



Super-Resolution of Positive Sources on an Arbitrarily Fine Grid

Veniamin I. Morgenshtern¹ 

Received: 15 May 2020 / Revised: 29 March 2021 / Accepted: 25 April 2021 /
Published online: 29 November 2021

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract

In super-resolution it is necessary to locate with high precision point sources from noisy observations of the spectrum of the signal at low frequencies capped by f_{l_0} . In the case when the point sources are positive and are located on a grid, it has been recently established that the super-resolution problem can be solved via linear programming in a stable manner and that the method is nearly optimal in the minimax sense. The quality of the reconstruction critically depends on the Rayleigh regularity of the support of the signal; that is, on the maximum number of sources that can occur within an interval of side length about $1/f_{l_0}$. This work extends the earlier result and shows that the conclusion continues to hold when the locations of the point sources are arbitrary, i.e., the grid is arbitrarily fine. The proof relies on new interpolation constructions in Fourier analysis.

Keywords Super-resolution · Sparsity · Inverse problem · Convex optimization · Linear programming · Single-molecule super-resolution microscopy · Spectrum extrapolation

Mathematics Subject Classification 65T40 · 65T99 · 65Z05

Communicated by Gabriel Peyre.

✉ Veniamin I. Morgenshtern
veniamin.morgenshtern@fau.de

¹ Chair of Multimedia Communications and Signal Processing, University of Erlangen-Nuremberg, 91058 Erlangen, Germany

1 Introduction

The super-resolution problem of positive sources (see Fig. 1) consists of recovering a high-frequency signal

$$x(w) = \sum_i x_i \delta(w - w_i) \tag{1}$$

consisting of positive point sources (*spikes*, for short) located at unknown positions $w_i \in [0, 1)$ and of unknown intensity $x_i > 0$; $\delta(\cdot)$ is the Dirac delta function. The signal is observed through a convolution measurement of the form

$$s(v) = \int k_{l_0}(v - w)x(w)dw + z(v), \tag{2}$$

where $k_{l_0}(\cdot)$ is a low-frequency kernel that erases the high-frequency components of the signal and $z(\cdot)$ is noise.

This problem arises in single-molecule super-resolution microscopy [7,22,33]. In this application, w_i 's encode the unknown locations of fluorescent molecules, x_i is proportional to the number of photons emitted by the i th molecule during the observation time. Crucially, the number of photons is a nonnegative number, leading to the assumption $x_i > 0$, which makes the problem much simpler. Assume that light of wavelength λ_{l_0} is emitted by the molecules. Due to diffraction of light, the high-frequency spatial details of the signal are destroyed, no matter how perfect or large the microscope is. At the detector we record a blurred version of the signal, no details smaller than about λ_{l_0} are visible. To restate this mathematically: the function $k_{l_0}(\cdot)$ models the (PSF) of the microscope; due to diffraction of light the PSF is band-limited to $f_{l_0} \triangleq 1/\lambda_{l_0}$. The noise $z(v)$ represents all sources of noise in the system. For example, the thermal noise at the detector, the Poisson quantum mechanical noise due to photon quantization in low-intensity imaging, and the noise originating from the imperfect knowledge of the PSF in the optical system. We refer the interested reader to [43], where the connection to super-resolution microscopy is worked out in details.

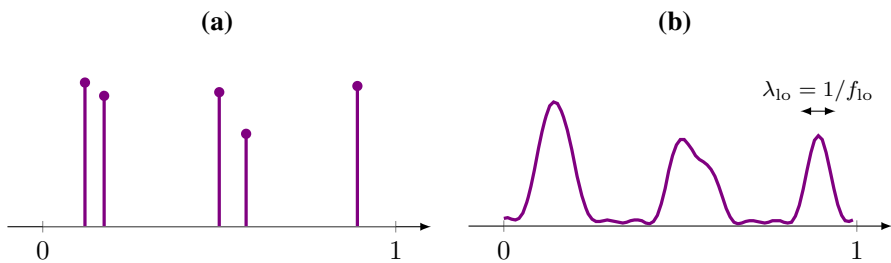


Fig. 1 Microscope as a low-pass filter: signal in **a** and convolution measurement in **b** (Color figure online)

1.1 Discrete Model

In the earlier work [43] a discrete analog of the model in (1) and (2) has been considered. The signal is modeled by a discrete vector $\mathbf{x} = [x_0 \dots x_{N-1}]^T \in \mathbb{R}^N$, where N is the number of elements in the grid, corresponding to partitioning the interval $w_i \in [0, 1)$ into N equispaced segments. Each nonzero element in \mathbf{x} corresponds to one spike in (1). The PSF is modeled by matrix \mathbf{Q} that implements an ideal low-pass filter in the sense that it has a flat spectrum with a sharp cut-off at f_{l_0} . Formally,

$$\mathbf{Q} = \mathbf{F}^H \hat{\mathbf{Q}} \mathbf{F}, \tag{3}$$

where \mathbf{F} is the $N \times N$ discrete Fourier transform matrix

$$[\mathbf{F}]_{k,l} = \frac{1}{\sqrt{N}} e^{-i2\pi kl/N}$$

and $\hat{\mathbf{Q}} = \text{diag}([\hat{Q}_{-N/2+1} \dots \hat{Q}_{N/2}]^T)$ with

$$\hat{Q}_k = \begin{cases} 1, & k = -f_{l_0}, \dots, f_{l_0}, \\ 0, & \text{otherwise.} \end{cases} \tag{4}$$

The wavelength $\lambda_{l_0} = 1/f_{l_0}$ gives the width of the convolution kernel represented by \mathbf{Q} . This kernel is called the Dirichlet kernel. We assume throughout the paper that N is even for simplicity.

Translated to discrete setting the model in (2) becomes

$$\mathbf{s} = \mathbf{Q}\mathbf{x} + \mathbf{z}. \tag{5}$$

1.2 Recovery Algorithm

Our recovery method from the observations \mathbf{s} in (5) is extremely simple: solve

$$\hat{\mathbf{x}} = \arg \min_{\tilde{\mathbf{x}}} \|\mathbf{s} - \mathbf{Q}\tilde{\mathbf{x}}\|_1 \quad \text{subject to} \quad \tilde{\mathbf{x}} \geq \mathbf{0}. \tag{CVX}$$

In other words, we are looking for a set of positive spikes such that the mismatch in received intensities is minimum. Note that this method does not make use of any knowledge other than the observations \mathbf{s} and the PSF \mathbf{Q} . Furthermore, (CVX) is a simple convex optimization program, which can be recast as a linear program since both \mathbf{x} and \mathbf{Q} are real valued.

1.3 Rayleigh Regularity

Consider the discrete signal $\mathbf{x} \in \mathbb{R}^N$ as samples on the grid $\{0, 1/N, \dots, 1 - 1/N\} \subset \mathbb{T}$, where \mathbb{T} is the circle in 1D, i.e., the interval $[0, 1)$ with 0 and 1 identified. For $a, b \in$

\mathbb{T} , the wrap-around distance between a and b is $|b - a| \triangleq \min(b - a \pmod 1, a - b \pmod 1)$; for an interval $[a, b] \subset \mathbb{T}$, its wrap-around length is $|b - a|$.

We introduce a definition of Rayleigh regularity inspired by [23, Def. 1]. Let $\text{supp}(\mathbf{x}) \triangleq \{l/N : x_l > 0\}$ denote the support of the discrete signal. As we shall see, our ability to super-resolve the signal \mathbf{x} , will be fundamentally determined by how regular $\text{supp}(\mathbf{x})$ is in the following sense.

Definition 1 (Rayleigh regularity) We say that the set of points $\mathcal{V} \subset \mathbb{T}$ is Rayleigh-regular with parameters (d, r) and write $\mathcal{V} \in \mathcal{R}(d, r)$ if it may be partitioned as $\mathcal{V} = \mathcal{V}_1 \cup \dots \cup \mathcal{V}_r$, where the \mathcal{V}_i 's are disjoint, and each obeys a separation constraint:

1. for all $1 \leq i < j \leq r$, $\mathcal{V}_i \cap \mathcal{V}_j = \emptyset$;
2. for all intervals $\mathcal{D} \subset \mathbb{T}$ of wrap-around length $|\mathcal{D}| = d$ and all i ,

$$|\mathcal{V}_i \cap \mathcal{D}| \leq 1,$$

where $|\mathcal{V}_i \cap \mathcal{D}|$ denotes the cardinality of the set $\mathcal{V}_i \cap \mathcal{D}$.

In this paper we are interested in super-resolving signals with Rayleigh-regular support: $\text{supp}(\mathbf{x}) \in \mathcal{R}(d, r)$. Such signals are illustrated in Fig. 2.

As we will discuss, in the special case when $\text{supp}(\mathbf{x}) \in \mathcal{R}(\tilde{c}\lambda_{10}, 1)$ [i.e., when $r = 1$] with \tilde{c} a bit larger than one (as in Fig. 2a), the super-resolution problem is particularly easy. In this case we will say that the spikes in \mathbf{x} are *well-separated*.

1.4 Discrete Stability Estimates

The main result of the earlier work [43] is the following proposition.

Proposition 1 Assume $\mathbf{x} \geq \mathbf{0}$ and $\text{supp}(\mathbf{x}) \in \mathcal{R}(\kappa\lambda_{10}r, r)$ with $\kappa \triangleq 1.87$ and $f_{10} \geq 128r$. Assume that the observations \mathbf{s} are given by (5). Then the solution $\hat{\mathbf{x}}$ to (CVX) obeys

$$\underbrace{\|\hat{\mathbf{x}} - \mathbf{x}\|_1}_{\text{Error}} \leq C_d(r) \cdot \left(\frac{N}{f_{10}}\right)^{2r} \cdot \|\mathbf{z}\|_1 = C_d(r) \cdot \text{DSRF}^{2r} \cdot \|\mathbf{z}\|_1, \tag{6}$$

where $C_d(r)$ only depends on r (if $\text{DSRF} \geq 3.03/r$, it can be taken to be $C_d(r) = r^{2r} \cdot 4 \cdot 17^r$).

The ratio $\text{DSRF} \triangleq N/f_{10}$ is called the *discrete super-resolution factor*; this is the ratio between the scale at which we observe the data, $1/f_{10}$, and the scale of the finest details in the data, $1/N$.

1.5 Breakdown of Discrete Stability Estimates

In practice, signals do not belong to a discrete grid. In order to accurately approximate the continuous model in (1) we might need to make the grid very fine, i.e., take N large.

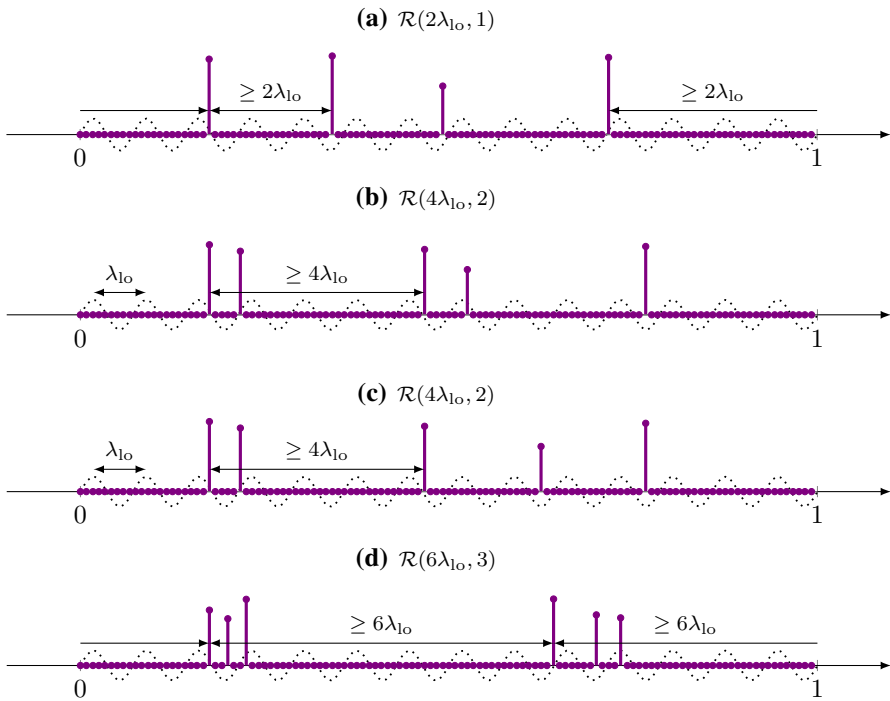


Fig. 2 Examples of discrete N dimensional signals whose support belongs to the Rayleigh classes $\mathcal{R}(2\lambda_{l_0}, 1)$ in **a**, $\mathcal{R}(4\lambda_{l_0}, 2)$ in **b** and in **c**, $\mathcal{R}(6\lambda_{l_0}, 3)$ in **d** depicted on the grid $\{0, 1/N, \dots, 1 - 1/N\} \subset \mathbb{T}$. Note that the signals in **b** and in **c** both have support in $\mathcal{R}(4\lambda_{l_0}, 2)$. In general, Rayleigh regularity *does not* require that all spikes in the signal are arranged into separated clusters as is the case in **b** and in **d**. The sine wave $\sin(2\pi f_{l_0} t)$ at the highest visible frequency is shown by the dotted line for reference. Here, $N = 120$ and $f_{l_0} = 14$, so that $\lambda_{l_0} = 1/14$. By periodicity, the endpoints are identified (Color figure online)

The problem is that the theoretical result in (6) becomes meaningless when f_{l_0} and $\|\mathbf{z}\|_1$ remain fixed, and $N \rightarrow \infty$. Indeed, observe that (6) guarantees accurate signal recovery when the right-hand side of (6) is much smaller than $\|\mathbf{x}\|_1$. When $N \rightarrow \infty$ with f_{l_0} fixed, then $\text{DSRF}^{2r} \rightarrow \infty$ very quickly, so that the right-hand side of (6) becomes larger than $\|\mathbf{x}\|_1$, even for very small noise.

This is expected. Consider the hypothetical situation illustrated in Fig. 3. The true signal \mathbf{x} consists of three spikes as depicted in Fig. 3a in solid purple. The grid is very fine (N is large); the PSF is wide as shown in Fig. 3b (the solid purple curve, with characteristic width λ_{l_0} , represents $\mathbf{s} = \mathbf{Q}\mathbf{x} + \mathbf{z}$ with \mathbf{x} from Fig. 3a); and the data is noisy. Imagine an algorithm produced an estimate $\hat{\mathbf{x}}_{\text{good}}$ as depicted in Fig. 3c by the dashed blue spikes. The estimate $\hat{\mathbf{x}}_{\text{good}}$ is excellent: the dashed blue spikes are located in the neighboring discrete bins to the corresponding solid purple ground truth spikes, the magnitudes are estimated perfectly. In the presence of noise we cannot hope for infinite resolution, so for large N , we should be happy if we were able to obtain $\hat{\mathbf{x}}_{\text{good}}$ as in Fig. 3c. Yet,

$$\|\hat{\mathbf{x}}_{\text{good}} - \mathbf{x}\|_1 = 2\|\mathbf{x}\|_1,$$

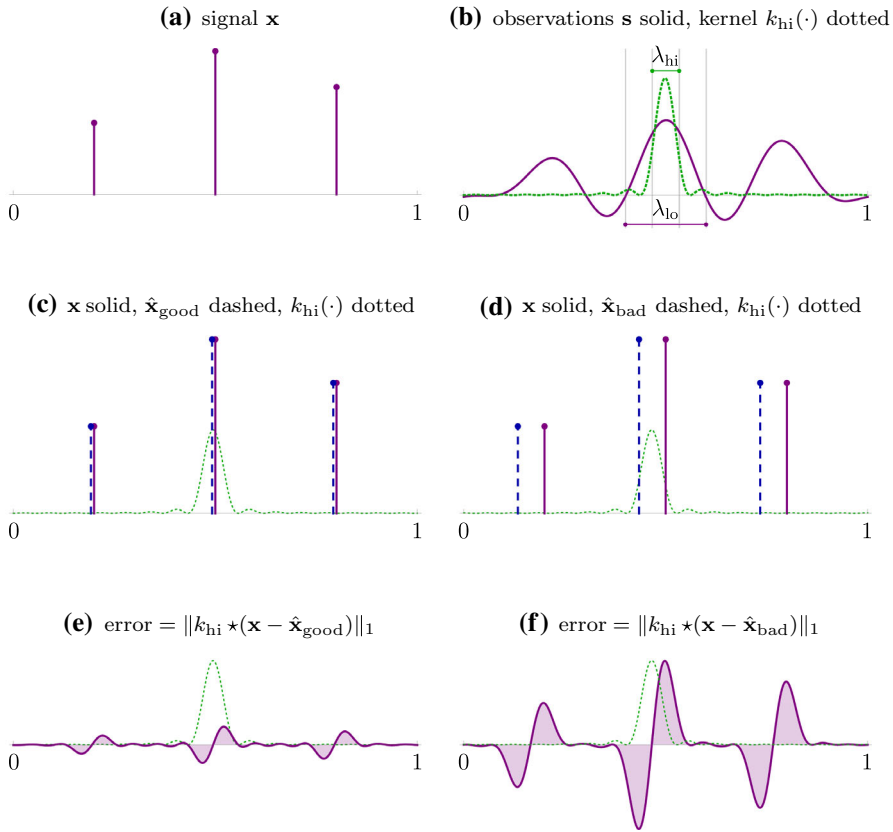


Fig. 3 Measuring the estimation error when the grid is very fine (N is large) (Color figure online)

i.e., the estimation error is about as large as it can possibly be. We conclude that the reason why the result in (6) becomes meaningless when f_{l_0} and $\|\mathbf{z}\|_1$ remain fixed and $N \rightarrow \infty$ is that the error metric $\|\hat{\mathbf{x}} - \mathbf{x}\|_1$ becomes inadequate. We need a more *forgiving* error metric that should penalize small localization errors on the fine grid mildly.

We will explain in Sect. 2 how to construct the more forgiving error metric and how to change the definition of super-resolution factor accordingly. With these modifications we can generalize Proposition 1 and formulate the stability estimates in Theorem 1 that remain meaningful even when $N \rightarrow \infty$ and f_{l_0} and $\|\mathbf{z}\|_1$ are fixed. With the appropriate new definitions, the result in Theorem 1 is nearly identical to that in Proposition 1. Surprisingly, the proof technique necessary to obtain Theorem 1 is much harder than the trick that was sufficient to prove Proposition 1. The proof relies on new trigonometric interpolation constructions that constitute the main mathematical contribution of this paper.

2 Main Results

2.1 Measuring the Reconstruction Error

To avoid penalizing the estimators that produce spikes very close to the original spikes on the fine grid, a natural approach is to convolve the difference $\hat{\mathbf{x}} - \mathbf{x}$ with a nonnegative kernel $k_{hi}(\cdot)$ of width λ_{hi} (represented by the dotted green line in Fig. 3b) before computing the ℓ_1 norm:

$$\text{error} = \|k_{hi} \star (\hat{\mathbf{x}} - \mathbf{x})\|_1,$$

where

$$[k_{hi} \star (\hat{\mathbf{x}} - \mathbf{x})]_n \triangleq \sum_{m=0}^{N-1} k_{hi} \left(\frac{n-m}{N} \right) h_m$$

and $\mathbf{h} = [h_0, h_1, \dots, h_{N-1}]^T \triangleq \hat{\mathbf{x}} - \mathbf{x}$ is the difference vector. The new error metric is illustrated in Fig. 3c–f. When the estimated spikes are closer than λ_{hi} to the original spikes, as is the case for $\hat{\mathbf{x}} = \hat{\mathbf{x}}_{\text{good}}$ in Fig. 3c, the error, represented by the area of the shaded region in Fig. 3e, is very small. Conversely, when the estimated spikes are further than λ_{hi} from the original spikes, as is the case for $\hat{\mathbf{x}} = \hat{\mathbf{x}}_{\text{bad}}$ in Fig. 3d, we have, $\text{error} = \|k_{hi} \star (\hat{\mathbf{x}} - \mathbf{x})\|_1 \approx 2\|\mathbf{x}\|_1$, so that the error is large, as illustrated in Fig. 3f.

The width, λ_{hi} , of the kernel $k_{hi}(\cdot)$ is a parameter of the theory. This parameter will be chosen to be (i) larger (or equal to) the finest scale of the data, $\lambda_{hi} \geq 1/N$, and, simultaneously, (ii) smaller than the native resolution of the observations, $\lambda_{hi} < \lambda_{lo}$. Having chosen λ_{hi} , we define the *super-resolution factor* as:

$$\text{SRF} \triangleq \frac{\lambda_{lo}}{\lambda_{hi}}.$$

The SRF will play the same role in our theory as the DSRF played in Proposition 1. In Fig. 3b, the SRF is the ratio between λ_{lo} , the width of the kernel \mathbf{Q} , and λ_{hi} , the width of the kernel $k_{hi}(\cdot)$.

To be concrete, a reasonable situation might be: $\lambda_{lo} = 1/10$, $1/N = 1/1000$ so that $\text{DSRF} = 100$. This makes the right-hand side of (6) huge so that the stability estimate is useless. Now, choose $\lambda_{hi} = 1/100$ so that $\text{SRF} = 10$, which is much smaller than DSRF. The main result of this paper, Theorem 1 below, shows that we can upper-bound the error $\|k_{hi} \star (\hat{\mathbf{x}} - \mathbf{x})\|_1$ in terms of SRF^{2r} , which is much smaller than DSRF^{2r} , keeping the bound tight for realistic values of the noise.

For $k_{hi}(\cdot)$, in this paper we use the Fejér kernel:

$$k_{hi}(t) \triangleq \frac{1}{N} \frac{1}{f_{hi} + 1} \left(\frac{\sin(\pi(f_{hi} + 1)t)}{\sin(\pi t)} \right)^2, \quad f_{hi} = 1/\lambda_{hi}. \tag{7}$$

The normalization is such that

$$\sum_{n=0}^{N-1} k_{hi} \left(\frac{n}{N} \right) = 1, \tag{8}$$

which ensures that the “energy” in the error is preserved in the sense that $\text{error} = \|k_{hi} \star (\hat{\mathbf{x}} - \mathbf{x})\|_1 \approx 2\|\mathbf{x}\|_1$ whenever the estimated spikes in $\hat{\mathbf{x}}$ are far away from the true spikes in \mathbf{x} . The concrete form of the kernel $k_{hi}(\cdot)$ is not important. Our results hold for any other periodic nonnegative high-resolution kernel as long as it satisfies conditions (106) and (107) below.

When $N \rightarrow \infty$, the error metric defined here becomes the one used in [12] in the analysis of the continuous super-resolution problem. Compared to [12], the key novelty of this paper is that the results in [12] apply only when the spikes in the signal are well-separated [$\text{supp}(\mathbf{x}) \in \mathcal{R}(1.87\lambda_{lo}, 1)$] as in Fig. 2a, a stringent assumption. In this paper we don’t assume that the spikes are well-separated and our results also hold for signals with $\text{supp}(\mathbf{x}) \in \mathcal{R}(1.87\lambda_{lo}r, r)$ and $r > 1$ as in Fig. 2b–d. The price we pay is that our results are only valid for nonnegative signals, whereas the results in [12] are valid for complex-valued signals.

2.2 Stability Estimate on an Arbitrarily Fine Grid

In this paper we prove the following theorem.

Theorem 1 *Assume $\mathbf{x} \geq \mathbf{0}$ and $\text{supp}(\mathbf{x}) \in \mathcal{R}(\kappa\lambda_{lo}r, r)$ with $\kappa \triangleq 1.87$ and $f_{lo} \geq 128r$. Assume $\lambda_{hi} < \lambda_{lo}$, $\lambda_{hi} \geq 1/N$, and $\text{SRF} > 12$. Assume, in addition, that the elements of $\text{supp}(\mathbf{x})$ are separated by at least $2\lambda_{hi}$: if $t, t' \in \text{supp}(\mathbf{x})$ with $t \neq t'$, then $|t - t'| \geq 2\lambda_{hi}$, where $|\cdot|$ is the wrap-around distance on \mathbb{T} . Assume that the observations \mathbf{s} are given by (5). Then the solution $\hat{\mathbf{x}}$ to (CVX) obeys*

$$\|k_{hi} \star (\hat{\mathbf{x}} - \mathbf{x})\|_1 \leq C(r)\text{SRF}^{2r} \|\mathbf{z}\|_1, \tag{9}$$

where $C(r) \triangleq r^{2r+4}c^{r+1}$ and the positive numerical constant c is defined in (128) below.

The theorem is proven in Sects. 5, 6, 7 and in the appendices. Before we embark on the proof, we discuss the significance and the accuracy of the result.

2.2.1 Significance of the Result

Theorem 1 gives essentially the same stability estimate for an arbitrarily fine grid as Proposition 1 does for a discrete grid. With the new definition for error metric, λ_{hi} in Theorem 1 plays the same role as the grid segment size, $1/N$, played in Proposition 1. In turn, the grid segment size, $1/N$, in Theorem 1 may be arbitrarily small without affecting the stability estimate at all. The only thing that changes when N grows is that it becomes numerically harder to solve (CVX).

2.2.2 Tightness

The result is information-theoretically tight in the following sense. It is possible to prove a converse theorem (see [43, Sec. 2.3]) that says that the best possible algorithm in the worst case (the minimax setting) cannot achieve stability estimate in (9) with super-resolution factor dependence better than SRF^{2r-1} . In other words, the exponent of SRF in (9) is near-optimal.

We have made no attempt to optimize $C(r)$. Finding the tightest possible $C(r)$ is an important open problem, which seems to be hard to address with the mathematical techniques developed in this paper.

2.2.3 Mathematical Novelty

The reader might expect that since Theorem 1 is so similar to Proposition 1, the proof of Theorem 1 is a minor modification of the work done in [43]. Perhaps surprisingly, this is not the case.

The proof technique in [43] relied on a simple and elegant trigonometric interpolation construction reviewed in Sect. 6.2. In this paper, in addition, we had to develop a flexible set of techniques that allowed us to build trigonometric polynomials with specific interpolation properties. These techniques—that constitute the main mathematical contribution of this paper—are presented in Sects. 6.3, 6.4, and in Appendix 2 “Proof of Lemma 5”. We believe that the new techniques are interesting in their own right and may be useful in other projects.

2.2.4 Separation by $2\lambda_{\text{hi}}$

Theorem 1 requires the assumption that no two spikes in \mathbf{x} are closer than $2\lambda_{\text{hi}}$. It is important to contrast this assumption with the separation assumption in [12, 13]. The results in [13] hold only when no two spikes in \mathbf{x} are closer than $1.87\lambda_{\text{lo}}$ (the spikes are well-separated). Our separation requirement is much weaker than the one needed in [12, 13]: we require the separation at the scale of λ_{hi} whereas the results in [12, 13] need separation on the scale of λ_{lo} . Since the whole point of super-resolution is to reconstruct the original signal with accuracy about $\lambda_{\text{hi}} \ll \lambda_{\text{lo}}$, our assumption is mild, whereas the assumption in [12, 13] is restrictive.

Further, it follows from the proof of Theorem 1 that the $2\lambda_{\text{hi}}$ separation requirement may be relaxed to, for example, $\lambda_{\text{hi}}/2$, or, more generally, to $\lambda_{\text{hi}}/\beta$ for any $\beta > 1$. The result in Theorem 1 will not change, except that the constant $C(r)$ will now depend on β . Specifically, the result will read:

$$\|k_{\text{hi}} \star (\hat{\mathbf{x}} - \mathbf{x})\|_1 \leq r^{2r+4} c^{r+1} \beta^{2r} \text{SRF}^{2r} \|\mathbf{z}\|_1.$$

To keep the proof of Theorem 1 as clean as possible, we decided to stick with the $2\lambda_{\text{hi}}$ separation assumption in the theorem.

Finally, it is not clear if the separation assumption of the form $\lambda_{\text{hi}}/\beta$ is fundamentally necessary. Certainly, it is necessary for the proof technique developed in this paper. It is an open problem to either find a proof of Theorem 1 that does not rely on this

assumption, or to prove a converse result showing that this assumption is unavoidable. Note that there is no *explicit* separation assumption in Proposition 1; however, since the spikes are on the grid, the separation assumption at the scale of $1/N$ is made *implicitly*.

2.2.5 Density Constant

We next discuss the following question: can the constant $\kappa = 1.87$ in Theorem 1 be made smaller without changing the result? The answer is “probably yes”. Specifically, our proof builds upon Lemmas 1 and 2 below. The lemmas generalize [12, Lm. 2.4, Lm. 2.5, Sec. 2.5] and their proof exploits a construction developed in [13]. The specific value for $\kappa = 1.87$ comes from the construction borrowed from [13]. An improved construction has recently been reported in [28] leading to a smaller value $\kappa = 1.26$. To keep this paper as simple as possible, we decided not to accommodate this improvement. To do so, one would need to change Lemmas 1 and 2 below and the proof of Lemma 2 in Appendix 1 all other derivations in this paper will remain unchanged. The constant $C(r)$ in Theorem 1 would need to be updated accordingly.

We expect that there is a trade-off: the larger κ is, the smaller the constant $C(r)$ can be made. However, our estimates do not provide the smallest possible constant. Hence, we cannot analyze the trade-off.

Finally, as explained in [43, Sec. 2.3.1], $\kappa > 1$ is a fundamental limit, so our result is within the factor 1.87 from the optimum.

2.2.6 Gridless Super-Resolution

It has been shown in [8,12,13] that under the assumption that spikes are separated by at least $1.87\lambda_{10}$ (well-separated spikes), one can solve the gridless super-resolution problem in which the spikes have completely arbitrary locations on \mathbb{T} (no need for the $1/N$ discretization). It turns out that in the gridless setup one needs to solve an infinite-dimensional, but convex, total-variation-minimization problem (see [13, eq. (1.4)]). Surprisingly, if one works in the dual domain and uses the idea of lifting, the equivalent problem becomes finite-dimensional and, therefore, may be solved on the computer. The solution to the original problem may then be reconstructed by duality. This approach is explained in [13, Sec. 4].

The approach, by now standard, may be carried over to the problem considered in this paper, where we work with a nonnegative signal \mathbf{x} and the spikes need not be well-separated. The same trigonometric polynomials that certify optimality of (CVX) and lead to Theorem 1 may also be used to prove stability of the corresponding gridless algorithm.

The reason why we chose to focus on the arbitrarily fine grid and not to discuss the gridless problem in details is the following *practical* consideration. In applications, for example in super-resolution microscopy, there is no real difference between the gridless problem and the problem with a very fine grid. The real sources have some finite nonzero size, perhaps small. Therefore, in practice, one has a choice between solving (CVX) on a sufficiently fine grid or solving the infinite-dimensional total-variation-minimization problem via lifting. To solve (CVX) with N variables efficiently, one

would use a first-order solver whose complexity is dominated by repeated multiplications by \mathbf{Q} , \mathbf{Q}^\top . Using (3) one would implement \mathbf{Q} via the fast Fourier transform so that each matrix multiplication takes $\mathcal{O}(N \log N + f_{\text{lo}})$ multiplications. The gridless approach via lifting, in its standard implementation, requires one to solve a semidefinite convex optimization problem (see [13, eq. (4.3)]) with $\mathcal{O}(f_{\text{lo}}^2)$ variables. The complexity of the gridless approach does not depend on N at all, a very nice property. However, the necessity to deal with a semidefinite problem with $\mathcal{O}(f_{\text{lo}}^2)$ variables make it more costly than solving (CVX) on a sufficiently fine grid, for example, in the important applications in super-resolution microscopy. Having said this, recently very interesting new approaches to solve the gridless problem much faster have been developed. One idea is to construct a solver for the semidefinite problem that directly leverages the structure of the problem [15], another idea is to construct primal Frank-Wolfe based solvers with heuristics [10,20,29]. It would be interesting to see a rigorous study comparing the fine grid methods with the gridless methods in applied super-resolution microscopy problems. We refer to [20, Sec. 5] for fascinating work in this direction.

2.2.7 General PSFs

The sharp rectangular frequency cut-off of \mathbf{Q} in (4) corresponds to the Dirichlet PSF $k_{\text{lo}}(\cdot)$ in (2). The Dirichlet kernel takes negative values (as shown in Fig. 3b in solid purple), whereas all PSFs in microscopy take nonnegative values (as shown in Fig. 1). The simplest PSF that takes nonnegative values is the Fejèr kernel. The spectrum of \mathbf{Q} that corresponds to the Fejèr kernel has a triangular decay of \hat{q}_k in (4) as in [43, eq. 13] and in (196). The results for the rectangular spectrum can be translated into the results for the triangular spectrum (in fact for the spectrum of any reasonable shape) using the idea of spectrum equalization. We refer the reader to [43] for a detailed explanation on how this can be done. In this paper we focus on the basic case in (4) only.

2.2.8 The Use of ℓ_1 -norm in Stability Estimates

A careful reader may ask why do we use the ℓ_1 -norm in (9) and not the more usual ℓ_2 -norm. Interestingly, as explained in [43, Sect. 1.1], in super-resolution microscopy $\|\mathbf{x}\|_1$, and not $\|\mathbf{x}\|_2$, has the meaning of cumulative emitted intensity or the total energy of light emitted per second. Therefore, $\|k_{\text{hi}} \star (\hat{\mathbf{x}} - \mathbf{x})\|_1$ is a natural choice as it corresponds to the error (at high resolution) in cumulative emitted intensity, exactly the quantity one would like to minimize in microscopy.

The other reason for the choice of the ℓ_1 -norm is mathematical feasibility. It would be interesting to obtain similar bounds for the ℓ_2 -norm as this might be relevant for some applications. However, the dual certificates (see below) required to obtain such bounds need to have completely different properties, and, hence, new mathematical constructions seem to be necessary for such an extension.

3 Literature Review and Innovations

3.1 Prior Art

Prony's method Prony's method [46] is an algebraic approach for solving the gridless super-resolution problem from *noiseless* data when the number of spikes is known a priori. The observations \mathbf{s} are used to form a trigonometric polynomial, whose roots coincide with the spike locations. The trigonometric polynomial is then factored, thus revealing those locations, and the amplitudes estimated by solving a system of linear equations. In the noiseless case, Prony's method recovers \mathbf{x} perfectly provided that $\|\mathbf{x}\|_0 \leq f_{10}$. Here and below, $\|\cdot\|_0$ is the pseudo-norm that denotes the number of nonzero elements in a vector. No further Rayleigh regularity assumption on the signal support is needed. With noise, however, the performance of Prony's method degrades sharply. The difficulty comes from the fact that the roots of a trigonometric polynomial constructed by an algebraic method can shift dramatically even with small changes in the data. See [6] for quantitative analysis showing that Prony's method is very sensitive to noise. Therefore, a crucial problem is to find a method for solving the super-resolution problem in the *presence of noise* whose performance decays gracefully with the amount of noise.

Fundamental limits In the pioneering work [23], Donoho studied limits of performance for the super-resolution problem and recognized the importance of Rayleigh regularity as the fundamental property that determines how easy it is to super-resolve the signal. He analyzed an *intractable* exhaustive search algorithm and demonstrated that assuming $\text{supp}(\mathbf{x}) \in \mathcal{R}(2\lambda_{10}r, r)$, the estimator, $\hat{\mathbf{x}}$, produced by this algorithm satisfies:

$$\|\hat{\mathbf{x}} - \mathbf{x}\|_2 \leq \tilde{C}(r) \text{SRF}^{2r+1} \|\mathbf{z}\|_2. \quad (10)$$

The algorithm proposed by Donoho may only be applied to vectors \mathbf{x} with very few dimensions. Therefore, the fundamental problem posed by Donoho is to find an *efficient* algorithm that is stable in the sense of (10). Donoho has also proven a converse to (10): the SRF dependence in (10) cannot be better than SRF^{2r-1} even for the best possible algorithm in the worst-case scenario (the minimax setting). The results of Donoho have been recently (partially) improved in [4, 17] where for the same intractable algorithm the following stability estimate was derived:

$$\|\hat{\mathbf{x}} - \mathbf{x}\|_2 \leq \tilde{C}(r, \|\mathbf{x}\|_0) \text{SRF}^{2r-1} \|\mathbf{z}\|_2. \quad (11)$$

The result is sharp in the sense that the SRF dependence matches Donoho's converse. The weakness is that $\tilde{C}(r, \|\mathbf{x}\|_0)$ depends on the total number of spikes in the signal, which may be very large. This weakness has recently been (partially) removed in [5], where a model with clustered spikes is considered that is somewhat similar, but distinct from Rayleigh regularity; stability estimates for the spike locations and (complex-valued) spike magnitudes are derived and the estimate for the magnitudes scales as $\text{SRF}^{2r-1} \|\mathbf{z}\|_2$. Note further that the stability estimates in (10), (11) are

expressed in terms of ℓ_2 norms, whereas our stability estimates in (9) are expressed in terms of ℓ_1 norms.

Other works [6,53,54] study the stability of the super-resolution problem in the presence of noise, but likewise do not provide a tractable algorithm to perform recovery. Work in [31,50,51] analyzes the detection and separation of two closely-spaced spikes, but does not generalize to the case when there are more than two spikes in the signal. Recent papers [39,40] analyze the problem of accurately recovering the *number* of spikes from the low-pass measurements.

Super-resolution of well-separated spikes Progress towards resolving the question posed in [23] in the general situation where $\mathbf{x} \in \mathbb{C}^N$ —in this paper we consider the case $\mathbf{x} \geq \mathbf{0}$ only—has been made in [12,13,28]. The sharpest from this series of results [28] implies the following. Assume $\text{supp}(\mathbf{x}) \in \mathcal{R}(1.26\lambda_{10}, 1)$, then the solution to ℓ_1 -minimization problem

$$\hat{\mathbf{x}} = \arg \min_{\tilde{\mathbf{x}}} \|\tilde{\mathbf{x}}\|_1 \quad \text{subject to} \quad \|\mathbf{s} - \mathbf{Q}\tilde{\mathbf{x}}\|_1 \leq \delta \quad (\text{L1})$$

with δ chosen so that $\|\mathbf{z}\|_1 \leq \delta$ satisfies

$$\|\mathbf{x} - \hat{\mathbf{x}}\|_1 \leq \tilde{c} \cdot \text{SRF}^2,$$

where \tilde{c} is a numerical constant. The requirement $\text{supp}(\mathbf{x}) \in \mathcal{R}(1.26\lambda_{10}, 1)$ (well-separated spikes in our terminology) is restrictive because it means that the signal \mathbf{x} cannot contain spikes that are at a distance less than $1.26\lambda_{10}$. This is a limitation for many applications including single-molecule microscopy, as it is usually understood that the goal of super-resolution is to distinguish spikes that are (significantly) closer than the Rayleigh diffraction limit, i.e., at a fraction of λ_{10} apart. Unfortunately, if there are spikes at a distance smaller than λ_{10} , ℓ_1 minimization does not, in general, return the correct solution even if there is no noise. The central question therefore is: which algorithms and under which assumptions are able to super-resolve signals *robustly* when the distance between some of the spikes may be *substantially smaller* than λ_{10} ?

On a similar line of research, see [56] and [57] for related results on the denoising of line spectra and on the recovery of sparse signals from a random subset of their low-pass Fourier coefficients. The accuracy of support detection for well-separated spikes is analyzed in [2,27].

Noise-aware algebraic methods Many noise-aware versions of Prony's method are used frequently in engineering applications, for example in radar (see [52, Ch. 6]). The most popular methods are MUSIC and its numerous variations [3,9,11,45,49,58], matrix-pencil [32], and ESPRIT [44,47]. For more details on algebraic methods we refer the reader to the excellent book [52, Ch. 4]. It is important to point out that unlike convex optimization based methods like (L1), algebraic methods do *not* need the spikes to be well-separated (\mathbf{x} may contain spikes closer than λ_{10}) even when the signal is complex-valued, at least in the noiseless case.

The stability of noise-aware algebraic methods is an active area of research. Asymptotic results (at high signal-to-noise ratio) on the stability of MUSIC in the presence of Gaussian noise are derived in [16,55]. More recently, some steps towards analyzing

MUSIC and matrix-pencil in a non-asymptotic regime have been taken in [38] and in [42], respectively.

Especially important is the question of stability of algebraic methods when the spikes are not well-separated. Substantial progress in understanding this for MUSIC and ESPRIT algorithms has been made by Li and Liao in the last two years [35–37]. See also [34] for a simplified exposition of ideas in [36] and some extensions. The authors considered a *separated cluster model* for spike locations; the model is similar to Rayleigh regularity in spirit, but is more restrictive. For example, the signals depicted in Fig. 2b and c are both Rayleigh-regular with $r = 2$ and $d = 4\lambda_{10}$. At the same time, the signal in Fig. 2b has spike clusters that are separated by about $3\lambda_{10}$, but the signal in Fig. 2c has spike clusters that are separated by *only* about $2\lambda_{10}$. The separation between clusters determines the stability guarantee of the reconstruction in the presence of noise. Hence, the theory based on the separated cluster model predicts that the signal in Fig. 2b is easier to super-resolve than the signal in Fig. 2c. Intuitively, it is not clear why this should be the case. The theory based on Rayleigh regularity provides identical stability guarantees for the signals in Fig. 2b and c. Continuing this argument, it is possible to construct examples when the guarantees based on separated cluster model will be weak, yet the guarantees based on Rayleigh-regularity will be strong; the examples where the reverse is true do not exist. For MUSIC in [35,36] and for ESPRIT in [37], assuming Gaussian noise and making a further (restrictive) assumption $f_{10} \gtrsim \|\mathbf{x}\|_0^2$, the authors derived bounds on signal-to-noise ratio in terms of SRF^{2r-2} and a factor that depend on f_{10} so that the correct signal support recovery is guaranteed. There is still a large gap between these stability estimates and the minimax converse results. For example, for ESPRIT, the gap is a factor proportional to f_{10} , which may be very large for high-dimensional signals [37]. Hence, the problem of finding a super-resolution method for complex-valued signals that performs well empirically and has sharp theoretical stability estimates in the case when the spikes are not well-separated is still open.

Super-resolution of nonnegative signals The case of nonnegative signal, $\mathbf{x} \geq \mathbf{0}$, was analyzed in [24], see also [30] for a shorter exposition of the same idea. It is proven in [24] that as long as $\|\mathbf{x}\|_0 \leq f_{10}$, one can recover \mathbf{x} by solving a simple convex feasibility problem in the noiseless setting. In the presence of noise, [24] does not provide sharp estimates: it does not reveal the correct SRF dependence in the stability estimate.

More recently, the authors of [48] generalized [24] to the case of more general point spread functions and sampling patterns in the noiseless case; further refinements have been obtained in [25]. The corresponding noisy case has been studied in [26]. Being very general, the results of [26] do not appear to be sharp enough to reveal the fundamental dependence between the stability of the algorithm, the regularity of the signal, and the super-resolution factor.

Most relevant to this work is the earlier paper [43] where Proposition 1 has been proven. The key question remained: what happens if the grid becomes arbitrarily fine or when there is no grid at all (the gridless setting). Some progress towards answering this question has since been made in [19] where stability estimates for the detection of signal support have been expressed in terms of $\text{SRF}^{2\|\mathbf{x}\|_0-1}$. Note that $\|\mathbf{x}\|_0$ may be arbitrarily large for high-dimensional signals, and so the bounds in [19] become

highly suboptimal for the practically relevant case in which the spikes are distributed in a regular way in the signal. It was shown in [18, Sect. 2.4] that the results of [19] may be generalized to a separated cluster model, in which case the $\|\mathbf{x}\|_0$ in the estimate above is substituted with r , the number of spikes in one cluster. However, the result in [18, Sect. 2.4, Theorem 4] requires the clusters to be far enough, and there is no quantification of how far the clusters need to be. In other words, the clusters need to be arbitrarily far away from one another. In contrast, the Rayleigh regularity concept in this paper specifies exactly how far the clusters need to be for stability estimates to hold and, further, this distance is tight to within a factor of 1.87 [43, Sect. 2.3.1]. Also see the discussion above explaining that Rayleigh regularity is generally a more forgiving requirement than the separated cluster model.

3.2 Innovations

The innovations in this paper may be summarized as follows:

- Generalization of the results of [43] to the case when the grid is arbitrarily fine.
- Seamless connection between the super-resolution results for the discrete grid and the results for the gridless (continuous) setting. This has theoretical as well as practical implications.
- Mathematically the paper builds on the ideas from [12] and [43] and develops these methods further. The interpolation constructions in Lemmas 4 and 5 are new. These constructions may be of independent interest and may be useful for other problems.

4 Notation

Sets are denoted by calligraphic letters \mathcal{A} , \mathcal{B} , and so on. Boldface letters \mathbf{A} , \mathbf{B} , \dots and \mathbf{a} , \mathbf{b} , \dots denote matrices and vectors, respectively. The element in the i th row and j th column of a matrix \mathbf{A} is a_{ij} or $[\mathbf{A}]_{i,j}$, and the i th element of a vector \mathbf{a} is a_i or $[\mathbf{a}]_i$. For a vector \mathbf{a} , $\text{diag}(\mathbf{a})$ stands for the diagonal matrix that has the entries of \mathbf{a} on its main diagonal. The vector of all zeros is denoted $\mathbf{0}$. The superscript \top stands for transposition. For $a, b \in \mathbb{T}$, the wrap-around distance between a and b is $|b - a| \triangleq \min(b - a \bmod 1, a - b \bmod 1)$; for an interval $[a, b] \subset \mathbb{T}$, its wrap-around length is $|b - a|$. For a finite set \mathcal{I} , we write $|\mathcal{I}|$ for the cardinality. Note that the notation $|\cdot|$ is overloaded. For $x \in \mathbb{R}$, $\lceil x \rceil \triangleq \min\{m \in \mathbb{Z} \mid m \geq x\}$. We use $[l:k]$ to designate the set of natural numbers $\{l, l+1, \dots, k\}$. For a vector $\mathbf{a} \in \mathbb{C}^n$, $\|\mathbf{a}\|_1 = \sum_{j=0}^{n-1} |a_j|$ denotes the ℓ_1 norm; $\|\mathbf{a}\|_2 = (\sum_{j=0}^{n-1} a_j^2)^{1/2}$ denotes the ℓ_2 norm; $\|\mathbf{a}\|_\infty = \max_j |a_j|$ denotes the ℓ_∞ norm; and $\|\mathbf{a}\|_0$ denotes the number of nonzero elements in \mathbf{a} . For a function $f(\cdot) : \mathbb{R} \rightarrow \mathbb{R}$, $\|f(\cdot)\|_\infty = \max_{t \in \mathbb{R}} |f(t)|$. The indicator function is denoted as $I[\cdot]$, it is equal to one if the condition in the brackets is satisfied and zero otherwise. We use c with various subindexes and superindexes to denote *positive numerical* constants; to track things simpler, we use the convention that the numerical constants with the subscript u , like c_{u1} , satisfy $c_{u1} > 1$, and the numerical constants with subscript l , like c_{l1} , satisfy $0 < c_{l1} < 1$. Throughout the paper we use the convention: f_{i0} denotes

the frequency cut-off of the measured data [see (4)], $\lambda_{10} = 1/f_{10}$ is the corresponding wavelength; f_c denotes an abstract frequency cut-off (this value changes in different places in the paper) and $\lambda_c = 1/f_c$ is the corresponding wavelength. To simplify writing, we follow the conventions: $\prod_{i=1}^r a_i = 1$ and $\{a_1, \dots, a_r\} = \emptyset$ when $r = 0$.

5 Structure of the Proof

Previous results in the field [12, 13, 43] suggest that Theorem 1 may be proven by constructing an appropriate dual certificate. Specifically, consider a compressed sensing type problem in which an unknown signal is being recovered from an incomplete set of noiseless linear measurements using a convex optimization procedure with prior constraints on the signal (such as nonnegativity) and regularization (such as ℓ_1 norm minimization); see, for example, [14] for a short exposition of the classical ideas. Then the dual certificate is a vector from the range of the adjoint of the measurement matrix that satisfies specific interpolation conditions determined by the properties of the true unknown signal (such as sparsity) and by the type of prior constraints and regularization used. Geometrically, the existence of the dual certificate guarantees that the null space of the measurement matrix is oriented in a favorable way: starting from the true unknown signal it is not possible to move along the null space of the measurement matrix while satisfying the prior constraints and making the solution more regularized. Therefore, the true unknown signal is the optimal solution of the convex optimization problem. In other words, the existence of the dual certificate guarantees that the convex optimization problem recovers the correct solution. In the presence of noise a similar approach allows one to prove stability of the reconstruction, but the dual certificate often must satisfy additional properties [43].

In our problem, since the measurement operator is a low-pass kernel, the dual certificate is a real-valued trigonometric polynomial frequency-limited to f_{10} with additional properties. In fact, since we work on an arbitrarily fine grid, similar to [12], we will need three trigonometric polynomials instead of one, each with its own properties; they will be called $q_0(\cdot)$, $q_1(\cdot)$, and $q_2(\cdot)$. These dual trigonometric polynomials are constructed in Lemmas 3, 4, and 5 in Sect. 6; $q_0(\cdot)$ is borrowed from [43], $q_1(\cdot)$ and $q_2(\cdot)$ are new—they are the main mathematical contribution of this paper. In Sect. 7 we use $q_0(\cdot)$, $q_1(\cdot)$, and $q_2(\cdot)$ to derive the stability estimates and prove Theorem 1.

We invite the reader unfamiliar with the concept of dual certificates in convex optimization or the use of dual certificates in super-resolution problems to study the short proof of [43, Lm. 1, pp. 426–427] before reading this paper further. The derivations in Sect. 7 generalize [43, Lm. 1] to the arbitrarily fine grid setting, but they are much more involved.

Some calculations in this paper are complicated, but we tried to present the key new ideas in a simple way. At the first pass through the paper we suggest that the reader studies Sects. 6.1–6.2; then focuses on the formulations of Lemmas 4 and 5 and the new constructions in Sects. 6.3.1 and in Appendix “Proof of Lemma 5-construction” skips the details in Sects. 6.3.2–6.3.6 and in the remainder of Appendix “Proof of Lemma 5” and finally studies the stability estimates in Sect. 7. After this, return to the technical details in Sects. 6.3.2–6.3.6 and in Appendix “Proof of Lemma 5”.

6 Dual Certificates

Throughout the paper we will use the following definitions. Define the error vector

$$\mathbf{h} = [h_0, \dots, h_{N-1}]^T \triangleq \hat{\mathbf{x}} - \mathbf{x}$$

and the set of points where the error vector takes on negative values

$$\mathcal{T} = \{t_1, \dots, t_S\} \triangleq \{m/N : h_m < 0\}. \tag{12}$$

The points are ordered according to $t_1 < \dots < t_S$. Recall, $\hat{\mathbf{x}} \geq \mathbf{0}$ and $\mathbf{x} \geq \mathbf{0}$. Therefore, h_m can only take on negative values on $\text{supp}(\mathbf{x})$, which implies $\mathcal{T} \subset \text{supp}(\mathbf{x})$. Since $\text{supp}(\mathbf{x}) \in \mathcal{R}(\kappa\lambda_{lo}r, r)$ and since the elements of $\text{supp}(\mathbf{x})$ are separated by at least $2\lambda_{hi}$, it follows $\mathcal{T} \in \mathcal{R}(\kappa\lambda_{lo}r, r)$ and the elements of \mathcal{T} are also separated by at least $2\lambda_{hi}$. As we will see below, the dual trigonometric polynomials $q_0(\cdot)$, $q_1(\cdot)$, and $q_2(\cdot)$ need to satisfy specific interpolation constraints on \mathcal{T} .

Throughout the paper we will use the following neighborhood notations.

Definition 2 For $\tau \in \mathbb{T}$, $\delta > 0$,

$$\mathcal{N}(\delta, \tau) \triangleq \{t \in \mathbb{T} : |t - \tau| \leq \delta\},$$

where $|\cdot|$ denotes the wrap-around distance on \mathbb{T} . Above, $\mathcal{N}(\cdot, \cdot)$ stands for ‘‘near’’ (i.e., the points near τ).

For a set $\mathcal{V} \subset \mathbb{T}$ and $\delta > 0$,

$$\mathcal{N}(\delta, \mathcal{V}) \triangleq \cup_{\tau \in \mathcal{V}} \mathcal{N}(\delta, \tau),$$

$$\mathcal{F}(\delta, \mathcal{V}) \triangleq \mathbb{T} \setminus \mathcal{N}(\delta, \mathcal{V}).$$

Above, $\mathcal{F}(\cdot, \cdot)$ stands for ‘‘far’’ (i.e., the points far from \mathcal{V}).

6.1 Building Blocks

The following two lemmas serve as common building blocks for the construction of trigonometric polynomials $q_0(\cdot)$, $q_1(\cdot)$, and $q_2(\cdot)$.

Lemma 1 allows us to construct a trigonometric polynomial frequency-limited to f_c that *interpolates zeros* at well-separated points as illustrated in Fig. 4a.

Lemma 1 Let $\lambda_c \in (0, 1/128)$, set $f_c \triangleq 1/\lambda_c$. Consider a collection of points $v_1 < v_2 < \dots < v_V$, define $\mathcal{V} \triangleq \{v_1, v_2, \dots, v_V\}$ and assume $\mathcal{V} \in \mathcal{R}(\kappa\lambda_c, 1)$. Then, there exists a real-valued trigonometric polynomial $q(\cdot) = q_{\lambda_c, \mathcal{V}}(\cdot)$ that satisfied the following properties.

1. Frequency limitation to f_c : $q(t) = \sum_{k=-f_c}^{f_c} \hat{q}_k e^{-i2\pi kt}$ for some $\hat{q}_k \in \mathbb{C}$.
2. Zero values and zero derivatives on \mathcal{V} : for all $v \in \mathcal{V}$, $q(v) = q'(v) = 0$.
3. Uniform confinement between zero and one: for all $\tau \in \mathbb{R}$, $0 \leq q(\tau) \leq 1$.

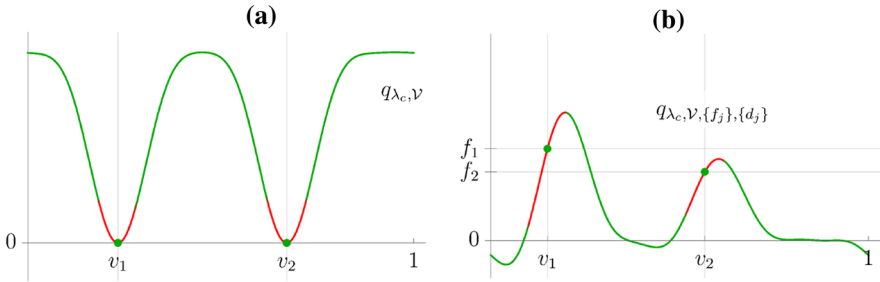


Fig. 4 **a** Illustration of Lemma 1. Trigonometric polynomial frequency-limited to $f_c = 6$ interpolates zeros at well-separated points $\{v_1, v_2\} \in \mathcal{R}(2.5\lambda_c, 1)$. Specifically, $q_{\lambda_c, \mathcal{V}}(v_j) = q'_{\lambda_c, \mathcal{V}}(v_j) = 0$ and the curvature in the neighborhoods of v_1 and v_2 is controlled (indicated in red) according to (13). **b** Illustration of Lemma 2. Trigonometric polynomial frequency-limited to $f_c = 6$ interpolates values f_1 and f_2 at well-separated points $\{v_1, v_2\} \in \mathcal{R}(2.5\lambda_c, 1)$. Specifically, $q_{\lambda_c, \mathcal{V}, \{f_j\}, \{d_j\}}(v_j) = f_j$ and the derivatives at v_1 and v_2 are constrained (indicated in red) according to $q'_{\lambda_c, \mathcal{V}, \{f_j\}, \{d_j\}}(v_j) = d_j$ (Color figure online)

4. *Quadratic behavior near \mathcal{V} : for all $v \in \mathcal{V}$ and for all $\tau \in \mathcal{N}(\Delta\lambda_c, v)$*

$$\frac{c_l(v - \tau)^2}{\lambda_c^2} \leq q(\tau) \leq \frac{c_u(v - \tau)^2}{\lambda_c^2}. \tag{13}$$

5. *Boundedness away from zero far from \mathcal{V} : for all $\tau \in \mathcal{F}(\Delta\lambda_c, \mathcal{V})$, $q(\tau) \geq c_{l1} > 0$.*

6. *Uniform confinement of the derivative: $\|q'(\cdot)\|_\infty \leq 2\pi/\lambda_c$.*

7. *Uniform confinement of the second derivative: $\|q''(\cdot)\|_\infty \leq 4\pi^2/\lambda_c^2$.*

Above, all the constants are positive numerical constants. Specifically,

$$\begin{aligned} \Delta &\triangleq 0.17, \quad c_l \triangleq 0.029 \\ c_u &\triangleq 2\pi^2, \quad c_{l1} \triangleq \Delta^2 c_l = 8.3 \times 10^{-4}. \end{aligned} \tag{14}$$

Proof This lemma is a direct consequence of the technique developed in [13]. Let $q_{CFG}(\cdot)$ be the trigonometric polynomial constructed as in [13, eq. (2.4)] to interpolate -1 on \mathcal{V} . Then, according to [13, Lm. 2.4, Lm. 2.5, Sec. 2.5], $q(\cdot) = 0.5(q_{CFG}(\cdot) + 1)$ satisfies Properties 1, 2, 3, 5 of the lemma, and the lower bound in (13). Since, by Property 3, $\|q(\cdot)\|_\infty \leq 1$, Properties 6 and 7 follow by applying (129) [Bernstein theorem]. Finally, the upper bound in (13) follows from Property 2 and Property 7 by (195) [Mean Value theorem]. \square

Lemma 2 allows us to construct a trigonometric polynomial frequency-limited to f_c that *interpolates arbitrary values* and has constrained derivatives at well-separated points as illustrated in Fig. 4b.

Lemma 2 *Let $\lambda_c \in (0, 1/128)$, set $f_c \triangleq 1/\lambda_c$. Consider a collection of points $v_1 < v_2 < \dots < v_V$, define $\mathcal{V} \triangleq \{v_1, v_2, \dots, v_V\}$ and assume $\mathcal{V} \in \mathcal{R}(\kappa\lambda_c, 1)$. Consider two sets of real numbers $\{f_1, f_2, \dots, f_V\}$ and $\{d_1, d_2, \dots, d_V\}$ that satisfy*

$$|f_j| \leq 1 \quad \text{and} \quad |d_j| \leq \frac{1}{\lambda_c} \tag{15}$$

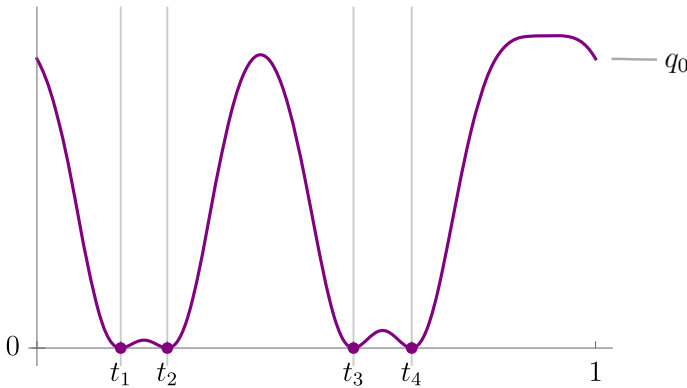


Fig. 5 Illustration of Lemma 3. Trigonometric polynomial frequency-limited to $f_{l_0} = 12$ interpolates zeros on Rayleigh-regular set $\mathcal{T} = \{t_1, t_2, t_3, t_4\} \in \mathcal{R}(5\lambda_{l_0}, 2)$ and bounces away from zeros “quickly”: the curvature in the neighborhoods of each point t_i is “high” in the sense of (17). In the figure, $|t_3 - t_1| \geq 5\lambda_{l_0} = 5/12$, $|t_4 - t_2| \geq 5\lambda_{l_0} = 5/12$, $|t_2 - t_1| \sim 2\lambda_{hi}$, $|t_3 - t_4| \sim 2\lambda_{hi}$ (Color figure online)

for all $j = 1, \dots, V$. Then, there exists a real-valued trigonometric polynomial $q(\cdot) = q_{\lambda_c, \mathcal{V}, \{f_j\}, \{d_j\}}(\cdot)$ that satisfies the following properties.

1. Frequency limitation to f_c : $q(t) = \sum_{k=-f_c}^{f_c} \hat{q}_k e^{-i2\pi kt}$ for some $\hat{q}_k \in \mathbb{C}$.
2. Constrained values and derivatives on \mathcal{V} : for all $j = 1, \dots, V$,

$$q(v_j) = f_j \text{ and } q'(v_j) = d_j.$$

3. Uniform confinement: $\|q(\cdot)\|_\infty \leq c_{u0}$.
4. Uniform confinement of the derivative: $\|q'(\cdot)\|_\infty \leq c_{u1}/\lambda_c$.
5. Uniform confinement of the second derivative: $\|q''(\cdot)\|_\infty \leq c_{u2}/\lambda_c^2$.

Above, c_{u0} , c_{u1} , and c_{u2} are positive numerical constants that are defined in the proof of the lemma in Appendix 1 “Proof of Lemma 2”.

The proof of the lemma generalizes the results in [13, Lm. 2.4, Lm. 2.5, Sec. 2.5] slightly in several technical aspects; it is given in Appendix 1 “Proof of Lemma 2”.

6.2 Dual Certificate $q_0(\cdot)$

We are now ready to construct the trigonometric polynomial $q_0(\cdot)$. This trigonometric polynomial, illustrated in Fig. 5, is frequency-limited to f_{l_0} , interpolates zeros on a Rayleigh-regular set, is confined between zero and one, and quickly grows around its zeros.

The key difference between the trigonometric polynomial $q_0(\cdot)$ and the building block $q_{\lambda_c, \mathcal{V}}(\cdot)$ constructed in Lemma 1 is that the points where $q_0(\cdot)$ must take zero values may belong to a Rayleigh-regular set from a class $\mathcal{R}(d, r)$ with $r > 1$. Zeros of $q_0(\cdot)$ may be close, whereas zeros of $q_{\lambda_c, \mathcal{V}}(\cdot)$ are well-separated (compare Figs. 4a to 5). This is the reason why the technique of [13] and [12] that was used to prove Lemma 1 cannot be applied directly to construct $q_0(\cdot)$.

Lemma 3 *There exists a real-valued trigonometric polynomial $q_0(\cdot)$ that satisfies the following properties.*

1. *Frequency limitation to f_{10} : $q_0(t) = \sum_{k=-f_{10}}^{f_{10}} \hat{q}_{0,k} e^{-i2\pi kt}$ for some $\hat{q}_{0,k} \in \mathbb{C}$.*
2. *Zero values and zero derivatives on \mathcal{T} : for all $t \in \mathcal{T}$, $q_0(t) = q'_0(t) = 0$.*
3. *Uniform confinement between zero and one: for all $\tau \in \mathbb{R}$, $0 \leq q_0(\tau) \leq 1$.*
4. *Controlled behavior near \mathcal{T} : Take $\tau \in \mathcal{N}(r\Delta\lambda_{10}, \mathcal{T})$, where $\Delta = 0.17$ as before. Let $\{v_1^\tau, \dots, v_{\hat{r}}^\tau\} \triangleq \mathcal{N}(r\Delta\lambda_{10}, \tau) \cap \mathcal{T}$. [Note: since $\mathcal{T} \in \mathcal{R}(r\kappa\lambda_{10}, r)$ and $\Delta < \kappa$, it follows that $1 \leq \hat{r} \leq r$.] Set $v^\tau \triangleq \arg \min_{v \in \{v_1^\tau, \dots, v_{\hat{r}}^\tau\}} |v - \tau|$. Then, the following estimates hold.*

(a) *Lower bound:*

$$q_0(\tau) \geq c_{l2}^r \frac{\prod_{l=1}^{\hat{r}} (v_l^\tau - \tau)^2}{(r\lambda_{10})^{2\hat{r}}} \tag{16}$$

$$\geq c_{l2}^r \frac{(v^\tau - \tau)^2 \lambda_{hi}^{2(r-1)}}{(r\lambda_{10})^{2r}}. \tag{17}$$

(b) *Upper bound:*

$$q_0(\tau) \leq c_u^{\hat{r}} \frac{\prod_{l=1}^{\hat{r}} (v_l^\tau - \tau)^2}{(r\lambda_{10})^{2\hat{r}}}. \tag{18}$$

5. *Boundedness away from zero far from \mathcal{T} : for all $\tau \in \mathcal{F}(r\Delta\lambda_{10}, \mathcal{T})$,*

$$q_0(\tau) \geq c_{l1}^r > 0. \tag{19}$$

6. *Fast growth immediately away from \mathcal{T} : for all $\tau \in \mathcal{F}(\lambda_{hi}, \mathcal{T})$,*

$$q_0(\tau) \geq c_l^r \frac{\lambda_{hi}^{2r}}{(r\lambda_{10})^{2r}}.$$

Above, c_{l2} is a positive numerical constant, defined in the proof below.

The trick to prove this lemma is the main contribution of the earlier paper [43]. The key observation is the following. It is possible to construct the nonnegative trigonometric polynomial $q_0(\cdot)$ frequency-limited to f_{10} that is zero on all the points of the set $\mathcal{T} \in \mathcal{R}(r\kappa\lambda_{10}, r)$ as a product of r trigonometric polynomials. Each of these trigonometric polynomials is zero on a set that belongs to $\mathcal{R}(\kappa\lambda_{10}r, 1)$ and is constructed via Lemma 1. We reproduce the proof below because it motivates the new construction in Sect. 6.3.

Proof Set

$$\mathcal{T}_k \triangleq \left\{ t_{jr+k} : j \in [0 : \lfloor (S-1)/r \rfloor] \right\}, \quad k = 1, \dots, r. \tag{20}$$

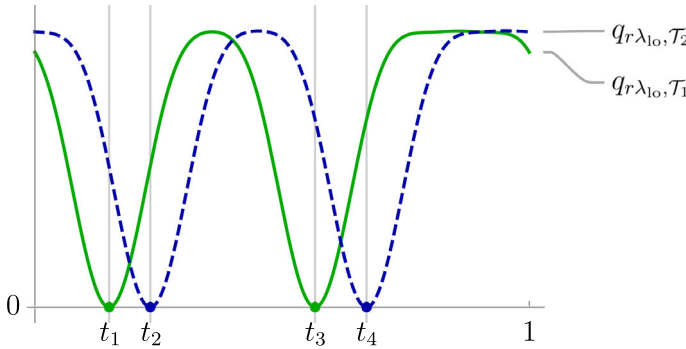


Fig. 6 Illustration of the proof of Lemma 3. The set $\mathcal{T} = \{t_1, t_2, t_3, t_4\}$ is Rayleigh-regular: $\mathcal{T} \in \mathcal{R}(5\lambda_{10}, 2)$, with $r = 2$ and $\lambda_{10} = 1/12$. The idea is to split this set as $\mathcal{T} = \mathcal{T}_1 \cup \mathcal{T}_2$ with $\mathcal{T}_1 = \{t_1, t_3\}$ and $\mathcal{T}_2 = \{t_2, t_4\}$ and observe $\mathcal{T}_i \in \mathcal{R}(5\lambda_{10}, 1)$. The trigonometric polynomials are frequency-limited to $f_{10}/2 = 6$ and satisfy the interpolation constraints $q_{r\lambda_{10}, \mathcal{T}_1}(t) = q'_{r\lambda_{10}, \mathcal{T}_1}(t) = 0$ for all $t \in \mathcal{T}_1$ and $q_{r\lambda_{10}, \mathcal{T}_2}(t) = q'_{r\lambda_{10}, \mathcal{T}_2}(t) = 0$ for all $t \in \mathcal{T}_2$. Then, $q_0(\cdot) = (q_{r\lambda_{10}, \mathcal{T}_1} \times q_{r\lambda_{10}, \mathcal{T}_2})(\cdot)$ satisfies $q_0(t) = q'_0(t) = 0$ for all $t \in \mathcal{T}$ and is frequency-limited to $2 \times f_{10}/2 = 12$. The trigonometric polynomial $q_0(\cdot)$ is displayed in Fig. 5. In the figure, $|t_2 - t_1| \sim 2\lambda_{hi}$, $|t_3 - t_4| \sim 2\lambda_{hi}$ (Color figure online)

Observe that $\mathcal{T} = \mathcal{T}_1 \cup \dots \cup \mathcal{T}_r$ and $\mathcal{T}_k \in \mathcal{R}(\kappa\lambda_{10}r, 1)$. Set

$$q_0(t) \triangleq q_{r\lambda_{10}, \mathcal{T}_1}(t) \times \dots \times q_{r\lambda_{10}, \mathcal{T}_r}(t), \tag{21}$$

where $q_{r\lambda_{10}, \mathcal{T}_k}(\cdot)$, $k = 1, \dots, r$, are the trigonometric polynomials constructed¹ via Lemma 1 with $\lambda_c = r\lambda_{10}$ and $\mathcal{V} = \mathcal{T}_k \in \mathcal{R}(\kappa\lambda_{10}r, 1)$. The idea of this construction for $r = 2$ is illustrated in Fig. 6.

It remains to verify that Properties 1–6 are satisfied. Broadly, this follows from (21) and Lemma 1; the details are given below.

Property 1 is satisfied because each of trigonometric polynomials $q_{r\lambda_{10}, \mathcal{T}_k}(\cdot)$, $k = 1, \dots, r$ is frequency-limited to f_{10}/r . Hence, the product in (21) is frequency-limited to $r(f_{10}/r) = f_{10}$.

Properties 2 and 3 follow from (21) and from Lemma 1, Properties 2 and 3, respectively.

To prove (16) we lower-bound the terms in (21) separately as follows. Assume that $k \in \{1, \dots, r\}$ is such that $\mathcal{N}(r\Delta\lambda_{10}, \tau) \cap \mathcal{T}_k \neq \emptyset$, i.e., there exist $l \in \{1, \dots, \hat{r}\}$ that satisfies $v_l^\tau \in \mathcal{T}_k$. In this case, we use the left-hand side of (13) to write

$$q_{r\lambda_{10}, \mathcal{T}_k}(\tau) \geq c_l \frac{(v_l^\tau - \tau)^2}{(r\lambda_{10})^2}. \tag{22}$$

¹ Strictly speaking this requires that the frequency limitation of $q_{r\lambda_{10}, \mathcal{T}_k}(\cdot)$, f_{10}/r , is an integer. In the rest of the paper, for simplicity, we will make this additional assumption. If this assumption is not satisfied, we can simply substitute f_{10} with $\lfloor f_{10}/r \rfloor r$ and repeat all the arguments in the paper, leading only to a small increase in the density constant 1.87 in Theorem 1.

Note that there are exactly \hat{r} such terms in (21). Assume that $k \in \{1, \dots, r\}$ is such that $\mathcal{N}(r\Delta\lambda_{10}, \tau) \cap \mathcal{T}_k = \emptyset$. In this case, use Lemma 1, Property 5, to write

$$q_{r\lambda_{10}, \mathcal{T}_k}(\tau) \geq c_{l1}. \tag{23}$$

Note that there are exactly $r - \hat{r}$ such terms in (21). The desired bound (16) is obtained by plugging (22) and (23) into (21) and setting $c_{l2} \triangleq \min(c_l, c_{l1})$.

Bound (17) follows because the elements of \mathcal{T} are separated by at least $2\lambda_{hi}$ and because $\lambda_{hi}/\lambda_{10} < 1$.

To prove (18) we upper-bound the terms in (21) separately as follows. Assume that $k \in \{1, \dots, r\}$ is such that $\mathcal{N}(r\Delta\lambda_{10}, \tau) \cap \mathcal{T}_k \neq \emptyset$, i.e., there exist $l \in \{1, \dots, \hat{r}\}$ that satisfies $v_l^\tau \in \mathcal{T}_k$. In this case, we use the right-hand side of (13) to write

$$q_{r\lambda_{10}, \mathcal{T}_k}(\tau) \leq c_u \frac{(v_l^\tau - \tau)^2}{(r\lambda_{10})^2}. \tag{24}$$

Assume that $k \in \{1, \dots, r\}$ is such that $\mathcal{N}(r\Delta\lambda_{10}, \tau) \cap \mathcal{T}_k = \emptyset$. In this case, we use Lemma 1, Property 3, to write

$$q_{r\lambda_{10}, \mathcal{T}_k}(\tau) \leq 1. \tag{25}$$

The desired bound (18) is obtained by plugging (24) and (25) into (21).

Property 5 follows by (21) and Lemma 1, Property 5.

Finally, Property 6 follows from (21), (13), Lemma 1, Property 5, and (14). \square

6.3 Dual Certificate $q_1(\cdot)$

We are now ready to construct the trigonometric polynomial $q_1(\cdot)$. This construction and its analysis is the main mathematical contribution of this paper. Trigonometric polynomial $q_1(\cdot)$, illustrated in Fig. 7, is frequency-limited to f_{10} and, on the points $t_j \in \mathcal{T}$, $q_1(\cdot)$ interpolates the set of signs

$$s_j \triangleq \text{sign} \left(\sum_{m/N \in \mathcal{N}(\lambda_{hi}, t_j)} h_m \right), \quad j = 1, \dots, S, \tag{26}$$

at a (low) level $\rho/2$, $\rho = (\lambda_{hi}/\lambda_{10})^{2r} \ll 1$. The behavior of $q_1(\cdot)$ is controlled by $q_0(\cdot)$ as explained in Lemma 4 below.

Lemma 4 *Set $\rho \triangleq \lambda_{hi}^{2r}/\lambda_{10}^{2r}$. Then, there exists a real-valued trigonometric polynomial $q_1(\cdot)$ that satisfies the following properties.*

1. *Frequency limitation to f_{10} : $q_1(t) = \sum_{k=-f_{10}}^{f_{10}} \hat{q}_{1,k} e^{-i2\pi kt}$ for some $\hat{q}_{1,k} \in \mathbb{C}$.*
2. *Constrained sign pattern (at level ρ) on \mathcal{T} and controlled behavior near \mathcal{T} : for all $j = 1, \dots, S$ and all $\tau \in \mathcal{N}(\lambda_{hi}, t_j)$,*

$$\left| q_1(\tau) - \frac{\rho s_j}{2} \right| \leq r^{2r+4} c_{u27}^{r+1} q_0(\tau), \tag{27}$$

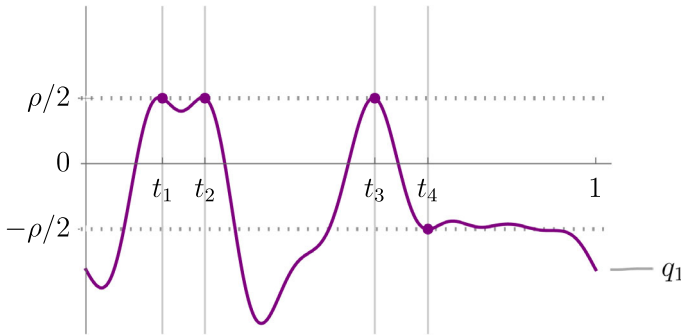


Fig. 7 Illustration of Lemma 4. Trigonometric polynomial frequency-limited to $f_{l_0} = 12$ interpolates the sign pattern $\{s_1, s_2, s_3, s_4\} = \{+1, +1, +1, -1\}$ at a (low) level $\rho/2$. Specifically, $q_1(t_i) = s_i \rho/2$ and $q'_1(t_i) = 0$. The set $\mathcal{T} = \{t_1, t_2, t_3, t_4\}$ is Rayleigh-regular: $\mathcal{T} \in \mathcal{R}(5\lambda_{l_0}, 2)$ with $\lambda_{l_0} = 1/f_{l_0}$ and $|t_2 - t_1| \sim 2\lambda_{hi}$, $|t_3 - t_4| \sim 2\lambda_{hi}$ (Color figure online)

where s_j are defined² in (26). Since $q_0(\tau) = 0$ for $\tau \in \mathcal{T}$, (27) implies, in particular, that $q_1(\cdot)$ interpolates the sign pattern in (26) at level $\rho/2$ on \mathcal{T} .

- 3. Uniform confinement: $\|q_1(\cdot)\|_\infty \leq r^{2r+1} c_{u55}^r$.
- 4. Boundedness far from \mathcal{T} : for all $\tau \in \mathcal{F}(\lambda_{hi}, \mathcal{T})$,

$$|q_1(\tau)| \leq r^{2r+2} c_{u27}^r c_{u29}^r q_0(\tau), \tag{28}$$

The positive numerical constants c_{u27} , c_{u55} , and c_{u29} are defined in the proof below.

Discussion Let's compare $q_1(\cdot)$ illustrated in Fig. 7 to $q_{\lambda_c, \mathcal{V}, \{f_j\}, \{d_j\}}(\cdot)$ constructed in Lemma 2 and illustrated in Fig. 4b. In $q_{\lambda_c, \mathcal{V}, \{f_j\}, \{d_j\}}(\cdot)$, the behavior at a well-separated set of points is independently controlled: the trigonometric polynomial can take arbitrary values (between -1 and 1). Reminder: we say that the points are well-separated if the distances between the points are no smaller than $\sim \tilde{c}/f_c$, where f_c is the frequency limitation of the trigonometric polynomial under consideration and \tilde{c} is a bit larger than 1 . In the case of $q_1(\cdot)$, the points where the behavior is controlled are *not* well-separated as illustrated on Fig. 7: $|t_2 - t_1| \sim 2\lambda_{hi} \ll 1/f_{l_0}$, $|t_3 - t_4| \sim 2\lambda_{hi} \ll 1/f_{l_0}$. Therefore, by Bernstein theorem (see Theorem 2), the behavior of $q_1(\cdot)$ at nearby points cannot be controlled *independently*. To be concrete: suppose we require that $q_1(t_1) = -1$ and $q_1(t_2) = +1$. Since the points t_1 and t_2 are separated by about $2\lambda_{hi} \ll 1/f_{l_0}$ (not well-separated), Bernstein theorem says that these two requirements cannot be satisfied simultaneously. Indeed, since $\|q_1(\cdot)\|_\infty \leq \tilde{C}(r) = r^{2r+1} c_{u55}$, by (129), $\|q'_1(\cdot)\|_\infty \leq 2\pi \tilde{C}(r) f_{l_0}$. If the two requirements had been satisfied simultaneously, the derivative of $q_1(\cdot)$ between the points t_1 and t_2 would have been about $(q_1(t_2) - q_1(t_1))/(2\lambda_{hi}) = 2/(2\lambda_{hi}) = f_{hi} \gg 2\pi \tilde{C}(r) f_{l_0}$ (we are assuming that SRF is large). However, if we require that $q_1(t_1) = -\rho$ and $q_1(t_2) = +\rho$ and ρ is small enough, the two requirements may be satisfied simultaneously. This is the reason why ρ is set to $\lambda_{hi}^{2r}/\lambda_{l_0}^{2r} \ll 1$ in the formulation of Lemma 4.

² The lemma is valid for an arbitrary sign pattern, we formulate it for the sign pattern defined in (26) for concreteness.

Let's compare $q_1(\cdot)$ to the trigonometric polynomial $q_0(\cdot)$ constructed in Lemma 3 and illustrated in Fig. 5. In both trigonometric polynomials the behavior is controlled on a Rayleigh-regular set, whose points are not well-separated in general. The difference is that $q_0(\cdot)$ takes *the same* value (zero) on all the points of the Rayleigh-regular set. This allows us to use the multiplication trick illustrated in Fig. 6 to prove Lemma 3. In the case of $q_1(\cdot)$ this does not work because we need to interpolate an *arbitrary* sign pattern on the Rayleigh-regular set. A method to resolve this problem, presented next, is the main mathematical contribution of this paper.

Proof Lemma 4 is proven in Sects. 6.3.1–6.3.6 below. □

6.3.1 Construction

We first describe how the trigonometric polynomial $q_1(\cdot)$ is constructed. In Sects. 6.3.2–6.3.6 we prove that the construction is valid and that it satisfies the required Properties 1–4.

Recall, $\mathcal{T} = \{t_1, \dots, t_S\}$ is defined in (12) and, as before, define $\mathcal{T}_k, k = 1, \dots, r$, as in (20); remember that $\mathcal{T} = \mathcal{T}_1 \cup \dots \cup \mathcal{T}_r$ and $\mathcal{T}_k \in \mathcal{R}(\kappa\lambda_{l_0}r, 1)$. Set $\eta_j = \rho(s_j + 1)/2$ for $j = 1, \dots, S$.

We will construct the trigonometric polynomial $q_1(\cdot)$ as a (shifted) *sum* of r trigonometric polynomials $\{\phi_k(\cdot)\}_{k=1}^r$ (see Fig. 8):

$$q_1(t) = \sum_{k=1}^r \phi_k(t) - \rho/2. \tag{29}$$

Each of the trigonometric polynomials $\{\phi_k(\cdot)\}_{k=1}^r$ is frequency-limited to f_{l_0} ,

$$\phi_k(t) = \sum_{l=-f_{l_0}}^{f_{l_0}} \hat{\phi}_{k,l} e^{-i2\pi lt} \quad \text{for some } \hat{\phi}_{k,l} \in \mathbb{C} \tag{30}$$

and is constructed separately to satisfy the following interpolation constraints on \mathcal{T} :

$$\phi_k(t_l) = \begin{cases} \eta_l, & \text{if } t_l \in \mathcal{T}_k, \\ 0, & \text{if } t_l \in \mathcal{T}_k^c \triangleq \mathcal{T} \setminus \mathcal{T}_k, \end{cases} \tag{31}$$

$$\phi'_k(t) = 0 \text{ for all } t \in \mathcal{T}. \tag{32}$$

Constraints (31), (32), and definition (29) guarantee that for all $l = 1, \dots, S$

$$q_1(t_l) = \rho s_l/2, \tag{33}$$

$$q'_1(t_l) = 0. \tag{34}$$

To develop intuition, observe that (30) and (29) guarantee that Property 1 is satisfied. Further, observe that the interpolation constraints (33) and (34) are needed for (27) to hold because $q_0(t) = q'_0(t) = 0$ for all $t \in \mathcal{T}$.

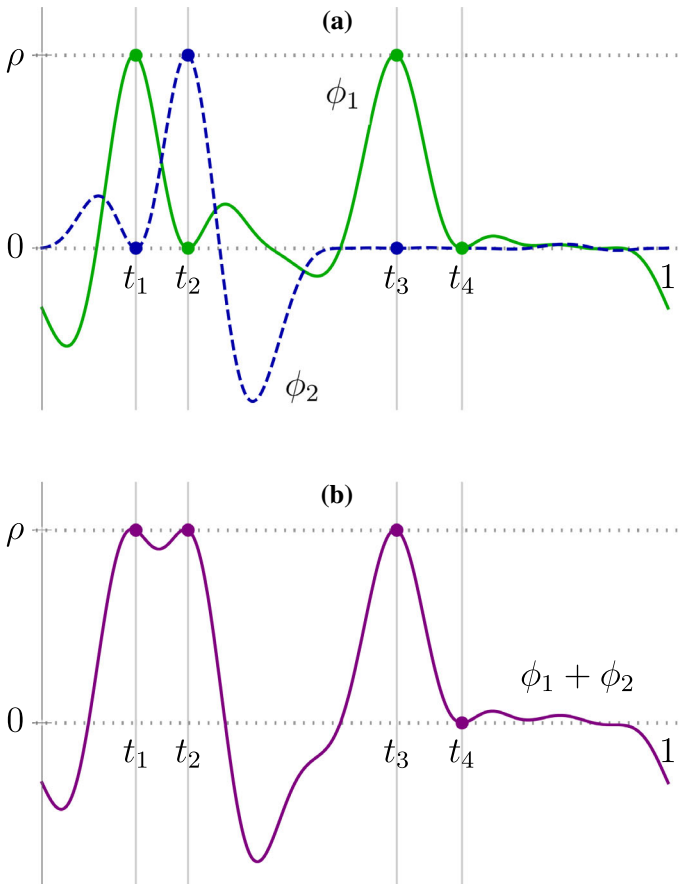


Fig. 8 Construction of the trigonometric polynomial $q_1(\cdot)$ (displayed in Fig. 7) with target sign pattern $\{s_1, s_2, s_3, s_4\} = \{+1, +1, +1, -1\}$. **a** trigonometric polynomials $\phi_1(\cdot)$ and $\phi_2(\cdot)$ satisfy interpolation constraints (31) and (32) as indicated by the points highlighted in bold. Specifically, $\phi_1(t_1) = \phi_1(t_3) = \rho$, $\phi_1(t_2) = \phi_1(t_4) = 0$, $\phi_2(t_1) = \phi_2(t_3) = \phi_2(t_4) = 0$ and $\phi_2(t_2) = \rho$; and further $\phi_i'(t_j) = 0$. **b** the sum of $\phi_1(\cdot)$ and $\phi_2(\cdot)$ that, after shifting down by $\rho/2$, is equal to $q_1(\cdot)$. In this figure, $\phi_1(\cdot)$ and $\phi_2(\cdot)$ are frequency-limited to $f_{lo} = 12$; $\mathcal{T} \in \mathcal{R}(5\lambda_{lo}, 2)$ with $\lambda_{lo} = 1/f_{lo}$ is represented as $\mathcal{T} = \mathcal{T}_1 \cup \mathcal{T}_2$ with $\mathcal{T}_1 = \{t_1, t_3\} \in \mathcal{R}(5\lambda_{lo}, 1)$, $\mathcal{T}_2 = \{t_2, t_4\} \in \mathcal{R}(5\lambda_{lo}, 1)$; $|t_2 - t_1| \sim 2\lambda_{hi}$, $|t_3 - t_4| \sim 2\lambda_{hi}$ (Color figure online)

For $r = 2$ the construction is illustrated in Fig. 8. The trigonometric polynomials $\phi_1(\cdot)$ and $\phi_2(\cdot)$ are displayed in Fig. 8a; they satisfy the interpolation constraints (31) and (32) as indicated by the points highlighted in bold. When we compute $(\phi_1 + \phi_2)(\cdot)$ we obtain the trigonometric polynomial displayed in Fig. 8b, which, when shifted down by $\rho/2$, is equal to the desired $q_1(\cdot)$ displayed in Fig. 7.

The difficulty remains: how to construct trigonometric polynomials $\phi_k(\cdot)$? Set

$$\mathcal{T}_k^0 \triangleq \left\{ t_{jr+k} : j \in [0: \lfloor (S-1)/r \rfloor] \text{ and } \eta_{jr+k} = 0 \right\}$$

and $\mathcal{T}_k^+ \triangleq \mathcal{T}_k \setminus \mathcal{T}_k^0$ for $k = 1, \dots, r$. The idea now is to construct $\phi_k(\cdot)$ as a *product* of two trigonometric polynomials (see Fig. 9):

$$\phi_k(t) \triangleq \phi_{0,k}(t) \times \phi_{+,k}(t). \tag{35}$$

The first term in the product is defined as

$$\phi_{0,k}(t) \triangleq \prod_{1 \leq l \leq r, l \neq k} q_{r\lambda_{l0}, \mathcal{T}_l}(t), \tag{36}$$

where $q_{r\lambda_{l0}, \mathcal{T}_l}(\cdot)$, $l = 1, \dots, r$, are the trigonometric polynomials constructed via Lemma 1 with $\lambda_c = r\lambda_{l0}$ and $\mathcal{V} = \mathcal{T}_l \in \mathcal{R}(\kappa\lambda_{l0}r, 1)$. Observe similarity to the trigonometric polynomial in (21); the difference is that the k th term is missing from the product.

The second term in the product,

$$\phi_{+,k}(t) \triangleq r^{2r} c_{u8}^r q_{r\lambda_{l0}, \mathcal{T}_k, \{f_j\}, \{d_j\}}(t) \tag{37}$$

is a (rescaled) trigonometric polynomial $q_{r\lambda_{l0}, \mathcal{T}_k, \{f_j\}, \{d_j\}}(\cdot)$ constructed via Lemma 2 with $\lambda_c = r\lambda_{l0}$, and $\mathcal{V} = \mathcal{T}_k \in \mathcal{R}(\kappa\lambda_{l0}r, 1)$ and c_{u8} is a positive numerical constant defined in (62) below. Further, the function-values and derivatives of $q_{r\lambda_{l0}, \mathcal{T}_k, \{f_j\}, \{d_j\}}(\cdot)$ are constrained on $\mathcal{T}_k = \mathcal{T}_k^0 \cup \mathcal{T}_k^+$ so that $\phi_{+,k}(\cdot)$ satisfies the following:

$$\phi_{+,k}(t) = \begin{cases} 0, & t \in \mathcal{T}_k^0, \\ \rho \frac{1}{\phi_{0,k}(t)}, & t \in \mathcal{T}_k^+, \end{cases} \tag{38}$$

$$\phi'_{+,k}(t) = \begin{cases} 0, & t \in \mathcal{T}_k^0, \\ -\rho \frac{\phi'_{0,k}(t)}{\phi_{0,k}^2(t)}, & t \in \mathcal{T}_k^+. \end{cases} \tag{39}$$

We will prove in Sect. 6.3.3 below, that this specification is valid, in the sense that the corresponding function values and derivatives of $q_{r\lambda_{l0}, \mathcal{T}_k, \{f_j\}, \{d_j\}}(\cdot)$ on \mathcal{T}_k satisfy requirements (15) of Lemma 2.

It follows from (35), (36), (38), Lemma 1, Properties 2, 4, and 5 that the interpolation constraint (31) is satisfied:

$$\begin{aligned} \phi_k(t) &= \underbrace{\phi_{0,k}(t)}_0 \phi_{+,k}(t) = 0 \text{ for all } t \in \mathcal{T}_k^c, \\ \phi_k(t) &= \phi_{0,k}(t) \underbrace{\phi_{+,k}(t)}_0 = 0 \text{ for all } t \in \mathcal{T}_k^0, \\ \phi_k(t) &= \phi_{0,k}(t) \phi_{+,k}(t) = \rho \text{ for all } t \in \mathcal{T}_k^+. \end{aligned}$$

Next, by (35),

$$\phi'_k(t) = \phi'_{0,k}(t) \phi_{+,k}(t) + \phi_{0,k}(t) \phi'_{+,k}(t).$$

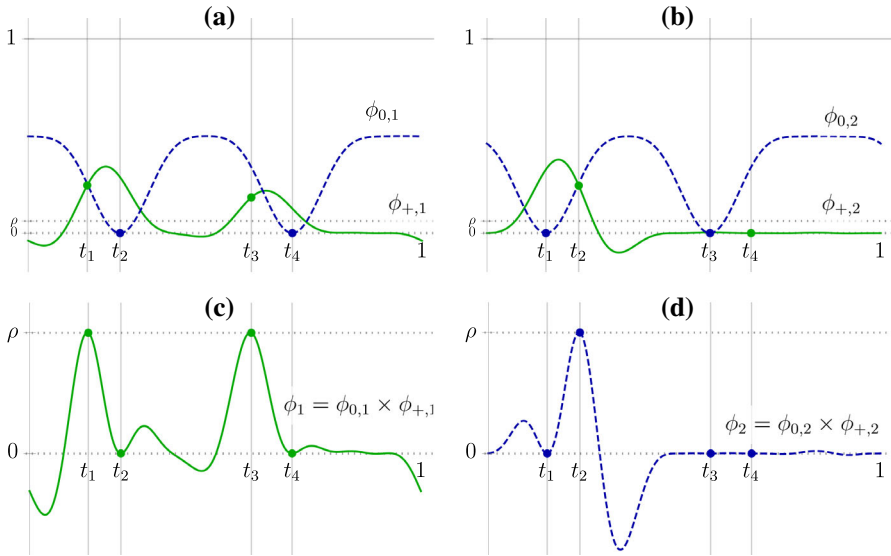


Fig. 9 Left column: constructing $\phi_1(\cdot)$ as a product of $\phi_{0,1}(\cdot)$ and $\phi_{+,1}(\cdot)$. The trigonometric polynomial $\phi_{0,1}(\cdot)$ is constrained to take zero values (bold blue points) on $\mathcal{T}_2 = \{t_2, t_4\}$ and it is strictly positive everywhere else. The function values of $\phi_{0,1}(\cdot)$ on $\mathcal{T}_1 = \{t_1, t_3\}$ are unconstrained. The trigonometric polynomial $\phi_{+,1}(\cdot)$, in turn, is only constrained on \mathcal{T}_1 (bold green points). In this case $\mathcal{T}_1 = \mathcal{T}_1^0 \cup \mathcal{T}_1^+$ with $\mathcal{T}_1^+ = \{t_1, t_3\}$ and $\mathcal{T}_1^0 = \emptyset$. The function values and derivatives of $\phi_{+,1}(\cdot)$ are constrained on \mathcal{T}_1 to “compensate” for the function values and derivatives of $\phi_{0,1}(\cdot)$ on \mathcal{T}_1 in the sense of (38) and (39). The compensation is such that once the two polynomials are multiplied we obtain $\phi_1(\cdot)$ with the local maxima at level ρ on \mathcal{T}_1 as shown in (c) (bold green points). The local minima of $\phi_1(\cdot)$ on \mathcal{T}_2 are produced “automatically”, because $\phi_{0,1}(\cdot)$ has zeros on \mathcal{T}_2 . Note that the function values and the derivatives of $\phi_{+,1}(\cdot)$ can be controlled at t_1 and t_3 independently, because these two points are well-separated and this would have been impossible if these points where not well-separated. **Right column:** constructing $\phi_2(\cdot)$ as a product of $\phi_{0,2}(\cdot)$ and $\phi_{+,2}(\cdot)$. The construction is similar, with the roles of \mathcal{T}_1 and \mathcal{T}_2 reversed. The difference is that in this case $\mathcal{T}_2 = \mathcal{T}_2^0 \cup \mathcal{T}_2^+$ with $\mathcal{T}_2^+ = \{t_2\}$ and $\mathcal{T}_2^0 = \{t_4\}$. Since \mathcal{T}_2^0 is nonempty, we set $\phi_{+,2}(t_4) = \phi'_{+,2}(t_4) = 0$. Finally: observe that the scale in **a** and **b** is different from the scale in **c** and **d**; the level ρ is marked for reference in **a** and **b** by a dotted line just above the zero line. The fact that $\rho = 1/\text{SRF}^{2r} \ll 1$ is responsible for the noise amplification. The setup is the same as in Figs. 7 and 8 (Color figure online)

Therefore, by (38), (39), Lemma 1, Properties 2, 4, and 5, the interpolation constraint (32) is satisfied:

$$\begin{aligned} \phi'_k(t) &= \underbrace{\phi'_{0,k}(t)}_0 \phi_{+,k}(t) + \underbrace{\phi_{0,k}(t)}_0 \phi'_{+,k}(t) = 0, \text{ for all } t \in \mathcal{T}_k^c, \\ \phi'_k(t) &= \phi'_{0,k}(t) \underbrace{\phi_{+,k}(t)}_0 + \phi_{0,k}(t) \underbrace{\phi'_{+,k}(t)}_0 = 0, \text{ for all } t \in \mathcal{T}_k^0, \\ \phi'_k(t) &= \phi'_{0,k}(t) \phi_{+,k}(t) + \phi_{0,k}(t) \phi'_{+,k}(t) \\ &= \phi'_{0,k}(t) \rho \frac{1}{\phi_{0,k}(t)} - \phi_{0,k}(t) \rho \frac{\phi'_{0,k}(t)}{\phi_{0,k}^2(t)} = 0, \text{ for all } t \in \mathcal{T}_k^+. \end{aligned}$$

Finally, (30) follows from (35) because $\phi_{0,k}(\cdot)$ in (36) is frequency-limited to $(r - 1)/(\lambda_{10}r)$ [Lemma 1, Property 1] and $\phi_{+,k}(\cdot)$ in (37) is frequency-limited to $1/(\lambda_{10}r)$ [Lemma 2, Property 1] so that $\phi_k(\cdot)$ is frequency-limited to $(r - 1)/(\lambda_{10}r) + 1/(\lambda_{10}r) = 1/\lambda_{10} = f_{10}$. Therefore, by (29), $q_1(\cdot)$ is also frequency-limited to f_{10} , which proves Property 1.

For $r = 2$ the construction is illustrated in Fig. 9. In Fig. 9a trigonometric polynomials $\phi_{0,1}(\cdot)$ and $\phi_{+,1}(\cdot)$ are displayed; $\phi_{0,1}(t) = 0$ for $t \in \mathcal{T}_2$ as indicated by the bold blue points; $\phi_{+,1}(\cdot)$ satisfies the interpolation constraints (38) and (39) on \mathcal{T}_1 as indicated by the bold green points. When we compute $\phi_1(\cdot) = (\phi_{0,1} \times \phi_{+,1})(\cdot)$ we obtain the trigonometric polynomial in Fig. 9c. The same process is displayed in Fig. 9b and d for $\phi_{0,2}(\cdot)$ and $\phi_{+,2}(\cdot)$. The trigonometric polynomials $\phi_1(\cdot)$ and $\phi_2(\cdot)$ in Fig. 9c and d are the same ones as in Fig. 8a.

6.3.2 Properties of $\phi_{0,k}(\cdot)$

We will now record useful properties of $\phi_{0,k}(\cdot)$ that are needed in the proof below. For $r = 1$, according to (36), $\phi_{0,k}(t) = 1$ for all t . For $r > 1$, the following properties hold.

1. Controlled behavior near \mathcal{T}_k^c : Take $\tau \in \mathcal{N}(r\Delta\lambda_{10}, \mathcal{T}_k^c)$. Let

$$\{v_1^\tau, \dots, v_{\hat{r}}^\tau\} \triangleq \mathcal{N}(r\Delta\lambda_{10}, \tau) \cap \mathcal{T}_k^c.$$

Note that since $\mathcal{T}_k^c \in \mathcal{R}(\kappa\lambda_{10}r, r - 1)$ and $\Delta < \kappa$, it follows that $1 \leq \hat{r} \leq r - 1$. Then, the following estimates hold.

- (a) Lower bound:

$$\phi_{0,k}(\tau) \geq c_{l_2}^{r-1} \frac{\prod_{l=1}^{\hat{r}} (v_l^\tau - \tau)^2}{(r\lambda_{10})^{2\hat{r}}}. \tag{40}$$

- (b) Upper bound:

$$\phi_{0,k}(\tau) \leq c_u^{\hat{r}} \frac{\prod_{l=1}^{\hat{r}} (v_l^\tau - \tau)^2}{(r\lambda_{10})^{2\hat{r}}}. \tag{41}$$

- (c) Upper bound on modulus of the first derivative:

$$\begin{aligned} |\phi'_{0,k}(\tau)| &\leq \sum_{m=1}^{\hat{r}} c_{u3}^{\hat{r}} \frac{\prod_{\substack{l=1 \leq l \leq \hat{r} \\ l \neq m}} (v_l^\tau - \tau)^2 |v_m^\tau - \tau|}{(r\lambda_{10})^{2\hat{r}}} \\ &\quad + (r - 1 - \hat{r}) c_{u3}^{\hat{r}+1} \frac{\prod_{l=1}^{\hat{r}} (v_l^\tau - \tau)^2}{(r\lambda_{10})^{2\hat{r}+1}}, \end{aligned} \tag{42}$$

where c_{u3} is a positive numerical constant defined in the proof below.

(d) Upper bound on modulus of the second derivative:

$$\begin{aligned}
 & |\phi''_{0,k}(\tau)| \\
 & \leq \sum_{1 \leq m \leq \hat{r}} \sum_{\substack{1 \leq m' \leq \hat{r} \\ m \neq m'}} I[\hat{r} \geq 2] c_{u3}^{\hat{r}} \frac{\prod_{\substack{1 \leq l \leq \hat{r} \\ l \neq m, l \neq m'}} (v_l^\tau - \tau)^2}{(r\lambda_{10})^{2(\hat{r}-2)}} \frac{|v_m^\tau - \tau|}{(r\lambda_{10})^2} \frac{|v_{m'}^\tau - \tau|}{(r\lambda_{10})^2} \\
 & + 2(r-1-\hat{r}) \sum_{1 \leq m \leq \hat{r}} c_{u3}^{\hat{r}+1} \frac{\prod_{\substack{1 \leq l \leq \hat{r} \\ l \neq m}} (v_l^\tau - \tau)^2}{(r\lambda_{10})^{2(\hat{r}-1)}} \frac{|v_m^\tau - \tau|}{(r\lambda_{10})^2} \frac{1}{r\lambda_{10}} \\
 & + (r-1-\hat{r})^2 c_{u3}^{\hat{r}+2} \frac{\prod_{l=1}^{\hat{r}} (v_l^\tau - \tau)^2}{(r\lambda_{10})^{2\hat{r}}} \frac{1}{(r\lambda_{10})^2} \\
 & + \sum_{1 \leq m \leq \hat{r}} c_{u3}^{\hat{r}} \frac{\prod_{\substack{1 \leq l \leq \hat{r} \\ l \neq m}} (v_l^\tau - \tau)^2}{(r\lambda_{10})^{2(\hat{r}-1)}} \frac{1}{(r\lambda_{10})^2} \\
 & + (r-1-\hat{r}) c_{u3}^{\hat{r}+1} \frac{\prod_{l=1}^{\hat{r}} (v_l^\tau - \tau)^2}{(r\lambda_{10})^{2\hat{r}}} \frac{1}{(r\lambda_{10})^2}. \tag{43}
 \end{aligned}$$

2. Boundedness away from zero far from \mathcal{T}_k^c : for all $\tau \in \mathcal{F}(r\Delta\lambda_{10}, \mathcal{T}_k^c)$,

$$\phi_{0,k}(\tau) \geq c_{l1}^{r-1} > 0. \tag{44}$$

3. Uniform confinement of the derivative:

$$\|\phi'_{0,k}(\cdot)\|_\infty \leq 2\pi/\lambda_{10}. \tag{45}$$

4. Uniform confinement of the second derivative:

$$\|\phi''_{0,k}(\cdot)\|_\infty \leq c_{u5}/\lambda_{10}^2. \tag{46}$$

5. Fast growth immediately away from \mathcal{T}_k^c : for all $\tau \in \mathcal{F}(\lambda_{hi}, \mathcal{T}_k^c)$,

$$\phi_{0,k}(\tau) \geq c_l^{r-1} \frac{\lambda_{hi}^{2(r-1)}}{(r\lambda_{10})^{2(r-1)}}. \tag{47}$$

Next, we give the proofs of the properties.

Proof of properties 1a–1b These properties are derived in the same way as Properties 4a and 4b in Lemma 3.

Proof of property 1c To prove (42), observe

$$|\phi'_{0,k}(\tau)| = \left| \left[\prod_{\substack{1 \leq m \leq r \\ m \neq k}} q_{r\lambda_{10}, \mathcal{T}_m}(\tau) \right]' \right| \leq \sum_{\substack{1 \leq m \leq r \\ m \neq k}} \left| \prod_{\substack{1 \leq j \leq r \\ j \neq k, j \neq m}} q_{r\lambda_{10}, \mathcal{T}_j}(\tau) \right| |q'_{r\lambda_{10}, \mathcal{T}_m}(\tau)|. \tag{48}$$

Above, we applied the chain rule for derivative to (36) and used the triangle inequality.

To upper-bound the sum in (48), we upper-bound quantities $|q_{r\lambda_{10}, \mathcal{T}_j}(\tau)|$ and $|q'_{r\lambda_{10}, \mathcal{T}_m}(\tau)|$ separately. To upper-bound $|q_{r\lambda_{10}, \mathcal{T}_j}(\tau)|$ we use the same bounds as in (24) and (25). To upper-bound $|q'_{r\lambda_{10}, \mathcal{T}_m}(\tau)|$ we use a similar strategy as follows. Assume that m is such that $\mathcal{N}(r\Delta\lambda_{10}, \tau) \cap \mathcal{T}_m \neq \emptyset$, i.e., there exist $l \in \{1, \dots, \hat{r}\}$ that satisfies $v_l^\tau \in \mathcal{T}_m$. In this case, according to Lemma 1, Property 2, $q'_{r\lambda_{10}, \mathcal{T}_m}(v_l^\tau) = 0$ and according to Lemma 1, Property 7, $|q''_{r\lambda_{10}, \mathcal{T}_m}(t)| \leq 4\pi^2/(r\lambda_{10})^2$ for all t . This, by (194) [Mean Value theorem], gives the following bound:

$$|q'_{r\lambda_{10}, \mathcal{T}_m}(\tau)| \leq c_{u2} \frac{|v_l^\tau - \tau|}{(r\lambda_{10})^2}. \tag{49}$$

Assume that m is such that $\mathcal{N}(r\Delta\lambda_{10}, \tau) \cap \mathcal{T}_m = \emptyset$. In this case, we use Lemma 1, Property 6, to write

$$|q'_{r\lambda_{10}, \mathcal{T}_m}(\tau)| \leq 2\pi \frac{1}{r\lambda_{10}}. \tag{50}$$

Plugging the estimates for $|q_{r\lambda_{10}, \mathcal{T}_j}(\tau)|$ [(24) and (25)], (49), and (50) into (48), setting $c_{u3} \triangleq \max(2\pi, 4\pi^2, c_u)$ we obtain (42).

Proof of property 1d To prove (43), observe

$$\begin{aligned} |\phi''_{0,k}(\tau)| &= \left| \left[\prod_{\substack{1 \leq m \leq r \\ m \neq k}} q_{r\lambda_{10}, \mathcal{T}_m}(\tau) \right]'' \right| \\ &\leq \sum_{\substack{1 \leq m \leq r \\ m \neq k}} \sum_{\substack{1 \leq m' \leq r \\ m' \neq k, m' \neq m}} \left| \prod_{\substack{1 \leq j \leq r \\ j \neq k, j \neq m, j \neq m'}} q_{r\lambda_{10}, \mathcal{T}_j}(\tau) \right| |q'_{r\lambda_{10}, \mathcal{T}_m}(\tau)| |q'_{r\lambda_{10}, \mathcal{T}_{m'}}(\tau)| \\ &\quad + \sum_{\substack{1 \leq m \leq r \\ m \neq k}} \left| \prod_{\substack{1 \leq j \leq r \\ j \neq k, j \neq m}} q_{r\lambda_{10}, \mathcal{T}_j}(\tau) \right| |q''_{r\lambda_{10}, \mathcal{T}_m}(\tau)|. \end{aligned} \tag{51}$$

To upper-bound the sum in (51), we upper-bound the quantities $|q_{r\lambda_{10}, \mathcal{T}_j}(\tau)|$, $|q'_{r\lambda_{10}, \mathcal{T}_m}(\tau)|$, and $|q''_{r\lambda_{10}, \mathcal{T}_m}(\tau)|$ separately. To upper-bound $|q_{r\lambda_{10}, \mathcal{T}_j}(\tau)|$ and $|q'_{r\lambda_{10}, \mathcal{T}_m}(\tau)|$ we use estimates (24), (25) and (49), (50), respectively. To upper-bound $|q''_{r\lambda_{10}, \mathcal{T}_m}(\tau)|$ we use Lemma 1, Property 7, to write

$$|q''_{r\lambda_{10}, \mathcal{T}_m}(\tau)| \leq 4\pi^2 \frac{1}{(r\lambda_{10})^2}.$$

Plugging these estimates into (51), we obtain (43).

Proof of properties 2–5 Property 2 follows by (36) and Lemma 1, Property 5. Property 3 follow from (48) and from Lemma 1, Property 6:

$$|\phi'_{0,k}(\tau)| \leq (r - 1) \frac{2\pi}{r\lambda_{10}} < \frac{2\pi}{\lambda_{10}}.$$

Property 4 follow from (51) and from Lemma 1, Properties 6 and 7:

$$|\phi''_{0,k}(\tau)| \leq (r - 1)(r - 2) \frac{4\pi^2}{(r\lambda_{10})^2} + (r - 1) \frac{4\pi^2}{(r\lambda_{10})^2} < \frac{8\pi^2}{\lambda_{10}^2},$$

where we defined $c_{u5} \triangleq 8\pi^2$. Finally, Property 5 follows from (36), (13), Lemma 1, Property 5, and (14).

6.3.3 Existence of $\phi_{+,k}(\cdot)$

In this subsection, we check that trigonometric polynomial $\phi_{+,k}(\cdot)$ that satisfies (38) and (39) can indeed be defined according to (37) with $q_{r\lambda_{10}, \mathcal{T}_k, \{f_j\}, \{d_j\}}$ constructed via Lemma 2 with $\lambda_c = r\lambda_{10}$ and $\mathcal{V} = \mathcal{T}_k \in \mathcal{R}(\kappa\lambda_{10}r, 1)$. To this end, we need to show that the constraints on the function values $\{f_j\}$ and on the derivatives $\{d_j\}$ that are implied by the constraints (38) and (39) satisfy requirements (15) of Lemma 2.

First consider the case $r = 1$. As already discussed, in this case $\phi_{0,k}(t) = 1$ for all t , and, therefore, $\phi'_{0,k}(t) = 0$ for all t . Plugging these values into (38) and (39) we see from (37) that the requirements (15) of Lemma 2 are satisfied.

Next, consider the case $r > 1$.

For $t \in \mathcal{T}_k^0$, by (37), (38), (39), $q_{r\lambda_{10}, \mathcal{T}_k}(t) = q'_{r\lambda_{10}, \mathcal{T}_k}(t) = 0$ so that requirements (15) of Lemma 2, are satisfied.

To check that requirements (15) are also satisfied for $t \in \mathcal{T}_k^+$, we need to find upper bounds on $|\phi_{+,k}(\cdot)|$ and $|\phi'_{+,k}(\cdot)|$.

Take $t \in \mathcal{T}_k^+$ and observe:

$$|\phi_{+,k}(t)| \stackrel{(a)}{=} \left| \rho \frac{1}{\phi_{0,k}(t)} \right| \stackrel{(b)}{\leq} \frac{\lambda_{hi}^{2r}}{\lambda_{lo}^{2r}} \frac{1}{\frac{\lambda_{hi}^{2(r-1)}}{(r\lambda_{10})^{2(r-1)}}} \frac{1}{c_l^{r-1}} \stackrel{(c)}{\leq} r^{2(r-1)} \frac{1}{c_l^r} \frac{\lambda_{hi}^2}{\lambda_{lo}^2} \tag{52}$$

$$\leq r^{2r} \frac{1}{c_l^r}. \tag{53}$$

Above, (a) follows by (38); (b) follows by (47) which is valid because $t \in \mathcal{T}_k^+$ implies $t \in \mathcal{F}(\lambda_{hi}, \mathcal{T}_k^c)$; (c) follows because $c_l < 1$.

Next, take $t \in \mathcal{T}_k^+$ and observe, according to (39),

$$|\phi'_{+,k}(t)| = \left| \rho \frac{\phi'_{0,k}(t)}{\phi_{0,k}^2(t)} \right|. \tag{54}$$

Consider two cases.

Case 1: $t \in \mathcal{F}(r\Delta\lambda_{10}, \mathcal{T}_k^c)$. Then, by (44), $\phi_{0,k}(t) \geq c_{l1}^{r-1}$, and, by (45), $|\phi'_{0,k}(t)| \leq 2\pi/\lambda_{10}$. Plugging these estimates into (54) we obtain

$$|\phi'_{+,k}(t)| \leq \left| \rho \frac{2\pi}{c_{l1}^{r-1}} \frac{1}{\lambda_{10}} \right| \stackrel{(a)}{\leq} \frac{2\pi}{c_{l1}^{2r-2}} \frac{\lambda_{hi}}{\lambda_{lo}^2} \stackrel{(b)}{<} r^{2r-1} \left(\frac{2\pi}{c_{l1}^2} \right)^r \frac{\lambda_{hi}}{\lambda_{lo}^2}. \tag{55}$$

Above, in (a) we used $\rho = (\lambda_{hi}/\lambda_{10})^{2r} \leq \lambda_{hi}/\lambda_{10}$; (b) is a crude inequality where we used $c_{l1} < 1$.

Case 2: $t \in \mathcal{N}(r\Delta\lambda_{10}, \mathcal{T}_k^c)$. In this case set $\{v_1, \dots, v_{\hat{r}}\} \triangleq \mathcal{T}_k^c \cap \mathcal{N}(r\Delta\lambda_{10}, t)$ and note $1 \leq \hat{r} \leq r - 1$. Hence, by (40):

$$\phi_{0,k}(t) \geq c_{l2}^{r-1} \frac{\prod_{j=1}^{\hat{r}} (v_j - t)^2}{(r\lambda_{10})^{2\hat{r}}}. \tag{56}$$

By (42):

$$|\phi'_{0,k}(t)| \leq \sum_{m=1}^{\hat{r}} c_{u3}^{\hat{r}} \frac{\prod_{\substack{1 \leq l \leq \hat{r} \\ l \neq m}} (v_l - t)^2 |v_m - t|}{(r\lambda_{10})^{2\hat{r}}} + (r - 1 - \hat{r}) c_{u3}^{\hat{r}+1} \frac{\prod_{l=1}^{\hat{r}} (v_l - t)^2}{(r\lambda_{10})^{2\hat{r}+1}}. \tag{57}$$

Plugging (56) and (57) into (54):

$$\begin{aligned} \frac{|\phi'_{0,k}(t)|}{\phi_{0,k}^2(t)} &\leq \sum_{m=1}^{\hat{r}} \frac{c_{u3}^{\hat{r}}}{c_{l2}^{2(r-1)}} \frac{(r\lambda_{10})^{2\hat{r}}}{\prod_{\substack{1 \leq j \leq \hat{r} \\ j \neq r}} (v_j - t)^2 |v_m - t|^3} \\ &\quad + (r - 1 - \hat{r}) \frac{c_{u3}^{\hat{r}+1}}{c_{l2}^{2(r-1)}} \frac{(r\lambda_{10})^{2\hat{r}-1}}{\prod_{j=1}^{\hat{r}} (v_j - t)^2} \\ &\stackrel{(a)}{\leq} r^{2\hat{r}+1} \frac{c_{u3}^{\hat{r}+1}}{c_{l2}^{2(r-1)}} \left(\frac{\lambda_{lo}^{2\hat{r}}}{\lambda_{hi}^{2\hat{r}+1}} + \frac{\lambda_{lo}^{2\hat{r}-1}}{\lambda_{hi}^{2\hat{r}}} \right) \\ &\stackrel{(b)}{\leq} r^{2r-1} \frac{c_{u3}^r}{c_{l2}^{2(r-1)}} \left(\frac{\lambda_{lo}^{2r-2}}{\lambda_{hi}^{2r-1}} + \frac{\lambda_{lo}^{2r-3}}{\lambda_{hi}^{2r-2}} \right) \end{aligned}$$

$$\stackrel{(c)}{\leq} 2r^{2r-1} \left(\frac{c_{u3}}{c_{l2}^2} \right)^r \frac{\lambda_{l0}^{2r-2}}{\lambda_{hi}^{2r-1}} \stackrel{(d)}{\leq} r^{2r-1} c_{u6}^r \frac{\lambda_{l0}^{2r-2}}{\lambda_{hi}^{2r-1}}. \tag{58}$$

Above, in (a) we used that $|v_j - t| \geq 2\lambda_{hi}$ for all $j = 1, \dots, \hat{r}, r - 1 - \hat{r} \leq r$, and $c_{u3} > 1$; in (b) we used that $\hat{r} \leq r - 1$, $\lambda_{l0}/\lambda_{hi} > 1$, and $c_{u3} > 1$; in (c) we used $\lambda_{l0}/\lambda_{hi} > 1$ and $c_{l2} < 1$; in (d) we defined $c_{u6} \triangleq 2c_{u3}/c_{l2}^2$. Plugging the estimate (58) into (54),

$$|\phi'_{+,k}(t)| \leq r^{2r-1} c_{u6}^r \frac{\lambda_{hi}^{2r} \lambda_{l0}^{2r-2}}{\lambda_{l0}^{2r} \lambda_{hi}^{2r-1}} = r^{2r-1} c_{u6}^r \frac{\lambda_{hi}}{\lambda_{l0}^2}. \tag{59}$$

Combining (55) and (59) we find that for all $t \in \mathcal{T}_k^+$,

$$|\phi'_{+,k}(t)| \leq r^{2r-1} c_{u7}^r \frac{\lambda_{hi}}{\lambda_{l0}^2} \tag{60}$$

$$\leq r^{2r-1} c_{u7}^r \frac{1}{\lambda_{l0}}, \tag{61}$$

where we defined $c_{u7} \triangleq \max(c_{u6}, c_{u1}/c_{l1}^2)$.

It follows from (53) and (61) that the function values and derivatives of $q_{r\lambda_{l0}, \mathcal{T}_k}(t) = \phi_{+,k}(t)/(r^{2r} c_{u8}^r)$ with

$$c_{u8} \triangleq \max(c_{u7}, 1/c_l) \tag{62}$$

satisfy requirements (15) of Lemma 2 on \mathcal{T}_k^+ . We conclude that $\phi_{+,k}(\cdot)$ can indeed be defined according to (37). According to Properties 3, 4, and 5 of Lemma 2, and (37), $\phi_{+,k}(\cdot)$ satisfies the following properties:

$$\|\phi_{+,k}(\cdot)\|_\infty \leq r^{2r} c_{u8}^r c_{u0}, \tag{63}$$

$$\|\phi'_{+,k}(\cdot)\|_\infty \leq r^{2r-1} c_{u8}^r c_{u1} \frac{1}{\lambda_{l0}}, \tag{64}$$

$$\|\phi''_{+,k}(\cdot)\|_\infty \leq r^{2r-2} c_{u8}^r c_{u2} \frac{1}{\lambda_{l0}^2}. \tag{65}$$

6.3.4 Proof of Property 2

Take $j \in \{1, \dots, S\}$ and consider $t_j \in \mathcal{T}$. There exists a unique $l \in \{1, \dots, r\}$ such that $t_j \in \mathcal{T}_l$. We will show that for all $\tau \in \mathcal{N}(\lambda_{hi}, t_j)$

$$|\phi_l(\tau) - \eta_j| \leq r^{2r+3} c_{u24}^{r+1} q_0(\tau) \tag{66}$$

and

$$|\phi_k(\tau)| \leq r^{2r+3} c_{u26}^{r+1} q_0(\tau), \text{ for } k \in \{1, \dots, r\}, k \neq l, \tag{67}$$

where the positive numerical constants c_{u24} and c_{u26} are defined below.

From this we will conclude that

$$\begin{aligned} \left|q_1(\tau) - \frac{\rho S_j}{2}\right| &= \left|\sum_{k=1}^r \phi_k(\tau) - \frac{\rho}{2} - \frac{\rho S_j}{2}\right| \\ &= \left|\sum_{k=1}^r \phi_k(\tau) - \eta_j\right| \\ &\leq \sum_{k \neq l}^r |\phi_k(\tau)| + |\phi_l(\tau) - \eta_j| \leq r^{2r+4} c_{u27}^{r+1} q_0(\tau) \end{aligned}$$

with $c_{u27} \triangleq 2 \max(c_{u24}, c_{u26})$, as desired.

To prove (66) and (67), recall, by (31) and (32):

$$|\phi_l(t_j) - \eta_j| = 0 = q_0(t_j), \tag{68}$$

$$\phi_k(t_j) = 0 = q_0(t_j), \text{ for } k \in \{1, \dots, r\}, k \neq l, \tag{69}$$

$$\phi'_k(t_j) = 0 = q'_0(t_j), \text{ for } k \in \{1, \dots, r\}. \tag{70}$$

Hence, in order to prove the bounds in (67) and (66), we will derive upper bounds on the second derivatives $|\phi''_k(\tau)|$, $k \in \{1, \dots, r\}$, valid for all $\tau \in \mathcal{N}(\lambda_{hi}, t_j)$ and use the Mean Value theorem (see Theorem 3).

Taking the second derivative of (35) and applying the triangle inequality we find:

$$|\phi''_k(\tau)| \leq \underbrace{|\phi''_{0,k}(\tau)| |\phi_{+,k}(\tau)|}_{E_1(\tau)} + 2 \underbrace{|\phi'_{0,k}(\tau)| |\phi'_{+,k}(\tau)|}_{E_2(\tau)} + \underbrace{|\phi_{0,k}(\tau)| |\phi''_{+,k}(\tau)|}_{E_3(\tau)}. \tag{71}$$

In the derivation below we upper-bound the terms separately.

We will need the following notations. Set $\{v_1^\tau, \dots, v_{\tilde{r}}^\tau\} \triangleq \mathcal{N}(r\Delta\lambda_{lo}, \tau) \cap \mathcal{T}_k^c$ and set $\{v_1, \dots, v_{\tilde{r}}\} \triangleq \mathcal{N}(r\Delta\lambda_{lo} - \lambda_{hi}, t_j) \cap \mathcal{T}_k^c$. Note that the set $\{v_1, \dots, v_{\tilde{r}}\}$ does not depend on τ and also $\{v_1, \dots, v_{\tilde{r}}\} \subset \{v_1^\tau, \dots, v_{\tilde{r}}^\tau\}$ so that $\tilde{r} \leq \hat{r}$.

The remainder of the proof of Property 2 is organized as follows. First, consider the case $t_j \in \mathcal{T}_k$ and prove (66), next consider the case $t_j \in \mathcal{T}_k^c$ and prove (67).

Proof of (66): case $t_j \in \mathcal{T}_k$. Bounding $E_1(\tau)$: By (195) [Mean Value theorem] and the triangle inequality we can write

$$|\phi_{+,k}(\tau)| \leq |\phi_{+,k}(t_j)| + |\phi'_{+,k}(t_j)| |\tau - t_j| + \frac{1}{2} |\phi''_{+,k}(\tau_m)| (\tau - t_j)^2$$

with $\tau_m \in (t_j, \tau)$. Next, we use (38) and (52) to upper-bound $|\phi_{+,k}(t_j)|$ by the right-hand side of (52); use (39) and (60) to upper bound $|\phi'_{+,k}(t_j)|$ by the right-hand side of (60); use (65) to upper-bound $|\phi''_{+,k}(\tau_m)|$. With these estimates we can further

upper-bound $|\phi_{+,k}(\tau)|$ as follows:

$$|\phi_{+,k}(\tau)| \leq r^{2r-1} c_{u9}^r \left(\frac{\lambda_{hi}^2}{\lambda_{lo}^2} + \frac{\lambda_{hi}}{\lambda_{lo}^2} |\tau - t_j| + \frac{1}{\lambda_{lo}^2} (\tau - t_j)^2 \right) \leq r^{2r-1} c_{u10}^r \frac{\lambda_{hi}^2}{\lambda_{lo}^2}. \tag{72}$$

Above, we defined $c_{u9} \triangleq \max(1/c_l, c_{u7}, c_{u8}c_{u2})$, $c_{u10} \triangleq 3c_{u9}$, and used $|\tau - t_j| \leq \lambda_{hi}$.

Assume $\tilde{r} \geq 1$ (the case $\tilde{r} = 0$ will be treated separately below) so that $\hat{r} \geq 1$ and $\tau \in \mathcal{N}(r\Delta\lambda_{lo}, \mathcal{T}_k^c)$, which implies that we can use (43) to upper-bound $|\phi''_{0,k}(\tau)|$:

$$\begin{aligned} |\phi''_{0,k}(\tau)| &\leq \sum_{1 \leq m \leq \hat{r}} \sum_{\substack{1 \leq m' \leq \hat{r} \\ m \neq m'}} I[\hat{r} \geq 2] c_{u3}^{\hat{r}} \frac{\prod_{\substack{1 \leq l \leq \hat{r} \\ l \neq m, l \neq m'}} (v_l^\tau - \tau)^2}{(r\lambda_{lo})^{2(\hat{r}-2)}} \frac{|v_m^\tau - \tau|}{(r\lambda_{lo})^2} \frac{|v_{m'}^\tau - \tau|}{(r\lambda_{lo})^2} \\ &+ 2(r-1-\hat{r}) \sum_{1 \leq m \leq \hat{r}} c_{u3}^{\hat{r}+1} \frac{\prod_{\substack{1 \leq l \leq \hat{r} \\ l \neq m}} (v_l^\tau - \tau)^2}{(r\lambda_{lo})^{2(\hat{r}-1)}} \frac{|v_m^\tau - \tau|}{(r\lambda_{lo})^2} \frac{1}{r\lambda_{lo}} \\ &+ (r-1-\hat{r})^2 c_{u3}^{\hat{r}+2} \frac{\prod_{l=1}^{\hat{r}} (v_l^\tau - \tau)^2}{(r\lambda_{lo})^{2\hat{r}}} \frac{1}{(r\lambda_{lo})^2} \\ &+ \sum_{1 \leq m \leq \hat{r}} c_{u3}^{\hat{r}} \frac{\prod_{\substack{1 \leq l \leq \hat{r} \\ l \neq m}} (v_l^\tau - \tau)^2}{(r\lambda_{lo})^{2(\hat{r}-1)}} \frac{1}{(r\lambda_{lo})^2} \\ &+ (r-1-\hat{r}) c_{u3}^{\hat{r}+1} \frac{\prod_{l=1}^{\hat{r}} (v_l^\tau - \tau)^2}{(r\lambda_{lo})^{2\hat{r}}} \frac{1}{(r\lambda_{lo})^2}. \end{aligned} \tag{73}$$

Multiplying (72) and (73) and simplifying we obtain the following upper bound on E_1 :

$$\begin{aligned} E_1(\tau) = |\phi''_{0,k}(\tau)| |\phi_{+,k}(\tau)| &\stackrel{(a)}{\leq} r^{2r+1} c_{u11}^{r+1} \frac{\prod_{1 \leq l \leq \hat{r}} (v_l^\tau - \tau)^2}{(r\lambda_{lo})^{2\hat{r}}} \frac{1}{\lambda_{lo}^2} \\ &\stackrel{(b)}{\leq} r^{2r+1} c_{u11}^{r+1} \frac{\prod_{1 \leq l \leq \hat{r}} (v_l - \tau)^2}{(r\lambda_{lo})^{2\hat{r}}} \frac{1}{\lambda_{lo}^2}. \end{aligned} \tag{74}$$

Above, in (a) we used (multiple times) the bound $\lambda_{hi} \leq |v_l^\tau - \tau|$, which is true for all $l \in \{1, \dots, \hat{r}\}$ (follows because the elements of \mathcal{T} are separated by at least $2\lambda_{hi}$), used $\lambda_{hi}/\lambda_{lo} < 1$, and defined $c_{u11} \triangleq \max(6c_{u3}c_{u10}, c_{u8}c_{u0}c_{u5})$; in (b) we used the fact that $|v_l^\tau - \tau|/(\lambda_{lo}r) \leq \Delta < 1$ for all $l \in \{1, \dots, \hat{r}\}$.

For the case $\tilde{r} = 0$, the upper-bound (74) also holds by (46) and (63).

Bounding $E_2(\tau)$: By (194) [Mean Value theorem] we can write

$$|\phi'_{+,k}(\tau)| \leq |\phi'_{+,k}(t_j)| + |\phi''_{+,k}(\tau_m)| |\tau - t_j|$$

with $\tau_m \in (t_j, \tau)$. Next, we use (39) and (60) to upper-bound $|\phi'_{+,k}(t_j)|$ by the right-hand side of (60); use (65) to upper-bound $|\phi''_{+,k}(\tau_m)|$. With these estimates we can further upper-bound $|\phi'_{+,k}(\tau)|$ as follows:

$$|\phi'_{+,k}(\tau)| \leq r^{2r-1} c_{u12}^r \left(\frac{\lambda_{hi}}{\lambda_{lo}^2} + \frac{1}{\lambda_{lo}^2} |\tau - t_j| \right) \leq r^{2r-1} c_{u13}^r \frac{\lambda_{hi}}{\lambda_{lo}^2}. \tag{75}$$

Above, we defined $c_{u12} \triangleq \max(c_{u7}, c_{u2}c_{u8})$, $c_{u13} \triangleq 2c_{u12}$, and used $|\tau - t_j| \leq \lambda_{hi}$. Assume $\tilde{r} \geq 1$ (the case $\tilde{r} = 0$ will be treated separately below) so that $\hat{r} \geq 1$ and $\tau \in \mathcal{N}(r\Delta\lambda_{lo}, \mathcal{T}_k^c)$, which implies that we can use (42) to upper-bound $|\phi'_{0,k}(t)|$:

$$|\phi'_{0,k}(\tau)| \leq \sum_{m=1}^{\hat{r}} c_{u3}^{\hat{r}} \frac{\prod_{\substack{l \leq \hat{r} \\ l \neq m}} (v_l^\tau - \tau)^2 |v_m^\tau - \tau|}{(r\lambda_{lo})^{2\hat{r}}} + (r-1-\hat{r})c_{u3}^{\hat{r}+1} \frac{\prod_{l=1}^{\hat{r}} (v_l^\tau - \tau)^2}{(r\lambda_{lo})^{2\hat{r}+1}}. \tag{76}$$

Multiplying (75) and (76) and simplifying we obtain the following upper bound on E_2 :

$$\begin{aligned} E_2(\tau) &= |\phi'_{0,k}(\tau)| |\phi'_{+,k}(\tau)| \stackrel{(a)}{\leq} r^{2r} c_{u14}^r \frac{\prod_{1 \leq l \leq \hat{r}} (v_l^\tau - \tau)^2}{(r\lambda_{lo})^{2\hat{r}}} \frac{1}{\lambda_{lo}^2} \\ &\stackrel{(b)}{\leq} r^{2r} c_{u14}^r \frac{\prod_{1 \leq l \leq \hat{r}} (v_l - \tau)^2}{(r\lambda_{lo})^{2\hat{r}}} \frac{1}{\lambda_{lo}^2}. \end{aligned} \tag{77}$$

Above, in (a) we used the bound $\lambda_{hi} \leq |v_l^\tau - \tau|$, which is true for all $l \in \{1, \dots, \hat{r}\}$ (follows because the elements of \mathcal{T} are separated by at least $2\lambda_{hi}$), used $\lambda_{hi}/\lambda_{lo} < 1$, and defined $c_{u14} \triangleq \max(2c_{u13}c_{u3}, 2\pi c_{u8}c_{u1})$; in (b) we used the fact that $|v_l^\tau - \tau|/(\lambda_{lo}r) \leq \Delta < 1$ for all $l \in \{1, \dots, \hat{r}\}$.

For the case $\tilde{r} = 0$, the upper bound (77) also holds by (45) and (64).

Bounding $E_3(\tau)$: By (65),

$$|\phi''_{+,k}(\tau)| \leq r^{2r-2} c_{u8}^r c_{u2} \frac{1}{\lambda_{lo}^2}. \tag{78}$$

Assume $\tilde{r} \geq 1$ (the case $\tilde{r} = 0$ will be treated separately below) so that $\hat{r} \geq 1$ and $\tau \in \mathcal{N}(r\Delta\lambda_{lo}, \mathcal{T}_k^c)$, which implies that we can use (41) to upper-bound $|\phi_{0,k}(\tau)|$:

$$|\phi_{0,k}(\tau)| \leq c_u^{\hat{r}} \frac{\prod_{l=1}^{\hat{r}} (v_l^\tau - \tau)^2}{(r\lambda_{lo})^{2\hat{r}}}. \tag{79}$$

Multiplying (78) and (79) and simplifying we obtain the following upper bound on E_3 :

$$\begin{aligned}
 E_3(\tau) = |\phi_{0,k}(\tau)| |\phi''_{+,k}(\tau)| &\stackrel{(a)}{\leq} r^{2r-2} c_{u15}^r \frac{\prod_{l=1}^{\hat{r}} (v_l^\tau - \tau)^2}{(r\lambda_{10})^{2\hat{r}}} \frac{1}{\lambda_{10}^2} \\
 &\stackrel{(b)}{\leq} r^{2r-2} c_{u15}^r \frac{\prod_{1 \leq l \leq \tilde{r}} (v_l - \tau)^2}{(r\lambda_{10})^{2\tilde{r}}} \frac{1}{\lambda_{10}^2}. \tag{80}
 \end{aligned}$$

Above, (a) we defined $c_{u15} \triangleq c_{u8}c_{u2}c_u$; in (b) we use the fact that $|v_l^\tau - \tau|/(\lambda_{10}r) \leq \Delta < 1$ for all $l \in \{1, \dots, \hat{r}\}$.

For the case $\tilde{r} = 0$, the upper-bound (80) also holds by (78) because by (36) and Lemma 1, Property 3, $|\phi_{0,k}(\tau)| < 1$ and because $c_u > 1$ and $c_{u2} > 1$.

From (71), (74), (77), and (80) we conclude that

$$|\phi''_k(\tau)| \leq r^{2r+1} c_{u16}^{r+1} \frac{\prod_{1 \leq l \leq \tilde{r}} (v_l - \tau)^2}{(r\lambda_{10})^{2\tilde{r}}} \frac{1}{\lambda_{10}^2}, \tag{81}$$

where we defined $c_{u16} \triangleq 4 \max(c_{u11}, c_{u14}, c_{u15})$.

Putting pieces together: On the one hand, by (195) [Mean Value theorem], using (68), (70), (81) and we can write for all $\tau \in \mathcal{N}(\lambda_{hi}, t_j)$:

$$\begin{aligned}
 |\phi_l(\tau) - \eta_j| &\stackrel{(a)}{\leq} \frac{1}{2} r^{2r+1} c_{u16}^{r+1} \frac{\prod_{1 \leq l \leq \tilde{r}} (v_l - \tau_m)^2}{(r\lambda_{10})^{2\tilde{r}}} \frac{(\tau - t_j)^2}{\lambda_{10}^2} \\
 &\stackrel{(b)}{\leq} \frac{1}{2} r^{2r+1} c_{u16}^{r+1} 2^{\tilde{r}} \frac{\prod_{1 \leq l \leq \tilde{r}} (v_l - \tau)^2}{(r\lambda_{10})^{2\tilde{r}}} \frac{(\tau - t_j)^2}{\lambda_{10}^2} \\
 &\stackrel{(c)}{\leq} r^{2r+3} c_{u23}^{r+1} \frac{\prod_{1 \leq l \leq \tilde{r}} (v_l - \tau)^2}{(r\lambda_{10})^{2\tilde{r}}} \frac{(\tau - t_j)^2}{(r\lambda_{10})^2}. \tag{82}
 \end{aligned}$$

Above, in (a) $\tau_m \in (t_j, \tau)$; in (b) we used that $|v_l - \tau_m| < |v_l - \tau| + \lambda_{hi} < 2|v_l - \tau|$, which is true because $\tau \in \mathcal{N}(\lambda_{hi}, t_j)$ and because the elements of \mathcal{T} are separated by at least $2\lambda_{hi}$; in (c) we defined $c_{u23} \triangleq 2c_{u16}$.

On the other hand, let $\{u_1^\tau, \dots, u_{\tilde{r}}^\tau\} \triangleq \mathcal{N}(r\Delta\lambda_{10}, \tau) \cap \mathcal{T}$. Then, by (16),

$$\begin{aligned}
 q_0(\tau) &\geq c_{l2}^r \frac{\prod_{l=1}^{\tilde{r}} (u_l^\tau - \tau)^2}{(r\lambda_{10})^{2\tilde{r}}} \\
 &\stackrel{(a)}{\geq} c_{l2}^r \frac{\prod_{1 \leq l \leq \tilde{r}} (v_l - \tau)^2}{(r\lambda_{10})^{2\tilde{r}}} \frac{(\tau - t_j)^2}{(r\lambda_{10})^2} \underbrace{\left(\frac{(r\Delta\lambda_{10} - 2\lambda_{hi})^2}{(r\lambda_{10})^2} \right)^{\tilde{r}-\tilde{r}-1}}_{P_1} \\
 &\stackrel{(b)}{\geq} c_{l3}^r \frac{\prod_{1 \leq l \leq \tilde{r}} (v_l - \tau)^2}{(r\lambda_{10})^{2\tilde{r}}} \frac{(\tau - t_j)^2}{(r\lambda_{10})^2}. \tag{83}
 \end{aligned}$$

Above, in (a) we use the fact that $\{v_1, \dots, v_{\tilde{r}}\} \cup \{t_j\} \subset \{u_1^\tau, \dots, u_{\tilde{r}}^\tau\}$ and the fact that by construction of the set $\{v_1, \dots, v_{\tilde{r}}\}$ it follows that if, for some $k, u_k^\tau \notin \{v_1, \dots, v_{\tilde{r}}\} \cup \{t_j\}$, then $|u_k^\tau - \tau| \geq r\Delta\lambda_{lo} - 2\lambda_{hi}$; in (b) we used the assumption $\text{SRF} \geq 12$ so that $\lambda_{hi} \leq \lambda_{lo}/12$ and therefore $r\Delta\lambda_{lo} - 2\lambda_{hi} \geq r(\Delta - 1/6)\lambda_{lo}$, used that $0 < \Delta - 1/6 < 1$, which implies that $P_1 \geq (\Delta - 1/6)^{2r}$, and defined $c_{l3} \triangleq c_{l2}(\Delta - 1/6)^2$ that satisfies $0 < c_{l3} < 1$.

The bound (66) follows from (82) and (83) by defining $c_{u24} \triangleq c_{u23}/c_{l3}$.

Proof of (67): case $t_j \in \mathcal{T}_k^c$. We only need to consider this case when $r > 1$. Indeed, when $r = 1$, the sum in (29) only contains one element, $\phi_l(\cdot)$, and, necessarily, $t_j \in \mathcal{I}_l$ because \mathcal{I}_l^c is empty.

In this case t_j is one of the elements among $\{v_1, \dots, v_{\tilde{r}}\} \subset \{v_1^\tau, \dots, v_{\tilde{r}}^\tau\}$; in other words, $t_j = v_{\tilde{m}}^\tau = v_{\tilde{m}}^\tau$ for some $1 \leq \tilde{m} \leq \tilde{r}, 1 \leq \hat{m} \leq \hat{r}$. The set $\mathcal{T}_k \cap \mathcal{N}(r\Delta\lambda_{lo} - \lambda_{hi}, t_j)$ is either empty or contains exactly one element. Let $b \triangleq |\mathcal{T}_k \cap \mathcal{N}(r\Delta\lambda_{lo} - \lambda_{hi}, t_j)|$. In the case when $b = 1$, let $\{\tilde{t}\} \triangleq \mathcal{T}_k \cap \mathcal{N}(r\Delta\lambda_{lo} - \lambda_{hi}, t_j)$.

Bounding $E_1(\tau)$: Consider the case $b = 1$. By (195) [Mean Value theorem] we can write

$$|\phi_{+,k}(\tau)| \leq |\phi_{+,k}(\tilde{t})| + |\phi'_{+,k}(\tilde{t})||\tau - \tilde{t}| + \frac{1}{2}|\phi''_{+,k}(\tau_m)|(\tau - \tilde{t})^2$$

with $\tau_m \in (\tilde{t}, \tau)$. Next, we use (38) and (52) to upper-bound $|\phi_{+,k}(\tilde{t})|$ by the right-hand side of (52); use (39) and (60) to upper-bound $|\phi'_{+,k}(\tilde{t})|$ by the right-hand side of (60); use (65) to upper-bound $|\phi''_{+,k}(\tau_m)|$. With these estimates we can further upper-bound $|\phi_{+,k}(\tau)|$ as follows:

$$\begin{aligned} |\phi_{+,k}(\tau)| &\leq r^{2r-1}c_{u9}^r \left(\frac{\lambda_{hi}^2}{\lambda_{lo}^2} + \frac{\lambda_{hi}}{\lambda_{lo}^2}|\tau - \tilde{t}| + \frac{1}{\lambda_{lo}^2}(\tau - \tilde{t})^2 \right) \\ &\leq r^{2r+1}c_{u10}^r \frac{(\tilde{t} - \tau)^2}{(r\lambda_{lo})^2} = r^{2r+1}c_{u10}^r \left[\frac{(\tilde{t} - \tau)^2}{(r\lambda_{lo})^2} \right]^{l(b=1)}, \end{aligned} \tag{84}$$

where we used that $\lambda_{hi} \leq |\tilde{t} - \tau|$ because the elements of \mathcal{T} are separated by at least $2\lambda_{hi}$ and $\tau \in \mathcal{N}(\lambda_{hi}, t_j)$ with $\tilde{t} \neq t_j$. According to (63) the upper bound (84) also holds for $b = 0$.

Since $t_j \in \mathcal{T}_k^c$ and $\tau \in \mathcal{N}(\lambda_{hi}, t_j)$, it follows $\tau \in \mathcal{N}(r\Delta\lambda_{lo}, \mathcal{T}_k^c)$ so that $\hat{r} \geq 1$, which implies that we can use (43) to upper-bound $|\phi''_{0,k}(\tau)|$:

$$\begin{aligned} |\phi''_{0,k}(\tau)| &\leq \sum_{1 \leq m \leq \hat{r}} \sum_{\substack{1 \leq m' \leq \hat{r} \\ m \neq m'}} I[\hat{r} \geq 2]c_{u3}^{\hat{r}} \frac{\prod_{\substack{1 \leq l \leq \hat{r} \\ l \neq m, l \neq m'}} (v_l^\tau - \tau)^2}{(r\lambda_{lo})^{2(\hat{r}-2)}} \frac{|v_m^\tau - \tau|}{(r\lambda_{lo})^2} \frac{|v_{m'}^\tau - \tau|}{(r\lambda_{lo})^2} \\ &\quad + 2(r - 1 - \hat{r}) \sum_{1 \leq m \leq \hat{r}} c_{u3}^{\hat{r}+1} \frac{\prod_{\substack{1 \leq l \leq \hat{r} \\ l \neq m}} (v_l^\tau - \tau)^2}{(r\lambda_{lo})^{2(\hat{r}-1)}} \frac{|v_m^\tau - \tau|}{(r\lambda_{lo})^2} \frac{1}{r\lambda_{lo}} \end{aligned}$$

$$\begin{aligned}
 &+ (r - 1 - \hat{r})^2 c_{u3}^{\hat{r}+2} \frac{\prod_{l=1}^{\hat{r}} (v_l^\tau - \tau)^2}{(r\lambda_{10})^{2\hat{r}}} \frac{1}{(r\lambda_{10})^2} \\
 &+ \sum_{1 \leq m \leq \hat{r}} c_{u3}^{\hat{r}} \frac{\prod_{\substack{1 \leq l \leq \hat{r} \\ l \neq m}} (v_l^\tau - \tau)^2}{(r\lambda_{10})^{2(\hat{r}-1)}} \frac{1}{(r\lambda_{10})^2} \\
 &+ (r - 1 - \hat{r}) c_{u3}^{\hat{r}+1} \frac{\prod_{l=1}^{\hat{r}} (v_l^\tau - \tau)^2}{(r\lambda_{10})^{2\hat{r}}} \frac{1}{(r\lambda_{10})^2} \\
 &\stackrel{(a)}{\leq} r^2 c_{u17}^{r+1} \frac{\prod_{1 \leq l \leq \hat{r}, l \neq \hat{m}} (v_l^\tau - \tau)^2}{(r\lambda_{10})^{2\hat{r}}}. \tag{85}
 \end{aligned}$$

Above, in (a) we used (multiple times) the fact that $|v_m^\tau - \tau| \leq |v_l^\tau - \tau|$ for all $l \in \{1, \dots, \hat{r}\}$, the fact that $|v_l^\tau - \tau|/(\lambda_{10}r) \leq \Delta < 1$ for all $l \in \{1, \dots, \hat{r}\}$, the fact $\hat{r} \leq r - 1$, and defined $c_{u17} \triangleq 6c_{u3}$.

Multiplying (84) and (85) and simplifying we obtain the following upper bound on E_1 :

$$\begin{aligned}
 E_1(\tau) &= |\phi''_{0,k}(\tau)| |\phi_{+,k}(\tau)| \\
 &\stackrel{(a)}{\leq} r^{2r+1} c_{u18}^{r+1} \left[\frac{(\tilde{t} - \tau)^2}{(r\lambda_{10})^2} \right]^{I[b=1]} \frac{\prod_{1 \leq l \leq \hat{r}, l \neq \hat{m}} (v_l^\tau - \tau)^2}{(r\lambda_{10})^{2(\hat{r}-1)}} \frac{1}{\lambda_{10}^2} \\
 &\stackrel{(b)}{\leq} r^{2r+1} c_{u18}^{r+1} \left[\frac{(\tilde{t} - \tau)^2}{(r\lambda_{10})^2} \right]^{I[b=1]} \frac{\prod_{1 \leq l \leq \hat{r}, l \neq \hat{m}} (v_l - \tau)^2}{(r\lambda_{10})^{2(\hat{r}-1)}} \frac{1}{\lambda_{10}^2}. \tag{86}
 \end{aligned}$$

Above, in (a) we defined $c_{u18} \triangleq c_{u10}c_{u17}$; in (b) we used that $|v_l^\tau - \tau|/(\lambda_{10}r) \leq \Delta < 1$ for all $l \in \{1, \dots, \hat{r}\}$.

Bounding $E_2(\tau)$: Consider the case $b = 1$. By (194) [Mean Value theorem] we can write

$$|\phi'_{+,k}(\tau)| \leq |\phi'_{+,k}(\tilde{t})| + |\phi''_{+,k}(\tau_m)| |\tau - \tilde{t}|$$

with $\tau_m \in (\tilde{t}, \tau)$. Next, we use (39) and (60) to upper-bound $|\phi'_{+,k}(\tilde{t})|$ by the right-hand side of (60); use (65) to upper-bound $|\phi''_{+,k}(\tau_m)|$. With these estimates we can further upper-bound $|\phi'_{+,k}(\tau)|$ as follows:

$$\begin{aligned}
 |\phi'_{+,k}(\tau)| &\leq r^{2r-1} c_{u12}^r \left(\frac{\lambda_{hi}}{\lambda_{10}^2} + \frac{1}{\lambda_{10}^2} |\tau - \tilde{t}| \right) \leq r^{2r} c_{u13}^r \frac{|\tau - \tilde{t}|}{r\lambda_{10}} \frac{1}{\lambda_{10}} \\
 &= r^{2r} c_{u13}^r \left[\frac{|\tau - \tilde{t}|}{r\lambda_{10}} \right]^{I[b=1]} \frac{1}{\lambda_{10}}, \tag{87}
 \end{aligned}$$

where we used that $\lambda_{hi} \leq |\tilde{t} - \tau|$. According to (64) the upper bound (87) also holds for $b = 0$.

Since $t_j \in \mathcal{T}_k^c$ and $\tau \in \mathcal{N}(\lambda_{hi}, t_j)$, it follows $\tau \in \mathcal{N}(r\Delta\lambda_{10}, \mathcal{T}_k^c)$ so that $\hat{r} \geq 1$, which implies that we can use (42) to upper-bound $|\phi'_{0,k}(\tau)|$:

$$\begin{aligned}
 |\phi'_{0,k}(\tau)| &\leq \sum_{m=1}^{\hat{r}} c_{u3}^{\hat{r}} \frac{\prod_{\substack{1 \leq l \leq \hat{r} \\ l \neq m}} (v_l^\tau - \tau)^2 |v_m^\tau - \tau|}{(r\lambda_{10})^{2\hat{r}}} + (r-1-\hat{r})c_{u3}^{\hat{r}+1} \frac{\prod_{l=1}^{\hat{r}} (v_l^\tau - \tau)^2}{(r\lambda_{10})^{2\hat{r}+1}} \\
 &\stackrel{(a)}{\leq} c_{u19}^r \frac{\prod_{1 \leq l \leq \hat{r}, l \neq \hat{m}} (v_l^\tau - \tau)^2}{(r\lambda_{10})^{2(\hat{r}-1)}} \left[\frac{|\tilde{t} - \tau|}{r\lambda_{10}} \right]^{I[b=1]} \frac{1}{\lambda_{10}}. \tag{88}
 \end{aligned}$$

Above, in (a) we used the fact that $|v_m^\tau - \tau| \leq |v_l^\tau - \tau|$ for all $l \in \{1, \dots, \hat{r}\}$, the fact that $|v_l^\tau - \tau|/(\lambda_{10}r) \leq \Delta < 1$ for all $l \in \{1, \dots, \hat{r}\}$, and the fact that $|v_m^\tau - \tau| \leq |\tilde{t} - \tau|$, and defined $c_{u19} \triangleq 2c_{u3}$.

Multiplying (87) and (88) and simplifying we obtain the following upper bound on E_2 :

$$\begin{aligned}
 E_2(\tau) &= |\phi'_{0,k}(\tau)| |\phi'_{+,k}(\tau)| \\
 &\stackrel{(a)}{\leq} r^{2r} c_{u20}^r \left[\frac{(\tilde{t} - \tau)^2}{(r\lambda_{10})^2} \right]^{I[b=1]} \frac{\prod_{1 \leq l \leq \hat{r}, l \neq \hat{m}} (v_l^\tau - \tau)^2}{(r\lambda_{10})^{2(\hat{r}-1)}} \frac{1}{\lambda_{10}^2} \\
 &\stackrel{(b)}{\leq} r^{2r} c_{u20}^r \left[\frac{(\tilde{t} - \tau)^2}{(r\lambda_{10})^2} \right]^{I[b=1]} \frac{\prod_{1 \leq l \leq \hat{r}, l \neq \hat{m}} (v_l - \tau)^2}{(r\lambda_{10})^{2(\hat{r}-1)}} \frac{1}{\lambda_{10}^2}. \tag{89}
 \end{aligned}$$

Above, in (a) we defined $c_{u20} \triangleq c_{u13}c_{u19}$; in (b) we used the fact that $|v_l^\tau - \tau|/(\lambda_{10}r) \leq \Delta < 1$ for all $l \in \{1, \dots, \hat{r}\}$.

Bounding $E_3(\tau)$: By (65),

$$|\phi''_{+,k}(\tau)| \leq r^{2r-2} c_{u8}^r c_{u2} \frac{1}{\lambda_{10}^2}. \tag{90}$$

Since $t_j \in \mathcal{T}_k^c$ and $\tau \in \mathcal{N}(\lambda_{hi}, t_j)$, it follows $\tau \in \mathcal{N}(r\Delta\lambda_{10}, \mathcal{T}_k^c)$ so that $\hat{r} \geq 1$, which implies that we can use (41) to upper-bound $|\phi_{0,k}(\tau)|$:

$$|\phi_{0,k}(\tau)| \leq c_u^{\hat{r}} \frac{\prod_{l=1}^{\hat{r}} (v_l^\tau - \tau)^2}{(r\lambda_{10})^{2\hat{r}}} \stackrel{(a)}{\leq} c_u^{\hat{r}} \left[\frac{(\tilde{t} - \tau)^2}{(r\lambda_{10})^2} \right]^{I[b=1]} \frac{\prod_{1 \leq l \leq \hat{r}, l \neq \hat{m}} (v_l^\tau - \tau)^2}{(r\lambda_{10})^{2(\hat{r}-1)}}. \tag{91}$$

Above, in (a) we used the fact that $|v_m^\tau - \tau| \leq |v_l^\tau - \tau|$ for all $l \in \{1, \dots, \hat{r}\}$, the fact that $|v_l^\tau - \tau|/(\lambda_{10}r) \leq \Delta < 1$ for all $l \in \{1, \dots, \hat{r}\}$, and the fact that $|v_m^\tau - \tau| \leq |\tilde{t} - \tau|$. Multiplying (90) and (91) and simplifying we obtain the following upper bound on E_3 :

$$E_3(\tau) = |\phi_{0,k}(\tau)| |\phi''_{+,k}(\tau)|$$

$$\begin{aligned}
 &\stackrel{(a)}{\leq} r^{2r-2} c_{u21}^r \left[\frac{(\tilde{t} - \tau)^2}{(r\lambda_{10})^2} \right]^{I[b=1]} \frac{\prod_{1 \leq l \leq \tilde{r}, l \neq \tilde{m}} (v_l^\tau - \tau)^2}{(r\lambda_{10})^{2(\tilde{r}-1)}} \frac{1}{\lambda_{10}^2} \\
 &\stackrel{(b)}{\leq} r^{2r-2} c_{u21}^r \left[\frac{(\tilde{t} - \tau)^2}{(r\lambda_{10})^2} \right]^{I[b=1]} \frac{\prod_{1 \leq l \leq \tilde{r}, l \neq \tilde{m}} (v_l - \tau)^2}{(r\lambda_{10})^{2(\tilde{r}-1)}} \frac{1}{\lambda_{10}^2}. \tag{92}
 \end{aligned}$$

Above, in (a) we defined $c_{u21} \triangleq c_{u8}c_{u2}c_u$; in (b) we used that $|v_l^\tau - \tau|/(\lambda_{10}r) \leq \Delta < 1$ for all $l \in \{1, \dots, \tilde{r}\}$.

From (71), (86), (89), and (92) we conclude that

$$|\phi_k''(\tau)| \leq r^{2r+1} c_{u22}^{r+1} \left[\frac{(\tilde{t} - \tau)^2}{(r\lambda_{10})^2} \right]^{I[b=1]} \frac{\prod_{1 \leq l \leq \tilde{r}, l \neq \tilde{m}} (v_l - \tau)^2}{(r\lambda_{10})^{2(\tilde{r}-1)}} \frac{1}{\lambda_{10}^2}, \tag{93}$$

where we defined $c_{u22} \triangleq 4 \max(c_{u18}, c_{u20}, c_{u21})$.

Putting pieces together: On the one hand, by (195) [Mean Value theorem], using (69), (70), (93), and we can write for all $\tau \in \mathcal{N}(\lambda_{hi}, t_j)$:

$$\begin{aligned}
 |\phi_k(\tau)| &\stackrel{(a)}{\leq} \frac{1}{2} r^{2r+1} c_{u22}^{r+1} \left[\frac{(\tilde{t} - \tau_m)^2}{(r\lambda_{10})^2} \right]^{I[b=1]} \frac{\prod_{1 \leq l \leq \tilde{r}, l \neq \tilde{m}} (v_l - \tau_m)^2}{(r\lambda_{10})^{2(\tilde{r}-1)}} \frac{(\tau - t_j)^2}{\lambda_{10}^2} \\
 &\stackrel{(b)}{\leq} \frac{1}{2} r^{2r+1} c_{u22}^{r+1} 2^{\tilde{r}} \left[\frac{(\tilde{t} - \tau)^2}{(r\lambda_{10})^2} \right]^{I[b=1]} \frac{\prod_{1 \leq l \leq \tilde{r}, l \neq \tilde{m}} (v_l - \tau)^2}{(r\lambda_{10})^{2(\tilde{r}-1)}} \frac{(\tau - t_j)^2}{\lambda_{10}^2} \\
 &\stackrel{(c)}{\leq} r^{2r+3} c_{u25}^{r+1} \left[\frac{(\tilde{t} - \tau)^2}{(r\lambda_{10})^2} \right]^{I[b=1]} \frac{\prod_{1 \leq l \leq \tilde{r}} (v_l - \tau)^2}{(r\lambda_{10})^{2\tilde{r}}}. \tag{94}
 \end{aligned}$$

Above, in (a) $\tau_m \in (t_j, \tau)$; in (b) we used the fact that, for $l \neq \tilde{m}$, $|v_l - \tau_m| < |v_l - \tau| + \lambda_{hi} < 2|v_l - \tau|$ and $|\tilde{t} - \tau_m| < |\tilde{t} - \tau| + \lambda_{hi} < 2|\tilde{t} - \tau|$, which is true because $\tau \in \mathcal{N}(\lambda_{hi}, t_j)$ and because the elements of \mathcal{T} are separated by at least $2\lambda_{hi}$; in (c) we defined $c_{u25} \triangleq 2c_{u22}$ and used the fact that $t_j = v_{\tilde{m}}$.

On the other hand, let $\{u_1^\tau, \dots, u_{\tilde{r}}^\tau\} \triangleq \mathcal{N}(r\Delta\lambda_{10}, \tau) \cap \mathcal{T}$. Then by (16),

$$\begin{aligned}
 q_0(\tau) &\geq c_{l2}^r \frac{\prod_{l=1}^{\tilde{r}} (u_l^\tau - \tau)^2}{(r\lambda_{10})^{2\tilde{r}}} \\
 &\stackrel{(a)}{\geq} c_{l2}^r \left[\frac{(\tilde{t} - \tau)^2}{(r\lambda_{10})^2} \right]^{I[b=1]} \frac{\prod_{1 \leq l \leq \tilde{r}} (v_l - \tau)^2}{(r\lambda_{10})^{2\tilde{r}}} \underbrace{\left(\frac{(r\Delta\lambda_{10} - 2\lambda_{hi})^2}{(r\lambda_{10})^2} \right)^{\tilde{r} - \tilde{r} - I[b=1]}}_{P_2} \\
 &\stackrel{(b)}{\geq} c_{l3}^r \left[\frac{(\tilde{t} - \tau)^2}{(r\lambda_{10})^2} \right]^{I[b=1]} \frac{\prod_{1 \leq l \leq \tilde{r}} (v_l - \tau)^2}{(r\lambda_{10})^{2\tilde{r}}}. \tag{95}
 \end{aligned}$$

Above, in (a) we used the fact that $\{v_1, \dots, v_{\tilde{r}}\} \subset \{u_1^\tau, \dots, u_{\tilde{r}}^\tau\}$, the fact that if $b = 1$, then $\tilde{t} \in \{u_1^\tau, \dots, u_{\tilde{r}}^\tau\}$, and the fact that by construction of the set $\{v_1, \dots, v_{\tilde{r}}\}$ it follows that if, for some k , $u_k^\tau \notin \{v_1, \dots, v_{\tilde{r}}\}$ and $u_k^\tau \neq \tilde{t}$, then $|u_k^\tau - \tau| \geq r\Delta\lambda_{10} - 2\lambda_{hi}$; in (b)

we used the assumption $SRF \geq 12$ so that $\lambda_{hi} \leq \lambda_{l0}/12$ and therefore $r\Delta\lambda_{l0} - 2\lambda_{hi} \geq r(\Delta - 1/6)\lambda_{l0}$, used that $0 < \Delta - 1/6 < 1$, which implies that $P_2 \geq (\Delta - 1/6)^{2r}$.

The bound (67) follows from (94) and (95) by defining $c_{u26} \triangleq c_{u25}/c_{l3}$.

6.3.5 Proof of Property 3

By (29) and the triangle inequality:

$$\begin{aligned} \|q_1(\cdot)\|_\infty &\leq \rho/2 + r \max_{1 \leq k \leq r} \|\phi_k(\cdot)\|_\infty \\ &\stackrel{(a)}{\leq} \rho/2 + r \max_{1 \leq k \leq r} \|\phi_{+,k}(\cdot)\|_\infty \\ &\stackrel{(b)}{=} \rho/2 + r^{2r+1} c_{u8}^r \max_{1 \leq k \leq r} \|q_{r\lambda_{l0}, \mathcal{T}_k, \{f_j\}, \{d_j\}}(\cdot)\|_\infty \\ &\stackrel{(c)}{\leq} \rho/2 + r^{2r+1} c_{u0} c_{u8}^r \stackrel{(d)}{\leq} r^{2r+1} c_{u55}^r. \end{aligned}$$

Above, in (a) we used (35) and the fact that by (36) and Lemma 1, Property 3, $\|\phi_{0,k}(\cdot)\|_\infty \leq 1$; in (b) we used (37); in (c) we used Lemma 2, Property 3; in (d) we defined $c_{u55} \triangleq 2c_{u0}c_{u8}$ and used the fact that $\rho/2 < 1 < c_{u0}c_{u8}$.

6.3.6 Proof of Property 4

Take $\tau \in \mathcal{F}(\lambda_{hi}, \mathcal{T})$. As above, let $\{u_1^\tau, \dots, u_{\check{r}}^\tau\} \triangleq \mathcal{N}(r\Delta\lambda_{l0}, \tau) \cap \mathcal{T}$. Then by (16),

$$q_0(\tau) \geq c_{l2}^{\check{r}} \frac{\prod_{l=1}^{\check{r}} (u_l^\tau - \tau)^2}{(r\lambda_{l0})^{2\check{r}}}. \tag{96}$$

By (19) this bound is also valid when $\check{r} = 0$.

Fix k . If $\tau \in \mathcal{N}(r\Delta\lambda_{l0}, \mathcal{T}_k^c)$, then we can use (41) to upper-bound $|\phi_{0,k}(\tau)|$:

$$|\phi_{0,k}(\tau)| \leq c_u^{\hat{r}} \frac{\prod_{l=1}^{\hat{r}} (v_l^\tau - \tau)^2}{(r\lambda_{l0})^{2\hat{r}}}, \tag{97}$$

where, as before, $\{v_1^\tau, \dots, v_{\hat{r}}^\tau\} \triangleq \mathcal{N}(r\Delta\lambda_{l0}, \tau) \cap \mathcal{T}_k^c$. If $\tau \notin \mathcal{N}(r\Delta\lambda_{l0}, \mathcal{T}_k^c)$, we will use that by (36) and by Lemma 1, Property 3,

$$|\phi_{0,k}(\tau)| \leq 1. \tag{98}$$

The set $\mathcal{T}_k \cap \mathcal{N}(r\Delta\lambda_{l0}, \tau)$ is either empty or contains exactly one element. Let $b \triangleq |\mathcal{T}_k \cap \mathcal{N}(r\Delta\lambda_{l0}, \tau)|$ denote the size of this set; when $b = 1$, let $\{\tilde{t}\} \triangleq \mathcal{T}_k \cap \mathcal{N}(r\Delta\lambda_{l0}, \tau)$. Following the steps that lead to (84), we obtain

$$|\phi_{+,k}(\tau)| \leq r^{2r+1} c_{u10}^r \left[\frac{(\tilde{t} - \tau)^2}{(r\lambda_{l0})^2} \right]^{l[b=1]} \tag{99}$$

and the bound is valid for both cases $b = 0$ and $b = 1$.

Case $\hat{r} \geq 1$: Then, $\{u_1^\tau, \dots, u_{\check{r}}^\tau\} = \{v_1^\tau, \dots, v_{\check{r}}^\tau\} \cup \{\tilde{t}\}$ if $b = 1$, and $\{u_1^\tau, \dots, u_{\check{r}}^\tau\} = \{v_1^\tau, \dots, v_{\check{r}}^\tau\}$ if $b = 0$. Therefore,

$$\begin{aligned}
 |\phi_k(\tau)| &= |\phi_{0,k}(\tau)| |\phi_{+,k}(\tau)| \\
 &\stackrel{(a)}{\leq} r^{2r+1} c_{u10}^r c_u^{\hat{r}} \left[\frac{(\tilde{t} - \tau)^2}{(r\lambda_{l0})^2} \right]^{I[b=1]} \frac{\prod_{l=1}^{\hat{r}} (v_l^\tau - \tau)^2}{(r\lambda_{l0})^{2\hat{r}}} \\
 &= r^{2r+1} c_{u10}^r c_u^{\hat{r}} \frac{\prod_{l=1}^{\check{r}} (u_l^\tau - \tau)^2}{(r\lambda_{l0})^{2\check{r}}} \stackrel{(b)}{\leq} r^{2r+1} c_{u28}^r q_0(\tau). \tag{100}
 \end{aligned}$$

Above, (a) follows by (97) and (99); (b) follows by (96) with $c_{u28} \triangleq c_{u10} c_u / c_{l2}$.

Case $\hat{r} = 0$: Then, $\check{r} = 1$ and $\{u_{\check{r}}^\tau\} = \{\tilde{t}\}$ if $b = 1$ and $\check{r} = 0$ if $b = 0$. Therefore,

$$\begin{aligned}
 |\phi_k(\tau)| &= |\phi_{0,k}(\tau)| |\phi_{+,k}(\tau)| \\
 &\stackrel{(a)}{\leq} c_{u10}^r r^{2r+1} \left[\frac{(\tilde{t} - \tau)^2}{(r\lambda_{l0})^2} \right]^{I[b=1]} \\
 &= r^{2r+1} c_{u10}^r \frac{\prod_{l=1}^{\check{r}} (u_l^\tau - \tau)^2}{(r\lambda_{l0})^{2\check{r}}} \stackrel{(b)}{\leq} r^{2r+1} c_{u28}^r q_0(\tau). \tag{101}
 \end{aligned}$$

Above, (a) follows by (98) and (99); (b) follows by (96) because $c_u > 1$.

By Lemma 3, Property 6,

$$\frac{\rho}{2} = \frac{\lambda_{hi}^{2r}}{2\lambda_{lo}^{2r}} \leq \frac{r^{2r}}{c_l^r} q_0(\tau). \tag{102}$$

Therefore, by (29), (100), (101), (102),

$$|q_1(\tau)| \leq \sum_{k=1}^r |\phi_k(\tau)| + \rho/2 \leq r^{2r+2} c_{u29}^r q_0(\tau),$$

where we defined $c_{u29} \triangleq c_{u28} + 1/c_l$. □

6.4 Dual Certificate $q_2(\cdot)$

Finally, we construct the trigonometric polynomial $q_2(\cdot)$. This trigonometric polynomial is conceptually similar to $q_1(\cdot)$. The difference is that in $q_1(\cdot)$ we control the function values on \mathcal{T} and the derivatives on \mathcal{T} are zero; in $q_2(\cdot)$ we control the derivatives on \mathcal{T} and the function values on \mathcal{T} are zero.

Specifically, on the point $t_j \in \mathcal{T}$, $q_2(\cdot)$ is approximated by a linear function whose derivative is controlled by the sign

$$s'_j \triangleq \text{sign} \left(\sum_{m/N \in \mathcal{N}(\lambda_{hi}, t_j)} \left(\frac{m}{N} - t_j \right) h_m \right), \quad j = 1, \dots, S, \tag{103}$$

as explained in Lemma 5 below.

Lemma 5 *Set $\gamma \triangleq \rho/\lambda_{hi} = \lambda_{hi}^{2r-1}/\lambda_{lo}^{2r}$. Then, there exists a real-valued trigonometric polynomial $q_2(\cdot)$ that satisfies the following properties.*

1. *Frequency limitation to f_{lo} : $q_2(t) = \sum_{k=-f_{lo}}^{f_{lo}} \hat{q}_{2,k} e^{-i2\pi kt}$ for some $\hat{q}_{2,k} \in \mathbb{C}$.*
2. *Constrained derivative on \mathcal{T} and controlled behavior near \mathcal{T} : for all $j = 1, \dots, S$ and all $\tau \in \mathcal{N}(\lambda_{hi}, t_j)$,*

$$\left| q_2(\tau) - \gamma s'_j(\tau - t_j) \right| \leq r^{2r+4} c_{u34}^{r+1} q_0(\tau), \tag{104}$$

where s'_j are defined³ in (103). Since $q_0(\tau) = 0$ for $\tau \in \mathcal{T}$, (104) implies, in particular, that the derivative of $q_2(\cdot)$ interpolates the sign pattern in (103) scaled by γ on \mathcal{T} .

3. *Uniform confinement: $\|q_2(\cdot)\|_\infty \leq r^{2r+1} c_{u56}^r$.*
4. *Boundedness far from \mathcal{T} : for all $\tau \in \mathcal{F}(\lambda_{hi}, \mathcal{T})$,*

$$|q_2(\tau)| \leq r^{2r+2} c_{u52}^r q_0(\tau). \tag{105}$$

The positive numerical constants c_{u34} , c_{u56} , and c_{u52} are defined in the proof below.

The proof of the lemma parallels that of Lemma 4 but contains some important differences; it is given in Appendix 2 “Proof of Lemma 5”.

7 Stability Estimates

In this section we use the dual trigonometric polynomials $q_0(\cdot)$, $q_1(\cdot)$, and $q_2(\cdot)$ to prove Theorem 1.

We will use the fact that the high-resolution kernel $k_{hi}(\cdot)$ satisfies the following estimates:

$$\sum_{n=0}^{N-1} \left| k'_{hi} \left(\frac{n}{N} \right) \right| \leq \frac{c'_k}{\lambda_{hi}}, \tag{106}$$

$$\frac{1}{2} \sum_{n=0}^{N-1} \sup_{u \in \mathcal{N}(\lambda_{hi}, n/N)} |k''_{hi}(u)| \leq \frac{c''_k}{\lambda_{hi}^2}, \tag{107}$$

³ The lemma is valid for arbitrary sign pattern, we formulate it for the sign pattern defined in (103) for concreteness.

where c'_k and c''_k are positive numerical constants. The bounds are proven in Appendix 4 “Properties of Fejér Kernel”.

We will use the following shorthand notations:

$$\begin{aligned} \mathcal{N}_{hi} &\triangleq \cup_{t \in \mathcal{T}} \mathcal{N}(\lambda_{hi}, t), \\ \mathcal{F}_{hi} &\triangleq \mathbb{T} \setminus \mathcal{N}_{hi}. \end{aligned}$$

7.1 Basic Estimates

We begin by decomposing the error $\|k_{hi} \star (\hat{\mathbf{x}} - \mathbf{x})\|_1$ into a sum of simpler terms; each of the terms will then be upper-bounded separately:

$$\begin{aligned} \|k_{hi} \star (\hat{\mathbf{x}} - \mathbf{x})\|_1 &= \sum_{n=0}^{N-1} \left| \sum_{m=0}^{N-1} k_{hi} \left(\frac{n-m}{N} \right) h_m \right| \\ &\leq \sum_{n=0}^{N-1} \left| \sum_{m/N \in \mathcal{F}_{hi}} k_{hi} \left(\frac{n-m}{N} \right) h_m \right| \\ &\quad + \sum_{n=0}^{N-1} \left| \sum_{m/N \in \mathcal{N}_{hi}} k_{hi} \left(\frac{n-m}{N} \right) h_m \right|. \end{aligned} \tag{108}$$

The first term in (108) can be written as follows:

$$\begin{aligned} \sum_{n=0}^{N-1} \left| \sum_{m/N \in \mathcal{F}_{hi}} k_{hi} \left(\frac{n-m}{N} \right) h_m \right| &\stackrel{(a)}{=} \sum_{m/N \in \mathcal{F}_{hi}} \left(\sum_{n=0}^{N-1} k_{hi} \left(\frac{n-m}{N} \right) \right) h_m \\ &\stackrel{(b)}{=} \sum_{m/N \in \mathcal{F}_{hi}} \left(\sum_{n=0}^{N-1} k_{hi} \left(\frac{n}{N} \right) \right) h_m \\ &\stackrel{(c)}{=} \underbrace{\sum_{m/N \in \mathcal{F}_{hi}} h_m}_{A_0}. \end{aligned} \tag{109}$$

Above, (a) follows because $h_m \geq 0$ for $m/N \in \mathcal{F}_{hi}$ and $k_{hi}(\cdot) \geq 0$; (b) follows by periodicity of $k_{hi}(\cdot)$; (c) follows by (8).

The second term in (108) can be upper-bounded as follows:

$$\sum_{n=0}^{N-1} \left| \sum_{m/N \in \mathcal{N}_{hi}} k_{hi} \left(\frac{n-m}{N} \right) h_m \right| \stackrel{(a)}{=} \sum_{n=0}^{N-1} \left| \sum_{j=1}^S \sum_{m/N \in \mathcal{N}(\lambda_{hi}, t_j)} k_{hi} \left(\frac{n-m}{N} \right) h_m \right|$$

$$\stackrel{(b)}{\leq} \underbrace{\sum_{n=0}^{N-1} \sum_{j=1}^S \left| \sum_{m/N \in \mathcal{N}(\lambda_{hi}, t_j)} k_{hi} \left(\frac{n-m}{N} \right) h_m \right|}_B. \tag{110}$$

Above, (a) follows because the sets $\mathcal{N}(\lambda_{hi}, t_j)$ do not intersect; (b) follows by the triangle inequality. To upper-bound B in (110) we will use that for all $\tau \in \mathcal{N}(\lambda_{hi}, t_j) \cap \mathbb{T}$ and all $t \in \mathbb{T}$,

$$|k_{hi}(t - \tau) - k_{hi}(t - t_j) - k'_{hi}(t - t_j)(t_j - \tau)| \leq \sup_{u \in \mathcal{N}(\lambda_{hi}, t - t_j)} \frac{1}{2} |k''_{hi}(u)| (t_j - \tau)^2. \tag{111}$$

The inequality follows by expanding $k_{hi}(t - \tau)$ in Taylor series in τ around $\tau = t_j$ up to first order and writing the remainder in Lagrange form. We have:

$$\begin{aligned} & \left| \sum_{m/N \in \mathcal{N}(\lambda_{hi}, t_j)} k_{hi} \left(\frac{n-m}{N} \right) h_m \right| \\ & \stackrel{(a)}{\leq} \left| \sum_{m/N \in \mathcal{N}(\lambda_{hi}, t_j)} k_{hi} \left(\frac{n}{N} - t_j \right) h_m \right| \\ & \quad + \left| \sum_{m/N \in \mathcal{N}(\lambda_{hi}, t_j)} k'_{hi} \left(\frac{n}{N} - t_j \right) \left(t_j - \frac{m}{N} \right) h_m \right| \\ & \quad + \sum_{m/N \in \mathcal{N}(\lambda_{hi}, t_j)} \left| k_{hi} \left(\frac{n-m}{N} \right) - k_{hi} \left(\frac{n}{N} - t_j \right) - k'_{hi} \left(\frac{n}{N} - t_j \right) \left(t_j - \frac{m}{N} \right) \right| |h_m| \\ & \stackrel{(b)}{\leq} k_{hi} \left(\frac{n}{N} - t_j \right) \left| \sum_{m/N \in \mathcal{N}(\lambda_{hi}, t_j)} h_m \right| \\ & \quad + \left| k'_{hi} \left(\frac{n}{N} - t_j \right) \right| \left| \sum_{m/N \in \mathcal{N}(\lambda_{hi}, t_j)} \left(t_j - \frac{m}{N} \right) h_m \right| \\ & \quad + \sum_{m/N \in \mathcal{N}(\lambda_{hi}, t_j)} \sup_{u \in \mathcal{N}(\lambda_{hi}, n/N - t_j)} \frac{1}{2} |k''_{hi}(u)| \left(t_j - \frac{m}{N} \right)^2 |h_m|. \end{aligned} \tag{112}$$

Above, (a) follows by adding and subtracting the corresponding terms and applying the triangle inequality; (b) follows by (111) with $t = n/N$ and $\tau = m/N$ and because $k_{hi}(\cdot) \geq 0$.

Using (112) we can upper-bound B in (110) as follows:

$$\begin{aligned}
 B &\leq \sum_{j=1}^S \left(\sum_{n=0}^{N-1} k_{\text{hi}} \left(\frac{n}{N} - t_j \right) \right) \left| \sum_{m/N \in \mathcal{N}(\lambda_{\text{hi}}, t_j)} h_m \right| \\
 &\quad + \sum_{j=1}^S \left(\sum_{n=0}^{N-1} \left| k'_{\text{hi}} \left(\frac{n}{N} - t_j \right) \right| \right) \left| \sum_{m/N \in \mathcal{N}(\lambda_{\text{hi}}, t_j)} \left(t_j - \frac{m}{N} \right) h_m \right| \\
 &\quad + \sum_{j=1}^S \sum_{n=0}^{N-1} \sup_{u \in \mathcal{N}(\lambda_{\text{hi}}, n/N - t_j)} \frac{1}{2} |k''_{\text{hi}}(u)| \sum_{m/N \in \mathcal{N}(\lambda_{\text{hi}}, t_j)} \left(t_j - \frac{m}{N} \right)^2 |h_m| \\
 &\stackrel{(a)}{=} \left(\sum_{n=0}^{N-1} k_{\text{hi}} \left(\frac{n}{N} \right) \right) \sum_{j=1}^S \left| \sum_{m/N \in \mathcal{N}(\lambda_{\text{hi}}, t_j)} h_m \right| \\
 &\quad + \left(\sum_{n=0}^{N-1} \left| k'_{\text{hi}} \left(\frac{n}{N} \right) \right| \right) \sum_{j=1}^S \left| \sum_{m/N \in \mathcal{N}(\lambda_{\text{hi}}, t_j)} \left(t_j - \frac{m}{N} \right) h_m \right| \\
 &\quad + \left(\sum_{n=0}^{N-1} \sup_{u \in \mathcal{N}(\lambda_{\text{hi}}, n/N)} \frac{1}{2} |k''_{\text{hi}}(u)| \right) \sum_{j=1}^S \sum_{m/N \in \mathcal{N}(\lambda_{\text{hi}}, t_j)} \left(t_j - \frac{m}{N} \right)^2 |h_m| \\
 &\stackrel{(b)}{\leq} \underbrace{\sum_{j=1}^S \left| \sum_{m/N \in \mathcal{N}(\lambda_{\text{hi}}, t_j)} h_m \right|}_{A_1} + c'_k \frac{1}{\lambda_{\text{hi}}} \underbrace{\sum_{j=1}^S \left| \sum_{m/N \in \mathcal{N}(\lambda_{\text{hi}}, t_j)} \left(t_j - \frac{m}{N} \right) h_m \right|}_{A_2} \\
 &\quad + c''_k \underbrace{\frac{1}{\lambda_{\text{hi}}^2} \sum_{j=1}^S \sum_{m/N \in \mathcal{N}(\lambda_{\text{hi}}, t_j)} \left(t_j - \frac{m}{N} \right)^2 |h_m|}_{A_3}. \tag{113}
 \end{aligned}$$

Above, (a) follows by periodicity of $k_{\text{hi}}(\cdot)$; (b) follows by (8), (106), (107).

To complete the proof of Theorem 1, it remains to upper-bound each of the terms $A_0, A_1, A_2,$ and A_3 by $\sim C(r) \text{SRF}^{2r} \|\mathbf{z}\|_1$. To do this we will use extended duality arguments that will rely on the trigonometric polynomials $q_0(\cdot), q_1(\cdot),$ and $q_2(\cdot)$.

7.2 Upper Bound on A_0

In this section we use the trigonometric polynomial $q_0(\cdot)$ from Lemma 3 to upper-bound A_0 . Let

$$\mathbf{q}^0 = [q_0^0, \dots, q_{N-1}^0]^\top \triangleq [q_0(l/N) : l \in [0 : N - 1]]^\top$$

be the vector that consists of the samples of $q_0(\cdot)$.

On the one hand,

$$\begin{aligned}
 \langle \mathbf{q}^0, \mathbf{h} \rangle &\stackrel{(a)}{=} \langle \mathbf{Q}\mathbf{q}^0, \mathbf{h} \rangle \stackrel{(b)}{=} \langle \mathbf{q}^0, \mathbf{Q}\mathbf{h} \rangle \\
 &\stackrel{(c)}{\leq} \|\mathbf{q}^0\|_\infty \|\mathbf{Q}\mathbf{h}\|_1 \\
 &\stackrel{(d)}{\leq} \|\mathbf{Q}\hat{\mathbf{x}} - \mathbf{s} + \mathbf{s} - \mathbf{Q}\mathbf{x}\|_1 \\
 &\stackrel{(e)}{\leq} \|\mathbf{Q}\hat{\mathbf{x}} - \mathbf{s}\|_1 + \|\mathbf{s} - \mathbf{Q}\mathbf{x}\|_1 \stackrel{(f)}{\leq} 2\|\mathbf{z}\|_1. \tag{114}
 \end{aligned}$$

Above, (a) follows because by Lemma 3, Property 1, $q_0(\cdot)$ is frequency-limited to f_{lo} , and, therefore, the vector of its samples is also frequency limited (in discrete sense) so that $\mathbf{q}^0 = \mathbf{Q}\mathbf{q}^0$; (b) follows because \mathbf{Q} is self-adjoint; (c) follows by Cauchy-Schwartz inequality; (d) follows by Lemma 3, Property 3; (e) follows by the triangle inequality; (f) follows since (CVX) implies $\|\mathbf{Q}\hat{\mathbf{x}} - \mathbf{s}\|_1 \leq \|\mathbf{Q}\mathbf{x} - \mathbf{s}\|_1$ and, by assumption, $\mathbf{s} = \mathbf{Q}\mathbf{x} + \mathbf{z}$.

On the other hand,

$$\langle \mathbf{q}^0, \mathbf{h} \rangle \stackrel{(a)}{=} \sum_{m=0}^{N-1} q_m^0 |h_m| \stackrel{(b)}{\geq} \sum_{m/N \in \mathcal{F}_{hi}} q_m^0 |h_m| \stackrel{(c)}{\geq} c_l^r \frac{\lambda_{hi}^{2r}}{(r\lambda_{lo})^{2r}} \sum_{m/N \in \mathcal{F}_{hi}} |h_m|. \tag{115}$$

Above, (a) follows because, by construction, $q_0(t) = 0$ for all $t \in \mathcal{T}$, which means that $h_m < 0$ implies $q_m^0 = q_0(m/N) = 0$, so that $q_m^0 h_m \geq 0$ for $m = 0, \dots, N - 1$; (b) follows because all terms in the sum are nonnegative; (c) follows from Lemma 3, Property 6. From (114) and (115), we conclude that

$$A_0 = \sum_{m/N \in \mathcal{F}_{hi}} |h_m| \leq \frac{r^{2r}}{c_l^r} \text{SRF}^{2r} \|\mathbf{z}\|_1, \tag{116}$$

where the equality follows because $h_m \geq 0$ for $m/N \in \mathcal{F}_{hi}$ and we remind the reader that $c_l < 1$.

7.3 Upper Bound on A_3

In this section we use \mathbf{q}^0 to upper-bound A_3 . We have,

$$\begin{aligned}
 2\|\mathbf{z}\|_1 &\stackrel{(a)}{\geq} \sum_{m=0}^{N-1} q_m^0 |h_m| \stackrel{(b)}{\geq} \sum_{j=1}^S \sum_{m/N \in \mathcal{N}(\lambda_{hi}, t_j)} q_m^0 |h_m| \\
 &\stackrel{(c)}{\geq} \sum_{j=1}^S \sum_{m/N \in \mathcal{N}(\lambda_{hi}, t_j)} c_{l2}^r \frac{(t_j - m/N)^2 \lambda_{hi}^{2(r-1)}}{(r\lambda_{lo})^{2r}} |h_m|.
 \end{aligned}$$

Above, (a) follows from (114) because $q_m^0 h_m \geq 0$ for $m = 0, \dots, N - 1$; (b) follows because the sets $\mathcal{N}(\lambda_{hi}, t_j)$ do not intersect since the elements of \mathcal{T} are separated by at least $2\lambda_{hi}$; (c) follows from (17). Hence,

$$A_3 = \frac{1}{\lambda_{hi}^2} \sum_{j=1}^S \sum_{m/N \in \mathcal{N}(\lambda_{hi}, t_j)} \left(t_j - \frac{m}{N}\right)^2 |h_m| \leq \frac{r^{2r}}{c_{l2}^r} \text{SRF}^{2r} 2 \|\mathbf{z}\|_1 \tag{117}$$

and we remind the reader that $c_{l2} < 1$.

7.4 Upper Bound on A_1

In this section we use trigonometric polynomial $q_1(\cdot)$ constructed in Lemma 4 to upper-bound A_1 . Set

$$\mathbf{q}^1 = [q_0^1, \dots, q_{N-1}^1]^T \triangleq [q_1(l/N) : l \in [0 : N - 1]]^T.$$

We now proceed as follows:

$$\begin{aligned} A_1 &= \sum_{j=1}^S \left| \sum_{m/N \in \mathcal{N}(\lambda_{hi}, t_j)} h_m \right| \\ &\stackrel{(a)}{=} \frac{2}{\rho} \sum_{j=1}^S \sum_{m/N \in \mathcal{N}(\lambda_{hi}, t_j)} \frac{\rho S j}{2} h_m \\ &\stackrel{(b)}{=} \frac{2}{\rho} \left| \sum_{j=1}^S \sum_{m/N \in \mathcal{N}(\lambda_{hi}, t_j)} \left(\frac{\rho S j}{2} - q_m^1\right) h_m + \sum_{j=1}^S \sum_{m/N \in \mathcal{N}(\lambda_{hi}, t_j)} q_m^1 h_m \right| \\ &\stackrel{(c)}{\leq} \underbrace{\frac{2}{\rho} \sum_{j=1}^S \sum_{m/N \in \mathcal{N}(\lambda_{hi}, t_j)} \left| \frac{\rho S j}{2} - q_m^1 \right| |h_m|}_{A_{11}} + \underbrace{\frac{2}{\rho} \left| \sum_{m/N \in \mathcal{N}_{hi}} q_m^1 h_m \right|}_{A_{12}}. \tag{118} \end{aligned}$$

Above, (a) follows by (26); (b) follows by adding and subtracting the corresponding term and because the expression in (a) is nonnegative; (c) follows by the triangle inequality and because the sets $\mathcal{N}(\lambda_{hi}, t_j)$ do not intersect since the elements of \mathcal{T} are separated by at least $2\lambda_{hi}$. Next, we upper-bound the terms A_{11} and A_{12} separately.

The first term in (118), A_{11} , can be upper-bounded as follows:

$$\begin{aligned} A_{11} &\stackrel{(a)}{\leq} r^{2r+4} c_{u27}^{r+1} \sum_{m/N \in \mathcal{N}_{hi}} q_m^0 |h_m| \\ &\stackrel{(b)}{=} r^{2r+4} c_{u27}^{r+1} \sum_{m/N \in \mathcal{N}_{hi}} q_m^0 h_m \end{aligned}$$

$$\stackrel{(c)}{\leq} r^{2r+4} c_{u27}^{r+1} \sum_{m=0}^{N-1} q_m^0 h_m \stackrel{(d)}{\leq} r^{2r+4} 2 c_{u27}^{r+1} \|\mathbf{z}\|_1. \tag{119}$$

Above, (a) follows by (27) and because the sets $\mathcal{N}(\lambda_{hi}, t_j)$ do not intersect; (b) follows because $h_m < 0$ implies $q_m^0 = 0$; (c) follows because $q_m^0 h_m \geq 0$ for $m = 0, \dots, N - 1$; (d) follows by (114).

Following exactly the same steps as in (114), changing q_m^0 to q_m^1 , and using in step (d) that by Lemma 4, Property 3, $\|\mathbf{q}^1\|_\infty \leq r^{2r+1} c_{u55}^r$, we obtain:

$$\left| \sum_{m=0}^{N-1} q_m^1 h_m \right| \leq 2r^{2r+1} c_{u55}^r \|\mathbf{z}\|_1. \tag{120}$$

Using this, the second term in (118), A_{12} , can be upper-bounded as follows

$$\begin{aligned} A_{12} &\stackrel{(a)}{\leq} \left| \sum_{m=0}^{N-1} q_m^1 h_m \right| + \left| \sum_{m/N \in \mathcal{F}_{hi}} q_m^1 h_m \right| \\ &\stackrel{(b)}{\leq} 2r^{2r+1} c_{u55}^r \|\mathbf{z}\|_1 + \sum_{m/N \in \mathcal{F}_{hi}} |q_m^1| |h_m| \\ &\stackrel{(c)}{\leq} 2r^{2r+1} c_{u55}^r \|\mathbf{z}\|_1 + r^{2r+2} c_{u29}^r \sum_{m/N \in \mathcal{F}_{hi}} q_m^0 |h_m| \\ &\stackrel{(d)}{=} 2r^{2r+1} c_{u55}^r \|\mathbf{z}\|_1 + r^{2r+2} c_{u29}^r \sum_{m/N \in \mathcal{F}_{hi}} q_m^0 h_m \\ &\stackrel{(e)}{\leq} 4r^{2r+2} c_{u57}^r \|\mathbf{z}\|_1. \end{aligned} \tag{121}$$

Above, (a) follow by the triangle inequality and because \mathcal{F}_{hi} is complementary to \mathcal{N}_{hi} ; (b) follow by (120) and by the triangle inequality; (c) follow by (28); (d) follows because $h_m > 0$ for $m/N \in \mathcal{F}_{hi}$; (e) follows by (114) because $q_m^0 h_m \geq 0$ for $m = 0, \dots, N - 1$, because $c_{u29} > 1$, and by defining $c_{u57} \triangleq \max(c_{u55}, c_{u29})$.

Substituting (119) and (121) into (118), using that $1/\rho = \text{SRF}^{2r}$, we finally obtain

$$A_1 = \sum_{j=1}^S \left| \sum_{m/N \in \mathcal{N}(\lambda_{hi}, t_j)} h_m \right| \leq r^{2r+4} c_{u53}^{r+1} \text{SRF}^{2r} \|\mathbf{z}\|_1, \tag{122}$$

where we defined $c_{u53} \triangleq 12 \max(c_{u27}, c_{u57})$.

7.5 Upper Bound on A_2

In this section we use trigonometric polynomial $q_2(\cdot)$ to upper-bound A_2 . Set

$$\mathbf{q}^2 = [q_0^2, \dots, q_{N-1}^2]^\top \triangleq [q_2(l/N) : l \in [0 : N - 1]]^\top.$$

We now proceed as follows:

$$\begin{aligned} A_2 &= \frac{1}{\lambda_{hi}} \sum_{j=1}^S \left| \sum_{m/N \in \mathcal{N}(\lambda_{hi}, t_j)} \left(\frac{m}{N} - t_j \right) h_m \right| \\ &\stackrel{(a)}{=} \frac{1}{\rho} \sum_{j=1}^S \sum_{m/N \in \mathcal{N}(\lambda_{hi}, t_j)} \gamma s'_j \left(\frac{m}{N} - t_j \right) h_m \\ &\stackrel{(b)}{=} \frac{1}{\rho} \left| \sum_{j=1}^S \sum_{m/N \in \mathcal{N}(\lambda_{hi}, t_j)} \left(\gamma s'_j \left(\frac{m}{N} - t_j \right) - q_m^2 \right) h_m + \sum_{j=1}^S \sum_{m/N \in \mathcal{N}(\lambda_{hi}, t_j)} q_m^2 h_m \right| \\ &\stackrel{(c)}{\leq} \frac{1}{\rho} \underbrace{\sum_{j=1}^S \sum_{m/N \in \mathcal{N}(\lambda_{hi}, t_j)} \left| \gamma s'_j \left(\frac{m}{N} - t_j \right) - q_m^2 \right| |h_m|}_{A_{21}} + \frac{1}{\rho} \underbrace{\left| \sum_{m/N \in \mathcal{N}_{hi}} q_m^2 h_m \right|}_{A_{22}}. \end{aligned} \tag{123}$$

Above, (a) follows by (103) and because $\gamma = \rho/\lambda_{hi}$; (b) follows by adding and subtracting the corresponding term and because the expression in (a) is nonnegative; (c) follows by the triangle inequality and because the sets $\mathcal{N}(\lambda_{hi}, t_j)$ do not intersect. Next, we upper-bound the terms A_{21} and A_{22} separately.

The first term in (123), A_{21} , can be upper-bounded as follows

$$\begin{aligned} A_{21} &\stackrel{(a)}{\leq} r^{2r+4} c_{u34}^{r+1} \sum_{m/N \in \mathcal{N}_{hi}} q_m^0 |h_m| \stackrel{(b)}{=} r^{2r+4} c_{u34}^{r+1} \sum_{m/N \in \mathcal{N}_{hi}} q_m^0 h_m \\ &\stackrel{(c)}{\leq} r^{2r+4} c_{u34}^{r+1} \sum_{m=0}^{N-1} q_m^0 h_m \stackrel{(d)}{\leq} 2r^{2r+4} c_{u34}^{r+1} \|\mathbf{z}\|_1. \end{aligned} \tag{124}$$

Above, (a) follows by (104) and because the sets $\mathcal{N}(\lambda_{hi}, t_j)$ do not intersect; (b) follows because $h_m < 0$ implies $q_m^0 = 0$; (c) follows because $q_m^0 h_m \geq 0$ for $m = 0, \dots, N - 1$; (d) follows by (114).

Following exactly the same steps as in (114), changing q_m^0 to q_m^2 , and using that by Lemma 5, Property 3, $\|\mathbf{q}_2\|_\infty \leq r^{2r+1} c_{u56}^r$, we obtain:

$$\left| \sum_{m=0}^{N-1} q_m^2 h_m \right| \leq 2r^{2r+1} c_{u56}^r \|\mathbf{z}\|_1. \tag{125}$$

Using this, the second term in (123), A_{22} , can be upper-bounded as follows

$$\begin{aligned}
 A_{22} &\stackrel{(a)}{\leq} \left| \sum_{m=0}^{N-1} q_m^2 h_m \right| + \left| \sum_{m/N \in \mathcal{F}_{hi}} q_m^2 h_m \right| \\
 &\stackrel{(b)}{\leq} 2r^{2r+1} c_{u56}^r \|\mathbf{z}\|_1 + \sum_{m/N \in \mathcal{F}_{hi}} |q_m^2| |h_m| \\
 &\stackrel{(c)}{\leq} 2r^{2r+1} c_{u56}^r \|\mathbf{z}\|_1 + r^{2r+2} c_{u52}^r \sum_{m/N \in \mathcal{F}_{hi}} q_m^0 |h_m| \\
 &\stackrel{(d)}{=} 2r^{2r+1} c_{u56}^r \|\mathbf{z}\|_1 + r^{2r+2} c_{u52}^r \sum_{m/N \in \mathcal{F}_{hi}} q_m^0 h_m \stackrel{(e)}{\leq} 4r^{2r+2} c_{u58}^r \|\mathbf{z}\|_1. \tag{126}
 \end{aligned}$$

Above, (a) follow by the triangle inequality and because \mathcal{F}_{hi} is complementary to \mathcal{N}_{hi} ; (b) follow by (125) and by the triangle inequality; (c) follow by (105); (d) follows because $h_m > 0$ for $m/N \in \mathcal{F}_{hi}$; (e) follows by (114) because $q_m^0 h_m \geq 0$ for $m = 0, \dots, N - 1$, because $c_{u52} > 1$, and by defining $c_{u58} \triangleq \max(c_{u56}, c_{u52})$.

Substituting (124) and (126) into (123), using that $1/\rho = \text{SRF}^{2r}$, we finally obtain:

$$A_2 = \frac{1}{\lambda_{hi}} \sum_{j=1}^S \left| \sum_{m/N \in \mathcal{N}(\lambda_{hi}, t_j)} h_m \right| \leq r^{2r+4} c_{u54}^{r+1} \text{SRF}^{2r} \|\mathbf{z}\|_1, \tag{127}$$

where we defined $c_{u54} \triangleq 6 \max(c_{u34}, c_{u58})$.

7.6 Putting Pieces Together

Substituting (116) into (109); substituting (117), (122), (127) into (113) and the result into (110); then, substituting (109) and (110) into (108), and defining

$$c \triangleq 4 \max(1/c_l, c_k''/c_{l2}, c_{u53}, c_k' c_{u54}) \tag{128}$$

we obtain the desired bound (9) and complete the proof of Theorem 1. □

8 Connection to Bernstein Theorem

The famous Bernstein theorem states the following [21, Ch. 4, eq. (1.1)].

Theorem 2 (Bernstein) Consider a trigonometric polynomial frequency-limited to f_c :

$$q(t) = \sum_{k=-f_c}^{f_c} \hat{q}_k e^{-i2\pi kt}. \text{ Then,}$$

$$\|q'(\cdot)\|_\infty \leq 2\pi f_c \|q(\cdot)\|_\infty. \tag{129}$$

In other words, if a trigonometric polynomial is uniformly bounded, its derivative cannot be too large anywhere.

Bernstein theorem helped us construct trigonometric polynomials $q_0(\cdot)$, $q_1(\cdot)$, and $q_2(\cdot)$ with the required properties by telling us what may be achievable and what is forbidden. We now describe these connections to provide more intuition about our constructions.

Independent control Consider $q(\cdot) = q_{\lambda_c, \mathcal{V}, \{f_j\}, \{d_j\}}(\cdot)$ in Fig. 4. Since we require $\|q(\cdot)\|_\infty \leq c_{u0}$, then, by Bernstein theorem, $\|q'(\cdot)\|_\infty \leq 2\pi c_{u0} f_c$. Suppose, $q(v_1) = 0$. How large $q(v_2)$ may possibly be? Since $\|q'(\cdot)\|_\infty \leq 2\pi c_{u0} f_c$, we must have $q(v_2) \leq 2\pi c_{u0}(v_2 - v_1) f_c$. Now, if the points v_1 and v_2 are well-separated, i.e., if $v_2 - v_1$ is order λ_c , Bernstein theorem puts *no restrictions* on $q(v_2)$. However, if $v_2 - v_1 \ll \lambda_c$, then $|q(v_1) - q(v_2)| \leq 2\pi c_{u0}(v_2 - v_1) f_c \ll 1$. Generalizing: it may be possible to independently control $q(v_1)$ and $q(v_2)$ *only* if the points v_1 and v_2 are well-separated. This is the reason why $q_0(\cdot)$, $q_1(\cdot)$, and $q_2(\cdot)$ are constructed in an *interlaced* way. We control the building blocks on sets of interlaced points that are well-separated, then we multiply the resulting trigonometric polynomials. See (21) and Fig. 6 for an easy example of interlacing; see (35), (36), Figs. 8, and 9 for a more sophisticated example of interlacing.

For readers familiar with using ℓ_1 -minimization for super-resolution of real-valued (spikes may be positive *and* negative) and complex-valued signals [13]: Bernstein theorem is responsible for the fact that ℓ_1 -minimization fails when the spikes are not well-separated (closer than λ_{l0} to one another). The dual certificate in the real-valued case is a trigonometric polynomial $q(\cdot)$ with $\|q(\cdot)\|_\infty \leq 1$ that interpolates the *sign* of the spikes in the signal. If, say $q(v_1) = -1$, and $v_2 - v_1 \ll \lambda_c$, it is not possible that $q(v_2) = +1$ because $\|q'(\cdot)\|_\infty \leq 2\pi f_c$. The required dual trigonometric polynomial does not exist and the algorithm fails.

In contrast to the real-valued case, consider our trigonometric polynomial $q_1(\cdot)$, displayed in Fig. 7. Here, we interpolate the sign of the sequence s_1, s_2, s_3, s_4 at a set of points t_1, t_2, t_3, t_4 that are *not* well-separated. How is that possible? The difference is that we interpolate the sign sequence at a *low* level $\rho = (\lambda_{hi}/\lambda_{lo})^{2r} \ll 1$, i.e., we interpolate the points $s_i \rho/2$, and not the points s_i . The transitions $q_1(\cdot)$ needs to make between the points, are small; for example $|q_1(t_3) - q_1(t_4)| = \rho \ll 1$ and this is not disallowed by Bernstein theorem.

High curvature As should be clear by now, the curvature of the building block $q_{\lambda_c, \mathcal{V}}(\cdot)$ in the vicinity of its zeros expressed by (13) (see also the sections marked in red in Fig. 4a) determines the noise amplification in our bounds. How curvy can $q_{\lambda_c, \mathcal{V}}(\cdot)$ possibly be? Since $\|q_{\lambda_c, \mathcal{V}}(\cdot)\|_\infty \leq 1$, applying Bernstein theorem twice, we conclude that the second derivative must satisfy $\|q''_{\lambda_c, \mathcal{V}}(\cdot)\|_\infty \leq 4\pi^2 f_c^2$. Therefore, for $v \in \mathcal{V}$, it must hold that $q_{\lambda_c, \mathcal{V}}(v - \tau) \leq 2\pi^2(v - \tau)^2/\lambda_c^2$. We conclude that the curvature of $q_{\lambda_c, \mathcal{V}}(\cdot)$ in (13) depends on λ_c in an optimal way (up to a constant). This leads to the near-optimal stability estimate in Theorem 1.

9 Numerical Results

In this section we describe a simple numerical experiment that demonstrates how the bounds developed in this paper reflect the true error accurately in the setting where the error metric used in the previous work [43] leads to an unreasonably pessimistic conclusion.

Set $f_{10} = 9$. For $\text{DSRF} \in \{8, 16, 32, 64, 128\}$ set $N = f_{10} \cdot \text{DSRF}$. For each N we generate 50 signals from each of the classes $\mathcal{R}(2\lambda_{10}, 1)$ and $\mathcal{R}(4\lambda_{10}, 2)$. To focus on the worst-case setup described by the theorems, the signals from $\mathcal{R}(2\lambda_{10}, 1)$ were generated so that pairs of spikes are forced to be close together. Specifically, for signals from $\mathcal{R}(4\lambda_{10}, 2)$, each spike has a pair that is no further than $3/N$ away (assuming the signal is depicted on the $(0, 1]$ interval as in Fig. 2). All spikes were chosen to have equal magnitude set to one. The observations are generated according to $\mathbf{s} = \mathbf{Q}\mathbf{x} + \mathbf{z}$. The noise is taken to be low-pass filtered Gaussian: $\mathbf{z} = c \cdot \mathbf{Q}\mathbf{z}_G$, where \mathbf{z}_G has independent identically distributed entries that are standard Gaussian with mean zero and unit variance, the constant c controls the signal-to-noise ratio (SNR) taken to be $\text{SNR} = \|\mathbf{x}\|_1 / \|\mathbf{z}\|_1 = 100000$ in all simulations.

In Fig. 10a we depict $\log\left(\frac{1}{50} \sum_{i=1}^{50} \|\mathbf{x}_i - \hat{\mathbf{x}}_i\|_1 / \|\mathbf{s}_i - \mathbf{Q}\mathbf{x}_i\|_1\right)$, i.e., the log of the noise amplification factor (NAF), averaged over the 50 trials, as a function of $\log(\text{DSRF})$. Observe that on the log-log scale the data are well approximated by straight lines (displayed). We found the slopes of these lines to be 1.02 for $\mathcal{R}(2\lambda_{10}, 1)$ data and 2.85 for $\mathcal{R}(4\lambda_{10}, 2)$ data. This matches the conclusions of Proposition 1 and the corresponding converse [43, Sec. 2.3], which together predict $(2r - 1) \log(\text{SRF}) \leq \log(\|\mathbf{x} - \hat{\mathbf{x}}\|_1 / \|\mathbf{s} - \mathbf{Q}\mathbf{x}\|_1) \leq (2r) \log(\text{SRF})$. Note that for $r = 2$, the exponent in the NAF, 2.85, estimated in the simulation, happened to be a bit more optimistic than the range between $2r - 1 = 3$ and $2r = 4$ predicted by theory, presumably because the Gaussian noise is not the worst case. In Fig. 10b, for reference, we depict the average relative error, $\frac{1}{50} \sum_{i=1}^{50} \|\mathbf{x}_i - \hat{\mathbf{x}}_i\|_1 / \|\mathbf{x}_i\|_1$ and observe that for all DSRF values and for both ensembles, the average relative error is significantly below one, implying that the reconstructed signals are high quality.

Next, consider a 20 times higher super-resolution factor $\text{DSRF}_f \in \{160, 320, 640, 1280, 2560\}$ with $N_f = f_{10} \cdot \text{DSRF}_f = 20N$ (' f ' stands for fine-scale) and repeat the above experiment with all other parameters kept the same. We use the same random signals from the experiment above, when depicted on the interval $(0, 1]$. Since the grid is now 20 times finer, in the vector representation the index of each nonzero element of \mathbf{x} has been multiplied by 20. To be sure: if before the first and the second elements of \mathbf{x} were equal to one, all others being zero, now the 20th and the 40th elements of \mathbf{x} are equal to one, all others being zero. This construction guarantees that for signals from $\mathcal{R}(4\lambda_{10}, 2)$, each spike has a pair that is no further than $3 \cdot 20/N_f = 3/N$ away (assuming the signal is depicted on the $(0, 1]$ interval as in Fig. 2), precisely as before.

In Fig. 10c we depict $\log\left(\frac{1}{50} \sum_{i=1}^{50} \|\mathbf{x}_i - \hat{\mathbf{x}}_i\|_1 / \|\mathbf{s}_i - \mathbf{Q}\mathbf{x}_i\|_1\right)$, i.e., the log of the NAF, averaged over the 50 trials as a function of $\log(\text{DSRF}_f)$ for the fine-scale experiment. The curve for signals from $\mathcal{R}(2\lambda_{10}, 1)$ [blue diamonds] looks very similar to how it did for the coarse grid in Fig. 10a. In contrast, the curve for signals from

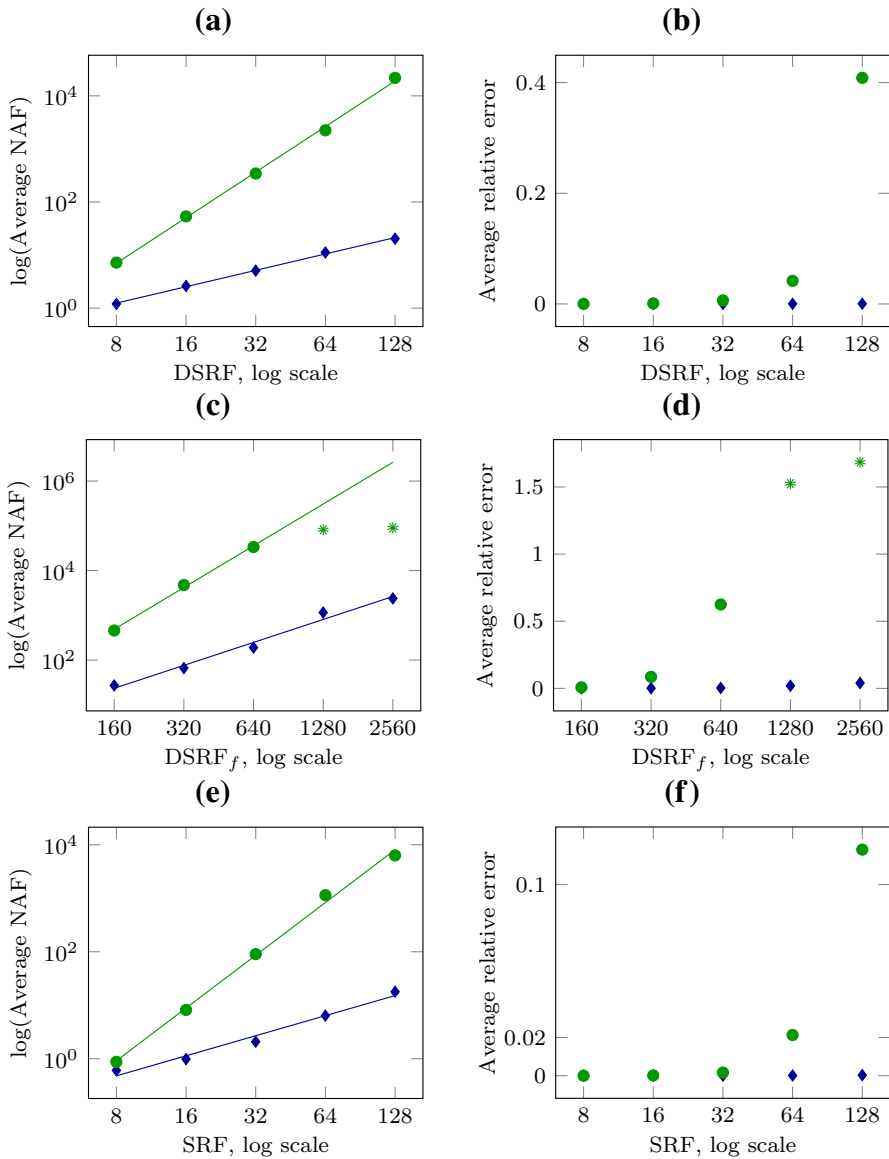


Fig. 10 In **a** and **c**, $\log(\frac{1}{50} \sum_{i=1}^{50} \|\mathbf{x}_i - \hat{\mathbf{x}}_i\|_1 / \|\mathbf{s}_i - \mathbf{Q}\mathbf{x}_i\|_1)$, as a function of $\log(\text{DSRF})$ and $\log(\text{DSRF}_f)$, respectively. In **e**, $\log(\frac{1}{50} \sum_{i=1}^{50} \|k_{\text{hi}} \star (\mathbf{x}_i - \hat{\mathbf{x}}_i)\|_1 / \|\mathbf{s}_i - \mathbf{Q}\mathbf{x}_i\|_1)$, as a function of $\log(\text{SRF})$. In **b** and **d**, $\frac{1}{50} \sum_{i=1}^{50} \|\mathbf{x}_i - \hat{\mathbf{x}}_i\|_1 / \|\mathbf{x}_i\|_1$, as a function of $\log(\text{DSRF})$ and $\log(\text{DSRF}_f)$, respectively. In **f**, $\frac{1}{50} \sum_{i=1}^{50} \|k_{\text{hi}} \star (\mathbf{x}_i - \hat{\mathbf{x}}_i)\|_1 / \|\mathbf{x}_i\|_1$, as a function of $\log(\text{DSRF})$. In all plots, blue diamonds represent signals from class $\mathcal{R}(2\lambda_{10}, 1)$ and green circles (and green stars) represent signals from class $\mathcal{R}(4\lambda_{10}, 2)$ (Color figure online)

$\mathcal{R}(4\lambda_{10}, 2)$ [green circles and stars] now looks different: it saturates for large values of DSRF_f ($\text{DSRF}_f \in \{1280, 2560\}$), as marked by the green stars). By looking at Fig. 10d, where we depict the average relative error $\frac{1}{50} \sum_{i=1}^{50} \|\mathbf{x}_i - \hat{\mathbf{x}}_i\|_1 / \|\mathbf{x}_i\|_1$ as a function of $\log(\text{DSRF}_f)$, we see that for the large values of DSRF_f , the norm of the error, $\|\mathbf{x}_i - \hat{\mathbf{x}}_i\|_1$, is comparable to the norm of the signal $\|\mathbf{x}_i\|_1$, so that the error is nearly as large as it can possibly be (see the points marked by green stars). Does this mean that the algorithm produced bad reconstructions? The point of this paper is that the reconstructions are still good, even for the high values of $\text{DSRF}_f \in \{1280, 2560\}$, if the error is measured at the appropriate scale. To see this, in Fig. 10f we depict $\frac{1}{50} \sum_{i=1}^{50} \|k_{\text{hi}} \star (\mathbf{x}_i - \hat{\mathbf{x}}_i)\|_1 / \|\mathbf{x}_i\|_1$ with $\lambda_{\text{hi}} = 1/(\text{DSRF} \cdot f_{10})$ so that $\text{SRF} = \lambda_{10}/\lambda_{\text{hi}} = \text{DSRF}$, matching the resolution of the first experiment, as a function of $\log(\text{SRF})$. We observe that $\|k_{\text{hi}} \star (\mathbf{x}_i - \hat{\mathbf{x}}_i)\|_1 \ll \|\mathbf{x}_i\|_1$ for all values of SRF . In other words, if the error is measured at the scale λ_{hi} the reconstruction is good for all values of SRF . For example, for $\text{DSRF}_f = 1280$ in Fig. 10d the error on average is 1.5 times larger than the signal (in 11-norm), but for the corresponding $\text{SRF} = 64$ in Fig. 10f, the 11-norm of the error on average is only 0.02 of the 11-norm of the signal—small. By comparing Fig. 10d and f, we infer that in the fine-scale case and for high $\text{DSRF} \in \{1280, 2560\}$ the reconstruction algorithm disperses the spikes, but the estimated dispersed spikes are still concentrated around the correct locations with dispersion smaller than λ_{hi} (recall illustration in Fig. 3). Next, in Fig. 10e we depict $\log\left(\frac{1}{50} \sum_{i=1}^{50} \|k_{\text{hi}} \star (\mathbf{x}_i - \hat{\mathbf{x}}_i)\|_1 / \|\mathbf{s}_i - \mathbf{Q}\mathbf{x}_i\|_1\right)$ as a function of $\log(\text{SRF})$. We see that the saturation effect disappeared and, as in the first experiment, on the log-log scale the data are well approximated by straight lines (displayed). We found the slopes of these lines to be 1.25 for $\mathcal{R}(2\lambda_{10}, 1)$ data and 3.28 for $\mathcal{R}(4\lambda_{10}, 2)$ data. This demonstrates that the bounds in Theorem 1 together with the converse [43, Sec. 2.3] reflect the true error on the scale λ_{hi} accurately: $(2r - 1) \log(\text{SRF}) \leq \log(\|k_{\text{hi}} \star (\mathbf{x} - \hat{\mathbf{x}})\|_1 / \|\mathbf{s} - \mathbf{Q}\mathbf{x}\|_1) \leq (2r) \log(\text{SRF})$, i.e., the exponents in the NAF are predicted to be between 1 and 2 for $r = 1$ and between 3 and 4 for $r = 2$, as observed in the experiment.

10 Conclusion

When a signal is positive and Rayleigh-regular, then linear programming solves the super-resolution problem with near-optimal worst-case performance. This result holds independently on how fine the discretization grid is, approximating the continuum arbitrarily closely. The proof relies on new trigonometric interpolation constructions; the underlying ideas might be useful for other problems.

Finding an efficient algorithm that solves the same problem with a near-optimal worst-case performance for complex-valued signals is still an open problem. Despite recent work that derives stability estimates for MUSIC and ESPRIT algorithms in certain cases, the question of how far are these algorithms from the optimal performance is not yet answered completely.

In a different direction, all results in this paper are valid for the 1D model. The discrete results have been generalized to the 2D model in [43]. It would be interesting to see if the techniques developed in this paper may also be extended to the 2D model.

Acknowledgements V. M. was supported by the Simons Foundation when developing the early ideas that led to this work. The author is grateful to Emmanuel Candès for inspiring and useful discussions.

Proof of Lemma 2

Let

$$g(t) \triangleq \left[\frac{\sin(\pi(f_c/2 + 1)t)}{(f_c/2 + 1)\sin(\pi t)} \right]^4$$

and set

$$q(t) \triangleq \sum_{j=1}^V \alpha_j g(t - v_j) + \beta_j g'(t - v_j),$$

where $\{\alpha_j\}_{j=1}^V$ and $\{\beta_j\}_{j=1}^V$ are free coefficients that will be determined in the following. Because $g(\cdot)$ is frequency-limited to $[-f_c, f_c]$ [cf. (7), (196)], $q(\cdot)$ satisfies Property 1. Note,

$$q'(t) = \sum_{j=1}^V \alpha_j g'(t - v_j) + \beta_j g''(t - v_j).$$

Define matrices $\mathbf{D}_0, \mathbf{D}_1, \mathbf{D}_2 \in \mathbb{R}^{V \times V}$ with the elements

$$[\mathbf{D}_0]_{jk} = g(v_j - v_k), \quad [\mathbf{D}_1]_{jk} = g'(v_j - v_k), \quad [\mathbf{D}_2]_{jk} = g''(v_j - v_k).$$

To satisfy the interpolation constraints in Property 2 we define $\alpha = [\alpha_1, \dots, \alpha_V]^T$, $\beta = [\beta_1, \dots, \beta_V]^T$, $\mathbf{f} = [f_1, \dots, f_V]^T$, $\mathbf{d} = [d_1, \dots, d_V]^T$, demand

$$\underbrace{\begin{bmatrix} \mathbf{D}_0 & \mathbf{D}_1 \\ \mathbf{D}_1 & \mathbf{D}_2 \end{bmatrix}}_{\mathbf{D}} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} \mathbf{f} \\ \mathbf{d} \end{bmatrix}$$

and solve for α and β . It can be verified that \mathbf{D}_0 and \mathbf{D}_2 are both invertible; the corresponding Schur complements

$$\begin{aligned} \mathbf{E} &= \mathbf{D}_2 - \mathbf{D}_1 \mathbf{D}_0^{-1} \mathbf{D}_1 \\ \mathbf{F} &= \mathbf{D}_0 - \mathbf{D}_1 \mathbf{D}_2^{-1} \mathbf{D}_1 \end{aligned}$$

are well defined and are also both invertible (see [13, Sec. 2.3.1, pp. 925–926], [12, App. B, p. 1249] for the relevant results). Therefore, \mathbf{D} is invertible and the inverse can be written as [41, Sec. 9.11.3.(2)]

$$\mathbf{D}^{-1} = \begin{bmatrix} \mathbf{F}^{-1} & -\mathbf{D}_0^{-1}\mathbf{D}_1\mathbf{E}^{-1} \\ -\mathbf{D}_2^{-1}\mathbf{D}_1\mathbf{F}^{-1} & \mathbf{E}^{-1} \end{bmatrix}.$$

We know (see [13, Sec. 2.3.1, pp. 925–926], [12, App. B, p. 1249]):

$$\|\mathbf{D}_1\|_\infty \leq \frac{\tilde{c}_1}{\lambda_c}, \tag{130}$$

$$\|\mathbf{D}_0^{-1}\|_\infty \leq \frac{1}{1 - \tilde{c}_0} = \bar{c}_0, \tag{131}$$

$$\|\mathbf{E}^{-1}\|_\infty \leq \left(\frac{\pi^2 f_c (f_c + 4)}{3} - \left(\tilde{c}_2 + \frac{\tilde{c}_1^2}{1 - \tilde{c}_0} \right) f_c^2 \right)^{-1} \leq \bar{c}_E \lambda_c^2, \tag{132}$$

$$\begin{aligned} \|\mathbf{F}^{-1}\|_\infty &\leq \bar{c}_F, \\ \|\mathbf{D}_2^{-1}\|_\infty &\leq \bar{c}_2 \lambda_c^2. \end{aligned} \tag{133}$$

Above, $\|\mathbf{A}\|_\infty$ is the infinity norm of a matrix defined as

$$\|\mathbf{A}\|_\infty = \max_{\|\mathbf{y}\|_\infty=1} \|\mathbf{A}\mathbf{y}\|_\infty = \max_i \sum_j |a_{ij}|$$

and

$$\begin{aligned} \tilde{c}_0 &\leq 0.007, \quad \tilde{c}_1 \leq 0.08, \quad \tilde{c}_2 \leq 1.06, \\ \bar{c}_0 &\leq 1.008, \quad \bar{c}_E \leq 0.47, \quad \bar{c}_2 \leq 0.43, \quad \bar{c}_F \leq 1.009. \end{aligned}$$

Now we have

$$\begin{aligned} \|\alpha\|_\infty &= \|\mathbf{F}^{-1}\mathbf{f} - \mathbf{D}_0^{-1}\mathbf{D}_1\mathbf{E}^{-1}\mathbf{d}\|_\infty \\ &\leq \|\mathbf{F}^{-1}\mathbf{f}\|_\infty + \|\mathbf{D}_0^{-1}\mathbf{D}_1\mathbf{E}^{-1}\mathbf{d}\|_\infty \\ &\leq \|\mathbf{F}^{-1}\|_\infty \|\mathbf{f}\|_\infty + \|\mathbf{D}_0^{-1}\mathbf{D}_1\mathbf{E}^{-1}\|_\infty \|\mathbf{d}\|_\infty \\ &\stackrel{(a)}{\leq} \|\mathbf{F}^{-1}\|_\infty + \frac{1}{\lambda_c} \|\mathbf{D}_0^{-1}\|_\infty \|\mathbf{D}_1\|_\infty \|\mathbf{E}^{-1}\|_\infty \\ &\stackrel{(b)}{\leq} \bar{c}_F + \bar{c}_0 \tilde{c}_1 \bar{c}_E \triangleq c_\alpha. \end{aligned}$$

Above, (a) follows because $|f_j| \leq 1$ and $|d_j| \leq 1/\lambda_c$ for all $j = 1, \dots, V$; (b) follows by (130), (131), (132), and (133); and c_α can be upper-bounded as $c_\alpha \leq 1.05$. Similarly,

$$\|\beta\|_\infty = \|-\mathbf{D}_2^{-1}\mathbf{D}_1\mathbf{F}^{-1}\mathbf{f} + \mathbf{E}^{-1}\mathbf{d}\|_\infty$$

$$\begin{aligned}
 &\leq \|-\mathbf{D}_2^{-1}\mathbf{D}_1\mathbf{F}^{-1}\mathbf{f}\|_\infty + \|\mathbf{E}^{-1}\mathbf{d}\|_\infty \\
 &\leq \|-\mathbf{D}_2^{-1}\mathbf{D}_1\mathbf{F}^{-1}\|_\infty\|\mathbf{f}\|_\infty + \|\mathbf{E}^{-1}\|_\infty\|\mathbf{d}\|_\infty \\
 &\leq \|\mathbf{D}_2^{-1}\|_\infty\|\mathbf{D}_1\|_\infty\|\mathbf{F}^{-1}\|_\infty + \frac{1}{\lambda_c}\|\mathbf{E}^{-1}\|_\infty \\
 &\leq c_\beta\lambda_c
 \end{aligned}$$

with $c_\beta \triangleq \bar{c}_2\bar{c}_1\bar{c}_F + \bar{c}_E$ that can be upper-bounded as $c_\beta \leq 0.51$.

The following lemma, proven in the end of this section, records bounds on $\sum_{j=1}^V |g^{(l)}(t - v_j)|, l = 0, 1$.

Lemma 6 *The following estimates hold:*

$$\begin{aligned}
 \sum_{j=1}^V |g(t - v_j)| &\leq c_{s0}, \\
 \sum_{j=1}^V |g'(t - v_j)| &\leq c_{s1}/\lambda_c,
 \end{aligned}$$

where c_{s0}, c_{s1} are positive numerical constants defined in the proof of the lemma below.

Using the bounds we obtain the required estimates as follows. Observe,

$$\begin{aligned}
 |q(t)| &= \left| \sum_{j=1}^V \alpha_j g(t - v_j) + \beta_j g'(t - v_j) \right| \\
 &\leq \|\alpha\|_\infty \underbrace{\sum_{j=1}^V |g(t - v_j)|}_{\leq c_{s0}} + \|\beta\|_\infty \underbrace{\sum_{j=1}^V |g'(t - v_j)|}_{\leq c_{s1}/\lambda_c} \\
 &\leq c_\alpha c_{s0} + c_\beta c_{s1} \triangleq c_{u0}.
 \end{aligned} \tag{134}$$

This proves Property 3. Property 4 in the lemma follows from (134) by (129) [Bernstein theorem], using that $q'(t)$ is also a trigonometric polynomial frequency-limited to f_c :

$$|q'(t)| \leq c_{u1} f_c, \quad c_{u1} \triangleq 2\pi c_{u0}. \tag{135}$$

In turn, Property 5 follows from (135) by (129) [Bernstein theorem], using that $q''(t)$ is also a trigonometric polynomial frequency-limited to f_c :

$$|q''(t)| \leq c_{u2} f_c^2, \quad c_{u2} \triangleq 4\pi^2 c_{u0}.$$

This completes the proof of Lemma 2. □

Proof of Lemma 6 For all $t \in [-1/2, 1/2]$, we have the following bounds [13, Sec. 2.3.2, p. 928]:

$$|g(t)| \leq 1, \tag{136}$$

$$|g'(t)| \leq \frac{\pi^2}{3} f_c(f_c + 4)|t|. \tag{137}$$

For all t with $\lambda_c/2 \leq |t| \leq 1/2$ and $l = 0, 1$, by inspection it follows from [13, Lm. 2.6] that the following bound holds:

$$\left|g^{(l)}(t)\right| \leq \frac{\pi^l c_g^l}{(f_c + 2)^{4-l} t^4} \tag{138}$$

with $c_g^0 \triangleq 1, c_g^1 \triangleq 6$. [To obtain this result from [13, Lm. 2.6], observe that, in the terminology of [13], for all t with $\lambda_c/2 \leq |t| \leq \sqrt{2}/\pi, b(t) < 2a(t)$ and $a(t) < 1$.]

Define $u_{k_j} \triangleq t - v_j$, ordered in such a way that $|u_1| < \dots < |u_V|$. Since $\{v_1, v_1, \dots, v_V\} \in \mathcal{R}(\kappa\lambda_c, 1)$, we have

$$|u_{2j}| > \lambda_c \kappa(j - 1) \text{ and } |u_{2j-1}| > \lambda_c \kappa(j - 1), \text{ for } j \geq 2, \tag{139}$$

and also,

$$|u_j| > \frac{\lambda_c}{2} \kappa j, \text{ for } j > 1. \tag{140}$$

First,

$$\begin{aligned} \sum_{j=1}^V |g(t - v_j)| &= \sum_{j=1}^V |g(u_j)| \stackrel{(a)}{\leq} 2 + \frac{1}{(f_c + 2)^4} \sum_{j=3}^V \frac{1}{u_j^4} \\ &\stackrel{(b)}{\leq} 2 + \frac{2}{(f_c + 2)^4} \sum_{j=1}^{\lceil V/2 \rceil} \frac{1}{(\lambda_c \kappa j)^4} \\ &\leq 2 + \frac{1}{f_c^4} \frac{2}{(\lambda_c \kappa)^4} \sum_{j=1}^{\infty} \frac{1}{j^4} \stackrel{(c)}{\leq} 2 + \frac{2}{\kappa^4} \frac{4}{3} \stackrel{(d)}{\leq} c_{s0}. \end{aligned}$$

Above, in (a) we used (136) to bound the terms for $j = 1, 2$ and used (138) to bound the terms for $j > 2$ [(138) is applicable because $u_3 > \lambda_c \kappa/2 > \lambda_c/2$ since $\kappa=1.87$]; in (b) we used (139); in (c) we used

$$\sum_{j=1}^{\infty} \frac{1}{j^4} = \frac{\pi^2}{90} \leq \frac{4}{3};$$

and in (d) we defined $c_{s0} \triangleq 2.22$.

Second,

$$\begin{aligned}
 \sum_{j=1}^V |g'(t - v_j)| &= |g'(u_1)| + \sum_{j=2}^V |g'(u_j)| \\
 &\stackrel{(a)}{\leq} \frac{\pi^2}{3} f_c(f_c + 4) \frac{\lambda_c}{2} \kappa + \frac{1}{(f_c + 2)^3} \frac{6\pi}{((\lambda_c/2)\kappa)^4} + \sum_{j=2}^V |g'(u_j)| \\
 &\stackrel{(b)}{\leq} \frac{\pi^2}{3} f_c(f_c + 4) \frac{\lambda_c}{2} \kappa + \frac{1}{(f_c + 2)^3} \frac{6\pi}{((\lambda_c/2)\kappa)^4} + \frac{1}{(f_c + 2)^3} \sum_{j=2}^V \frac{6\pi}{u_j^4} \\
 &\stackrel{(c)}{\leq} \frac{\pi^2}{3} f_c(f_c + 4) \frac{\lambda_c}{2} \kappa + \frac{1}{(f_c + 2)^3} \sum_{j=1}^V \frac{6\pi}{((\lambda_c/2)\kappa j)^4} \\
 &\leq \frac{\pi^2}{6} (f_c + 4)\kappa + \frac{6\pi}{f_c^3} \frac{2^4}{(\lambda_c \kappa)^4} \sum_{j=1}^{\infty} \frac{1}{j^4} \\
 &\stackrel{(d)}{\leq} \frac{\pi^2}{6} (f_c + 4)\kappa + 35f_c \leq 12.4 + 38.1 f_c \stackrel{(e)}{\leq} c_{s1} f_c.
 \end{aligned}$$

Above, in (a) we used that if $|u_1| < \lambda_c \kappa / 2$, then $|g'(u_1)| < (\pi^2/3) f_c(f_c + 4)\lambda_c \kappa / 2$ by (137) and otherwise $|g'(u_1)| < 6\pi / [(f_c + 2)^3 ((\lambda_c/2)\kappa)^4]$ by (138); in (b) we used (138), which is applicable because $|u_j| > \lambda_c \kappa > \lambda_c / 2$ by (140); in (c) we used (140); in (d) we used $\sum_{j=1}^{\infty} 1/j^4 < 4/3$; in (e) we used that $f_c > 128$ and defined $c_{s1} \triangleq 38.2$. □

Proof of Lemma 5

Construction

We first describe how the trigonometric polynomial $q_2(\cdot)$ is constructed. In the following four subsections, we prove that the construction is valid and that it satisfies the required Properties 1–4.

Recall, $\mathcal{T} = \{t_1, \dots, t_S\}$ is defined in (12) and, as before, define $\mathcal{T}_k, k = 1, \dots, r$, as in (20); remember that $\mathcal{T} = \mathcal{T}_1 \cup \dots \cup \mathcal{T}_r$ and $\mathcal{T}_k \in \mathcal{R}(\kappa \lambda_{10} r, 1)$.

We will construct the trigonometric polynomial $q_2(\cdot)$ as a (shifted) sum of r trigonometric polynomials $\{\phi_k(\cdot)\}_{k=1}^r$:

$$q_2(t) = \sum_{k=1}^r \phi_k(t) - \rho. \tag{141}$$

Note that we are overloading the notations here and $\phi_k(\cdot)$ in this section are *different* from $\phi_k(\cdot)$ in Sect. 6.3.1. Each of the trigonometric polynomials $\{\phi_k(\cdot)\}_{k=1}^r$ is

frequency-limited to f_{10} ,

$$\phi_k(t) = \sum_{l=-f_{10}}^{f_{10}} \hat{\phi}_{k,l} e^{-i2\pi lt} \quad \text{for some } \hat{\phi}_{k,l} \in \mathbb{C} \tag{142}$$

and is constructed separately to satisfy the following interpolation constraints on \mathcal{T} :

$$\phi_k(t_l) = \begin{cases} \rho, & \text{if } t_l \in \mathcal{T}_k, \\ 0, & \text{if } t_l \in \mathcal{T}_k^c \triangleq \mathcal{T} \setminus \mathcal{T}_k, \end{cases} \tag{143}$$

$$\phi'_k(t_l) = \begin{cases} \gamma s'_l, & \text{if } t_l \in \mathcal{T}_k, \\ 0, & \text{if } t_l \in \mathcal{T}_k^c. \end{cases} \tag{144}$$

Constraints (143), (144), and definition (141) guarantee, for all $l = 1, \dots, S$,

$$q_2(t_l) = 0, \tag{145}$$

$$q'_2(t_l) = \gamma s'_l. \tag{146}$$

To develop intuition about our construction, observe that (142) and (141) guarantee that Property 1 is satisfied. Further, observe that interpolation constraints (145) and (146) are needed for (104) to hold because $q_0(t) = q'_0(t) = 0$ for all $t \in \mathcal{T}$.

Next, we explain how to construct the trigonometric polynomials $\phi_k(\cdot)$, $k = 1, \dots, r$. The idea is to construct $\phi_k(\cdot)$ as a product of two trigonometric polynomials:

$$\phi_k(t) \triangleq \phi_{0,k}(t) \times \phi_{+,k}(t). \tag{147}$$

The first term in the product is defined as

$$\phi_{0,k}(t) \triangleq \prod_{l \neq k} q_{r\lambda_{10}, \mathcal{T}_l}(t), \tag{148}$$

where $q_{r\lambda_{10}, \mathcal{T}_l}(\cdot)$, $l = 1, \dots, r$ are the trigonometric polynomials constructed via Lemma 1 with $\lambda_c = r\lambda_{10}$ and $\mathcal{V} = \mathcal{T}_l \in \mathcal{R}(\kappa\lambda_{10}r, 1)$. The second term,

$$\phi_{+,k}(t) \triangleq r^{2r} c_{u31}^r q_{r\lambda_{10}, \mathcal{T}_k, \{f_j\}, \{d_j\}}(t) \tag{149}$$

is a (rescaled) trigonometric polynomial $q_{r\lambda_{10}, \mathcal{T}_k, \{f_j\}, \{d_j\}}(\cdot)$ constructed via Lemma 2 with $\lambda_c = r\lambda_{10}$, and $\mathcal{V} = \mathcal{T}_k \in \mathcal{R}(\kappa\lambda_{10}r, 1)$ and c_{u31} is a positive numerical constant defined in (160) below. Further, the function-values and derivatives of $q_{r\lambda_{10}, \mathcal{T}_k, \{f_j\}, \{d_j\}}(\cdot)$ are constrained on \mathcal{T}_k so that $\phi_{+,k}(\cdot)$ satisfies the following:

$$\phi_{+,k}(t) = \rho \frac{1}{\phi_{0,k}(t)}, \quad \text{for all } t \in \mathcal{T}_k, \tag{150}$$

$$\phi'_{+,k}(t_l) = -\rho \frac{\phi'_{0,k}(t_l)}{\phi^2_{0,k}(t_l)} + \frac{\gamma s'_l}{\phi_{0,k}(t_l)}, \text{ for all } t_l \in \mathcal{T}_k. \tag{151}$$

Note that $\phi_{0,k}(\cdot)$ in this section is *identical* to $\phi_{0,k}(\cdot)$ defined in Sect. 6.3.1 and, therefore, satisfies all the properties derived in Sect. 6.3.2; $\phi_{+,k}(\cdot)$ in this section is *different* from $\phi_{+,k}(\cdot)$ in Sect. 6.3.1 and the notation is overloaded.

We will prove in the next subsection, that this specification is valid, in the sense that the corresponding function values and derivatives of $q_{r\lambda_{10}, \mathcal{T}_k, \{f_j\}, \{d_j\}}(\cdot)$ on \mathcal{T}_k satisfy requirements (15) of Lemma 2.

It follows from (147), (148), (150), Lemma 1, Properties 2, 4, and 5, that the interpolation constraint (143) is satisfied:

$$\begin{aligned} \phi_k(t) &= \phi_{0,k}(t)\phi_{+,k}(t) = \rho, \text{ for all } t \in \mathcal{T}_k, \\ \phi_k(t) &= \underbrace{\phi_{0,k}(t)}_0 \phi_{+,k}(t) = 0, \text{ for all } t \in \mathcal{T}_k^c. \end{aligned}$$

Next, by (147),

$$\phi'_k(t) = \phi'_{0,k}(t)\phi_{+,k}(t) + \phi_{0,k}(t)\phi'_{+,k}(t).$$

Therefore, by (150) and (151),

$$\begin{aligned} \phi'_k(t_l) &= \phi'_{0,k}(t_l)\phi_{+,k}(t_l) + \phi_{0,k}(t_l)\phi'_{+,k}(t_l) \\ &= \phi'_{0,k}(t_l)\rho \frac{1}{\phi_{0,k}(t_l)} + \phi_{0,k}(t_l) \left(-\rho \frac{\phi'_{0,k}(t_l)}{\phi^2_{0,k}(t_l)} + \frac{\gamma s'_l}{\phi_{0,k}(t_l)} \right) = \gamma s'_l, \end{aligned}$$

for all $t_l \in \mathcal{T}_k$. Further, by (148), Lemma 1, Property 2,

$$\phi'_k(t) = \underbrace{\phi'_{0,k}(t)}_0 \phi_{+,k}(t) + \underbrace{\phi_{0,k}(t)}_0 \phi'_{+,k}(t) = 0,$$

for all $t \in \mathcal{T}_k^c$. We conclude that the interpolation constraint (144) is satisfied.

Finally, (142) follows from (147) because $\phi_{0,k}(\cdot)$ in (148) is frequency-limited to $(r - 1)/(\lambda_{10}r)$ [Lemma 1, Property 1] and $\phi_{+,k}(\cdot)$ in (149) is frequency-limited to $1/(\lambda_{10}r)$ [Lemma 2, Property 1] so that $\phi_k(\cdot)$ is frequency-limited to $(r - 1)/(\lambda_{10}r) + 1/(\lambda_{10}r) = 1/\lambda_{10} = f_{10}$. Therefore, by (141), $q_2(\cdot)$ is also frequency-limited to f_{10} , which proves Property 1.

Existence of $\phi_{+,k}(\cdot)$

In this subsection, we check that the trigonometric polynomial $\phi_{+,k}(\cdot)$ that satisfies (150) and (151) can be defined according to (149) with $q_{r\lambda_{10}, \mathcal{T}_k, \{f_j\}, \{d_j\}}(\cdot)$ constructed via Lemma 2 with $\lambda_c = r\lambda_{10}$ and $\mathcal{V} = \mathcal{T}_k \in \mathcal{R}(\kappa\lambda_{10}r, 1)$. To this end, we need to show that the constraints on the function values $\{f_j\}$ and on the derivatives

$\{d_j\}$ that are implied by the constraints (150) and (151) satisfy requirements (15) of Lemma 2.

First consider the case $r = 1$. As already discussed, in this case $\phi_{0,k}(t) = 1$ for all t , and, therefore, $\phi'_{0,k}(t) = 0$ for all t . Plugging these values into (150) and (151) we see from (149) that the requirements (15) of Lemma 2 are satisfied.

Next, consider the case $r > 1$.

To check that requirements (15) are satisfied for $t \in \mathcal{T}_k$, we need to find upper bounds on $|\phi_{+,k}(t)|$ and on $|\phi'_{+,k}(t)|$.

Take $t \in \mathcal{T}_k$ and observe:

$$|\phi_{+,k}(t)| \stackrel{(a)}{=} \left| \rho \frac{1}{\phi_{0,k}(t)} \right| \stackrel{(b)}{\leq} \frac{\lambda_{hi}^{2r}}{\lambda_{lo}^{2r}} \frac{1}{\frac{\lambda_{hi}^{2(r-1)}}{(r\lambda_{lo})^{2(r-1)}}} \frac{1}{c_l^{r-1}}$$

$$\stackrel{(c)}{\leq} r^{2(r-1)} \frac{1}{c_l^r} \frac{\lambda_{hi}^2}{\lambda_{lo}^2} \tag{152}$$

$$\leq r^{2r} \frac{1}{c_l^r}. \tag{153}$$

Above, (a) follows by (150); (b) follows by (47), which is valid because $t \in \mathcal{T}_k$ implies $t \in \mathcal{F}(\lambda_{hi}, \mathcal{T}_k^c)$; (c) follows because $c_l < 1$.

Next, take $t \in \mathcal{T}_k$ and observe, according to (151),

$$|\phi'_{+,k}(t)| \leq \left| \rho \frac{\phi'_{0,k}(t)}{\phi_{0,k}^2(t)} \right| + \left| \frac{\gamma}{\phi_{0,k}(t)} \right|. \tag{154}$$

The first term above can be upper-bounded following exactly the same steps that lead from (54) to (60). This gives:

$$\left| \rho \frac{\phi'_{0,k}(t)}{\phi_{0,k}^2(t)} \right| \leq r^{2r-1} c_{u7}^r \frac{\lambda_{hi}}{\lambda_{lo}^2}. \tag{155}$$

To upper-bound the second term in (154), consider two cases.

Case 1: $t \in \mathcal{F}(r\Delta\lambda_{lo}, \mathcal{T}_k^c)$. Then, by (44), $\phi_{0,k}(t) \geq c_{l1}^{r-1}$ and, therefore,

$$\left| \frac{\gamma}{\phi_{0,k}(t)} \right| \leq \frac{\gamma}{c_{l1}^{r-1}} = \frac{1}{c_{l1}^{r-1}} \frac{\lambda_{hi}^{2r-1}}{\lambda_{lo}^{2r}} \leq r^{2r-1} \left(\frac{1}{c_{l1}} \right)^r \frac{\lambda_{hi}}{\lambda_{lo}^2}. \tag{156}$$

Above, in the last (crude) inequality we used that $c_{l1} < 1$ and that $\lambda_{hi}/\lambda_{lo} < 1$.

Case 2: $t \in \mathcal{N}(r\Delta\lambda_{lo}, \mathcal{T}_k^c)$. In this case set $\{v_1, \dots, v_{\hat{r}}\} \triangleq \mathcal{T}_k^c \cap \mathcal{N}(r\Delta\lambda_{lo}, t)$ and note that $1 \leq \hat{r} \leq r - 1$. Hence, by (40),

$$\phi_{0,k}(t) \geq c_{l2}^{r-1} \frac{\prod_{j=1}^{\hat{r}} (v_j - t)^2}{(r\lambda_{lo})^{2\hat{r}}} \geq c_{l2}^{r-1} \frac{\lambda_{hi}^{2\hat{r}}}{(r\lambda_{lo})^{2\hat{r}}},$$

where we used that $|v_j - t| \geq 2\lambda_{hi}$ for all $j = 1, \dots, \hat{r}$ because all elements of \mathcal{T} are separated by at least $2\lambda_{hi}$. Therefore,

$$\left| \frac{\gamma}{\phi_{0,k}(t)} \right| \leq \frac{\gamma}{c_{l2}^{r-1}} \frac{(r\lambda_{lo})^{2\hat{r}}}{\lambda_{hi}^{2\hat{r}}} = \frac{1}{c_{l2}^{r-1}} \frac{\lambda_{hi}^{2r-1}}{\lambda_{lo}^{2r}} \frac{(r\lambda_{lo})^{2\hat{r}}}{\lambda_{hi}^{2\hat{r}}} \leq r^{2r-1} \left(\frac{1}{c_{l2}^2} \right)^r \frac{\lambda_{hi}}{\lambda_{lo}^2}. \tag{157}$$

Above, in the last (crude) inequality we used that $c_{l2} < 1$, $\lambda_{hi}/\lambda_{lo} < 1$, and $\hat{r} \leq r - 1$. Plugging (155), (156), and (157) into (154) we obtain

$$|\phi'_{+,k}(t)| \leq r^{2r-1} c_{u30}^r \frac{\lambda_{hi}}{\lambda_{lo}^2} \tag{158}$$

$$\leq r^{2r-1} c_{u30}^r \frac{1}{\lambda_{lo}}, \tag{159}$$

where we used that $\lambda_{hi}/\lambda_{lo} < 1$ and defined $c_{u30} \triangleq 2 \max(c_{u7}, 1/c_{l1}^2, 1/c_{l2}^2)$.

It follows from (153) and (159) that the function values and the derivatives of $q_{r\lambda_{lo}, \mathcal{T}_k}(t) = \phi_{+,k}(t)/(r^{2r} c_{u31}^r)$ with

$$c_{u31} \triangleq \max(c_{u30}, 1/c_l) \tag{160}$$

satisfy requirements (15) of Lemma 2 on \mathcal{T}_k . We conclude that $\phi_{+,k}(\cdot)$ can indeed be defined according to (149). According to Properties 3, 4, and 5 of Lemma 2, and (149), $\phi_{+,k}(\cdot)$ satisfies the following properties:

$$\|\phi_{+,k}(\cdot)\|_\infty \leq r^{2r} c_{u31}^r c_{u0}, \tag{161}$$

$$\|\phi'_{+,k}(\cdot)\|_\infty \leq r^{2r-1} c_{u31}^r c_{u1} \frac{1}{\lambda_{lo}}, \tag{162}$$

$$\|\phi''_{+,k}(\cdot)\|_\infty \leq r^{2r-2} c_{u31}^r c_{u2} \frac{1}{\lambda_{lo}^2}. \tag{163}$$

Proof of Property 2

Take $j \in \{1, \dots, S\}$ and consider $t_j \in \mathcal{T}$. There exists a unique $l \in \{1, \dots, r\}$ such that $t_j \in \mathcal{T}_l$. We will show that for all $\tau \in \mathcal{N}(\lambda_{hi}, t_j)$

$$\left| \phi_l(\tau) - \rho - \gamma s'_j(\tau - t_j) \right| \leq r^{2r+3} c_{u44}^{r+1} q_0(\tau) \tag{164}$$

and

$$|\phi_k(\tau)| \leq r^{2r+3} c_{u50}^{r+1} q_0(\tau), \text{ for } k \in \{1, \dots, r\}, k \neq l, \tag{165}$$

where the positive numerical constants c_{u44} and c_{u50} are defined below.

From this we will conclude:

$$\begin{aligned}
 |q_2(\tau) - \gamma s'_j(\tau - t_j)| &= \left| \sum_{k=1}^r \phi_k(\tau) - \rho - \gamma s'_j(\tau - t_j) \right| \\
 &\leq \sum_{k \neq l} |\phi_k(\tau)| + |\phi_l(\tau) - \rho - \gamma s'_j(\tau - t_j)| \\
 &\leq r^{2r+4} c_{u34}^{r+1} q_0(\tau),
 \end{aligned}$$

with $c_{u34} \triangleq 2 \max(c_{u44}, c_{u50})$, as desired.

To prove (164) and (165) recall, by (143) and (144),

$$\phi_l(\tau) - \rho - \gamma s'_j(\tau - t_j) \Big|_{\tau=t_j} = 0 = q_0(t_j), \tag{166}$$

$$\frac{d(\phi_l(\tau) - \rho - \gamma s'_j(\tau - t_j))}{d\tau} \Big|_{\tau=t_j} = \phi'_l(t_j) - \gamma s'_j = 0 = q'_0(t_j), \tag{167}$$

$$\phi_k(t_j) = 0 = q_0(t_j) \text{ for } k \in \{1, \dots, r\}, k \neq l, \tag{168}$$

$$\phi'_k(t_j) = 0 = q'_0(t_j) \text{ for } k \in \{1, \dots, r\}, k \neq l. \tag{169}$$

Hence, in order to prove the bounds in (165) and (164), we will derive upper bounds on the second derivatives $|\phi''_k(\tau)|$, $k \in \{1, \dots, r\}$, valid for all $\tau \in \mathcal{N}(\lambda_{hi}, t_j)$, and use the Mean Value theorem (see Theorem 3).

Taking the second derivative of (147) and applying the triangle inequality we find:

$$|\phi''_k(\tau)| \leq \underbrace{|\phi''_{0,k}(\tau)| |\phi_{+,k}(\tau)|}_{E_1(\tau)} + 2 \underbrace{|\phi'_{0,k}(\tau)| |\phi'_{+,k}(\tau)|}_{E_2(\tau)} + \underbrace{|\phi_{0,k}(\tau)| |\phi''_{+,k}(\tau)|}_{E_3(\tau)}. \tag{170}$$

In the derivation below we upper-bound the terms separately.

We will need the following notations. Set $\{v_1^\tau, \dots, v_{\tilde{r}}^\tau\} \triangleq \mathcal{N}(r\Delta\lambda_{lo}, \tau) \cap \mathcal{T}_k^c$. Also, set $\{v_1, \dots, v_{\tilde{r}}\} \triangleq \mathcal{N}(r\Delta\lambda_{lo} - \lambda_{hi}, t_j) \cap \mathcal{T}_k^c$. Note that the set $\{v_1, \dots, v_{\tilde{r}}\}$ does not depend on τ and also $\{v_1, \dots, v_{\tilde{r}}\} \subset \{v_1^\tau, \dots, v_{\tilde{r}}^\tau\}$ so that $\tilde{r} \leq \hat{r}$.

The remainder of the proof of Property 2 is organized as follows. First, consider the case $t_j \in \mathcal{T}_k$ and prove (164), next consider the case $t_j \in \mathcal{T}_k^c$ and prove (165).

Proof of (164): case $t_j \in \mathcal{T}_k$.

Bounding $E_1(\tau)$: By (195) [Mean Value theorem] and the triangle inequality we can write

$$|\phi_{+,k}(\tau)| \leq |\phi_{+,k}(t_j)| + |\phi'_{+,k}(t_j)| |\tau - t_j| + \frac{1}{2} |\phi''_{+,k}(\tau_m)| (\tau - t_j)^2$$

with $\tau_m \in (t_j, \tau)$. Next, we use (152) to upper-bound $|\phi_{+,k}(t_j)|$; use (158) to upper-bound $|\phi'_{+,k}(t_j)|$; use (163) to upper-bound $|\phi''_{+,k}(\tau_m)|$. With these estimates we can

further upper-bound $|\phi_{+,k}(\tau)|$ as follows:

$$\begin{aligned}
 |\phi_{+,k}(\tau)| &\leq r^{2r-1} c_{u35}^r \left(\frac{\lambda_{hi}^2}{\lambda_{lo}^2} + \frac{\lambda_{hi}}{\lambda_{lo}^2} |\tau - t_j| + \frac{1}{\lambda_{lo}^2} (\tau - t_j)^2 \right) \\
 &\leq r^{2r-1} c_{u36}^r \frac{\lambda_{hi}^2}{\lambda_{lo}^2}.
 \end{aligned}
 \tag{171}$$

Above, we defined $c_{u35} \triangleq \max(1/c_l, c_{u30}, c_{u31}c_{u2})$, $c_{u36} \triangleq 3c_{u35}$, and used $|\tau - t_j| \leq \lambda_{hi}$.

Assume $\tilde{r} \geq 1$ (the case $\tilde{r} = 0$ will be treated separately below) so that $\hat{r} \geq 1$ and $\tau \in \mathcal{N}(r\Delta\lambda_{lo}, \mathcal{T}_k^c)$, which implies that $|\phi''_{0,k}(\tau)|$ can be upper-bounded by (73).

Multiplying (171) and (73) and simplifying we obtain the following upper bound on E_1 :

$$\begin{aligned}
 E_1(\tau) = |\phi''_{0,k}(\tau)| |\phi_{+,k}(\tau)| &\stackrel{(a)}{\leq} r^{2r+1} c_{u37}^{r+1} \frac{\prod_{1 \leq l \leq \hat{r}} (v_l^\tau - \tau)^2}{(r\lambda_{lo})^{2\hat{r}}} \frac{1}{\lambda_{lo}^2} \\
 &\stackrel{(b)}{\leq} r^{2r+1} c_{u37}^{r+1} \frac{\prod_{1 \leq l \leq \hat{r}} (v_l - \tau)^2}{(r\lambda_{lo})^{2\hat{r}}} \frac{1}{\lambda_{lo}^2}.
 \end{aligned}
 \tag{172}$$

Above, in (a) we used (multiple times) the bound $\lambda_{hi} \leq |v_l^\tau - \tau|$, which is true for all $l \in \{1, \dots, \hat{r}\}$ (follows because the elements of \mathcal{T} are separated by at least $2\lambda_{hi}$), used $\lambda_{hi}/\lambda_{lo} < 1$, and defined $c_{u37} \triangleq \max(6c_{u3}c_{u36}, c_{u31}c_{u0}c_{u5})$; in (b) we used the fact that $|v_l^\tau - \tau|/(\lambda_{lo}r) \leq \Delta < 1$ for all $l \in \{1, \dots, \hat{r}\}$.

For the case $\tilde{r} = 0$, the upper bound (74) also holds by (46) and (161).

Bounding $E_2(\tau)$: By (194) [Mean Value theorem] we can write

$$|\phi'_{+,k}(\tau)| \leq |\phi'_{+,k}(t_j)| + |\phi''_{+,k}(\tau_m)| |\tau - t_j|$$

with $\tau_m \in (t_j, \tau)$. Next, we use (158) to upper-bound $|\phi'_{+,k}(t_j)|$; use (163) to upper-bound $|\phi''_{+,k}(\tau_m)|$. With these estimates we can further upper-bound $|\phi'_{+,k}(\tau)|$ as follows:

$$|\phi'_{+,k}(\tau)| \leq r^{2r-1} c_{u38}^r \left(\frac{\lambda_{hi}}{\lambda_{lo}^2} + \frac{1}{\lambda_{lo}^2} |\tau - t_j| \right) \leq r^{2r-1} c_{u39}^r \frac{\lambda_{hi}}{\lambda_{lo}^2}. \tag{173}$$

Above, we defined $c_{u38} \triangleq \max(c_{u30}, c_{u2}c_{u31})$, $c_{u39} \triangleq 2c_{u38}$, and used $|\tau - t_j| \leq \lambda_{hi}$.

Assume $\tilde{r} \geq 1$ (the case $\tilde{r} = 0$ will be treated separately below) so that $\hat{r} \geq 1$ and $\tau \in \mathcal{N}(r\Delta\lambda_{lo}, \mathcal{T}_k^c)$, which implies that $|\phi'_{0,k}(t)|$ can be upper-bounded by (76).

Multiplying (173) and (76) and simplifying, we obtain the following upper bound on E_2 :

$$\begin{aligned}
 E_2(\tau) &= |\phi'_{0,k}(\tau)| |\phi'_{+,k}(\tau)| \stackrel{(a)}{\leq} r^{2r} c_{u40}^r \frac{\prod_{1 \leq l \leq \hat{r}} (v_l^\tau - \tau)^2}{(r\lambda_{10})^{2\hat{r}}} \frac{1}{\lambda_{10}^2} \\
 &\stackrel{(b)}{\leq} r^{2r} c_{u40}^r \frac{\prod_{1 \leq l \leq \hat{r}} (v_l - \tau)^2}{(r\lambda_{10})^{2\hat{r}}} \frac{1}{\lambda_{10}^2}. \tag{174}
 \end{aligned}$$

Above, in (a) we used the bound $\lambda_{hi} \leq |v_l^\tau - \tau|$, which is true for all $l \in \{1, \dots, \hat{r}\}$ (follows because the elements of \mathcal{T} are separated by at least $2\lambda_{hi}$), used $\lambda_{hi}/\lambda_{10} < 1$, and defined $c_{u40} \triangleq \max(2c_{u39}c_{u3}, 2\pi c_{u31}c_{u1})$; in (b) we use the fact that $|v_l^\tau - \tau|/(\lambda_{10}r) \leq \Delta < 1$ for all $l \in \{1, \dots, \hat{r}\}$.

For the case $\tilde{r} = 0$, the upper bound (174) also holds by (45) and (162).

Bounding $E_3(\tau)$: By (163),

$$|\phi''_{+,k}(\tau)| \leq r^{2r-2} c_{u31}^r c_{u2} \frac{1}{\lambda_{10}^2}. \tag{175}$$

Assume $\tilde{r} \geq 1$ (the case $\tilde{r} = 0$ will be treated separately below) so that $\hat{r} \geq 1$ and $\tau \in \mathcal{N}(r\Delta\lambda_{10}, \mathcal{T}_k^c)$, which implies that $|\phi_{0,k}(\tau)|$ can be upper-bounded by (79). Multiplying (175) and (79) and simplifying, we obtain the following upper bound on E_3 :

$$\begin{aligned}
 E_3(\tau) &= |\phi_{0,k}(\tau)| |\phi''_{+,k}(\tau)| \stackrel{(a)}{\leq} r^{2r-2} c_{u41}^r \frac{\prod_{l=1}^{\hat{r}} (v_l^\tau - \tau)^2}{(r\lambda_{10})^{2\hat{r}}} \frac{1}{\lambda_{10}^2} \\
 &\stackrel{(b)}{\leq} r^{2r-2} c_{u41}^r \frac{\prod_{1 \leq l \leq \tilde{r}} (v_l - \tau)^2}{(r\lambda_{10})^{2\tilde{r}}} \frac{1}{\lambda_{10}^2}. \tag{176}
 \end{aligned}$$

Above, (a) we defined $c_{u41} \triangleq c_{u31}c_{u2}c_u$; in (b) we use the fact that $|v_l^\tau - \tau|/(\lambda_{10}r) \leq \Delta < 1$ for all $l \in \{1, \dots, \hat{r}\}$.

For the case $\tilde{r} = 0$, the upper bound (176) also holds by (175) because by (148) and Lemma 1, Property 3, $|\phi_{0,k}(\tau)| < 1$ and because $c_u > 1$ and $c_{u2} > 1$.

From (170), (172), (174), and (176) we conclude that

$$|\phi''_k(\tau)| \leq r^{2r+1} c_{u42}^{r+1} \frac{\prod_{1 \leq l \leq \tilde{r}} (v_l - \tau)^2}{(r\lambda_{10})^{2\tilde{r}}} \frac{1}{\lambda_{10}^2}, \tag{177}$$

where we defined $c_{u42} \triangleq 4 \max(c_{u37}, c_{u40}, c_{u41})$.

Putting pieces together: Applying (195) [Mean Value theorem] to the function $f(\cdot) = \phi_l(\cdot) - \rho - \gamma s'_j(\cdot - t_j)$ with $a = t_j$ and $b = \tau$ and using (166), (167), (177) and we can write for all $\tau \in \mathcal{N}(\lambda_{hi}, t_j)$:

$$\begin{aligned}
 \left| \phi_l(\tau) - \rho - \gamma s'_j(\tau - t_j) \right| &\stackrel{(a)}{\leq} \frac{1}{2} r^{2r+1} c_{u42}^{r+1} \frac{\prod_{1 \leq l \leq \tilde{r}} (v_l - \tau_m)^2}{(r\lambda_{10})^{2\tilde{r}}} \frac{(\tau - t_j)^2}{\lambda_{10}^2} \\
 &\stackrel{(b)}{\leq} \frac{1}{2} r^{2r+1} c_{u42}^{r+1} 2^{\tilde{r}} \frac{\prod_{1 \leq l \leq \tilde{r}} (v_l - \tau)^2}{(r\lambda_{10})^{2\tilde{r}}} \frac{(\tau - t_j)^2}{\lambda_{10}^2}
 \end{aligned}$$

$$\stackrel{(c)}{\leq} r^{2r+3} c_{u43}^{r+1} \frac{\prod_{1 \leq l \leq \tilde{r}} (v_l - \tau)^2}{(r\lambda_{10})^{2\tilde{r}}} \frac{(\tau - t_j)^2}{(r\lambda_{10})^2}. \tag{178}$$

Above, in (a) $\tau_m \in (t_j, \tau)$; in (b) we used that $|v_l - \tau_m| < |v_l - \tau| + \lambda_{hi} < 2|v_l - \tau|$, which is true because $\tau \in \mathcal{N}(\lambda_{hi}, t_j)$ and because the elements of \mathcal{T} are separated by at least $2\lambda_{hi}$; in (c) we defined $c_{u43} \triangleq 2c_{u42}$.

The bound (164) follows from (178) and (83) by defining $c_{u44} \triangleq c_{u43}/c_{l3}$.

Proof of (165): case $t_j \in \mathcal{T}_k^c$. We only need to consider this case when $r > 1$. Indeed, when $r = 1$, the sum in (141) only contains one element, $\phi_l(\cdot)$, and, necessarily, $t_j \in \mathcal{T}_l$ because \mathcal{T}_l^c is empty.

In this case t_j is one of the elements among $\{v_1, \dots, v_{\tilde{r}}\} \subset \{v_1^{\tilde{r}}, \dots, v_{\tilde{r}}^{\tilde{r}}\}$; in other words, $t_j = v_{\tilde{m}} = v_{\tilde{m}}^{\tilde{r}}$ for some $1 \leq \tilde{m} \leq \tilde{r}, 1 \leq \hat{m} \leq \hat{r}$. The set $\mathcal{T}_k \cap \mathcal{N}(r\Delta\lambda_{10} - \lambda_{hi}, t_j)$ is either empty or contains exactly one element. Let $b \triangleq |\mathcal{T}_k \cap \mathcal{N}(r\Delta\lambda_{10} - \lambda_{hi}, t_j)|$.

In the case when $b = 1$, let $\{\tilde{t}\} \triangleq \mathcal{T}_k \cap \mathcal{N}(r\Delta\lambda_{10} - \lambda_{hi}, t_j)$.

Bounding $E_1(\tau)$: Consider the case $b = 1$. By (195) [Mean Value theorem] we can write:

$$|\phi_{+,k}(\tau)| \leq |\phi_{+,k}(\tilde{t})| + |\phi'_{+,k}(\tilde{t})| |\tau - \tilde{t}| + \frac{1}{2} |\phi''_{+,k}(\tau_m)| (\tau - \tilde{t})^2$$

with $\tau_m \in (\tilde{t}, \tau)$. Next, we use (152) to upper-bound $|\phi_{+,k}(\tilde{t})|$; use (158) to upper-bound $|\phi'_{+,k}(\tilde{t})|$; use (163) to upper-bound $|\phi''_{+,k}(\tau_m)|$. With these estimates we can further upper-bound $|\phi_{+,k}(\tau)|$ as follows:

$$\begin{aligned} |\phi_{+,k}(\tau)| &\leq r^{2r-1} c_{u35}^r \left(\frac{\lambda_{hi}^2}{\lambda_{10}^2} + \frac{\lambda_{hi}}{\lambda_{10}^2} |\tau - \tilde{t}| + \frac{1}{\lambda_{10}^2} (\tau - \tilde{t})^2 \right) \\ &\leq r^{2r+1} c_{u36}^r \frac{(\tilde{t} - \tau)^2}{(r\lambda_{10})^2} = r^{2r+1} c_{u36}^r \left[\frac{(\tilde{t} - \tau)^2}{(r\lambda_{10})^2} \right]^{I[b=1]}, \end{aligned} \tag{179}$$

where we used that $\lambda_{hi} \leq |\tilde{t} - \tau|$ because the elements of \mathcal{T} are separated by at least $2\lambda_{hi}$ and $\tau \in \mathcal{N}(\lambda_{hi}, t_j)$ with $\tilde{t} \neq t_j$. According to (161) the upper bound (179) also holds for $b = 0$.

Since $t_j \in \mathcal{T}_k^c$ and $\tau \in \mathcal{N}(\lambda_{hi}, t_j)$, it follows $\tau \in \mathcal{N}(r\Delta\lambda_{10}, \mathcal{T}_k^c)$ so that $\hat{r} \geq 1$, which implies that $|\phi''_{0,k}(\tau)|$ can be upper-bounded by (85).

Multiplying (179) and (85) and simplifying, we obtain the following upper bound on E_1 :

$$\begin{aligned} E_1(\tau) &= |\phi''_{0,k}(\tau)| |\phi_{+,k}(\tau)| \\ &\stackrel{(a)}{\leq} r^{2r+1} c_{u45}^{r+1} \left[\frac{(\tilde{t} - \tau)^2}{(r\lambda_{10})^2} \right]^{I[b=1]} \frac{\prod_{1 \leq l \leq \hat{r}, l \neq \hat{m}} (v_l^{\tilde{r}} - \tau)^2}{(r\lambda_{10})^{2(\hat{r}-1)}} \frac{1}{\lambda_{10}^2} \\ &\stackrel{(b)}{\leq} r^{2r+1} c_{u45}^{r+1} \left[\frac{(\tilde{t} - \tau)^2}{(r\lambda_{10})^2} \right]^{I[b=1]} \frac{\prod_{1 \leq l \leq \tilde{r}, l \neq \tilde{m}} (v_l - \tau)^2}{(r\lambda_{10})^{2(\tilde{r}-1)}} \frac{1}{\lambda_{10}^2}. \end{aligned} \tag{180}$$

Above, in (a) we defined $c_{u45} \triangleq c_{u36}c_{u17}$; in (b) we used that $|v_l^\tau - \tau|/(\lambda_{10}r) \leq \Delta < 1$ for all $l \in \{1, \dots, \hat{r}\}$.

Bounding $E_2(\tau)$: Consider the case $b = 1$. By (194) [Mean Value theorem] we can write:

$$|\phi'_{+,k}(\tau)| \leq |\phi'_{+,k}(\tilde{t})| + |\phi''_{+,k}(\tau_m)| |\tau - \tilde{t}|$$

with $\tau_m \in (\tilde{t}, \tau)$. Next, we use (158) to upper-bound $|\phi'_{+,k}(\tilde{t})|$; use (163) to upper-bound $|\phi''_{+,k}(\tau_m)|$. With these estimates we can further upper-bound $|\phi'_{+,k}(\tau)|$ as follows:

$$\begin{aligned} |\phi'_{+,k}(\tau)| &\leq r^{2r-1} c_{u38}^r \left(\frac{\lambda_{hi}}{\lambda_{10}^2} + \frac{1}{\lambda_{10}^2} |\tau - \tilde{t}| \right) \\ &\leq r^{2r} c_{u39}^r \frac{|\tau - \tilde{t}|}{r\lambda_{10}} \frac{1}{\lambda_{10}} = r^{2r} c_{u39}^r \left[\frac{|\tau - \tilde{t}|}{r\lambda_{10}} \right]^{I[b=1]} \frac{1}{\lambda_{10}}, \end{aligned} \tag{181}$$

where we used that $\lambda_{hi} \leq |\tilde{t} - \tau|$. According to (162) the upper bound (181) also holds for $b = 0$.

Since $t_j \in \mathcal{T}_k^c$ and $\tau \in \mathcal{N}(\lambda_{hi}, t_j)$, it follows $\tau \in \mathcal{N}(r\Delta\lambda_{10}, \mathcal{T}_k^c)$ so that $\hat{r} \geq 1$, which implies that $|\phi'_{0,k}(\tau)|$ can be upper-bounded by (88).

Multiplying (181) and (88) and simplifying we obtain the following upper bound on E_2 :

$$\begin{aligned} E_2(\tau) &= |\phi'_{0,k}(\tau)| |\phi'_{+,k}(\tau)| \\ &\stackrel{(a)}{\leq} r^{2r} c_{u46}^r \left[\frac{(\tilde{t} - \tau)^2}{(r\lambda_{10})^2} \right]^{I[b=1]} \frac{\prod_{1 \leq l \leq \hat{r}, l \neq \hat{m}} (v_l^\tau - \tau)^2}{(r\lambda_{10})^{2(\hat{r}-1)}} \frac{1}{\lambda_{10}^2} \\ &\stackrel{(b)}{\leq} r^{2r} c_{u46}^r \left[\frac{(\tilde{t} - \tau)^2}{(r\lambda_{10})^2} \right]^{I[b=1]} \frac{\prod_{1 \leq l \leq \hat{r}, l \neq \hat{m}} (v_l - \tau)^2}{(r\lambda_{10})^{2(\hat{r}-1)}} \frac{1}{\lambda_{10}^2}. \end{aligned} \tag{182}$$

Above, in (a) we defined $c_{u46} \triangleq c_{u39}c_{u19}$; in (b) we used that $|v_l^\tau - \tau|/(\lambda_{10}r) \leq \Delta < 1$ for all $l \in \{1, \dots, \hat{r}\}$.

Bounding $E_3(\tau)$: By (163),

$$|\phi''_{+,k}(\tau)| \leq r^{2r-2} c_{u31}^r c_{u2} \frac{1}{\lambda_{10}^2}. \tag{183}$$

Since $t_j \in \mathcal{T}_k^c$ and $\tau \in \mathcal{N}(\lambda_{hi}, t_j)$, it follows $\tau \in \mathcal{N}(r\Delta\lambda_{10}, \mathcal{T}_k^c)$ so that $\hat{r} \geq 1$, which implies that $|\phi_{0,k}(\tau)|$ can be upper-bounded as in (91). Multiplying (183) and (91) and simplifying, we obtain the following upper bound on E_3 :

$$E_3(\tau) = |\phi_{0,k}(\tau)| |\phi''_{+,k}(\tau)|$$

$$\begin{aligned}
 &\stackrel{(a)}{\leq} r^{2r-2} c_{u47}^r \left[\frac{(\tilde{t} - \tau)^2}{(r\lambda_{10})^2} \right]^{I[b=1]} \frac{\prod_{1 \leq l \leq \tilde{r}, l \neq \tilde{m}} (v_l^\tau - \tau)^2}{(r\lambda_{10})^{2(\tilde{r}-1)}} \frac{1}{\lambda_{10}^2} \\
 &\stackrel{(b)}{\leq} r^{2r-2} c_{u47}^r \left[\frac{(\tilde{t} - \tau)^2}{(r\lambda_{10})^2} \right]^{I[b=1]} \frac{\prod_{1 \leq l \leq \tilde{r}, l \neq \tilde{m}} (v_l - \tau)^2}{(r\lambda_{10})^{2(\tilde{r}-1)}} \frac{1}{\lambda_{10}^2}. \tag{184}
 \end{aligned}$$

Above, in (a) we defined $c_{u47} \triangleq c_{u31}c_{u2}c_u$; in (b) we used the fact that $|v_l^\tau - \tau|/(\lambda_{10}r) \leq \Delta < 1$ for all $l \in \{1, \dots, \tilde{r}\}$.

From (170), (180), (182), and (184) we conclude that

$$|\phi_k''(\tau)| \leq r^{2r+1} c_{u48}^{r+1} \left[\frac{(\tilde{t} - \tau)^2}{(r\lambda_{10})^2} \right]^{I[b=1]} \frac{\prod_{1 \leq l \leq \tilde{r}, l \neq \tilde{m}} (v_l - \tau)^2}{(r\lambda_{10})^{2(\tilde{r}-1)}} \frac{1}{\lambda_{10}^2}, \tag{185}$$

where we defined $c_{u48} \triangleq 4 \max(c_{u45}, c_{u46}, c_{u47})$.

Putting pieces together: Applying (195) [Mean Value theorem] to the function $f(\cdot) = \phi_l(\cdot)$ with $a = t_j$ and $b = \tau$ and using (168), (169), (185), and we can write for all $\tau \in \mathcal{N}(\lambda_{hi}, t_j)$:

$$\begin{aligned}
 |\phi_k(\tau)| &\stackrel{(a)}{\leq} \frac{1}{2} r^{2r+1} c_{u48}^{r+1} \left[\frac{(\tilde{t} - \tau_m)^2}{(r\lambda_{10})^2} \right]^{I[b=1]} \frac{\prod_{1 \leq l \leq \tilde{r}, l \neq \tilde{m}} (v_l - \tau_m)^2}{(r\lambda_{10})^{2(\tilde{r}-1)}} \frac{(\tau - t_j)^2}{\lambda_{10}^2} \\
 &\stackrel{(b)}{\leq} \frac{1}{2} r^{2r+1} c_{u48}^{r+1} 2^{\tilde{r}} \left[\frac{(\tilde{t} - \tau)^2}{(r\lambda_{10})^2} \right]^{I[b=1]} \frac{\prod_{1 \leq l \leq \tilde{r}, l \neq \tilde{m}} (v_l - \tau)^2}{(r\lambda_{10})^{2(\tilde{r}-1)}} \frac{(\tau - t_j)^2}{\lambda_{10}^2} \\
 &\stackrel{(c)}{\leq} r^{2r+3} c_{u49}^{r+1} \left[\frac{(\tilde{t} - \tau)^2}{(r\lambda_{10})^2} \right]^{I[b=1]} \frac{\prod_{1 \leq l \leq \tilde{r}} (v_l - \tau)^2}{(r\lambda_{10})^{2\tilde{r}}}. \tag{186}
 \end{aligned}$$

Above, in (a) $\tau_m \in (t_j, \tau)$; in (b) we used the fact that, for $l \neq \tilde{m}$, $|v_l - \tau_m| < |v_l - \tau| + \lambda_{hi} < 2|v_l - \tau|$ and $|\tilde{t} - \tau_m| < |\tilde{t} - \tau| + \lambda_{hi} < 2|\tilde{t} - \tau|$, which is true because $\tau \in \mathcal{N}(\lambda_{hi}, t_j)$ and because the elements of \mathcal{T} are separated by at least $2\lambda_{hi}$; in (c) we defined $c_{u49} \triangleq 2c_{u48}$ and used the fact that $t_j = v_{\tilde{m}}$.

The bound (165) follows from (186) and (95) by defining $c_{u50} \triangleq c_{u49}/c_{l3}$.

Proof of Property 3

By (141) and the triangle inequality:

$$\begin{aligned}
 \|q_2(\cdot)\|_\infty &\leq \rho + r \max_{1 \leq k \leq r} \|\phi_k(\cdot)\|_\infty \\
 &\stackrel{(a)}{\leq} \rho + r \max_{1 \leq k \leq r} \|\phi_{+,k}(\cdot)\|_\infty \\
 &\stackrel{(b)}{=} \rho + r^{2r+1} c_{u31}^r \max_{1 \leq k \leq r} \|q_{r\lambda_{10}, \mathcal{T}_k, \{f_j\}, \{d_j\}}(\cdot)\|_\infty \\
 &\stackrel{(c)}{\leq} \rho + r^{2r+1} c_{u31}^r c_{u0} \stackrel{(d)}{\leq} r^{2r+1} c_{u56}^r.
 \end{aligned}$$

Above, in (a) we used (147) and the fact that by (148) and Lemma 1, Property 3, $\|\phi_{0,k}(\cdot)\|_\infty \leq 1$; in (b) we used (149); in (c) we used Lemma 2, Property 3; in (d) we defined $c_{u56} \triangleq 2c_{u0}c_{u31}$ and used the fact that $\rho < 1 < c_{u0}c_{u31}$.

Proof of Property 4

Take $\tau \in \mathcal{F}(\lambda_{hi}, T)$. As above, let $\{u_1^\tau, \dots, u_{\check{r}}^\tau\} \triangleq \mathcal{N}(r\Delta\lambda_{l0}, \tau) \cap \mathcal{T}$. Then by (16),

$$q_0(\tau) \geq c_{l2}^{\check{r}} \frac{\prod_{l=1}^{\check{r}} (u_l^\tau - \tau)^2}{(r\lambda_{l0})^{2\check{r}}}. \tag{187}$$

By (19) this bound is also valid when $\check{r} = 0$.

Fix k . If $\tau \in \mathcal{N}(r\Delta\lambda_{l0}, \mathcal{T}_k^c)$, then we can use (41) to upper-bound $|\phi_{0,k}(\tau)|$:

$$|\phi_{0,k}(\tau)| \leq c_u^{\hat{r}} \frac{\prod_{l=1}^{\hat{r}} (v_l^\tau - \tau)^2}{(r\lambda_{l0})^{2\hat{r}}}, \tag{188}$$

where, as before, $\{v_1^\tau, \dots, v_{\hat{r}}^\tau\} \triangleq \mathcal{N}(r\Delta\lambda_{l0}, \tau) \cap \mathcal{T}_k^c$. If $\tau \notin \mathcal{N}(r\Delta\lambda_{l0}, \mathcal{T}_k^c)$, we will use that by (36) and by Lemma 1, Property 3,

$$|\phi_{0,k}(\tau)| \leq 1. \tag{189}$$

The set $\mathcal{T}_k \cap \mathcal{N}(r\Delta\lambda_{l0}, \tau)$ is either empty or contains exactly one element. Let $b \triangleq |\mathcal{T}_k \cap \mathcal{N}(r\Delta\lambda_{l0}, \tau)|$ denote the size of this set; when $b = 1$, let $\{\tilde{\tau}\} \triangleq \mathcal{T}_k \cap \mathcal{N}(r\Delta\lambda_{l0}, \tau)$. Following the steps that lead to (179), we obtain:

$$|\phi_{+,k}(\tau)| \leq r^{2r+1} c_{u36}^r \left[\frac{(\tilde{\tau} - \tau)^2}{(r\lambda_{l0})^2} \right]^{I[b=1]} \tag{190}$$

and the bound is valid for both cases $b = 0$ and $b = 1$.

Case $\hat{r} \geq 1$: Then, $\{u_1^\tau, \dots, u_{\check{r}}^\tau\} = \{v_1^\tau, \dots, v_{\hat{r}}^\tau\} \cup \{\tilde{\tau}\}$ if $b = 1$, and $\{u_1^\tau, \dots, u_{\check{r}}^\tau\} = \{v_1^\tau, \dots, v_{\hat{r}}^\tau\}$ if $b = 0$. Therefore,

$$\begin{aligned} |\phi_k(\tau)| &= |\phi_{0,k}(\tau)| |\phi_{+,k}(\tau)| \stackrel{(a)}{\leq} r^{2r+1} c_{u36}^r c_u^{\hat{r}} \left[\frac{(\tilde{\tau} - \tau)^2}{(r\lambda_{l0})^2} \right]^{I[b=1]} \frac{\prod_{l=1}^{\hat{r}} (v_l^\tau - \tau)^2}{(r\lambda_{l0})^{2\hat{r}}} \\ &= r^{2r+1} c_{u36}^r c_u^{\hat{r}} \frac{\prod_{l=1}^{\check{r}} (u_l^\tau - \tau)^2}{(r\lambda_{l0})^{2\check{r}}} \stackrel{(b)}{\leq} r^{2r+1} c_{u51}^r q_0(\tau). \end{aligned} \tag{191}$$

Above, (a) follows by (188) and (190); (b) follows by (187) with $c_{u51} \triangleq c_{u36}c_u/c_{l2}$.

Case $\hat{r} = 0$: Then, $\check{r} = 1$ and $\{u_{\check{r}}^\tau\} = \{\tilde{\tau}\}$ if $b = 1$ and $\check{r} = 0$ if $b = 0$. Therefore,

$$|\phi_k(\tau)| = |\phi_{0,k}(\tau)| |\phi_{+,k}(\tau)| \stackrel{(a)}{\leq} r^{2r+1} c_{u36}^r \left[\frac{(\tilde{\tau} - \tau)^2}{(r\lambda_{l0})^2} \right]^{I[b=1]}$$

$$= r^{2r+1} c_{u36}^r \frac{\prod_{l=1}^{\tilde{r}} (u_l^{\tilde{r}} - \tau)^2}{(r\lambda_{l0})^{2\tilde{r}}} \stackrel{(b)}{\leq} r^{2r+1} c_{u51}^r q_0(\tau). \tag{192}$$

Above, (a) follows by (189) and (190); (b) follows by (187) because $c_u > 1$.

By Lemma 3, Property 6,

$$\rho = \frac{\lambda_{hi}^{2r}}{\lambda_{lo}^{2r}} \leq r^{2r} \frac{1}{c_l^r} q_0(\tau). \tag{193}$$

Therefore, by (141), (191), (192), (193),

$$|q_2(\tau)| \leq \sum_{k=1}^r |\phi_k(\tau)| + \rho \leq r^{2r+2} c_{u52}^r q_0(\tau),$$

where we defined $c_{u52} \triangleq c_{u51} + 1/c_l$. □

Mean Value Theorem

We repeatedly use the Taylor series approximation with the remained expressed via the Mean Value theorem [1, p. 880, 25.2.25] given below for the convenience of the reader.

Theorem 3 *Assume that $f(t)$ is twice differentiable on the interval $[a, b]$. Then, there exists $t_1 \in (a, b)$ such that*

$$f(b) = f(a) + f'(t_1)(b - a). \tag{194}$$

and there exists $t_2 \in (a, b)$ such that

$$f(b) = f(a) + f'(a)(b - a) + \frac{f''(t_2)}{2}(b - a)^2. \tag{195}$$

Properties of Fejér Kernel

The results proven in subsections below are analogous to the results in [12, eq. (1.11) and eq. (2.6)] with the difference that here we need bounds on sums and in [12] bounds on the corresponding integrals are provided.

Below, we will need uniform upper bounds on $|k_{hi}(\cdot)|$, $|k'_{hi}(\cdot)|$, and $|k''_{hi}(\cdot)|$; these are derived next.

Fejér kernel (7) can be written as a Fourier sum as follows:

$$k_{hi}(t) = \frac{1}{N} \sum_{|k| \leq f_{hi}} \left(1 - \frac{|k|}{f_{hi} + 1} \right) e^{i2\pi tk}. \tag{196}$$

Taking the absolute value of both sides in (196) and applying the triangle inequality we find:

$$|k_{\text{hi}}(t)| \leq \frac{1}{N} \sum_{|k| \leq f_{\text{hi}}} \left(1 - \frac{|k|}{f_{\text{hi}} + 1} \right) = \frac{1}{N} (1 + f_{\text{hi}}).$$

Above, the equality follows by summing up the simple series.

Differentiating (196) we obtain:

$$k'_{\text{hi}}(t) = \frac{i2\pi}{N} \sum_{|k| \leq f_{\text{hi}}} \left(1 - \frac{|k|}{f_{\text{hi}} + 1} \right) k e^{i2\pi tk}. \quad (197)$$

Taking the absolute value of both sides in (197) and applying the triangle inequality we find:

$$|k'_{\text{hi}}(t)| \leq \frac{2\pi}{N} \sum_{|k| \leq f_{\text{hi}}} |k| \left(1 - \frac{|k|}{f_{\text{hi}} + 1} \right) \stackrel{(a)}{=} \frac{2\pi}{3N} f_{\text{hi}} (2 + f_{\text{hi}}) \stackrel{(b)}{\leq} \frac{2\pi}{N} f_{\text{hi}}^2. \quad (198)$$

Above, (a) follows by summing up the simple power series; (b) follows because $f_{\text{hi}} > 1$.

Differentiating (197) we obtain:

$$k''_{\text{hi}}(t) = -\frac{(2\pi)^2}{N} \sum_{|k| \leq f_{\text{hi}}} \left(1 - \frac{|k|}{f_{\text{hi}} + 1} \right) k^2 e^{i2\pi tk}. \quad (199)$$

Taking the absolute value of both sides in (199) and applying the triangle inequality we find:

$$|k''_{\text{hi}}(t)| \leq \frac{(2\pi)^2}{N} \sum_{|k| \leq f_{\text{hi}}} k^2 \left(1 - \frac{|k|}{f_{\text{hi}} + 1} \right) \stackrel{(a)}{=} \frac{2\pi^2}{3N} f_{\text{hi}} (2 + 3f_{\text{hi}} + f_{\text{hi}}^2) \stackrel{(b)}{\leq} \frac{4\pi^2}{N} f_{\text{hi}}^3. \quad (200)$$

Above, (a) follows by summing up the simple power series; (b) follows because $f_{\text{hi}} > 1$.

The bounds derived below in this section are crude in the sense that no attempt has been made to obtain the tightest possible constants; for this reason some of the steps below may appear unnecessarily wasteful.

Proof of (106)

Begin by splitting the summation interval and recombining the terms in the following way:

$$\begin{aligned} \sum_{n=0}^{N-1} \left| k'_{hi} \left(\frac{n}{N} \right) \right| &= \sum_{\frac{n}{N} \in [0, \lambda_{hi})} \left| k'_{hi} \left(\frac{n}{N} \right) \right| + \sum_{\frac{n}{N} \in [\lambda_{hi}, \frac{1}{2})} \left| k'_{hi} \left(\frac{n}{N} \right) \right| \\ &\quad + \sum_{\frac{n}{N} \in [\frac{1}{2}, 1 - \lambda_{hi})} \left| k'_{hi} \left(\frac{n}{N} \right) \right| + \sum_{\frac{n}{N} \in [1 - \lambda_{hi}, 1)} \left| k'_{hi} \left(\frac{n}{N} \right) \right| \\ &\leq 2 \sum_{\frac{n}{N} \in [0, \lambda_{hi} + \frac{1}{N})} \left| k'_{hi} \left(\frac{n}{N} \right) \right| + 2 \sum_{\frac{n}{N} \in [\lambda_{hi}, \frac{1}{2} + \frac{1}{N})} \left| k'_{hi} \left(\frac{n}{N} \right) \right|. \end{aligned} \tag{201}$$

Above, the inequality follows by symmetry of $k'_{hi}(\cdot)$ around $1/2$. Next, we upper-bound the two sums separately.

To upper-bound the first sum in (201) we proceed as follows:

$$\sum_{\frac{n}{N} \in [0, \lambda_{hi} + \frac{1}{N})} \left| k'_{hi} \left(\frac{n}{N} \right) \right| \leq (\lambda_{hi} + 1/N) N \max_t |k'_{hi}(t)| \stackrel{(a)}{\leq} 2\lambda_{hi} 2\pi f_{hi}^2 \stackrel{(b)}{=} 4\pi/\lambda_{hi}. \tag{202}$$

Above, in (a) we used (198) and the assumption that $1/N \leq \lambda_{hi}$; in (b) we used that $f_{hi} = 1/\lambda_{hi}$.

To upper-bound the second term in (201) we observe that $|k'_{hi}(\cdot)|$ can be upper-bounded as follows for $t \in [0, 0.55]$:

$$\begin{aligned} |k'_{hi}(t)| &\stackrel{(a)}{=} \frac{2\pi \sin(\pi(f_{hi} + 1)t)}{N} \left| \frac{\cos(\pi(f_{hi} + 1)t)}{\sin^2(\pi t)} - \frac{\cos(\pi t) \sin(\pi(f_{hi} + 1)t)}{(f_{hi} + 1) \sin^3(\pi t)} \right| \\ &\stackrel{(b)}{\leq} \frac{1}{N} \left(\frac{2\pi}{\sin^2(\pi t)} + \frac{2\pi}{(f_{hi} + 1) \sin^3(\pi t)} \right) \stackrel{(c)}{\leq} \frac{1}{N} \left(\frac{2}{(f_{hi} + 1)t^3} + \frac{2}{t^2} \right). \end{aligned} \tag{203}$$

Above, (a) follows by differentiating $k_{hi}(\cdot)$ in (7); in (b) we used the triangle inequality and the fact that $|\sin(\cdot)| \leq 1$, $|\cos(\cdot)| \leq 1$; in (c) we used the inequalities $\sin(\pi t)^2 \geq \pi t^2$ and $\sin(\pi t)^3 \geq \pi t^3$ for $t \in [0, 0.55]$. Therefore,

$$\begin{aligned} \sum_{\frac{n}{N} \in [\lambda_{hi}, \frac{1}{2} + \frac{1}{N})} \left| k'_{hi} \left(\frac{n}{N} \right) \right| &\stackrel{(a)}{\leq} \frac{2}{1 + f_{hi}} \frac{1}{N} \sum_{\frac{n}{N} \in [\lambda_{hi}, \frac{1}{2} + \frac{1}{N})} \frac{1}{(n/N)^3} + \frac{2}{N} \sum_{\frac{n}{N} \in [\lambda_{hi}, \frac{1}{2} + \frac{1}{N})} \frac{1}{(n/N)^2} \end{aligned}$$

$$\begin{aligned}
 &\stackrel{(b)}{\leq} \frac{2}{1 + f_{hi}} \left(\frac{1}{N} \frac{1}{\lambda_{hi}^3} + \int_{\lambda_{hi}}^{1/2+1/N} \frac{1}{t^3} dt \right) + \frac{2}{N} \frac{1}{\lambda_{hi}^2} + 2 \int_{\lambda_{hi}}^{1/2+1/N} \frac{1}{t^2} dt \\
 &\leq \frac{2}{1 + f_{hi}} \left(\frac{1}{N} \frac{1}{\lambda_{hi}^3} + \int_{\lambda_{hi}}^{\infty} \frac{1}{t^3} dt \right) + \frac{2}{N} \frac{1}{\lambda_{hi}^2} + 2 \int_{\lambda_{hi}}^{\infty} \frac{1}{t^2} dt \\
 &= \frac{2}{1 + f_{hi}} \left(\frac{1}{N} \frac{1}{\lambda_{hi}^3} + \frac{1}{2\lambda_{hi}^2} \right) + \frac{2}{N} \frac{1}{\lambda_{hi}^2} + \frac{2}{\lambda_{hi}} \\
 &\stackrel{(c)}{\leq} \frac{3}{1 + f_{hi}} \frac{1}{\lambda_{hi}^2} + \frac{4}{\lambda_{hi}} < \frac{7}{\lambda_{hi}}. \tag{204}
 \end{aligned}$$

Above, (a) follows from (203) because $1/2 + 1/N < 0.55$; in (b) the bound for the first term follows because the function $1/t^3$ is monotonically decreasing and the bound for the second term follows because the function $1/t^2$ is monotonically decreasing; (c) follows because $1/N \leq \lambda_{hi}$.

Finally, plugging (202) and (204) into (201) and defining $c'_k = 8\pi + 14$ we obtain (106).

Proof of (107)

Begin by splitting the summation interval and recombining the terms in the following way:

$$\begin{aligned}
 \frac{1}{2} \sum_{n=0}^{N-1} \sup_{u \in \mathcal{N}(\lambda_{hi}, \frac{n}{N})} |k''_{hi}(u)| &= \frac{1}{2} \sum_{\frac{n}{N} \in [0, 2\lambda_{hi})} \sup_{u \in \mathcal{N}(\lambda_{hi}, \frac{n}{N})} |k''_{hi}(u)| \\
 &\quad + \frac{1}{2} \sum_{\frac{n}{N} \in [2\lambda_{hi}, 1/2)} \sup_{u \in \mathcal{N}(\lambda_{hi}, \frac{n}{N})} |k''_{hi}(u)| \\
 &\quad + \frac{1}{2} \sum_{\frac{n}{N} \in [1/2, 1-2\lambda_{hi})} \sup_{u \in \mathcal{N}(\lambda_{hi}, \frac{n}{N})} |k''_{hi}(u)| \\
 &\quad + \frac{1}{2} \sum_{\frac{n}{N} \in [1-2\lambda_{hi}, 1)} \sup_{u \in \mathcal{N}(\lambda_{hi}, \frac{n}{N})} |k''_{hi}(u)| \\
 &\leq \sum_{\frac{n}{N} \in [0, 2\lambda_{hi} + \frac{1}{N})} \sup_{u \in \mathcal{N}(\lambda_{hi}, \frac{n}{N})} |k''_{hi}(u)| \\
 &\quad + \sum_{\frac{n}{N} \in [2\lambda_{hi}, \frac{1}{2} + \frac{1}{N})} \sup_{u \in \mathcal{N}(\lambda_{hi}, \frac{n}{N})} |k''_{hi}(u)|. \tag{205}
 \end{aligned}$$

Above, the inequality follows by symmetry of $\sup_{u \in \mathcal{N}(\lambda_{hi}, \cdot)} |k''_{hi}(u)|$ around $1/2$. Next, we upper-bound the two sums separately.

To upper-bound the first sum in (205) we proceed as follows:

$$\sum_{\frac{n}{N} \in [0, 2\lambda_{hi} + \frac{1}{N})} \sup_{u \in \mathcal{N}(\lambda_{hi}, \frac{n}{N})} |k''_{hi}(u)| \leq \left(2\lambda_{hi} + \frac{1}{N}\right) N \max_t |k''_{hi}(t)|$$

$$\stackrel{(a)}{\leq} 3\lambda_{hi} 4\pi^2 f_{hi}^3 \stackrel{(b)}{=} \frac{12\pi^2}{\lambda_{hi}^2}. \tag{206}$$

Above, in (a) we used (200) and the assumption that $1/N \leq \lambda_{hi}$; in (b) we used that $f_{hi} = 1/\lambda_{hi}$.

To upper-bound the second term in (205) we differentiate $k_{hi}(\cdot)$ in (7) twice to obtain:

$$k''_{hi}(t) = \frac{\pi^2 (\sin(2\pi t) \sin(2\pi f_{hi}t) - \cos(2\pi t) \cos(2\pi f_{hi}t) - 2 \cos(2\pi f_{hi}t))}{N (f_{hi} + 1) \sin^4(\pi t)}$$

$$+ \frac{\pi^2 (\cos(2\pi t) + 2)}{N (f_{hi} + 1) \sin^4(\pi t)}$$

$$- \frac{4\pi^2 f_{hi} \sin(\pi(2f_{hi} + 1)t)}{N (f_{hi} + 1) \sin^3(\pi t)} + \frac{2\pi^2 f_{hi}^2 \cos(2\pi(f_{hi} + 1)t)}{N (f_{hi} + 1) \sin^2(\pi t)}.$$

This leads to the following upper bound on $|k''_{hi}(t)|$ for $t \in [0, 0.55]$:

$$|k''_{hi}(t)| \stackrel{(a)}{\leq} \frac{1}{N} \left(\frac{7\pi^2}{(f_{hi} + 1) \sin^4(\pi t)} + \frac{4\pi^2 f_{hi}}{(f_{hi} + 1) \sin^3(\pi t)} + \frac{2\pi^2 f_{hi}^2}{(f_{hi} + 1) \sin^2(\pi t)} \right)$$

$$\stackrel{(b)}{\leq} \frac{1}{N} \left(\frac{7\pi}{f_{hi}t^4} + \frac{4\pi}{t^3} + \frac{2\pi f_{hi}}{t^2} \right). \tag{207}$$

Above, in (a) we used the triangle inequality and the fact that $|\sin(\cdot)| \leq 1$, $|\cos(\cdot)| \leq 1$; in (b) we used the inequalities $\sin(\pi t)^2 \geq \pi t^2$, $\sin(\pi t)^3 \geq \pi t^3$, $\sin(\pi t)^4 \geq \pi t^4$ for $t \in [0, 0.55]$. Next observe that since the right-hand side of (207) is monotonically decreasing for $t > 0$ we have for $t \in (\lambda_{hi}, 0.55]$:

$$\sup_{u \in \mathcal{N}(\lambda_{hi}, t)} |k''_{hi}(u)| \leq \frac{1}{N} \left(\frac{7\pi}{f_{hi}(t - \lambda_{hi})^4} + \frac{4\pi}{(t - \lambda_{hi})^3} + \frac{2\pi f_{hi}}{(t - \lambda_{hi})^2} \right). \tag{208}$$

Therefore, the second term in (205) can be upper-bounded as follows:

$$\sum_{\frac{n}{N} \in [2\lambda_{hi}, \frac{1}{2} + \frac{1}{N})} \sup_{u \in \mathcal{N}(\lambda_{hi}, \frac{n}{N})} |k''_{hi}(u)|$$

$$\stackrel{(a)}{\leq} \frac{7\pi}{f_{hi}} \frac{1}{N} \sum_{\frac{n}{N} \in [2\lambda_{hi}, \frac{1}{2} + \frac{1}{N})} \frac{1}{(n/N - \lambda_{hi})^4}$$

$$\begin{aligned}
 &+ 4\pi \frac{1}{N} \sum_{\frac{n}{N} \in [2\lambda_{hi}, \frac{1}{2} + \frac{1}{N})} \frac{1}{(n/N - \lambda_{hi})^3} \\
 &+ 2\pi f_{hi} \frac{1}{N} \sum_{\frac{n}{N} \in [2\lambda_{hi}, \frac{1}{2} + \frac{1}{N})} \frac{1}{(n/N - \lambda_{hi})^2} \\
 \stackrel{(b)}{\leq} &\frac{7\pi}{f_{hi}} \left(\frac{1}{N} \frac{1}{(2\lambda_{hi} - \lambda_{hi})^4} + \int_{2\lambda_{hi}}^{1/2+1/N} \frac{1}{(t - \lambda_{hi})^4} dt \right) \\
 &+ 4\pi \left(\frac{1}{N} \frac{1}{(2\lambda_{hi} - \lambda_{hi})^3} + \int_{2\lambda_{hi}}^{1/2+1/N} \frac{1}{(t - \lambda_{hi})^3} dt \right) \\
 &+ 2\pi f_{hi} \left(\frac{1}{N} \frac{1}{(2\lambda_{hi} - \lambda_{hi})^2} + \int_{2\lambda_{hi}}^{1/2+1/N} \frac{1}{(t - \lambda_{hi})^2} dt \right) \\
 \stackrel{(c)}{\leq} &\frac{7\pi}{f_{hi}} \left(\frac{1}{N} \frac{1}{\lambda_{hi}^4} + \int_{\lambda_{hi}}^{\infty} \frac{1}{t^4} dt \right) + 4\pi \left(\frac{1}{N} \frac{1}{\lambda_{hi}^3} + \int_{\lambda_{hi}}^{\infty} \frac{1}{t^3} dt \right) \\
 &+ 2\pi f_{hi} \left(\frac{1}{N} \frac{1}{\lambda_{hi}^2} + \int_{\lambda_{hi}}^{\infty} \frac{1}{t^2} dt \right) \\
 = &\frac{7\pi}{f_{hi}} \left(\frac{1}{N} \frac{1}{\lambda_{hi}^4} + \frac{1}{3\lambda_{hi}^3} \right) + 4\pi \left(\frac{1}{N} \frac{1}{\lambda_{hi}^3} + \frac{1}{2\lambda_{hi}^2} \right) \\
 &+ 2\pi f_{hi} \left(\frac{1}{N} \frac{1}{\lambda_{hi}^2} + \frac{1}{\lambda_{hi}} \right) \\
 \stackrel{(d)}{\leq} &\left(\frac{28}{3}\pi + 6\pi + 4\pi \right) \frac{1}{\lambda_{hi}^2} = \frac{58\pi}{3} \frac{1}{\lambda_{hi}^2}. \tag{209}
 \end{aligned}$$

Above, (a) follows from (208) because $1/2 + 1/N < 0.55$; (b) follows because the functions $1/(t - \lambda_{hi})^4$, $1/(t - \lambda_{hi})^3$, and $1/(t - \lambda_{hi})^2$ are monotonically decreasing; (c) follows by changing the integration variable; (d) follows because $1/N \leq \lambda_{hi}$ and because $f_{hi} = 1/\lambda_{hi}$.

Finally, plugging (206) and (209) into (205) and defining $c_k'' \triangleq 12\pi^2 + 58\pi/3$ we obtain (107).

References

1. Abramowitz, M., Stegun, I.A.: Handbook of Mathematical Functions With Formulas, Graphs, and Mathematical Tables. Applied Mathematics Series, vol. 55, 10th edn. U.S. Government Printing Office, Washington, DC (1972)
2. Azaïš, J.M., de Castro, Y., Gamboa, F.: Spike detection from inaccurate samplings. Appl. Comput. Harmon. Anal. **38**(2), 177–195 (2015)
3. Barabell, A.: Improving the resolution performance of eigenstructure-based direction-finding algorithms. In: Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP), vol. 8, pp. 336–339 (1983)
4. Batenkov, D., Demanet, L., Goldman, G., Yomdin, Y.: Conditioning of partial nonuniform Fourier matrices with clustered nodes. SIAM J. Matrix Anal. Appl. **41**(1), 199–220 (2020)

5. Batenkov, D., Goldman, G., Yomdin, Y.: Super-resolution of near-colliding point sources. *Inf Inference* (2020). <https://doi.org/10.1093/imaiai/iaaa005>
6. Batenkov, D., Yomdin, Y.: On the accuracy of solving confluent Prony systems. *SIAM J. Appl. Math.* **73**(1), 134–154 (2013)
7. Betzig, E., Patterson, G.H., Sougrat, R., Lindwasser, O.W., Olenych, S., Bonifacino, J.S., Davidson, M.W., Lippincott-Schwartz, J., Hess, H.F.: Imaging intracellular fluorescent proteins at nanometer resolution. *Science* **313**(5793), 1642–1645 (2006)
8. Bhaskar, B.N., Tang, G., Recht, B.: Atomic norm denoising with applications to line spectral estimation. In: *Proc. Allerton Conf. Commun., Contr., Comput.*, 261–268 (2011)
9. Bienvenu, G.: Influence of the spatial coherence of the background noise on high resolution passive methods. In: *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, vol. 4, pp. 306–309 (1979)
10. Boyd, N., Schiebinger, G., Recht, B.: The alternating descent conditional gradient method for sparse inverse problems. *SIAM J. Optim.* **27**(2), 616–639 (2017)
11. Cadzow, J.A.: Signal enhancement—A composite property mapping algorithm. *IEEE Trans. Acoust. Speech Signal Process.* **36**(1), 49–62 (1988)
12. Candès, E.J., Fernandez-Granda, C.: Super-resolution from noisy data. *J. Fourier Anal. Appl.* **19**(6), 1229–1254 (2013)
13. Candès, E.J., Fernandez-Granda, C.: Towards a mathematical theory of super-resolution. *Commun. Pure Appl. Math.* **67**(6), 906–956 (2014)
14. Candès, E.J., Recht, B.: Simple bounds for recovering low-complexity models. *Mathematical Programming Series A* (2012)
15. Catala, P., Duval, V., Peyré, G.: A low-rank approach to off-the-grid sparse deconvolution. *J. Phys.* **904**, 012015 (2017)
16. Clergeot, H., Tressens, S., Ouamri, A.: Performance of high resolution frequencies estimation methods compared to the Cramer-Rao bounds. *IEEE Trans. Acoust. Speech Signal Process.* **37**(11), 1703–1720 (1989)
17. Demanet, L., Nguyen, N.: The recoverability limit for superresolution via sparsity. *CoRR* (2015). [arXiv:1502.01385](https://arxiv.org/abs/1502.01385)
18. Denoyelle, Q.: Theoretical and numerical analysis of super-resolution without grid. Ph.D. thesis, École Doctorale de Dauphine (2018)
19. Denoyelle, Q., Duval, V., Peyré, G.: Support recovery for sparse super-resolution of positive measures. *J. Fourier Anal. Appl.* **23**, 1153–1194 (2017)
20. Denoyelle, Q., Duval, V., Peyré, G., Soubies, E.: The sliding Frank-Wolfe algorithm and its application to super-resolution microscopy. *Inverse Problems* **36**(1), 014001 (2019)
21. DeVore, R.A., Lorentz, G.G.: *Constructive approximation*. Springer-Verlag, New York (1993)
22. Dickson, R.M., Cubitt, A.B., Tsien, R.Y., Moerner, W.: On/off blinking and switching behaviour of single molecules of green fluorescent protein. *Nature* **388**(6640), 355–358 (1997)
23. Donoho, D.L.: Superresolution via sparsity constraints. *SIAM J. Math. Anal.* **23**(5), 1309–1331 (1992)
24. Donoho, D.L., Johnstone, I.M., Hoch, J.C., Stern, A.S.: Maximum entropy and the nearly black object. *J. R. Stat. Soc. Ser. B* **54**(1), 41–81 (1992)
25. Duval, V.: A characterization of the non-degenerate source condition in super-resolution. *Inf. Inference* **9**(1), 235–269 (2019)
26. Eftekhari, A., Tanner, J., Thompson, A., Toader, B., Tyagi, H.: Sparse non-negative super-resolution—simplified and stabilised. *Appl. Comput. Harmon. Anal.* **50**, 216–280 (2021)
27. Fernandez-Granda, C.: Support detection in super-resolution. In: *10th international Conference on Sampling Theory and Applications (SampTA 2013)*, pp. 145–148. Bremen, Germany (2013)
28. Fernandez-Granda, C.: Super-resolution of point sources via convex programming. *Inf. Inference* **5**(3), 251–303 (2016)
29. Flinth, A., de Gournay, F., Weiss, P.: On the linear convergence rates of exchange and continuous methods for total variation minimization. *Math. Program.* (2020). <https://doi.org/10.1007/s10107-020-01530-0>
30. Fuchs, J.J.: Sparsity and uniqueness for some specific under-determined linear systems. In: *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, vol. 5, pp. v/729–v/732 (2005)
31. Helstrom, C.W.: The detection and resolution of optical signals. *IEEE Trans. Inf. Theory* **10**(4), 275–287 (1964)

32. Hua, Y., Sarkar, T.K.: Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise. *IEEE Trans. Acoust. Speech Signal Process.* **38**(5), 814–824 (1990)
33. Klar, T.A., Jakobs, S., Dyba, M., Egner, A., Hell, S.W.: Fluorescence microscopy with diffraction resolution barrier broken by stimulated emission. *Proc. Natl. Acad. Sci. USA* **97**(15), 8206–8210 (2000)
34. Kunis, S., Nagel, D.: On the smallest singular value of multivariate Vandermonde matrices with clustered nodes. *Linear Algebra Appl.* **604**, 1–20 (2020)
35. Li, W., Liao, W.: Conditioning of restricted Fourier matrices and super-resolution of MUSIC. In: 13th International Conference on Sampling Theory and Applications (SampTA), pp. 1–4 (2019)
36. Li, W., Liao, W.: Stable super-resolution limit and smallest singular value of restricted Fourier matrices. *Appl. Comput. Harmon. Anal.* **51**, 118–156 (2021)
37. Li, W., Liao, W., Fannjiang, A.: Super-resolution limit of the ESPRIT algorithm. *IEEE Trans. Inf. Theory* **66**(7), 4593–4608 (2020)
38. Liao, W., Fannjiang, A.: MUSIC for single-snapshot spectral estimation: stability and super-resolution. *IEEE Trans. Signal Process.* **63**(23), 6395–6406 (2015)
39. Liu, P., Zhang, H.: A theory of computational resolution limit for line spectral estimation (2020)
40. Liu, P., Zhang, H.: A mathematical theory of computational resolution limit in one dimension (2021)
41. Lütkepohl, H.: *Handbook of Matrices*. Wiley, Chichester (1996)
42. Moitra, A.: Super-resolution, extremal functions and the condition number of Vandermonde matrices. In: Proceedings of the Forty-Seventh Annual ACM Symposium on Theory of Computing, STOC '15, pp. 821–830 (2015)
43. Morgenshtern, V.I., Candès, E.J.: Super-resolution of positive sources: the discrete setup. *SIAM J. Imaging Sci.* **9**(1), 412–444 (2016)
44. Paulraj, A., Roy, R., Kailath, T.: A subspace rotation approach to signal parameter estimation. *Proc. IEEE* **74**(7), 1044–1046 (1986)
45. Pisarenko, V.F.: The retrieval of harmonics from a covariance function. *Geophys. J. Int.* **33**(3), 347–366 (1973)
46. Prony, R.: Essai expérimental et analytique. *J. l'Ecole Polytech. (Paris)* **1**(2), 24–76 (1795)
47. Roy, R., Kailath, T.: ESPRIT—estimation of signal parameters via rotational invariance techniques. *IEEE Trans. Acoust. Speech Signal Process.* **37**(7), 984–995 (1989)
48. Schiebinger, G., Robeva, E., Recht, B.: Superresolution without separation. *Inf. Inference* **7**(1), 1–30 (2017)
49. Schmidt, R.O.: Multiple emitter location and signal parameter estimation. *IEEE Trans. Antennas Propagat.* **AP-34**(3), 276–280 (1986)
50. Shahram, M., Milanfar, P.: Imaging below the diffraction limit: a statistical analysis. *IEEE Trans. Image Process.* **13**(5), 677–689 (2004)
51. Shahram, M., Milanfar, P.: On the resolvability of sinusoids with nearby frequencies in the presence of noise. *IEEE Trans. Signal Process.* **53**(7), 2579–2588 (2005)
52. Stoica, P., Moses, R.: *Spectral Analysis of Signals*. Prentice Hall, Hoboken (2005)
53. Stoica, P., Moses, R.L., Friedlander, B., Söderström, T.: Maximum likelihood estimation of the parameters of multiple sinusoids from noisy measurements. *IEEE Trans. Acoust. Speech Signal Process.* **37**(3), 378–392 (1989)
54. Stoica, P., Nehorai, A.: Statistical analysis of two nonlinear least-squares estimators of sine-wave parameters in the colored-noise case. *Circuits, Syst. Signal Process.* **8**(1), 3–15 (1989)
55. Stoica, P., Söderström, T.: Statistical analysis of MUSIC and subspace rotation estimates of sinusoidal frequencies. *IEEE Trans. Signal Process.* **39**(8), 1836–1847 (1991)
56. Tang, G., Bhaskar, B.N., Recht, B.: Near minimax line spectral estimation. *IEEE Trans. Inf. Theory* **61**(2), 499–512 (2015)
57. Tang, G., Bhaskar, B.N., Shah, P., Recht, B.: Compressed sensing off the grid. *IEEE Trans. Inf. Theory* **59**(11), 7465–7490 (2013)
58. Tufts, D.W., Kumaresan, R.: Estimation of frequencies of multiple sinusoids: making linear prediction perform like maximum likelihood. *Proc. IEEE* **70**(9), 975–989 (1982)