Super-Resolution of Complex Exponentials from Modulations with Known Waveforms

Zhihui Zhu[1], Manuel Lopez-Santillana[2], Michael B. Wakin[1]
[1]Department of Electrical Engineering, Colorado School of Mines, Golden, CO, 80401, USA
{zzhu, mwakin}@mines.edu
[2]Department of Applied Mathematics and Statistics, Colorado School of Mines, Golden, CO, 80401, USA
mlopezsa@mines.edu

***Abstract*—In this paper, we investigate parametric estimation of complex exponentials from modulations with known waveforms. This problem arises naturally in radar systems and wireless communications, especially in applications which suffer from multipath effects. Viewing the problem as a group sparse recovery, we recast it naturally into an atomic norm minimization, which has an equivalent semidefinite program (SDP) characterization and thus can be solved efficiently. We experimentally demonstrate the advantage of our approach when compared with a super-resolution method that does not consider multipath effects.**

***Keywords*—Atomic norm, super-resolution, group sparsity, multipath exploitation, semidefinite programming**

I. Introduction

We consider an acquired signal consisting of a superposition of the same point source convolved with different point spread functions. Mathematically, we consider the following parametric superpositon model for the acquired signal

$$x(t) = \sum_{l=1}^{L} (h(t)\xi_l(t)) \star g_l(t) \tag{1}$$

where $\star$ denotes the circular convolution operator, the signal of interest $h(t) = \sum_{k=1}^{K} \widetilde{\sigma}_k \delta(t - t_k)$ is a weighted superposition of spikes, $\xi_l(t)$ is an unknown (nonlinear) function which acts as an amplifier or attenuator, and $g_l(t)$ is a known point spread function. Our goal is to identify the spike locations $\{t_k\}$ and coefficients $\{\widetilde{\sigma}_k\}$ from low-frequency measurements of $x(t)$ (which are described formally in Section II). This model arises in applications such as radar imaging, DNA microarrays, and spike detection. A particular area that motivates this work is multipath exploitation in radar systems [1] and wireless communication [2]. We list two stylized applications below.

Multipath exploitation for urban radar imaging: The indirect multipath reflections of electromagnetic waves off of targets (in conjunction with the walls, floors, etc.) present a challenge in synthetic aperture radar (SAR) imaging [1, 3–6]. They result in ghost artifacts that can clutter the reconstructed image. Let $\widetilde{\sigma}_k$ and t_k respectively represent the complex reflectivity and locations of the k-th target. In this case, $l = 1$ corresponds to the direct path and $\xi_1(t) = 1, g_1(t) = \delta(t)$, i.e, there is no attenuator and convolution. The point spread function $g_l(t)$ corresponding to the l-th ($l \geq 2$) path is $\delta(t - \tau_l)$ where τ_l is the additive two-way travel time of the l-th path compared to the direct path. With prior information about the room geometry, the point spread function $g_l(t)$ (or τ_l) can be parametrized by the front and interior walls [1, 3]. The attenuator $\xi_l(t)$ is determined by the material of the wall and other factors and is usually modeled as unknown [1].

Spike detection for neural recording: Neuron spikes (action potentials) are often captured with a microelectrode tip which is surrounded by many neurons, and therefore, receives a mixture of neurons' electrical activities [7]. The neural recording—obtained through wireless neural recording systems—can be modeled as a superposition of returns due to radio signal reflection from surrounding materials [8], as in (1).

When there is only a single component with $L = 1$, (1) reduces to

$$x(t) = (h(t)\xi_1(t)) \star g_l(t).$$

Without loss of generality, suppose $\xi_1(t) = 1, g_1(t) = \delta(t)$. Otherwise we can deconvolve $x(t)$ with $g_1(t)$ and $\xi_1(t)$ can be absorbed into $h(t)$. Identifying the parameters in (1) from low-frequency measurements reduces to the super-resolution problem or line spectrum estimation (if we exchange the time and frequency domains), where one can apply conventional approaches for parameter estimation such as matrix pencil [9] and MUSIC [10]. Chandrasekaran et al. [11] propose to use the atomic norm, induced by the convex hull of a set of atoms, as the general convex penalty function for linear inverse problems. The atomic norm generates the l_1 norm for sparse recovery problems and the nuclear norm for low rank matrix recovery. Importantly, it provides a powerful framework to handle a dictionary with an infinite number of atoms. For super-resolution or line spectrum estimation, the atomic norm minimization approach exactly inverts the parameters when there is no noise [12–15] and a certain minimum separation condition is met. Recently, Zhu et al. [16] investigated SAR imaging via sparse atomic norm reconstruction. Atomic norm minimization has also been utilized for modal analysis [17]. Chi [18] and Yang et al. [19] utilized atomic norm minimization to exactly recover parameters of complex exponentials from their modulations with unknown waveforms via a lifting trick. Li and Chi [20] applied atomic norm minimization to simultaneously identify multiple sets of spikes from a superposition of their modulations with known point spread functions. In [21–23], the authors solved blind deconvolution problems with nuclear norm minimization by assuming the signals live in known subspaces or are observed using random masks. Our problem differs from [21–23] in that the modulations are known but the target signal is parameterized (by t).

The approach for estimating the spikes using sparse modeling in [3] (where the spikes correspond to the target locations) is to first divide the range of t uniformly to reduce the continuous parameter space into a finite set of grid points; then construct a dictionary for each component based on these grid points; and finally recover the spikes via sparse recovery algorithms. This approach suffers from the basis mismatch problem and has no theoretical guarantees. In this paper, we apply atomic norm techniques to invert the parameters in (1) from low-frequency measurements. The atomic norm is utilized to promote group sparsity and has an equivalent SDP characterization. Thus the problem can be solved efficiently using an off-the-shelf solver [24]. The outline of this paper is as follows. In Section II, the main problem is illustrated. Our approach is discussed in Section III. Section IV presents some simulations to support our proposed methods.

II. Problem Setup

Suppose $t_k \in [0,1]$ and $x(t)$ and $g_l(t)$ are supported on $[0,1]$ for all $k = 1,2,\ldots,K$ and $l = 1,2,\ldots,L$. We rewrite (1) as

$$x(t) = \sum_{l=1}^{L} (h(t)\xi_l(t)) \star g_l(t) = \sum_{l=1}^{L}\left(\sum_{k=1}^{K}\sigma_{lk} g_l(t \ominus t_k)\right),$$

where $\sigma_{lk} = \widetilde{\sigma}_k \xi_l(t_k)$ and $\ominus$ denotes the subtraction operator on the unit circle $[0,1]$. Taking the Fourier series of $x(t)$, we obtain the measurements

$$\boldsymbol{x}[n] = \sum_{l=1}^{L}\left(\sum_{k=1}^{K}\sigma_{lk}\boldsymbol{g}_l[n]e^{-j2\pi t_k n}\right) \tag{2}$$

for $n = -2M, -2M+1, \ldots, 2M$. Thus $\boldsymbol{x}$ only contains low-frequency information about x since we only observe the $4M+1$ lowest Fourier series coefficients. Denote $N = 4M+1$. Here $\boldsymbol{g}_l \in \mathbb{C}^N$ contains the lowest N Fourier series coefficients of $g_l(t)$ with elements

$$\boldsymbol{g}_l[n] = \int_0^1 g_l(t)e^{-j2\pi tn}dt$$

for $n = -2M, -2M+1, \ldots, 2M$.

We use $\Omega = \{t_1, \ldots, t_K\}$ to denote the unknown set of spike locations. Our goal is to estimate the spike locations Ω and $\{\sigma_{lk}, l = 1,\ldots,L, k = 1,\ldots,K\}$ from $\boldsymbol{x}$.

III. Our Approach

A. Atoms and Atomic Norm

Let $\boldsymbol{B}_l := \operatorname{diag}(\boldsymbol{g}_l)$ denote an $N \times N$ diagonal matrix with diagonal $\boldsymbol{g}_l$. Also let

$$\boldsymbol{e}_t := \begin{bmatrix} e^{-j2\pi t(-2M)} \\ \vdots \\ e^{-j2\pi t(2M)} \end{bmatrix} \in \mathbb{C}^N,\ t \in [0,1]$$

denote a length-N vector of samples from a discrete-time complex exponential signal with digital frequency t. We rewrite the measurements $\boldsymbol{x}$ in (2) with matrix notation

$$\boldsymbol{x} = \sum_{k=1}^{K}\sum_{l=1}^{L}\sigma_{lk}\boldsymbol{B}_l\boldsymbol{e}_{t_k} = \sum_{k=1}^{K} c_k\left(\sum_{l=1}^{L}\boldsymbol{\alpha}_k[l]\boldsymbol{B}_l\boldsymbol{e}_{t_k}\right), \tag{3}$$

where $c_k =: \sqrt{\sum_l |\sigma_{lk}|^2}$ and $\boldsymbol{\alpha}_k \in \mathbb{C}^L$ with elements $\boldsymbol{\alpha}_k[l] := \frac{\sigma_{kl}}{c_k}$ for all $k = 1,\ldots,K$ and $l = 1,\ldots,L$. By definition, we have $\|\boldsymbol{\alpha}_k\|_2 = 1$. So $\boldsymbol{x}$ can be viewed as a sparse combination of elements from the atomic set

$$\mathcal{A} := \left\{\boldsymbol{a}(t,\boldsymbol{\alpha}) = \sum_{l=1}^{L}\boldsymbol{\alpha}[l]\boldsymbol{B}_l\boldsymbol{e}_t, \boldsymbol{\alpha} \in \mathbb{C}^L, \|\boldsymbol{\alpha}\|_2 = 1\right\},$$

which can be viewed as an infinite dictionary governed by the parameters t and $\boldsymbol{\alpha}$. Note that the diagonal matrices $\boldsymbol{B}_l$ are fixed and known. The atomic norm of $\boldsymbol{x}$ is then defined as

$$\|\boldsymbol{x}\|_{\mathcal{A}} = \inf_{\substack{c_k \geq 0,\\ \|\boldsymbol{\alpha}_k\|_2 = 1\\ t_k \in [0,1]}}\left\{\sum_k c_k \mid \boldsymbol{x} = \sum_k c_k\boldsymbol{a}(t_k,\boldsymbol{\alpha}_k)\right\}, \tag{4}$$

which can be viewed as a penalty for promoting group sparsity of $\boldsymbol{x}$, i.e., representing $\boldsymbol{x}$ by picking as few items as possible from the group set

$$\{\{\boldsymbol{B}_l\boldsymbol{e}_t, l = 1,\ldots,L\}, t \in [0,1]\}.$$

B. Semidefinite Program Characterization

The following result indicates that the atomic norm $\|\boldsymbol{x}\|_{\mathcal{A}}$ admits an equivalent SDP characterization.

Theorem 1. *The atomic norm $\|\boldsymbol{x}\|_{\mathcal{A}}$ can be written equivalently as*

$$\begin{aligned}\|\boldsymbol{x}\|_{\mathcal{A}} = \inf_{\boldsymbol{u}\in\mathbb{C}^N, c}\Bigg\{&\frac{1}{2N}\operatorname{trace}(\operatorname{Toep}(\boldsymbol{u})) + \frac{1}{2}c \mid \\ &\operatorname{Toep}(\boldsymbol{u}) \succeq 0, \\ &\begin{bmatrix}\sum_l \boldsymbol{B}_l \operatorname{Toep}(\boldsymbol{u})\boldsymbol{B}_l^{\mathrm{H}} & \boldsymbol{x} \\ \boldsymbol{x}^H & c\end{bmatrix} \succeq 0\Bigg\}.\end{aligned} \tag{5}$$

Proof of Theorem 1: Denote the value of the right hand side as $\mathrm{SDP}(\boldsymbol{x})$. Suppose $\boldsymbol{x} = \sum_k c_k\boldsymbol{a}(f_k,\boldsymbol{\alpha}_k)$ with $c_k > 0$. Define $\boldsymbol{u} = \sum_k c_k\boldsymbol{e}_{t_k}$ and $c = \sum_k c_k$. We have $\operatorname{Toep}(\boldsymbol{u}) = \sum_k c_k\boldsymbol{e}_{t_k}\boldsymbol{e}_{t_k}^{\mathrm{H}} \succeq 0$. Note that

$$\begin{bmatrix}\sum_l \boldsymbol{B}_l \operatorname{Toep}(\boldsymbol{u})\boldsymbol{B}_l^{\mathrm{H}} & \boldsymbol{x} \\ \boldsymbol{x}^H & t\end{bmatrix} =$$

$$\sum_k c_k\begin{bmatrix}[\boldsymbol{B}_1\boldsymbol{e}_{t_k} \cdots \boldsymbol{B}_L\boldsymbol{e}_{t_k}] \\ \boldsymbol{\alpha}_k^{\mathrm{H}}\end{bmatrix}\begin{bmatrix}\begin{bmatrix}\boldsymbol{e}_{t_k}^{\mathrm{H}}\boldsymbol{B}_1^{\mathrm{H}} \\ \vdots \\ \boldsymbol{e}_{t_k}^{\mathrm{H}}\boldsymbol{B}_L^{\mathrm{H}}\end{bmatrix} & \boldsymbol{\alpha}_k\end{bmatrix} \succeq 0$$

and

$$\frac{1}{2N}\operatorname{trace}(\operatorname{Toep}(\boldsymbol{u})) + \frac{1}{2}t = \sum_k c_k.$$

Therefore, $\mathrm{SDP}(\boldsymbol{x}) \leq \|\boldsymbol{x}\|_{\mathcal{A}}$.

On the other hand, suppose for some $\boldsymbol{u}$ and $\boldsymbol{x}$,

$$\begin{bmatrix}\sum_l \boldsymbol{B}_l \operatorname{Toep}(\boldsymbol{u})\boldsymbol{B}_l^{\mathrm{H}} & \boldsymbol{x} \\ \boldsymbol{x}^H & c\end{bmatrix} \succeq 0, \quad \operatorname{Toep}(\boldsymbol{u}) \succeq 0.$$

By the Vandermonde decomposition lemma [14], we have

$$\operatorname{Toep}(\boldsymbol{u}) = \boldsymbol{V}\boldsymbol{D}\boldsymbol{V}^{\mathrm{H}} = \sum_k d_k\boldsymbol{e}_{t_k}\boldsymbol{e}_{t_k}^{\mathrm{H}}$$

with $\boldsymbol{D} = \mathrm{diag}(d_K)$, $d_k > 0$. It follows that $\boldsymbol{x}$ is in the range of $\boldsymbol{B}_l \boldsymbol{V}$, i.e., $\boldsymbol{x} = \sum_l \boldsymbol{B}_l \boldsymbol{V} \boldsymbol{w}_l$. Let $\gamma = \frac{c}{\sum_l \|\boldsymbol{w}_l\|_2^2}$. Now we have

$$\left[\begin{array}{cc} \sum_l \boldsymbol{B}_l \operatorname{Toep}(\boldsymbol{u}) \boldsymbol{B}_l^{\mathrm{H}} & \boldsymbol{x} \\ \boldsymbol{x}^H & c \end{array} \right] = \boldsymbol{A} \boldsymbol{\Sigma} \boldsymbol{A}^{\mathrm{H}} \succeq 0,$$

where

$$\boldsymbol{A} = \left[\begin{array}{cc} [\ \boldsymbol{B}_1 \boldsymbol{V} \ \ \cdots \ \ \boldsymbol{B}_L \boldsymbol{V}\ ] & \boldsymbol{0} \\ \boldsymbol{0} & [\ \boldsymbol{w}_1^{\mathrm{H}} \ \ \cdots \ \ \boldsymbol{w}_L^{\mathrm{H}}\ ] \end{array} \right],$$

$$\boldsymbol{\Sigma} = \left[\begin{array}{cc} \left[\begin{array}{ccc} \boldsymbol{D} & & \\ & \ddots & \\ & & \boldsymbol{D} \end{array} \right] & \mathbf{I} \\ \mathbf{I} & \gamma \mathbf{I} \end{array} \right].$$

This implies $\boldsymbol{\Sigma} \succeq 0$. It follows from the Schur complement lemma that

$$\gamma \mathbf{I} \geq \left[\begin{array}{ccc} \boldsymbol{D} & & \\ & \ddots & \\ & & \boldsymbol{D} \end{array} \right]^{-1}.$$

Thus, we have

$$c = \left(\sum_l \|\boldsymbol{w}_l\|_2^2 \right) \gamma \geq \sum_k \frac{\sum_l |\boldsymbol{w}_l[k]|^2}{d_k}.$$

It follows from the fact $\frac{1}{N} \operatorname{trace}(\operatorname{Toep}(\boldsymbol{u})) = \operatorname{trace}(\boldsymbol{D})$ that

$$\frac{1}{2N} \operatorname{trace}(\operatorname{Toep}(\boldsymbol{u})) + \frac{1}{2} c \geq \frac{1}{2} \sum_k d_k + \frac{1}{2} \sum_k \frac{\sum_l |\boldsymbol{w}_l[k]|^2}{d_k}$$
$$\geq \sqrt{\left(\sum_k d_k \right) \left(\sum_k \frac{\sum_l |\boldsymbol{w}_l[k]|^2}{d_k} \right)} \geq \|\boldsymbol{x}\|_{\mathcal{A}},$$

where the last line follows from the Cauchy-Schwartz inequality. ■

Rewrite the measurements $\boldsymbol{x}$ in (3) as

$$\boldsymbol{x} = \sum_{l=1}^{L} \boldsymbol{B}_l \left(\sum_{k=1}^{K} c_k \boldsymbol{\alpha}_k[l] \boldsymbol{e}_{t_k} \right) = \sum_{l=1}^{L} \boldsymbol{B}_l \widetilde{\boldsymbol{x}}_l$$

with $\widetilde{\boldsymbol{x}}_l = \sum_{k=1}^K c_k \boldsymbol{\alpha}_k[l] \boldsymbol{e}_{t_k}$. Denote $\widetilde{\boldsymbol{X}} = [\widetilde{\boldsymbol{x}}_1 \cdots \widetilde{\boldsymbol{x}}_L] = \sum_{k=1}^K c_k \boldsymbol{e}_{t_k} \boldsymbol{\alpha}_k^{\mathrm{T}}$, where T represents the nonconjugate transpose operator. This is the lifting scheme utilized in [18, 19], where a different atomic norm induced by a different atomic set is applied for the augmented matrix $\widetilde{\boldsymbol{X}}$. Note that the atomic norm utilized in [18, 19] for $\widetilde{\boldsymbol{X}}$ is equivalent to $\|\boldsymbol{x}\|_{\mathcal{A}}$, the atomic norm utilized for $\boldsymbol{x}$. Thus the SDP characterization for solving the corresponding atomic norm minimization in [18, 19] is also equivalent to $\|\boldsymbol{x}\|_{\mathcal{A}}$.

Proposition 1. *The SDP characterization in* (5) *has the following equivalent form*

$$\|\boldsymbol{x}\|_{\mathcal{A}} = \inf_{\substack{\boldsymbol{u} \in \mathbb{C}^N \\ \boldsymbol{C}, \boldsymbol{X}}} \left\{ \frac{1}{2N} \operatorname{trace}(\operatorname{Toep}(\boldsymbol{u})) + \frac{1}{2} \operatorname{trace}(\boldsymbol{C}) \mid \boldsymbol{x} = \mathcal{X}(\boldsymbol{X}), \left[\begin{array}{cc} \operatorname{Toep}(\boldsymbol{u}) & \boldsymbol{X} \\ \boldsymbol{X}^{\mathrm{H}} & \boldsymbol{C} \end{array} \right] \succeq 0 \right\}.$$

Here $\mathcal{X}(\boldsymbol{X}) = \sum_l \boldsymbol{B}_l \boldsymbol{x}_l$ with $\boldsymbol{x}_l$ being the l-th column of $\boldsymbol{X}$.

C. Recovery Guarantee

We can certify the optimality of minimizing the atomic norm defined in (4) using the following proposition.

Proposition 2. *Suppose $\boldsymbol{x}^\# = \sum_k c_k \boldsymbol{a}(t_k, \boldsymbol{\alpha}_k)$ with $c_k > 0$, $k = 1, 2, \ldots, K$ and $\{\boldsymbol{B}_l \boldsymbol{e}_{t_k}, l = 1, \ldots L, k = 1, \ldots, K\}$ are linearly independent. If there is a vector $\boldsymbol{p} \in \mathbb{C}^N$ such that the corresponding vector-valued dual polynomial $\boldsymbol{q}(t)[l] = \boldsymbol{e}_t^{\mathrm{H}} \boldsymbol{B}_l^{\mathrm{H}} \boldsymbol{p}$ satisfies*

$$\begin{aligned} &\boldsymbol{q}(t_k)[l] = \boldsymbol{\alpha}_k[l], \quad t_k \in \Omega, \\ &\|\boldsymbol{q}(t)\|_2 < 1, \quad t \notin \Omega, \end{aligned} \tag{6}$$

then $\boldsymbol{x}^\# = \sum_k c_k \boldsymbol{a}(f_k, \boldsymbol{\alpha}_k)$ is the unique atomic decomposition satisfying $\|\boldsymbol{x}^\#\|_{\mathcal{A}} = \sum_k c_k$.

The above optimality conditions are derived from the facts that the atomic norm minimization is convex with strong duality holding and that both primal and dual optimal values are attained. We omit the proof due to space limitations. We note that the construction of such a dual polynomial heavily depends on $\boldsymbol{g}_l$. Inspired by [12, 14], where the dual polynomial is constructed with the square of the Fejér kernel, we can construct a dual polynomial $\boldsymbol{q}(t)$ that satisfies (6) as long as $\boldsymbol{g}_l$ is populated from certain distributions.

Theorem 2. *[19] Suppose $\boldsymbol{x}^\# = \sum_k c_k \boldsymbol{a}(t_k, \boldsymbol{\alpha}_k)$ with $c_k > 0$, $k = 1, 2, \ldots, K$. Suppose*

$$\overline{\boldsymbol{g}}_m := [\ \boldsymbol{g}_1[m] \quad \boldsymbol{g}_2[m] \quad \cdots \quad \boldsymbol{g}_L[m]\ ]$$

for $m = -2M, \ldots, 2M$ are i.i.d. samples from a distribution $\mathcal{F}$ that satisfies the following two conditions

$$\mathrm{E}\, \boldsymbol{f}^{\mathrm{H}} \boldsymbol{f} = \mathbf{I}_L, \quad \max_{1 \leq l \leq L} |\boldsymbol{f}[l]|^2 \leq \mu(\mathcal{F}), \quad \boldsymbol{f} \sim \mathcal{F}. \tag{7}$$

Also assume $\Delta(\Omega) := \min_{k \neq k'} |t_k - t_{k'}|$, the smallest wrap-around distance between any pair of the spikes, is greater than $\frac{1}{M}$ and $M \geq 64$. Additionally, assume that $\boldsymbol{\alpha}_k$ are i.i.d. randomly generated from the uniform distribution on the complex unit sphere $\mathbb{CS}^{L-1}$. Then, there exists a numerical constant C such that

$$M \geq C \mu(\mathcal{F}) K L \log\left(\frac{MKL}{\delta}\right) \log\left(\frac{ML}{\delta}\right)$$

is sufficient to guarantee that $\boldsymbol{x}^\# = \sum_k c_k \boldsymbol{a}(f_k, \boldsymbol{\alpha}_k)$ is the unique atomic decomposition satisfying $\|\boldsymbol{x}^\#\|_{\mathcal{A}} = \sum_k c_k$ with probability at least $1 - \delta$.

The main idea is to construct a dual polynomial $\boldsymbol{q}(t)$ (which is similar to what is used in [19]) that satisfies (6).

Remark 1. The two conditions in (7) are referred as the isotropy and incoherence properties of $\mathcal{F}$ [18, 19]. We note that not all components need to be random. Without loss of generality, we suppose $\boldsymbol{f}[1] = 1$, $\boldsymbol{f} \sim \mathcal{F}$. Such a distribution $\mathcal{F}$ also satisfies (7) as long as the distribution of the other components satisfies certain conditions. In this case, $\boldsymbol{g}_1[m] = 1$ for all m which corresponds to the direct path in many applications that involve multipath. Theorem 2 guarantees that the spikes can be recoverd from $\boldsymbol{x}$ (a superpostion of signals from all paths) as long as the modulations $\boldsymbol{g}_l$ corresponding to the l-th path for all $l \geq 2$ jointly satisfy (7).

D. Localizing the spikes

The SDP formulation (5) can be used to recover the spikes. Suppose $\boldsymbol{u}$ is an optimal solution to (5). Then the Vandermonde decomposition of $\mathrm{Toep}(\boldsymbol{u})$ characterizes the spikes.

The dual norm of $\|\boldsymbol{x}\|_{\mathcal{A}}$ can be defined as

$$\|\boldsymbol{p}\|_{\mathcal{A}}^{*} = \sup_{\|\boldsymbol{x}\|_{\mathcal{A}}} \langle \boldsymbol{p}, \boldsymbol{x}\rangle_{\mathbb{R}} = \sup_{t\in[0,1]} \sqrt{\sum_{l} \left|\boldsymbol{e}_t^{\mathrm{H}} \boldsymbol{B}_l^{\mathrm{H}} \boldsymbol{p}\right|^2}.$$

The dual problem of minimizing the atomic norm (4) can be written as

$$\text{maximize } \langle \boldsymbol{x}, \boldsymbol{p}\rangle_{\mathbb{R}}\,, \quad \text{subject to } \|\boldsymbol{p}\|_{\mathcal{A}}^{*} \le 1 \tag{8}$$

which also has an equivalent SDP formulation.

The spike locations can alternatively be identified from $\widehat{\boldsymbol{p}}$, the optimal solution to (8). To be precise, consider the vector valued dual polynomial

$$\widehat{\boldsymbol{q}}(t)[l] = \boldsymbol{e}_f^{\mathrm{H}} \boldsymbol{B}_l^{\mathrm{H}} \widehat{\boldsymbol{p}}.$$

The set of frequencies can be obtained by finding the peaks of

$$\|\widehat{\boldsymbol{q}}(t)\|_2 : \widehat{\Omega} = \{t : \|\widehat{\boldsymbol{q}}(t)\|_2 = 1\}.$$

E. Revisiting the case $L = 1$

When $L = 1$, (1) reduces to $x(t) = h(t)$. Then the measurements in (3) simplify to $\boldsymbol{x} = \sum_k \sigma_{1k} \boldsymbol{e}_{t_k}$. Recovering the spikes from $\boldsymbol{x}$ is the super-resolution [12] or line spectrum estimation problem [14]. One can define an atomic set $\mathcal{A}' = \{\boldsymbol{e}_t, t \in [0,1]\}$ and obtain an equivalent SDP characterization (which is similar to (5), see [12, 14]) for the corresponding atomic norm. The dual norm is given by

$$\|\boldsymbol{p}\|_{\mathcal{A}'}^{*} = \sup_{t\in[0,1]} \left|\boldsymbol{e}_t^{\mathrm{H}} \boldsymbol{p}\right|.$$

Let $\widetilde{\boldsymbol{p}}$ denote the optimal solution to the dual problem

$$\text{maximize } \langle \boldsymbol{x}, \boldsymbol{p}\rangle_{\mathbb{R}}\,, \quad \text{subject to } \|\boldsymbol{p}\|_{\mathcal{A}'}^{*} \le 1.$$

The spikes can be localized by finding the peaks of $\widetilde{q}(t) = \boldsymbol{e}_t^{\mathrm{H}} \widetilde{\boldsymbol{p}}$ [12, 14].

F. Atomic Norm Soft Thresholding (AST)

We conclude this section with a discussion about the presence of noise in the measurements:

$$\boldsymbol{y}[n] = \boldsymbol{x}[n] + \boldsymbol{w}[n]$$

for $n = -2M, -2M+1, \ldots, 2M$. Here $\boldsymbol{w}[n]$ is additive noise. In this case, we can obtain an estimate $\widehat{\boldsymbol{x}}$ that solves the atomic norm soft thresholding (AST):

$$\underset{\widetilde{\boldsymbol{x}}\in\mathbb{C}^N}{\text{minimize}}\, \lambda\|\boldsymbol{x}\|_{\mathcal{A}} + \frac{1}{2}\|\widetilde{\boldsymbol{x}} - \boldsymbol{y}\|^2$$

where λ is an appropriately chosen regularization parameter that depends on the noise level [13, 16].

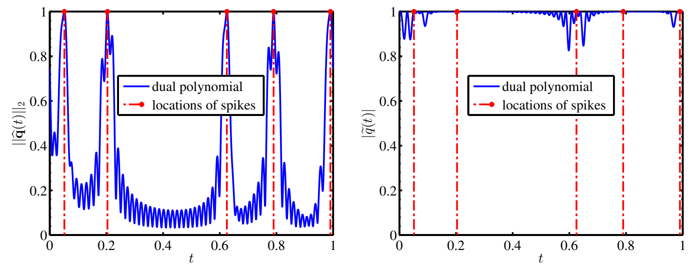

Figure 1. Illustration of dual polynomials. Left: $\|\widehat{\boldsymbol{q}}(t)\|_2$; Right: $|\widetilde{q}(t)|$. The dashed red lines represent the locations of spikes.

IV. Numerical Simulations

We present a synthetic experiment arising in [3] to support the proposed approach. Let $N = 64$. Without loss of generality, we set the measurement index $n \in \{0, \ldots, N-1\}$ instead of $n \in \{-2M, \ldots, 2M\}$. We generate $K = 5$ spikes with spike locations uniformly at random satisfying the minimum separation $\Delta(\Omega) > 1/N$ and coefficients generated with dynamic range of 10 and uniform phase. We generate $L = 2$ modulations with $\boldsymbol{g}_1[n] = 1$ and $\boldsymbol{g}_2[n] = e^{j2\pi n\tau}$, where $\tau = 0.013 \approx \frac{1}{N}$, i.e., the corresponding $g_1(t) = \delta(t)$ and $g_2(t) = \delta(t-\tau)$. We choose $\xi_1(t) = 1$ and $\xi_2(t_k)$ with i.i.d. standard Gaussian entries for $k = 1, \ldots, K$. Note that in this case

$$(h(t)\xi_1(t)) \star \overline{g}_1(t) = \sum_{k=1}^{K} \sigma_{1k}\delta(t - t_k)$$

and

$$(h(t)\xi_2(t)) \star \overline{g}_2(t) = \sum_{k=1}^{K} \sigma_{2k}\delta(t - t_k - \tau),$$

which implies we can alternatively apply a super-resolution algorithm [12] (also see Section III-E) to recover all the spikes $\{t_k, t_k + \tau\}_{k=1}^{K}$. Figure 1 illustrates the dual polynomials of $\|\widehat{\boldsymbol{q}}(t)\|_2$ and $|\widetilde{q}(t)|$ (see Sections III-D and III-E, respectively). We observe that the spikes can be localized correctly from the peaks of the dual polynomial $\|\widehat{\boldsymbol{q}}(t)\|_2$, while $|\widetilde{q}(t)|$ provides many spurious spikes.

Acknowledgment

This work was supported by NSF grant CCF-1409261.

References

[1] P. Setlur, T. Negishi, N. Devroye, and D. Erricolo, "Multipath exploitation in non-los urban synthetic aperture radar," *IEEE Jour. Selected Top. Sign. Proc.*, vol. 8, pp. 137–152, Feb 2014.

[2] W. U. Bajwa, J. Haupt, A. M. Sayeed, and R. Nowak, "Compressed channel sensing: A new approach to estimating sparse multipath channels," *Proc. IEEE*, vol. 98, no. 6, pp. 1058–1076, 2010.

[3] M. Leigsnering, F. Ahmad, M. Amin, and A. Zoubir, "Multipath exploitation in through-the-wall radar imaging using sparse reconstruction," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 50, no. 2, pp. 920–939, 2014.

[4] Z. Zhu and M. B. Wakin, "Wall clutter mitigation and target detection using Discrete Prolate Spheroidal Sequences," in *3rd Int. Workshop on Compressed Sensing Theory and its Applications to Radar, Sonar and Remote Sensing (CoSeRa)*, June 2015.

[5] F. Ahmad, J. Qian, and M. Amin, “Wall clutter mitigation using Discrete Prolate Spheroidal Sequences for sparse reconstruction of indoor stationary scenes,” *IEEE Trans. Geosci. Remote Sens.*, vol. 53, pp. 1549–1557, March 2015.

[6] Z. Zhu and M. B. Wakin, “On the dimensionality of wall and target return subspaces in through-the-wall radar imaging,” in *4th Int. Workshop on Compressed Sensing Theory and its Applications to Radar, Sonar and Remote Sensing (CoSeRa)*, September 2016.

[7] S. Shahid, J. Walker, and L. S. Smith, “A new spike detection algorithm for extracellular neural recordings,” *IEEE Trans. Biomed. Eng.*, vol. 57, no. 4, pp. 853–866, 2010.

[8] I. Obeid, M. A. Nicolelis, and P. D. Wolf, “A multichannel telemetry system for single unit neural recordings,” *J. Neurosci. Methods.*, vol. 133, no. 1?, pp. 33 – 38, 2004.

[9] Y. Hua and T. K. Sarkar, “Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise,” *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 38, no. 5, pp. 814–824, 1990.

[10] R. Schmidt, “Multiple emitter location and signal parameter estimation,” *IEEE Trans. Antennas Propag.*, vol. 34, no. 3, pp. 276–280, 1986.

[11] V. Chandrasekaran, B. Recht, P. A. Parrilo, and A. S. Willsky, “The convex geometry of linear inverse problems,” *Foundations Comput. Math.*, vol. 12, no. 6, pp. 805–849, 2012.

[12] E. J. Candès and C. Fernandez-Granda, “Towards a mathematical theory of super-resolution,” *Comm. Pure Appl. Math.*, vol. 67, no. 6, pp. 906–956, 2014.

[13] B. N. Bhaskar, G. Tang, and B. Recht, “Atomic norm denoising with applications to line spectral estimation,” *IEEE. Trans. Signal Process.*, vol. 61, no. 23, pp. 5987–5999, 2013.

[14] G. Tang, B. N. Bhaskar, P. Shah, and B. Recht, “Compressed sensing off the grid,” *IEEE Trans. Inf. Theory*, vol. 59, no. 11, pp. 7465–7490, 2013.

[15] Q. Li and G. Tang, “Approximate support recovery of atomic line spectral estimation: A tale of resolution and precision,” *arXiv preprint arXiv:1612.01459*, 2016.

[16] Z. Zhu, G. Tang, P. Setlur, S. Gogineni, M. Wakin, and M. Rangaswamy, “Super-resolution in SAR imaging: Analysis with the atomic norm,” in *Proc. IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)*, July 2016.

[17] S. Li, D. Yang, G. Tang, and M. B. Wakin, “Atomic norm minimization for modal analysis from random and compressed samples,” *arXiv preprint arXiv:1703.00938*, 2017.

[18] Y. Chi, “Guaranteed blind sparse spikes deconvolution via lifting and convex optimization,” *IEEE J. Selected Top. Sign. Proc.*, vol. 10, no. 4, pp. 782–794, 2016.

[19] D. Yang, G. Tang, and M. B. Wakin, “Super-resolution of complex exponentials from modulations with unknown waveforms,” to appear in *IEEE Trans. Inf. Theory*, 2016.

[20] Y. Li and Y. Chi, “Stable separation and super-resolution of mixture models,” *arXiv preprint arXiv:1506.07347*, 2015.

[21] A. Ahmed, B. Recht, and J. Romberg, “Blind deconvolution using convex programming,” *IEEE Trans. Inf. Theory*, vol. 60, no. 3, pp. 1711–1732, 2014.

[22] G. Tang and B. Recht, “Convex blind deconvolution with random masks,” in *Computational Optical Sensing and Imaging*, pp. CW4C–1, Optical Society of America, 2014.

[23] S. Bahmani and J. Romberg, “Lifting for blind deconvolution in random mask imaging: Identifiability and convex relaxation,” *SIAM J. Imaging Sci.*, vol. 8, no. 4, pp. 2203–2238, 2015.

[24] K.-C. Toh, M. J. Todd, and R. H. Tütüncü, “SDPT3—a MATLAB software package for semidefinite programming, version 1.3,” *Optimiz. Methods Softw.*, vol. 11, no. 1-4, pp. 545–581, 1999.