A Guaranteed Poly-Logarithmic Time Relaxation for the Line Spectral Estimation Problem

Maxime Ferreira Da Costa and Wei Dai
Department of Electrical and Electronic Engineering, Imperial College London, United Kingdom
Email: {maxime.ferreira, wei.dai1}@imperial.ac.uk

I. Background on line spectral estimation

The line spectral estimation problem aims to recover the frequencies of a complex time signal x that is assumed to be sparse in the spectral domain from its discrete measurements $y \in \mathbb{C}^n$, uniformly acquired at a sampling frequency $f_S \in \mathbb{R}^+$. More precisely, the time signal x is assumed to follow the s-spikes model given by

$$\forall t \in \mathbb{R}, \quad x(t) = \sum_{r=1}^{s} \alpha_r e^{i2\pi\xi_r t}, \tag{1}$$

whereby $\Xi = \{\xi_r\}_{1 \leq r \leq s}$ is the ordered set containing the s spectral components generating the signal x, and $\alpha = \{\alpha_r\}_{1 \leq r \leq s}$ the one of their associated complex amplitudes. The particularly of this model stands in the fact that the frequencies Ξ are drawn *continuously* on $[0, f_S)$ and *are not constrained belong to some finite discrete grid,* as opposed to discretization-based methods to tackle inverse problems.

This problem is ill-posed and there are infinitely many estimators of the spectral distribution $\hat{x}$ of x that are consistent with the measurement vector y. Among all those estimators, the one considered to be optimal in this spikes recovery context is the one returning a consistent spectral distribution $\hat{x}_0$ of $\hat{x}$ having the sparsest possible spectral support. Equivalently, this estimator can be defined by the output of the minimization program

$$\begin{aligned} \hat{x}_0 = \arg \min_{\hat{x} \in D_1} &\|\hat{x}\|_0 \\ \text{subject to} \quad & y = \mathcal{F}_n(\hat{x}), \end{aligned} \tag{2}$$

where D_1 denotes the space of absolutely integrable spectral distributions. The functions $\|\cdot\|_0$ and $\mathcal{F}_n(\cdot)$ are respectively the support counting pseudo-norm and the inverse discrete time Fourier transform whose expressions are given in Table I.

Program (2) is non-convex and difficult to solve in a direct approach due to the combinatorial nature of "L_0" minimization. A commonly proposed workaround consists in analysing the output of a *convex relaxation* of (2), obtained by swapping the cardinality cost function $\|\cdot\|_0$ into a minimization of the total-variation norm $|\cdot|(\mathbb{T}_{f_S})$ defined in Table I over D_1. This relaxation was proven to be tight in [1] and robust to noise in [2], provided that a sufficient separability criterion

$$\Delta_{\mathbb{T}_{f_S}}(\Xi) \geq \frac{4f_S}{n-1} \tag{3}$$

is respected, where $\Delta_{\mathbb{T}_{f_S}}(\cdot)$ is the minimal warp around distance on the rescaled elementary torus $\mathbb{T}_{f_S} = [0, f_S)$ between elements of a set defined in Table I. This bound was tightened later on in [3].

More interestingly, it has been shown in [4] that the tightness of the convex approach still holds with high probability when extracting independently at random a small number of observations and discarding the rest of it. The observation vector $y \in \mathbb{C}^m$ resulting from this random process is linked to the spectrum $\hat{x}$ of the probed signal by the linear relation $y = C_{\mathcal{I}}\mathcal{F}_n(\hat{x})$ where $C_{\mathcal{I}} \in \{0,1\}^{m \times n}$ is a boolean matrix whose rows are equal to $\{e_k^{\mathsf{T}}\}_{k \in \mathcal{I}}$ and where $\mathcal{I} \subseteq [\![0, n-1]\!]$ is the subset of cardinality m describing the indexes of the retained samples. In addition, it has been shown that the dual Lagrange program takes the form of the semidefinite program

$$\begin{aligned} (c_\star, H_\star) = \arg \max_{\substack{c \in \mathbb{C}^m \\ H \in \mathbb{C}^{n \times n}}} & \Re\left(y^{\mathsf{T}} c\right) \\ \text{subject to} \quad & \begin{bmatrix} H & q \\ q^* & 1 \end{bmatrix} \succeq 0 \\ & \mathcal{T}_n^*(H) = e_0 \\ & q = C_{\mathcal{I}}^* c, \end{aligned} \tag{4}$$

where $\mathcal{T}_n^*$ is the adjoint of the linear operator $\mathcal{T}_n$ and $\mathcal{T}_n(u)$ is the Toeplitz Hermitian matrix whose first row is equal to u for all $u \in \mathbb{C}^n$. Moreover, the polynomial of degree $n-1$ having for coefficients vector $q_\star = C_{\mathcal{I}}^* c_\star$ locates with high probability the frequencies supporting $\hat{x}_0$ around the unit circle.

II. Main contribution

The semidefinite program (4) remains of dimension n, which can be much greater than the number of observation m. Its output is computable in $\mathcal{O}(n^7)$ operations via the use of interior point solvers, which become intractable when n exceeds a few hundred. Our result complements the tightness guarantees of [4] by showing the existence of a semidefinite program of dimension m recovering the spectral support of $\hat{x}_0$ with high probability. Moreover since $m > \mathcal{O}(\log^2 n)$ has been guaranteed in [4] to produce a tight estimate of the spectral support, our program is computable in a *poly-logarithmic time* of the variable n. Our results are summarized by the following theorem, and relies on a novel extension of the theory of Gram parametrization of trigonometric polynomials to subspaces of polynomials [5].

Theorem 1. *Let $\mathcal{I}$ be a subset of cardinality m drawn uniformly at random in $[\![0, n-1]\!]$, and let $\mathcal{R}_{\mathcal{I}}$ the linear operator defined by $\mathcal{R}_{\mathcal{I}}(u) = C_{\mathcal{I}}\mathcal{T}_n(u)\,C_{\mathcal{I}}^*$ for all $u \in \mathbb{C}^n$. Suppose that x follows Model (1) and satisfies Condition (3). Moreover, suppose that the elements of α have phases drawn independently and uniformly at random in $[0, 2\pi)$. Consider any positive number $\delta > 0$. There exists a constant $C > 0$ such that if*

$$m \geq C \max\left\{\log^2 \frac{n}{\delta}, s \log \frac{s}{\delta} \log \frac{n}{\delta}\right\},$$

then the semidefinite program

$$\begin{aligned} (c_\star, S_\star) = \arg \max_{\substack{c \in \mathbb{C}^m \\ S \in \mathbb{C}^{m \times m}}} & \Re\left(y^{\mathsf{T}} c\right) \\ \textit{subject to} \quad & \begin{bmatrix} S & c \\ c^* & 1 \end{bmatrix} \succeq 0 \\ & \mathcal{R}_{\mathcal{I}}^*(S) = e_0 \end{aligned} \tag{5}$$

outputs with probability greater than $1-\delta$ a vector $c_\star \in \mathbb{C}^m$ such that $q_\star = C_{\mathcal{I}}^* c_\star$ induces a polynomial $Q_\star$ of degree $n-1$ locating the support of $\hat{x}_0$. Moreover, this program can be solved in $\mathcal{O}(m^3)$ operations via the alternating direction method of multipliers.

REFERENCES

[1] 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.

[2] ——, "Super-resolution from noisy data," *Journal of Fourier Analysis and Applications*, vol. 19, no. 6, pp. 1229–1254, 2013.

[3] C. Fernandez-Granda, "Super-resolution of point sources via convex programming," *Information and Inference*, vol. 5, no. 3, pp. 251–303, Sep. 2016.

[4] G. Tang, B. N. Bhaskar, P. Shah, and B. Recht, "Compressed sensing off the grid," *IEEE Transactions on Information Theory*, vol. 59, no. 11, pp. 7465–7490, Nov. 2013.

[5] M. Ferreira Da Costa and W. Dai, "Exact Dimensionality Reduction for Partial Line Spectra Estimation Problems," *arXiv:1609.03142*.

[6] 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, Jan. 2015.

[7] B. Dumitrescu, *Positive Trigonometric Polynomials and Signal Processing Applications*. Springer, 2010.

Function	Domain	Expression
$\mathcal{F}_n$	$D_1 \to \mathbb{C}^n$	$\forall k \in [\![0, n-1]\!]\,,\ \mathcal{F}_n(\hat{x})[k] = \int_{\mathbb{T}_{f_S}} e^{i2\pi\xi k} \mathrm{d}\hat{x}(\xi)$
$\Delta_{\mathbb{T}_{f_S}}$	$\wp(\mathbb{T}_{f_S}) \to \mathbb{R}^+$	$\Delta_{\mathbb{T}_{f_S}}(\Omega) = \inf_{(\xi,\xi')\in\Omega^2}\{\lvert\xi-\xi'\rvert : \xi \neq \xi'\}$
$\lVert\cdot\rVert_0$	$D_1 \to \mathbb{R}^+ \cup \{+\infty\}$	$\lVert\hat{x}\rVert_0 = \text{card}\{\hat{x}(\xi) \neq 0 :\ \xi \in \mathbb{T}_{f_S}\}$
$\lvert\cdot\rvert(\mathbb{T}_{f_S})$	$D_1 \to \mathbb{R}^+$	$\lvert\hat{x}\rvert(\mathbb{T}_{f_S}) = \sup_{h\in\mathcal{C}(\mathbb{T}_{f_S})}\left\{\Re\left[\int_{\mathbb{T}_{f_S}} \overline{h(\xi)}\mathrm{d}\hat{x}(\xi)\right] :\ \lVert h\rVert_\infty \leq 1\right\}$

Table I
MATHEMATICAL DEFINITIONS

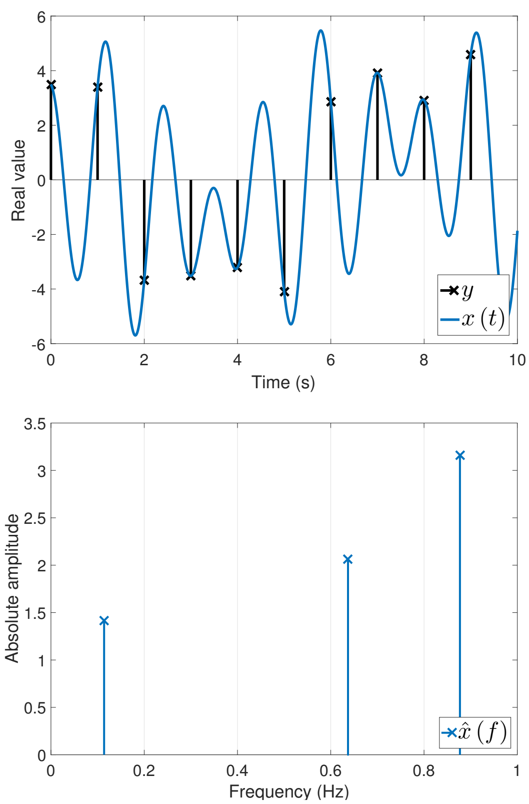

Figure 1. Time and spectral representation of a signal x following the spikes model with three spectral spikes and its measurement vector y when taking $n = 10$ observations at a frequency $f_S = 1\text{Hz}$.

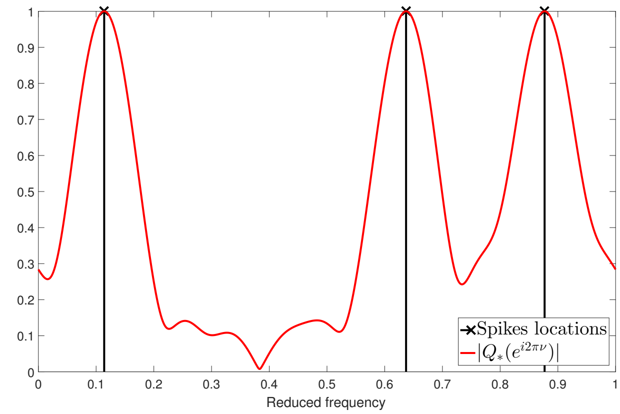

Figure 2. The optimal dual polynomial $Q_\star$ obtained by solving Program (5) when retaining entries of y with indexes in the set $\mathcal{I} = \{0, 3, 4, 6, 8, 9\}$ of cardinality $m = 6$. $Q_\star\left(e^{i2\pi\nu}\right)$ locates the frequencies of the signal x by reaching modulus 1 whenever $\nu \in \frac{1}{f_S}\Xi$.