A Super-Resolution Algorithm for Multiband Signal Identification

Zhihui Zhu, Dehui Yang, Michael B. Wakin, and Gongguo Tang
Department of Electrical Engineering, Colorado School of Mines
{zzhu, dyang, mwakin, gtang}@mines.edu

***Abstract*—Recent advances in convex optimization have led to super-resolution algorithms that provide exact frequency localization in multitone signals from limited time-domain samples. Such localization is accomplished by minimizing a certain atomic norm, which can be implemented in a semidefinite program. In this work, we consider the identification of multiband signals, which are comprised of multiple, unknown narrow bands of frequency content at multiple carrier frequencies. Integrating a basis of modulated discrete prolate spheroidal sequences (DPSS's) into the atomic norm minimization framework, we introduce a technique for estimating the unknown band positions based on limited time-domain samples of the signal.**

***Index Terms*—Atomic norm, discrete prolate spheroidal sequences, multiband identification, super-resolution**

I. Introduction

Super-resolution refers to the problem of identifying signal parameters with high resolution from limited measurements. The conventional super-resolution problem considers a multitone analog signal that can be expressed as a sum of J complex exponentials of various frequencies:

$$x(t) = \sum_{j=1}^{J} \theta_j e^{i2\pi F_j t}.$$

Suppose such a multitone signal $x(t)$ is bandlimited with bandlimit $\frac{B_{\text{nyq}}}{2}$ Hz, i.e., that $\max_j |F_j| \leq \frac{B_{\text{nyq}}}{2}$. Let $\boldsymbol{x}$ denote the length-N vector obtained by uniformly sampling $x(t)$ over the time interval $[0, NT_s)$ with sampling period $T_s \leq \frac{1}{B_{\text{nyq}}}$:

$$\boldsymbol{x}[n] = \sum_{j=1}^{J} \theta_j e^{i2\pi f_j n}, n = 0, 1, \ldots, N-1,$$

where $f_j = F_j T_s$. This model arises in problems such as radar signal processing with point targets [1], [2], line spectral estimation [3]–[5], and spike detection for neural recordings [6].

We consider a more realistic model in which $x(t)$ has a continuous-time Fourier transform (CTFT) supported on a union of several narrow bands

$$\mathbb{F} = \bigcup_{j=1}^{J} [F_j - B_{\text{band}_j}/2, F_j + B_{\text{band}_j}/2],$$

i.e.,

$$x(t) = \int_{\mathbb{F}} X(F) e^{j2\pi Ft} dF. \tag{1}$$

Here $X(F)$ denotes the CTFT of $x(t)$. The band centers are given by the frequencies $\{F_j\}_{j\in[J]}$ and the band widths are denoted by $\{B_{\text{band}_j}\}_{j\in[J]}$, where $[J]$ denotes the set $\{1, 2, \ldots, J\}$. Such multiband signal models arise in communications [7], radar signal processing with non-point targets [8], [9], and mitigation of narrowband interference [10], [11]. Our goal is to identify the band centers $\{F_j\}_{j\in[J]}$ from the finite-length samples (see (2)) of $x(t)$.

In [12], the authors consider recovering the finite-length samples of a multiband signal $x(t)$ from compressive measurements with a dictionary of multiband modulated discrete prolate spheroidal sequences (DPSS's) [13] dictionary. The representation ability of the multiband modulated DPSS dictionary for sampled multiband signals is further analyzed in [14]. However, the works [12], [14] assume that either the band centers are known or the bands are located in certain predefined grids. Instead, we attempt to identify the band centers which are assumed to be located arbitrarily as long as they satisfy a certain separation condition. In this paper, we apply atomic norm techniques to invert the parameters in (1) from finite-length samples. The atomic norm is utilized to promote group sparsity and has an equivalent semidefinite programming (SDP) characterization. Thus the problem can be solved efficiently using an off-the-shelf solver [15]. Our work differs from previous works on line spectral estimation [3], [4], [16] in that we assume that the CTFT of the corresponding analog signal occupies certain bandwidths, rather than that it is composed of several spikes.

The outline of this paper is as follows. The main problem is illustrated in Section II. Our approach based on atomic norm minimization is discussed in Section III. Section IV discusses the development of a recovery guarantee for the atomic norm minimization approach. Section V presents some simulations to support our proposed method.

II. Problem Setup

To begin, we define

$$\boldsymbol{e}_f = \begin{bmatrix} 1 & e^{i2\pi f1} & \cdots & e^{i2\pi f(N-1)} \end{bmatrix}^T \in \mathbb{C}^N$$

as a length-N sampled complex sinusoid of normalized frequency $f \in [-\frac{1}{2}, \frac{1}{2}]$. We assume each band has the same bandwidth, i.e., $B_{\text{band}_j} = B_{\text{band}}$ for all $j \in [J]$. Consider a signal vector $\boldsymbol{x}$ that arises by sampling the analog signal $x(t)$ over the time interval $[0, NT_s)$ with sampling period T_s. We

assume T_s is chosen to satisfy the minimum Nyquist sampling rate, which means

$$T_s \leq \frac{1}{B_{\text{nyq}}} := \frac{1}{2\max_{j\in[J]}\{|F_j \pm B_{\text{band}}/2|\}}.$$

Under these assumptions, the sampled multiband signal $\boldsymbol{x}$ can be expressed as an integral of sampled pure tones (i.e., discrete-time sinusoids)

$$\boldsymbol{x} = \int_{\mathbb{W}} \widetilde{x}(f)\boldsymbol{e}_f \, df, \tag{2}$$

where the digital frequency f is integrated over a union of intervals

$$\mathbb{W} := T_s\mathbb{F} = \bigcup_{j=1}^{J}[f_j - W, f_j + W] \subseteq \left[-\frac{1}{2}, \frac{1}{2}\right].$$

Here, $W = \frac{B_{\text{band}}T_s}{2}$, $f_j = T_sF_j$ and the weighting function $\widetilde{x}(f) = \frac{1}{T_s}X(f/T_s)$ equals the scaled CTFT of the analog signal and corresponds to the discrete-time Fourier transform (DTFT) of its infinite sample sequence. However, we stress that our interest concerns the finite-length sample vector $\boldsymbol{x}$ and not the infinite sample sequence. Spectral analysis of $\boldsymbol{x}$ is complicated by its finite nature: conventional frequency analysis based on the discrete Fourier transform (DFT) will suffer from familiar "leakage" artifacts, making it difficult, for example, to reliably identify the band centers f_j. However, taking each $\boldsymbol{e}_{f_j}$ to act as a modulator to the center frequency f_j and viewing each $\boldsymbol{g}_j$ as samples of a baseband signal formed by integrating over a collection of sinusoids $\boldsymbol{e}_f$ with $f \in [-W, W]$, we can express the vector $\boldsymbol{x}$:

$$\boldsymbol{x} = \sum_{j=1}^{J}\left(\int_{f_j-W}^{f_j+W} \widetilde{x}(f)\boldsymbol{e}_f \, df\right) = \sum_{j=1}^{J} \boldsymbol{e}_{f_j} \odot \boldsymbol{g}_j \tag{3}$$

with

$$\boldsymbol{g}_j = \int_{-W}^{W} \widetilde{x}(f + f_j)\boldsymbol{e}_f \, df. \tag{4}$$

Here $\odot$ represents the elementwise (Hadamard) product between two length-N vectors. On its surface, the observation model in (3) describes a collection of sinusoids with unknown frequencies f_j modulated by different unknown waveforms $\boldsymbol{g}_j$ [17], [18]. Our goal is to identify the frequencies $\Omega = \{f_1, \ldots, f_J\}$ and the unknowns waveforms $\boldsymbol{g}_1, \ldots, \boldsymbol{g}_J$.

III. Our Approach

A. DPSS basis

It is known [12], [14] that vectors $\boldsymbol{g}_j$ formed from integrating sinusoids over a narrow range of frequencies (as in (4)) can be approximated—to a very high degree of accuracy—using a basis constructed from the DPSS's [13] from time-frequency analysis. Given $W \in (0, \frac{1}{2})$, the DPSS vectors $\{\boldsymbol{s}_{N,W}^{(\ell)}\}_{\ell\in[N]}$ are length-N vectors whose DTFT have a certain concentration in the digital frequency band $[-W, W]$. Now define

$$\boldsymbol{S} := \begin{bmatrix} \boldsymbol{s}_{N,W}^{(1)} & \cdots & \boldsymbol{s}_{N,W}^{(L)} \end{bmatrix} \tag{5}$$

to contain the first L DPSS vectors for some value of $L \in \{1, 2, \ldots, N\}$ that we can choose as desired. The columns of $\boldsymbol{S}$ are orthonormal. Throughout the paper, for any matrix $\boldsymbol{S}$ with orthonormal columns, we use

$$\mathcal{P}_{\boldsymbol{S}} := \mathbf{I}_N - \boldsymbol{S}\boldsymbol{S}^{\text{H}}$$

to denote an orthogonal projection from $\mathbb{C}^N$ to the *orthogonal complement* of the subspace formed by the columns of $\boldsymbol{S}$. Taking $L \approx 2NW$, the dictionary $\boldsymbol{S}$ provides very accurate approximations (in a mean square error (MSE) sense) for all sampled sinusoids in the targeted band [14], [19], [20].

Theorem 1. *[14], [19], [20] Fix $W \in (0, \frac{1}{2})$. Let $\boldsymbol{S}$ be an $N \times L$ orthobasis defined in (5).*

- *(Asymptotic guarantee) For fixed $\eta \in (0, \frac{1}{2})$. choose $L = 2NW(1+\eta)$. Then there exist positive constants C_1, C_2 (where C_1, C_2 may depend on W and ϵ) such that for all $N \geq N_0$*

$$||\mathcal{P}_{\boldsymbol{S}}\boldsymbol{e}_f||_2 \leq C_1 N^{5/4} e^{-C_2 N}, \ \forall \ f \in [-W, W].$$

- *(Nonasymptotic guarantee) For any $\epsilon \in (0, \frac{1}{2})$, the orthobasis $\boldsymbol{S}$ satisfies*

$$\frac{\|\mathcal{P}_{\boldsymbol{S}}\boldsymbol{e}_f\|_2^2}{\|\boldsymbol{e}_f\|_2^2} \leq \epsilon$$

for all $f \in [-W, W]$ with

$$L = \left\lceil C_N \log\left(\frac{60\pi C_N}{\epsilon^2}\right)\right\rceil + 1,$$

and

$$\int_{-W}^{W} \frac{\|\mathcal{P}_{\boldsymbol{S}}\boldsymbol{e}_f\|_2^2}{\|\boldsymbol{e}_f\|_2^2} df \leq \epsilon$$

with

$$L = \left\lceil C_N \log\left(\frac{15C_N}{N\epsilon}\right)\right\rceil + 1.$$

Here $C_N = \frac{4}{\pi^2}\log(8N) + 6$.

It follows that, with $L \approx 2NW$ DPSS vectors (which can be much smaller than N), one obtains an efficient basis $\boldsymbol{S} \in \mathbb{R}^{N\times L}$ for representing most sampled baseband signals such as $\boldsymbol{g}_j$ in (4).

B. Atoms and Atomic Norm

Utilizing the fact that each $\boldsymbol{g}_j \approx \boldsymbol{S}\boldsymbol{\sigma}_j$ for some $\boldsymbol{\sigma}_j \in \mathbb{C}^L$, we can rewrite (3) as

$$\boldsymbol{x} \approx \sum_{j=1}^{J}(\boldsymbol{S}\boldsymbol{\sigma}_j) \odot \boldsymbol{e}_{f_j} = \sum_{j=1}^{J} \|\boldsymbol{\sigma}_j\|_2(\boldsymbol{S}\,\text{sign}(\boldsymbol{\sigma}_j)) \odot \boldsymbol{e}_{f_j},$$

where $\text{sign}(\boldsymbol{a}) = \frac{\boldsymbol{a}}{\|\boldsymbol{a}\|_2}$ denotes the sign of a vector. Thus, $\boldsymbol{x}$ can be viewed as a sparse combination of elements from the atomic set

$$\mathcal{A} := \left\{\boldsymbol{a}(f, \boldsymbol{\alpha}) = (\boldsymbol{S}\boldsymbol{\alpha}) \odot \boldsymbol{e}_f, \boldsymbol{\alpha} \in \mathbb{C}^L, \|\boldsymbol{\alpha}\|_2 = 1\right\}.$$

The atomic norm of $\boldsymbol{x}$ is then defined as

$$\|\boldsymbol{x}\|_{\mathcal{A}} = \inf_{\substack{c_j \geq 0, \\ \|\boldsymbol{\alpha}_j\|_2 = 1 \\ f_j \in [-\frac{1}{2}, \frac{1}{2})}} \left\{ \sum_j c_j : \boldsymbol{x} = \sum_j c_j \boldsymbol{a}(f_j, \boldsymbol{\alpha}_j) \right\}, \quad (6)$$

which can be used as a regularizer for promoting a certain notion of simplicity or sparsity of $\boldsymbol{x}$, in particular by representing $\boldsymbol{x}$ using as few items as possible from the atomic set $\mathcal{A}$. The atomic norm (6) has the following equivalent form [17], [18]

$$\|\boldsymbol{x}\|_{\mathcal{A}} = \inf_{\substack{\boldsymbol{u} \in \mathbb{C}^N \\ \boldsymbol{C}, \boldsymbol{X}}} \left\{ \frac{1}{2N} \operatorname{trace}(\operatorname{Toep}(\boldsymbol{u})) + \frac{1}{2} \operatorname{trace}(\boldsymbol{C}) \;\middle|\; \boldsymbol{x} = \mathcal{S}(\boldsymbol{X}), \begin{bmatrix} \operatorname{Toep}(\boldsymbol{u}) & \boldsymbol{X} \\ \boldsymbol{X}^{\mathrm{H}} & \boldsymbol{C} \end{bmatrix} \succeq 0 \right\}. \quad (7)$$

Here $\mathcal{S}(\boldsymbol{X}) = \sum_\ell \boldsymbol{s}_{N,W}^{(\ell)} \odot \boldsymbol{x}_\ell$ with $\boldsymbol{x}_\ell$ being the ℓ-th column of $\boldsymbol{X}$.

C. Localizing the band centers

As we explain below, the SDP formulation (7) can be used to recover the band centers $\{f_j\}_{j \in [J]}$. Suppose $\boldsymbol{u}$ is an optimal solution to (7). Then the Vandermonde decomposition of $\operatorname{Toep}(\boldsymbol{u})$ characterizes the band centers.

The dual norm of $\|\boldsymbol{p}\|_{\mathcal{A}}$ is defined as

$$\|\boldsymbol{p}\|_{\mathcal{A}}^* = \sup_{\|\boldsymbol{y}\|_{\mathcal{A}} \leq 1} \langle \boldsymbol{p}, \boldsymbol{y} \rangle_{\mathbb{R}} = \sup_{f \in [-\frac{1}{2}, \frac{1}{2}]} \left\| \boldsymbol{S}^{\mathrm{H}} (\boldsymbol{e}_{-f} \odot \boldsymbol{p}) \right\|_2.$$

The dual problem of computing the atomic norm (6) can be written as

$$\underset{\boldsymbol{p} \in \mathbb{C}^N}{\text{maximize}} \; \langle \boldsymbol{x}, \boldsymbol{p} \rangle_{\mathbb{R}}, \quad \text{subject to } \|\boldsymbol{p}\|_{\mathcal{A}}^* \leq 1 \quad (8)$$

which also has an equivalent SDP formulation.

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

$$\widehat{\boldsymbol{q}}(f) = \boldsymbol{S}^{\mathrm{H}} (\boldsymbol{e}_{-f} \odot \widehat{\boldsymbol{p}}). \quad (9)$$

The set of frequencies can be obtained by finding the peaks of $\|\widehat{\boldsymbol{q}}(f)\|_2$:

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

IV. Recovery Guarantee

We can certify the optimality of computing the atomic norm defined in (6) using the following proposition. Let $\boldsymbol{E}_f := \operatorname{diag}(\boldsymbol{e}_f)$ denote an $N \times N$ diagonal matrix for any $f \in [-\frac{1}{2}, \frac{1}{2})$.

Proposition 1. *Suppose $\boldsymbol{x} = \sum_j (\boldsymbol{S}\boldsymbol{\sigma}_j) \odot \boldsymbol{e}_{f_j}$ with $\|\boldsymbol{\sigma}_j\| > 0$, $j = 1, 2, \ldots, J$. Also suppose the columns $\left\{ \boldsymbol{E}_{f_j} \boldsymbol{s}_{N,W}^{(\ell)}, j = 1, \ldots, J, \ell = 1, \ldots, L \right\}$ are linearly independent. If there is a vector $\boldsymbol{p} \in \mathbb{C}^N$ such that the corresponding vector-valued dual polynomial $\boldsymbol{q}(f) = \boldsymbol{S}^{\mathrm{H}}(\boldsymbol{e}_{-f} \odot \boldsymbol{p})$ satisfies*

$$\begin{aligned} &\boldsymbol{q}(f_j) = \operatorname{sign}(\boldsymbol{\sigma}_j), \quad f_j \in \Omega, \\ &\|\boldsymbol{q}(f)\|_2 < 1, \quad f \notin \Omega, \end{aligned} \quad (10)$$

then $\boldsymbol{x} = \sum_j (\boldsymbol{S}\boldsymbol{\sigma}_j) \odot \boldsymbol{e}_{f_j}$ is the unique atomic decomposition satisfying $\|\boldsymbol{x}\|_{\mathcal{A}} = \sum_j \|\boldsymbol{\sigma}_j\|_2$.

The above optimality conditions are derived from the facts that the atomic norm minimization is convex, strong duality holds, and thus both primal and dual optimal values are attainable. Inspired by [3], [4], where the dual polynomial is constructed with the square of the Fejér kernel, [17] and [18] construct a dual polynomial $\boldsymbol{q}(f)$ that satisfies (10) when the orthobasis $\boldsymbol{S}$ is populated from certain random distributions.

However, here the DPSS basis $\boldsymbol{S}$ is not a random orthobasis and it has particular structural properties. The following result establishes that the DPSS vectors are approximately orthogonal to the modulated ones.

Lemma 1. *[12, Lemma 5.1] Fix $\epsilon \in (0, 1)$. Let $L = \lfloor 2NW(1-\epsilon) \rfloor$, and let $\boldsymbol{S}$ be the orthobasis as defined in (5). Then there exist constants C_1, C_2 (where C_1, C_2 may depend on W and ϵ) such that for all $N \geq N_0$ and $|f| \geq 2W$*

$$\left| \langle \boldsymbol{s}_{N,W}^{(\ell)}, \boldsymbol{E}_f \boldsymbol{s}_{N,W}^{(\ell')} \rangle \right| \leq 3\sqrt{C_1} e^{-\frac{C_2}{2} N}, \forall \ell, \ell' \in [L]$$

and

$$\left\| \boldsymbol{S}^{\mathrm{H}} \boldsymbol{E}_f \boldsymbol{S} \right\|_2 \leq 3N\sqrt{C_1} e^{-\frac{C_2 N}{2}}.$$

Rather than utilizing the Fejér kernel as in [3], [4], we construct a pre-certificate by simply solving the following problem:

$$\begin{aligned} &\text{minimize } \frac{1}{2} \|\boldsymbol{p}\|_2^2 \\ &\text{subject to } \boldsymbol{q}(f_j) = \operatorname{sign}(\boldsymbol{\sigma}_j), j \in [J]. \end{aligned} \quad (11)$$

With this, we can rewrite $\boldsymbol{e}_f \odot \boldsymbol{p} = \boldsymbol{E}_f \boldsymbol{p}$. Now the equality constraints in (11) can be written as a linear system of equations:

$$\underbrace{\begin{bmatrix} \boldsymbol{S}^{\mathrm{H}} \boldsymbol{E}_{f_1}^{\mathrm{H}} \\ \vdots \\ \boldsymbol{S}^{\mathrm{H}} \boldsymbol{E}_{f_J}^{\mathrm{H}} \end{bmatrix}}_{\boldsymbol{A}} \boldsymbol{p} = \underbrace{\begin{bmatrix} \operatorname{sign}(\boldsymbol{\sigma}_1) \\ \vdots \\ \operatorname{sign}(\boldsymbol{\sigma}_J) \end{bmatrix}}_{\boldsymbol{\lambda}}$$

The optimality condition ensures that the optimal $\boldsymbol{p}$ has the form

$$\boldsymbol{p} = \boldsymbol{A}^{\mathrm{H}} \boldsymbol{\alpha} = \sum_{j=1}^{J} \boldsymbol{E}_{f_j} \boldsymbol{S} \boldsymbol{\alpha}_j$$

for some $\boldsymbol{\alpha} = \begin{bmatrix} \boldsymbol{\alpha}_1^{\mathrm{H}} & \cdots & \boldsymbol{\alpha}_J^{\mathrm{H}} \end{bmatrix}^{\mathrm{H}}$ with $\boldsymbol{\alpha}_j \in \mathbb{C}^L$ satisfying the normal equation

$$\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}} \boldsymbol{\alpha} = \boldsymbol{\lambda}$$

which gives (when $\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}}$ is invertible)

$$\begin{aligned} \boldsymbol{\alpha} &= (\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}})^{-1} \boldsymbol{\lambda}, \\ \boldsymbol{p} &= \boldsymbol{A}^{\mathrm{H}} \boldsymbol{\alpha} = \boldsymbol{A}^{\mathrm{H}} (\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}})^{-1} \boldsymbol{\lambda}. \end{aligned}$$

With this form of $\boldsymbol{p}$, we obtain the dual polynomial

$$\boldsymbol{q}(f) = \boldsymbol{S}^{\mathrm{H}} \boldsymbol{E}_f^{\mathrm{H}} \boldsymbol{p} = \boldsymbol{S}^{\mathrm{H}} \boldsymbol{E}_f^{\mathrm{H}} \boldsymbol{A}^{\mathrm{H}} (\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}})^{-1} \boldsymbol{\lambda}. \quad (12)$$

To certify the optimality of computing the atomic norm (6), it remains to show $\|\boldsymbol{q}(f)\|_2 < 1,\ f \notin \{f_1, f_2, \ldots, f_J\}$.

A. Simple case: $\boldsymbol{S}^{\mathrm{H}}\boldsymbol{E}_f\boldsymbol{S} = \boldsymbol{0}$ if $|f| \geq 2W$

Lemma (1) shows that $\boldsymbol{S}$ is nearly orthogonal to $\boldsymbol{E}_f\boldsymbol{S}$ for any $|f| \geq 2W$. To simplify analysis, we could hypothetically assume $\boldsymbol{S}$ is exactly orthogonal to $\boldsymbol{E}_f\boldsymbol{S}$, i.e., $\boldsymbol{S}^{\mathrm{H}}\boldsymbol{E}_f\boldsymbol{S} = \boldsymbol{0}$ for any $|f| \geq 2W$. Under this case, the following result establishes conditions under which the dual polynomial obtained by (11) certifies the success of the atomic norm minimization.

Proposition 2. *Let $\boldsymbol{x} = \sum_j (\boldsymbol{S}\boldsymbol{\sigma}_j) \odot \boldsymbol{e}_{f_j}$ with $\|\boldsymbol{\sigma}_j\| > 0$. Additionally, assume $\operatorname{sign}(\boldsymbol{\sigma}_j)$ are drawn independently and identically (i.i.d.) from the uniform distribution on the complex unit circle and*

$$\Delta_f := \min_{j \neq j'} \operatorname{dist}(f_j - f_{j'}) \geq 4W,$$

where the distance $\operatorname{dist}(f_j - f_{j'})$ is understood as the warp-around distance on the unit circle. Also suppose $\boldsymbol{S}^{\mathrm{H}}\boldsymbol{E}_f\boldsymbol{S} = \boldsymbol{0}$ when $|f| \geq 2W$. Then with high probability, $\boldsymbol{x} = \sum_j (\boldsymbol{S}\boldsymbol{\sigma}_j) \odot \boldsymbol{e}_{f_j}$ is the unique atomic decomposition of $\boldsymbol{x}$.

Proof: Note that when $\boldsymbol{S}^{\mathrm{H}}\boldsymbol{E}_f\boldsymbol{S} = \boldsymbol{0}$ if $|f| \geq 2W$ and $\Delta_f \geq 4W$, we have

$$\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}} = \mathbf{I}$$

and the dual polynomial (12) becomes

$$\boldsymbol{q}(f) = \boldsymbol{S}^{\mathrm{H}}\boldsymbol{E}_f^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}(\boldsymbol{A}\boldsymbol{A}^{\mathrm{H}})^{-1}\boldsymbol{\lambda} = \sum_j \boldsymbol{S}^{\mathrm{H}}\boldsymbol{E}_{f-f_j}^{\mathrm{H}}\boldsymbol{S}\boldsymbol{\lambda}_j.$$

We first consider the case where f is away from the bands, i.e., $\operatorname{dist}(f, f_j) \geq 2W$ for all $j \in [J]$. Then

$$\boldsymbol{q}(f) = \sum_j \boldsymbol{S}^{\mathrm{H}}\boldsymbol{E}_{f-f_j}^{\mathrm{H}}\boldsymbol{S}\boldsymbol{\lambda}_j = \boldsymbol{0}$$

since $\boldsymbol{S}^{\mathrm{H}}\boldsymbol{E}_{f-f_j}^{\mathrm{H}}\boldsymbol{S} = \boldsymbol{0}$.

Now consider the case where f is within one band; that is $\operatorname{dist}(f - f_j) \leq 2W$ for some $j \in \{1, 2, \ldots, J\}$. In this case, we have

$$\boldsymbol{q}(f) = \boldsymbol{S}^{\mathrm{H}}\boldsymbol{E}_{f-f_j}^{\mathrm{H}}\boldsymbol{S}\operatorname{sign}(\boldsymbol{\sigma}_j)$$

since $\operatorname{dist}(f, f_{j'}) > 2W$ for all $j' \neq j$. Without loss of generality, let $f_j = 0$. To show $\|\boldsymbol{q}(f)\|_2 < 1$, we first show that $\sigma_{\min}(\boldsymbol{S}^{\mathrm{H}}\boldsymbol{E}_f^{\mathrm{H}}\boldsymbol{S})$ (the smallest singular value of $\boldsymbol{S}^{\mathrm{H}}\boldsymbol{E}_f^{\mathrm{H}}\boldsymbol{S}$) is strictly less than 1 for any $f \neq 0$. We prove this by contradiction. Note that both $\boldsymbol{S}$ and $\boldsymbol{E}_f\boldsymbol{S}$ are orthonormal matrices with L columns. This implies that $\sigma_{\min}(\boldsymbol{S}^{\mathrm{H}}\boldsymbol{E}_f\boldsymbol{S}) < 1$ as long as $\operatorname{Range}(\boldsymbol{S}) \neq \operatorname{Range}(\boldsymbol{E}_f\boldsymbol{S})$. Now we suppose $\operatorname{Range}(\boldsymbol{S}) = \operatorname{Range}(\boldsymbol{E}_f\boldsymbol{S})$ for some $f \neq 0$, which indicates that $\boldsymbol{E}_f\boldsymbol{S} = \boldsymbol{S}\boldsymbol{R}$ for some $L \times L$ orthonormal matrix $\boldsymbol{R}$. Then we have

$$\begin{aligned}\operatorname{Range}(\boldsymbol{E}_f(\boldsymbol{E}_f\boldsymbol{S})) &= \operatorname{Range}(\boldsymbol{E}_f\boldsymbol{S}\boldsymbol{R}) = \operatorname{Range}(\boldsymbol{E}_f\boldsymbol{S}) \\ &= \operatorname{Range}(\boldsymbol{S}),\end{aligned}$$

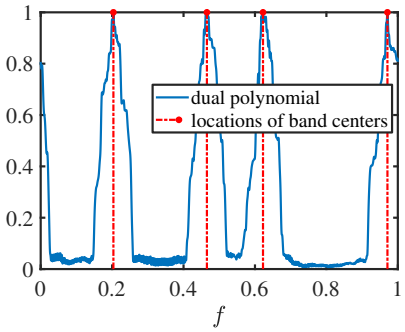

Figure 1. Dual polynomial $\widehat{q}(f)$ (see (9)) for multiband signal identification. Here, the dashed red lines represent the true frequencies, which are correctly localized by examining the peaks of the dual polynomial.

where the second equality follows because $\boldsymbol{R}$ is an orthonormal matrix. Similarly, we have

$$\operatorname{Range}(\underbrace{\boldsymbol{E}_f(\cdots(\boldsymbol{E}_f}_{\lceil \frac{2W}{f} \rceil}\boldsymbol{S}))) = \cdots = \operatorname{Range}(\boldsymbol{S}),$$

which contradicts to the fact $\boldsymbol{S}^{\mathrm{H}}\boldsymbol{E}_{f\lceil \frac{2W}{f} \rceil}^{\mathrm{H}}\boldsymbol{S} = \boldsymbol{0}$. The proof is completed by noting that

$$\|\boldsymbol{q}(f)\| = \|\boldsymbol{S}^{\mathrm{H}}\boldsymbol{E}_{f-f_j}^{\mathrm{H}}\boldsymbol{S}\operatorname{sign}(\boldsymbol{\sigma}_j)\| < 1$$

with high probability if $\operatorname{sign}(\boldsymbol{\sigma}_j)$ is drawn i.i.d. from the uniform distribution on the complex unit circle $\{\boldsymbol{\sigma} \in \mathbb{C}^L : \|\boldsymbol{\sigma}\| = 1\}$. ■

We note that Proposition 2 is premised on the idealized assumption that $\boldsymbol{S}^{\mathrm{H}}\boldsymbol{E}_f\boldsymbol{S} = \boldsymbol{0}$ for any $|f| \geq 2W$, which holds only approximately, not exactly. Thus, we save for future work the task of formalizing Proposition 2 using the DPSS basis properties outlined in Lemma 1.

V. Numerical Simulations

As an illustration, we test on a time-limited vector $\boldsymbol{x}$ obtained by setting $N = 256$ and sampling an analog signal that has a multiband spectrum. In this multiband spectrum, the range of active digital frequencies $\mathbb{W}$ is a finite union of $J = 4$ small intervals of width $W = \frac{1}{32}$ within the interval $\left[-\frac{1}{2}, \frac{1}{2}\right)$. We randomly generate each vector $\boldsymbol{g}_j$ as a linear combination of the first $L = 5$ DPSS vectors. Figure 1 displays the dual polynomial obtained by solving the atomic norm minimization problem. We observe that the band centers are correctly localized by the peaks of the dual polynomial. Figure 2 shows the pre-certificate dual polynomial obtained by solving (11).

Acknowledgment

This work was supported by NSF grants CCF-1409261, CCF-1464205, and CCF-1704204.

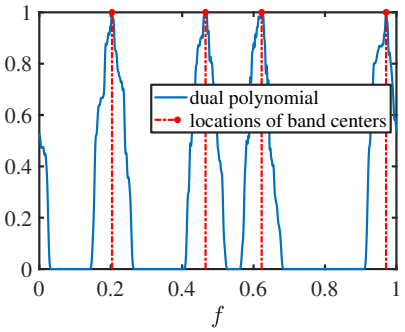

Figure 2. Dual polynomial constructed by solving (11).

REFERENCES

[1] E. Lagunas, M. Amin, F. Ahmad, and M. Najar, "Joint wall mitigation and compressive sensing for indoor image reconstruction," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, pp. 891–906, Feb 2013.

[2] Z. Zhu, G. Tang, P. Setlur, S. Gogineni, M. B. Wakin, and M. Rangaswamy, "Super-resolution in sar imaging: Analysis with the atomic norm," in *IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)*, pp. 1–5, 2016.

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

[4] 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.

[5] Q. Li and G. Tang, "Approximate support recovery of atomic line spectral estimation: A tale of resolution and precision," in *IEEE Global Conference on Signal and Information Processing*, 2016.

[6] 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.

[7] M. Wakin, S. Becker, E. Nakamura, M. Grant, E. Sovero, D. Ching, J. Yoo, J. Romberg, A. Emami-Neyestanak, and E. Candès, "A non-uniform sampler for wideband spectrally-sparse environments," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 2, no. 3, pp. 516–529, 2012.

[8] Z. Zhu and M. B. Wakin, "Wall clutter mitigation and target detection using discrete prolate spheroidal sequences," in *International Workshop on Compressed Sensing Theory and its Applications to Radar, Sonar and Remote Sensing (CoSeRa)*, 2015.

[9] Z. Zhu and M. B. Wakin, "On the dimensionality of wall and target return subspaces in through-the-wall radar imaging," in *International Workshop on Compressed Sensing Theory and its Applications to Radar, Sonar and Remote Sensing (CoSeRa)*, 2016.

[10] M. Davenport, S. Schnelle, J. Slavinsky, R. Baraniuk, M. Wakin, and P. Boufounos, "A wideband compressive radio receiver," in *Military Communications Conference (MILCOM)*, pp. 1193–1198, Oct 2010.

[11] M. A. Davenport and M. B. Wakin, "Reconstruction and cancellation of sampled multiband signals using Discrete Prolate Spheroidal Sequences," in *Proc of Workshop on Signal Processing with Adaptive Sparse Structured Representations (SPARS11)*, p. 61, 2011.

[12] M. A. Davenport and M. B. Wakin, "Compressive sensing of analog signals using Discrete Prolate Spheroidal Sequences," *Appl. Comput. Harmon. Anal.*, vol. 33, no. 3, pp. 438–472, 2012.

[13] D. Slepian, "Prolate Spheroidal Wave Functions, Fourier analysis, and uncertainty—V: The discrete case," *Bell Syst. Tech. J*, vol. 57, no. 5, pp. 1371–1430, 1978.

[14] Z. Zhu and M. B. Wakin, "Approximating sampled sinusoids and multiband signals using multiband modulated DPSS dictionaries," *J. Fourier Anal. Appl.*, vol. 23, pp. 1263–1310, Dec 2017.

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

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

[17] Y. Chi, "Guaranteed blind sparse spikes deconvolution via lifting and convex optimization," *IEEE J. Select. Topics Signal Process.*, vol. 10, no. 4, pp. 782–794, 2016.

[18] D. Yang, G. Tang, and M. B. Wakin, "Super-resolution of complex exponentials from modulations with unknown waveforms," *IEEE Trans. Inf. Theory*, vol. 62, pp. 5809–5830, Oct 2016.

[19] S. Karnik, Z. Zhu, M. B. Wakin, J. K. Romberg, and M. A. Davenport, "The fast Slepian transform," to appear in *Appl. Comp. Harm. Anal., arXiv preprint arXiv:1611.04950*.

[20] Z. Zhu, S. Karnik, M. B. Wakin, M. A. Davenport, and J. K. Romberg, "Fast orthogonal approximations of sampled sinusoids and bandlimited signals," in *IEEE Conf. Acous., Speech, Signal Process. (ICASSP)*, pp. 4511–4515, 2017.