SUPER-RESOLUTION OF NEAR-COLLIDING POINT SOURCES

DMITRY BATENKOV, GIL GOLDMAN, AND YOSEF YOMDIN

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

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

from their spectral measurements, known in a bandwidth Ω with absolute error not exceeding $\epsilon>0$. We consider the case when at most $p\leqslant d$ nodes $\{x_j\}$ of F form a cluster of size $h\ll\frac{1}{\Omega}$, while the rest of the nodes are well separated. Provided that $\epsilon\lessapprox(\Omega h)^{2p-1}$, we show that the minimax error rate for reconstruction of the cluster nodes is of order $(\Omega h)^{-2p+1}h\epsilon$, while for recovering the corresponding amplitudes $\{a_j\}$ the rate is of the order $(\Omega h)^{-2p+1}\epsilon$. Moreover, the corresponding minimax rates for the recovery of the non-clustered nodes and amplitudes are $\frac{\epsilon}{\Omega}$ and ϵ, respectively. Our numerical experiments show that the well-known Matrix Pencil method achieves the above accuracy bounds. These results suggest that stable super-resolution is possible in much more general situations than previously thought, and have implications for analyzing stability of super-resolution algorithms in this regime.

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 [35]. 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 numerical extrapolation, or analytic continuation ([28, 35, 9, 16], and references therein), however, such a prior is generally considered a weak one, due to the *logarithmic continuity* of the corresponding inverse problem. In recent years, considerable progress has been made in studying SR for *sparse signals*, which are frequently modelled as idealized "spike-trains"

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

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, 19].

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

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

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

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

2010 *Mathematics Subject Classification.* Primary 65H10, 94A12, 65J22.

Key words and phrases. Signal reconstruction, spike-trains, Fourier transform, Prony systems, sparsity, super-resolution.

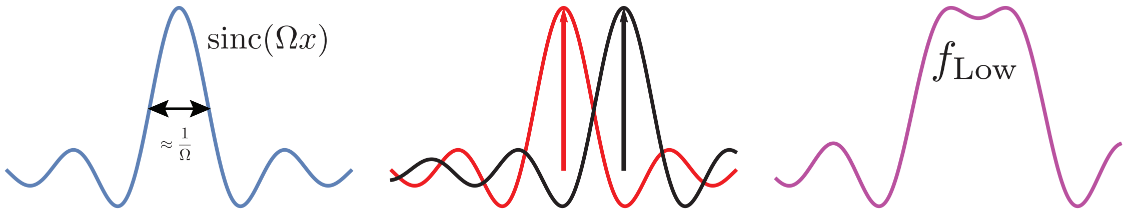


FIGURE 1. The Rayleigh limit. For a signal $F(x) = \sum_j a_j \delta(x - x_j)$, its low resolution version is given by

$$F_{\text{Low}}(x) = \mathcal{F}^{-1}\left(\mathcal{F}(F) \cdot \chi_{[-\Omega,\Omega]}\right) \asymp \sum_j a_j \,\mathrm{sinc}(\Omega(x - x_j)).$$

$F_{\text{Low}}(x)$ 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 $\frac{1}{\Omega}$.

The sparse SR problem is to estimate the parameters $\{a_j, x_j\}$ of the signal F from $\Phi(s)$. If $\epsilon = 0$, the problem can be solved exactly by a variety of parametric methods (Prony's method etc., see e.g. [44, 47] and Subsection 1.2 below). For $\epsilon > 0$, if $F' = F'(\Phi)$ is any reconstruction algorithm receiving Φ as an input and producing an estimate of the signal which satisfies (1.3), then, under an appropriate definition of the distance $\boldsymbol{\delta}(F, F')$, it is of great interest to have a good estimate of the "condition number" $\mathcal{K}$ such that

$$\boldsymbol{\delta}(F, F') \approx \mathcal{K}\epsilon. \tag{1.4}$$

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* $\Delta := \min_{1 \leqslant i < j \leqslant d} |x_i - x_j|$, 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 [47], as well as modern convex optimization based methods solve the problem under some sort of a separation condition of the form $\Delta \geqslant \frac{c}{\Omega}$ [19, 18, 24, 29, 23, 15, 5, 17, 46, 48], 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 (see Figure 1).

The sparse SR problem has appeared already in the work by R. Prony [44], 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, 47, 40, 41, 42, 50] and 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 $F' = \sum_{j=1}^{d} a'_j \delta(x - x'_j)$ satisfies $\min_{1 \leqslant j \leqslant d} |x'_j - x_j| \gtrapprox \Delta$, 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 [21]. It was shown that for *Rayleigh regular* measures supported on a grid of size Δ, the worst-case error from continuous measurements with a band-limit Ω and perturbation of size ϵ (in L_2 sense) scales like $\mathrm{SRF}^{\alpha}\epsilon$, where $\mathrm{SRF} = \frac{1}{\Omega\Delta} > 1$ is the so-called *super-resolution factor*, and α is related to the complexity of the signal. For measures which have on average d spikes per unit of time, α was shown to be between $2d - 1$ and $2d + 1$. In [20] 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 [33] 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 $\mathrm{SRF}^{2d-1}\epsilon$ 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 the 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 size $h \ll \frac{1}{\Omega}$ 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 2.2 below). Under these "p-clustered" assumptions, if the estimated signal F' has nodes x'_j and amplitudes a'_j, we show in Theorem 2.1 and Corollary 2.1 below that for small enough ϵ – and, in particular, for $\epsilon \lessapprox (\Omega h)^{2p-1}$, we have

$$|x_j - x'_j| \asymp \begin{cases} (\Omega h)^{-2p+1} h\epsilon, & \kappa \leqslant j \leqslant \kappa+p-1, \\ \frac{1}{\Omega}\epsilon & 1 \leqslant j < \kappa,\ \kappa+p \leqslant j \leqslant d, \end{cases}$$

$$|a_j - a'_j| \asymp \begin{cases} (\Omega h)^{-2p+1} \epsilon, & \kappa \leqslant j \leqslant \kappa+p-1, \\ \epsilon & 1 \leqslant j < \kappa,\ \kappa+p \leqslant j \leqslant d. \end{cases}$$

These bounds are minimax, i.e. for any estimation algorithm receiving as input Φ satisfying (1.3), and returning an estimate $F' = F'(\Phi)$ with $|\Phi(s) - \mathcal{F}(F')(s)| \leqslant \epsilon, \quad s \in [-\Omega,\Omega]$, the above estimates hold as upper bounds. On the other hand, for each p-clustered signal $F = \sum_{j=1}^d a_j \delta(x - x_j)$ there exists a p-clustered signal $F' = \sum_{j=1}^d a'_j \delta(x - x'_j)$ such that

(1) $|\mathcal{F}(F)(s) - \mathcal{F}(F')(s)| \leqslant \epsilon$, for $|s| \leqslant \Omega$;
(2) the above estimates hold as lower bounds.

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 h 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 super-resolution factor be defined, as before, by $\mathrm{SRF} := (\Omega h)^{-1}$, then the condition number of the cluster nodes scales like SRF^{2p-1} in the super-resolution regime $\mathrm{SRF} \gg 1$, 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 [2, 6, 10, 7, 1, 12, 11, 8]. 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 5.2, has been first proven in [8].

Our numerical experiments show that the above bounds are attained by Matrix Pencil, a well-known high-resolution algorithm.

1.4. **Related work and discussion.** Our main results generalize several previously available bounds for both on-grid and off-grid SR [20, 33, 7], replacing the overall sparsity d with the "local" sparsity p[1]. Compared with previous works, we also have an explicit control of the perturbation ϵ for which the stability bounds hold: $\epsilon \leqslant C \cdot (\Omega h)^{2p-1}$. So, given F satisfying the clustering assumptions and Ω, we can choose $\epsilon = c\,(\Omega h)^{2p-1}$ such that F can be accurately resolved, and c does not depend on Ω, h. But this also means that given $\epsilon > 0$, we can choose h_0 and Ω_0 such that $(\Omega_0 h_0)^{2p-1} \geqslant \frac{\epsilon}{c}$, and for any F satisfying the clustering assumptions with $h = h_0$ and $\Omega = \Omega_0$, the SR problem can be accurately solved. Therefore, we have a lower bound for the SRF values for which we can expect stable recovery, that reads $\mathrm{SRF} \gtrapprox \left(\frac{1}{\epsilon}\right)^{\frac{1}{2p-1}}$. 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 $h_\epsilon \lessapprox \frac{1}{\Omega}\epsilon^{\frac{1}{2p-1}}$, giving an upper bound for the attainable SRF values exactly matching the lower bound above. To summarize, we obtain the best scaling of the attainable resolution with clustered sparsity p and absolute perturbation ϵ:

$$\mathrm{SRF} \asymp \sqrt[2p-1]{\frac{1}{\epsilon}}. \tag{1.5}$$

This Hölder-type stability is much more favorable compared to *logarithmic* stability of 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 references therein.

Stable super-resolution in the "on-grid" setting of [20, 21, 33] is closely related to the smallest singular value of a certain class of Fourier-type matrices. Using the decimation technique, we have shown in a recent paper [8] that the asymptotic scaling of the condition number for on-grid super-resolution is SRF^{2p-1}, matching the off-grid setting of the present paper. This result extends and generalizes previously known bounds [3, 39, 14, 25], as well as recent works [32, 33].

Available studies of certain high-resolution algorithms such as MUSIC [34], ESPRIT/Matrix Pencil [22], Approximate Prony Method [43] and others do not provide rigorous performance guarantees in the case $\mathrm{SRF} > 1$. Our numerical experiments suggest that the Matrix Pencil is optimal in the high SRF regime, and we hope that our proof techniques may be used in deriving the stability limits of these and other methods in the super-resolution regime. 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.

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 Matrix Pencil algorithm. The proof of the main result, Theorem 2.1 is presented in the subsequent sections 4, 5 and 6.

1.6. **Acknowledgements.** The research of GG and YY is supported in part by the Minerva Foundation. DB is supported in part by AFOSR grant FA9550-17-1-0316, NSF grant DMS-1255203, and a grant from the MIT-Skolkovo initiative.

2. Minimax bounds for clustered super-resolution

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

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

[1]Our clustering model is distinct from Donoho's model of "sparse clumps" on a grid [21], and so the two results cannot be compared directly.

and identify signals F with their parameters $(\mathbf{a},\mathbf{x})\in\mathcal{P}$. 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 $\mathcal{P}_d$. For $F=(\mathbf{a},\mathbf{x}), F'=(\mathbf{a}',\mathbf{x}')\in\mathcal{P}_d$,

$$\boldsymbol{\delta}(F,F') := \|F-F'\| = \max\left(\|\mathbf{a}-\mathbf{a}'\|_\infty, \|\mathbf{x}-\mathbf{x}'\|_\infty\right).$$

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

Let $U\subset\mathcal{P}_d$. We consider the minimax error rate in estimating a signal $F\in U$ [2] from $\Phi(s)$ as in (1.3), with measurement error $\epsilon>0$. For each $F\in U$ and a measurement $\Phi(s)$, let $e(s)=\Phi(s)-\mathcal{F}(F)(s)$. Clearly, $|e(s)|\leqslant\epsilon$. Let $F'=F'(\Phi)$ denote any estimator of $F\in U$. The minimax error for U is defined as

$$\Lambda(\epsilon,U,\Omega) = \inf_{F'}\sup_{F\in U}\sup_{e:\ |e(s)|\leqslant\epsilon}\|F-F'\|.$$

Similarly the minimax errors of estimating the individual nodes, respectively, the amplitudes of $F=(\mathbf{a},\mathbf{x})\in U$ are defined by

$$\begin{aligned}\Lambda^{\mathbf{x},j}(\epsilon,U,\Omega) &= \inf_{F'=(\mathbf{a}',\mathbf{x}')}\ \sup_{F=(\mathbf{a},\mathbf{x})\in U}\ \sup_{e:\ |e(s)|\leqslant\epsilon}\|\mathbf{x}_j-\mathbf{x}'_j\|,\\ \Lambda^{\mathbf{a},j}(\epsilon,U,\Omega) &= \inf_{F'=(\mathbf{a}',\mathbf{x}')}\ \sup_{F=(\mathbf{a},\mathbf{x})\in U}\ \sup_{e:\ |e(s)|\leqslant\epsilon}\|\mathbf{a}_j-\mathbf{a}'_j\|.\end{aligned}$$

Let a signal $F\in\mathcal{P}_d$ as above be fixed. We define the ϵ-error set $E_\epsilon(F)$ to be the inverse image of an ϵ-cube under the Fourier transform mapping. Formally:

Definition 2.1. *The error set $E_{\epsilon,\Omega}(F)=E_\epsilon(F)\subset\mathcal{P}_d$ is the set consisting of all the signals $F'\in\mathcal{P}_d$ with*

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

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:

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

For any subset V of a metric space with metric $\boldsymbol{\delta}$, the diameter of V is

$$diam(V) = \sup_{v',v''\in V}\boldsymbol{\delta}(v',v'').$$

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 [37, 36, 38] (see also [21, 20, 33]).

[2]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.

Proposition 2.1. *For $U \subset \mathcal{P}_d$, $\Omega > 0$, $1 \leqslant j \leqslant d$ and $\epsilon > 0$ we have*

$$\frac{1}{2} \sup_{F:\ E_{\frac{1}{2}\epsilon,\Omega}(F) \subseteq U} diam\big(E_{\frac{1}{2}\epsilon,\Omega}(F)\big) \leqslant \Lambda(\epsilon, U, \Omega) \leqslant \sup_{F \in U} diam\big(E_{2\epsilon,\Omega}(F)\big) \tag{2.2}$$

$$\frac{1}{2} \sup_{F:\ E_{\frac{1}{2}\epsilon,\Omega}(F) \subseteq U} diam\big(E^{\mathbf{x},j}_{\frac{1}{2}\epsilon,\Omega}(F)\big) \leqslant \Lambda^{\mathbf{x},j}(\epsilon, U, \Omega) \leqslant \sup_{F \in U} diam\big(E^{\mathbf{x},j}_{2\epsilon,\Omega}(F)\big) \tag{2.3}$$

$$\frac{1}{2} \sup_{F:\ E_{\frac{1}{2}\epsilon,\Omega}(F) \subseteq U} diam\big(E^{\mathbf{a},j}_{\frac{1}{2}\epsilon,\Omega}(F)\big) \leqslant \Lambda^{\mathbf{a},j}(\epsilon, U, \Omega) \leqslant \sup_{F \in U} diam\big(E^{\mathbf{a},j}_{2\epsilon,\Omega}(F)\big) \tag{2.4}$$

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

Let $\epsilon > 0$. For the upper bound we note that for any $F \in U$ and a measurement Φ, for any estimator F' of F which satisfies $|\mathcal{F}(F')(s) - \Phi(s)| \leqslant \epsilon,\ s \in [-\Omega, \Omega]$, we have that $F' \in E_{2\epsilon}(F)$. Therefore $\|F - F'\| \leqslant diam\big(E_{2\epsilon}(F)\big)$.

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 $\|F^1 - F^2\| = diam\big(E_{\frac{1}{2}\epsilon}(F)\big) - \xi$. Let $\Phi = \mathcal{F}(F)$, and let $F' = F'(\Phi)$ be the output of a certain estimator corresponding to the input Φ. We have $\big|\Phi(s) - \mathcal{F}(F^1)(s)\big|,\ \big|\Phi(s) - \mathcal{F}(F^2)(s)\big| \leqslant \epsilon$. Consequently, there exist perturbations of size at most ϵ to F^1 and F^2, such that the output of the estimator in these two cases is F'. We have

$$diam\big(E_{\frac{1}{2}\epsilon}(F)\big) - \xi = \|F^1 - F^2\| \leqslant \|F^1 - F'\| + \|F^2 - F'\|,$$

so either $\|F^1 - F'\| \geqslant \frac{1}{2} diam\big(E_{\frac{1}{2}\epsilon}(F)\big) - \frac{\xi}{2}$ or $\|F^2 - F'\| \geqslant \frac{1}{2} diam\big(E_{\frac{1}{2}\epsilon}(F)\big) - \frac{\xi}{2}$. Since the estimator was arbitrary, the lower bound follows by letting $\xi \to 0$. ☐

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

First, for a node vector $\mathbf{x} \in \mathbb{R}^d$ with $x_1 < x_2, \ldots < x_d$, we denote by $T(\mathbf{x})$ its overall extent $T = T(\mathbf{x}) := x_d - x_1$.

Definition 2.2 (Uniform cluster configuration)**.** *Given $0 < \eta, \tau \leqslant 1$, a node vector $\mathbf{x} = (x_1, \ldots, x_d) \in \mathbb{R}^d$ with $T = T(\mathbf{x})$ is said to form a (p, h, T, τ, η)-cluster, if there exists a subset of p nodes $\mathbf{x}^c = \{x_\kappa, \ldots, x_{\kappa+p-1}\} \subset \mathbf{x}$, $p \geqslant 2$ which satisfies the following conditions:*

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

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

(2) for $x_\ell \in \mathbf{x} \backslash \mathbf{x}^c$ and $x_j \in \mathbf{x}$, $\ell \neq j$,

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

The main result of this paper is a uniform bound on $diam\left(E_{\epsilon,\Omega}(F)\right)$ and its coordinate projections for any signal F forming a cluster as above (and with non-vanishing amplitudes).

Theorem 2.1. *Let $F = (\mathbf{a}, \mathbf{x}) \in \mathcal{P}_d$, such that the nodes of $\mathbf{x}$ form a (p, h, T, τ, η)-cluster and furthermore, for $j = 1, .., d$ we have $0 < m \leqslant |a_j| \leqslant M < \infty$. Then there exist constants $C_1, \ldots, C_7$, depending only on d, p, τ, m, M, such that for each $\frac{C_6}{\eta T} \leqslant \Omega \leqslant \frac{C_7}{h}$ and $\epsilon \leqslant C_5(\Omega h)^{2p-1}$ it holds that*

(1) for each of the cluster nodes, $x_j \in \mathbf{x}^c$, we have

$$C_2 (\Omega h)^{-2p+1} h\epsilon \leqslant diam(E_{\epsilon,\Omega}^{\mathbf{x},j}(F)) \leqslant C_1 (\Omega h)^{-2p+1} h\epsilon$$
$$C_4 (\Omega h)^{-2p+1} \epsilon \leqslant diam(E_{\epsilon,\Omega}^{\mathbf{a},j}(F)) \leqslant C_3 (\Omega h)^{-2p+1} \epsilon;$$

(2) for each of the non-cluster nodes, $x_j \in \mathbf{x} \backslash \mathbf{x}^c$ we have

$$C_2 \frac{\epsilon}{\Omega} \leqslant diam(E_{\epsilon,\Omega}^{\mathbf{x},j}(F)) \leqslant C_1 \frac{\epsilon}{\Omega}$$
$$C_4 \epsilon \leqslant diam(E_{\epsilon,\Omega}^{\mathbf{a},j}(F)) \leqslant C_3 \epsilon.$$

Put SRF $:= (\Omega h)^{-1}$. By Proposition 2.1 we immediately obtain the estimates of the corresponding minimax error.

Corollary 2.1. *Let $U = U(p, d, h, \tau, \eta, T, m, M)$, Ω as defined in Theorem 2.1 above and such that U has non-empty interior. Then for $\epsilon \lessapprox (\Omega h)^{2p-1}$*

$$\Lambda^{\mathbf{x},j}(\epsilon, U, \Omega) \asymp \begin{cases} SRF^{2p-1} h\epsilon & x_j \in \mathbf{x}^c, \\ \frac{\epsilon}{\Omega} & x_j \in \mathbf{x} \backslash \mathbf{x}^c, \end{cases}$$
$$\Lambda^{\mathbf{a},j}(\epsilon, U, \Omega) \asymp \begin{cases} SRF^{2p-1} \epsilon & x_j \in \mathbf{x}^c, \\ \epsilon & x_j \in \mathbf{x} \backslash \mathbf{x}^c. \end{cases}$$

Proof. The upper bounds are a direct consequence of the upper bounds in Theorem 2.1 and Proposition 2.1.

To show the lower bounds, denote $U_\epsilon := \left\{ F \in U : E_{\frac{\epsilon}{2},\Omega}(F) \subseteq U \right\}$. Provided that U_ϵ is not empty, we get the lower bounds by applying the lower bounds of Theorem 2.1 to any $F \in U_\epsilon$, and then using Proposition 2.1. To ensure that U_ϵ is non-empty, pick $F_0 \in int(U)$, and choose $\epsilon \leqslant C_5' (\Omega h)^{2p-1}$ such that $diam(E_{\frac{\epsilon}{2},\Omega}(F_0)) < \delta$ where δ is the distance from F_0 to the boundary of U (and so $F_0 \in U_\epsilon$). This in turn, can always be done using the upper bound of Theorem 2.1. □

Remark 2.1. *In this paper we focus on estimating the diameter of the error set $E_\epsilon(F)$ and its projections. 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.*

3. Numerical optimality of Matrix Pencil algorithm

The main theoretical results of this paper, Theorem 2.1 and Corollary 2.1, establish the best possible scalings for the SR problem with clustered nodes. In this section we provide some numerical evidence that a certain "high-resolution" algorithm, the Matrix Pencil (MP) method [31, 30], attains these performance bounds.

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

3.1. **The Matrix Pencil method.** The Matrix Pencil algorithm assumes the data sequence $\tilde{\mathbf{m}} = \{\tilde{m}_k\}_{k=0,\dots,N-1}$ to be of the form

$$\tilde{m}_k = \underbrace{\sum_{j=1}^{d} b_j \exp(\imath \phi_j k)}_{:= m_k} + n_k, \quad b_j \in \mathbb{C},\ \phi_j \in (-\pi, \pi], \tag{3.1}$$

Algorithm 3.1: The Matrix Pencil algorithm

Input : Model order d

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

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

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

1 Compute the matrices $A := \widetilde{H}^{\uparrow}, B := \widetilde{H}_{\downarrow}$;

2 Compute the truncated Singular Value Decompositions (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 $\{\tilde{\phi}_j\} := \{\angle \tilde{z}_j\},\ j = 1, \ldots, d$;

5 Compute $\tilde{b}_j$ by solving the linear least squares problem

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

where $\tilde{V} = \tilde{V}(\tilde{\mathbf{z}})$ is the Vandermonde matrix $\tilde{V} = \left[\tilde{z}_j^k\right]_{k=0,\ldots,N-1}^{j=1,\ldots,d}$;

6 **return** the estimated $\tilde{\phi}_j$ and $\tilde{b}_j$.

where n_k is the noise term. Let the noiseless Hankel matrix be

$$H := \begin{bmatrix} m_0 & m_1 & \ldots & m_{N-L-1} \\ m_1 & m_2 & \ldots & m_{N-L} \\ \vdots & \iddots & \iddots & \vdots \\ m_L & m_{L+1} & \ldots & \ldots m_{N-1} \end{bmatrix} \in \mathbb{C}^{(L+1)\times(N-L)}, \tag{3.2}$$

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 that the numbers $z_j := \exp(\imath\phi_j)$ are the d nonzero generalized eigenvalues (i.e. rank-reducing numbers) of the pencil $H_{\downarrow} - zH^{\uparrow}$. If we now construct the noisy matrices $A := \widetilde{H}^{\uparrow}, B := \widetilde{H}_{\downarrow}$ from the available data $\tilde{\mathbf{m}}$, we could apparently just solve the Generalized Eigenvalue Problem with A, B. However, if $L > d$ then the pencil $B - zA$ is close to being singular, and so an additional step of low-rank approximation is required. We summarize the MP method in Algorithm 3.1, and the interested reader is referred to the widely available literature on the subject (e.g. [31, 30, 39, 47], and 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. *Choice of signal.* In our experiments presented below, we fixed a signal $F = (\mathbf{a}, \mathbf{x}) \in \mathcal{P}_d$ whose node vector forms a (p, h, T, τ, η)-cluster of size p with

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

as follows:

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

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

(3) The amplitude vector was fixed to $\mathbf{a} = (1, 1, \ldots, 1) \in \mathbb{C}^d$.

3.2.2. *Choice of perturbation.* In order to test the MP algorithm in the worst-case scenario, we selected the perturbation sequence $\{n_k\}$ to be in accordance with the construction of Section 6, and, in particular, of Theorem 6.1. Given a signal F, a subset $C \subseteq \{1, \ldots, d\}$ and noise level ϵ, the worst-case input sequence $\{\tilde{m}_k\}$ is the sequence of Fourier coefficients of a particular signal $F_{\epsilon,C} = (\mathbf{a}', \mathbf{x}') \in \mathcal{P}_d$ constructed according to Algorithm 3.2:

$$\tilde{m}_k = \sum_{j=1}^{d} \mathbf{a}'_j \exp(\imath \mathbf{x}'_j k), \quad k = 0, \ldots, N-1.$$

Algorithm 3.2: The worst-case perturbation signal

Input : Signal $F = (\mathbf{a}, \mathbf{x}) \in \mathcal{P}_d$
Input : Noise level ϵ
Input : A subset of the nodes to be perturbed: $C \subseteq \{1, \ldots, d\}$
Output: The perturbed signal $F_{\epsilon,C}$

1 Construct the moment vector $\mathbf{g} = \left(\sum_{j \in C} \mathbf{a}_j \mathbf{x}_j^k\right)_{k=0,1,\ldots,2|C|-1} \in \mathbb{R}^{2|C|}$;
2 Construct the vector $\mathbf{g}'$ to be equal to $\mathbf{g}$ except the last entry: $\mathbf{g}'_k = \mathbf{g}_k$ for $k = 0, 1, \ldots, 2|C|-2$ and $\mathbf{g}'_{2|C|-1} = \mathbf{g}_{2|C|-1} + \epsilon$;
3 Solve the Prony problem of order $|C|$ with the data $\mathbf{g}'$ (for ϵ small enough, a unique solution always exists – see Proposition A.3 and [12]), obtaining a signal $F' = (\mathbf{a}', \mathbf{x}') \in \mathcal{P}_{|C|}$;
4 Put

$$F_{\epsilon,C}(x) = \sum_{j \notin C} \mathbf{a}_j \delta(x - \mathbf{x}_j) + \sum_{j=1}^{|C|} \mathbf{a}'_j \delta(x - \mathbf{x}'_j);$$

return the signal $F_{\epsilon,C}$.

3.3. **Results.**

3.3.1. *Noise threshold for successful recovery.* In the first set of experiments, we investigated the noise threshold $\epsilon \lessapprox \mathrm{SRF}^{2-2p}$ for successful recovery, as predicted by the theory. We have performed 5000 random experiments (the randomness was in the choice of h, Ω, ϵ) according to Algorithm 3.3, recording the success/failure result of each such experiment. The results for $d = 4$ and $p = 2, 3$ are presented in Figure 2, and the theoretical scaling above is confirmed.

Although not covered by our current theory, it is of interest to establish the recovery threshold for every node separately. In Figure 3 we can see that for a non-cluster node, the threshold is approximately constant (i.e. does not depend on the SRF.)

3.3.2. *Error amplification factors.* In the second set of experiments, we measured the actual blowup factors $\mathcal{K}_{\mathbf{x},j}, \mathcal{K}_{\mathbf{a},j}$ as in Algorithm 3.3 (recall also (1.4)). We have performed 1000 random experiments (again, choosing ϵ, Ω, h from a pre-defined numerical range). The results are presented in Figure 4. The scalings of Theorem 2.1, in particular the dependence on SRF, is confirmed.

Algorithm 3.3: A single experiment

Input : $p, d, h, \Omega \equiv N, \epsilon$
Input : Perturbation set $C \subseteq \{1, \ldots, d\}$

Construct the signal F according to Subsection 3.2.1 ;
Construct the sequence $\tilde{m}_k$, $k = 0, \ldots, N-1$ according to Subsection 3.2.2 ;
Compute actual perturbation magnitude

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

Execute the MP method (Algorithm 3.1) and obtain $F_{MP} = (\mathbf{a}^{MP}, \mathbf{x}^{MP})$;
for *each* j **do**
    compute the error for node j:

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

    The success for node j is defined as

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

    if $Succ_j$ == **true then**
        let $\ell(j) := \arg\min_\ell |\mathbf{x}_j^{MP} - \mathbf{x}_\ell|$;
        compute normalized node blowup

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

        compute normalized amplitude blowup

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

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

4. Normalization

Instead of considering the general signal $F = (\mathbf{a}, \mathbf{x}) \in \mathcal{P}_d$ with $T(\mathbf{x}) = T$, we shall assume that the node vector $\mathbf{x} = (x_1, \ldots, x_d)$ is normalized to the interval $\left[-\frac{1}{2}, \frac{1}{2}\right]$, so that $T(\mathbf{x}) = 1$ 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.

Define the shift transformation on $\mathcal{P}$.

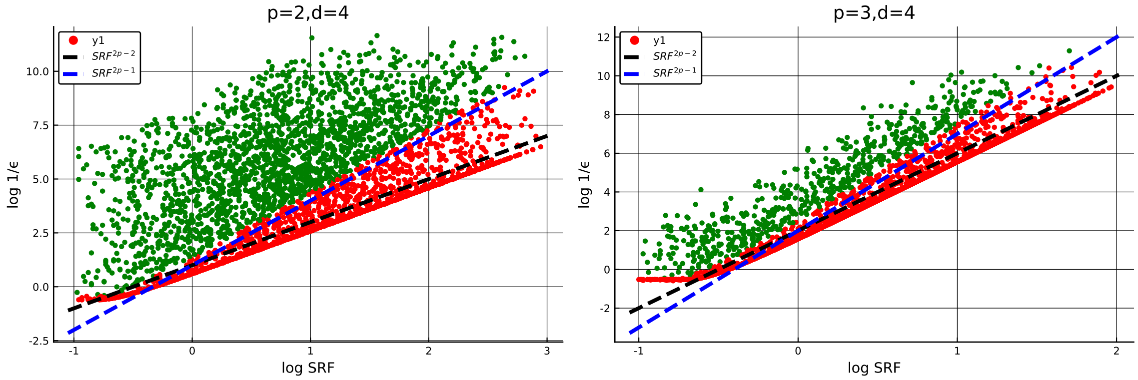
Definition 4.1. *We define $SH_\alpha : \mathcal{P}_d \to \mathcal{P}_d$. For $F = \sum_{j=1}^d a_j \delta(x - x_j) \in \mathcal{P}_d$ and $\alpha \in \mathbb{R}$,*

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

Define the scale transformation on $\mathcal{P}$.

Definition 4.2. *We define $SC_T : \mathcal{P}_d \to \mathcal{P}_d$. For $F = \sum_{j=1}^d a_j \delta(x - x_j) \in \mathcal{P}_d$ and $T > 0$,*

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



(A) $p = 2$ (B) $p = 3$

Figure 2. Phase transition for successful recovery, $d = 4$ and worst-case perturbations of the first p nodes. The recovery is successful (green dot) if Algorithm 3.3 returned success for each node. The relationship $\epsilon_{crit} \approx \mathrm{SRF}^{1-2p}$ for the critical value of ϵ can clearly be seen.

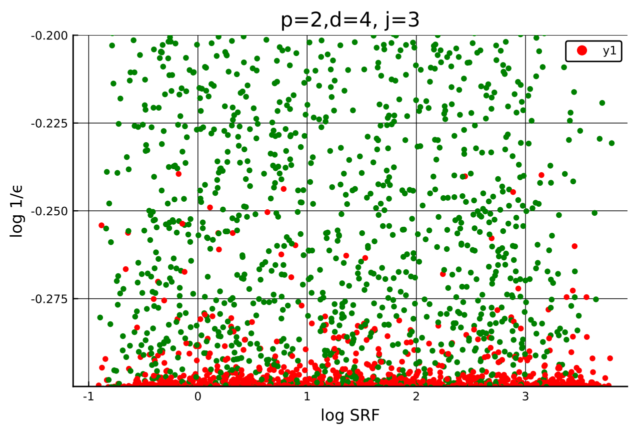


Figure 3. Phase transition for successful recovery of a non-cluster node. Comparing with Figure 2, the threshold is approximately constant $\epsilon_{crit} \approx const$. Here $p = 2, d = 4$, plotted is the successful recovery of node $j = 3$.

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

$$SH_\alpha(E_\epsilon(F)) = E_\epsilon(SH_\alpha(F)). \tag{4.1}$$

Therefore we can assume, without loss of generality, that the nodes of the signal F are centered at the origin, i.e. $x_d + x_1 = 0$.

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

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

Thus we have the following.

Proposition 4.1. *Suppose that $F = (\mathbf{a}, \mathbf{x}) \in \mathcal{P}_d$ such that $T(\mathbf{x}) = T$. Then for any $\epsilon > 0$ and $1 \leqslant j \leqslant d$ we have*

$$diam(E^{\mathbf{x},j}_{\epsilon,\Omega}(F)) = T diam\left(E^{\mathbf{x},j}_{\epsilon,\Omega T}\left(SC_T(F)\right)\right) \tag{4.3}$$

$$diam(E^{\mathbf{a},j}_{\epsilon,\Omega}(F)) = diam\left(E^{\mathbf{a},j}_{\epsilon,\Omega T}\left(SC_T(F)\right)\right) \tag{4.4}$$

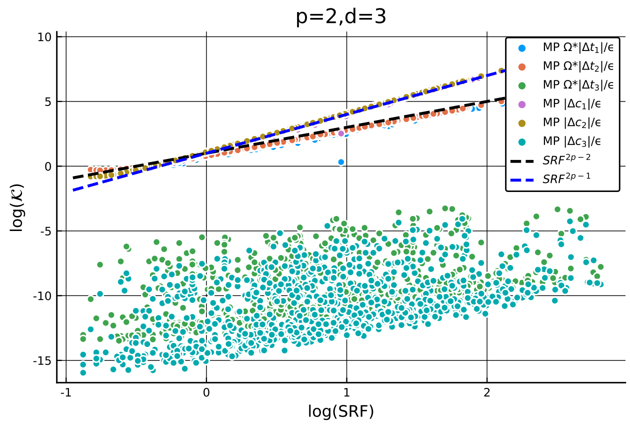


FIGURE 4. The error amplification factors. The signal F has $p = 2, d = 3$. For cluster nodes $j = 1, 2$ the node blowup scales like SRF^{2p-2}, while the amplitude blowup scales like SRF^{2p-1}. For the non-cluster node $j = 3$, both blowup factors are bounded by a constant.

5. Upper bound on $diam\big(E_\epsilon(F)\big)$

5.1. **Overview of the proof.** The proof of the upper bounds in Theorem 2.1, 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 $E_\epsilon(F) \subset \mathcal{P}_d$, is the inverse image of an (infinite-dimensional) ϵ-cube in the data space under the Fourier transform mapping $\mathcal{F}$ (recall (1.2) and Definition 2.1). 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 $FM_\lambda : \mathcal{P}_d \to \mathbb{C}^{2d}$, 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 $E_\epsilon(F)$ 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 2.1 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 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 5.1 (proved in Appendix C). These bounds essentially depend on the "effective separation" of each node in $\mathbf{x}$ from its neighbours, 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 exist 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 on 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 5.2.

(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 5.5, using the previously established general estimates from Proposition 5.1 and the quantitative inverse function theorem (Theorem B.1).

(4) Next, denoting $A = A_{R,\lambda} := FM_\lambda^{-1}(V)$, we show in Proposition 5.6 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 5.7 (proved in Appendix F), establishes that for a certain $\lambda^* \in \Lambda(\mathbf{x})$ and all possible permutations π and shifts $\ell \in \mathbb{Z}\backslash\{0\}$, 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* $E_{R,(\bar{\lambda})}$ is empty. From this fact it immediately follows that the original error set $E_\epsilon(F)$ with $\epsilon = R$ is contained in A_{R,λ^*} (Proposition 5.8). 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 2.1.*

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 $\bar{\mathcal{P}}_d$,

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

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 $FM_\lambda : \bar{\mathcal{P}}_d \cong \mathbb{C}^{2d} \to \mathbb{C}^{2d}$ by*

$$FM_\lambda((\mathbf{a}, \mathbf{x})) = \boldsymbol{\mu} = (\mu_0, \ldots, \mu_{2d-1}), \ \mu_k := \sum_{j=1}^d \mathbf{a}_j e^{2\pi i \mathbf{x}_j \lambda k}, \ k = 0, \ldots, 2d-1.$$

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*

$$\left\| FM_\lambda(F') - FM_\lambda(F) \right\| \leqslant \epsilon.$$

Similarly we denote by $E^{\mathbf{a},j}_{\epsilon,(\lambda)}(F)$, $E^{\mathbf{x},j}_{\epsilon,(\lambda)}(F)$ 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 $\lambda \leqslant \frac{\Omega}{2d-1}$ we have that $E_{\epsilon,\Omega}(F) \subseteq E_{\epsilon,(\lambda)}(F)$ giving

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

Hence, to prove the upper bound in Theorem 2.1, we shall show that there exists a *certain* subset $S \subseteq \left(0, \frac{\Omega}{2d-1}\right]$ such that for each $\lambda \in S$, $diam\left(E_{\epsilon,(\lambda)}(F)\right)$ 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 5.1 (Uniform Jacobian bounds)**.** *Let $F = (\mathbf{a}, \mathbf{x}) \in \bar{\mathcal{P}}_d$, $\mathbf{a} = (a_1, \ldots, a_d)$, $\mathbf{x} = (x_1, \ldots, x_d)$ and for $\lambda > 0$ let $z_1 = e^{2\pi i \lambda x_1}, \ldots, z_d = e^{2\pi i \lambda x_d}$. Suppose that for each $j = 1, \ldots, d$, we have $0 < \frac{m}{2} \leqslant |a_j|$ and $\frac{1}{2} \leqslant |z_j| \leqslant 2$ for some $m > 0$.*

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

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

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

$$\sum_{k=1}^{2d} |A_{j,k}| \leqslant K_1(\tilde{\eta}, d, p), \qquad x_j \in \mathbf{x} \backslash \mathbf{x}^c, \tag{5.2}$$

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

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

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

K_1, K_2, K_3, K_4 are constants with respect to $\lambda, \tilde{h}$, which are specified in the proof.

The proof of Proposition 5.1 is given in Appendix C.

5.3. **The existence of an admissible decimation.** In this section we shall prove the existence of a certain sub-interval $I \subset \left[\frac{1}{2}\frac{\Omega}{2d-1}, \frac{\Omega}{2d-1}\right]$, such that for every $\lambda \in I$, the corresponding inverse λ-decimation map FM_λ^{-1} has the best (smallest) possible norm (see Proposition 5.1 above), with respect to Ω, h.

Definition 5.3. *For each $x \in \mathbb{R}$ and $a > 0$ consider the operation $\mod \left(-\frac{a}{2}, \frac{a}{2}\right]$ defined as*

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

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

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

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

Definition 5.4. *For $\alpha, \beta \in \mathbb{C} \backslash \{0\}$, we define the angular distance between α, β as*

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

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

Lemma 5.1. *For $|x| = |y| = 1$, we have*

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

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 $|\theta| \leqslant \frac{\pi}{2}$ we have

$$\frac{2}{\pi}|\theta| \leqslant \sin|\theta| \leqslant |\theta|. \quad \square$$

Let $F = (\mathbf{a}, \mathbf{x}) \in \mathcal{P}_d$ such that the node vector $\mathbf{x} = (x_1, \ldots, x_d)$ forms a (p, h, T, τ, η)-cluster with $\mathbf{x}^c = \{x_\kappa, x_{\kappa+1}, \ldots, x_{\kappa+p-1}\}$. According to Proposition 5.1, the Jacobian norm of FM_λ^{-1} essentially depends on the "effective separation", or the minimal distance, between the "mapped" nodes $z_j(\lambda) = e^{2\pi i \lambda x_j}$. After a blowup by a factor of $\lambda \leqslant \frac{1}{2h}$, the pairwise angular distances $\angle(\cdot,\cdot)$ (and hence the euclidean distances) between the mapped cluster-nodes $z_\kappa, \ldots, z_{\kappa+p-1}$ 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 $z_1, \ldots, z_{\kappa-1}, z_{\kappa+p}, \ldots, z_d$ 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.

Proposition 5.2. *Let $F = (\mathbf{a}, \mathbf{x}) \in \mathcal{P}_d$, $\mathbf{x} = (x_1, \ldots, x_d) \subset [-\frac{1}{2}, \frac{1}{2}]$, such that the nodes of $\mathbf{x}$ form a $(p, h, 1, \tau, \eta)$-cluster with $\mathbf{x}^c = \{x_\kappa, x_{\kappa+1}, \ldots, x_{\kappa+p-1}\}$.*

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

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 $|I'| \geqslant (2d^2\eta)^{-1}$ such that for each $\lambda \in I'$:

(1) For all $x_\ell \in \mathbf{x} \backslash \mathbf{x}^c$ and $x_j \in \mathbf{x}$, $x_j \neq x_\ell$,

$$\angle(z_\ell(\lambda), z_j(\lambda)) \geqslant \frac{1}{d^2}. \tag{5.7}$$

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

$$\angle(z_j(\lambda), z_k(\lambda)) \geqslant 2\pi\lambda\tau h \geqslant \frac{\pi\tau}{2d-1}\Omega h. \tag{5.8}$$

Proof. Let us first prove that assertion (5.8) holds for any $\frac{1}{2}\frac{\Omega}{2d-1} \leqslant \lambda \leqslant \frac{\Omega}{2d-1}$.

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

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

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

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

By assumption $\lambda \geqslant \frac{1}{2}\frac{\Omega}{2d-1}$. Then, $\angle(z_j, z_k) \geqslant \frac{\pi\tau}{2d-1}\Omega h$. This concludes the proof of assertion (5.8).

In order to prove assertion (5.7), we shall require the following key estimate concerning the pairwise angular distance of any two nodes x_j, x_k.

Lemma 5.2 (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 blow-ups $z_j = z_j(\lambda) = e^{2\pi i\lambda x_j}, z_k = z_k(\lambda) = e^{2\pi i\lambda x_k}$. Then for $0 \leqslant \alpha \leqslant \pi$ and an interval $I = [a, b] \subset \mathbb{R}$, the set*

$$\Sigma^{\alpha}_{j,k}(I) = \left\{\lambda \in I : \angle\big(z_j(\lambda), z_k(\lambda)\big) \leqslant \alpha\right\} \tag{5.9}$$

is a union of N intervals $I_1, \ldots, I_N$ with $\left\lfloor \frac{|I|}{\Delta} \right\rfloor \leqslant N \leqslant \left\lceil \frac{|I|}{\Delta} \right\rceil + 1$, and

$$|I_j| \leqslant \frac{\alpha}{\pi}\frac{1}{\Delta}, \qquad j = 1, \ldots, N.$$

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

$$\angle(z_j(\lambda), z_k(\lambda)) = \left|\mathrm{Arg}\left(\frac{z_j(\lambda)}{z_k(\lambda)}\right)\right| = \left|\mathrm{Arg}(e^{2\pi i\lambda\Delta})\right|. \tag{5.10}$$

By equation (5.10) we have

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

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

$$S^{\alpha} = \left\{\lambda \in \mathbb{R} : -\frac{\alpha}{2\pi}\frac{1}{\Delta} \leqslant \left(\lambda \mod \left(-\frac{1}{2\Delta}, \frac{1}{2\Delta}\right]\right) \leqslant \frac{\alpha}{2\pi}\frac{1}{\Delta}\right\}. \tag{5.11}$$

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 $\left\lfloor \frac{|I|}{\Delta} \right\rfloor \leqslant N \leqslant \left\lceil \frac{|I|}{\Delta} \right\rceil + 1$ intervals of length smaller or equal to $\frac{\alpha}{\pi}\frac{1}{\Delta}$. This concludes the proof of Lemma 5.2. □

Using Lemma 5.2 we now prove that assertion (5.7) holds for any interval $I = [a, b] \subset \mathbb{R}$ of length $|I| = \frac{1}{\eta}$. Let I be such an interval. For each $0 < \alpha \leqslant \pi$ consider the set

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

We then have

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

where $\Sigma^{\alpha}_{\ell,j}$ are given by (5.9). By Lemma 5.2 each $\Sigma^{\alpha}_{\ell,j}(I)$ above is a union of at most $\left\lceil \frac{|I|}{\eta} \right\rceil + 1 = 2$ 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

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

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

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

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

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

By (5.13)

$$\nu\big((\Sigma^{\alpha'}(I))^c\big) \geqslant |I| - \frac{1}{2\eta} = \frac{1}{\eta} - \frac{1}{2\eta} = \frac{1}{2\eta}. \tag{5.14}$$

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

$$L = K + 1 = d(d-1) + 1 \leqslant d^2 \tag{5.15}$$

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

$$\frac{\nu\big((\Sigma^{\alpha'}(I))^c\big)}{L} \geqslant \frac{1}{d^2}\frac{1}{2\eta}.$$

We therefore conclude that $\big(\Sigma^{\alpha'}(I)\big)^c$ contains an interval of length greater or equal to $\frac{1}{d^2}\frac{1}{2\eta}$. This proves assertion (5.7) of Proposition 5.2. $\square$

5.4. **Error sets of admissible decimation maps.** Throughout this Section we fix a signal $F = (\mathbf{a}, \mathbf{x}) \in \mathcal{P}_d$, $\mathbf{a} = (a_1, \ldots, a_d)$, $\mathbf{x} = (x_1, \ldots, x_d) \subset \left[-\frac{1}{2}, \frac{1}{2}\right]$, such that the nodes of $\mathbf{x}$ form a $(p, h, 1, \tau, \eta)$-cluster with $\mathbf{x}^c = \{x_\kappa, x_{\kappa+1}, \ldots, x_{\kappa+p-1}\}$ and $\|\mathbf{a}\| \geqslant m > 0$. We also fix $\Omega > 0$ such that $\Omega h \leqslant \frac{1}{20d}$.

Proposition 5.2 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 5.5 (Admissible blowup factors)**.** *For each $F = (\mathbf{a}, \mathbf{x}) \in \mathcal{P}_d$, $\mathbf{x} = (x_1, \ldots, x_d)$, such that the nodes of $\mathbf{x}$ form a $(p, h, 1, \tau, \eta)$-cluster $\mathbf{x}^c$ and $z_j = z_j(\lambda) = e^{2\pi i \lambda x_j}$, $j = 1, \ldots, d$ and $\Omega > 0$, we define the set of admissible blowup factors $\Lambda(\mathbf{x}) = \Lambda_{\Omega,d}(\mathbf{x})$ as the set of all $\lambda \in \left[\frac{1}{2}\frac{\Omega}{2d-1}, \frac{\Omega}{2d-1}\right]$ satisfying:*

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

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

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

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

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

$$H_{\alpha,\beta}(F) = \left\{(\mathbf{a}', \mathbf{x}') \in \bar{\mathcal{P}}_d \ : \|\mathbf{a}' - \mathbf{a}\| \leqslant \alpha, \ \|\mathbf{x}' - \mathbf{x}\| \leqslant \beta\right\},$$

and by $H^{\mathrm{o}}_{\alpha,\beta}(F)$ the interior of $H_{\alpha,\beta}(F)$.

The following is proved in Appendix D.

Proposition 5.3 (One-to-one)**.** *For each $\lambda \in \Lambda(\mathbf{x})$ and $\Omega h < \frac{1}{20d}$, the map FM_λ is injective in the open polydisc $U = H^{\mathrm{o}}_{m,\frac{\tau h}{2\pi}}(F) \subset \bar{\mathcal{P}}_d$.*

Next we can estimate the Lipschitz constants of the inverse map FM_λ^{-1}, using the previously established general bounds in Proposition 5.1.

Proposition 5.4. *Let $H = H_{\frac{m}{2},\frac{\tau h}{4\pi}}(F) \subset U = H^{\mathrm{o}}_{m,\frac{\tau h}{2\pi}}(F)$. Then for each $F' \in H$, the inverse Jacobian $J_\lambda^{-1}(F') = \begin{bmatrix} A \\ \tilde{B} \end{bmatrix}$, $A, \tilde{B}$ are $d \times 2d$, satisfies:*

$$\sum_{k=1}^{2d} |A_{j,k}| \leqslant \tilde{C}, \qquad x_j \in \mathbf{x}\backslash\mathbf{x}^c, \tag{5.18}$$

$$\sum_{k=1}^{2d} |\tilde{B}_{j,k}| \leqslant \tilde{C}\frac{1}{\Omega}, \qquad x_j \in \mathbf{x}\backslash\mathbf{x}^c, \tag{5.19}$$

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

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

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

Proof. Let $F' = (\mathbf{a}', \mathbf{x}') \in H$, $\mathbf{a}' = (a_1', \ldots, a_d')$, $\mathbf{x}' = (x_1', \ldots, x_d')$. Let $z_j' = z_j'(\lambda) = e^{2\pi i \lambda x_j'}$ and let $z_j = z_j(\lambda) = e^{2\pi i \lambda x_j}$, $j = 1, \ldots, d$.

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

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

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

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

Then by (5.6)

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

We get that

$$|z_\ell' - z_j'| \geqslant |z_\ell - z_j| - |z_\ell' - z_\ell| - |z_j' - z_j| \geqslant |z_\ell - z_j| - 2\lambda\tau h \geqslant \frac{2}{\pi d^2} - 2\lambda\tau h.$$

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

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

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

$$|z'_\ell - z'_j| \geqslant \frac{1}{2d^2}. \tag{5.22}$$

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

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

Then by (5.6)

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

With a similar argument as above, we get that

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

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

$$|z'_j - z'_k| \geqslant 2\lambda\tau h \geqslant \frac{\tau}{2d-1}\Omega h. \tag{5.23}$$

Now using (5.22) and (5.23) we invoke Proposition 5.1 with $\tilde{h} = \frac{\tau}{2d-1}\Omega h$ and $\tilde{\eta} = \frac{1}{2d^2}$ and as a result prove Proposition 5.4 with

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

□

Proposition 5.5. *Let $U = H^{\circ}_{m,\frac{\tau h}{2\pi}}(F)$ and $H = H_{\frac{m}{2},\frac{\tau h}{4\pi}}(F) \subset U$. Let $\lambda \in \Lambda(\mathbf{x})$ and let $\boldsymbol{\mu} = FM_\lambda(F)$, then there exists a constant $C_5 = C_5(m,d,p,\tau)$ such that for $R = C_5(\Omega h)^{2p-1}$,*

$$FM_\lambda(H) \supseteq Q_R(\boldsymbol{\mu}).$$

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

$$FM_\lambda^{-1} : V \to U$$

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 $1 \leqslant j \leqslant d$, let $P_{\mathbf{a},j}, P_{\mathbf{x},j} : \bar{\mathcal{P}}_d \to \mathbb{C}$ be the projections onto the j^{th} amplitude and the j^{th} node coordinates respectively. Then FM_λ^{-1} is Lipschitz on $Q_R(\boldsymbol{\mu})$ with the following bounds:

$$\left|P_{\mathbf{a},j}FM_\lambda^{-1}(\boldsymbol{\mu}') - P_{\mathbf{a},j}FM_\lambda^{-1}(\boldsymbol{\mu}'')\right| \leqslant \tilde{C}_1\|\boldsymbol{\mu}'' - \boldsymbol{\mu}'\| \times \begin{cases}1 & x_j \in \mathbf{x}\backslash\mathbf{x}^c\\ (\Omega h)^{-2p+1} & x_j \in \mathbf{x}^c\end{cases},$$

$$\left|P_{\mathbf{x},j}FM_\lambda^{-1}(\boldsymbol{\mu}') - P_{\mathbf{x},j}FM_\lambda^{-1}(\boldsymbol{\mu}'')\right| \leqslant \tilde{C}_3\frac{1}{\Omega}\|\boldsymbol{\mu}'' - \boldsymbol{\mu}'\| \times \begin{cases}1 & x_j \in \mathbf{x}\backslash\mathbf{x}^c\\ (\Omega h)^{-2p+2} & x_j \in \mathbf{x}^c\end{cases},$$

for each $\boldsymbol{\mu}'', \boldsymbol{\mu}' \in Q_R(\boldsymbol{\mu})$, where $\tilde{C}_1 = \tilde{C}_1(m,d,p,\tau)$, $\tilde{C}_3 = \tilde{C}_3(m,d,p,\tau)$ are constants depending only on d,m,p,τ and $\tilde{C}_3 C_5 \leqslant 1$.

Proof. By Proposition 5.3 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 5.4 apply. Finally one can verify (using a similar argument as in the proof of Proposition 5.4) 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), (5.18), (5.21), and conclude that Proposition 5.5 holds with $\tilde{C}_1 = \tilde{C}_3 = \tilde{C}$ and $C_5 = \min\left(\frac{m}{2\tilde{C}}, \frac{\tau}{4\pi\tilde{C}}\right)$. $\square$

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 $\epsilon \leqslant R$ where $R = C_5(\Omega h)^{2p-1}$ is as specified in Proposition 5.5.

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

$$A_{\epsilon,\lambda}(F) = FM_\lambda^{-1}(Q_\epsilon(\boldsymbol{\mu})) \bigcap \mathcal{P}_d, \tag{5.24}$$

where $FM_\lambda^{-1} : V_\lambda \to U$ 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 5.1. *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 $F'' = (a, \mathbf{x}'')$ 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

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

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 shifts operations.

We denote the set of permutations of d elements by

$$\Pi = \Pi_d \subset \{\pi : \{1, \ldots, d\} \to \{1, \ldots, d\}\}.$$

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

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

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 5.6. *For each* $\lambda \in \Lambda(\mathbf{x})$ *and* $\epsilon \leqslant R$

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

5.5. **Proof of the upper bound.** Let $F = (\mathbf{a}, \mathbf{x}) \in \mathcal{P}_d$, $\mathbf{a} = (a_1, \ldots, a_d)$, $\mathbf{x} = (x_1, \ldots, x_d) \subset \left[-\frac{1}{2}, \frac{1}{2}\right]$, such that the nodes in $\mathbf{x}$ form a $(p, h, 1, \tau, \eta)$-cluster with $\mathbf{x}^c = \{x_\kappa, x_{\kappa+1}, \ldots, x_{\kappa+p-1}\}$, and also $\|\mathbf{a}\| \geqslant m > 0$.

Consider the set of the admissible blowup factors $\Lambda(\mathbf{x})$ (see Definition 5.5). By the analysis of Section 5.4, under the assumption that $\Omega h \leqslant \frac{1}{20d}$, the following assertions hold:

1. By Proposition 5.3 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 5.5 there exists a constant $C_5 = C_5(m, d, p, \tau)$ such for each $\lambda \in \Lambda(\mathbf{x})$, $V_\lambda = FM_\lambda(U)$ contains a cube $Q_R(\boldsymbol{\mu}_\lambda)$, where $\boldsymbol{\mu}_\lambda = FM_\lambda(F)$ and $R = C_5(\Omega h)^{2p-1}$.

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

$$A_{R,\lambda}(F) = FM_\lambda^{-1}(Q_R(\boldsymbol{\mu}_\lambda)) \bigcap \mathcal{P}_d.$$

The following intermediate claim is proved in Appendix F.

Proposition 5.7. *There exist positive constants* K_9, K_{10} *depending only on* d *such that for* $\frac{K_9}{\eta} \leqslant \Omega \leqslant \frac{K_{10}}{h}$ *the following holds. There exists* $\lambda \in \Lambda(\mathbf{x})$ *such that for each pair* $(\pi, \boldsymbol{\ell}) \in \Pi_d \times \left(\mathbb{Z}^d \backslash \{\mathbf{0}\}\right)$, *there exists* $\lambda_{\pi,\boldsymbol{\ell}} \in \Lambda(\mathbf{x})$ *for which*

$$\left(A^{\pi}_{R,\lambda}(F) + \frac{1}{\lambda}\boldsymbol{\ell} \right) \bigcap E_{R,(\lambda_{\pi,\boldsymbol{\ell}})}(F) = \varnothing. \tag{5.25}$$

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

Proposition 5.8. *There exists* $\lambda \in \Lambda(\mathbf{x})$ *such that*

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

Proof. Using Proposition 5.7 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 5.6:

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

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

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

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

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

By (5.27)

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

Then by (5.29) and (5.30)

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

For each pair $(\pi,\boldsymbol{\ell}) \in \Pi_d \times \left(\mathbb{Z}^d \backslash \{\mathbf{0}\}\right)$, let $\lambda_{\pi,\boldsymbol{\ell}} \in \Lambda(\mathbf{x})$ be the value asserted by Proposition 5.7, i.e. satisfying (5.25) for $\lambda = \lambda^*$. By this and by (5.31) we have

$$\begin{aligned} E_{R,\Omega}(F) &\subset \left(\bigcup_{\pi\in\Pi_d} A^{\pi}_{R,\lambda^*}(F) \right) \bigcup \left(\bigcup_{(\pi,\boldsymbol{\ell})\in\Pi_d\times(\mathbb{Z}^d\backslash\{\mathbf{0}\})} \left\{ \left(A^{\pi}_{R,\lambda^*}(F) + \frac{1}{\lambda^*}\boldsymbol{\ell} \right) \bigcap E_{R,(\lambda_{\pi,\boldsymbol{\ell}})}(F) \right\} \right) \\ &= \bigcup_{\pi\in\Pi_d} A^{\pi}_{R,\lambda^*}(F). \end{aligned} \tag{5.32}$$

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 $\lambda = \lambda^*$. ☐

We have everything in place to estimate the set $E_\epsilon(F)$ and its projections.

Proposition 5.9. *Let $F = (\mathbf{a},\mathbf{x}) \in \mathcal{P}_d$, $\mathbf{a} = (a_1,\ldots,a_d)$, $\mathbf{x} = (x_1,\ldots,x_d) \subset \left[-\frac{1}{2},\frac{1}{2}\right]$, such that the nodes in $\mathbf{x}$ form a $(p,h,1,\tau,\eta)$-cluster with $\mathbf{x}^c = \{x_\kappa, x_{\kappa+1},\ldots,x_{\kappa+p-1}\}$, and also $\|\mathbf{a}\| \geqslant m > 0$. Then there exist constants C_1, C_3, C_5, C_6, C_7, depending only on d,p,τ,m such that for each $\frac{C_6}{\eta} \leqslant \Omega \leqslant \frac{C_7}{h}$ and $\epsilon \leqslant C_5(\Omega h)^{2p-1}$ it holds that*

$$diam(E^{\mathbf{x},j}_{\epsilon,\Omega}(F)) \leqslant \begin{cases} C_1\frac{1}{\Omega}(\Omega h)^{-2p+2}\epsilon, & x_j \in \mathbf{x}^c \\ C_1\frac{1}{\Omega}\epsilon, & x_j \in \mathbf{x}\backslash\mathbf{x}^c, \end{cases}$$

$$diam(E^{\mathbf{a},j}_{\epsilon,\Omega}(F)) \leqslant \begin{cases} C_3(\Omega h)^{-2p+1}\epsilon, & x_j \in \mathbf{x}^c, \\ C_3\epsilon, & x_j \in \mathbf{x}\backslash\mathbf{x}^c. \end{cases}$$

Proof. Let $\epsilon \leqslant C_5(\Omega h)^{2p-1} = R$ and let $F' \in E_{\epsilon,\Omega}(F)$ with $F' = (\mathbf{a}',\mathbf{x}')$. Put $\boldsymbol{\mu}^* = FM_{\lambda^*}(F)$ where λ^* is given by Proposition 5.8. Thus, by (5.26)

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

Further denote $\boldsymbol{\mu}^+ = FM_{\lambda^*}(F')$. Furthermore, by Proposition 5.5, applied with $\boldsymbol{\mu} = \boldsymbol{\mu}' = \boldsymbol{\mu}^*$ and $\boldsymbol{\mu}'' = \boldsymbol{\mu}^+$, there exist constants $\tilde{C}_1 = \tilde{C}_1(m,d,p,\tau)$, $\tilde{C}_3 = \tilde{C}_3(m,d,p,\tau)$ such that

$$|\mathbf{a}_j - \mathbf{a}'_j| = \|P_{\mathbf{a},j}FM^{-1}_{\lambda^*}(\boldsymbol{\mu}^*) - P_{\mathbf{a},j}FM^{-1}_{\lambda^*}(\boldsymbol{\mu}^+)\| \leqslant \begin{cases} \tilde{C}_3(\Omega h)^{-2p+1}\epsilon, & x_j \in \mathbf{x}^c \\ \tilde{C}_3\epsilon, & x_j \in \mathbf{x}\backslash\mathbf{x}^c. \end{cases}$$

$$|\mathbf{x}_j - \mathbf{x}'_j| = \|P_{\mathbf{x},j}FM^{-1}_{\lambda^*}(\boldsymbol{\mu}^*) - P_{\mathbf{x},j}FM^{-1}_{\lambda^*}(\boldsymbol{\mu}^+)\| \leqslant \begin{cases} \tilde{C}_1\frac{1}{\Omega}(\Omega h)^{-2p+2}\epsilon, & x_j \in \mathbf{x}^c, \\ \tilde{C}_1\frac{1}{\Omega}\epsilon, & x_j \in \mathbf{x}\backslash\mathbf{x}^c. \end{cases}$$

Since F' was an arbitrary signal in $E_{\epsilon,\Omega}(F)$, we repeat the argument with $F'' \in E_{\epsilon,\Omega}(F)$ and finish the proof of Proposition 5.9 with $C_1 = 2\tilde{C}_1$, $C_3 = 2\tilde{C}_3$, C_5 which was defined above, $C_6 = K_9$ and $C_7 = \min\left(K_{10}, \frac{1}{20d}\right)$ where K_9, K_{10} are the constants in Proposition 5.7. ☐

Now we are ready to prove the upper bound of Theorem 2.1. Most of the work was already done in the course of the proof of Proposition 5.9. What is left is to apply the scale and shift properties of the Fourier transform.

Proof of Theorem 2.1, upper bound. Let $F = (\mathbf{a}, \mathbf{x}) \in \mathcal{P}_d$, $\mathbf{a} = (a_1, \ldots, a_d)$, $\mathbf{x} = (x_1, \ldots, x_d)$, such that the nodes of $\mathbf{x}$ form a (p, h, T, τ, η)-cluster with $\mathbf{x}^c = \{x_\kappa, x_{\kappa+1}, \ldots, x_{\kappa+p-1}\}$ and $\|\mathbf{a}\| \geqslant m > 0$. Let $\frac{C_6}{\eta T} \leqslant \Omega \leqslant \frac{C_7}{h}$ where $C_6 = C_6(d, p, \tau, m), C_7 = C_7(d, p, \tau, m)$ are the constants specified in Proposition 5.9.

The signal $SC_T(F) = (\mathbf{a}, \tilde{\mathbf{x}})$, $\tilde{\mathbf{x}} = (\tilde{x}_1, \ldots, \tilde{x}_d)$, $\tilde{x}_1 = \frac{x_1}{T}, \ldots, \tilde{x}_d = \frac{x_d}{T}$, is "normalized" such that $\tilde{x}_1, \ldots, \tilde{x}_d \in [-\frac{1}{2}, \frac{1}{2}]$ and $T(\mathbf{x}) = 1$. The node vector $\tilde{\mathbf{x}}$ forms a $(p, \frac{h}{T}, 1, \tau, \eta)$-cluster. Applying Proposition 5.9 for $\tilde{F} = SC_T(F)$, $\tilde{h} = \frac{h}{T}$, $\tilde{\Omega} = \Omega T \geqslant \frac{C_6}{\eta}$ and $\tilde{\Omega}\tilde{h} = \Omega h \leqslant C_7$, we conclude that there exist constants C_1, C_3, C_5 such that for any $\epsilon \leqslant C_5(\Omega h)^{2p-1}$

$$diam\big(E^{\mathbf{a},j}_{\epsilon,\Omega T}(SC_T(F))\big) \leqslant \begin{cases} C_3(\Omega h)^{-2p+1}\epsilon, & x_j \in \mathbf{x}^c \\ C_3\epsilon, & x_j \in \mathbf{x}\backslash\mathbf{x}^c, \end{cases}$$

$$diam\big(E^{\mathbf{x},j}_{\epsilon,\Omega T}(SC_T(F))\big) \leqslant \begin{cases} C_1\frac{1}{\Omega T}(\Omega h)^{-2p+2}\epsilon, & x_j \in \mathbf{x}^c, \\ C_1\frac{1}{\Omega T}\epsilon, & x_j \in \mathbf{x}\backslash\mathbf{x}^c. \end{cases}$$

Applying Proposition 4.1 we conclude the proof of the upper bounds of Theorem 2.1. □

6. Lower bound on $diam\big(E_\epsilon(F)\big)$

6.1. **Non-cluster nodes.** Let us first consider the easy case of a non-cluster node, $x_j \in \mathbf{x}\backslash\mathbf{x}^c$. Write the whole signal $F = (\mathbf{a}, \mathbf{x}) \in \mathcal{P}_d$, where $\|\mathbf{a}\| \leqslant M$, as

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

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 $G(x) = a_j'\delta(x - x_j') + F^o(x)$. For $|s| \leqslant \Omega$, the difference between the Fourier transforms of F and G satisfies

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

and therefore the proof is finished.

6.2. **Cluster nodes.** The lower bound for off-grid clustered super-resolution has been essentially proven in our earlier publication [1], however in a slightly weaker formulation than that of Theorem 2.1. Here we prove the stronger version, combining the method of [1] with recent results from [2] on the stability of the inverse algebraic moment problem for spike trains (see Appendix A). In this section all the constants $C_1, \ldots, K_1, \ldots$ are unrelated to those of the previous section.

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

Theorem 6.1. *Given the parameters $h > 0$, $0 < \tau \leqslant 1$, $0 < m \leqslant M < \infty$, let the signal $F = (\mathbf{a}, \mathbf{x}) \in \mathcal{P}_d$ form a single uniform cluster as follows:*

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

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

- *$m \leqslant |a_j| \leqslant M$, for $j = 1, \ldots, d$.*

Then there exist constants C_1, K_1, K_2, K_3, K_4 depending only on (d, τ, m, M) such that for every $\epsilon < C_1 h^{2d-1}$, 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* $k = 0, 1, \ldots, 2d-2$*, where* m_k *are given by* (A.1)*;*
(2) $m_{2d-1}(F_\epsilon) = m_{2d-1}(F) + \epsilon$*;*
(3) $K_1 h^{-2d+2}\epsilon \leqslant \max_{1\leqslant j\leqslant d} |\mathbf{x}_j - \mathbf{y}_j| \leqslant K_2 h^{-2d+2}\epsilon$*;*
(4) $K_3 h^{-2d+1}\epsilon \leqslant \max_{1\leqslant j\leqslant d} |\mathbf{b}_j - \mathbf{a}_j| \leqslant K_4 h^{-2d+1}\epsilon$*.*

Now let $F = (\mathbf{a}, \mathbf{x}) \in \mathcal{P}_d$ with $\mathbf{x} = (x_1, \ldots, x_d)$ and $\mathbf{a} = (a_1, \ldots, a_d)$ form an $(p, h, 1, \tau, \eta)$ cluster with cluster nodes $\mathbf{x}^c = (x_\kappa, \ldots, x_{\kappa+p-1})$ (according to Definition 2.2), such that $m \leqslant |a_j| \leqslant M$. 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. $x_\kappa + x_{\kappa+p-1} = 0$. Next, define a "blowup" of F^c by Ω as follows:

$$F^c_{(\Omega)} := SC_{\frac{1}{\Omega}}(F^c) = \sum_{x_j \in \mathbf{x}^c} a_j \delta(x - \Omega x_j). \tag{6.1}$$

Put $\tilde{d} = p, \tilde{h} := \Omega h$, and let $C_1 := C_1\left(\tilde{d}, \tau, m, M\right)$ as in Theorem 6.1. Let $\epsilon \leqslant C_1 (\Omega h)^{2p-1}$. Now we apply Theorem 6.1 with parameters $\tilde{d}, \tilde{h}, \tau, m, M, \tilde{\epsilon} = C_2\epsilon$ and the signal $F^c_{(\Omega)}$, where $C_2 \leqslant 1$ will be determined below. We obtain a signal $G^c_{(\Omega),\epsilon}$ such that the following hold for the difference $H := G^c_{(\Omega),\epsilon} - F^c_{(\Omega)}$:

$$m_k(H) = 0, \quad k = 0, 1, \ldots, 2p-2, \tag{6.2}$$
$$m_{2p-1}(H) = \tilde{\epsilon}; \tag{6.3}$$

while also

$$\left|P_{\mathbf{x},j}\left(G^c_{(\Omega),\epsilon}\right) - P_{\mathbf{x},j}\left(F^c_{(\Omega)}\right)\right| \geqslant K_1 (\Omega h)^{-2p+2}\tilde{\epsilon}, \quad j = 1, \ldots, p, \tag{6.4}$$
$$\left|P_{\mathbf{x},j}\left(G^c_{(\Omega),\epsilon}\right) - P_{\mathbf{x},j}\left(F^c_{(\Omega)}\right)\right| \leqslant K_2 (\Omega h)^{-2p+2}\tilde{\epsilon}, \quad j = 1, \ldots, p, \tag{6.5}$$
$$\left|P_{\mathbf{a},j}\left(G^c_{(\Omega),\epsilon}\right) - P_{\mathbf{a},j}\left(F^c_{(\Omega)}\right)\right| \geqslant K_3 (\Omega h)^{-2p+1}\tilde{\epsilon}, \quad j = 1, \ldots, p. \tag{6.6}$$

Now put

$$F^c_{(\Omega),\epsilon} := SC_\Omega\left(G^c_{(\Omega),\epsilon}\right).$$

Applying the inverse blow-up to the above inequalities, we obtain in fact that

$$\left|P_{\mathbf{x},j}\left(F^c_{(\Omega),\epsilon}\right) - P_{\mathbf{x},j}(F^c)\right| \geqslant \frac{K_1}{\Omega} (\Omega h)^{-2p+2}\tilde{\epsilon}, \quad j = 1, \ldots, p, \tag{6.7}$$
$$\left|P_{\mathbf{a},j}\left(F^c_{(\Omega),\epsilon}\right) - P_{\mathbf{a},j}(F^c)\right| \geqslant K_3 (\Omega h)^{-2p+1}\tilde{\epsilon}, \quad j = 1, \ldots, p. \tag{6.8}$$

From the above definitions we have $H_\Omega := SC_\Omega(H) = F^c_{(\Omega),\epsilon} - F^c$. Let us now show that there is a choice of C_2 such that

$$|\mathcal{F}(H_\Omega)(s)| \leqslant \epsilon, \quad |s| \leqslant \Omega. \tag{6.9}$$

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]):

$$\mathcal{F}(H)(\omega)=\sum_{k=0}^{\infty}\frac{1}{k!}m_k\left(H\right)\left(-2\pi\imath\omega\right)^k. \tag{6.10}$$

Next we apply the following easy corollary of the Turán's First Theorem [49, Theorem 6.1], appearing in [13, Theorem 3.1], using the recurrence relation satisfied by the moments of H according to Proposition A.2.

Theorem 6.2. *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 $k\geqslant 2p$ we have the so-called "Taylor domination" property*

$$|m_k(H)|\,R^k\leqslant\left(\frac{2ek}{2p}\right)^{2p}\max_{\ell=0,1,\dots,2p-1}|m_\ell\left(H\right)|\,R^\ell. \tag{6.11}$$

Proposition 6.1. *Assume that $\Omega h\leqslant C_3$, then the constant R in Theorem 6.2 satisfies $R\geqslant C_4$, where C_4 does not depend on Ω,h.*

Proof. Recall that $H=G^c_{(\Omega),\epsilon}-F^c_{(\Omega)}$. The nodes of $F^c_{(\Omega)}$ are, by construction, inside the interval $\left[-\frac{\Omega h}{2},\frac{\Omega h}{2}\right]$. The nodes of $G^c_{(\Omega),\epsilon}$, by (6.5), satisfy

$$\begin{aligned}\left|P_{\mathbf{x},j}\left(G^c_{(\Omega),\epsilon}\right)\right|&\leqslant\frac{\Omega h}{2}+K_2\left(\Omega h\right)^{-2p+2}\tilde{\epsilon}\\&\leqslant\frac{\Omega h}{2}+K_2\left(\Omega h\right)^{-2p+2}C_1\left(\Omega h\right)^{2p-1}\\&=\left(\Omega h\right)\left(C_1K_2+\frac{1}{2}\right).\end{aligned}$$

This concludes the proof with $C_4:=\frac{2}{C_3\left(C_1K_2+\frac{1}{2}\right)}$. ∎

Therefore, by (6.11), (6.2) and (6.3) we have for $k\geqslant 2p$

$$\begin{aligned}|m_k\left(H\right)|&\leqslant\left(\frac{e}{p}\right)^{2p}k^{2p}R^{2p-1-k}\tilde{\epsilon}\\&\leqslant C_5C_4^{2p-1-k}k^{2p}\tilde{\epsilon}.\end{aligned}$$

Now plugging this into (6.10) we obtain

$$|\mathcal{F}\left(H\right)\left(\omega\right)|\leqslant\frac{\tilde{\epsilon}\left|2\pi\omega\right|^{2p-1}}{(2p-1)!}+C_5C_4^{2p-1}\tilde{\epsilon}\sum_{k\geqslant 2p}\left(\frac{2\pi\left|\omega\right|}{C_4}\right)^k\frac{k^{2p}}{k!}.$$

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

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

We can therefore choose $C_2:=\min\left(1,\frac{1}{C_7}\right)$ to ensure that

$$|\mathcal{F}\left(H\right)\left(\omega\right)|\leqslant\epsilon,\quad|\omega|\leqslant 1,$$

which shows (6.9).

Finally, construct the signal $F_\epsilon:=F^{nc}+F^c_{(\Omega),\epsilon}$. Combining (6.9), together with (6.7) and (6.8) finishes the proof of the case $x_j\in\mathbf{x}^c$. ∎

References

[1] Andrey Akinshin, Dmitry Batenkov, and Yosef Yomdin. Accuracy of spike-train Fourier reconstruction for colliding nodes. In *2015 International Conference on Sampling Theory and Applications (SampTA)*, pages 617–621. IEEE, 2015.

[2] Andrey Akinshin, Gil Goldman, and Yosef Yomdin. Geometry of error amplification in solving Prony system with near-colliding nodes. *arXiv:1701.04058 [math]*, January 2017.

[3] Céline Aubel and Helmut Bölcskei. Vandermonde matrices with nodes in the unit disk and the large sieve. *Applied and Computational Harmonic Analysis*, August 2017.

[4] Jon R Auton and Michael L Van Blaricum. Investigation of procedures for automatic resonance extraction from noisy transient electromagnetics data. *Math. Notes*, 1:79, 1981.

[5] Jean-Marc Azaïs, Yohann de Castro, and Fabrice Gamboa. Spike detection from inaccurate samplings. *Applied and Computational Harmonic Analysis*, 38(2):177–195, March 2015.

[6] Dmitry Batenkov. Accurate solution of near-colliding Prony systems via decimation and homotopy continuation. *Theoretical Computer Science*, 681:27–40, June 2017.

[7] Dmitry Batenkov. Stability and super-resolution of generalized spike recovery. *Applied and Computational Harmonic Analysis*, 45(2):299–323, September 2018.

[8] Dmitry Batenkov, Laurent Demanet, Gil Goldman, and Yosef Yomdin. Stability of partial Fourier matrices with clustered nodes. *arXiv:1809.00658 [cs, math]*, September 2018.

[9] Dmitry Batenkov, Laurent Demanet, and Hrushikesh N Mhaskar. Stable soft extrapolation of entire functions. *Inverse Problems*, 35(1):015011, January 2019.

[10] Dmitry Batenkov, Gil Goldman, Yehonatan Salman, and Yosef Yomdin. Algebraic geometry of error amplification: the Prony leaves. *arXiv preprint arXiv:1702.05338*, 2017.

[11] Dmitry Batenkov and Yosef Yomdin. On the accuracy of solving confluent Prony systems. *SIAM Journal on Applied Mathematics*, 73(1):134–154, 2013.

[12] Dmitry Batenkov and Yosef Yomdin. Geometry and singularities of the Prony mapping. In *Proceedings of 12th International Workshop on Real and Complex Singularities*, volume 10, pages 1–25, 2014.

[13] Dmitry Batenkov and Yosef Yomdin. Taylor domination, Turán lemma, and Poincaré-Perron sequences. In Boris Mordukhovich, Simeon Reich, and Alexander Zaslavski, editors, *Contemporary Mathematics*, volume 659, pages 1–15. American Mathematical Society, Providence, Rhode Island, 2016.

[14] F.S.V. Bazán. Conditioning of rectangular Vandermonde matrices with nodes in the unit disk. *SIAM Journal on Matrix Analysis and Applications*, 21:679, 2000.

[15] John J. Benedetto and Weilin Li. Super-resolution by means of Beurling minimal extrapolation. *Applied and Computational Harmonic Analysis*, May 2018.

[16] M. Bertero and P. Boccacci. *Introduction to Inverse Problems in Imaging*. Taylor & Francis, 1998.

[17] B. N. Bhaskar, G. Tang, and B. Recht. Atomic Norm Denoising With Applications to Line Spectral Estimation. *IEEE Transactions on Signal Processing*, 61(23):5987–5999, December 2013.

[18] Emmanuel J Candès and Carlos Fernandez-Granda. Super-resolution from noisy data. *Journal of Fourier Analysis and Applications*, 19(6):1229–1254, 2013.

[19] Emmanuel J Candès and Carlos Fernandez-Granda. Towards a mathematical theory of super-resolution. *Communications on Pure and Applied Mathematics*, 67(6):906–956, 2014.

[20] Laurent Demanet and Nam Nguyen. The recoverability limit for superresolution via sparsity. *arXiv preprint arXiv:1502.01385*, 2015.

[21] David L Donoho. Superresolution via sparsity constraints. *SIAM journal on mathematical analysis*, 23(5):1309–1331, 1992.

[22] Albert Fannjiang. Compressive Spectral Estimation with Single-Snapshot ESPRIT: Stability and Resolution. *arXiv:1607.01827 [cs, math]*, July 2016.

[23] Carlos Fernandez-Granda. Support detection in super-resolution. In *Proceedings of the 10th International Conference on Sampling Theory and Applications (SampTA 2013)*, pages 145–148, 2013.

[24] Carlos Fernandez-Granda. Super-resolution of point sources via convex programming. *Information and Inference*, page iaw005, 2016.

[25] Paulo Ferreira. *Superresolution, the Recovery of Missing Samples, and Vandermonde Matrices on the Unit Circle*. 1999.

[26] Walter Gautschi. On inverses of Vandermonde and confluent Vandermonde matrices. *Numerische Mathematik*, 4(1):117–123, 1962.

[27] Walter Gautschi. On inverses of Vandermonde and confluent Vandermonde matrices. ii. *Numerische Mathematik*, 5(1):425–430, 1963.

[28] Joseph W. Goodman. *Introduction to Fourier Optics*. Roberts and Company Publishers, 2005.

[29] Reinhard Heckel, Veniamin I Morgenshtern, and Mahdi Soltanolkotabi. Super-resolution radar. *Information and Inference: A Journal of the IMA*, 5(1):22–75, 2016.

[30] Y. Hua and T. K. Sarkar. Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 38(5):814–824, May 1990.

[31] Y. Hua and T. K. Sarkar. On SVD for estimating generalized eigenvalues of singular matrix pencil in noise. *IEEE Transactions on Signal Processing*, 39(4):892–900, April 1991.

[32] Stefan Kunis and Dominik Nagel. On the condition number of Vandermonde matrices with pairs of nearly-colliding nodes. *arXiv:1812.08645 [math]*, December 2018.

[33] Weilin Li and Wenjing Liao. Stable super-resolution limit and smallest singular value of restricted Fourier matrices. *arXiv:1709.03146 [cs, math]*, September 2017.

[34] Wenjing Liao and Albert Fannjiang. Music for single-snapshot spectral estimation: Stability and super-resolution. *Applied and Computational Harmonic Analysis*, 40(1):33–67, 2016.

[35] Jari Lindberg. Mathematical concepts of optical superresolution. *Journal of Optics*, 14(8):083001, 2012.

[36] C. A. Micchelli and T. J. Rivlin. A Survey of Optimal Recovery. In *Optimal Estimation in Approximation Theory*, The IBM Research Symposia Series, pages 1–54. Springer, Boston, MA, 1977.

[37] C. A. Micchelli and T. J. Rivlin. Lectures on optimal recovery. In *Numerical Analysis Lancaster 1984*, pages 21–93. Springer, 1985.

[38] C. A. Micchelli, T. J. Rivlin, and S. Winograd. The optimal recovery of smooth functions. *Numerische Mathematik*, 26(2):191–200, June 1976.

[39] Ankur Moitra. Super-resolution, Extremal Functions and the Condition Number of Vandermonde Matrices. In *Proceedings of the Forty-Seventh Annual ACM on Symposium on Theory of Computing*, STOC '15, pages 821–830, New York, NY, USA, 2015. ACM.

[40] Victor Pereyra and Godela Scherer. *Exponential Data Fitting and Its Applications*. Bentham Science Publishers, January 2010.

[41] Thomas Peter and Gerlind Plonka. A generalized Prony method for reconstruction of sparse sums of eigenfunctions of linear operators. *Inverse Problems*, 29(2):025001, 2013.

[42] Gerlind Plonka and Manfred Tasche. Prony methods for recovery of structured functions. *GAMM-Mitteilungen*, 37(2):239–258, 2014.

[43] D. Potts and M. Tasche. Parameter estimation for exponential sums by approximate Prony method. *Signal Processing*, 90(5):1631–1642, 2010.

[44] R. Prony. Essai experimental et analytique. *J. Ec. Polytech.(Paris)*, 2:24–76, 1795.

[45] R Michael Range. *Holomorphic functions and integral representations in several complex variables*, volume 108. Springer Science & Business Media, 2013.

[46] Geoffrey Schiebinger, Elina Robeva, and Benjamin Recht. Superresolution without separation. *Information and Inference: A Journal of the IMA*, 7(1):1–30, March 2018.

[47] P. Stoica and R.L. Moses. *Spectral Analysis of Signals*. Pearson/Prentice Hall, 2005.

[48] G. Tang, B. N. Bhaskar, and B. Recht. Near Minimax Line Spectral Estimation. *IEEE Transactions on Information Theory*, 61(1):499–512, January 2015.

[49] P. Turán, G. Halász, and J. Pintz. *On a New Method of Analysis and Its Applications*. Wiley-Interscience, 1984.

[50] Martin Vetterli, Pina Marziliano, and Thierry Blu. Sampling signals with finite rate of innovation. *IEEE transactions on Signal Processing*, 50(6):1417–1428, 2002.

Appendix 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*

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

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

$$PM_k(a_1, \ldots, a_d, w_1, \ldots, w_d) := \sum_{j=1}^{d} a_j w_j^k, \qquad k = 0, 1, \ldots, 2d-1. \tag{A.2}$$

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 $\mu = (\mu_0, \ldots, \mu_{2d-1}) \in \mathbb{C}^{2d}$,

$$PM_k(a_1, \ldots, a_d, z_1, \ldots, z_d) = \mu_k, \qquad k = 0, 1, \ldots, 2d-1. \tag{A.3}$$

The following fact can be found in the literature about Prony systems and Padé approximation (see e.g. [12] Propositions 3.2 and 3.3).

Proposition A.1. *If a solution $(a_1, \ldots, a_d, z_1, \ldots, z_d)$ to System (A.3) exists with $a_j \neq 0$, $j = 1, \ldots, d$ and for $1 \leqslant j < k \leqslant d$, $z_j \neq z_k$, 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 $\nu := \{\nu_k\}_{k \in \mathbb{N}}$ be given by*

$$\nu_k := PM_k(a_1, \ldots, a_d, z_1, \ldots, z_d).$$

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

$$\sum_{\ell=0}^{d} \nu_{k+\ell} c_\ell = 0, \tag{A.4}$$

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

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

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

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

□

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 $\|a\| > 0$. Then the parameters $(\mathbf{a}, \mathbf{x})$ can be recovered exactly by the following procedure:*

(1) *Construct the* $d \times (d+1)$ *Hankel matrix* $H = [m_{i+j}]_{0 \leqslant i \leqslant d-1}^{0 \leqslant j \leqslant d}$;
(2) *Find a nonzero 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. [12]. □

Appendix B. Quantitative Inverse Function Theorem

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

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

$$Q_r(\mathbf{a}) = Q_{r,d}(\mathbf{a}) = \left\{ \mathbf{x} \in \mathbb{C}^d : \|\mathbf{a} - \mathbf{x}\| \leqslant r \right\}.$$

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 $\mathbf{a}$,

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

For $j = 1, \dots, d$, we denote by $P_j : \mathbb{C}^d \to \mathbb{C}$ 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.

Theorem B.1. *Let* $U \subseteq \mathbb{C}^d$ *be open. Let* $f : U \to \mathbb{C}^d$ *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* $\mathbf{x} \in H(\mathbf{a})$,

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

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* $f^{-1} : V \to U$ *is holomorphic in an open neighborhood of* $Q_R(\mathbf{b})$.
(2) *For each* $j = 1, \dots, d$, $f_j^{-1} = P_j f^{-1} : Q_R(\mathbf{b}) \to \mathbb{C}^d$ *is Lipschitz on* $Q_R(\mathbf{b})$ *with*

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

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 $f : U \to V$ is an injection, then $f^{-1} : V \to U$ 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. [3]

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.

[3] 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. [45], 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 $j = 1, \ldots, d$,

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

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 $\hat{j} \in \{1, \ldots, d\}$ such that $\left|f_{\hat{j}}^{-1}(\mathbf{b}+\mathbf{h}) - f_{\hat{j}}^{-1}(\mathbf{b})\right| = r_{\hat{j}}$. Then by equation (B.1)

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

Hence $R' \geqslant \frac{r_{\hat{j}}}{\alpha_{\hat{j}}} \geqslant \min(\frac{r_1}{\alpha_1}, \ldots, \frac{r_d}{\alpha_d}) = R$. 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

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

This proves item (2) of the Theorem. □

Appendix C. Norm bounds on the inverse Jacobian matrix

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

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

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

Denote the left hand matrix in the factorization (C.1) by $U_{2d} = U_{2d}(z_1, \ldots, z_d)$. The matrix U_{2d} is an instance of a confluent Vandermonde matrix, whose inverses have been extensively studied in [26, 27, 7]. 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 [27].

Theorem C.1 (Gautschi, [27], eqs. (3.10), (3.12))**.** *For $z_1, \ldots, z_d \in \mathbb{C}$ pairwise distinct, put*

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

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

$$\sum_{k=1}^{2d} |A_{j,k}| \leqslant (1 + 2(1 + |z_j|)|\Delta_j|)\Gamma_j, \qquad j = 1, .., d, \tag{C.2}$$

$$\sum_{k=1}^{2d} |B_{j,k}| \leqslant (1 + |z_j|)\Gamma_j, \qquad j = 1, .., d, \tag{C.3}$$

where

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

Proof of Proposition 5.1. By the factorization (C.1)

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

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

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

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

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

By Theorem C.1

$$\sum_{k=1}^{2d} |A_{j,k}| \leqslant (1 + 2(1 + |z_j|)|\Delta_j|)\Gamma_j, \qquad j = 1, .., d, \tag{C.5}$$

$$\sum_{k=1}^{2d} |B_{j,k}| \leqslant (1 + |z_j|)\Gamma_j, \qquad j = 1, .., d, \tag{C.6}$$

where

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

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

  By assumptions we have

$$|z_\ell - z_j| \geqslant \tilde{\eta}, \qquad \forall x_\ell \in \mathbf{x} \backslash \mathbf{x}^c, x_j \in \mathbf{x}, \ell \neq j.$$

  Then we obtain

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

while

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

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

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

and

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

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

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,

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

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

This completes the proof of equations (5.2) and (5.3) of Proposition 5.1.

- 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

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

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

$$
\Delta_j = \sum_{\ell=1,\ell\neq j}^{d} \frac{1}{|z_j - z_\ell|} \leqslant \frac{d-1}{\tilde{h}},
\tag{C.12}
$$

while

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

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

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

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

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,

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

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

This completes the proof of equations (5.4) and (5.5) of Proposition 5.1.

□

Appendix D. Proof of Proposition 5.3

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

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

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

$$
FM_\lambda = PM \circ g_\lambda. \tag{D.2}
$$

Put

$$
W := g_\lambda(H^{\mathrm{o}}_{m,\frac{\tau h}{2\pi}}(F)) = g_\lambda(U).
$$

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 $\mathbf{v}', \mathbf{v}'' \in W = g_\lambda(H^{\mathrm{o}}_{m,\frac{\tau h}{2\pi}}(F)) = g_\lambda(U)$, with $\mathbf{v}' = (\mathbf{a}', \mathbf{z}')$, $\mathbf{a}' = (a'_1, \ldots, a'_d)$, $\mathbf{z}' = (z'_1, \ldots, z'_d)$, $\mathbf{v}'' = (\mathbf{a}'', \mathbf{z}'')$, $\mathbf{a}'' = (a''_1, \ldots, a''_d)$, $\mathbf{z}'' = (z''_1, \ldots, z''_d)$, and $\mathbf{v}' \neq \mathbf{v}''$, it holds that:*

(1) $a'_j \neq 0$ for $j = 1, \ldots, d$.

(2) $z'_j \neq z'_k$ for each $1 \leqslant j < k \leqslant d$.

(3) $z'_j \neq z''_k$ for all $1 \leqslant j < k \leqslant d$.

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

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

We now prove assertions 2 and 3.

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

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

$$|z_j - z_k| \geqslant 4\lambda\tau h, \qquad 1 \leqslant j < k \leqslant d. \tag{D.3}$$

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

$$\frac{\pi}{2} > \frac{1}{d^2} > 2\pi\lambda\tau h. \tag{D.4}$$

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

$$\angle(z_j, z_k) \geqslant 2\pi\lambda\tau h. \tag{D.5}$$

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

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

Next we claim that

$$W \subset H^{\mathrm{o}}_{m,2\lambda\tau h}(\mathbf{a}, \mathbf{z}) = \left\{(\mathbf{a}', \mathbf{z}') \in \mathbb{C}^{2d} \ : \|\mathbf{a}' - \mathbf{a}\| < m, \ \|\mathbf{z}' - \mathbf{z}\| < 2\lambda\tau h\right\}. \tag{D.6}$$

Let $(\mathbf{a}''', \mathbf{x}''') \in H^{\mathrm{o}}_{m,\frac{\tau h}{2\pi}}(F)$. To show (D.6), we need to verify that $g_\lambda(\mathbf{a}''', \mathbf{x}''') \in H^{\mathrm{o}}_{m,2\lambda\tau h}(\mathbf{a}, \mathbf{z})$. For this purpose put $g_\lambda(\mathbf{a}''', \mathbf{x}''') = (\mathbf{a}''', \mathbf{z}''')$, $\mathbf{z}''' = (e^{2\pi i \lambda x'''_1}, \ldots, e^{2\pi i \lambda x'''_d})$. Then using the integral mean value bound, for any $j = 1, \ldots, d$,

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

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

We now prove assertion 2.

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

$$|z_j - z_k| \leqslant |z_j - z'_j| + |z_k - z'_j| < 4\lambda\tau h,$$

which is a contradiction to (D.3).

Finally we prove assertion 3.

Assume by contradiction that for $1 \leqslant j < k \leqslant d$, $z'_j = z''_k$. 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

$$|z_j - z_k| \leqslant |z_j - z'_j| + |z_k - z'_j| < 4\lambda\tau h,$$

which is a contradiction to (D.3).

This completes the proof of Proposition D.1. $\square$

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 $H^{\circ}_{m,\frac{1}{2\lambda}}(F)$.*

Proof. Let $(\mathbf{a}', \mathbf{x}'), (\mathbf{a}'', \mathbf{x}'') \in H^{\circ}_{m,\frac{1}{2\lambda}}(F)$ 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 $k = 1, \ldots, d$, $g_k(a_1, \ldots, a_d, x_1, \ldots, x_d) = a_k$ therefore $\mathbf{a}'' = \mathbf{a}'$.

For coordinates $d + 1, \ldots, 2d$,

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

Fix a certain $1 \leqslant j \leqslant d$ 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 $(\mathbf{a}', \mathbf{x}'), (\mathbf{a}'', \mathbf{x}'') \in H^{\circ}_{m,\frac{1}{2\lambda}}(F)$ 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 $\lambda \leqslant \frac{\Omega}{2d-1}$ and $\Omega h \leqslant \frac{1}{20d}$ then $\frac{1}{\lambda} > h$. Using the former, $U = H^{\circ}_{m,\frac{\tau h}{2\pi}}(F) \subset H^{\circ}_{m,\frac{1}{2\lambda}}(F)$ 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 5.3. $\square$

Appendix E. Proof of Proposition 5.6

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

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

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

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

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

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

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 $\mathbf{z}'' = (z''_1,\ldots,z''_d)$, $z''_j = e^{2\pi i \lambda x''_j}$ for $j = 1,\ldots,d$. 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 $PM\left((\mathbf{a},\mathbf{z})\right) = \boldsymbol{\mu}'$ is given by

$$\left\{ (\mathbf{a}''^{\pi},\mathbf{z}''^{\pi}) : \pi \in \Pi_d \right\}. \tag{E.1}$$

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

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

Finally since $x''_1,\ldots,x''_d$ are real, the set of all solutions to the equation $g_\lambda\left((\mathbf{a},\mathbf{x})\right) = (\mathbf{a}''^{\pi},\mathbf{z}''^{\pi})$ 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 $\left(\mathbf{a}''^{\pi},\mathbf{x}''^{\pi} + \frac{1}{\lambda}\boldsymbol{\ell}\right)$ for some $\pi \in \Pi_d$ and $\boldsymbol{\ell} \in \mathbb{Z}^d$.

This concludes the proof of Proposition 5.6. □

Appendix F. Proof of Proposition 5.7

Within the course of the proof we will make appropriate assumptions of the form $\frac{C'}{\eta} \leqslant \Omega \leqslant \frac{C''}{h}$, 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 $\Omega \geqslant \frac{2(2d-1)}{\eta}$. 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 5.2 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

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

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

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

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*

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

Proposition F.1. *There exists constants K_{11}, K_{12} depending only on d such that for $\frac{K_{11}}{\eta} \leqslant \Omega \leqslant \frac{K_{12}}{h}$ the following holds. For each $3h < |c| \leqslant \frac{\eta}{6}$, 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 $k \in \mathbb{Z}$*

$$\left|c - \frac{k}{\lambda}\right| > 3h. \tag{F.2}$$

We now complete the proof of Proposition 5.7 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

$$\left|x_i - x_j + \frac{n}{\lambda^*}\right| > (32d^4)^{-1}\frac{1}{\lambda_1}, \qquad \text{for all } n \in \mathbb{Z}. \tag{F.3}$$

Put

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

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

$$|c_{i,j}| \geqslant \eta. \tag{F.4}$$

Put $I = \tilde{I}_1^{-1}$, $c = c_{i,j}\lambda_1$, $\epsilon = (32d^4)^{-1}$ 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 $\Omega \geqslant \frac{2}{\eta d}$ we have that $a \geqslant \frac{1}{2}$. With $d \geqslant 2$ by assumption we have that $\epsilon = \frac{1}{32d^4} < \frac{1}{100d^2}$. With $\Omega \geqslant \frac{2}{\eta d}$ we have that

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

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

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

Then

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

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

$$\nu\Big(\Big\{\lambda \in I_1 : \exists k \in \mathbb{Z} \text{ such that } \left|\frac{k}{\lambda} - c_{i,j}\right| \leqslant (32d^4)^{-1}\frac{1}{\lambda_1}\Big\}\Big) < \frac{2}{d^2}|I_1|. \tag{F.6}$$

Define the set

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

Then using (F.6) and the union bound

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

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 5.7.

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

$$\left(A^{\tilde{\pi}}_{R,\lambda^*}(F) + \frac{1}{\lambda^*} \tilde{\boldsymbol{\ell}} \right) \cap \left(A^{\pi}_{R,\lambda_{\tilde{\pi},\tilde{\boldsymbol{\ell}}}}(F) + \frac{1}{\lambda_{\tilde{\pi},\tilde{\boldsymbol{\ell}}}} \boldsymbol{\ell} \right) = \varnothing. \tag{F.8}$$

Proposition 5.7 will then follow by Proposition 5.6.

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

$$\left(A_{R,\lambda^*}(F) + \frac{1}{\lambda^*} \tilde{\boldsymbol{\ell}} \right) \cap \left(A^{\pi}_{R,\lambda_{\tilde{\boldsymbol{\ell}}}}(F) + \frac{1}{\lambda_{\tilde{\boldsymbol{\ell}}}} \boldsymbol{\ell} \right) = \varnothing. \tag{F.9}$$

Fix i such that $\tilde{\boldsymbol{\ell}}_i \neq 0$ and set $n = \tilde{\boldsymbol{\ell}}_i$. Assume that $x_i \in \mathbf{x}^c$, and one can verify that the case where $x_i \in \mathbf{x} \backslash \mathbf{x}^c$ 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 5.5. For each $F' = (\mathbf{a}', \mathbf{x}') \in A_{R,\lambda}(F)$ with $\mathbf{x}' = (x'_1, \ldots, x'_d)$,

$$\left| x'_j - x_j \right| \leqslant \tilde{C}_3 \frac{1}{\Omega} (\Omega h)^{-2p+2} R \leqslant h, \qquad j = 1, \ldots, d. \tag{F.10}$$

We consider the following mutually exclusive and collectively exhaustive cases:

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

Put $c = \frac{n}{\lambda^*}$. Then under the assumption of this case and with $\Omega \geqslant \frac{d}{3h}$ we have that $3h < |c| \leqslant \frac{\eta}{6}$. We can therefore apply Proposition F.1 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

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

Put

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

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

$$|c_j| = |x_i + \frac{n}{\lambda^*} - x_j| \geqslant |x_i - x_j| - \frac{n}{\lambda^*} \geqslant \eta - \frac{n}{\lambda^*} \geqslant \eta - \frac{\eta}{6} \geqslant \frac{5}{6}\eta, \tag{F.12}$$

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 $|c| \geqslant \frac{5}{6}\eta\lambda_2$. Using the former, one can validate that there exists positive constants $C'(d), C''(d)$ such that if $\frac{C'(d)}{\eta} \leqslant \Omega \leqslant \frac{C''(d)}{h}$, then I, c, ϵ, α meet the conditions of Lemma F.1. We then invoke Lemma F.1 and get that

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

Then

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

By the above and using Lemma (F.2)

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

Define the set

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

Using the union bound and (F.13)

$$\nu(E) < |I_2|. \tag{F.14}$$

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 $\lambda_{\tilde{\boldsymbol{\ell}}} = \lambda$. 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

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

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

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

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

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

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

Put (as above)

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

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 $\frac{C'(d)}{\eta} \leqslant \Omega \leqslant \frac{C''(d)}{h}$, then I, c, ϵ, α meet the conditions of Lemma F.1. We then invoke Lemma F.1 and get that

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

Then

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

By the above and using Lemma (F.2)

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

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 $\frac{C'(d)}{\eta} \leqslant \Omega \leqslant \frac{C''(d)}{h}$, then I, c, ϵ, α meet the conditions of Lemma F.1. Invoking it and using Lemma (F.2) we have that

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

Define the set

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

Using the union bound and (F.18)

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

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

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

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

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

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 $y \in \mathbf{x}\backslash\mathbf{x}^c$ such that $|x_i + \frac{n}{\lambda^*} - y| \leqslant \frac{\eta}{6}$. Therefore let j be the index of the non-cluster node for which we have $|x_i + \frac{n}{\lambda^*} - x_j| \leqslant \frac{\eta}{6}$. 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

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

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

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

We now invoke Proposition F.1 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 $k \in \mathbb{Z}$

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

Put

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

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

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

Put $I = \tilde{I}_3^{-1}$, $c = c_\ell \lambda_3$, $\epsilon = 2h\lambda_3$ and $\alpha = \frac{1}{2d}$. Then with the above $|c| \geqslant \frac{5}{6}\eta\lambda_3$ 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 $\frac{C'}{\eta} \leqslant \Omega \leqslant \frac{C''}{h}$ where C', C'' are constants depending only on d. Invoking Lemma F.1 with I, c, ϵ, α we get that

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

Then

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

By the above and using Lemma (F.2)

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

Define the set

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

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 $1 \leqslant \ell \leqslant d, \ell \neq j$,

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

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 $1 \leqslant \ell \leqslant d$

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

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

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

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

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

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

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

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

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

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

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

$$\frac{c}{\lfloor c\rfloor + N} \leqslant a < \frac{c}{\lfloor c\rfloor + N - 1}. \tag{F.23}$$

Then

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

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

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

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

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

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

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

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

$$N \leqslant \lfloor c\rfloor + 2. \tag{F.26}$$

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

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

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

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

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

This completes the proof of Lemma F.1. $\square$

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

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

Using the above

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

This completes the proof of Lemma F.2. □

Proof of Proposition F.1. Without loss of generality assume that $c > 0$ and put $T = c\lambda_1$.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

For any integer $k \leqslant \lfloor T \rceil$ we have using (F.32) that

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

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

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

where in the 3^{rd} inequality we used (F.33), in the 4^{th} inequality we used both $\beta \leqslant \frac{1}{8}$ and $0 \leqslant \alpha \leqslant \lambda_1$, and in last inequality we used $\Omega h \leqslant \frac{d}{24}$.

We then conclude that when Ωh is small enough, each $\lambda(\alpha)$ with $0 \leqslant \alpha \leqslant \lambda_1$ 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 $0 \leqslant \alpha \leqslant \lambda_1$ such that

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

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

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

satisfies (F.33).

We conclude from the above that for

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

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

Now we recall that by Proposition 5.2, 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

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

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

$$\frac{\lambda_1}{4\lceil T\rceil} - \frac{6\lambda_1^2 h}{T} \geqslant \frac{\lambda_1}{T}\left(\frac{1}{4} - 6\lambda_1 h\right) \geqslant \frac{6}{\eta}\left(\frac{1}{4} - 6\lambda_1 h\right) > \frac{1}{\eta},$$

where in the penultimate inequality we used the proposition assumption that $\frac{\eta}{6} \geqslant c = \frac{T}{\lambda_1}$ 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 $\Omega h < \frac{d}{72}$:

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

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 5.2 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 (F.2) this completes the proof of the case $\beta \leqslant \frac{1}{8}$ of Proposition (F.1).

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

Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
E-mail address: batenkov@mit.edu

Department of Mathematics, The Weizmann Institute of Science, Rehovot 76100, Israel
E-mail address: gil.goldman@weizmann.ac.il

Department of Mathematics, The Weizmann Institute of Science, Rehovot 76100, Israel
E-mail address: yosef.yomdin@weizmann.ac.il