, any of these mapped non-cluster nodes might coincide with, or be very close to, a certain mapped cluster node, or yet another mapped non-cluster node.

While this situation might occur for some values of λ , we will now show that there exist certain sets of λ s for which this does not happen. We shall require the following key estimate concerning the pairwise angular distance between any two mapped nodes.

LEMMA 5.7 (A uniform blowup of two nodes). Let $x_j, x_k \in \mathbb{R}$, $x_j \neq x_k$, and let $\Delta = |x_j - x_k|$. Consider the following blowups $z_j = z_j(\lambda) = e^{2\pi i \lambda x_j}$, $z_k = z_k(\lambda) = e^{2\pi i \lambda x_k}$. Then for $0 \leq \alpha \leq \pi$ and an interval $I = [a, b] \subset \mathbb{R}$, the set

$$\Sigma_{j,k}^\alpha(I) = \left\{ \lambda \in I : \angle(z_j(\lambda), z_k(\lambda)) \leq \alpha \right\} \quad (5.7)$$

is a union of N intervals $I_1, \dots, I_N$ with $\lfloor |I|\Delta \rfloor \leq N \leq \lfloor |I|\Delta \rfloor + 1$, and

$$|I_j| \leq \frac{\alpha}{\pi} \frac{1}{\Delta}, \quad j = 1, \dots, N.$$

Proof. For each $\lambda \in I$ we have

$$\angle(z_j(\lambda), z_k(\lambda)) = \left| \operatorname{Arg} \left(\frac{z_j(\lambda)}{z_k(\lambda)} \right) \right| = \left| \operatorname{Arg}(e^{2\pi i \lambda \Delta}) \right|. \quad (5.8)$$

By equation (5.8) we have

$$\begin{aligned} & \left\{ \lambda \in I : \angle(z_j(\lambda), z_k(\lambda)) \leq \alpha \right\} = \\ & \left\{ \lambda \in I : \left| \operatorname{Arg}(e^{2\pi i \lambda \Delta}) \right| \leq \alpha \right\} = \\ & \left\{ \lambda \in I : |2\pi \lambda \Delta \bmod (-\pi, \pi)| \leq \alpha \right\} = \\ & \left\{ \lambda \in I : -\alpha \leq (2\pi \lambda \Delta \bmod (-\pi, \pi)) \leq \alpha \right\} = \\ & \left\{ \lambda \in I : -\frac{\alpha}{2\pi} \frac{1}{\Delta} \leq \left(\lambda \bmod \left(-\frac{1}{2\Delta}, \frac{1}{2\Delta} \right) \right) \leq \frac{\alpha}{2\pi} \frac{1}{\Delta} \right\}. \end{aligned}$$

The last set above can be written as $I \cap S^\alpha$ where

$$S^\alpha = \left\{ \lambda \in \mathbb{R} : -\frac{\alpha}{2\pi} \frac{1}{\Delta} \leq \left(\lambda \bmod \left(-\frac{1}{2\Delta}, \frac{1}{2\Delta} \right) \right) \leq \frac{\alpha}{2\pi} \frac{1}{\Delta} \right\}. \quad (5.9)$$

Define the interval $I^\alpha = \left[-\frac{\alpha}{2\pi} \frac{1}{\Delta}, \frac{\alpha}{2\pi} \frac{1}{\Delta} \right]$. Then the set S^α is a union of intervals of length $\frac{\alpha}{\pi} \frac{1}{\Delta}$ as follows

$$S^\alpha = \bigcup_{\ell \in \mathbb{Z}} \left(I^\alpha + \frac{\ell}{\Delta} \right) = \bigcup_{\ell \in \mathbb{Z}} \left\{ \lambda + \frac{\ell}{\Delta} : \lambda \in I^\alpha \right\}.$$

The intersection of S^α with any interval I is then a union of $\lfloor |I|\Delta \rfloor \leq N \leq \lfloor |I|\Delta \rfloor + 1$ intervals of length smaller or equal to $\frac{\alpha}{\pi} \frac{1}{\Delta}$. This concludes the proof of Lemma 5.7. $\square$

Now, we state and prove the main result of this subsection.

PROPOSITION 5.8 Let $F = (\mathbf{a}, \mathbf{x}) \in \mathcal{P}_d$, $\mathbf{x} = (x_1, \dots, x_d) \subset [-\frac{1}{2}, \frac{1}{2}]$, such that $\mathbf{x}$ forms a $(p, h, 1, \tau, \eta)$ -clustered configuration with $\mathbf{x}^c = \{x_\kappa, x_{\kappa+1}, \dots, x_{\kappa+p-1}\}$.

Let $\Omega \leq \frac{2d-1}{2} \cdot \frac{1}{h}$. For each $\lambda > 0$ let $z_1(\lambda) = e^{2\pi i \lambda x_1}, \dots, z_d(\lambda) = e^{2\pi i \lambda x_d}$.

Then, each interval $I \subset \left[\frac{1}{2} \frac{\Omega}{2d-1}, \frac{\Omega}{2d-1} \right]$ of length $|I| = \frac{1}{\eta}$ contains a sub-interval $I' \subset I$ of length $|I'| \geq (2d^2\eta)^{-1}$ such that for each $\lambda \in I'$:

1. For all $x_\ell \in \mathbf{x} \setminus \mathbf{x}^c$ and $x_j \in \mathbf{x}$, $x_j \neq x_\ell$,

$$\angle(z_\ell(\lambda), z_j(\lambda)) \geq \frac{1}{d^2}. \quad (5.10)$$

2. For all $x_j, x_k \in \mathbf{x}^c, x_k \neq x_j$,

$$\angle(z_j(\lambda), z_k(\lambda)) \geq 2\pi\lambda\tau h \geq \frac{\pi\tau}{2d-1}\Omega h. \quad (5.11)$$

Proof. Let us first prove that assertion (5.11) holds for any $\frac{1}{2}\frac{\Omega}{2d-1} \leq \lambda \leq \frac{\Omega}{2d-1}$.

Let $x_j, x_k, j > k$ be two cluster nodes. The angular distance between the mapped cluster nodes $z_j = z_j(\lambda) = e^{2\pi i\lambda x_j}, z_k = z_k(\lambda) = e^{2\pi i\lambda x_k}$, is

$$\angle(z_j, z_k) = \left| \text{Arg}(e^{2\pi i\lambda(x_j - x_k)}) \right|.$$

By assumption $\Omega h \leq \frac{2d-1}{2}$, then $\lambda \leq \frac{1}{2h}$ and then $0 \leq 2\pi\lambda(x_j - x_k) \leq 2\pi\lambda h \leq \pi$. With this, we have

$$\angle(z_j, z_k) = 2\pi\lambda(x_j - x_k) \geq 2\pi\lambda\tau h.$$

By assumption $\lambda \geq \frac{1}{2}\frac{\Omega}{2d-1}$. Then, $\angle(z_j, z_k) \geq \frac{\pi\tau}{2d-1}\Omega h$. This concludes the proof of assertion (5.11).

Using Lemma 5.7, we now prove that assertion (5.10) holds for any interval $I = [a, b] \subset \mathbb{R}$ of length $|I| = \frac{1}{\eta}$. Let I be such an interval. For each $0 < \alpha \leq \pi$, consider the set

$$\Sigma^\alpha(I) = \left\{ \lambda \in I : \exists x_\ell \in \mathbf{x} \setminus \mathbf{x}^c \text{ s.t. } \min_{1 \leq j \leq d, j \neq \ell} \angle(z_\ell(\lambda), z_j(\lambda)) \leq \alpha \right\}.$$

We then have

$$\Sigma^\alpha(I) = \bigcup_{x_\ell \in \mathbf{x} \setminus \mathbf{x}^c} \bigcup_{x_j \neq x_\ell} \Sigma_{\ell,j}^\alpha(I),$$

where $\Sigma_{\ell,j}^\alpha$ are given by (5.7). By Lemma 5.7, each $\Sigma_{\ell,j}^\alpha(I)$ above is a union of at most $\lfloor |I|\eta \rfloor + 1 = 2$ intervals, the length of each interval is at most $\frac{\alpha}{\pi} \frac{1}{\eta}$. Therefore, $\Sigma^\alpha(I)$ is a union of at most $K = \binom{d}{2} 2 = d(d-1)$ intervals. Moreover, let ν denote the Lebesgue measure on $\mathbb{R}$, then

$$\nu(\Sigma^\alpha(I)) \leq K \frac{\alpha}{\pi} \frac{1}{\eta} \leq d(d-1) \frac{\alpha}{\pi} \frac{1}{\eta} \leq d^2 \alpha \frac{1}{2\eta}. \quad (5.12)$$

Put $\alpha' = \frac{1}{d^2}$ then by (5.12)

$$\nu(\Sigma^{\alpha'}(I)) \leq \frac{1}{2\eta}. \quad (5.13)$$

Now, consider the complement set of $\Sigma^{\alpha'}(I)$ with respect to I ,

$$(\Sigma^{\alpha'}(I))^c = \left\{ \lambda \in I : \forall x_\ell \in \mathbf{x} \setminus \mathbf{x}^c, \min_{1 \leq j \leq d, j \neq \ell} \angle(z_\ell(\lambda), z_j(\lambda)) > \frac{1}{d^2} \right\}.$$

By (5.13)

$$v((\Sigma^{\alpha'}(I))^c) \geq |I| - \frac{1}{2\eta} = \frac{1}{\eta} - \frac{1}{2\eta} = \frac{1}{2\eta}. \quad (5.14)$$

In addition, since $\Sigma^{\alpha'}(I)$ is a union of at most $K = d(d-1)$ intervals, then $(\Sigma^{\alpha'}(I))^c$ is a union of at most

$$L = K + 1 = d(d-1) + 1 \leq d^2 \quad (5.15)$$

intervals. Using (5.14) and (5.15), the average size of these intervals is bounded as follows:

$$\frac{v((\Sigma^{\alpha'}(I))^c)}{L} \geq \frac{1}{d^2} \frac{1}{2\eta}.$$

We therefore conclude that $(\Sigma^{\alpha'}(I))^c$ contains an interval of length greater than or equal to $\frac{1}{d^2} \frac{1}{2\eta}$. This proves assertion (5.10) of Proposition 5.8. $\square$

5.4 Error sets of admissible decimation maps

Throughout this section, we fix a signal $F = (\mathbf{a}, \mathbf{x}) \in \mathcal{P}_d$, $\mathbf{a} = (a_1, \dots, a_d)$, $\mathbf{x} = (x_1, \dots, x_d) \subset [-\frac{1}{2}, \frac{1}{2}]$, such that $\mathbf{x}$ forms a $(p, h, 1, \tau, \eta)$ -clustered configuration, with $\mathbf{x}^c = \{x_\kappa, x_{\kappa+1}, \dots, x_{\kappa+p-1}\}$ and $\|\mathbf{a}\| \geq m > 0$. We also fix $\Omega > 0$ such that $\Omega h \leq \frac{1}{20d}$.

Proposition 5.8 demonstrated the existence of certain λ -decimation maps which achieve good separation of the non-cluster nodes. We define the set $\Lambda(\mathbf{x})$ to consist of all such admissible λ s, as follows.

DEFINITION 5.9 (Admissible blowup factors). For each $F = (\mathbf{a}, \mathbf{x}) \in \mathcal{P}_d$, $\mathbf{x} = (x_1, \dots, x_d)$, such that $\mathbf{x}$ forms a $(p, h, 1, \tau, \eta)$ -clustered configuration and $z_j = z_j(\lambda) = e^{2\pi i \lambda x_j}$, $j = 1, \dots, d$ and $\Omega > 0$, we define the set of admissible blowup factors $\Lambda(\mathbf{x}) = \Lambda_{\Omega, d}(\mathbf{x})$ as the set of all $\lambda \in [\frac{1}{2} \frac{\Omega}{2d-1}, \frac{\Omega}{2d-1}]$ satisfying:

1. For all $\ell \neq j$ such that $x_\ell \in \mathbf{x} \setminus \mathbf{x}^c$ and $x_j \in \mathbf{x}$,

$$\angle(z_\ell(\lambda), z_j(\lambda)) \geq \frac{1}{d^2}. \quad (5.16)$$

2. For all $j \neq k$ such that $x_j, x_k \in \mathbf{x}^c$

$$\angle(z_j(\lambda), z_k(\lambda)) \geq 2\pi \lambda \tau h \geq \frac{\pi}{2d-1} \Omega \tau h. \quad (5.17)$$

5.4.1 The local geometry of admissible decimation maps The next result gives an explicit description of a neighborhood around F where the map FM_λ is injective (and, therefore, we can speak about a local inverse).

DEFINITION 5.10 For each $\alpha, \beta > 0$ we denote by $H_{\alpha,\beta}(F)$ the closed polydisc

$$H_{\alpha,\beta}(F) = \{(\mathbf{a}', \mathbf{x}') \in \tilde{\mathcal{P}}_d \mid \|\mathbf{a}' - \mathbf{a}\| \leq \alpha, \|\mathbf{x}' - \mathbf{x}\| \leq \beta\},$$

and by $H_{\alpha,\beta}^0(F)$ the interior of $H_{\alpha,\beta}(F)$.

The following is proved in Appendix D.

PROPOSITION 5.11 (One-to-one). For each $\lambda \in \Lambda(\mathbf{x})$ the map FM_λ is injective in the open polydisc $U = H_{m, \frac{\tau h}{2\pi}}^0(F) \subset \tilde{\mathcal{P}}_d$.

Next, we can estimate the Lipschitz constants of the inverse map FM_λ^{-1} , using the previously established general bounds in Proposition 5.3.

PROPOSITION 5.12 Let $H = H_{\frac{m}{2}, \frac{\tau h}{4\pi}}(F) \subset U = H_{m, \frac{\tau h}{2\pi}}^0(F)$. Then, for each $F' \in H$:

1. The Jacobian matrix of FM_λ at F' , denoted by $J_\lambda(F')$, is non-degenerate.
2. Put $J_\lambda^{-1}(F') = \begin{bmatrix} A \\ \tilde{B} \end{bmatrix}$, where $A, \tilde{B}$ are $d \times 2d$. Then, the ℓ_1 norms of the rows of the blocks $A, \tilde{B}$ are bounded as follows:

$$\sum_{k=1}^{2d} |A_{j,k}| \leq \tilde{C}, \quad x_j \in \mathbf{x} \setminus \mathbf{x}^c, \quad (5.18)$$

$$\sum_{k=1}^{2d} |\tilde{B}_{j,k}| \leq \tilde{C} \frac{1}{\Omega}, \quad x_j \in \mathbf{x} \setminus \mathbf{x}^c, \quad (5.19)$$

$$\sum_{k=1}^{2d} |A_{j,k}| \leq \tilde{C}(\Omega \tau h)^{-2p+1}, \quad x_j \in \mathbf{x}^c, \quad (5.20)$$

$$\sum_{k=1}^{2d} |\tilde{B}_{j,k}| \leq \tilde{C} \frac{1}{\Omega} (\Omega \tau h)^{-2p+2}, \quad x_j \in \mathbf{x}^c, \quad (5.21)$$

where $\tilde{C} = \tilde{C}(m, d, p)$ is a constant depending only on d, m, p .

Proof. Let $F' = (\mathbf{a}', \mathbf{x}') \in H$, $\mathbf{a}' = (a'_1, \dots, a'_d)$, $\mathbf{x}' = (x'_1, \dots, x'_d)$. Let $z'_j = z'_j(\lambda) = e^{2\pi i \lambda x'_j}$, and let $z_j = z_j(\lambda) = e^{2\pi i \lambda x_j}$, $j = 1, \dots, d$.

By the integral mean value theorem, for each $j = 1, \dots, d$,

$$|z'_j - z_j| = \left| e^{2\pi i \lambda x'_j} - e^{2\pi i \lambda x_j} \right| \leq \lambda \tau h.$$

Let $\ell \neq j$ such that $x_\ell \in \mathbf{x} \setminus \mathbf{x}^c$ and $x_j \in \mathbf{x}$. Since $\lambda \in \Lambda(x)$,

$$\angle(z_\ell, z_j) \geq \frac{1}{d^2}.$$

Then by (5.6)

$$|z_\ell - z_j| \geq \frac{2}{\pi d^2}.$$

We get that

$$|z'_\ell - z'_j| \geq |z_\ell - z_j| - |z'_\ell - z_\ell| - |z'_j - z_j| \geq |z_\ell - z_j| - 2\lambda\tau h \geq \frac{2}{\pi d^2} - 2\lambda\tau h.$$

With $\Omega h \leq \frac{1}{20d}$ and $\lambda \leq \frac{\Omega}{2d-1}$ by assumption, we have that $2\lambda\tau h \leq \frac{1}{3\pi d^2}$ then

$$|z'_\ell - z'_j| \geq \frac{2}{\pi d^2} - 2\lambda\tau h \geq \frac{2}{\pi d^2} - \frac{1}{3\pi d^2} \geq \frac{1}{2d^2}.$$

We conclude that for each $\ell \neq j$ such that $x_\ell \in \mathbf{x} \setminus \mathbf{x}^c$ and $x_j \in \mathbf{x}$

$$|z'_\ell - z'_j| \geq \frac{1}{2d^2}. \quad (5.22)$$

Let $j \neq k$ such that $x_j, x_k \in \mathbf{x}^c$. $\lambda \in \Lambda(\mathbf{x})$ then

$$\angle(z_j, z_k) \geq 2\pi\lambda\tau h.$$

Then by (5.6)

$$|z_j - z_k| \geq 4\lambda\tau h.$$

With a similar argument as above, we get that

$$|z'_j - z'_k| \geq |z_j - z_k| - 2\lambda\tau h \geq 2\lambda\tau h.$$

Using $\lambda \in \Lambda(\mathbf{x}) \Rightarrow \lambda \geq \frac{\Omega}{2(2d-1)}$, we conclude that for each $j \neq k$ such that $x_j, x_k \in \mathbf{x}^c$

$$|z'_j - z'_k| \geq 2\lambda\tau h \geq \frac{1}{2d-1} \Omega\tau h. \quad (5.23)$$

Now, using (5.22) and (5.23), we invoke Proposition 5.3 with $\tilde{h} = \frac{1}{2d-1} \Omega\tau h$ and $\tilde{\eta} = \frac{1}{2d^2}$, and as a result, prove Proposition 5.12 with

$$\tilde{C} = (2d-1)^{2p-1} \max \left[K_1 \left(\frac{1}{2d^2}, d, p \right), K_2 \left(m, \frac{1}{2d^2}, d, p \right), K_3 \left(\frac{1}{2d^2}, d, p \right), K_4 \left(m, \frac{1}{2d^2}, d, p \right) \right].$$

□

DEFINITION 5.13 For $\mathbf{v} \in \mathbb{C}^d$ and $r > 0$, we denote by $Q_r(\mathbf{v})$ the closed cube of radius r centered at $\mathbf{v}$:

$$Q_r(\mathbf{v}) = Q_{r,d}(\mathbf{v}) = \left\{ \mathbf{u} \in \mathbb{C}^d : \|\mathbf{u} - \mathbf{v}\| \leq r \right\}.$$

PROPOSITION 5.14 Let $U = H_{m, \frac{\tau h}{2\pi}}^o(F)$ and $H = H_{\frac{m}{2}, \frac{\tau h}{4\pi}}(F) \subset U$. Let $\lambda \in \Lambda(\mathbf{x})$ and let $\mu_\lambda = FM_\lambda(F)$, then there exists a constant $\tilde{C}_3 = \tilde{C}_3(m, d, p)$ such that for $R = \tilde{C}_3(\Omega \tau h)^{2p-1}$,

$$FM_\lambda(H) \supseteq Q_R(\mu_\lambda).$$

Furthermore for $V_\lambda = FM_\lambda(U)$ let

$$FM_\lambda^{-1} : V_\lambda \rightarrow U$$

be the local inverse of FM_λ , i.e. for all $F' \in U$ we have $FM_\lambda^{-1}(FM_\lambda(F')) = F'$. For each $1 \leq j \leq d$, let $P_{\mathbf{a}_j}, P_{\mathbf{x}_j} : \tilde{\mathcal{P}}_d \rightarrow \mathbb{C}$ be the projections onto the j^{th} amplitude and the j^{th} node coordinates, respectively. Then FM_λ^{-1} is Lipschitz on $Q_R(\mu_\lambda)$ with the following bounds:

$$\begin{aligned} \left| P_{\mathbf{x}_j} FM_\lambda^{-1}(\mu') - P_{\mathbf{x}_j} FM_\lambda^{-1}(\mu'') \right| &\leq \tilde{C}_1 \frac{1}{\Omega} \|\mu'' - \mu'\| \times \begin{cases} 1 & x_j \in \mathbf{x} \setminus \mathbf{x}^c \\ (\Omega \tau h)^{-2p+2} & x_j \in \mathbf{x}^c \end{cases}, \\ \left| P_{\mathbf{a}_j} FM_\lambda^{-1}(\mu') - P_{\mathbf{a}_j} FM_\lambda^{-1}(\mu'') \right| &\leq \tilde{C}_2 \|\mu'' - \mu'\| \times \begin{cases} 1 & x_j \in \mathbf{x} \setminus \mathbf{x}^c \\ (\Omega \tau h)^{-2p+1} & x_j \in \mathbf{x}^c \end{cases}, \end{aligned}$$

for each $\mu'', \mu' \in Q_R(\mu_\lambda)$, where $\tilde{C}_1 = \tilde{C}_1(m, d, p)$, $\tilde{C}_2 = \tilde{C}_2(m, d, p)$ are constants depending only on d, m, p and $\tilde{C}_1 \tilde{C}_3 \leq 1$.

Proof. By Proposition 5.11, FM_λ is injective in the open neighborhood U of the polydisc $H = H_{\frac{m}{2}, \frac{\tau h}{4\pi}}(F)$. In addition, for each $F' \in H$ the inverse Jacobian norm bounds derived in Proposition 5.12 apply. Finally, one can verify (using a similar argument as in the proof of Proposition 5.12) that $J_\lambda(F')$ is non-degenerate for each $F' \in U$. We can therefore invoke Theorem B.1 with U, H and $f = FM_\lambda$ and the bounds (5.18), (5.19), (5.20), (5.21) and conclude that Proposition 5.14 holds with $\tilde{C}_1 = \tilde{C}_2 = \tilde{C}$ and $\tilde{C}_3 = \min\left(\frac{m}{2\tilde{C}}, \frac{1}{4\pi\tilde{C}}\right)$. $\square$

5.4.2 *The global geometry of admissible decimation maps* In this subsection, we give a global description of the geometry of the error set $E_{\epsilon,(\lambda)}(F)$ for any $\lambda \in \Lambda(\mathbf{x})$ and for $\epsilon \leq R$ where $R = \tilde{C}_3(\Omega \tau h)^{2p-1}$, and $\tilde{C}_3$ is as specified in Proposition 5.14.

For each $\lambda \in \Lambda(\mathbf{x})$ let $\mu_\lambda = FM_\lambda(F)$, and put

$$A_{\epsilon,\lambda}(F) = FM_\lambda^{-1}(Q_\epsilon(\mu_\lambda)) \cap \mathcal{P}_d, \quad (5.24)$$

where $FM_\lambda^{-1} : V_\lambda \rightarrow U$ is the local inverse of FM_λ on U .

Observe that $A_{\epsilon,\lambda}(F) \subset E_{\epsilon,(\lambda)}(F)$. The analysis of this subsection will reveal that globally $E_{\epsilon,(\lambda)}(F)$ is made from certain periodic repetitions of the set $A_{\epsilon,\lambda}(F)$ and its permutations.

Consider the following example.

EXAMPLE 5.15 Let $F(x) = \delta(x - \frac{1}{10}) + \delta(x - \frac{2}{10})$, and let $\lambda = \frac{10}{3}$. Applying FM_λ on F , we get that

$$FM_\lambda(F) = (2, e^{\frac{2\pi}{3}i} + e^{-\frac{2\pi}{3}i}, e^{-\frac{2\pi}{3}i} + e^{\frac{2\pi}{3}i}, 2) = (2, -1, -1, 2).$$

If we set $F = (\mathbf{a}, \mathbf{x})$ with $\mathbf{a} = (a_1, a_2) = (1, 1)$ and $\mathbf{x} = (x_1, x_2) = (\frac{1}{10}, \frac{2}{10})$, then clearly the signal $F' = (\mathbf{a}, \mathbf{x}')$, $\mathbf{x}' = (x_2, x_1) = (\frac{2}{10}, \frac{1}{10})$, that is attained by permuting the nodes of the signal F , satisfies that $FM_\lambda(F) = FM_\lambda(F')$. Observe that $F' \notin \mathcal{P}_2$ since its nodes are not in ascending order (a condition that was posed on $\mathcal{P}_d$ to avoid redundant solutions). However, the signal $F'' = (\mathbf{a}, \mathbf{x}'')$ with $\mathbf{x}'' = \mathbf{x}' - \frac{1}{\lambda}(1, 0) = \mathbf{x}' - \frac{3}{10}(1, 0) = (-\frac{1}{10}, \frac{1}{10})$, is in $\mathcal{P}_2$ and it holds that $FM_\lambda(F) = FM_\lambda(F'')$.

One can verify that the set of signals $G \in \mathcal{P}_2$, which satisfies $FM_\lambda(G) = FM_\lambda(F)$ is given by

$$\left\{ G = (\mathbf{a}, \mathbf{y}) \in \mathcal{P}_2 : \mathbf{y} = \mathbf{x} + \frac{1}{\lambda}(n_1, n_2), \quad n_1, n_2 \in \mathbb{Z} \right\} \cup \left\{ G = (\mathbf{a}, \mathbf{y}) \in \mathcal{P}_2 : \mathbf{y} = \mathbf{x}' + \frac{1}{\lambda}(n_1, n_2), \quad n_1, n_2 \in \mathbb{Z} \right\}.$$

In order to formalize the statement regarding the global structure of $E_{\epsilon,(\lambda)}(F)$, which is essentially a generalization of the example above, we require some notation regarding permutation and shift operations.

We denote the set of permutations of d elements by

$$\Pi = \Pi_d \subset \{\pi : \{1, \dots, d\} \rightarrow \{1, \dots, d\}\}.$$

For a vector $\mathbf{x} = (x_1, \dots, x_d) \in \mathbb{C}^d$ and a permutation π , we denote by $\mathbf{x}^\pi$ the vector attained by permuting the coordinates of $\mathbf{x}$ according to π

$$\mathbf{x}^\pi = (x_{\pi(1)}, \dots, x_{\pi(d)}).$$

For a set $A \subseteq \mathcal{P}_d$ and a permutation $\pi \in \Pi_d$, we denote by A^π the set attained from A by permuting the nodes and amplitudes of each signal in A according to π

$$A^\pi = \{(\mathbf{a}^\pi, \mathbf{x}^\pi) : (\mathbf{a}, \mathbf{x}) \in A\}.$$

The following proposition gives a description of the global geometry of $E_{\epsilon,(\lambda)}(F)$. Its proof is presented in Appendix E.

PROPOSITION 5.16 For each $\lambda \in \Lambda(\mathbf{x})$ and $\epsilon \leq R$

$$E_{\epsilon,(\lambda)}(F) = \left(\bigcup_{\pi \in \Pi_d} \bigcup_{\ell \in \mathbb{Z}^d} A_{\epsilon, \lambda}^\pi(F) + \frac{1}{\lambda} \ell \right) \cap \mathcal{P}_d.$$

5.5 Proof of the upper bound

Fix $F = (\mathbf{a}, \mathbf{x}) \in \mathcal{P}_d$, $\mathbf{a} = (a_1, \dots, a_d)$, $\mathbf{x} = (x_1, \dots, x_d) \subset \left[-\frac{1}{2}, \frac{1}{2}\right]$, such that $\mathbf{x}$ forms a $(p, h, 1, \tau, \eta)$ -clustered configuration with $\mathbf{x}^c = \{x_\kappa, x_{\kappa+1}, \dots, x_{\kappa+p-1}\}$, and $\|\mathbf{a}\| \geq m > 0$.

Consider the set of the admissible blowup factors $\Lambda(\mathbf{x})$ (see Definition 5.9). By the analysis of Section 5.4, under the assumption that $\Omega h \leq \frac{1}{20d}$, the following assertions hold:

1. By Proposition 5.11, there exists a neighborhood U of F such that for each $\lambda \in \Lambda(\mathbf{x})$, FM_λ is one-to-one on U .
2. By Proposition 5.14, there exists a constant $\tilde{C}_3 = \tilde{C}_3(m, d, p)$ such for each $\lambda \in \Lambda(\mathbf{x})$, $V_\lambda = FM_\lambda(U)$ contains a cube $Q_R(\mu_\lambda)$, where $\mu_\lambda = FM_\lambda(F)$ and $R = \tilde{C}_3(\Omega \tau h)^{2p-1}$.

For each $\lambda \in \Lambda(\mathbf{x})$ consider the local inverse $FM_\lambda^{-1} : V_\lambda \rightarrow U$ and let (as above)

$$A_{R,\lambda}(F) = FM_\lambda^{-1}(Q_R(\mu_\lambda)) \cap \mathcal{P}_d.$$

The following intermediate claim is proved in Appendix F.

PROPOSITION 5.17 There exist positive constants K_9 and $K_{10} \leq \frac{1}{20d}$ depending only on d , such that for $\frac{K_9}{\eta} \leq \Omega \leq \frac{K_{10}}{h}$ the following holds. There exists $\lambda \in \Lambda(\mathbf{x})$ such that for each pair $(\pi, \ell) \in \Pi_d \times (\mathbb{Z}^d \setminus \{\mathbf{0}\})$, there exists $\lambda_{\pi,\ell} \in \Lambda(\mathbf{x})$ for which

$$\left(A_{R,\lambda}^\pi(F) + \frac{1}{\lambda}\ell\right) \cap E_{R,(\lambda_{\pi,\ell})}(F) = \emptyset. \quad (5.25)$$

With a bit of additional work, we obtain the main geometric result regarding the error set $E_{\epsilon,\Omega}(F)$.

PROPOSITION 5.18 Let Ω as in Proposition 5.17, then there exists $\lambda \in \Lambda(\mathbf{x})$ such that

$$E_{R,\Omega}(F) \subset A_{R,\lambda}(F). \quad (5.26)$$

Proof. Using Proposition 5.17 fix $\lambda^* \in \Lambda(\mathbf{x})$ which satisfies (5.25). We will prove that λ^* satisfies (5.26).

For each $\lambda \in \Lambda(\mathbf{x})$, we have the following result due to Proposition 5.16:

$$E_{R,(\lambda)}(F) \subset \bigcup_{\pi \in \Pi_d} \bigcup_{\ell \in \mathbb{Z}^d} \left(A_{R,\lambda}^\pi(F) + \frac{1}{\lambda}\ell\right). \quad (5.27)$$

Putting $\epsilon = R$ in (5.1), we obtain

$$E_{R,\Omega}(F) \subseteq \bigcap_{\lambda \in (0, \frac{\Omega}{2d-1}]} E_{R,(\lambda)}(F). \quad (5.28)$$

We then obtain (5.26) from (5.25), (5.27) and (5.28) by algebra of sets calculation as follows:

First by (5.28)

$$E_{R,\Omega}(F) \subseteq \bigcap_{\lambda \in (0, \frac{\Omega}{2d-1}]} E_{R,(\lambda)}(F) = E_{R,(\lambda^*)}(F) \cap \left(\bigcap_{\lambda \in (0, \frac{\Omega}{2d-1}]} E_{R,(\lambda)}(F) \right). \quad (5.29)$$

By (5.27)

$$E_{R,(\lambda^*)}(F) \subset \bigcup_{\pi \in \Pi_d, \ell \in \mathbb{Z}^d} \left(A_{R,\lambda^*}^\pi(F) + \frac{1}{\lambda^*} \ell \right). \quad (5.30)$$

Then by (5.29) and (5.30)

$$E_{R,\Omega}(F) \subseteq \left(\bigcup_{\pi \in \Pi_d, \ell \in \mathbb{Z}^d} A_{R,\lambda^*}^\pi(F) + \frac{1}{\lambda^*} \ell \right) \cap \left(\bigcap_{\lambda \in (0, \frac{\Omega}{2d-1}]} E_{R,(\lambda)}(F) \right). \quad (5.31)$$

For each pair $(\pi, \ell) \in \Pi_d \times (\mathbb{Z}^d \setminus \{\mathbf{0}\})$, let $\lambda_{\pi, \ell} \in \Lambda(\mathbf{x})$ be the value asserted by Proposition 5.17, i.e. satisfying (5.25) for $\lambda = \lambda^*$. By this and by (5.31) we have

$$\begin{aligned} E_{R,\Omega}(F) &\subset \left(\bigcup_{\pi \in \Pi_d} A_{R,\lambda^*}^\pi(F) \right) \cup \left(\bigcup_{(\pi, \ell) \in \Pi_d \times (\mathbb{Z}^d \setminus \{\mathbf{0}\})} \left\{ \left(A_{R,\lambda^*}^\pi(F) + \frac{1}{\lambda^*} \ell \right) \cap E_{R,(\lambda_{\pi, \ell})}(F) \right\} \right) \\ &= \bigcup_{\pi \in \Pi_d} A_{R,\lambda^*}^\pi(F). \end{aligned} \quad (5.32)$$

By definition $E_{R,\Omega}(F) \subset \mathcal{P}_d$ where we assume a canonical ascending order of the nodes. Then, we conclude from (5.32) that $E_{R,\Omega}(F) \subset A_{R,\lambda^*}(F)$ which proves (5.26) for $\lambda = \lambda^*$. $\square$

We have everything in place to estimate the diameter of the set $E_{\epsilon, \Omega}(F)$ and its projections.

PROPOSITION 5.19 Let $F = (\mathbf{a}, \mathbf{x}) \in \mathcal{P}_d$, $\mathbf{x} \subset \left[-\frac{1}{2}, \frac{1}{2}\right]$, such that $\mathbf{x}$ forms a $(p, h, 1, \tau, \eta)$ -clustered configuration and $\|\mathbf{a}\| \geq m > 0$. Then, there exist positive constants $C_1, \dots, C_5$, depending only on d, p, m , such that for each $\frac{C_4}{\eta} \leq \Omega \leq \frac{C_5}{h}$ and $\epsilon \leq C_3(\Omega \tau h)^{2p-1}$, it holds that:

$$\begin{aligned} \text{diam}(E_{\epsilon, \Omega}^{\mathbf{x}_j}(F)) &\leq \begin{cases} C_1 \frac{1}{\Omega} (\Omega \tau h)^{-2p+2} \epsilon, & x_j \in \mathbf{x}^c, \\ C_1 \frac{1}{\Omega} \epsilon, & x_j \in \mathbf{x} \setminus \mathbf{x}^c, \end{cases} \\ \text{diam}(E_{\epsilon, \Omega}^{\mathbf{a}_j}(F)) &\leq \begin{cases} C_2 (\Omega \tau h)^{-2p+1} \epsilon, & x_j \in \mathbf{x}^c, \\ C_2 \epsilon, & x_j \in \mathbf{x} \setminus \mathbf{x}^c. \end{cases} \end{aligned}$$

Proof. Let Ω be such that $\frac{K_9}{\eta} \leq \Omega \leq \frac{K_{10}}{h}$, where $K_9 = K_9(d)$, $K_{10} = K_{10}(d)$ are the constants specified in Proposition 5.17. Let $\epsilon \leq \tilde{C}_3(\Omega \tau h)^{2p-1} = R$, where $\tilde{C}_3 = \tilde{C}_3(m, d, p)$ is as specified in Proposition

5.14. Let $F' \in E_{\epsilon, \Omega}(F)$ with $F' = (\mathbf{a}', \mathbf{x}')$. Using Proposition 5.18 fix $\lambda^* \in \Lambda(\mathbf{x})$ which satisfies (5.26), and put $\mu^* = FM_{\lambda^*}(F)$. Consequently

$$F' \in A_{R, \lambda^*}(F) = FM_{\lambda^*}^{-1}(Q_R(\mu^*)) \cap \mathcal{P}_d.$$

Put $\mu' = FM_{\lambda^*}(F')$. By Proposition 5.14, there exist constants $\tilde{C}_1 = \tilde{C}_1(m, d, p)$, $\tilde{C}_2 = \tilde{C}_2(m, d, p)$ such that

$$\begin{aligned} |\mathbf{x}_j - \mathbf{x}'_j| &= \|P_{\mathbf{x}_j} FM_{\lambda^*}^{-1}(\mu^*) - P_{\mathbf{x}_j} FM_{\lambda^*}^{-1}(\mu')\| \leq \begin{cases} \tilde{C}_1 \frac{1}{\Omega} (\Omega h)^{-2p+2} \epsilon, & x_j \in \mathbf{x}^c, \\ \tilde{C}_1 \frac{1}{\Omega} \epsilon, & x_j \in \mathbf{x} \setminus \mathbf{x}^c. \end{cases} \\ |\mathbf{a}_j - \mathbf{a}'_j| &= \|P_{\mathbf{a}_j} FM_{\lambda^*}^{-1}(\mu^*) - P_{\mathbf{a}_j} FM_{\lambda^*}^{-1}(\mu')\| \leq \begin{cases} \tilde{C}_2 (\Omega h)^{-2p+1} \epsilon, & x_j \in \mathbf{x}^c \\ \tilde{C}_2 \epsilon, & x_j \in \mathbf{x} \setminus \mathbf{x}^c. \end{cases} \end{aligned}$$

Since F' was an arbitrary signal in $E_{\epsilon, \Omega}(F)$, we repeat the above argument with $F'' \in E_{\epsilon, \Omega}(F)$ and consequently prove Proposition 5.19 with $C_1 = 2\tilde{C}_1$, $C_2 = 2\tilde{C}_2$, $C_3 = \tilde{C}_3$, $C_4 = K_9$ and $C_5 = K_{10}$. $\square$

We are now in a position to prove Theorem 2.6, essentially by combining Proposition 5.19 with Proposition 4.3.

Proof of Theorem 2.6 Let $F = (\mathbf{a}, \mathbf{x}) \in \mathcal{P}_d$ such that $\mathbf{x}$ forms a (p, h, T, τ, η) -clustered configuration and $\|\mathbf{a}\| \geq m > 0$. Let $\frac{C_4}{\eta T} \leq \Omega \leq \frac{C_5}{h}$ where $C_4 = C_4(d, p, m)$, $C_5 = C_5(d, p, m)$ are the constants specified in Proposition 5.19.

Put $\alpha = (x_1 + x_d)/2$. The signal $SC_T(SH_\alpha(F)) = (\mathbf{a}, \tilde{\mathbf{x}})$, $\tilde{\mathbf{x}} = (\tilde{x}_1, \dots, \tilde{x}_d)$, $\tilde{x}_1 = \frac{x_1 - \alpha}{T}, \dots, \tilde{x}_d = \frac{x_d - \alpha}{T}$ is normalized such that $\tilde{x}_1, \dots, \tilde{x}_d \in [-\frac{1}{2}, \frac{1}{2}]$. The node vector $\tilde{\mathbf{x}}$ forms a $(p, \frac{h}{T}, 1, \tau, \eta)$ -clustered configuration. Applying Proposition 5.19 for $\tilde{F} = SC_T(SH_\alpha(F))$, $\tilde{h} = \frac{h}{T}$, $\tilde{\Omega} = \Omega T \geq \frac{C_4}{\eta}$ and $\tilde{\Omega} \tilde{h} = \Omega h \leq C_5$, we conclude that there exist constants C_1, C_2, C_3 , depending only on d, p, m , such that for any $\epsilon \leq C_3(\Omega \tau h)^{2p-1}$

$$\begin{aligned} \text{diam}(E_{\epsilon, \Omega T}^{\mathbf{x}_j}(SC_T(SH_\alpha(F)))) &\leq \begin{cases} C_1 \frac{1}{\Omega T} (\Omega \tau h)^{-2p+2} \epsilon, & x_j \in \mathbf{x}^c, \\ C_1 \frac{1}{\Omega T} \epsilon, & x_j \in \mathbf{x} \setminus \mathbf{x}^c, \end{cases} \\ \text{diam}(E_{\epsilon, \Omega T}^{\mathbf{a}_j}(SC_T(SH_\alpha(F)))) &\leq \begin{cases} C_2 (\Omega \tau h)^{-2p+1} \epsilon, & x_j \in \mathbf{x}^c, \\ C_2 \epsilon, & x_j \in \mathbf{x} \setminus \mathbf{x}^c. \end{cases} \end{aligned}$$

Applying Proposition 4.3, we conclude the proof Theorem 2.6. $\square$

6. Lower bounds

In this section all the constants $c_1, \dots, k_1, \dots, K_1, \dots$ are unrelated to those of the previous section.

The main technical result we need is the following.

PROPOSITION 6.1 Let $F = (\mathbf{a}, \mathbf{x}) \in \mathcal{P}_d$, such that $\mathbf{x}$ forms a $(p, h, 1, \tau, \eta)$ -clustered configuration, with cluster nodes $\mathbf{x}^c = (x_1, \dots, x_p)$ (according to Definition 2.5), and with $\mathbf{a} \in \mathbb{R}^d$ satisfying $m \leq \|\mathbf{a}\| \leq M$.

Then, there exist constants c_1, k_1, k_2 , depending only on (d, τ, m, M) , such that for all $\epsilon < c_1(\Omega h)^{2p-1}$ and $\Omega h \leq 2$, there exists a signal $F_\epsilon \in \mathcal{P}_d$ satisfying, for some $j_1, j_2 \in \{1, \dots, p\}$,

$$\left| P_{\mathbf{x}_{j_1}}(F_\epsilon) - P_{\mathbf{x}_{j_1}}(F) \right| \geq \frac{k_1}{\Omega} (\Omega h)^{-2p+2} \epsilon, \quad (6.1)$$

$$\left| P_{\mathbf{a}_{j_2}}(F_\epsilon) - P_{\mathbf{a}_{j_2}}(F) \right| \geq k_2 (\Omega h)^{-2p+1} \epsilon, \quad (6.2)$$

$$\left| \mathcal{F}(F_\epsilon)(s) - \mathcal{F}(F)(s) \right| \leq \epsilon, \quad |s| \leq \Omega. \quad (6.3)$$

Assuming validity of Proposition 6.1, let us prove Theorem 2.7.

Proof of Theorem 2.7 Let $\mathbf{a} \in \mathbb{R}^d$ be any real amplitude vector satisfying $m \leq \|\mathbf{a}\| \leq M$. Let Ω, h satisfy $\Omega h \leq 2$, and choose $\mathbf{x}$ to be the configuration with cluster nodes

$$\mathbf{x}^c = (x_1 = 0, x_1 = \tau h, \dots, x_p = (p-1)\tau h),$$

with the rest of the nodes equally spaced in $((p-1)\tau h, 1)$. Now, denote $h' = (p-1)\tau h$ and $\tau' = \frac{1}{p-1}$. Clearly, $\mathbf{x}$ is a $(p, h', 1, \tau', \eta)$ -clustered configuration for all sufficiently small h (for instance, $h < \frac{1}{d} < 1 - \eta(d-p+1)$). Now, we apply Proposition 6.1 with the signal $F = (\mathbf{a}, \mathbf{x})$. Since τ' does not depend on τ , and therefore, the constants c_1, k_1, k_2 depend only on d, p, m, M , we conclude that for $\epsilon < c_1(p-1)^{2p-1}(\Omega \tau h)^{2p-1}$ and $\Omega h < \frac{2}{(p-1)\tau}$, there exist $j_1, j_2 \in \{1, \dots, p\}$ such that

$$\text{diam}(E_{\epsilon, \Omega}^{\mathbf{x}_{j_1}}(F)) \geq \frac{k_1}{\Omega} (p-1)^{-2p+2} \epsilon (\Omega \tau h)^{-2p+2},$$

$$\text{diam}(E_{\epsilon, \Omega}^{\mathbf{a}_{j_2}}(F)) \geq k_2 \epsilon (p-1)^{-2p+1} (\Omega \tau h)^{-2p+1}.$$

Now, we consider the case of a non-cluster node, $x_j \in \mathbf{x} \setminus \mathbf{x}^c$. Let $F = (\mathbf{a}, \mathbf{x})$ be the signal above. Decompose F as follows:

$$F(x) = a_j \delta(x - x_j) + \underbrace{\sum_{\ell \neq j} a_\ell \delta(x - x_\ell)}_{F^0}.$$

Now, let ϵ be fixed. Define $a'_j = a_j + \frac{\epsilon}{2}$ and $x'_j = x_j + \frac{\epsilon}{4\pi\Omega M}$. Put $F'_j(x) = a'_j \delta(x - x'_j) + F^0(x)$. For $|s| \leq \Omega$, the difference between the Fourier transforms of F and F'_j satisfies

$$\begin{aligned} \left| \mathcal{F}(F)(s) - \mathcal{F}(F'_j)(s) \right| &= \left| a_j e^{2\pi i x_j s} - a'_j e^{2\pi i x'_j s} \right| \\ &\leq \left| a_j e^{2\pi i x_j s} \left(1 - e^{2\pi i \frac{\epsilon}{4\pi\Omega M} s} \right) \right| + \left| a'_j - a_j \right| \\ &\leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon. \end{aligned}$$

Since the constants do not depend on τ at all, and the above construction of F'_j can be repeated for each $j \notin \{\kappa, \dots, \kappa + p - 1\}$, the proof of the non-cluster node case is finished.

Again, the case of general T follows by rescaling and applying Proposition 4.3 (as was done in the proof of Theorem 2.6).

This finishes the proof of Theorem 2.7 with $C'_1 = \max\left(\frac{k_1}{(p-1)^{2p-2}}, \frac{1}{4\pi M}\right)$, $C'_2 = \max\left(\frac{1}{2}, \frac{k_2}{(p-1)^{2p-1}}\right)$, $C'_3 = c_1(p-1)^{2p-1}$, $C'_4 = \frac{1}{d}$ and $C'_5 = 2$. $\square$

In the rest of this section, we prove Proposition 6.1.

We start by stating the following result which has been shown in (2, Theorems 4.1 and 4.2).

THEOREM 6.2 Given the parameters $0 < h \leq 2$, $0 < \tau \leq 1$, $0 < m \leq M < \infty$, let the signal $F = (\mathbf{a}, \mathbf{x}) \in \mathcal{P}_d$ with $\mathbf{a} \in \mathbb{R}^d$ form a single uniform cluster as follows:

- (centered) $x_d = -x_1$;
- (uniform) for $1 \leq j < k \leq d$ we have

$$\tau h \leq |x_j - x_k| \leq h;$$

- $m \leq \|a_j\| \leq M$.

Then, there exist constants $K_1, \dots, K_5$ depending only on (d, τ, m, M) such that for every $\epsilon < K_5 h^{2d-1}$, there exists a signal $F_\epsilon = (\mathbf{b}, \mathbf{y}) \in \mathcal{P}_d$ such that

1. $m_k(F) = m_k(F_\epsilon)$ for $k = 0, 1, \dots, 2d-2$, where m_k are given by (A1);
2. $m_{2d-1}(F_\epsilon) = m_{2d-1}(F) + \epsilon$;
3. $K_1 h^{-2d+2} \epsilon \leq \|\mathbf{x} - \mathbf{y}\| \leq K_2 h^{-2d+2} \epsilon$;
4. $K_3 h^{-2d+1} \epsilon \leq \|\mathbf{b} - \mathbf{a}\| \leq K_4 h^{-2d+1} \epsilon$.

Proof of Proposition 6.1 Define F^c and F^{nc} to be the cluster and the non-cluster part of F correspondingly, i.e.

$$F^c = \sum_{x_j \in \mathbf{x}^c} a_j \delta(x - x_j),$$

$$F^{nc} = \sum_{x_j \in \mathbf{x} \setminus \mathbf{x}^c} a_j \delta(x - x_j).$$

Without loss of generality, suppose that F^c is centered, i.e. $x_1 + x_p = 0$. Next, define a blowup of F^c by Ω as follows:

$$F_{(\Omega)}^c = S C_{\frac{1}{\Omega}}(F^c) = \sum_{x_j \in \mathbf{x}^c} a_j \delta(x - \Omega x_j). \quad (6.4)$$

Put $\tilde{d} = p, \tilde{h} = \Omega h$, and let $c_1 = K_5(\tilde{d}, \tau, m, M)$ as in Theorem 6.2. Let $\epsilon \leq c_1(\Omega h)^{2p-1}$. Now, we apply Theorem 6.2 with parameters $\tilde{d}, \tilde{h}, \tau, m, M, \tilde{\epsilon} = c_2 \epsilon$ and the signal $F_{(\Omega)}^c$, where $c_2 \leq 1$ will be determined below. We obtain a signal $G_{(\Omega), \epsilon}^c$ such that the following hold for the difference

$$H = G_{(\Omega),\epsilon}^c - F_{(\Omega)}^c:$$

$$m_k(H) = 0, \quad k = 0, 1, \dots, 2p-2, \quad (6.5)$$

$$m_{2p-1}(H) = \tilde{\epsilon}; \quad (6.6)$$

while also, for some $j_1, j_2 \in \{1, \dots, p\}$

$$\left| P_{\mathbf{x}j_1} \left(G_{(\Omega),\epsilon}^c \right) - P_{\mathbf{x}j_1} \left(F_{(\Omega)}^c \right) \right| \geq K_1 (\Omega h)^{-2p+2} \tilde{\epsilon}, \quad (6.7)$$

$$\left| P_{\mathbf{x}j} \left(G_{(\Omega),\epsilon}^c \right) - P_{\mathbf{x}j} \left(F_{(\Omega)}^c \right) \right| \leq K_2 (\Omega h)^{-2p+2} \tilde{\epsilon}, \quad j = 1, \dots, p, \quad (6.8)$$

$$\left| P_{\mathbf{a}j_2} \left(G_{(\Omega),\epsilon}^c \right) - P_{\mathbf{a}j_2} \left(F_{(\Omega)}^c \right) \right| \geq K_3 (\Omega h)^{-2p+1} \tilde{\epsilon}. \quad (6.9)$$

Now put

$$F_{(\Omega),\epsilon}^c = SC_{\Omega} \left(G_{(\Omega),\epsilon}^c \right).$$

Applying the inverse blowup to the above inequalities, we obtain in fact that

$$\left| P_{\mathbf{x}j_1} \left(F_{(\Omega),\epsilon}^c \right) - P_{\mathbf{x}j_1} \left(F^c \right) \right| \geq \frac{K_1}{\Omega} (\Omega h)^{-2p+2} \tilde{\epsilon}, \quad (6.10)$$

$$\left| P_{\mathbf{a}j_2} \left(F_{(\Omega),\epsilon}^c \right) - P_{\mathbf{a}j_2} \left(F^c \right) \right| \geq K_3 (\Omega h)^{-2p+1} \tilde{\epsilon}. \quad (6.11)$$

From the above definitions, we have $H_{\Omega} = SC_{\Omega}(H) = F_{(\Omega),\epsilon}^c - F^c$. Let us now show that there is a choice of c_3 such that

$$\left| \mathcal{F}(H_{\Omega})(s) \right| \leq \epsilon, \quad |s| \leq \Omega. \quad (6.12)$$

Put $\omega = s/\Omega$, then

$$\mathcal{F}(H_{\Omega})(s) = \mathcal{F}(H)(\omega).$$

Now, we employ the fact that the Fourier transform of a spike train has Taylor series coefficients precisely equal to its algebraic moments (see [1, Proposition 3.1]):

$$\mathcal{F}(H)(\omega) = \sum_{k=0}^{\infty} \frac{1}{k!} m_k(H) (-2\pi i \omega)^k. \quad (6.13)$$

Next, we apply the following easy corollary of the Turán's First Theorem [56, Theorem 6.1], appearing in [14, Theorem 3.1], using the recurrence relation satisfied by the moments of H according to Proposition A.2.

THEOREM 6.3 Let $H = \sum_{j=1}^{2p} \beta_j \delta(x - t_j)$, and put $R = \min_{j=1, \dots, 2p} |t_j|^{-1} > 0$. Then, for all $k \geq 2p$, we have the so-called Taylor domination property

$$|m_k(H)| R^k \leq \left(\frac{2ek}{2p} \right)^{2p} \max_{\ell=0,1,\dots,2p-1} |m_\ell(H)| R^\ell. \quad (6.14)$$

PROPOSITION 6.4 The constant R in Theorem 6.3 satisfies $R \geq C_4$, where C_4 does not depend on Ω, h .

Proof. Recall that $H = G_{(\Omega),\epsilon}^c - F_{(\Omega)}^c$. The nodes of $F_{(\Omega)}^c$ are, by construction, inside the interval $[-\frac{\Omega h}{2}, \frac{\Omega h}{2}]$. The nodes of $G_{(\Omega),\epsilon}^c$, by (6.8), satisfy

$$\begin{aligned} |P_{\mathbf{x},j}(G_{(\Omega),\epsilon}^c)| &\leq \frac{\Omega h}{2} + K_2 (\Omega h)^{-2p+2} \tilde{\epsilon} \\ &\leq \frac{\Omega h}{2} + K_2 (\Omega h)^{-2p+2} c_1 (\Omega h)^{2p-1} \\ &= (\Omega h) \left(c_1 K_2 + \frac{1}{2} \right). \end{aligned}$$

Since $\Omega h \leq 2$ by assumption, this concludes the proof with $C_4 = \frac{1}{2(c_1 K_2 + \frac{1}{2})}$. □

Therefore, by (6.14), (6.5) and (6.6) we have for $k \geq 2p$

$$\begin{aligned} |m_k(H)| &\leq \left(\frac{e}{p} \right)^{2p} k^{2p} R^{2p-1-k} \tilde{\epsilon} \\ &\leq C_5 C_4^{2p-1-k} k^{2p} \tilde{\epsilon}. \end{aligned}$$

Now, plugging this into (6.13) we obtain

$$|\mathcal{F}(H)(\omega)| \leq \frac{\tilde{\epsilon} |2\pi\omega|^{2p-1}}{(2p-1)!} + C_5 C_4^{2p-1} \tilde{\epsilon} \sum_{k \geq 2p} \left(\frac{2\pi|\omega|}{C_4} \right)^k \frac{k^{2p}}{k!}.$$

Put $\zeta = \frac{2\pi|\omega|}{C_4}$, then, since $|\omega| \leq 1$,

$$\begin{aligned} |\mathcal{F}(H)(\omega)| &\leq C_6 \tilde{\epsilon} \sum_{k \geq 2p-1} \zeta^k \frac{k^{2p}}{k!} \\ &\leq C_7 \tilde{\epsilon}. \end{aligned}$$

We can therefore choose $c_2 = \min\left(1, \frac{1}{C_7}\right)$ to ensure that

$$|\mathcal{F}(H)(\omega)| \leq \epsilon, \quad |\omega| \leq 1,$$

which shows (6.12).

Finally, construct the signal $F_\epsilon = F^{nc} + F_{(\Omega),\epsilon}^c$. Combining (6.12), together with (6.10) and (6.11) finishes the proof of Proposition 6.1 with $k_1 = K_1$ and $k_2 = K_3$. $\square$

Acknowledgements

The authors would like to thank the anonymous reviewers for helping them improve the quality of the presentation. The authors would also like to thank Weilin Li for offering a nice technical correction.

Funding

Air Force Office of Scientific Research [FA9550-17-1-0316]; the National Science Foundation [DMS-12552-03]; and the Minerva Foundation.

REFERENCES

1. AKINSHIN, A., BATENKOV, D. & YOMDIN, Y. (2015) Accuracy of spike-train Fourier reconstruction for colliding nodes. In *2015 International Conference on Sampling Theory and Applications (SampTA)*. NYC, NY, USA, IEEE, pp. 617–621.
2. AKINSHIN, A., GOLDMAN, G. & YOMDIN, Y. (2017) Geometry of error amplification in solving Prony system with near-colliding nodes. Preprint arXiv:1701.04058 [math].
3. AUBEL, C. & BÖLCSKEI, H. (2017) Vandermonde matrices with nodes in the unit disk and the large sieve. *Applied and Computational Harmonic Analysis*, **47**, pp. 53–86, Jul. 2019. doi: [10.1016/j.acha.2017.07.006](https://doi.org/10.1016/j.acha.2017.07.006).
4. AUTON, J. R. & VAN BLARICUM, M. L. (1981) Investigation of procedures for automatic resonance extraction from noisy transient electromagnetics data. *Math. Notes*, **1**, 79.
5. AZAÏS, J. M., DE CASTRO, Y. & GAMBOA, F. (2015) Spike detection from inaccurate samplings. *Appl. Comput. Harmon. Anal.*, **38**, 177–195.
6. BATENKOV, D. (2017) Accurate solution of near-colliding Prony systems via decimation and homotopy continuation. *Theoret. Comput. Sci.*, **681**, 27–40.
7. BATENKOV, D. (2018) Stability and super-resolution of generalized spike recovery. *Appl. Comput. Harmon. Anal.*, **45**, 299–323.
8. BATENKOV, D., DEMANET, L., GOLDMAN, G. & YOMDIN, Y. (2020) Conditioning of partial nonuniform Fourier matrices with clustered nodes. *SIAM J. Matrix Anal. Appl.*, **44**, 199–220.
9. BATENKOV, D., DEMANET, L. & MHASKAR, H. N. (2019) Stable soft extrapolation of entire functions. *Inverse Problems*, **35**, pp.1–29.
10. BATENKOV, D., DIEDERICH, B., GOLDMAN, G. & YOMDIN, Y. (2019) The spectral properties of Vandermonde matrices with clustered nodes. Preprint arXiv:1909.01927 [cs, math].
11. GOLDMAN, G., SALMAN, Y. & YOMDIN, Y. (2018) Geometry and singularities of Prony varieties. *Methods Appl. Anal.*, **25**, 257–275.
12. BATENKOV, D. & YOMDIN, Y. (2013) On the accuracy of solving confluent Prony systems. *SIAM J. Appl. Math.*, **73**, 134–154.
13. BATENKOV, D. (2014) Geometry and singularities of the Prony mapping. In *Proceedings of 12th International Workshop on Real and Complex Singularities*, vol. **10**, pp. 1–25.
14. BATENKOV, D. (2016) Taylor Domination, Turán Lemma, and Poincaré-Perron Sequences. In *Contemporary Mathematics*, vol. **659**. (B. MORDUKHOVICH, S. REICH & A. ZASLAVSKI eds.). Providence, Rhode Island: American Mathematical Society, pp. 1–15.
15. BAZÁN, F. (2000) Conditioning of rectangular Vandermonde matrices with nodes in the unit disk. *SIAM J. Matrix Anal. Appl.*, **21**, 679.
16. BECKERMAN, B., GOLUB, G. H. & LABAHN, G. (2007) On the numerical condition of a generalized Hankel eigenvalue problem. *Numer. Math.*, **106**, 41–68.

17. BENEDETTO, J. J. & LI, W. (2018) Super-resolution by means of Beurling minimal extrapolation. *Appl. Comput. Harmon. Anal.* **48**, 218–241. doi [10.1016/j.acha.2018.05.002](https://doi.org/10.1016/j.acha.2018.05.002).
18. BERTERO, M. & BOCCACCI, P. (1998) *Introduction to Inverse Problems in Imaging*. Institute of Physics Publishing, Bristol and Philadelphia.
19. BHASKAR, B. N., TANG, G. & RECHT, B. (2013) Atomic norm denoising with applications to line spectral estimation. *IEEE Trans. Signal Process.*, **61**, 5987–5999.
20. CANDÈS, E. J. & FERNANDEZ-GRANDA, C. (2013) Super-resolution from noisy data. *J. Fourier Anal. Appl.*, **19**, 1229–1254.
21. CANDÈS, E. J. & FERNANDEZ-GRANDA, C. (2014) Towards a mathematical theory of super-resolution. *Comm. Pure Appl. Math.*, **67**, 906–956.
22. CUYT, A. & LEE, W. S. (2018) How to get high resolution results from sparse and coarsely sampled data. *Appl. Comput. Harmon. Anal.* **48**, 1066–1087. doi [10.1016/j.acha.2018.10.001](https://doi.org/10.1016/j.acha.2018.10.001)
23. CUYT, A., TSAI, M. N., VERHOYE, M. & LEE, W. S. (2018) Faint and clustered components in exponential analysis. *Appl. Math. Comput.*, **327**, 93–103.
24. DEMANET, L. & NGUYEN, N. (2015) The recoverability limit for superresolution via sparsity. Preprint arXiv 1502.01385.
25. DENOYELLE, Q., DUVAL, V. & PEYRÉ, G. (2015) Support recovery for sparse deconvolution of positive measures. Preprint arXiv:1506.08264 [cs, math].
26. DONOHO, D. L. (1992) Superresolution via sparsity constraints. *SIAM J. Math. Anal.*, **23**, 1309–1331.
27. FERNANDEZ-GRANDA, C. (2013) Support detection in super-resolution. In *Proceedings of the 10th International Conference on Sampling Theory and Applications (SampTA 2013)*, pp. 145–148.
28. FERNANDEZ-GRANDA, C. (2016) Super-resolution of point sources via convex programming. *Information and Inference: A Journal of the IMA*, **5**, 251–303, <https://doi.org/10.1093/imaiai/iaw005>.
29. FERREIRA, P. (1999) 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. <http://sweet.ua.pt/pjf/PDF/Ferreira99h.pdf>
30. GAUTSCHI, W. (1962) On inverses of Vandermonde and confluent Vandermonde matrices. *Numer. Math.*, **4**, 117–123.
31. GAUTSCHI, W. (1963) On inverses of Vandermonde and confluent Vandermonde matrices II. *Numer. Math.*, **5**, 425–430.
32. GOLUB, G. H., MILANFAR, P. & VARAH, J. (1999) A stable numerical method for inverting shape from moments. *SIAM J. Sci. Comput.*, **21**, 1222–1243.
33. GOLUB, G. H. & PEREYRA, V. (1973) The differentiation of pseudo-inverses and nonlinear least squares problems whose variables separate. *SIAM J. Numer. Anal.*, **10**, 413–432.
34. GOODMAN, J. W. (2005) *Introduction to Fourier Optics*. Roberts and Company Publishers 4950 S. Yosemite Street, F 2 # 197 Greenwood Village, CO 80111, UNITED STATES.
35. HECKEL, R., MORGENTHAU, V. I. & SOLTANOLKOTABI, M. (2016) Super-resolution radar. *Inf. Inference*, **5**, 22–75.
36. HUA, Y. & SARKAR, T. K. (1990) Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise. *IEEE Trans. Acoust. Speech Signal Process.*, **38**, 814–824.
37. HUA, Y. & SARKAR, T. K. (1991) On SVD for estimating generalized eigenvalues of singular matrix pencil in noise. *IEEE Trans. Signal Process.*, **39**, 892–900.
38. KUNIS, S. & NAGEL, D. (2018) On the condition number of Vandermonde matrices with pairs of nearly-colliding nodes. Preprint arXiv:1812.08645[math].
39. LI, W. & LIAO, W. (2018) Stable super-resolution limit and smallest singular value of restricted Fourier matrices. Preprint arXiv:1709.03146v2 [cs, math].
40. LI, W., LIAO, W. & FANNJIANG, A. (2019) Super-resolution limit of the ESPRIT algorithm. arXiv:1905.03782v3 [cs, math]. <https://ieeexplore.ieee.org/abstract/document/9000636>
41. LINDBERG, J. (2012) Mathematical concepts of optical superresolution. *J. Opt.*, **14**, 1–23.

42. MICCHELLI, C. A. & RIVLIN, T. J. (1977) A survey of optimal recovery. In *Optimal Estimation in Approximation Theory, The IBM Research Symposia Series*. Boston, MA: Springer, pp. 1–54.
43. MICCHELLI, C. A. & RIVLIN, T. J. (1985) Lectures on optimal recovery. In TURNER, P. R. (eds) *Numerical Analysis Lancaster 1984*. Springer, Berlin, Heidelberg pp. 21–93. doi [10.1007/BFb0075157](https://doi.org/10.1007/BFb0075157)
44. MICCHELLI, C. A., RIVLIN, T. J. & WINOGRAD, S. (1976) The optimal recovery of smooth functions. *Numer. Math.*, **26**, 191–200.
45. MOITRA, A. (2015) 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*. New York, NY, USA: ACM, pp. 821–830.
46. MORGENSHTERN, V. I. & CANDES, E. J. (2016) Super-resolution of positive sources: the discrete setup. *SIAM J. Imaging Sci.*, **9**, 412–444.
47. O'LEARY, D. P. & RUST, B. W. (2013) Variable projection for nonlinear least squares problems. *Comput. Optim. Appl.*, **54**, 579–593.
48. PEREYRA, V. & SCHERER, G. (2010) *Exponential Data Fitting and Its Applications*. Bentham Science Publishers, Sharjah, the United Arab Emirates.
49. PETER, T. & PLONKA, G. (2013) A generalized Prony method for reconstruction of sparse sums of eigenfunctions of linear operators. *Inverse Problems*, **29**, 1–21.
50. PLONKA, G. & TASCHE, M. (2014) Prony methods for recovery of structured functions. *GAMM-Mitt.*, **37**, 239–258.
51. PRONY, R. (1795) Essai experimental et analytique. *J. Éc. Polytech. Math.*, **2**, 24–76.
52. RANGE, R. M. (2013) *Holomorphic Functions and Integral Representations in Several Complex Variables*, vol. **108**. Springer Science & Business Media New York.
53. SCHIEBINGER, G., ROBEVA, E. & RECHT, B. (2018) Superresolution without separation. *Inf. Inference*, **7**, 1–30.
54. STOICA, P. & MOSES, R. (2005) *Spectral Analysis of Signals*. Pearson/Prentice Hall Upper Saddle River, New Jersey 07458.
55. TANG, G., BHASKAR, B. N. & RECHT, B. (2015) Near minimax line spectral estimation. *IEEE Trans. Inform. Theory*, **61**, 499–512.
56. TURÁN, P., HALÁSZ, G. & PINTZ, J. (1984) *On a New Method of Analysis and Its Applications*. Wiley-Interscience, Hoboken, New Jersey.
57. VETTERLI, M., MARZILIANO, P. & BLU, T. (2002) Sampling signals with finite rate of innovation. *IEEE Trans. Signal Process.*, **50**, 1417–1428.

A. Algebraic Prony system

The so-called Prony system of equations relates the parameters of the signal F as in (1.1) and its *algebraic moments*

$$m_k(F) = \int F(x)x^k dx = \sum_{j=1}^d a_j x_j^k, \quad k = 0, 1, \dots, \quad (\text{A.1})$$

Extending the above to arbitrary complex nodes and amplitudes, we define the *Prony map* $PM : \mathbb{C}^{2d} \rightarrow \mathbb{C}^{2d}$ as follows:

$$PM_k(a_1, \dots, a_d, w_1, \dots, w_d) = \sum_{j=1}^d a_j w_j^k, \quad k = 0, 1, \dots, 2d - 1. \quad (\text{A.2})$$

Now, consider the system of equations defined by PM , i.e. with unknowns $\{a_j, z_j\}_{j=1}^d \in \mathbb{C}^{2d}$ and a given right-hand side $\mu = (\mu_0, \dots, \mu_{2d-1}) \in \mathbb{C}^{2d}$,

$$PM_k(a_1, \dots, a_d, z_1, \dots, z_d) = \mu_k, \quad k = 0, 1, \dots, 2d-1. \quad (\text{A.3})$$

The following fact can be found in the literature about Prony systems and Padé approximation (see e.g. (13) Propositions 3.2 and 3.3).

PROPOSITION A.1 If a solution $(a_1, \dots, a_d, z_1, \dots, z_d)$ to System (A.3) exists with $a_j \neq 0$, $j = 1, \dots, d$ and for $1 \leq j < k \leq d$, $z_j \neq z_k$, it is unique up to a permutation of the nodes $\{z_j\}$ and corresponding amplitudes $\{a_j\}$.

Clearly, the definition of PM_k is valid for arbitrary integer $k \in \mathbb{N}$. The next fact is very well known, and it is the basis of Prony's method of solving (A.3).

PROPOSITION A.2 Let the sequence $v = \{v_k\}_{k \in \mathbb{N}}$ be given by

$$v_k = PM_k(a_1, \dots, a_d, z_1, \dots, z_d).$$

Then, each consecutive $d+1$ elements of v satisfy the following linear recurrence relation:

$$\sum_{\ell=0}^d v_{k+\ell} c_\ell = 0, \quad (\text{A.4})$$

where the constants $\{c_\ell\}_{\ell=0}^d$ are the coefficients of the (monic) polynomial with roots $\{z_1, \dots, z_d\}$ (the 'Prony polynomial'), i.e.

$$Q(z) = \prod_{j=1}^d (z - z_j) \equiv \sum_{\ell=0}^d c_\ell z^\ell. \quad (\text{A.5})$$

Proof. Let $k \in \mathbb{N}$, then

$$\begin{aligned} \sum_{\ell=0}^d v_{k+\ell} c_\ell &= \sum_{\ell=0}^d c_\ell \sum_{j=1}^d a_j z_j^{k+\ell} \\ &= \sum_{j=1}^d a_j z_j^k Q(z_j) = 0. \end{aligned}$$

□

PROPOSITION A.3 (Prony's method). Let there be given the algebraic moments $\{m_k(F)\}_{k=0}^{2d-1}$ of the signal $F = (\mathbf{a}, \mathbf{x})$ where the nodes of $\mathbf{x}$ are pairwise distinct and $\|\mathbf{a}\| > 0$. Then, the parameters $(\mathbf{a}, \mathbf{x})$ can be recovered exactly by the following procedure:

1. Construct the $d \times (d+1)$ Hankel matrix $H = [m_{i+j}]_{0 \leq i \leq d-1}^{0 \leq j \leq d}$.
2. Find a non-zero vector $\mathbf{c}$ in the null-space of H .
3. Find $\mathbf{x}_j$ to be the roots of the Prony polynomial (A.5), whose coefficient vector is $\mathbf{c}$.

4. Find the amplitudes $\mathbf{a}$ by solving the linear system $V\mathbf{a} = \mathbf{m}$, where V is the Vandermonde matrix

$$V = \left[\mathbf{x}_j^k \right]_{k=0, \dots, d-1}^{j=1, \dots, d}.$$

Proof. See e.g. [13]. □

B. Quantitative Inverse Function Theorem

Here, we prove a certain quantitative version of the inverse function theorem, which applies to holomorphic mappings $\mathbb{C}^d \rightarrow \mathbb{C}^d$ (here d is a generic parameter).

For $\mathbf{a} \in \mathbb{C}^d$ and $r_1, \dots, r_d > 0$, let $H_{r_1, \dots, r_d}(\mathbf{a}) \subset \mathbb{C}^d$ be the closed polydisc centered at $\mathbf{a}$,

$$H_{r_1, \dots, r_d}(\mathbf{a}) = \{\mathbf{x} \in \mathbb{C}^d : |\mathbf{x}_j - \mathbf{a}_j| \leq r_j, \text{ for all } j = 1, \dots, d\}.$$

For $j = 1, \dots, d$, we denote by $P_j : \mathbb{C}^d \rightarrow \mathbb{C}$ the orthogonal projection onto the j^{th} coordinate. With some abuse of notation, we will also treat P_j as the $d \times d$ matrix representing this projection.

Finally, recall Definition 5.13 of the hypercube $Q_\cdot$.

THEOREM B.1 Let $U \subseteq \mathbb{C}^d$ be open. Let $f : U \rightarrow \mathbb{C}^d$ be a holomorphic injection with an invertible Jacobian $J(\mathbf{x})$, for all $\mathbf{x} \in U$. For $\mathbf{a} \in U$ and $r_1, \dots, r_d > 0$, let $H(\mathbf{a}) = H_{r_1, \dots, r_d}(\mathbf{a}) \subset U$ be such that for all $\mathbf{x} \in H(\mathbf{a})$,

$$\sum_{k=1}^d |J_{j,k}^{-1}(\mathbf{x})| \leq \alpha_j, \quad j = 1, \dots, d.$$

Put $\mathbf{b} = f(\mathbf{a})$ and $f(U) = V$. Then:

1. For $R = \min(\frac{r_1}{\alpha_1}, \dots, \frac{r_d}{\alpha_d})$, $Q_R(\mathbf{b}) \subseteq f(H(\mathbf{a}))$ and $f^{-1} : V \rightarrow U$ is holomorphic in an open neighborhood of $Q_R(\mathbf{b})$.
2. For each $j = 1, \dots, d$, $f_j^{-1} = P_j f^{-1} : Q_R(\mathbf{b}) \rightarrow \mathbb{C}^d$ is Lipschitz on $Q_R(\mathbf{b})$ with

$$|f_j^{-1}(\mathbf{y}'') - f_j^{-1}(\mathbf{y}')| \leq \alpha_j \|\mathbf{y}'' - \mathbf{y}'\|,$$

for each $\mathbf{y}', \mathbf{y}'' \in Q_R(\mathbf{b})$.

Proof. First, we show that $f(U) = V$ is open and f^{-1} is holomorphic and provides a homeomorphism between U and V .

By assumption $f : U \rightarrow V$ is an injection, then $f^{-1} : V \rightarrow U$ is well defined. By assumption, f is continuously differentiable with non-degenerate Jacobians $J(x)$ for all $x \in U$. Then, by the inverse function theorem V is open and f^{-1} is continuously differentiable on V . We conclude that f is a biholomorphism between U and V .⁶

We now show that for $R = \min(\frac{r_1}{\alpha_1}, \dots, \frac{r_d}{\alpha_d})$, $Q_R(\mathbf{b}) \subseteq f(H(\mathbf{a}))$. f is a homeomorphism between U and V ; hence, $S = f(H(\mathbf{a}))$ is a compact subset of V . We take $Q_R(\mathbf{b}) \subseteq S$ as the maximal cube centered at $\mathbf{b}$ that is contained in S .

⁶ It is an interesting fact that the condition that f has non-degenerate Jacobians on U can be dropped. Contrary to a real version of Theorem B.1 where this condition is necessary, it is true that if f is holomorphic and an injection on the open set U then f is biholomorphism between U and $f(U)$ (see e.g. (52), discussion at page 23).

Then, there exists a point $\mathbf{p}$ such that $\mathbf{p} \in \partial S \cap \partial Q_{R'}(\mathbf{b})$. Put $\mathbf{h} = \mathbf{p} - \mathbf{b}$. f^{-1} is continuously differentiable on $V \supset Q_{R'}(\mathbf{b})$, we can therefore apply the mean value theorem in integral form and obtain (here, the integral is applied to each component of the inverse Jacobian matrix)

$$f^{-1}(\mathbf{b} + \mathbf{h}) - f^{-1}(\mathbf{b}) = \left(\int_0^1 J^{-1}(\mathbf{b} + t\mathbf{h}) dt \right) \mathbf{h}.$$

Then, for each coordinate $j = 1, \dots, d$,

$$f_j^{-1}(\mathbf{b} + \mathbf{h}) - f_j^{-1}(\mathbf{b}) = \left(\int_0^1 P_j J^{-1}(\mathbf{b} + t\mathbf{h}) dt \right) \mathbf{h}. \quad (\text{B.1})$$

f is a homeomorphism between U and V hence f^{-1} maps the boundary of S into boundary of $f^{-1}(S) = Q_r(\mathbf{a})$. Therefore, there exists a coordinate $\hat{j} \in \{1, \dots, d\}$ such that

$$|f_{\hat{j}}^{-1}(\mathbf{b} + \mathbf{h}) - f_{\hat{j}}^{-1}(\mathbf{b})| = r_{\hat{j}}.$$

Then, by equation (B.1)

$$r_{\hat{j}} = |f_{\hat{j}}^{-1}(\mathbf{b} + \mathbf{h}) - f_{\hat{j}}^{-1}(\mathbf{b})| = \left| \left(\int_0^1 P_{\hat{j}} J^{-1}(\mathbf{b} + t\mathbf{h}) dt \right) \mathbf{h} \right| \leq \alpha_{\hat{j}} \|\mathbf{h}\| = \alpha_{\hat{j}} R'.$$

Hence, $R' \geq \frac{r_{\hat{j}}}{\alpha_{\hat{j}}} \geq \min(\frac{r_1}{\alpha_1}, \dots, \frac{r_d}{\alpha_d}) = R$. We get that

$$Q_R(\mathbf{b}) \subseteq Q_{R'}(\mathbf{b}) \subseteq S = f(H(\mathbf{a})).$$

Since we already argued that $V \supset f(H(\mathbf{a})) \supseteq Q_R(\mathbf{b})$ is open, then clearly f^{-1} is holomorphic in an open neighborhood of $Q_R(\mathbf{b})$. This proves item (1) of Theorem B.1.

The second item of the Theorem is proved with a similar argument: let $\mathbf{y}'', \mathbf{y}' \in Q_R(\mathbf{b})$ and put $\mathbf{h}' = \mathbf{y}'' - \mathbf{y}'$. Applying again the mean value theorem

$$|f_j^{-1}(\mathbf{y}' + \mathbf{h}') - f_j^{-1}(\mathbf{y}')| = \left| \left(\int_0^1 P_j J^{-1}(\mathbf{y}' + t\mathbf{h}') dt \right) \mathbf{h}' \right| \leq \alpha_j \|\mathbf{h}'\|.$$

This proves item (2) of the Theorem. $\square$

C. Norm bounds on the inverse Jacobian matrix

Let $F = (\mathbf{a}, \mathbf{x}) \in \tilde{\mathcal{D}}_d$, $\mathbf{a} = (a_1, \dots, a_d)$, $\mathbf{x} = (x_1, \dots, x_d)$. Put $z_j = z_j(\lambda) = e^{2\pi i \lambda x_j}$, $j = 1, \dots, d$. By direct computation, the Jacobian matrix $J = J_\lambda(F) = J_\lambda(\mathbf{a}, \mathbf{x})$, of FM_λ at F is given by

$$J_\lambda(\mathbf{a}, \mathbf{x}) = \begin{bmatrix} 1 & \dots & 1 & 0 & \dots & 0 \\ z_1 & \dots & z_d & 1 & \dots & 1 \\ z_1^2 & \dots & z_d^2 & 2z_1 & \dots & 2z_d \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ z_1^{2d-1} & \dots & z_d^{2d-1} & (2d-1)z_1^{2d-2} & \dots & (2d-1)z_d^{2d-2} \end{bmatrix} \begin{bmatrix} I_d & 0 \\ 0 & D \end{bmatrix}, \quad (\text{C.1})$$

where D is a $d \times d$ diagonal matrix, $D_{jj} = a_j 2\pi i \lambda z_j$, $j = 1, \dots, d$, and I_d is the $d \times d$ identity matrix.

Denote the left-hand matrix in the factorization (C.1) by $U_{2d} = U_{2d}(z_1, \dots, z_d)$. The matrix U_{2d} is an instance of a confluent Vandermonde matrix, whose inverses have been extensively studied in [7, 30,

[31]. In particular, the elements of U_{2d}^{-1} can be constructed using the coefficients of polynomials from an appropriate Hermite interpolation scheme. Consequently, we have the following result due to [31].

THEOREM C.1 (Gautschi [31]; equations (3.10), (3.12)). For $z_1, \dots, z_d \in \mathbb{C}$ pairwise distinct, put

$$U_{2d}^{-1}(z_1, \dots, z_d) = \begin{bmatrix} A \\ B \end{bmatrix},$$

where A, B are $d \times 2d$. Then, we have the following upper bounds on the 1-norm of the rows of the blocks A, B

$$\sum_{k=1}^{2d} |A_{j,k}| \leq (1 + 2(1 + |z_j|)|\Delta_j|)\Gamma_j, \quad j = 1, \dots, d, \quad (\text{C.2})$$

$$\sum_{k=1}^{2d} |B_{j,k}| \leq (1 + |z_j|)\Gamma_j, \quad j = 1, \dots, d, \quad (\text{C.3})$$

where

$$\Delta_j = \sum_{\ell=1, \ell \neq j}^d \frac{1}{|z_j - z_\ell|}, \quad \Gamma_j = \left(\prod_{\ell=1, \ell \neq j}^d \frac{1 + |z_\ell|}{|z_j - z_\ell|} \right)^2.$$

Proof. Proof of Proposition 5.3 By the factorization (C.1)

$$J_\lambda(F) = U_{2d}(z_1, \dots, z_d) \begin{bmatrix} I_d & 0 \\ 0 & D \end{bmatrix},$$

where $z_1 = e^{2\pi i \lambda x_1}, \dots, z_d = e^{2\pi i \lambda x_d}$ and $D = D(z_1, \dots, z_d)$ is the $d \times d$ diagonal matrix, $D_{jj} = a_j 2\pi i \lambda z_j$, $j = 1, \dots, d$.

By assumption, the mapped nodes $\{z_j\}$ are pairwise distinct, and so it immediately follows that $J_\lambda(F)$ is non-degenerate.

Put $U_{2d}^{-1} = U_{2d}^{-1}(z_1, \dots, z_d) = \begin{bmatrix} A \\ B \end{bmatrix}$, where A, B are $d \times 2d$. Put $\tilde{B} = D^{-1}B$. Then

$$J_\lambda^{-1}(F) = \begin{bmatrix} A \\ \tilde{B} \end{bmatrix}. \quad (\text{C.4})$$

By Theorem C.1

$$\sum_{k=1}^{2d} |A_{j,k}| \leq (1 + 2(1 + |z_j|)|\Delta_j|)\Gamma_j, \quad j = 1, \dots, d, \quad (\text{C.5})$$

$$\sum_{k=1}^{2d} |B_{j,k}| \leq (1 + |z_j|)\Gamma_j, \quad j = 1, \dots, d, \quad (\text{C.6})$$

where

$$\Delta_j = \sum_{\ell=1, \ell \neq j}^d \frac{1}{|z_j - z_\ell|}, \quad \Gamma_j = \left(\prod_{\ell=1, \ell \neq j}^d \frac{1 + |z_\ell|}{|z_j - z_\ell|} \right)^2.$$

Non-cluster node Let ℓ be such that $x_\ell \in \mathbf{x} \setminus \mathbf{x}^c$.

By assumptions we have

$$|z_\ell - z_j| \geq \tilde{\eta}, \quad \forall x_\ell \in \mathbf{x} \setminus \mathbf{x}^c, x_j \in \mathbf{x}, \ell \neq j.$$

Then we obtain

$$\Delta_\ell = \sum_{j=1, j \neq \ell}^d \frac{1}{|z_\ell - z_j|} \leq \frac{d-1}{\tilde{\eta}} = K_5(\tilde{\eta}, d), \quad (\text{C.7})$$

while

$$\begin{aligned} \Gamma_\ell &= \left(\prod_{j=1, j \neq \ell}^d \frac{1 + |z_j|}{|z_\ell - z_j|} \right)^2 \leq \left(3^{d-1} \prod_{j=1, j \neq \ell}^d \frac{1}{|z_\ell - z_j|} \right)^2 \\ &\leq \left(3^{d-1} \frac{\tilde{\eta}^{-d+1}}{\left(\lfloor \frac{d-p}{2} \rfloor! \right)^2} \right)^2 \\ &= \left(\left(\frac{3}{\tilde{\eta}} \right)^{d-1} \frac{1}{\left(\lfloor \frac{d-p}{2} \rfloor! \right)^2} \right)^2 \\ &= K_6(\tilde{\eta}, d, p). \end{aligned} \quad (\text{C.8})$$

Inserting equations (C.7) and (C.8) into (C.5) and (C.6), we get

$$\sum_{k=1}^{2d} |A_{\ell,k}| \leq (1 + 2(1 + |z_\ell|)|\Delta_\ell|)\Gamma_\ell \leq (1 + 6K_5)K_6 = K_1(\tilde{\eta}, d, p), \quad (\text{C.9})$$

and

$$\sum_{k=1}^{2d} |B_{\ell,k}| \leq (1 + |z_\ell|)\Gamma_\ell \leq 3K_6 = K_7(\tilde{\eta}, d, p), \quad (\text{C.10})$$

for each ℓ such that $x_\ell \in \mathbf{x} \setminus \mathbf{x}^c$.

Now, we are ready to bound the norms of rows of the blocks $A, \tilde{B}$ for each non-cluster node index.

For the block A , such bound is given in equation (C.9).

For the block $\tilde{B}$, we have, using equation C.10,

$$\sum_{k=1}^{2d} |\tilde{B}_{\ell,k}| = \sum_{k=1}^{2d} |(a_\ell 2\pi i \lambda z_\ell)^{-1}| |B_{\ell,k}| \leq \frac{2K_7}{\pi m} \frac{1}{\lambda} = K_2(m, \tilde{\eta}, d, p) \frac{1}{\lambda}, \quad (\text{C.11})$$

for each ℓ such that $x_\ell \in \mathbf{x} \setminus \mathbf{x}^c$.

This completes the proof of equations (5.2) and (5.3) of Proposition 5.3.

Cluster node

We now bound the norm of each row of $J_\lambda^{-1}(F)$ at an index corresponding to a cluster node.

By assumptions

$$\begin{aligned} |z_j - z_k| &\geq \tilde{h}, & \forall x_j, x_k \in \mathbf{x}^c, j \neq k, \\ |z_j - z_\ell| &\geq \tilde{\eta}, & \forall x_j \in \mathbf{x}^c, x_\ell \in \mathbf{x} \setminus \mathbf{x}^c. \end{aligned}$$

Then, for each j such that $x_j \in \mathbf{x}^c$

$$\Delta_j = \sum_{\ell=1, \ell \neq j}^d \frac{1}{|z_j - z_\ell|} \leq \frac{d-1}{\tilde{h}}, \quad (\text{C.12})$$

while

$$\begin{aligned} \Gamma_j &= \left(\prod_{\ell=1, \ell \neq j}^d \frac{1 + |z_\ell|}{|z_j - z_\ell|} \right)^2 \leq \left(3^{d-1} \prod_{\ell=1, \ell \neq j}^d \frac{1}{|z_j - z_\ell|} \right)^2 \\ &\leq \left(3^{d-1} \frac{\tilde{\eta}^{-d+p} \tilde{h}^{-p+1}}{\left(\lfloor \frac{d-p}{2} \rfloor! \right)^2} \right)^2 \\ &= K_8(\tilde{\eta}, d, p) \tilde{h}^{-2p+2}, \end{aligned} \quad (\text{C.13})$$

$$\text{where } K_8(\tilde{\eta}, d, p) = \left(3^{d-1} \frac{\tilde{\eta}^{-d+p}}{\left(\lfloor \frac{d-p}{2} \rfloor! \right)^2} \right)^2.$$

Inserting equations (C.12) and (C.13) into (C.5) and (C.6), we get

$$\begin{aligned} \sum_{k=1}^{2d} |A_{j,k}| &\leq (1 + 2(1 + |z_j|) \Delta_j) \Gamma_j \leq 7(d-1) K_8 \tilde{h}^{-2p+1} \\ &= K_3(\tilde{\eta}, d, p) \tilde{h}^{-2p+1}, \end{aligned} \quad (\text{C.14})$$

$$\begin{aligned} \sum_{k=1}^{2d} |B_{j,k}| &\leq (1 + |z_j|) \Gamma_j \leq 3 K_8 \tilde{h}^{-2p+2} \\ &= K_9(\tilde{\eta}, d, p) \tilde{h}^{-2p+2}, \end{aligned} \quad (\text{C.15})$$

for each j such that $x_j \in \mathbf{x}^c$.

We now bound the norms of rows of the blocks $A, \tilde{B}$ for each cluster node index.

For the block A , the bound was given in equation (C.14).

For the block $\tilde{B}$, we have, using equation C.15,

$$\begin{aligned} \sum_{k=1}^{2d} |\tilde{B}_{j,k}| &= \sum_{k=1}^{2d} |(a_j 2\pi i \lambda z_j)^{-1} B_{j,k}| \leq \frac{2K_9}{\pi m} \frac{1}{\lambda} \tilde{h}^{-2p+2} \\ &= K_4(\tilde{\eta}, d, p, m) \frac{1}{\lambda} \tilde{h}^{-2p+2}, \end{aligned} \quad (\text{C.16})$$

for each j such that $x_j \in \mathbf{x}^c$.

This completes the proof of equations (5.4) and (5.5) of Proposition 5.3. $\square$

D. Proof of Proposition 5.11

Proof. Let the map $g = g_\lambda : \tilde{\mathcal{P}}_d \simeq \mathbb{C}^{2d} \rightarrow \mathbb{C}^{2d}$ be defined as

$$\begin{aligned} g_k(a_1, \dots, a_d, x_1, \dots, x_d) &= a_k, & k &= 1, \dots, d, \\ g_{d+k}(a_1, \dots, a_d, x_1, \dots, x_d) &= e^{2\pi i \lambda x_k}, & k &= 1, \dots, d. \end{aligned} \quad (\text{D.1})$$

Consider the definition of the Prony map PM from (A.2). We thus have

$$FM_\lambda = PM \circ g_\lambda. \quad (\text{D.2})$$

Put

$$W = g_\lambda(H_{m, \frac{\tau h}{2\pi}}^0(F)) = g_\lambda(U).$$

We will show that g_λ is injective on U and that PM is injective on W .

First, we show that PM is injective on W .

Proposition A.1 gives sufficient conditions for PM to be one-to-one on a subset of $\mathbb{C}^{2d}$, the next Proposition asserts that these conditions hold for W .

PROPOSITION D.1 Let $\lambda \in \Lambda(\mathbf{x})$. Then for each $\mathbf{v}', \mathbf{v}'' \in W = g_\lambda(H_{m, \frac{\tau h}{2\pi}}^0(F)) = g_\lambda(U)$, with $\mathbf{v}' = (\mathbf{a}', \mathbf{z}')$, $\mathbf{a}' = (a'_1, \dots, a'_d)$, $\mathbf{z}' = (z'_1, \dots, z'_d)$, $\mathbf{v}'' = (\mathbf{a}'', \mathbf{z}'')$, $\mathbf{a}'' = (a''_1, \dots, a''_d)$, $\mathbf{z}'' = (z''_1, \dots, z''_d)$ and $\mathbf{v}' \neq \mathbf{v}''$, it holds that:

1. $a'_j \neq 0$ for $j = 1, \dots, d$.
2. $z'_j \neq z'_k$ for each $1 \leq j < k \leq d$.
3. $z'_j \neq z''_k$ for all $1 \leq j < k \leq d$.

Proof. Let $\lambda \in \Lambda(\mathbf{x})$ and let $\mathbf{v}', \mathbf{v}'' \in g_\lambda(H_{m, \frac{\tau h}{2\pi}}^0(F))$ as specified in Proposition D.1.

The first assertion is apparent from the fact that $\|\mathbf{a}' - \mathbf{a}\| < m$ and the assumption that $|a_j| \geq m$ for $j = 1, \dots, d$.

We now prove assertions 2 and 3.

Let $\mathbf{z} = (z_1, \dots, z_d)$, with $z_1 = e^{2\pi i \lambda x_1}, \dots, z_d = e^{2\pi i \lambda x_d}$.

As a first step, we argue that for each pair of mapped nodes z_j, z_k , $1 \leq j < k \leq d$,

$$|z_j - z_k| \geq 4\lambda \tau h, \quad 1 \leq j < k \leq d. \quad (\text{D.3})$$

Indeed, with the assumption that $\Omega h \leq \frac{1}{20d}$ we have that

$$\frac{\pi}{2} > \frac{1}{d^2} > 2\pi\lambda\tau h. \quad (\text{D.4})$$

By (D.4) and since $\lambda \in \Lambda(\mathbf{x})$

$$\angle(z_j, z_k) \geq 2\pi\lambda\tau h. \quad (\text{D.5})$$

Then by (D.4), (D.5) and (5.6)

$$|z_j - z_k| \geq 4\lambda\tau h.$$

Next, we claim that

$$W \subset H_{m, 2\lambda\tau h}^0(\mathbf{a}, \mathbf{z}) = \left\{ (\mathbf{a}', \mathbf{z}') \in \mathbb{C}^{2d} \mid \|\mathbf{a}' - \mathbf{a}\| < m, \|\mathbf{z}' - \mathbf{z}\| < 2\lambda\tau h \right\}. \quad (\text{D.6})$$

Let $(\mathbf{a}''', \mathbf{x}''') \in H_{m, \frac{\tau h}{2\pi}}^0(F)$. To show (D.6), we need to verify that $g_\lambda(\mathbf{a}''', \mathbf{x}''') \in H_{m, 2\lambda\tau h}^0(\mathbf{a}, \mathbf{z})$. For this purpose, put $g_\lambda(\mathbf{a}''', \mathbf{x}''') = (\mathbf{a}''', \mathbf{z}''')$, $\mathbf{z}''' = (e^{2\pi i \lambda x_1'''}, \dots, e^{2\pi i \lambda x_d'''})$. Then, using the integral mean value bound, for any $j = 1, \dots, d$,

$$\begin{aligned} \left| e^{2\pi i \lambda x_j'''} - e^{2\pi i \lambda x_j} \right| &\leq \max_{c \in [x_j + t(x_j''' - x_j) : t \in [0, 1]]} \left| \frac{d}{dx} e^{2\pi i \lambda x} \right|_c \left| \frac{\tau h}{2\pi} \right| \\ &\leq \lambda \tau h e^{\lambda h} \\ &< 2\lambda\tau h, \end{aligned}$$

where in the last step, we used the assumption $\Omega h \leq \frac{1}{20d}$ and the fact that $\lambda \leq \frac{\Omega}{2d-1}$, which then implies that $e^{\lambda h} < 2$. This in turn proves (D.6).

We now prove assertion 2.

Let $1 \leq j < k \leq d$ and assume by contradiction that $z_j' = z_k'$. By (D.6), $(\mathbf{a}', \mathbf{z}') \in H_{m, 2\lambda\tau h}^0(\mathbf{a}, \mathbf{z})$ then $|z_j - z_j'| < 2\lambda\tau h$ and $|z_k - z_j'| = |z_k - z_k'| < 2\lambda\tau h$. Then

$$|z_j - z_k| \leq |z_j - z_j'| + |z_k - z_j'| < 4\lambda\tau h,$$

which is a contradiction to (D.3).

Finally, we prove assertion 3.

Assume by contradiction that for $1 \leq j < k \leq d$, $z_j' = z_k''$. By (D.6) $|z_j - z_j'| < 2\lambda\tau h$. By assumption $|z_k - z_j'| = |z_k - z_k''|$ then by (D.6) $|z_k - z_j'| < 2\lambda\tau h$. Using these

$$|z_j - z_k| \leq |z_j - z_j'| + |z_k - z_j'| < 4\lambda\tau h,$$

which is a contradiction to (D.3).

This completes the proof of Proposition D.1. □

Now, by Propositions D.1 and A.1 we have that PM is injective on W .

We now show that g_λ is injective on U .

PROPOSITION D.2 For each $\lambda > 0$, the map g_λ is injective in the polydisc $H_{m, \frac{1}{2\lambda}}^0(F)$.

Proof. Let $(\mathbf{a}', \mathbf{x}'), (\mathbf{a}'', \mathbf{x}'') \in H_{m, \frac{1}{2\lambda}}^0(F)$ such that $g(\mathbf{a}'', \mathbf{x}'') = g(\mathbf{a}', \mathbf{x}')$. We will show that $(\mathbf{a}', \mathbf{x}') = (\mathbf{a}'', \mathbf{x}'')$.

For the amplitudes coordinates $k = 1, \dots, d$, $g_k(a_1, \dots, a_d, x_1, \dots, x_d) = a_k$ therefore $\mathbf{a}'' = \mathbf{a}'$.
For coordinates $d + 1, \dots, 2d$,

$$g_{d+j}(a_1, \dots, a_d, x_1, \dots, x_d) = g_{d+j}(x_j) = e^{2\pi i \lambda x_j}, \quad j = 1, \dots, d.$$

Fix a certain $1 \leq j \leq d$ and set $x'_j = \alpha'_j + \beta'_j i$, $\alpha'_j, \beta'_j \in \mathbb{R}$. The set of complex numbers $w = \alpha + \beta i$ such that $g_{d+j}(w) = g_{d+j}(x'_j) = e^{2\pi i \lambda x'_j}$ is equal to

$$S_j = \left\{ \alpha + \beta i : \beta = \beta'_j, \alpha = \alpha'_j + \frac{\ell}{\lambda}, \forall \ell \in \mathbb{Z} \right\}.$$

Since $(\mathbf{a}', \mathbf{x}'), (\mathbf{a}'', \mathbf{x}'') \in H_{m, \frac{1}{2\lambda}}^0(F)$ implies that $|x'_j - x''_j| < \frac{1}{\lambda}$ then $x''_j = x'_j$ and because j was chosen arbitrarily we have $\mathbf{x}'' = \mathbf{x}'$. $\square$

By assumption $\lambda \leq \frac{\Omega}{2d-1}$ and $\Omega h \leq \frac{1}{20d}$ then $\frac{1}{\lambda} > h$. Using the former, $U = H_{m, \frac{\tau h}{2\pi}}^0(F) \subset H_{m, \frac{1}{2\lambda}}^0(F)$ then by Proposition D.2 g_λ is injective on U .

We have shown that g_λ is injective on U and that PM is injective on $W = g_\lambda(U)$ then by (D.2) FM_λ is injective on U .

This completes the proof of Proposition 5.11. $\square$

E. Proof of Proposition 5.16

Proof. First, observe that if $F' \in \mathcal{P}_d$ is of the form $F' = (\mathbf{a}'^\pi, \mathbf{x}'^\pi) + \frac{1}{\lambda} \ell$, with $\pi \in \Pi_d$ and $\ell \in \mathbb{Z}^d$, and $(\mathbf{a}', \mathbf{x}') \in A_{\epsilon, \lambda}(F)$ then

$$\begin{aligned} FM_\lambda(F') &= FM_\lambda \left(\left(\mathbf{a}'^\pi, \mathbf{x}'^\pi + \frac{1}{\lambda} \ell \right) \right) \\ &= \sum_{j=1}^d \mathbf{a}'_{\pi(j)} e^{2\pi i \lambda (\mathbf{x}'_{\pi(j)} + \frac{\ell_j}{\lambda})} \\ &= \sum_{j=1}^d \mathbf{a}'_{\pi(j)} e^{2\pi i \lambda \mathbf{x}'_{\pi(j)}} \\ &= \sum_{j=1}^d \mathbf{a}'_j e^{2\pi i \lambda \mathbf{x}'_j} \\ &= FM_\lambda((\mathbf{a}', \mathbf{x}')). \end{aligned}$$

Since by definition of $A_{\epsilon, \lambda}(F)$ (see equation (5.24)), $(\mathbf{a}', \mathbf{x}') \in A_{\epsilon, \lambda}(F)$ implies that $(\mathbf{a}', \mathbf{x}') \in E_{\epsilon, (\lambda)}(F)$, then the above shows that

$$E_{\epsilon, (\lambda)}(F) \supseteq \left(\bigcup_{\pi \in \Pi_d} \bigcup_{\ell \in \mathbb{Z}^d} A_{\epsilon, \lambda}^\pi(F) + \frac{1}{\lambda} \ell \right) \cap \mathcal{P}_d.$$

For the other direction, let $F' = (\mathbf{a}', \mathbf{y}') \in E_{\epsilon, (\lambda)}(F)$ with $\mathbf{a}' = (a'_1, \dots, a'_d)$ and $\mathbf{y}' = (y'_1, \dots, y'_d)$. Put $\mu' = FM_\lambda(F')$, then $\mu' \in Q_\epsilon(\mu_\lambda)$ (with $\mu_\lambda = FM_\lambda(F)$ as above).

By definition of the set $A_{\epsilon, \lambda}(F)$, there exists a signal $F'' \in A_{\epsilon, \lambda}(F)$ such that $FM_\lambda(F'') = \mu'$, and put $F'' = (\mathbf{a}'', \mathbf{x}'')$ with $\mathbf{a}'' = (a''_1, \dots, a''_d)$ and $\mathbf{x}'' = (x''_1, \dots, x''_d)$.

Recall that by (D.2) (see (A.2) and (D.1))

$$FM_\lambda = PM \circ g_\lambda.$$

Put $g_\lambda(F'') = (\mathbf{a}'', \mathbf{z}'')$ with $\mathbf{z}'' = (z''_1, \dots, z''_d)$, $z''_j = e^{2\pi i \lambda x''_j}$ for $j = 1, \dots, d$. By Proposition D.1 each point in $W = g_\lambda(U)$ has non-vanishing amplitudes and pairwise distinct nodes. We have that $F'' \in A_{\epsilon, \lambda}(F) \subseteq U$ and hence $(\mathbf{a}'', \mathbf{z}'')$ satisfies the above properties. Then, by Proposition A.1, the set of all solutions to the equation $PM((\mathbf{a}, \mathbf{z})) = \mu'$ is given by

$$\{(\mathbf{a}''^\pi, \mathbf{z}''^\pi) : \pi \in \Pi_d\}. \quad (\text{E.1})$$

By (E.1) there exists $\pi \in \Pi_d$ such that

$$g_\lambda(F') = g_\lambda((\mathbf{a}', \mathbf{y}')) = (\mathbf{a}''^\pi, \mathbf{z}''^\pi).$$

Finally, since $x''_1, \dots, x''_d$ are real, the set of all solutions to the equation $g_\lambda((\mathbf{a}, \mathbf{x})) = (\mathbf{a}''^\pi, \mathbf{z}''^\pi)$ is given by

$$\left\{(\mathbf{a}''^\pi, \mathbf{x}''^\pi + \frac{1}{\lambda} \boldsymbol{\ell}) : \boldsymbol{\ell} \in \mathbb{Z}^d\right\}.$$

By the above, F' is of the form $(\mathbf{a}''^\pi, \mathbf{x}''^\pi + \frac{1}{\lambda} \boldsymbol{\ell})$ for some $\pi \in \Pi_d$ and $\boldsymbol{\ell} \in \mathbb{Z}^d$.

This concludes the proof of Proposition 5.16. $\square$

F. Proof of Proposition 5.17

Within the course of the proof, we will make appropriate assumptions of the form $\frac{C'}{\eta} \leq \Omega \leq \frac{C''}{h}$, with C', C'' being constants depending only on d , for which some arguments of the proof hold. It is to be understood that K_9 is the maximum of the constants C' and K_{10} is the minimum of the constants C'' .

Assume that $\Omega \geq \frac{2(2d-1)}{\eta}$. Then, the length of the interval $\left[\frac{1}{2} \frac{\Omega}{2d-1}, \frac{\Omega}{2d-1}\right]$ is larger than $\frac{1}{\eta}$ and by Proposition 5.8 there exists an interval $I \subseteq \left[\frac{1}{2} \frac{\Omega}{2d-1}, \frac{1}{2} \frac{\Omega}{2d-1} + \frac{1}{\eta}\right]$ such that

$$I \subset \Lambda(\mathbf{x}), \quad |I| = (2d^2\eta)^{-1}. \quad (\text{F.1})$$

Fix

$$I_1 = [\lambda_1, \lambda_1 + (2d^2\eta)^{-1}] \subseteq \Lambda(\mathbf{x}) \cap \left[\frac{1}{2} \frac{\Omega}{2d-1}, \frac{1}{2} \frac{\Omega}{2d-1} + \frac{1}{\eta}\right]$$

to be the sub-interval of $\Lambda(\mathbf{x}) \cap \left[\frac{1}{2} \frac{\Omega}{2d-1}, \frac{1}{2} \frac{\Omega}{2d-1} + \frac{1}{\eta}\right]$ with the minimal starting point λ_1 which satisfies (F.1). We will show that there exists $\lambda \in I_1$ that satisfies (5.25).

We require the following intermediate results.

As in Section 5.3, we denote by ν the Lebesgue measure on $\mathbb{R}$.

LEMMA F.1 Let $\frac{1}{2} \leq a < 1$ and $I = [a, 1]$. Then, for each $\epsilon, \alpha, c \in \mathbb{R}$ such that $0 < \alpha \leq 1$, $0 < \epsilon \leq \frac{1}{100}\alpha$ and $|c| \geq 8\frac{\epsilon}{\alpha|I|}$, it holds that

$$\nu(\{x \in I : \exists k \in \mathbb{Z} \text{ such that } |kx - c| \leq \epsilon\}) < \alpha|I|.$$

LEMMA F.2 Consider the interval $[a, b] \subset (0, \infty)$ and let $S \subseteq [a, b]$ be a union of N disjoint sub-intervals $S = \bigcup_{i=1}^N [a_i, b_i]$. Set $I^{-1} = [\frac{1}{b}, \frac{1}{a}]$ and $S^{-1} = \bigcup_{i=1}^N [\frac{1}{b_i}, \frac{1}{a_i}]$. Then

$$\frac{\nu(S)}{\nu(I)} \leq \frac{b}{a} \frac{\nu(S^{-1})}{\nu(I^{-1})}.$$

PROPOSITION F.3 There exists constants K_{11}, K_{12} depending only on d such that for $\frac{K_{11}}{\eta} \leq \Omega \leq \frac{K_{12}}{h}$ the following holds. For each $3h < |c| \leq \frac{\eta}{6}$, there exists an interval $I \subset \Lambda(\mathbf{x})$ of length $|I| = (2d^2\eta)^{-1}$ such that for all $\lambda \in I$ and for all $k \in \mathbb{Z}$

$$\left| c - \frac{k}{\lambda} \right| > 3h. \quad (\text{F.2})$$

We now complete the proof of Proposition 5.17 using the claims above and provide their proofs thereafter.

Step 1:

First, it is shown, using Lemma F.1 and Lemma F.2, that there exists $\lambda^* \in I_1$ such that for all pair of distinct nodes i, j with not both x_i, x_j in $\mathbf{x}^c$, it holds that

$$\left| x_i - x_j + \frac{n}{\lambda^*} \right| > (32d^4)^{-1} \frac{1}{\lambda^*}, \quad \text{for all } n \in \mathbb{Z}. \quad (\text{F.3})$$

Put

$$I_1^{-1} = \left[\frac{1}{\lambda_1 + (d^2 2\eta)^{-1}}, \frac{1}{\lambda_1} \right], \quad \tilde{I}_1^{-1} = \lambda_1 I_1^{-1} = \left[\frac{\lambda_1}{\lambda_1 + (d^2 2\eta)^{-1}}, 1 \right].$$

Fix any distinct indices i, j such that not both x_i, x_j are in $\mathbf{x}^c$. Put $c_{ij} = x_i - x_j$ and observe that under the cluster assumption

$$|c_{ij}| \geq \eta. \quad (\text{F.4})$$

Put $I = \tilde{I}_1^{-1}$, $c = c_{ij}\lambda_1$, $\epsilon = (32d^4)^{-1}$ and $\alpha = \frac{1}{d^2}$. We now validate that under appropriate assumptions on the size of Ω , we have that I, c, ϵ, α satisfy the conditions of Lemma F.1. Put a as the left end point of the interval I then with $\Omega \geq \frac{2}{\eta d}$ we have that $a \geq \frac{1}{2}$. With $d \geq 2$ by assumption we have that $\epsilon = \frac{1}{32d^4} < \frac{1}{100d^2}$. With $\Omega \geq \frac{2}{\eta d}$ we have that

$$|I| = |\tilde{I}_1^{-1}| \geq (4d^2\eta\lambda_1)^{-1}. \quad (\text{F.5})$$

Now, with (F.4) and (F.5), we have that $|c| \geq 8\frac{\epsilon}{\alpha|I|}$. Having validated the conditions of Lemma F.1 hold for I, c, ϵ, α we now invoke it and get that

$$\nu(\{t \in \tilde{I}_1^{-1} : \exists k \in \mathbb{Z} \text{ such that } |kt - c_{ij}\lambda_1| \leq (32d^4)^{-1}\}) < \frac{1}{d^2} |\tilde{I}_1^{-1}|.$$

Then

$$\nu(\{t \in I_1^{-1} : \exists k \in \mathbb{Z} \text{ such that } |kt - c_{ij}| \leq (32d^4)^{-1} \frac{1}{\lambda_1}\}) < \frac{1}{d^2} |I_1^{-1}|.$$

Now, we apply Lemma F.2 and conclude from the above that

$$\nu\left(\left\{\lambda \in I_1 : \exists k \in \mathbb{Z} \text{ such that } \left|\frac{k}{\lambda} - c_{ij}\right| \leq (32d^4)^{-1} \frac{1}{\lambda_1}\right\}\right) < \frac{2}{d^2} |I_1|. \quad (\text{F.6})$$

Define the set

$$E = \bigcup_{\substack{1 \leq i < j \leq d \\ \neg(x_i \in \mathbf{x}^c \wedge x_j \in \mathbf{x}^c)}} \left\{\lambda \in I_1 : \exists k \in \mathbb{Z} \text{ such that } \left|\frac{k}{\lambda} - c_{ij}\right| \leq (32d^4)^{-1} \frac{1}{\lambda_1}\right\}.$$

Then, using (F.6) and the union bound

$$\nu(E) < \binom{d}{2} \frac{2}{d^2} |I_1| < |I_1|. \quad (\text{F.7})$$

We conclude from (F.7) that there exists $\lambda^* \in I_1$ which satisfies (F.3).

Step 2:

Now, we show that in fact λ^* satisfies (5.25), i.e. it satisfies the condition of Proposition 5.17.

Let $(\tilde{\pi}, \tilde{\ell}) \in \Pi_d \times (\mathbb{Z}^d \setminus \{\mathbf{0}\})$. We will show that there exists $\lambda_{\tilde{\pi}, \tilde{\ell}} \in \Lambda(\mathbf{x})$ such that for all $\pi \in \Pi_d$ and for all $\ell \in \mathbb{Z}^d$

$$\left(A_{R, \lambda^*}^{\tilde{\pi}}(F) + \frac{1}{\lambda^*} \tilde{\ell}\right) \cap \left(A_{R, \lambda_{\tilde{\pi}, \tilde{\ell}}}^{\pi}(F) + \frac{1}{\lambda_{\tilde{\pi}, \tilde{\ell}}} \ell\right) = \emptyset. \quad (\text{F.8})$$

Proposition 5.17 will then follow by Proposition 5.16.

We can assume without loss of generality that $\tilde{\pi} = id$. Accordingly, we put $A_{R, \lambda^*}^{\tilde{\pi}}(F) = A_{R, \lambda^*}(F)$, and we will prove that there exists $\lambda_{\tilde{\ell}} \in \Lambda(\mathbf{x})$ such that for all $\pi \in \Pi_d$ and for all $\ell \in \mathbb{Z}^d$

$$\left(A_{R, \lambda^*}(F) + \frac{1}{\lambda^*} \tilde{\ell}\right) \cap \left(A_{R, \lambda_{\tilde{\ell}}}^{\pi}(F) + \frac{1}{\lambda_{\tilde{\ell}}} \ell\right) = \emptyset. \quad (\text{F.9})$$

Fix i such that $\tilde{\ell}_i \neq 0$ and set $n = \tilde{\ell}_i$. Assume that $x_i \in \mathbf{x}^c$, and one can verify that the case where $x_i \in \mathbf{x} \setminus \mathbf{x}^c$ is proved using a similar argument to the one that is given below.

In the cases considered below, we will use the following fact about the ‘radius’ of the set $A_{R, \lambda}(F)$ for each $\lambda \in \Lambda(\mathbf{x})$, established in Proposition 5.14. For each $F' = (\mathbf{a}', \mathbf{x}') \in A_{R, \lambda}(F)$ with $\mathbf{x}' = (x'_1, \dots, x'_d)$,

$$\left|x'_j - x_j\right| \leq \tilde{C}_1 \frac{1}{\Omega} (\Omega \tau h)^{-2p+2} R \leq h, \quad j = 1, \dots, d. \quad (\text{F.10})$$

We consider the following mutually exclusive and collectively exhaustive cases:

Case 1: $\frac{n}{\lambda^*} \leq \frac{\eta}{6}$.

Put $c = \frac{n}{\lambda^*}$. Then, under the assumption of this case and with $\Omega \geq \frac{d}{3h}$ we have that $3h < |c| \leq \frac{\eta}{6}$. We can therefore apply Proposition F.3 for c and (under appropriate further assumptions on Ω) get that there exists an interval $I_2 \subset \Lambda(\mathbf{x})$ of length $|I_2| = (2d^2\eta)^{-1}$, such that for all $\lambda \in I_2$ and for all $k \in \mathbb{Z}$ it holds that

$$\left|c - \frac{k}{\lambda}\right| = \left|\frac{n}{\lambda^*} - \frac{k}{\lambda}\right| > 3h. \quad (\text{F.11})$$

Put

$$I_2 = [\lambda_2, \lambda_2 + (d^2 2\eta)^{-1}], \quad I_2^{-1} = \left[\frac{1}{\lambda_2 + (d^2 2\eta)^{-1}}, \frac{1}{\lambda_2}\right], \quad \tilde{I}_2^{-1} = \lambda_2 I_2^{-1}.$$

Let $1 \leq j \leq d$ be any index such that $x_j \in \mathbf{x} \setminus \mathbf{x}^c$. Put $c_j = (x_i + \frac{n}{\lambda^*} - x_j)$. Then

$$|c_j| = |x_i + \frac{n}{\lambda^*} - x_j| \geq |x_i - x_j| - \frac{n}{\lambda^*} \geq \eta - \frac{n}{\lambda^*} \geq \eta - \frac{\eta}{6} \geq \frac{5}{6}\eta, \quad (\text{F.12})$$

where in the second inequality, we used the fact that x_j is a non-cluster node, and in the third inequality, we used the assumption of case 1.

Put $I = I_2^{-1}$, $c = c_j \lambda_2$, $\epsilon = 2h \lambda_2$ and $\alpha = \frac{1}{2d}$. By (F.12) we have that $|c| \geq \frac{5}{6}\eta \lambda_2$. Using the former, one can validate that there exists positive constants $C'(d)$, $C''(d)$ such that if $\frac{C'(d)}{\eta} \leq \Omega \leq \frac{C''(d)}{h}$, then I, c, ϵ, α meet the conditions of Lemma F.1. We then invoke Lemma F.1 and get that

$$v\left(\left\{t \in \tilde{I}_2^{-1} : \exists k \in \mathbb{Z} \text{ such that } |kt - c_j \lambda_2| \leq 2h \lambda_2\right\}\right) < \frac{1}{2d} |\tilde{I}_2^{-1}|.$$

Then

$$v\left(\left\{t \in I_2^{-1} : \exists k \in \mathbb{Z} \text{ such that } |kt - c_j| \leq 2h\right\}\right) < \frac{1}{2d} |I_2^{-1}|.$$

By the above and using Lemma (F.2)

$$v\left(\left\{\lambda \in I_2 : \exists k \in \mathbb{Z} \text{ such that } \left|\frac{k}{\lambda} - c_j\right| \leq 2h\right\}\right) < \frac{1}{d} |I_2|. \quad (\text{F.13})$$

Define the set

$$E = \bigcup_{\substack{1 \leq j \leq d, \\ x_j \notin \mathbf{x}^c}} \left\{\lambda \in I_2 : \exists k \in \mathbb{Z} \text{ such that } \left|\frac{k}{\lambda} - c_j\right| \leq 2h\right\}.$$

Using the union bound and (F.13)

$$v(E) < |I_2|. \quad (\text{F.14})$$

We conclude from the above that there exists $\lambda \in I_2$ such that for any non-cluster node x_j and for any $k \in \mathbb{Z}$

$$\left|x_i + \frac{n}{\lambda^*} - x_j - \frac{k}{\lambda}\right| > 2h.$$

On the other hand, we have that for all $k \in \mathbb{Z}$ (see (F.11))

$$\left|\frac{n}{\lambda^*} - \frac{k}{\lambda}\right| > 3h.$$

Fix $\lambda_{\tilde{\ell}} = \lambda$. Then, using the above, for any $\pi \in \Pi_d$ and any $k \in \mathbb{Z}$, if $x_{\pi(i)}$ is a cluster node then

$$\left|x_i + \frac{n}{\lambda^*} - x_{\pi(i)} - \frac{k}{\lambda_{\tilde{\ell}}}\right| \geq \left|\frac{n}{\lambda^*} - \frac{k}{\lambda_{\tilde{\ell}}}\right| - |x_i - x_{\pi(i)}| > 3h - h = 2h, \quad (\text{F.15})$$

and if $x_{\pi(i)}$ is a non-cluster node then

$$\left|x_i + \frac{n}{\lambda^*} - x_{\pi(i)} - \frac{k}{\lambda_{\tilde{\ell}}}\right| > 2h. \quad (\text{F.16})$$

Now by combining (F.10), (F.15) and (F.16), we get that $\lambda_{\tilde{\ell}}$ satisfies (F.9). This completes the proof of case 1.

Case 2: $\frac{n}{\lambda^*} > \frac{\eta}{6}$ and $\forall y \in \mathbf{x} \setminus \mathbf{x}^c : |x_i + \frac{n}{\lambda^*} - y| > \frac{\eta}{6}$.

We show that in this case, there exists $\lambda \in I_1$ such that $\lambda_{\tilde{\ell}} = \lambda$ satisfies (F.9).

Put (as above)

$$I_1^{-1} = \left[\frac{1}{\lambda_1 + (d^2 2\eta)^{-1}}, \frac{1}{\lambda_1} \right], \quad \tilde{I}_1^{-1} = \lambda_1 I_1^{-1} = \left[\frac{\lambda_1}{\lambda_1 + (d^2 2\eta)^{-1}}, 1 \right].$$

Put $I = \tilde{I}_1^{-1}$, $c = \frac{n}{\lambda^*} \lambda_1$, $\epsilon = 3h\lambda_1$ and $\alpha = \frac{1}{4}$. By the assumptions of this case, we have $\frac{n}{\lambda^*} > \frac{\eta}{6}$, then $c = \frac{n}{\lambda^*} \lambda_1 > \frac{\eta}{6} \lambda_1$. Using the former, one can validate that there exist positive constants $C'(d)$, $C''(d)$ such that if $\frac{C'(d)}{\eta} \leq \Omega \leq \frac{C''(d)}{h}$, then I, c, ϵ, α meet the conditions of Lemma F.1. We then invoke Lemma F.1 and get that

$$\nu \left(\left\{ t \in \tilde{I}_1^{-1} : \exists k \in \mathbb{Z} \text{ such that } \left| kt - \frac{n}{\lambda^*} \lambda_1 \right| \leq 3h\lambda_1 \right\} \right) < \frac{1}{4} |\tilde{I}_1^{-1}|.$$

Then

$$\nu \left(\left\{ t \in I_1^{-1} : \exists k \in \mathbb{Z} \text{ such that } \left| kt - \frac{n}{\lambda^*} \right| \leq 3h \right\} \right) < \frac{1}{4} |I_1^{-1}|.$$

By the above and using Lemma (F.2)

$$\nu \left(\left\{ \lambda \in I_1 : \exists k \in \mathbb{Z} \text{ such that } \left| \frac{k}{\lambda} - \frac{n}{\lambda^*} \right| \leq 3h \right\} \right) < \frac{1}{2} |I_1|. \quad (\text{F.17})$$

Now, for any index j such that x_j is a non-cluster node put $c_j = x_i + \frac{n}{\lambda^*} - x_j$. Put $I = \tilde{I}_1^{-1}$, $c = c_j \lambda_1$, $\epsilon = 2h\lambda_1$ and $\alpha = \frac{1}{4d}$. Then, by the assumptions of this case $|c| > \frac{\eta}{6} \lambda_1$ and with this one can validate that there exist positive constants $C'(d)$, $C''(d)$ such that if $\frac{C'(d)}{\eta} \leq \Omega \leq \frac{C''(d)}{h}$, then I, c, ϵ, α meet the conditions of Lemma F.1. Invoking it and using Lemma (F.2) we have that

$$\nu \left(\left\{ \lambda \in I_1 : \exists k \in \mathbb{Z} \text{ such that } \left| \frac{k}{\lambda} - c_j \right| \leq 2h \right\} \right) < \frac{1}{2d} |I_1|. \quad (\text{F.18})$$

Define the set

$$E = \bigcup_{\substack{1 \leq j \leq d, \\ x_j \notin \mathbf{x}^c}} \left\{ \lambda \in I_1 : \exists k \in \mathbb{Z} \text{ such that } \left| \frac{k}{\lambda} - c_j \right| \leq 2h \right\}.$$

Using the union bound and (F.18)

$$\nu(E) < \frac{1}{2} |I_1|. \quad (\text{F.19})$$

Now, combining (F.17) and (F.19), we get that there exists $\lambda \in I_1$ such that for all $k \in \mathbb{Z}$

$$\begin{aligned} \left| \frac{k}{\lambda} - \frac{n}{\lambda^*} \right| &> 3h, \\ \left| x_i + \frac{n}{\lambda^*} - x_j - \frac{k}{\lambda} \right| &> 2h, \quad \forall x_j \in \mathbf{x} \setminus \mathbf{x}^c. \end{aligned}$$

Finally, setting $\lambda_{\tilde{t}} = \lambda$ we get from the above and (F.10) that $\lambda_{\tilde{t}}$ satisfies (F.9).

Case 3: $\frac{n}{\lambda^*} > \frac{\eta}{6}$ and $\exists y \in \mathbf{x} \setminus \mathbf{x}^c : |x_i + \frac{n}{\lambda^*} - y| \leq \frac{\eta}{6}$.

First, we note that since the non-cluster nodes are each separated from any other node by at least η , there can be at most one node $y \in \mathbf{x} \setminus \mathbf{x}^c$ such that $|x_i + \frac{n}{\lambda^*} - y| \leq \frac{\eta}{6}$. Therefore, let j be the index of the non-cluster node for which we have $|x_i + \frac{n}{\lambda^*} - x_j| \leq \frac{\eta}{6}$. By the choice of λ^* , we also have that

$|x_i + \frac{n}{\lambda^*} - x_j| > (32d^4)^{-1} \frac{1}{\lambda_1}$ (see (F.3)). We conclude that

$$(32d^4)^{-1} \frac{1}{\lambda_1} \leq |x_i + \frac{n}{\lambda^*} - x_j| \leq \frac{\eta}{6},$$

and for $\Omega \leq \frac{1}{96d^3h}$ we then have that

$$3h < |x_i + \frac{n}{\lambda^*} - x_j| \leq \frac{\eta}{6}.$$

We now invoke Proposition F.3 and get that there exists an interval $I_3 \in \Lambda(\mathbf{x})$ of length $|I_3| = (2d^2\eta)^{-1}$ such that for all $\lambda \in I_3$ and for all $k \in \mathbb{Z}$

$$\left| x_i + \frac{n}{\lambda^*} - x_j - \frac{k}{\lambda} \right| > 3h. \quad (\text{F.20})$$

Put

$$I_3 = [\lambda_3, \lambda_3 + (2d^2\eta)^{-1}], \quad I_3^{-1} = \left[\frac{1}{\lambda_3 + (d^2 2\eta)^{-1}}, \frac{1}{\lambda_3} \right], \quad \tilde{I}_3^{-1} = \lambda_3 I_3^{-1}.$$

For each index $1 \leq \ell \leq d$, $\ell \neq j$ put $c_\ell = x_i + \frac{n}{\lambda^*} - x_\ell$ and note that

$$|c_\ell| = |x_i + \frac{n}{\lambda^*} - x_j + x_j - x_\ell| \geq |x_j - x_\ell| - |x_i + \frac{n}{\lambda^*} - x_j| \geq \frac{5}{6}\eta.$$

Put $I = \tilde{I}_3^{-1}$, $c = c_\ell \lambda_3$, $\epsilon = 2h\lambda_3$ and $\alpha = \frac{1}{2d}$. Then, with the above $|c| \geq \frac{5}{6}\eta\lambda_3$ and then following similar computations as in the previous cases (see cases 1, 2), one can validate that I, c, ϵ, α meet the conditions of Lemma F.1 for $\frac{C'}{\eta} \leq \Omega \leq \frac{C''}{h}$ where C', C'' are constants depending only on d . Invoking Lemma F.1 with I, c, ϵ, α we get that

$$\nu\left(\left\{t \in \tilde{I}_3^{-1} : \exists k \in \mathbb{Z} \text{ such that } |kt - c_\ell \lambda_3| \leq 2h\lambda_3\right\}\right) < \frac{1}{2d}|\tilde{I}_3^{-1}|.$$

Then

$$\nu\left(\left\{t \in I_3^{-1} : \exists k \in \mathbb{Z} \text{ such that } |kt - c_\ell| \leq 2h\right\}\right) < \frac{1}{2d}|I_3^{-1}|.$$

By the above and using Lemma (F.2)

$$\nu\left(\left\{\lambda \in I_3 : \exists k \in \mathbb{Z} \text{ such that } \left|\frac{k}{\lambda} - c_\ell\right| \leq 2h\right\}\right) < \frac{1}{d}|I_3|. \quad (\text{F.21})$$

Define the set

$$E = \bigcup_{1 \leq \ell \leq d, \ell \neq j} \left\{ \lambda \in I_3 : \exists k \in \mathbb{Z} \text{ such that } \left|\frac{k}{\lambda} - c_\ell\right| \leq 2h \right\}.$$

Using the union bound and (F.21)

$$\nu(E) < |I_3|.$$

We conclude from the above that there exists $\lambda \in I_3$ such that for all $k \in \mathbb{Z}$ and for any index $1 \leq \ell \leq d$, $\ell \neq j$,

$$\left| x_i + \frac{n}{\lambda^*} - x_\ell - \frac{k}{\lambda} \right| > 2h. \quad (\text{F.22})$$

Put $\lambda_{\tilde{\ell}} = \lambda$. Recall that I_3 satisfies (F.20). Then, with (F.20) and (F.22) $\lambda_{\tilde{\ell}}$ satisfies that for all $k \in \mathbb{Z}$ and for any index $1 \leq \ell \leq d$

$$\left| x_i + \frac{n}{\lambda^*} - x_\ell - \frac{k}{\lambda_{\tilde{\ell}}} \right| > 2h.$$

Using the above and (F.10), we get that that $\lambda_{\tilde{\ell}}$ satisfies (F.9). □

We now prove the intermediate claims: Lemma F.1, Lemma F.2 and Proposition F.3.

Proof of Lemma F.1. Let a, ϵ, α, c and $I = [a, 1]$ as specified in Lemma F.1. Without loss of generality, we assume that $c > 0$, consequently it is sufficient to prove that

$$\nu(\{x \in I : \exists k \in \mathbb{N} \text{ such that } |kx - c| \leq \epsilon\}) < \alpha|I|.$$

If $0 < c < 2$ then one can verify that

$$\nu(\{x \in I : \exists k \in \mathbb{N} \text{ such that } |kx - c| \leq \epsilon\}) \leq 2\epsilon.$$

Then, under this condition and with the assumption that $c \geq 8\frac{\epsilon}{\alpha|I|}$, we have that $2\epsilon < \alpha|I|$, therefore

$$\nu(\{x \in I : \exists k \in \mathbb{N} \text{ such that } |kx - c| \leq \epsilon\}) \leq 2\epsilon < \alpha|I|.$$

We now prove the case $c \geq 2$.

Let $N \in \mathbb{N}$ be the unique integer such that

$$\frac{c}{\lfloor c \rfloor + N} \leq a < \frac{c}{\lfloor c \rfloor + N - 1}. \quad (\text{F.23})$$

Then

$$\nu(\{x \in I : \exists k \in \mathbb{Z} \text{ such that } |kx - c| \leq \epsilon\}) \leq \sum_{k=0}^N \frac{2\epsilon}{\lfloor c \rfloor + k} = 2\epsilon \sum_{k=0}^N \frac{1}{\lfloor c \rfloor + k}. \quad (\text{F.24})$$

If $N \leq 2$ then with $c \geq 8\frac{\epsilon}{\alpha|I|}$

$$2\epsilon \sum_{k=0}^N \frac{1}{\lfloor c \rfloor + k} \leq 2\epsilon \sum_{k=0}^2 \frac{1}{\lfloor c \rfloor + k} < 8\frac{\epsilon}{c} \leq \alpha|I|.$$

Combining (F.24) with the above proves the claim for this case.

We are left to prove the case $N \geq 3, c \geq 2$.

For H_n the n^{th} partial sum of the Harmonic series we have that

$$\log(n) + \gamma < H_n < \log(n+1) + \gamma,$$

where $\log$ is the base 2 logarithm. Then

$$\begin{aligned} 2\epsilon \sum_{k=0}^N \frac{1}{\lfloor c \rfloor + k} &\leq 2\epsilon (\log(\lfloor c \rfloor + N + 1) - \log(\lfloor c \rfloor - 1)) \\ &= 2\epsilon \log \left(\frac{\lfloor c \rfloor + N + 1}{\lfloor c \rfloor - 1} \right) \\ &= 2\epsilon \log \left(1 + \frac{N + 2}{\lfloor c \rfloor - 1} \right). \end{aligned} \quad (\text{F.25})$$

Using (F.23) and since by assumption $a \geq \frac{1}{2}$ we have that

$$N \leq \lfloor c \rfloor + 2. \quad (\text{F.26})$$

Then by (F.23) and (F.26) (and assuming $N \geq 3$, $c \geq 2$)

$$|I| = 1 - a \geq \frac{N-2}{\lfloor c \rfloor + N - 1} \geq \frac{N-2}{2\lfloor c \rfloor + 1} \geq \frac{1}{5} \frac{(N+2)}{2\lfloor c \rfloor + 1} \geq \frac{1}{25} \frac{(N+2)}{\lfloor c \rfloor - 1}. \quad (\text{F.27})$$

Inserting (F.27) into (F.25) and using the assumption that $100\epsilon \leq \alpha$

$$\begin{aligned} 2\epsilon \log \left(1 + \frac{N+2}{\lfloor c \rfloor - 1} \right) &\leq 2\epsilon \log (1 + 25|I|) \\ &= 2\epsilon \log(e) \ln (1 + 25|I|) \\ &< 100\epsilon |I| \\ &\leq \alpha |I|, \end{aligned} \quad (\text{F.28})$$

which then proves the claim using (F.24) and (F.25). □

This completes the proof of Lemma F.1.

Proof of Lemma F.2. For any sub-interval $[c, d] \subseteq I$ we have that

$$\frac{\nu([c, d])}{\nu(I)} = \frac{d - c}{b - a} = \frac{cd \frac{1}{c} - \frac{1}{d}}{ab \frac{1}{a} - \frac{1}{b}} \leq \frac{b \nu\left(\left[\frac{1}{d}, \frac{1}{c}\right]\right)}{a \nu(I^{-1})}. \quad (\text{F.29})$$

Using the above

$$\frac{\nu(S)}{\nu(I)} = \sum_i \frac{\nu([a_i, b_i])}{\nu(I)} \leq \frac{b}{a} \sum_i \frac{\nu\left(\left[\frac{1}{b_i}, \frac{1}{a_i}\right]\right)}{\nu(I^{-1})} = \frac{b}{a} \nu(S^{-1}).$$

This completes the proof of Lemma F.2. □

Proof of Proposition F.3. Without loss of generality, assume that $c > 0$ and put $T = c\lambda_1$.

We will use the following inequality repeatably below. For each $k \geq 0$ and $0 \leq \alpha \leq \lambda_1$ we have

$$\frac{k\alpha}{2\lambda_1^2} \leq k \left(\frac{1}{\lambda_1} - \frac{1}{\lambda_1 + \alpha} \right) \leq \frac{k\alpha}{\lambda_1^2}. \quad (\text{F.30})$$

Put $\beta = T - \lfloor T \rfloor$ and consider the following cases:

Case 1: $\frac{1}{8} \leq \beta \leq \frac{7}{8}$.

We show that in this case $I = I_1 \subset \Lambda(\mathbf{x})$ satisfies (F.2) provided that $\Omega h < \frac{d}{96}$ and $\Omega \geq \frac{4}{d\eta}$. To see this, recall that $I_1 = [\lambda_1, \lambda_1 + (2d^2\eta)^{-1}]$. Put $\lambda(\alpha) = \lambda_1 + \alpha$, $0 \leq \alpha \leq (2d^2\eta)^{-1}$. We have that for each integer $k \leq \lfloor T \rfloor$

$$\left| c - \frac{k}{\lambda(\alpha)} \right| = \frac{T}{\lambda_1} - \frac{k}{\lambda(\alpha)} \geq \frac{\beta}{\lambda_1} \geq \frac{1}{8\lambda_1}.$$

On the other hand, for each integer $k \geq \lceil T \rceil$

$$\begin{aligned}
 \left| c - \frac{k}{\lambda(\alpha)} \right| &\geq \frac{k-T}{\lambda_1} - k \left(\frac{1}{\lambda_1} - \frac{1}{\lambda(\alpha)} \right) \\
 &\geq \frac{k-T}{\lambda_1} - \frac{k\alpha}{\lambda_1^2} = (k-T) \left(\frac{1}{\lambda_1} - \frac{\alpha}{\lambda_1^2} \right) - \frac{T\alpha}{\lambda_1^2} \\
 &\geq (1-\beta) \left(\frac{1}{\lambda_1} - \frac{\alpha}{\lambda_1^2} \right) - \frac{T\alpha}{\lambda_1^2} \\
 &\geq \frac{1}{8} \left(\frac{1}{\lambda_1} - \frac{\alpha}{\lambda_1^2} \right) - \frac{T\alpha}{\lambda_1^2},
 \end{aligned} \tag{F.31}$$

where in the second inequality, we used (F.30). Using $\Omega \geq \frac{4}{d\eta} \Rightarrow \frac{\alpha}{\lambda_1} \leq \frac{1}{2}$, $\frac{T}{\lambda_1} \leq \frac{\eta}{6}$ and $\Omega h < \frac{d}{96}$ we have that

$$\frac{1}{8} \left(\frac{1}{\lambda_1} - \frac{\alpha}{\lambda_1^2} \right) - \frac{T\alpha}{\lambda_1^2} \geq \frac{1}{16\lambda_1} - \frac{1}{32\lambda_1} = \frac{1}{32\lambda_1} > 3h.$$

We conclude from the above that for $\frac{1}{8} \leq \beta \leq \frac{7}{8}$ (and under the assumptions on Ω and Ωh) $I = I_1 \subset \Lambda(\mathbf{x})$ satisfies (F.2).

Case 2: $\beta \leq \frac{1}{8}$.

First, if $\lfloor T \rfloor = 0$ we show that $I = I_1 \subset \Lambda(\mathbf{x})$ satisfies (F.2) for $\Omega h \leq \frac{d}{8}$. For $k = 0$

$$\left| c - \frac{k}{\lambda} \right| = c > 3h.$$

For $k > 0$ and $\lambda \in I_1$

$$\left| c - \frac{k}{\lambda} \right| = \left| \frac{\beta}{\lambda_1} - \frac{k}{\lambda} \right| \geq \frac{1}{\lambda} - \frac{\beta}{\lambda_1} \geq \frac{1}{2\lambda_1} - \frac{1}{8\lambda_1} = \frac{3}{8\lambda_1} > 3h,$$

where in the last inequality, we used the assumption that $\Omega h \leq \frac{d}{8}$.

Now, assume that $\lfloor T \rfloor > 0$ and consider the next inequalities

$$T \left(\frac{1}{\lambda_1} - \frac{1}{\lambda(\alpha)} \right) > 3h, \tag{F.32}$$

$$\lfloor T \rfloor \left(\frac{1}{\lambda_1} - \frac{1}{\lambda(\alpha)} \right) < \frac{1}{4\lambda_1}. \tag{F.33}$$

We show that if for $0 \leq \alpha \leq \lambda_1$, $\lambda(\alpha)$ satisfies both (F.32) and (F.33) then $\lambda(\alpha)$ satisfies (F.2), provided that $\Omega h \leq \frac{d}{24}$.

For any integer $k \leq \lfloor T \rfloor$ we have, using (F.32), that

$$\frac{T}{\lambda_1} - \frac{k}{\lambda(\alpha)} \geq T \left(\frac{1}{\lambda_1} - \frac{1}{\lambda(\alpha)} \right) > 3h.$$

For any integer $k > \lfloor T \rfloor$

$$\begin{aligned}
 \frac{k}{\lambda(\alpha)} - \frac{T}{\lambda_1} &\geq \frac{\lfloor T \rfloor}{\lambda(\alpha)} - \frac{T}{\lambda_1} + \frac{1}{\lambda(\alpha)} \\
 &\geq -\lfloor T \rfloor \left(\frac{1}{\lambda_1} - \frac{1}{\lambda(\alpha)} \right) - \frac{\beta}{\lambda_1} + \frac{1}{\lambda(\alpha)} \\
 &> -\frac{1}{4\lambda_1} - \frac{\beta}{\lambda_1} + \frac{1}{\lambda(\alpha)} \\
 &\geq -\frac{3}{8\lambda_1} + \frac{1}{2\lambda_1} \\
 &\geq \frac{1}{8\lambda_1} \\
 &\geq 3h,
 \end{aligned}$$

where in the third inequality, we used (F.33), in the fourth inequality, we used both $\beta \leq \frac{1}{8}$ and $0 \leq \alpha \leq \lambda_1$, and in the last inequality, we used $\Omega h \leq \frac{d}{24}$.

We then conclude that when Ωh is small enough, each $\lambda(\alpha)$ with $0 \leq \alpha \leq \lambda_1$ which satisfies both (F.32) and (F.33) satisfies (F.2). We now solve (F.32) and (F.33) for α . By (F.30) $\frac{T\alpha}{2\lambda_1^2} > 3h \Rightarrow T \left(\frac{1}{\lambda_1} - \frac{1}{\lambda(\alpha)} \right) > 3h$, then each $0 \leq \alpha \leq \lambda_1$ such that

$$\alpha > \frac{6\lambda_1^2 h}{T}$$

satisfies (F.32). By (F.30) $\frac{\lfloor T \rfloor \alpha}{\lambda_1^2} < \frac{1}{4\lambda_1} \Rightarrow \lfloor T \rfloor \left(\frac{1}{\lambda_1} - \frac{1}{\lambda(\alpha)} \right) < \frac{1}{4\lambda_1}$, then each $0 \leq \alpha \leq \lambda_1$ such that

$$\alpha < \frac{\lambda_1}{4\lfloor T \rfloor},$$

satisfies (F.33).

We conclude from the above that for

$$\alpha \in \left(\frac{6\lambda_1^2 h}{T}, \frac{\lambda_1}{4\lfloor T \rfloor} \right) = I_3,$$

$\lambda(\alpha)$ satisfies (F.2).

Now, we recall that by Proposition 5.8, every interval $I' \subset \left[\frac{1}{2} \frac{\Omega}{2d-1}, \frac{\Omega}{2d-1} \right]$ of size $\frac{1}{\eta}$ contains a sub-interval I of size $(2d^2\eta)^{-1}$ such that $I \subset \Lambda(\mathbf{x})$. Put $I_4 = \lambda_1 + I_3$ and $I_5 = I_4 \cap \left[\frac{1}{2} \frac{\Omega}{2d-1}, \frac{\Omega}{2d-1} \right]$. We will now validate that $|I_5| > \frac{1}{\eta}$ for $\Omega h < \frac{d}{72}$. To prove that, we show that

$$\lambda_1 + \frac{6\lambda_1^2 h}{T} + \frac{1}{\eta} < \min \left(\lambda_1 + \frac{\lambda_1}{4\lfloor T \rfloor}, \frac{\Omega}{2d-1} \right).$$

First, we show that $\lambda_1 + \frac{6\lambda_1^2 h}{T} + \frac{1}{\eta} < \lambda_1 + \frac{\lambda_1}{4[T]}$:

$$\frac{\lambda_1}{4[T]} - \frac{6\lambda_1^2 h}{T} \geq \frac{\lambda_1}{T} \left(\frac{1}{4} - 6\lambda_1 h \right) \geq \frac{6}{\eta} \left(\frac{1}{4} - 6\lambda_1 h \right) > \frac{1}{\eta},$$

where in the penultimate inequality, we used the proposition assumption that $\frac{\eta}{6} \geq c = \frac{T}{\lambda_1}$, and in the last inequality, we used $\Omega h < \frac{d}{72}$. Next, we show that $\lambda_1 + \frac{6\lambda_1^2 h}{T} + \frac{1}{\eta} < \frac{\Omega}{2d-1}$ for $\Omega > \frac{5(2d-1)}{\eta}$ and $\Omega h < \frac{d}{72}$:

$$\begin{aligned} \lambda_1 + \frac{6\lambda_1^2 h}{T} + \frac{1}{\eta} &\leq \lambda_1 (1 + 6\lambda_1 h) + \frac{1}{\eta} \leq \frac{13}{12} \lambda_1 + \frac{1}{\eta} \\ &\leq \frac{13}{12} \left(\frac{\Omega}{2(2d-1)} + \frac{1}{\eta} \right) + \frac{1}{\eta} < \frac{\Omega}{2d-1}. \end{aligned}$$

We conclude that $|I_5| > \frac{1}{\eta}$ and $I_5 \subset \left[\frac{1}{2} \frac{\Omega}{2d-1}, \frac{\Omega}{2d-1} \right]$ then by Proposition 5.8 I_5 contains a sub-interval I of size $(2d^2\eta)^{-1}$ such that $I \subset \Lambda(\mathbf{x})$. Since by construction I_5 satisfies (F2), this completes the proof of the case $\beta \leq \frac{1}{8}$ of Proposition (F3).

We are left to prove the case $\frac{7}{8} \leq \beta$. This case is proved similarly to the case $\beta \leq \frac{1}{8}$. We therefore omit the proof of this case. $\square$