

On the Stability of Super-Resolution and a Beurling–Selberg Type Extremal Problem

Maxime Ferreira Da Costa and Urbashi Mitra

Department of Electrical and Computer Engineering, University of Southern California, USA

Abstract—Super-resolution estimation is the problem of recovering a stream of spikes (point sources) from the noisy observation of a few numbers of its first trigonometric moments. The performance of super-resolution is recognized to be intimately related to the separation between the spikes to recover. A novel notion of stability of the Fisher information matrix (FIM) of the super-resolution problem is introduced when the minimal eigenvalue of the FIM is *not* asymptotically vanishing. The regime where the minimal separation is inversely proportional to the number of acquired moments is considered. It is shown that there is a separation threshold above which the eigenvalues of the FIM can be bounded by a quantity that does not depend on the number of moments. The proof relies on characterizing the connection between the stability of the FIM and a generalization of the Beurling–Selberg box approximation problem.

I. INTRODUCTION

The classical formulation of super-resolution consists of recovering a stream of point sources (or spikes), characterized by their amplitudes and positions from noisy and distorted measurements. This popular setup serves as a parametric model for many inverse problems in applied and experimental sciences, such as spectrum analysis, system identification, radar, sonar, and optical imaging [1], [2], as well as wireless communications and sensing systems. When the distortion is modeled by a low-pass, bandlimited, shift-invariant point spread function (PSF), the super-resolution problem is also referred to as *line spectral estimation*, considered herein.

While super-resolution has been persistently studied from both theoretical and algorithmic perspectives for several decades, many open related statistical challenges remain. In particular, a complete stability analysis in the presence of noise has not occurred. Herein, the study of the stability of the Fisher information matrix (FIM) of the super-resolution problem, which is defined by the non-vanishing asymptotic of its smallest eigenvalue when the problem dimension becomes large, is undertaken. The existence of a minimal separation between the sources above which the FIM is stable is established. Through the Cramér–Rao lower bound (CRLB), a new algorithm-free statistical resolution limit to stably recover the point sources is provided. The proof relies on relating the extremal singular values of some generalized Vandermonde matrices, with a novel extension of the Beurling–Selberg box approximation problem.

This work has been funded in part by one or more of the following: Cisco Foundation 1980393, ONR N00014–15–1–2550, ONR 503400–78050, NSF CCF–1410009, NSF CCF–1817200, NSF CCF–2008927, Swedish Research Council 2018–04359, ARO W911NF1910269, and DOE DE–SC0021417.

Authors' emails: {mferreira, ubli}@usc.edu

A. Contributions and Organization of the Paper

Section II describes the signal model and defines the super-resolution model that would be considered in this work. After a review of the existing statistical and algorithmic resolution bounds, the notion of stability of the FIM is presented in Definition 1 and connected to the CRLB. The main result of this work is that the stability of the FIM for super-resolution is guaranteed whenever the separation parameter of the problem is greater than 3.54. This result is presented in Proposition 3. The key technical result relies on the bounding of the extremal singular values of the sensitivity matrix of the problem. Further, relevant bounds are established in Lemma 4 as a function of the separation between the point sources.

The remainder of the paper provides the proof of Lemma 4. In Section III, a connection between the Beurling–Selberg box approximation problem and the condition number of Vandermonde matrices with nodes on the unit circle is recalled from the literature. Based on this observation, a novel box approximation problem by bandlimited functions is defined, and a relationship between the solution of this problem and the conditioning of the sensitivity matrix is given in Proposition 6. A pair of sub-optimal solutions to the box approximation problem is constructed whenever the separation parameter is large enough, yielding the numerical bounds of Proposition 3. A detailed proof of Proposition 6 is given in Section IV. Conclusions and future works are drawn in Section V.

B. Notation and Definitions

Vectors of \mathbb{C}^N and matrices of $\mathbb{C}^{N \times r}$ are denoted by boldface letters \mathbf{a} and capital boldface letters \mathbf{A} , respectively. The minimal (resp. maximal) eigenvalues and singular values of a matrix \mathbf{A} are denoted $\lambda_{\min}(\mathbf{A})$, $\sigma_{\min}(\mathbf{A})$ (resp. $\lambda_{\max}(\mathbf{A})$, $\sigma_{\max}(\mathbf{A})$). For any function $g \in L_1(\mathbb{R})$, we denote by \hat{g} its *continuous time Fourier transform*, defined as

$$\hat{g}(u) = \int_{\mathbb{R}} g(t) e^{-i2\pi ut} dt, \quad \forall u \in \mathbb{R}. \quad (1)$$

A function $g \in L_1(\mathbb{R})$ is said to be β -bandlimited if for all u such that $|u| > \beta$, we have $\hat{g}(u) = 0$. For any $\alpha > 0$, I_α is the indicator function of the interval $[-\alpha/2, \alpha/2]$, i.e.,

$$I_\alpha(t) = \begin{cases} 1 & \text{if } |t| \leq \alpha/2, \\ 0 & \text{otherwise.} \end{cases} \quad (2)$$

which we call the *box function*. We let $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ to be the *unidimensional torus*. Let $N = 2n + 1$ be an odd number,

then for any $\tau \in \mathbb{T}$, we denote by $\mathbf{v}_0(\tau) \in \mathbb{C}^N$ and $\mathbf{v}_1(\tau) = \frac{d\mathbf{v}_0(\tau)}{d\tau} / \left\| \frac{d\mathbf{v}_0(\tau)}{d\tau} \right\|_2 \in \mathbb{C}^N$ the unit norm vectors

$$\mathbf{v}_0(\tau) = \frac{1}{\sqrt{N}} \left[e^{-i2\pi(-n)\tau}, \dots, e^{-i2\pi n\tau} \right]^\top \quad (3a)$$

$$\mathbf{v}_1(\tau) = \frac{C_N}{\sqrt{N}} \left[-i2\pi(-n)e^{-i2\pi(-n)\tau}, \dots, -i2\pi ne^{-i2\pi n\tau} \right]^\top, \quad (3b)$$

where $C_N = \sqrt{\frac{3}{\pi^2(N-1)(N+1)}}$ is a normalization constant. For any vector $\boldsymbol{\tau} = [\tau_1, \dots, \tau_r]^\top \in \mathbb{T}^r$, we define by $\mathbf{V}_0(\boldsymbol{\tau}), \mathbf{V}_1(\boldsymbol{\tau}) \in \mathbb{C}^{N \times r}$, the *generalized Vandermonde matrices*

$$\mathbf{V}_0(\boldsymbol{\tau}) = [\mathbf{v}_0(\tau_1), \dots, \mathbf{v}_0(\tau_r)], \quad (4a)$$

$$\mathbf{V}_1(\boldsymbol{\tau}) = [\mathbf{v}_1(\tau_1), \dots, \mathbf{v}_1(\tau_r)]. \quad (4b)$$

The concatenation $\mathbf{W}(\boldsymbol{\tau}) = [\mathbf{V}_0(\boldsymbol{\tau}), \mathbf{V}_1(\boldsymbol{\tau})] \in \mathbb{C}^{N \times 2r}$ is referred to as the *sensitivity matrix*. Finally, the *wrap-around distance* $\Delta(\boldsymbol{\tau})$ is defined by the minimal distance between two pairs of points in $\boldsymbol{\tau}$ over the torus, i.e. $\Delta(\boldsymbol{\tau}) \triangleq \min_{\ell \neq \ell'} \inf_{j \in \mathbb{Z}} |\tau_\ell - \tau_{\ell'} + j|$.

II. THE STABILITY OF SUPER-RESOLUTION

A. Signal Model

Consider r sources characterized by a vector $\mathbf{c} = [c_1, \dots, c_r]^\top \in \mathbb{C}^r$ of r complex amplitudes, and a vector $\boldsymbol{\tau} = [\tau_1, \dots, \tau_r]^\top \in \mathbb{T}^r$. Our super-resolution problem is to estimate the $2r$ parameters $\{\mathbf{c}, \boldsymbol{\tau}\}$ of the continuous time domain signal $x(t) = \sum_{\ell=1}^r c_\ell \delta(t - \tau_\ell)$ from the noisy observation $\mathbf{y} = [y_{-n}, \dots, y_0, \dots, y_n] \in \mathbb{C}^N$ of the form

$$\begin{aligned} \mathbf{y} &= [\hat{x}(-n), \dots, \hat{x}(n)]^\top + \mathbf{w} \\ &= \sum_{\ell=1}^r c_\ell \mathbf{v}_0(\tau_\ell) + \mathbf{w} = \mathbf{V}_0(\boldsymbol{\tau})\mathbf{c} + \mathbf{w}, \end{aligned} \quad (5)$$

where $\mathbf{w} \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I}_N)$ is assumed to be white Gaussian noise with variance σ^2 , and $r \leq n$ to ensure the uniqueness of the solution with r components in absence of noise [3]. Up to a scaling, it is assumed that $1 \leq |c_\ell| \leq \kappa$ of any $\ell = 1, \dots, r$, where κ is the *dynamic range* of the problem. As the vectors \mathbf{c} and $\boldsymbol{\tau}$ are of different units, and that the statistical error of an estimator of τ is expected to be inversely proportional to the number of measurement N , we denote by $\tilde{\boldsymbol{\tau}} = C_N^{-1} \boldsymbol{\tau}$ the normalized location of the points. We seek to recover, without loss of generality, the set of parameters $\boldsymbol{\theta} = \{\mathbf{c}, \tilde{\boldsymbol{\tau}}\}$.

B. Stability of Super-Resolution and Resolution Limits

An important statistical analysis goal for super-resolution is to guarantee that an estimator $\hat{\boldsymbol{\theta}}$ remains reasonably close to the ground truth $\boldsymbol{\theta}$. A variety of definitions exists in the literature to characterize the stability of super-resolution. All have been shown to be related to the separation parameter $\alpha \triangleq N\Delta(\boldsymbol{\tau})$ between the sources, as empirically established by Rayleigh (see e.g. [4]). We give a brief review of the key prior results in this context. When the sources are on a discrete grid, the stability of super-resolution has been shown

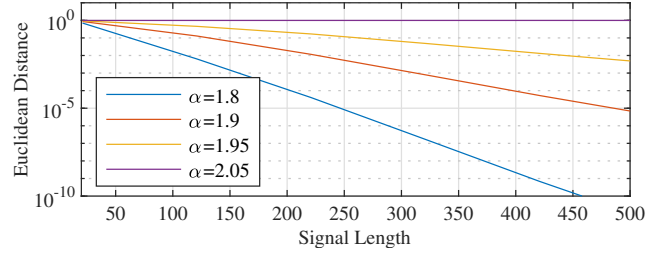


Figure 1. Euclidean distance $\|\mathbf{x}_1 - \mathbf{x}_2\|_2$ between two distinct noiseless signals with separation parameter α and $r = N/6$ sources, for different values of α and of the signal length N . Herein, the minimal cross-separation between the sets of locations τ_1 and τ_2 composing \mathbf{x}_1 and \mathbf{x}_2 is set to $\alpha/2$.

to be related to the Rayleigh index of the support set [5], [6]. In the last decade, there has been a resurgence in interest in characterizing the resolution limit for sources lying *off-the-grid*. It is well-established that some instances of super-resolution are statistically unidentifiable under a fixed noise level σ^2 when N grows large if the separation parameter satisfies $\alpha < 2$. That is, two ill-separated sets of distinct parameters can lead to arbitrarily close signals. This phenomena is highlighted in Figure 1 for the construction of two signals with $r = N/6$ regularly spaced sources, and can be explained by a phase transition in the conditioning of the matrix $\mathbf{V}_0(\boldsymbol{\tau})$ whenever $\alpha < 1$ [7], [8]. Additional studies of this property are proposed in the context of colliding pairs of sources [9]–[11], and the impact of the bandwidth selection was highlighted in [12].

Other lines of work aim to characterize instead the resolution limits of specific estimators for the super-resolution problem. For instance, stability guarantees for the MUSIC algorithm are provided as a function of the separation parameter in [13] in the noisy setting. The atomic norm denoiser [14]–[16] (a.k.a Beurling-LASSO) can fail to recover $\boldsymbol{\theta}$ in the absence of noise whenever $\alpha < 2$ [17], and is guaranteed to succeed for $\alpha \geq 2.56$ [14], [18]. Stability and denoising performance under Gaussian white noise are given for a separation $\alpha \geq 5.01$ in [19]. Finally, the support stability [20], ensuring that the estimate $\hat{\boldsymbol{\theta}}$ as the same number of components as of the ground truth $\boldsymbol{\theta}$, is established for $\alpha \geq 2.27$ in [21].

C. Stability of the Fisher Information Matrix

In this paper, we associate the stability of super-resolution with the property of the Fisher information matrix $\mathbf{J}(\boldsymbol{\theta})$ that its smallest eigenvalue is strictly bounded away from 0. A formal definition of stability is given in the following.

Definition 1 (Stability of the Fisher Information Matrix). *The super-resolution problem is to be stable for a separation parameter α and a dynamic range κ if and only if there exists a constant $C(\alpha, \kappa)$ independent of N such that for any set of parameters $\boldsymbol{\theta}$ with $N\Delta(\boldsymbol{\tau}) \geq \alpha$ and $1 \leq \min\{c\} \leq \max\{c\} \leq \kappa$*

$$\lambda_{\min}(\mathbf{J}(\boldsymbol{\theta})) \geq \sigma^{-2} C(\alpha, \kappa). \quad (6)$$

When read in the light of the CRLB, which states that the covariance $\Sigma_{\hat{\boldsymbol{\theta}}}$ of an unbiased estimator $\hat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$ given \mathbf{y} is such that $\Sigma_{\hat{\boldsymbol{\theta}}} - \mathbf{J}(\boldsymbol{\theta})^{-1}$ is a positive semi-definite Hermitian matrix

(see e.g. [22], [23]), Definition 1 can be reinterpreted in terms of the stability of estimates of linear forms as follows.

Proposition 2 (Bounded CRLB Property). *The super-resolution problem is stable in the sense of Definition 1 if and only if the CRLB of any linear form $z(\theta)$ on the parameters θ of the kind $z(\theta) = q^H \theta$ verifies*

$$\text{CRLB}(z(\theta); \mathbf{y}) \leq \sigma^2 C(\alpha, \kappa)^{-1} \|\mathbf{q}\|_2^2, \quad (7)$$

independently of N .

The contrapositive of Proposition 2 induces that the super-resolution problem is *not* stable if and only if there exists a linear form on the parameters that cannot be reliably estimated in the limit $N \rightarrow \infty$. Hence, Definition 1 is stronger than requiring the stability of each individual parameter composing θ , which is commonly adopted in the literature. Yet, the notion of stability considered in this work doesn't guarantee the actual existence of a consistent estimator of a given linear form. The main result of this paper, in the following proposition, establishes stability in the sense of the Fisher information provided that the separation parameter α is sufficiently large.

Proposition 3 (Sufficient Condition for Stability). *Suppose that $\alpha \geq 3.54$, then under Gaussian white noise $\mathbf{w} \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I})$, and for any $\kappa \geq 1$ the super-resolution problem is stable in the sense of Definition 1. Moreover, there exist a non-decreasing function $h_-(\alpha) > 0$ and a non-increasing function $h_+(\alpha) > 0$ such that*

$$\lambda_{\min}(\mathbf{J}(\theta)) \geq \sigma^{-2} h_-(\alpha) > 0 \quad (8a)$$

$$\lambda_{\max}(\mathbf{J}(\theta)) \leq \sigma^{-2} \kappa^2 h_+(\alpha). \quad (8b)$$

Figure 2 compares the empirical realizations of the extremal eigenvalues of the FIM with the functions $h_-(\cdot)$ and $h_+(\cdot)$ that will be later constructed in Section III. The evaluation $h_-(3.54) > 0$ suggests that $\alpha > 3.54$ is *sufficient* to guarantee the stability in the sense of Definition 1, however there is no reason to believe this bound will be tight. Instead, the numerical experiments suggest a phase transition on $\lambda_{\min}(\mathbf{J}(\theta))$ whenever $\alpha < 2$ as the eigenvalue rapidly vanishes to 0. This observation also coincides with the limit discussed in Section II-B. We leave a sharpening of the bounds for future work. The sequel is devoted to proving Proposition 3.

D. Proof of Proposition 3

We start the analysis by deriving the FIM $\mathbf{J}(\theta)$ (see e.g. [22], [24] of the super-resolution problem (5))

$$\mathbf{J}(\theta) = \sigma^{-2} \text{diag}(\mathbf{1}, \mathbf{c})^H \mathbf{W}(\tau)^H \mathbf{W}(\tau) \text{diag}(\mathbf{1}, \mathbf{c}), \quad (9)$$

where $\text{diag}(\mathbf{1}, \mathbf{c})$ denotes the $2r \times 2r$ diagonal matrix whose first r diagonal entries are equal to 1 and the following r are equal to \mathbf{c} . Some properties of this matrix were studied in [25], [26]. Equation (9) immediately yields that $\mathbf{J}(\theta)$ is a positive semi-definite matrix with the following bounds on its extremal eigenvalues

$$\begin{aligned} \lambda_{\min}(\mathbf{J}(\theta)) &\geq \sigma^{-2} \min(1, |c_1|, \dots, |c_r|)^2 \sigma_{\min}(\mathbf{W}(\tau))^2 \\ &\geq \sigma^{-2} \sigma_{\min}(\mathbf{W}(\tau))^2 \end{aligned} \quad (10a)$$

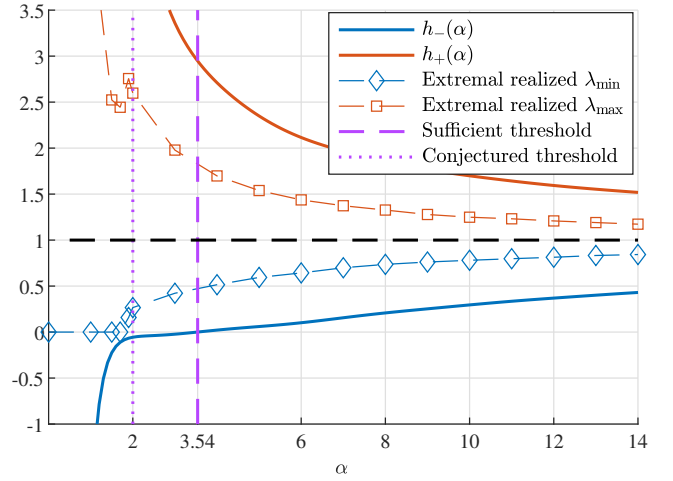


Figure 2. The theoretical bounds $h_-(\alpha)$ and $h_+(\alpha)$ versus the experimental extremal realizations of $\lambda_{\min}(\mathbf{J}(\theta))$ and $\lambda_{\max}(\mathbf{J}(\theta))$ when $N\Delta(\tau) = \alpha$. Parameter values are $N = 1001$, $\kappa = 1$ and $\sigma^2 = 1$. Experimental results are computed over 200 randomized trials.

$$\begin{aligned} \lambda_{\max}(\mathbf{J}(\theta)) &\leq \sigma^{-2} \max(1, |c_1|, \dots, |c_r|)^2 \sigma_{\max}(\mathbf{W}(\tau))^2 \\ &\leq \sigma^{-2} \kappa^2 \sigma_{\max}(\mathbf{W}(\tau))^2. \end{aligned} \quad (10b)$$

Therefore, the crux of the problem consists of ensuring the existence of the two functions $h_-(\cdot)$ and $h_+(\cdot)$, satisfying the conditions stated in Proposition 3 and that minorize (resp. majorize) the quantities $\sigma_{\min}(\mathbf{W}(\tau))^2$ (resp. $\sigma_{\max}(\mathbf{W}(\tau))^2$). This goal is accomplished in the following lemma and proved in Section III after a detour on the Beurling–Selberg extremal approximation problem.

Lemma 4 (Conditioning of $\mathbf{W}(\tau)$). *There exist a non-decreasing function $h_-(\cdot)$ and a non-increasing function $h_+(\cdot)$ such that for any N and any $\tau \in \mathbb{T}^r$ verifying $3.54 \leq \alpha \leq N\Delta(\tau)$, we have the following inequalities*

$$\sigma_{\min}(\mathbf{W}(\tau))^2 \geq h_-(\alpha) > 0 \quad (11a)$$

$$\sigma_{\max}(\mathbf{W}(\tau))^2 \leq h_+(\alpha). \quad (11b)$$

The proof of Proposition 3 immediately follows from (10) and Lemma 4. \square

III. RELATIONSHIP WITH BANDLIMITED APPROXIMATION

A. The Beurling–Selberg Box Approximation Problem

The Beurling–Selberg extremal box approximation problem consists of finding a minorant and a majorant to the box function I_α that are 1-bandlimited while minimizing the $L_1(\mathbb{R})$ distance to I_α [27], [28]. Selberg proposed a construction of two approximation functions $B_-^\alpha(\cdot)$ and $B_+^\alpha(\cdot)$ by leveraging the properties of Beurling's extremal approximant of the signum function [29]. In particular, those two functions are said to be *extremal* in the sense that they achieve $\hat{B}_-^\alpha(0) = 1 - \alpha^{-1}$ and $\hat{B}_+^\alpha(0) = 1 + \alpha^{-1}$ and that any 1-bandlimited minorant F_-^α (resp. majorant F_+^α) of I_α satisfies $\hat{F}_-(0) \leq 1 - \alpha^{-1}$ (resp. $\hat{F}_+(0) \geq 1 + \alpha^{-1}$). Those properties are particularly interesting because they can be used to bound, in a fairly elegant manner,

the singular values of Vandermonde matrices $\mathbf{V}_0(\tau)$ for any τ with separation parameter $\alpha = N\Delta(\tau)$ through (see e.g. [30])

$$\sigma_{\min}(\mathbf{V}_0(\tau))^2 \geq 1 - \alpha^{-1} \quad (12a)$$

$$\sigma_{\max}(\mathbf{V}_0(\tau))^2 \leq 1 + \alpha^{-1}. \quad (12b)$$

Henceforth, given a fixed $\alpha > 1$, it is immediately inferred from (12) that $\sigma_{\min}(\mathbf{V}_0(\tau))^2$ is non-vanishing, and that $\sigma_{\max}(\mathbf{V}_0(\tau))^2$ is upper bounded by a quantity independent of N for any values of $\alpha > 0$.

B. Higher Order Box Approximations

With the purpose of demonstrating Lemma 4, we introduce a novel *higher order* box approximation problem, which imposes an additional assumption on the decay rate of the minorizing and majorizing functions of I_α . The minorant set \mathcal{E}_α^- and minorant set \mathcal{E}_α^+ are defined as follows:

Definition 5 (2nd order minorant and majorant set).

Minorant set: Given $\alpha > 0$, we let $\mathcal{E}_\alpha^- \subset L_1(\mathbb{R})$ the sets of functions $\mathcal{X}_- \in L_1(\mathbb{R})$ satisfying the following properties:

- 1) \mathcal{X}_- is 1-bandlimited;
- 2) \mathcal{X}_- minorizes I_α , i.e. $\mathcal{X}_-(t) \leq I_\alpha(t)$ for all t ;
- 3) $\mathcal{X}_-(t) = \mathcal{O}(t^{-(3+\varepsilon)})$ when $|t| \rightarrow \infty$ for some $\varepsilon > 0$;
- 4) $\hat{\mathcal{X}}'_-(0) = 0$.

Majorant set: Given $\alpha > 0$, we let $\mathcal{E}_\alpha^+ \subset L_1(\mathbb{R})$ the sets of functions $\mathcal{X}_+ \in L_1(\mathbb{R})$ satisfying the following properties:

- 1) \mathcal{X}_+ is 1-bandlimited;
- 2) \mathcal{X}_+ majorizes I_α , i.e. $\mathcal{X}_+(t) \geq I_\alpha(t)$ for all t ;
- 3) $\mathcal{X}_+(t) = \mathcal{O}(t^{-(3+\varepsilon)})$ when $|t| \rightarrow \infty$ for some $\varepsilon > 0$;
- 4) $\hat{\mathcal{X}}'_+(0) = 0$.

We note that, by the Riemann–Lebesgue lemma, the third condition in both cases guarantees the Fourier transform of the functions in \mathcal{E}_α^- and \mathcal{E}_α^+ to be at least twice differentiable, which is needed in the fourth condition. The sets proposed in Definition 5 only differ from the original Beurling–Selberg approximation problem in the third and fourth assumptions. However, it can be easily checked from their definitions [29] that $B_\alpha^-(\cdot)$ and $B_\alpha^+(\cdot)$ do not belong to \mathcal{E}_α^- and \mathcal{E}_α^+ , respectively, as they decay at an asymptotic rate $B_\alpha^-(t) \sim B_\alpha^+(t) \sim Ct^{-2}$ when $|t| \rightarrow \infty$, thus fail to meet the third assumption.

The following proposition is at the core of the proof of Proposition 3 and establishes a relationship between the functions belonging to the set \mathcal{E}_α^- and \mathcal{E}_α^+ and the extremal singular values of the matrix $\mathbf{W}(\tau)$.

Proposition 6 (Conditioning of $\mathbf{W}(\tau)$ via Bandlimited Functions). *Let $N = 2n + 1$ be an odd integer. For any fixed $\alpha > 0$, let $\tau \in \mathbb{T}^r$, be such that $N\Delta(\tau) \geq \alpha$. Then, for any function $\mathcal{X}_- \in \mathcal{E}_\alpha^-$ and $\mathcal{X}_+ \in \mathcal{E}_\alpha^+$, we have*

$$\sigma_{\min}(\mathbf{W}(\tau))^2 \geq \min \left\{ \alpha^{-1} \hat{\mathcal{X}}_-(0), -\frac{3}{\pi^2} \alpha^{-3} \hat{\mathcal{X}}''_-(0) \right\} \quad (13a)$$

$$\sigma_{\max}(\mathbf{W}(\tau))^2 \leq \max \left\{ \alpha^{-1} \hat{\mathcal{X}}_+(0), -\frac{3}{\pi^2} \alpha^{-3} \hat{\mathcal{X}}''_+(0) \right\}. \quad (13b)$$

A few remarks are of order regarding the statement of Proposition 6. The crux of the problem is two fold. First, the extremal singular values can be controlled given the existence of functions within the the sets \mathcal{E}_α^- and \mathcal{E}_α^+ , hence a strategy to prove Proposition 3 is to show that those two sets are not empty for a large enough value of α . Second, noticing that $\hat{I}_\alpha(0) = \int_{\mathbb{R}} I_\alpha(t) dt = \alpha$ and $\hat{I}_\alpha''(0) = \int_{\mathbb{R}} -4\pi^2 t^2 I_\alpha(t) dt = -\frac{\pi^2}{3} \alpha^3$, the scaling of the bounds (13) can be considered to be tight in the sense that the right sides of both equations with the functions $\mathcal{X}_- = \mathcal{X}_+ = I_\alpha$ equal 1. Hence, better bounds are achieved for functions in \mathcal{E}_α^- and \mathcal{E}_α^+ that are close to the box function I_α . The best bounds on the singular values of the sensitivity matrix $\mathbf{W}(\tau)$ in Proposition 6 are provided by the *extremal functions* that achieve the maximal and minimal values on the right-hand sides of equations (13a) and (13b) respectively. However, to our knowledge, there are no known derivations to date for those extremal functions. We seek to ameliorate that lack in the sequel.

C. Approximation via Polynomial Transforms

In this section, we construct a pair of suboptimal functions $G_\alpha^-(\cdot)$ and $G_\alpha^+(\cdot)$ belonging to the sets \mathcal{E}_α^- and \mathcal{E}_α^+ , respectively and that achieve a reasonable approximation of the box function I_α . We proceed by applying a transform \mathcal{P} to the Beurling–Selberg extremal functions discussed in III-A which preserves the first and second properties of \mathcal{E}_α^- and \mathcal{E}_α^+ in Definition 5 respectively. Mathematically, it is sufficient to ensure that \mathcal{P} satisfies the following properties:

- 1) For any 1-bandlimited function F , then $\mathcal{P}\{F\}$ is also 1-bandlimited;
- 2) If F minorizes (resp. majorizes) the box function I_α then $\mathcal{P}\{F\}$ also minorizes (resp. majorizes) I_α ;
- 3) If $F(t) \rightarrow 0$ at $|t| \rightarrow \infty$ then $\mathcal{P}\{F\}(t) = \mathcal{O}(F(t)^2)$ in the limit $|t| \rightarrow \infty$.

Polynomial transforms are a good fit to achieve those three conditions, since any polynomial P of degree d preserves the spectral support of F through the map $\mathcal{P}\{F\}(t) = P(F(\frac{t}{d}))$. It can be checked that, while not unique, the polynomial $P(t) = t^3$ is the minimal degree polynomial for which \mathcal{P} matches those three properties. This suggests a construction of $G_\alpha^-(\cdot)$ and $G_\alpha^+(\cdot)$ as in Lemma 7, whose proof is omitted for brevity.

Lemma 7. *Let $\alpha > 0$, and denote by G_α^- and G_α^+ the functions given by*

$$G_\alpha^-(t) \triangleq \mathcal{P}\{B_{\alpha/3}^-\}(t) = B_{\alpha/3}^-(t/3)^3 \quad (14a)$$

$$G_\alpha^+(t) \triangleq \mathcal{P}\{B_{\alpha/3}^+\}(t) = B_{\alpha/3}^+(t/3)^3, \quad (14b)$$

then we have that $G_\alpha^- \in \mathcal{E}_\alpha^-$ and $G_\alpha^+ \in \mathcal{E}_\alpha^+$.

Figure 3 shows the graph of these functions and their Fourier transforms for three different values of the separation parameter α .

We now have all the elements to prove Lemma 4 which is integral to the proof of Proposition 3. We let $h_-(\alpha)$ and $h_+(\alpha)$ be the value of the right-hand side of the bounds (13) applied to the functions $G_\alpha^-(\cdot) \in \mathcal{E}_\alpha^-$ and $G_\alpha^+(\cdot) \in \mathcal{E}_\alpha^+$ defined in (14), respectively. The numerical evaluation of those quantity shown

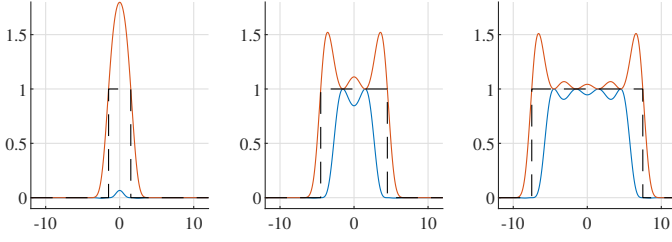


Figure 3. The graph of the functions $G_\alpha^-(t)$ (in blue), $G_\alpha^+(t)$ (in red) and $I_\alpha(t)$ (in black dashed lines). Left: $\alpha = 3$; Middle: $\alpha = 9$; Right: $\alpha = 15$

in Figure 2 ensures that $h_-(\alpha) > 0$ whenever $\alpha \geq 3.54$, which concludes on the desired statement. \square

IV. PROOF OF PROPOSITION 6

We focus on proving that (13a) holds, as (13b) can be derived in an analogous manner. Consider a vector $\boldsymbol{\tau} \in \mathbb{T}^r$ of an arbitrary number r of elements, and let $\alpha = N\Delta(\boldsymbol{\tau})$. We fix $\mathbf{u}_0, \mathbf{u}_1 \in \mathbb{C}^r$ and we let by $\mathbf{u} = [\mathbf{u}_0^\top, \mathbf{u}_1^\top]^\top \in \mathbb{C}^{2r}$, the concatenation of the two vectors. Assume that the function $\mathcal{X}_-(\cdot)$ is in the set \mathcal{E}_α^- . We start by introducing the auxiliary functions $\varphi_0(\cdot)$ and $\varphi_1(\cdot)$ defined for all $y \in \mathbb{R}$ by

$$\varphi_0(y) = \frac{1}{\sqrt{N}} \sum_{\ell=1}^r u_{0,\ell} e^{-i2\pi\tau_\ell y} \quad (15a)$$

$$\varphi_1(y) = -\frac{C_N}{\sqrt{N}} \sum_{\ell=1}^r u_{1,\ell} i2\pi\tau_\ell e^{-i2\pi\tau_\ell y}. \quad (15b)$$

and let by $S_p(\ell, \ell')$ for $p = 0, 1, 2$ the series

$$S_p(\ell, \ell') = \sum_{k=-\infty}^{\infty} \mathcal{X}_-(k\Delta(\boldsymbol{\tau})) (i2\pi k)^p e^{-i2\pi k(\tau_\ell - \tau_{\ell'})}. \quad (16)$$

Given these definitions, we can introduce \mathcal{X}_- to minorize the quantity $\|\mathbf{W}(\boldsymbol{\tau})\mathbf{u}\|_2^2$ as follows

$$\begin{aligned} \|\mathbf{W}(\boldsymbol{\tau})\mathbf{u}\|_2^2 &= \sum_{k=-n}^n |\varphi_0(k) + \varphi_1(k)|^2 \\ &= \sum_{k=-\infty}^{\infty} I_\alpha(k\Delta(\boldsymbol{\tau})) |\varphi_0(k) + \varphi_1(k)|^2 \\ &\geq \sum_{k=-\infty}^{\infty} \mathcal{X}_-(k\Delta(\boldsymbol{\tau})) |\varphi_0(k) + \varphi_1(k)|^2 \\ &= \frac{1}{N} \sum_{\ell=1}^r \sum_{\ell'=1}^r u_{0,\ell} u_{0,\ell'} S_0(\ell, \ell') \\ &\quad - \frac{2C_N}{N} \Re \left(\sum_{\ell=1}^r \sum_{\ell'=1}^r u_{0,\ell} u_{1,\ell'} S_1(\ell, \ell') \right) \\ &\quad + \frac{C_N^2}{N} \sum_{\ell=1}^r \sum_{\ell'=1}^r u_{1,\ell} u_{1,\ell'} S_2(\ell, \ell'), \quad (17) \end{aligned}$$

where the first inequality stems from the fact that $k\Delta(\boldsymbol{\tau}) \in [-\alpha/2, \alpha/2]$ if and only if $|k| \leq n$, and the second because $\mathcal{X}_-(t) \leq I_\alpha(t)$ for all $t \in \mathbb{R}$ by assumption. Moreover, the terms of the series $S_p(\ell, \ell')$ are equivalent to the function $h_p : x \mapsto \mathcal{X}_-(x\Delta(\boldsymbol{\tau})) (i2\pi x)^p e^{-i2\pi x(\tau_\ell - \tau_{\ell'})}$ when the argument is an integer. By the third assumption in Definition 5, the functions h_p are absolutely summable for $p = 0, 1, 2$ with continuous time Fourier transform $\hat{h}_p(u) = \Delta(\boldsymbol{\tau})^{-(1+p)} \hat{\mathcal{X}}_-^{(p)}(\Delta(\boldsymbol{\tau})^{-1}(u - \tau_\ell + \tau_{\ell'}))$, which is of finite support by the first assumption and bounded by the third one, therefore also absolutely summable. Hence, the Poisson summation formulae can be applied on the series $S_p(\ell, \ell')$. They yield for all pairs $(\ell, \ell') \in \{1, \dots, r\}^2$ and all $p = 0, 1, 2$

$$S_p(\ell, \ell') = \Delta(\boldsymbol{\tau})^{-(1+p)} \sum_{j=-\infty}^{\infty} \hat{\mathcal{X}}_-^{(p)}(\Delta(\boldsymbol{\tau})^{-1}(j - \tau_\ell + \tau_{\ell'})). \quad (18)$$

By our assumption on the separation, we have that $\Delta(\boldsymbol{\tau})^{-1}(j - \tau_\ell + \tau_{\ell'}) > 1$ for any $j \in \mathbb{Z}$ and any pair (ℓ, ℓ') unless $\ell = \ell'$ and $j = 0$. As $\mathcal{X}_-(\cdot)$ is 1-bandlimited, we conclude that the terms on the right hand side of (18) are all 0 unless $\ell = \ell'$ and $j = 0$. Thus, (18) reduces for $p = 0, 1, 2$ to

$$S_p(\ell, \ell') = \begin{cases} \Delta(\boldsymbol{\tau})^{-(p+1)} \hat{\mathcal{X}}_-^{(p)}(0) & \text{if } \ell = \ell', \\ 0 & \text{otherwise.} \end{cases} \quad (19)$$

Substituting (19) into (17) with $\hat{\mathcal{X}}_-^{(0)}(0) = 0$, $\frac{C_N^2}{N} \geq \frac{3}{\pi N^3}$ for all $N > 1$, and $\alpha = N\Delta(\boldsymbol{\tau})$ yields

$$\begin{aligned} \|\mathbf{W}(\boldsymbol{\tau})\mathbf{u}\|_2^2 &\geq (N\Delta(\boldsymbol{\tau}))^{-1} \hat{\mathcal{X}}_-(0) \|\mathbf{u}_0\|_2^2 - \frac{C_N^2}{N} \Delta(\boldsymbol{\tau})^{-3} \hat{\mathcal{X}}_-''(0) \|\mathbf{u}_1\|_2^2 \\ &= \alpha^{-1} \hat{\mathcal{X}}_-(0) \|\mathbf{u}_0\|_2^2 - \frac{3}{\pi^2} \alpha^{-3} \hat{\mathcal{X}}_-''(0) \|\mathbf{u}_1\|_2^2. \quad (20) \end{aligned}$$

Finally, the definition of the singular value and (20) imply

$$\begin{aligned} \sigma_{\min}(\mathbf{W}(\boldsymbol{\tau}))^2 &= \min_{\|\mathbf{u}\|_2=1} \|\mathbf{W}(\boldsymbol{\tau})\mathbf{u}\|_2^2 \\ &\geq \min \left\{ \alpha^{-1} \hat{\mathcal{X}}_-(0), -\frac{3}{\pi^2} \alpha^{-3} \hat{\mathcal{X}}_-''(0) \right\}, \quad (21) \end{aligned}$$

which yields our desired statement. \square

V. CONCLUSION AND FUTURE WORK

In the present work, we considered the stability of the Fisher information matrix of the super-resolution problem in the regime where the separation between the sources is inversely proportional to the number of measurements. Proposition 3 establishes the stability of the FIM whenever the separation parameter verifies $\alpha \geq 3.54$. This result was demonstrated by unveiling in Proposition 6 a connection between the singular values of the sensitivity matrix and the solutions of a novel Beurling–Selberg type approximation problem, for which we derived a pair of sub-optimal solutions. We leave for a sequel the study of the extremal solutions to this approximation problem in the hope of guaranteeing the stability of the FIM up to the conjectured threshold $\alpha > 2$.

REFERENCES

- [1] J. Lindberg, "Mathematical concepts of optical superresolution", *Journal of Optics - IOP Publishing*, vol. 14, no. 8, p. 83 001, 2012.
- [2] S.-H. Lee, J. Y. Shin, A. Lee, and C. Bustamante, "Counting single photoactivatable fluorescent molecules by photoactivated localization microscopy (PALM)", *Proceedings of the National Academy of Sciences*, vol. 109, no. 43, pp. 17 436–17 441, 2012.
- [3] J.-J. Fuchs, "Sparsity and uniqueness for some specific under-determined linear systems", in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 5, IEEE, 2005, pp. v–729.
- [4] M. Born and E. Wolf, *Principles of optics: electromagnetic theory of propagation, interference and diffraction of light*. Elsevier, 2013.
- [5] D. L. Donoho, "Superresolution via sparsity constraints", *SIAM journal on mathematical analysis*, vol. 23, no. 5, pp. 1309–1331, 1992.
- [6] L. Demanet and N. Nguyen, "The recoverability limit for superresolution via sparsity", *arXiv preprint arXiv:1502.01385*, 2015.
- [7] C. Aubel, D. Stotz, and H. Bölcskei, "A theory of super-resolution from short-time Fourier transform measurements", *Journal of Fourier Analysis and Applications*, vol. 24, no. 1, pp. 45–107, Feb. 2018.
- [8] A. Moitra, "Super-resolution, extremal functions and the condition number of Vandermonde matrices", in *Proceedings of the forty-seventh annual ACM symposium on Theory of computing*, ACM, 2015, pp. 821–830.
- [9] D. Batenkov, G. Goldman, and Y. Yomdin, "Super-resolution of near-colliding point sources", *Information and Inference: A Journal of the IMA*, vol. 10, no. 2, pp. 515–572, 2021.
- [10] S. Kunis and D. Nagel, "On the condition number of Vandermonde matrices with pairs of nearly-colliding nodes", *Numerical Algorithms*, vol. 87, no. 1, pp. 473–496, 2021.
- [11] B. Diederichs, "Well-Posedness of Sparse Frequency Estimation", *arXiv preprint arXiv:1905.08005*, 2019.
- [12] D. Batenkov, A. Bhandari, and T. Blu, "Rethinking super-resolution: the bandwidth selection problem", in *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, 2019, pp. 5087–5091.
- [13] W. Liao and A. Fannjiang, "MUSIC for single-snapshot spectral estimation: Stability and super-resolution", *Applied and Computational Harmonic Analysis*, vol. 40, no. 1, pp. 33–67, 2016.
- [14] E. J. Candès and C. Fernandez-Granda, "Towards a Mathematical Theory of Super-resolution", *Communications on Pure and Applied Mathematics*, vol. 67, no. 6, pp. 906–956, 2014.
- [15] Y. De Castro and F. Gamboa, "Exact reconstruction using Beurling minimal extrapolation", *Journal of Mathematical Analysis and Applications*, vol. 395, no. 1, pp. 336–354, 2012.
- [16] Y. Chi and M. Ferreira Da Costa, "Harnessing Sparsity Over the Continuum: Atomic norm minimization for superresolution", *IEEE Signal Processing Magazine*, vol. 37, no. 2, pp. 39–57, Mar. 2020.
- [17] M. Ferreira Da Costa and W. Dai, "A Tight Converse to the Spectral Resolution Limit via Convex Programming", in *2018 IEEE International Symposium on Information Theory (ISIT)*, Jun. 2018, pp. 901–905.
- [18] C. Fernandez-Granda, "Super-resolution of point sources via convex programming", *Information and Inference*, vol. 5, no. 3, pp. 251–303, Sep. 2016.
- [19] Q. Li and G. Tang, "Approximate support recovery of atomic line spectral estimation: A tale of resolution and precision", *Applied and Computational Harmonic Analysis*, 2018.
- [20] V. Duval and G. Peyré, "Exact support recovery for sparse spikes deconvolution", *Foundations of Computational Mathematics*, vol. 15, no. 5, pp. 1315–1355, Oct. 2015.
- [21] M. Ferreira Da Costa and Y. Chi, "On the Stable Resolution Limit of Total Variation Regularization for Spike Deconvolution", *IEEE Transactions on Information Theory*, 2020.
- [22] L. L. Scharf and L. T. McWhorter, "Geometry of the Cramér-Rao bound", *Signal Processing*, vol. 31, no. 3, pp. 301–311, 1993.
- [23] S. T. Smith, "Statistical resolution limits and the complexified Cramér-Rao bound", *IEEE Transactions on Signal Processing*, vol. 53, no. 5, pp. 1597–1609, May 2005.
- [24] P. Pakrooh, A. Pezeshki, L. L. Scharf, D. Cochran, and S. D. Howard, "Analysis of Fisher information and the Cramér-Rao bound for nonlinear parameter estimation after random compression", *IEEE Transactions on Signal Processing*, vol. 63, no. 23, pp. 6423–6428, 2015.
- [25] L. Li and T. P. Speed, "Parametric deconvolution of positive spike trains", *The Annals of Statistics*, vol. 28, no. 5, pp. 1279–1301, 2000.
- [26] T. Blu, P.-L. Dragotti, M. Vetterli, P. Marziliano, and L. Coulot, "Sparse sampling of signal innovations", *IEEE Signal Processing Magazine*, vol. 25, no. 2, pp. 31–40, 2008.
- [27] A. Selberg, *Collected papers. II*. Springer, 2014.
- [28] P. Akhilesh and D. Ramana, "A remark on the Beurling-Selberg function.", *Acta Mathematica Hungarica*, vol. 139, no. 4, 2013.
- [29] J. D. Vaaler, "Some extremal functions in Fourier analysis", *Bulletin (New Series) of The American Mathematical Society*, vol. 12, no. 2, pp. 183–216, 1985.
- [30] C. Aubel and H. Bölcskei, "Vandermonde matrices with nodes in the unit disk and the large sieve", *Applied and Computational Harmonic Analysis*, vol. 47, no. 1, pp. 53–86, 2019.