Support detection in super-resolution

Carlos Fernandez-Granda
Department of Electrical Engineering, Stanford University
Stanford, CA, USA
Email: cfgranda@stanford.edu

Abstract—We study the problem of super-resolving a superposition of point sources from noisy low-pass data with a cut-off frequency f_c. Solving a tractable convex program is shown to locate the elements of the support with high precision as long as they are separated by $2/f_c$ and the noise level is small with respect to the amplitude of the signal.

I. Introduction

The problem of super-resolution is of great importance in applications where the measuring process imposes a physical limit on the resolution of the available measurements. It is often the case that the signal of interest is well modeled as a superposition of point sources. Motivated by this, we consider a signal

$$x = \sum_j a_j \delta_{t_j}, \qquad (I.1)$$

consisting of a train of Dirac measures with complex amplitudes a_j located at different locations $\{t_j\}$ in the unit interval. Our aim is to estimate x from the lower end of its spectrum in the form of $n = 2f_c + 1$ Fourier series coefficients (f_c is an integer) perturbed by noise,

$$\begin{aligned} y(k) &= \int_0^1 e^{-i2\pi kt} x(\mathrm{d}t) + z(k) \\ &= \sum_j a_j e^{-i2\pi k t_j} + z(k), \end{aligned} \qquad (I.2)$$

for $k \in \mathbb{Z}$, $|k| \le f_c$. To ease notation, we write (I.2) as $y = \mathcal{F}_n x + z$. We model the perturbation $z \in \mathbb{C}^n$ as having bounded ℓ_2 norm,

$$||z||_2 \le \delta. \qquad (I.3)$$

The noise is otherwise arbitrary and can be adversarial.

Even if the signal x is very sparse, without further conditions to ensure that the support of x is not too clustered the super-resolution problem is hopelessly ill-posed. This can be checked numerically, but also formalized thanks to the seminal work of Slepian [4] on discrete prolate spheroidal sequences (see Section 3.2 of [3]). To avoid such extreme ill-posedness, we impose a lower bound on the minimum separation between the elements of the support of the signal.

Definition 1.1 (Minimum separation): Let $\mathbb{T}$ be the circle obtained by identifying the endpoints on $[0, 1]$.

For a family of points $T \subset \mathbb{T}$, the minimum separation is the closest distance between any two elements of T,

$$\Delta(T) = \inf_{(t,t') \in T \,:\, t \ne t'} |t - t'|. \qquad (I.4)$$

To recover x we propose minimizing the total variation of the estimate, a continuous analog to the ℓ_1 norm for discrete signals (see Appendix A in [3] for a rigorous definition), subject to data constraints:

$$\min_{\tilde{x}} ||\tilde{x}||_{\mathrm{TV}} \quad \text{subject to} \quad ||\mathcal{F}_n \tilde{x} - y||_2 \le \delta, \qquad (I.5)$$

where the minimization is carried out over the set of all finite complex measures $\tilde{x}$ supported on $[0, 1]$. For details on how to solve (I.5) using semidefinite programming see [2], [3].

Previous work established that if

$$\Delta(T) \ge \frac{2}{f_c} := 2\lambda_c \qquad (I.6)$$

TV-norm minimization achieves exact recovery in a noiseless setting [3]. Additionally, [2] characterized the reconstruction error for noisy measurements as the target resolution increases. In this work we study support detection using this method. If the original signal contains a spike of a certain amplitude we ask: *How accurately can we recover the position of the spike? How does the accuracy depend on the noise level, the amplitude of the spike and the amplitude of the signal at other locations?* These questions are not addressed by previous work and answering them requires non-trivial modifications to the arguments in [2] and [3]. Our main result establishes that convex programming is in fact a powerful method for support detection in super-resolution.

Theorem 1.2: Consider the noise model (I.3) *and assume the support T satisfies the minimum-separation condition* (I.6). *The solution to problem* (I.5)[1]

$$\hat{x} = \sum_{\hat{t}_k \in \widehat{T}} \hat{a}_k \delta_{\hat{t}_k}$$

---

[1]This solution can be shown to be an atomic measure with discrete support under very general conditions.

with support $\widehat{T}$ obeys the properties

(i): $\left| a_j - \sum\limits_{\left\{ \hat{t}_l \in \widehat{T}: \left| \hat{t}_l - t_j \right| \leq c\lambda_c \right\}} \hat{a}_l \right| \leq C_1 \delta \quad \forall t_j \in T,$

(ii): $\sum\limits_{\left\{ \hat{t}_l \in \widehat{T}, t_j \in T: \left| \hat{t}_l - t_j \right| \leq c\lambda_c \right\}} |\hat{a}_l| \left(\hat{t}_l - t_j \right)^2 \leq C_2 \lambda_c^2 \delta,$

(iii): $\sum\limits_{\left\{ \hat{t}_l \in \widehat{T}: \left| \hat{t}_l - t_j \right| > c\lambda_c \, \forall t_j \in T \right\}} |\hat{a}_l| \leq C_3 \delta,$

where C_1, C_2 and C_3 are positive numerical constants and $c = 0.1649$.

Properties (i) and (ii) show that the estimate clusters tightly around each element of the signal, whereas (iii) ensures that any spurious spikes detected by the algorithm have small amplitude. These bounds are essentially optimal for the case of adversarial noise, which can be highly concentrated. An intriguing consequence of our result is a bound on the support-detection error for a single spike that does not depend on the value of the signal at other locations.

Corollary 1.3: Under the conditions of Theorem 1.2, for any element t_i in the support of x such that $a_i > C_1 \delta$ there exists an element $\hat{t}_i$ in the support of the estimate $\hat{x}$ satisfying

$$\left| t_i - \hat{t}_i \right| \leq \sqrt{\frac{C_2 \delta}{|a_i| - C_1 \delta}} \lambda_c.$$

Despite the aliasing effect of the low-pass filter applied to the signal, the bound on the support-detection error *only depends on the amplitude of the corresponding spike* (and the noise level). This does not follow from previous analysis. In particular, the bound on the weighted $\mathcal{L}_1$ norm of the error derived in [2] does not allow to derive such local guarantees. A recent paper bounds the support-detection error of a related convex program in the presence of stochastic noise, but the bound depends on the amplitude of the solution rather than on the amplitude of the original spike [1]. As we explain below, obtaining detection guarantees that only depend on the amplitude of the spike of interest is made possible by the existence of a certain low-frequency polynomial, constructed in Lemma 2.2. This is the main technical contribution of the paper.

II. PROOF OF THEOREM 1.2

We begin with an intermediate result proved in Section II-A.

Lemma 2.1: Under the assumptions of Theorem 1.2

$$\sum_{\hat{t}_k \in \hat{T}} |\hat{a}_k| \min \left\{ C_a, \frac{C_b \, d \left(\hat{t}_k, T \right)}{\lambda_c^2} \right\} \leq 2\delta,$$

where C_a and C_b are positive numerical constants and

$$d \left(t, T \right) := \min_{t_i \in T} \left(t - t_i \right)^2.$$

Properties (ii) and (iii) are direct corollaries of Lemma (2.1). To establish property (i) we need the following key lemma, proved in Section II-B.

Lemma 2.2: Suppose T obeys condition (I.6) *and $f_c \geq 10$. Then for any $t_j \in T$ there exists a low-pass polynomial*

$$q_{t_j}(t) = \sum_{k=-f_c}^{f_c} b_k e^{i2\pi kt},$$

$b \in \mathbb{C}^n$, *such that $|q_{t_j}(t)| < 1$ for all $t \neq t_j$ and*

$$q_{t_j}(t_j) = 1, \qquad q_{t_j}(t_l) = 0 \quad t_l \in T \setminus \{t_j\},$$

$$|1 - q_{t_j}(t)| \leq \frac{C_1' \left(t - t_j \right)^2}{\lambda_c^2} \quad for \quad |t - t_j| \leq c\lambda_c,$$
$$\tag{II.1}$$

$$|q_{t_j}(t)| \leq \frac{C_1' \left(t - t_l \right)^2}{\lambda_c^2} \quad for \quad t_l \in T \setminus \{t_j\}, \ |t - t_l| \leq c\lambda_c,$$
$$\tag{II.2}$$

$$|q_{t_j}(t)| < C_2' \quad if \quad |t - t_l| > c\lambda_c \ \forall \ t_l \in T, \tag{II.3}$$

where $0 < c^2 C_2' \leq C_1' < 1$.

The polynomial q_{t_j} provided by this lemma is designed to satisfy $\int_{\mathbb{T}} q_{t_j}(t) x(\mathrm{d}t) = a_j$ and vanish on the rest of the support of the signal. This allows to decouple the estimation error at t_j from the amplitude of the rest of the spikes. Since x and $\hat{x}$ are feasible for (I.5), we can apply Parseval's Theorem and the Cauchy-Schwarz inequality to obtain

$$\left| \int_{\mathbb{T}} q_{t_j}(t) x(\mathrm{d}t) - \int_{\mathbb{T}} q_{t_j}(t) \hat{x}(\mathrm{d}t) \right|$$
$$= \left| \sum_{k=-f_c}^{f_c} b_k \mathcal{F}_n(x - \hat{x})_k \right|$$
$$\leq \left\| q_{t_j} \right\|_{\mathcal{L}_2} \left\| \mathcal{F}_n(x - \hat{x}) \right\|_2 \leq 2\delta, \tag{II.4}$$

where we have used that the absolute value and consequently the $\mathcal{L}_2$ norm of q_{t_j} is bounded by one. In addition, by Lemmas 2.2 and 2.1 we have

$$\left| \sum_{\left\{ k: \left| \hat{t}_k - t_j \right| > c\lambda_c \right\}} \hat{a}_k q_{t_j}(\hat{t}_k) + \sum_{\left\{ k: \left| \hat{t}_k - t_j \right| \leq c\lambda_c \right\}} \hat{a}_k \left(q_{t_j}(\hat{t}_k) - 1 \right) \right|$$

$$\leq \sum_{\left\{ k: \left| \hat{t}_k - t_j \right| > c\lambda_c \right\}} |\hat{a}_k| \left| q_{t_j}(\hat{t}_k) \right|$$

$$+ \sum_{\left\{ k: \left| \hat{t}_k - t_j \right| \leq c\lambda_c \right\}} |\hat{a}_k| \left| 1 - q_{t_j}(\hat{t}_k) \right|$$

$$\leq \sum_{\hat{t}_k \in \hat{T}} |\hat{a}_k| \min \left\{ C_2', \frac{C_1' d \left(\hat{t}_k, T \right)}{\lambda_c^2} \right\} \leq C\delta, \quad \text{(II.5)}$$

for a positive numerical constant C. Finally, Lemma 2.2, the triangle inequality, (II.4) and (II.5)

yield

$$
\left| a_j - \sum_{\{k:\, |\hat{t}_k - t_j| \le c\lambda_c\}} \hat{a}_k \right|
$$

$$
= \left| \int_{\mathbb{T}} q_{t_j}(t) x(\mathrm{d}t) - \int_{\mathbb{T}} q_{t_j}(t)\hat{x}(\mathrm{d}t) \right.
$$

$$
+ \sum_{\{k:\, |\hat{t}_k - t_j| > c\lambda_c\}} \hat{a}_k q_{t_j}(\hat{t}_k)
$$

$$
\left. + \sum_{\{k:\, |\hat{t}_k - t_j| \le c\lambda_c\}} \hat{a}_k \left(q_{t_j}(\hat{t}_k) - 1 \right) \right| \le C'\delta.
$$

for a positive numerical constant C'.

A. Proof of Lemma 2.1

The proof relies on a low-pass polynomial provided by Proposition 2.1 and Lemma 2.5 in [3].

Lemma 2.3: Let T obey (I.6). *For any $v \in \mathbb{C}^{|T|}$ such that $|v_j| = 1$ for all entries v_j there exists a low-pass polynomial $q(t) = \sum_{k=-f_c}^{f_c} d_k e^{i2\pi kt}$, $d \in \mathbb{C}^n$, satisfying*

$$
q(t_j) = v_j, \qquad t_j \in T,
$$
$$
|q(t)| < 1 - C_a, \qquad |t - t_j| > c\lambda_c \quad \forall t_j \in T,
$$
$$
|q(t)| \le 1 - \frac{C_b (t - t_j)^2}{\lambda_c^2}, \qquad |t - t_j| \le c\lambda_c,\, t_j \in T,
$$

with $0 < c^2 C_b \le C_a < 1$.

We set $v_j = \overline{a_j}/|a_j|$. The lemma implies

$$
\int_{\mathbb{T}} q(t)\hat{x}\,(\mathrm{d}t) \le \sum_k |\hat{a}_k|\,|q(\hat{t}_k)|
$$
$$
\le \sum_k \left(1 - \min\left\{ C_a, \frac{C_b\, d\left(\hat{t}_k, T\right)}{\lambda_c^2} \right\} \right) |\hat{a}_k|. \tag{II.6}
$$

The same argument used to prove (II.4) yields

$$
\left| \int_{\mathbb{T}} q(t)\hat{x}(\mathrm{d}t) - \int_{\mathbb{T}} q(t)x(\mathrm{d}t) \right| \le 2\delta.
$$

Now, taking into account that $\int_{\mathbb{T}} q(t)x\,(\mathrm{d}t) = \|x\|_{\mathrm{TV}}$ by construction and $\|\hat{x}\|_{\mathrm{TV}} \le \|x\|_{\mathrm{TV}}$, we have

$$
\int_{\mathbb{T}} q(t)\hat{x}\,(\mathrm{d}t)
$$
$$
= \int_{\mathbb{T}} q(t)x\,(\mathrm{d}t) + \int_{\mathbb{T}} q(t)\hat{x}(\mathrm{d}t) - \int_{\mathbb{T}} q(t)x(\mathrm{d}t)
$$
$$
\ge \|x\|_{\mathrm{TV}} - 2\delta \ge \|\hat{x}\|_{\mathrm{TV}} - 2\delta = \sum_k |\hat{a}_k| - 2\delta.
$$

Combining this with (II.6) completes the proof.

B. Proof of Lemma 2.2

We use a low-frequency kernel and its derivative to construct the desired polynomial exploiting the assumption that the support satisfies the minimum-separation condition (I.6). More precisely, we set

$$
q_{t_j}(t) = \sum_{t_k \in T} \alpha_k K(t - t_k) + \beta_k K^{(1)}(t - t_k), \tag{II.7}
$$

where $\alpha, \beta \in \mathbb{C}^{|T|}$ are coefficient vectors,

$$
K(t) = \left[\frac{\sin\left(\left(\frac{f_c}{2} + 1 \right) \pi t \right)}{\left(\frac{f_c}{2} + 1 \right) \sin(\pi t)} \right]^4, \quad t \in \mathbb{T} \setminus \{0\},
$$
$$
\tag{II.8}
$$

and $K(0) = 1$; here, $K^{(\ell)}$ is the ℓth derivative of K. Note that K, $K^{(1)}$ and, consequently, q_{t_j} are trigonometric polynomials of the required degree.

We impose $q_{t_j}(t_j) = 1$ and

$$
q_{t_j}(t_l) = 0, \quad t_l \in T/\{t_j\}, \qquad q'_{t_j}(t_k) = 0, \quad t_k \in T.
$$

We express these constraints in matrix form. Let $e_{t_j} \in \mathbb{R}^{|T|}$ denote the one-sparse vector with one nonzero entry at the position corresponding to t_j. Then,

$$
\begin{bmatrix} D_0 & D_1 \\ D_1 & D_2 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} e_{t_j} \\ 0 \end{bmatrix}, \quad \text{where}
$$

$$
(D_0)_{lk} = K(t_l - t_k), \quad (D_1)_{lk} = K^{(1)}(t_l - t_k),
$$
$$
(D_2)_{lk} = K^{(2)}(t_l - t_k),
$$

and l and k range from 1 to $|T|$. It is shown in Section 2.3.1 of [3] that under the minimum-separation condition this system is invertible. As a result α and β are well defined and q_{t_j} satisfies $q_{t_j}(t_j) = 1$ and $q_{t_j}(t_l) = 0$ for $t_l \in T/\{t_j\}$. The coefficient vectors can be expressed as

$$
\begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} I \\ -D_2^{-1}D_1 \end{bmatrix} S^{-1} e_{t_j}, \quad S := D_0 - D_1 D_2^{-1} D_1,
$$

where S is the Schur complement. Let $\|M\|_\infty$ denote the usual infinity norm of a matrix M defined as $\|M\|_\infty = \max_{\|x\|_\infty = 1} \|Mx\|_\infty = \max_l \sum_k |M_{lk}|$. We borrow some results from Section 2.3.1 in [3],

$$
\|I - S\|_\infty \le 8.747\,10^{-3},
$$
$$
\|S^{-1}\|_\infty \le 1 + 8.824\,10^{-3},
$$
$$
\|I - S^{-1}\|_\infty \le \|S^{-1}\|_\infty \|S - I\|_\infty \le 8.825\,10^{-3},
$$
$$
\|\alpha - e_{t_j}\|_\infty \le \|I - S^{-1}\|_\infty \|e_{t_j}\|_\infty
$$
$$
\le 8.825\,10^{-3}, \tag{II.9}
$$
$$
\|\beta\|_\infty \le 3.294\,10^{-2}\lambda_c. \tag{II.10}
$$

Lemma 2.6 in [3] allows to obtain

$$
K(t) \le \frac{1}{(f_c t)^4} \le 0.333, \quad K'(t) \le \frac{4\pi}{f_c^3 t^4} \le 4.18\, f_c,
$$

for $|t| > c\lambda_c$ as long $f_c \geq 10$. By the same lemma, if we set the minimum separation $\Delta_{\min}$ to $2/f_c$

$$\sum_{t_k \in T \setminus \{t_a, t_b\}} |K(t - t_k)|$$

$$\leq \sum_{l=0}^{\infty} \frac{1}{\left(f_c \Delta_{\min}(\frac{1}{2} + l)\right)^4} + \sum_{l=0}^{\infty} \frac{1}{(f_c \Delta_{\min} l)^4} \leq 1.083,$$

$$\sum_{t_k \in T \setminus \{t_a, t_b\}} \left|K^{(1)}(t - t_k)\right|$$

$$\leq \sum_{l=0}^{\infty} \frac{4\pi}{f_c^3 \left(\Delta_{\min}(\frac{1}{2} + l)\right)^4} + \sum_{j=0}^{\infty} \frac{4\pi}{f_c^3 (\Delta_{\min} l)^4} \leq 1.75 \, f_c,$$

where t_a and t_b are the two spikes nearest to t. Let t_i be the element of $T/\{t_j\}$ that is nearest to t. Combining these inequalities with (II.9) and (II.10) proves that

$$|q_{t_j}(t)| = \left| \sum_{t_k \in T} \alpha_k K(t - t_k) + \sum_{t_k \in T} \beta_k K^{(1)}(t - t_k) \right|$$

$$\leq |K(t - t_j)| + ||\alpha - e_{t_j}||_\infty \left(|K(t - t_j)| \right.$$

$$+ |K(t - t_i)| + \sum_{t_k \in T \setminus \{t_i, t_j\}} |K(t - t_k)| \Big)$$

$$+ ||\beta||_\infty \left(\left|K^{(1)}(t - t_j)\right| + \left|K^{(1)}(t - t_i)\right| \right.$$

$$+ \sum_{t_k \in T \setminus \{t_i, t_j\}} \left|K^{(1)}(t - t_k)\right| \Big) \leq 0.69,$$

if $|t - t_k| > c\lambda_c$ for all $t_k \in T$ so that (II.3) holds. The proof is completed by two lemmas which prove (II.1) and (II.2) and $|q_{t_j}(t)| < 1$ for any t. They rely on the following bounds borrowed from equation (2.25) in Section 2 of [3],

$$K(t) \geq 0.9539, \qquad K^{(2)}(t) \leq -2.923 \, f_c^2,$$

$$\left|K^{(1)}(t)\right| \leq 0.5595 \, f_c, \qquad \left|K^{(2)}(t)\right| \leq 3.393 \, f_c^2,$$

$$\left|K^{(3)}(t)\right| \leq 5.697 \, f_c^3,$$

(II.11)

and on the fact that, due to Lemma 2.7 in [3], for any $t_0 \in T$ and $t \in \mathbb{T}$ obeying $|t - t_0| \leq c\lambda_c$,

$$\sum_{t_k \in T \setminus \{t_0\}} \left|K^{(2)}(t - t_k)\right| \leq 1.06 \, f_c^2 \qquad \text{(II.12)}$$

$$\sum_{t_k \in T \setminus \{t_0\}} \left|K^{(3)}(t - t_k)\right| \leq 18.6 \, f_c^3. \qquad \text{(II.13)}$$

Lemma 2.4: For any t such that $|t - t_j| \leq c\lambda_c$,

$$1 - 4.07 (t - t_j)^2 f_c^2 \leq q_{t_j}(t) \leq 1 - 2.30 (t - t_j)^2 f_c^2.$$

Proof We assume without loss of generality that $t_j = 0$. By symmetry, it suffices to show the claim for $t \in (0, c\lambda_c]$. By (II.9), (II.10), (II.11), (II.12) and (II.13),

$$q_0''(t) = \sum_{t_k \in T} \alpha_k K^{(2)}(t - t_k) + \sum_{t_k \in T} \beta_k K^{(3)}(t - t_k)$$

$$\leq \left(1 + ||\alpha - e_{t_j}||_\infty\right) K^{(2)}(t)$$

$$+ ||\alpha - e_{t_j}||_\infty \sum_{t_k \in T \setminus \{0\}} \left|K^{(2)}(t - t_k)\right|$$

$$+ ||\beta||_\infty \left(\left|K^{(3)}(t)\right| + \sum_{t_k \in T \setminus \{0\}} \left|K^{(3)}(t - t_k)\right| \right)$$

$$\leq -2.30 \, f_c^2.$$

Similar computations yield $|q_0''(t)| \leq 4.07 \, f_c^2$. This together with $q_0(0) = 1$ and $q_0'(0) = 0$ implies the desired result. ∎

Lemma 2.5: For any $t_l \in T \setminus \{t_j\}$ and t obeying $|t - t_l| \leq c\lambda_c$, we have

$$|q_{t_j}(t)| \leq 16.64 (t - t_l)^2 f_c^2.$$

Proof We assume without loss of generality that $t_l = 0$ and prove the claim for $t \in (0, c\lambda_c]$. By (II.9), (II.10), (II.11), (II.12) and (II.13)

$$\left|q_{t_j}''(t)\right| = \left| \sum_{t_k \in T} \alpha_k K^{(2)}(t - t_k) + \sum_{t_k \in T} \beta_k K^{(3)}(t - t_k) \right|$$

$$\leq \left(1 + ||\alpha - e_{t_j}||_\infty\right) \left|K^{(2)}(t - t_j)\right|$$

$$+ ||\alpha - e_{t_j}||_\infty \left(\left|K^{(2)}(t)\right| + \sum_{t_k \in T \setminus \{0, t_j\}} \left|K^{(2)}(t - t_k)\right| \right)$$

$$+ ||\beta||_\infty \left(\left|K^{(3)}(t)\right| + \sum_{t_k \in T \setminus \{0\}} \left|K^{(3)}(t - t_k)\right| \right)$$

$$\leq 16.64 \, f_c^2,$$

since in the interval of interest $\left|K^{(2)}(t - t_j)\right| \leq \frac{18\pi^2}{f_c^2 (\Delta_{\min} - 0.16 f_c)^4} \leq 15.67 \, f_c^2$ due to Lemma 2.6 in [3]. This together with $q_{t_j}(0) = 0$ and $q_{t_j}'(0) = 0$ implies the desired result. ∎

REFERENCES

[1] J. M. Azais, Y. de Castro and F. Gamboa. Spike detection from innacurate samplings. Preprint.
[2] E. J. Candès and C. Fernandez-Granda. Super-resolution from noisy data. Preprint.
[3] E. J. Candès and C. Fernandez-Granda. Towards a mathematical theory of super-resolution. *Communications on Pure and Applied Mathematics*, to appear.
[4] D. Slepian. Prolate spheroidal wave functions, Fourier analysis, and uncertainty. V - The discrete case. *Bell System Technical Journal*, 57:1371–1430, 1978.