Support Stability of Spike Deconvolution via Total Variation Minimization

Maxime Ferreira Da Costa and Yuejie Chi

***Abstract*—Spike deconvolution is an inverse problem aiming at recovering point sources from their convolution with a point spread function (PSF). The stability of this problem in the presence of noise is long known to be closely related to the separation between those sources. It is therefore essential to characterize the resolution limit above which the point sources can be stably recovered from a given estimator, without spurious or missing sources from the estimate. In this paper, we establish the resolution limit above which the Beurling-LASSO estimator can stably recover two point sources, and show that the limit depends only on the PSF. Our result highlights the impact of PSF on the resolution limit in the noisy setting, which was not evident in previous studies of the noiseless setting. We further confirm our findings by comparing the theoretical limit with the empirical performance of the Beurling-LASSO estimator.**

I. Introduction

A. The resolution limits of spike deconvolution

The super-resolution problem consists of recovering a stream of one-dimensional point sources (or spikes) from distorted and noisy observations. This problem finds application in a broad variety of domains including, but not limited to, radar, sonar, optical imaging, wireless communications and sensing systems. The distortion degrading the point sources is often assumed to be of the form of a low-pass shift-invariant point spread function (PSF), whose bandwidth reflects the physical limitations of the measurement device involved in the acquisition of the point sources. In other words, super-resolution amounts to recovering the point sources from the observation of their convolution with the known PSF.

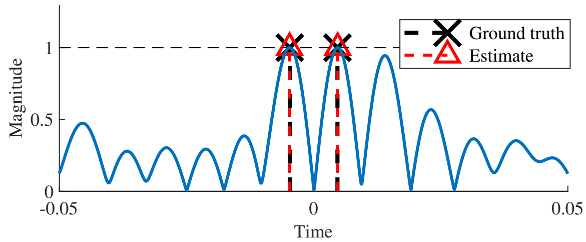
Of fundamental importance is the stability of the reconstruction from noisy observations, which has long been associated to the separation between the point sources. As an example, the Rayleigh limit (see e.g. [1]) is a popular criterion to characterize the *resolution limit* of the super-resolution problem. However, it stands as an empirical limit that does not rely on any statistical principle. Significant advances were made in recent years over a theoretical characterization of this limit from statistical perspectives. On one hand, super-resolution has been shown to be well-posed under a large enough separation of the sources, regardless of the estimator [2]. On the other hand, point sources are known to be statistically undistinguishable below a certain separation [3], [4]. Most of those results can be traced back to a phase transition phenomenon on the condition number of Vandermonde matrices with nodes on the unit circle [5]. However, they assume an asymptotic regime involving a diverging number of measurements, limiting their practical appeal.

In contrast, the *support stability* [6] has been proposed as a criterion to characterize the robustness of a broad range of parameter estimation problems, including super resolution. It is defined as the capability of an estimator to return exactly the same number of point sources as that of the ground truth, without spurious or missing elements. In view of the Rayleigh limit, it can be anticipated that the support stability of an estimator is directly related to the separation of the sources. Fig. 1 illustrates this intuition by plotting a ground truth measure containing two spikes and its reconstruction from the Beurling-LASSO estimator, a convex estimator based on the total variation (TV) minimization framework (defined in Section I-C), for two different separations of the sources, when the PSF is the ideal low-pass filter. Below a certain separation between the two ground truth sources, the estimator returns additional spurious point sources even at a high signal-to-noise ratio (SNR).

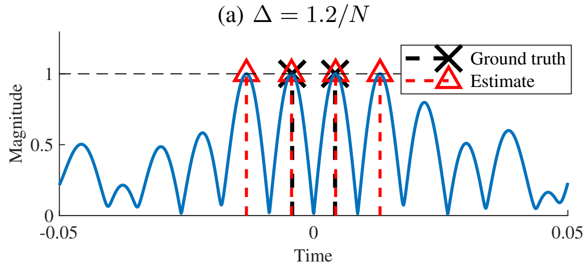
B. Contributions and organization of this paper

This paper focuses on characterizing the *support stability* of the Beurling-LASSO estimator [8] when the input signal is composed of two closely located point sources. Despite its apparent simplicity, this setup is of importance both in theory and in practice. In theory, it allows us to develop a deeper insight on the fundamental notion of *resolution limit* – the minimal distance above which two point sources are said to be distinguishable. In practice, it models the separation of a weak moving target from a strong clutter in radar [9], and accurate counting of the number of molecules in super-resolution fluorescence microscopy [10]. We establish in Theorem 2 a sufficient separation condition under which Beurling-LASSO is guaranteed to be support stable. Moreover this result highlights the dependency of the stable resolution

The authors are with the Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA 15213, USA (emails: {mferreira, yuejiechi}@cmu.edu). This work is supported in part by ONR under the grants N00014-18-1-2142 and N00014-19-1-2404, and by NSF under the CAREER grant ECCS-1818571.



(a) $\Delta = 1.2/N$



(b) $\Delta = 1.1/N$

Fig. 1. Support stability of the Beurling-LASSO estimator for reconstructing two point sources with different separations when the PSF is the ideal low-pass filter $g(\tau) = \mathrm{sinc}(\pi\tau)$. Here the number of samples is $N = 129$ and $\mathrm{SNR} = 40\mathrm{dB}$. The locations of point sources are estimated from (10) using the pick dual polynomial [7], drawn in blue. (a): $\Delta = 1.2/N$, the estimator returns exactly two spikes closely located to the ground truth and is support stable. (b): $\Delta = 1.1/N$, the estimator returns two additional spurious spikes and is not support stable.

for noisy super-resolution with respect to the PSF, which was not evident in the noiseless settings.

The rest of the paper is organized as follows. In Section II, we formulate the super-resolution problem and define the Beurling-LASSO estimator. Section III introduces the notion of support stability in Definition 1, and presents our main result in Theorem 2, with subsequent discussions in the light of related literature. Furthermore, the theoretical bound is corroborated by experimental simulations. A brief conclusion is drawn in Section IV.

C. Mathematical notations and definitions

Vectors in $\mathbb{C}^N$ are denoted by boldface letters such as $\boldsymbol{x}$. The Hilbert space of square integrable functions from $\mathbb{C}$ to $\mathbb{R}$ is denoted by L_2. We define by $\mathcal{M}(\mathbb{R})$ and $\mathcal{M}(\mathbb{T})$ the spaces of Radon measures defined over the reals and the torus $\mathbb{T} \sim \mathbb{R}/\mathbb{Z}$, respectively. The vector space of continuous functions from $\mathbb{T}$ to $\mathbb{C}$, denoted as $\mathcal{C}(\mathbb{T})$, is endowed with the supremum norm $\|\cdot\|_\infty$. The total variation norm $\|\cdot\|_{\mathrm{TV}}$ is defined as the dual norm of $\|\cdot\|_\infty$ and given for all $\mu \in \mathcal{M}(\mathbb{T})$ by

$$\|\mu\|_{\mathrm{TV}} = \sup_{\substack{h \in \mathcal{C}(\mathbb{T}) \\ \|h\|_\infty \le 1}} \Re \left[\int_{\mathbb{T}} \overline{h\,(t)} \mathrm{d}\mu\,(t) \right]. \quad (1)$$

II. Problem Formulation

A. Observation model

We consider a scenario where there are only two point sources to recover. Denoting by $\mathcal{M}(\mathbb{R})$ the set of complex Radon measures over the reals, the signal to resolve is modeled as a measure $\nu_\star \in \mathcal{M}(\mathbb{R})$ of the form

$$\nu_\star(\tau) = c_1\delta(\tau - \tau_1) + c_2\delta(\tau - \tau_2), \quad (2)$$

where $\delta(\cdot)$ is the Dirac measure, $\tau_1, \tau_2 \in \mathbb{R}$ are the time-domain locations of the two spikes and $c_1, c_2 \in \mathbb{C}\backslash\{0\}$ are their non-zero associated complex amplitudes. The continuous-time signal $x(\tau)$ resulted from the convolution of the ground truth measure $\nu_\star(\tau)$ with the PSF $g(\tau)$ writes as

$$\begin{aligned} x(\tau) &= (g * \nu_\star)(\tau) \\ &= c_1 g(\tau - \tau_1) + c_2 g(\tau - \tau_2), \quad \forall \tau \in \mathbb{R}, \end{aligned} \quad (3)$$

where $*$ denotes linear convolution. Furthermore, because of the needs of digital processing, one typically takes discrete-time measurements. An idealistic, yet credible approximation of many super-resolution problems encountered in practice is to consider measurements drawn from uniform sampling of the Fourier transform of $x(\tau)$. Let $\mathcal{F}(\cdot)$ be the Fourier transform of a measure in $\mathcal{M}(\mathbb{R})$, defined as

$$\mathcal{F}(\mu)(f) = \int_{\mathbb{R}} e^{-i2\pi f\tau} \mathrm{d}\mu(\tau), \quad \forall f \in \mathbb{R}, \text{ a.e..}$$

$\forall \mu \in \mathcal{M}(\mathbb{R})$. The Fourier-domain counterpart of the observation model (3) becomes

$$X(f) = G(f) \cdot \mathcal{F}(\nu_\star)(f), \quad \forall f \in \mathbb{R}, \text{ a.e.,}$$

where $X = \mathcal{F}(x)$, $G = \mathcal{F}(g)$ are the Fourier transforms of the signal $x(\tau)$ and the PSF $g(\tau)$, respectively. We assume that the PSF $g(\tau)$ is band-limited, with a bandwidth of $B > 0$. Therefore, $G(f) = 0$ for every f outside the interval $\left(-\frac{B}{2}, \frac{B}{2}\right)$. We further assume an odd number $N = 2n+1$ of measurements[1] is taken uniformly over the bandwidth $\left(-\frac{B}{2}, \frac{B}{2}\right)$. Therefore, the observation vector is given by $\boldsymbol{x} = \{x_k = X(kB/N)\}_{k=-n}^{n} \in \mathbb{C}^N$, corresponding to measuring $X(f)$ at frequencies $\{kB/N\}_{k=-n}^{n} \subset \left(-\frac{B}{2}, \frac{B}{2}\right)$.

For convenience, we introduce a normalized measure $\mu_\star \in \mathcal{M}(\mathbb{R})$ as $\mu_\star(t) = \frac{N}{B}\nu_\star(Nt/B)$ for all $t \in \mathbb{R}$, which by combining with (2) can be rewritten as,

$$\mu_\star(t) = c_1\delta(t - t_1) + c_2\delta(t - t_2), \quad (4)$$

where $t_1 = B\tau_1/N$ and $t_2 = B\tau_2/N$ are the normalized locations of the point sources. The observations $\boldsymbol{x}$ are linked to $\mu_\star$ by the linear relation

$$\boldsymbol{x} = \Phi_g(\mu_\star). \quad (5)$$

[1] An odd number of measurements is considered only for clarity and simplification purposes, and does not affect the generality of the results presented in this paper.

Here, the measurement operator Φ_g is defined by

$$\Phi_g : \mathcal{M}(\mathbb{R}) \to \mathbb{C}^N$$
$$\mu \mapsto \operatorname{diag}(\boldsymbol{g}) \left[\mathcal{F}(\mu)(-n), \ldots, \mathcal{F}(\mu)(n)\right]^\top, \tag{6}$$

where $\boldsymbol{g} = \{g_k = G(kB/N)\}_{k=-n}^{n} \in \mathbb{C}^N$ is the vector obtained by sampling the Fourier transform of the PSF $g(\tau)$ at frequencies $\{kB/N\}_{k=-n}^{n}$. Furthermore, notice that the observation operator Φ_g is invariant with respect to integer shifts of the underlying measure $\mu_\star$. Thus, one can only hope to identify $\mu_\star$ over $\mathcal{M}(\mathbb{T})$. Without loss of generality, the delays t_1, t_2 are normalized within the unit interval, i.e. $t_1, t_2 \in [-\frac{1}{2}, \frac{1}{2})$.[2]

In the presence of noise or measurement errors, we assume $\boldsymbol{x}$ is corrupted by an additive term $\boldsymbol{w}$. The observations are given as

$$\boldsymbol{z} = \boldsymbol{x} + \boldsymbol{w} = \Phi_g(\mu_\star) + \boldsymbol{w}, \tag{7}$$

where $\|\boldsymbol{w}\|_2 \leq \eta$ is assumed to be bounded for some noise level $\eta > 0$.

B. Reconstruction via total variation minimization

In the absence of noise, the super-resolution problem is defined as recovering $\mu_\star$ from the observations $\boldsymbol{x}$ and the PSF $g(\tau)$, yielding a linear inverse problem over the set of measures. Clearly, the problem is ill-posed as there are an uncountable infinity of measure μ that can explain the noiseless observations $\boldsymbol{x}$. Harnessing a sparsity prior on the ground truth measure $\mu_\star$, the ill-posedness of the problem can be addressed by searching for the measure $\widehat{\mu}_\natural$ with minimal support that is a solution of (5). Equivalently, denoting by $\|\cdot\|_0$ the "pseudo-norm" counting the potentially infinite cardinality of the support of a measure in $\mathcal{M}(\mathbb{T})$, the optimal estimator $\widehat{\mu}_\natural$ for the super-resolution problem is given by the solution of the optimization program

$$\widehat{\mu}_\natural = \underset{\mu \in \mathcal{M}(\mathbb{T})}{\arg\min} \|\mu\|_0 \quad \text{s.t.} \quad \boldsymbol{x} = \Phi_g(\mu), \tag{8}$$

which is known to have a unique solution equal to the ground truth $\mu_\star$ whenever the number of measurements N is at least twice as large as the number of spikes to recover [11].

An immediate drawback of the formulation (8) is the computational infeasibility of the minimization of the pseudo-norm $\|\cdot\|_0$. To overcome this issue, it is proposed instead in [12] to solve a convex relaxation of the estimator (8) to recover the ground truth measure. This is achieved by relaxing the cost function of (8) by the total variation norm, which is one of its convex surrogate. The total variation minimization of measures, also known as the atomic norm minimization [13], [14], is a convex optimization framework to regularize a variety of linear inverse problems over continuous dictionaries. The resulting TV estimator, denoted as $\widehat{\mu}_0$, is given by the convex optimization program

$$\widehat{\mu}_0 = \underset{\mu \in \mathcal{M}(\mathbb{T})}{\arg\min} \ \|\mu\|_{\mathrm{TV}} \quad \text{s.t.} \quad \boldsymbol{x} = \Phi_g(\mu). \tag{9}$$

Despite the infinite dimensionality of the space of Radon measure $\mathcal{M}(\mathbb{T})$, the program (9) can be computed efficiently by solving an associated semidefinite program in a finite dimension N (see e.g. [7]).

In the presence of noisy observations of the form (7), the Beurling-LASSO estimator $\widehat{\mu}_\lambda$ [8], also known as the atomic norm denoiser [7], can be used to recover the ground truth and its estimate $\widehat{\mu}_\lambda$ is written as

$$\widehat{\mu}_\lambda = \widehat{\mu}_\lambda(\boldsymbol{z}) = \underset{\mu \in \mathcal{M}(\mathbb{T})}{\arg\min} \ \frac{1}{2} \|\boldsymbol{z} - \Phi_g(\mu)\|_2^2 + \lambda \|\mu\|_{\mathrm{TV}}, \tag{10}$$

where $\lambda > 0$ is a regularization parameter drawing a trade-off between the TV norm of the estimate, as well as its fidelity to the observations.

III. Stability of the Beurling-LASSO

A. Support stability

Among the many possible metrics quantifying the stability of parameter estimation problems, the *support stability*, first introduced in [6], is of particular interest in the context of super-resolution. Roughly speaking, an estimator is said to be support stable if it outputs a measure containing the same number of point sources as that of the ground truth when the noise level is bounded. A formal definition of this criteria is given in the following.

Definition 1 (Support stability)**.** *Consider the observations $\boldsymbol{z} = \Phi_g(\mu_\star) + \boldsymbol{w}$. An estimator $\widehat{\mu} = \widehat{\mu}(\boldsymbol{z})$ based on $\boldsymbol{z}$ is said to be support stable for a given ground truth measure $\mu_\star$ of the form* (4) *if there exists $\eta > 0$ such that for all $\boldsymbol{w}$ with $\|\boldsymbol{w}\|_2 < \eta$, the estimate $\widehat{\mu}$ is a measure containing two spikes, i.e.*

$$\widehat{\mu}(\boldsymbol{z}) = \widehat{c}_1 \delta(t - \widehat{t}_1) + \widehat{c}_2 \delta(t - \widehat{t}_2),$$

and if the estimated parameters satisfy, up to a permutation Π of the indices: $\left|t_k - \widehat{t}_{\Pi(k)}\right|_{\mathbb{T}} = \mathcal{O}(\|\boldsymbol{w}\|_2)$ and $\left|c_k - \widehat{c}_{\Pi(k)}\right| = \mathcal{O}(\|\boldsymbol{w}\|_2)$ for $k = 1, 2$ in the limit of $\|\boldsymbol{w}\|_2 \to 0$.

As announced in Section I-A, the support stability of the Beurling-LASSO is excepted to be related to the separation of the sources defined as $\Delta = |t_2 - t_1|_{\mathbb{T}}$, where the distance is taken over the torus $\mathbb{T}$. This phenomena is highlighted by Fig. 1.

[2] Since $t_i = B\tau_i/N$, $i = 1, 2$, and assuming $\tau_i \in [-T/2, T/2)$, where T is the time window of interest, then the ambiguity constraint $t_i \in [-1/2, 1/2)$ suggests that the number of measurements should satisfy $N \geq T \cdot B$, the time-bandwidth product, to avoid aliasing.

B. Main results

We introduce the following auxiliary functions before stating the main result of this paper. We denote by $\kappa(\tau)$ the autocorrelation function of the PSF $g \in L_2$, which is defined as

$$\kappa(\tau) = \int_{\mathbb{R}} \overline{g(y)} g(\tau + y) \mathrm{d}y, \quad \forall \tau \in \mathbb{R}. \tag{11}$$

Additionally, the function $u_\beta, v_\beta \in L_2$, representing the sum and difference of two delayed version of the function κ by a shift $\beta \geq 0$ are given by

$$u_\beta(\tau) = \kappa(\tau - \frac{\beta}{2}) + \kappa(\tau + \frac{\beta}{2}), \tag{12a}$$

$$v_\beta(\tau) = \kappa(\tau - \frac{\beta}{2}) - \kappa(\tau + \frac{\beta}{2}). \tag{12b}$$

Our main contribution, summarized in Theorem 2, states that, under some mild smoothness and band-limited assumptions on the PSF $g(\tau)$, the Beurling-LASSO estimator is support stable over the set of two-spike measure provided that their separation $\Delta > \gamma^\star/N$ is large enough. Moreover the value of $\gamma^\star$ depends only on the PSF $g(\tau)$. We refer the reader to [15] for a stronger statement of this result, and its full proof.

Theorem 2 (Stable resolution limit of Beurling-LASSO)**.** *Suppose that the PSF g satisfies the following regularity conditions (H1) and (H2).*

(H1) *$g \in L_2$ is non-zero, real and three times differentiable, and verifies for some $\delta > 0$ and $C_\ell > 0$*

$$\left| g^{(\ell)}(\tau) \right| \leq \frac{C_\ell}{1 + |\tau|^{1+\delta}}, \quad \forall \tau \in \mathbb{R},\ \ell = 0, 1, 2, 3, \tag{13}$$

(H2) *$G = \mathcal{F}(g) \in L_2$ is band-limited within B, so that $G(f) = 0$, $\forall |f| > B/2$.*

Let $\gamma^\star$ the positive constant, depending only on the PSF g, defined as $\gamma^\star = \max\{\gamma_1^\star, \gamma_2^\star, \gamma_3^\star\} > 0$ with

$$\gamma_1^\star = B \sup_{\beta > 0} \left\{ \sup_{\tau \geq 0} |\tilde{r}_\beta(\tau)| > \tilde{r}_\beta\left(\frac{\beta}{2}\right) \right\}, \tag{14a}$$

$$\gamma_2^\star = B \sup_{\beta > 0} \left\{ \sup_{\tau \geq 0} |\tilde{s}_\beta(\tau)| > \tilde{s}_\beta\left(\frac{\beta}{2}\right) \right\}, \tag{14b}$$

$$\gamma_3^\star = B \sup_{\beta > 0} \left\{ -\kappa''(0)^2 + \kappa''(\beta)^2 - \kappa'(\beta)\kappa'''(\beta) \geq 0 \right\}, \tag{14c}$$

where the intermediate functions $\tilde{r}_\beta(\tau)$, $\tilde{s}_\beta(\tau)$ are further defined, for any $\beta > 0$ and $\tau \in \mathbb{R}$ as

$$\tilde{r}_\beta(\tau) = (-\kappa''(0) + \kappa''(\beta)) u_\beta(\tau) + \kappa'(\beta) v_\beta'(\tau) \tag{15a}$$

$$\tilde{s}_\beta(\tau) = (-\kappa''(0) - \kappa''(\beta)) v_\beta(\tau) - \kappa'(\beta) u_\beta'(\tau). \tag{15b}$$

Then there exists $N_0 \in \mathbb{N}$ such that, for every $N \geq N_0$ and every $\mu_\star$ of the form (4) *with*

$$|t_1 - t_2|_{\mathbb{T}} > \frac{\gamma^\star}{N}, \tag{16}$$

there exists $\alpha > 0$ such that the Beurling-LASSO estimator $\widehat{\mu}_\lambda(\Phi_g(\mu_\star) + \boldsymbol{w})$ with the regularization parameter $\lambda = \alpha^{-1} \|\boldsymbol{w}\|_2$ is support stable.

A few remarks are in order regarding the statement of Theorem 2. Firstly, it provides an explicit means to compute $\gamma^\star$, based on the evaluation of (14), for a given PSF satisfying the regularity conditions. The key quantities, $\gamma_k^\star$, $k = 1, 2, 3$ are suprema of continuous functions, and the complexity of the computation essentially depends on the variations and smoothness of the autocorrelation function κ. Moreover, the constants $\gamma_k^\star$s are invariant through a re-scaling the PSF via a transform $g(\tau) \leftarrow g(c\tau)$ for some $c > 0$. As a result $\gamma^\star$ can be evaluated for a PSF with nominal bandwidth $B = 1$. Table I provides the value of the constant $\gamma^\star$ for frequently encountered PSFs,[3] including the truncated Gaussian function, and the prolate spheroidal wave functions (PSWF)[4]. Fig. 2 illustrates how the constant $\gamma^\star$ increases while the temporal concentration of the truncated Gaussian function and the prolate spheroidal wave function degenerates.

Additionally, the separation condition (16) can be equivalently interpreted in terms of the delays τ_1, τ_2 of the unnormalized measure $\nu_\star$ as in (2) as

$$B |\tau_2 - \tau_1| > \gamma^\star, \tag{17}$$

provided the spikes to be localized are in the interval $(-N/(2B), N/(2B))$. Finally, our results allow arbitrary coefficients of the spikes, as long as they are sufficiently separated by $\gamma^\star$.

C. Connections to the related literature

The impact of the separation between the sources on the performance of the TV estimator (9) has been studied extensively in the noiseless setting. Exact recovery is first guaranteed in [12] under a separation of the sources $\Delta > 4/N$ provided that N is large enough. This result has been later improved to $\Delta \geq 2.56/N$ in [18]. On the other hand, it is known that TV regularization can fail for some signals with $\Delta < 2/N$ [19]. The previous constants are significantly larger than the one established in Table I as no prior assumptions on the number of sources are made in those works. When considering only

[3] Some of the PSFs listed in Table I, such as the ideal low-pass filter, do not satisfy the condition (13) of Theorem 2. Nevertheless, they can be handled by a stronger version of the result proposed in [15].

[4] The PSWF ψ_{τ_0} for the temporal concentration band $[-\tau_0, \tau_0]$ is defined as a function $g(\cdot)$ with a frequency band $\left(-\frac{1}{2}, \frac{1}{2}\right)$ and with $\|g\|_{L_2} = 1$ which maximizes the integral $\int_{-\tau_0}^{\tau_0} |g(\tau)|^2 \mathrm{d}\tau$ [16], [17].

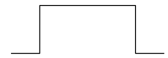
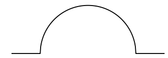
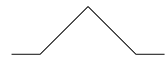
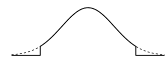
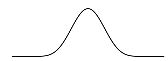
Point spread function	Fourier transform	$\gamma^\star$
Ideal low-pass: $\mathrm{sinc}(\pi\tau)$		1.132
Circular low-pass[5]: $J_0(\pi\tau)/\sqrt{\pi\tau}$		1.253
Triangular low-pass: $\mathrm{sinc}(\pi\tau/2)^2$		1.449
Truncated Gaussian: $e^{-\frac{\tau^2}{2\sigma^2}} * \mathrm{sinc}(\pi\tau)$		(see Fig. 2)
Prolate spheroidal wave function: $\psi_{\tau_0}(\tau)$		(see Fig. 2)

TABLE I
VALUES OF THE MINIMAL SEPARATION $\gamma^\star$ FOR COMMONLY ENCOUNTERED POINT SPREAD FUNCTIONS.

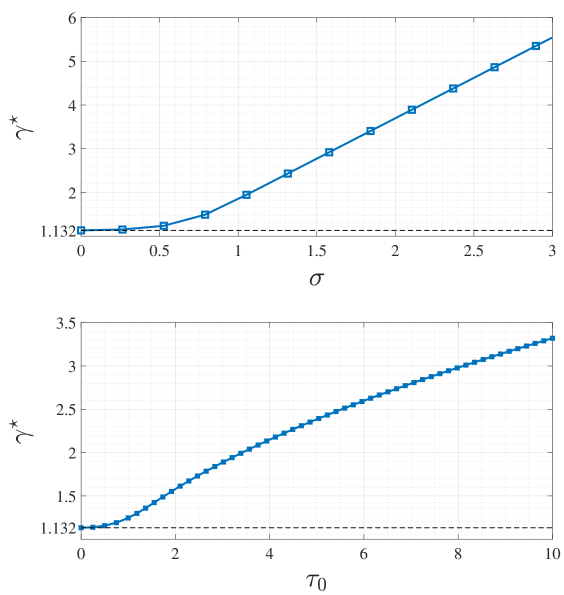


Fig. 2. Top: The stable resolution limit $\gamma^\star$ for a truncated Gaussian PSF for different values of the parameter σ. Bottom: The prolate spheroidal wave function of order zero ψ_{τ_0} for different widths of the concentration band $[-\tau_0, \tau_0]$.

two sources, as in (4), it is shown in [6] that a separation $\Delta > 1/N$ is necessary to guarantee exact recovery.

Guarantees in terms of support recovery have been given in the presence of noise in [20], [21], [22], by bounding the residual of $\widehat{\mu}$ outside the support of $\mu_\star$. Those bounds do not provide, however, a guarantee on the absence of spurious/missing point sources in the estimate. Under an extra white Gaussian noise assumption, a trade-off between the separation of the sources and the error of the parameters is highlighted in [23], when the PSF is the ideal low-pass filter.

More interestingly, in the context of this paper, the support stability of the Beurling-LASSO estimator is studied in [6] for a broad range of measurement operators using a "non-degenerate source condition". The analysis is based on the asymptotic behaviors of the dual solution of (10) when the noise level η and the regularization parameter λ simultaneously tend to 0. Yet, in the general case, it is challenging to explicitly verify this condition in terms of the separation parameter of the sources. In [24], [25], the support stability of reconstructing positive sources is considered without imposing a minimal separation condition. The proof of Theorem 2 is achieved by verifying the non-degenerate source condition for the two-spike case with arbitrary signs, which is already quite technical and non-trivial.

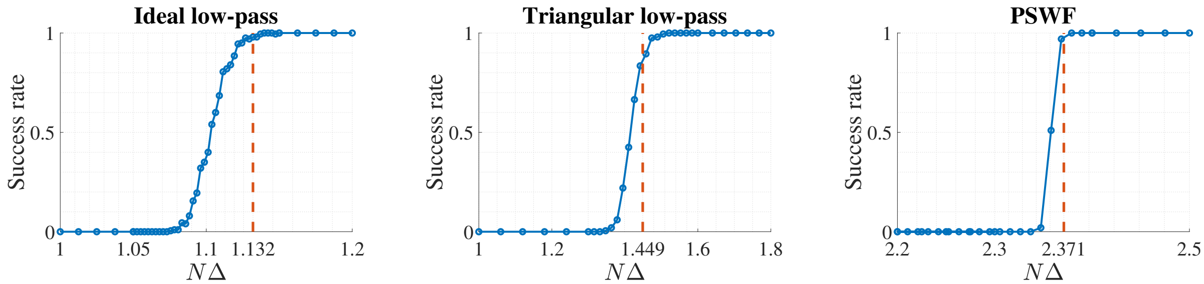
D. Numerical experiments

It is natural to compare the value of the constant $\gamma^\star$ anticipated by Theorem 2 with the empirical performance of the Beurling-LASSO estimator. Such a comparison is provided in Fig. 3 by varying the separation parameter $N\Delta$ for different PSFs. The empirical success rate suggests the existence of a phase transition on the support stability of the reconstruction around the value $\gamma^\star$, which supports the findings of our theory.

IV. CONCLUSIONS

This paper studies the support stability of the Beurling-LASSO estimator for estimating two closely located point sources and characterizes the resolution limit as a function of the PSF, above which the Beurling-LASSO estimator is support stable. Our result highlights and quantifies the role of PSF in noisy super resolution, which is not

[5] $J_0(\cdot)$ denotes the Bessel function of the first kind.



(a) $g(\tau) = \mathrm{sinc}(\pi\tau)$ (b) $g(\tau) = \mathrm{sinc}(\pi\tau/2)^2$ (c) $g(\tau) = \psi_5(\tau)$

Fig. 3. Empirical success rate for the Beurling-LASSO estimator to return a measure with two point sources, for three different PSFs, under additive white Gaussian noise, as a function of the separation parameter $N\Delta$. The support stability threshold $\gamma^\star$ predicted by Theorem 2 is shown in red. Here, we set $N = 101$, $\mathrm{SNR} = 40\mathrm{dB}$. The results are averaged over 200 trials.

evident in the study of the noiseless setting. In the future, it is worthwhile to further investigate the scenario with more than two point sources.

References

[1] M. Born and E. Wolf, *Principles of optics: electromagnetic theory of propagation, interference and diffraction of light*. Cambridge university press, 1999.

[2] B. Diederichs, "Well-Posedness of Sparse Frequency Estimation," *arXiv preprint arXiv:1905.08005*, 2019.

[3] 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.

[4] D. Batenkov, L. Demanet, G. Goldman, and Y. Yomdin, "Stability of partial Fourier matrices with clustered nodes," *arXiv preprint arXiv:1809.00658*, 2018.

[5] 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.

[6] 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.

[7] B. N. Bhaskar, G. Tang, and B. Recht, "Atomic norm denoising with applications to line spectral estimation," *IEEE Transactions on Signal Processing*, vol. 61, no. 23, pp. 5987–5999, 2013.

[8] 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.

[9] Y. Chi, A. Pezeshki, R. Calderbank, and S. Howard, "Range sidelobe suppression in a desired doppler band," in *2009 International Waveform Diversity and Design Conference*, Feb. 2009.

[10] 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.

[11] D. L. Donoho and M. Elad, "Optimally sparse representation in general (non-orthogonal) dictionaries via ℓ_1 minimization," *Proc. Natl. Acad. Sci.*, vol. 100, pp. 2197–2202, 2002.

[12] 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.

[13] V. Chandrasekaran, B. Recht, P. A. Parrilo, and A. S. Willsky, "The convex geometry of linear inverse problems," *Foundations of Computational Mathematics*, vol. 12, no. 6, pp. 805–849, 2012.

[14] Y. Chi and M. Ferreira Da Costa, "Harnessing sparsity over the continuum: Atomic norm minimization for super resolution," *arXiv preprint arXiv:1904.04283*, 2019.

[15] M. Ferreira Da Costa and Y. Chi, "On the stable resolution limit of total variation regularization for spike deconvolution," *arXiv preprint arXiv:1910.01629*, 2019.

[16] D. Slepian and H. O. Pollak, "Prolate spheroidal wave functions, Fourier analysis and uncertainty—I," *Bell System Technical Journal*, vol. 40, no. 1, pp. 43–63, 1961.

[17] I. C. Moore and M. Cada, "Prolate spheroidal wave functions, an introduction to the Slepian series and its properties," *Applied and Computational Harmonic Analysis*, vol. 16, no. 3, pp. 208–230, 2004.

[18] C. Fernandez-Granda, G. Tang, X. Wang, and L. Zheng, "Demixing sines and spikes: Robust spectral super-resolution in the presence of outliers," *Information and Inference: A Journal of the IMA*, vol. 7, no. 1, pp. 105–168, 2017.

[19] 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.

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

[21] G. Tang, B. N. Bhaskar, and B. Recht, "Near minimax line spectral estimation," *IEEE Transactions on Information Theory*, vol. 61, no. 1, pp. 499–512, 2015.

[22] V. I. Morgenshtern and E. J. Candes, "Super-resolution of positive sources: The discrete setup," *SIAM Journal on Imaging Sciences*, vol. 9, no. 1, pp. 412–444, 2016.

[23] 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.

[24] V. Duval, "A characterization of the non-degenerate source condition in super-resolution," *Information and Inference: A Journal of the IMA*, Mar. 2019.

[25] Q. Denoyelle, V. Duval, and G. Peyré, "Support recovery for sparse super-resolution of positive measures," *Journal of Fourier Analysis and Applications*, vol. 23, no. 5, pp. 1153–1194, 2017.