Blind Super-resolution of Sparse Spike Signals

Yuejie Chi
Department of Electrical and Computer Engineering
The Ohio State University

***Abstract*—In many applications, the observations can be modeled as a linear combination of a small number of scaled and shifted copies of a bandlimited point spread function, either determined by the nature or designed by the users. Examples include neural spike trains, returns in radar and sonar, images in astronomy and single-molecule microscopy, etc. It is of great interest to resolve the spike signal as accurate as possible from the observation. When the point spread function is assumed unknown, this problem is terribly ill-posed. This paper proposes a convex optimization framework based on minimization of the atomic norm for jointly spectrally-sparse ensembles to simultaneously estimate the point spread function as well as the spike signal with provable performance guarantees, by mildly constraining the point spread function lies in a known low-dimensional subspace with an unknown orientation. Numerical examples are provided to validate the effectiveness of the proposed approach.**

I. Introduction

Many applications encounter low-resolution observations of a spike signal $x(t)$ by passing it through a point spread function (PSF) $g(t)$, which is often band-limited, either determined by the nature or designed by the users. Alternatively, the observation is a linear combination of a small number of scaled and delayed copies of the PSF, where all parameters are continuous-valued. When the PSF $g(t)$ is known perfectly, many algorithms can be applied to retrieve the spike signal, from subspace methods such as ESPRIT [1] to convex optimization approaches [2]. However, in many applications, the PSF is not known a priori, and must be estimated together with the spike model. For example, in neural spike train decoding, the PSF of the neurons is determined by the nature and is unknown. Another example is blind channel estimation in wireless communications, where the transmitted signal is modulated by unknown information symbols, therefore is also unknown. Moreover, a related problem is *blind calibration* for uniform linear arrays [3], where the PSF becomes the calibration vector of the array instead. Conventional approaches for blind deconvolution are typically based on expectation maximization [3], [4], which may suffer from local minima and lack performance guarantees.

We consider the problem of blind spike deconvolution, that is to jointly estimate the PSF and the spike model, which is extremely ill-posed without further constraints. In this paper, we make the assumption that the PSF lies in a known low-dimensional subspace without knowing its orientation in the subspace. This assumption turns out to be quite flexible and holds in a sizable number of applications, see [5] for examples. Inspired by the application of lifting in solving bilinear inverse problems in the recent literature [5], [6], [7], we reformulated the blind spike deconvolution problem into minimization of the atomic norm of multiple measurement vectors [8], [9], which can be solved efficiently via semidefinite programming. We demonstrate that the spike signal can be localized with an infinite precision via constructing the dual polynomial of the proposed atomic norm minimization algorithm, without applying a discretization procedure. In particular, it is guaranteed that the proposed algorithm succeeds with high probability as long as the number of measurements is on the order of K^2L up to logarithmic factors, where K is the number of spikes, and L is the subspace dimension of the PSF, as long as the spikes are mildly separated, and the rows of the generating subspace of the PSF is drawn from a distribution obeying certain incoherence condition. Numerical examples are provided to corroborate our findings. Compared with [5], [6], [10], our work does not assume the spike signal to lie in a *known* subspace or sparsity in some predetermined dictionary, to avoid the consequence of basis mismatch [11]. This makes both the algorithms and performance analyses in [5], [6] inapplicable to our setting.

The rest of the paper is organized as follows. Section II describes the problem formulation. Section III describes the proposed algorithm and its performance analysis. Section IV provides numerical examples to demonstrate the performance of the proposed algorithms. Finally, we conclude in Section V.

II. Problem Formulation

Let the observation be

$$y(t) = x(t) * g(t) = \sum_{k=1}^{K} a_k g(t-\bar{\tau}_k), \tag{1}$$

where $*$ denotes the convolution. Here the spike signal $x(t)$ is given as $x(t)=\sum_{k=1}^{K} a_k\delta(t-\bar{\tau}_k)$, where K is the number of spikes, $a_k\in\mathbb{C}$ and $\bar{\tau}_k\in[0,T_{\max})$ are the complex amplitude and delay of the kth spike respectively, $1\leq k\leq K$, $T_{\max}$ is the maximum allowable delay. The bandwidth of the PSF $g(t)$ is denoted as $[-B_{\max}, B_{\max}]$. Taking the Fourier transform of (1), we have

$$Y(f) = X(f)G(f) = \left(\sum_{k=1}^{K} a_k e^{-j2\pi f\bar{\tau}_k}\right)\cdot G(f), \tag{2}$$

Email: chi.97@osu.edu. This work is supported in part by ONR under grant N00014-15-1-2387 and by NSF under grant CCF-1527456.

where $f \in [-B_{\max}, B_{\max}]$. By sampling (2) at $f_n = \frac{B_{\max} n}{2M}$, $n = -2M, \ldots, 2M$, we have

$$y_n = Y(f_n) = \left(\sum_{k=1}^{K} a_k e^{-j2\pi n \tau_k \frac{B_{\max} T_{\max}}{2M}} \right) \cdot G(f_n), \quad (3)$$

where $\tau_k = \bar{\tau}_k / T_{\max} \in [0, 1)$. From (3), it is straightforward to see that the number of samples needs to satisfy $2M \geq B_{\max} T_{\max}$ so that the delays $\tau_k \in [0, 1)$ can be uniquely identified. Without loss of generality, we will assume $2M = B_{\max} T_{\max}$, and consider the normalized delays $\tau_k \in [0, 1)$ in this paper, and denote $\mathcal{T} = \{\tau_k, 1 \leq k \leq K\}$ as the set of spike locations. We're interested in identification of the spike signal using as small M as possible.

We can now rewrite (3) as

$$y_n = g_n \cdot \left(\sum_{k=1}^{K} a_k e^{-j2\pi n \tau_k} \right) := g_n \cdot x_n, \quad (4)$$

where $g_n = G(f_n)$ and $x_n = \sum_{k=1}^{K} a_k e^{-j2\pi n \tau_k}$, $n = -2M, \ldots, 2M$. Denote $\mathcal{T} = \{\tau_k\}_{k=1}^{K}$. In the matrix form, we denote

$$\boldsymbol{y} = \operatorname{diag}(\boldsymbol{g}) \boldsymbol{x} = \operatorname{diag}(\boldsymbol{g}) \boldsymbol{V} \boldsymbol{a}, \quad (5)$$

where $\boldsymbol{y} = [y_{-2M}, \ldots, y_{2M}]^T$, $\boldsymbol{x} = [x_{-2M}, \ldots, x_{2M}]^T$, $\boldsymbol{g} = [g_{-2M}, \ldots, g_{2M}]^T$, $\boldsymbol{V} \in \mathbb{C}^{(4M+1) \times K}$ is the Vandermonde matrix with columns determined by τ_k's and $\boldsymbol{a} = [a_1, a_2, \ldots, a_K]^T$.

III. AtomicLift for Blind Spike Deconvolution

We consider the problem of jointly estimating the PSF and the spike model. This problem is extremely ill-posed without further constraints. In this paper, inspired by [5], we make the assumption that $\boldsymbol{g}$ lies in a known low-dimensional subspace without knowing its orientation in the subspace, given as $\boldsymbol{g} = \boldsymbol{B}\boldsymbol{h}$, where $\boldsymbol{B} \in \mathbb{C}^{(4M+1) \times L}$ is known, $\boldsymbol{h} \in \mathbb{C}^L$ is unknown, and L is much smaller than $4M+1$. This assumption turns out to be quite flexible and holds in a sizable number of applications, see [5] for examples.

Denote $\boldsymbol{B}^T = [\boldsymbol{b}_{-2M}, \ldots, \boldsymbol{b}_{2M}]$, where $\boldsymbol{b}_n \in \mathbb{C}^L$ is the nth column of the matrix $\boldsymbol{B}^T$. Using the lifting trick [5], [6], [7], we can rewrite y_n in (4) as

$$\begin{aligned} y_n &= \boldsymbol{b}_n^T \boldsymbol{h} x_n = \boldsymbol{b}_n^T \boldsymbol{h} \boldsymbol{e}_n^T \boldsymbol{x} \\ &= \boldsymbol{e}_n^T (\boldsymbol{x}\boldsymbol{h}^T) \boldsymbol{b}_n := \boldsymbol{e}_n^T \boldsymbol{Z}^\star \boldsymbol{b}_n = \langle \boldsymbol{Z}^\star, \boldsymbol{e}_n \boldsymbol{b}_n^H \rangle, \end{aligned} \quad (6)$$

where $\boldsymbol{e}_n$ is the n-th standard basis vector of $\mathbb{R}^{4M+1}$, and $\langle \boldsymbol{Y}, \boldsymbol{X} \rangle = \operatorname{Tr}(\boldsymbol{X}^H \boldsymbol{Y})$. Therefore, $\boldsymbol{y}$ can be regarded a set of linear measurements of the rank-one matrix:

$$\boldsymbol{Z}^\star = \boldsymbol{x}\boldsymbol{h}^T = \boldsymbol{V}\boldsymbol{a}\boldsymbol{h}^T = \sum_{k=1}^{K} a_k \boldsymbol{c}(\tau_k) \boldsymbol{h}^T \in \mathbb{C}^{(4M+1) \times L},$$

i.e. $\boldsymbol{y} = \mathcal{X}(\boldsymbol{Z}^\star)$, where $\mathcal{X}$ denotes the operator that performs the linear mapping (6).

The problem is now to recover the matrix $\boldsymbol{Z}^\star$ from $\boldsymbol{y}$, where it appears the number of unknowns is much more than the number of measurements. Our key observation is that each column of $\boldsymbol{Z}^\star$ is composed of K complex sinusoids with the same frequencies determined by $\boldsymbol{V}$. We then propose to motivate this joint spectral sparsity of $\boldsymbol{Z}^\star$ by minimizing the recently proposed atomic norm for jointly spectrally-sparse ensembles [8].

To proceed, define the set of atoms as

$$\mathcal{A} = \left\{ \boldsymbol{A}(\tau, \boldsymbol{u}) = \boldsymbol{c}(\tau) \boldsymbol{u}^H \in \mathbb{C}^{(4M+1) \times L} | \tau \in [0, 1), \|\boldsymbol{u}\|_2 = 1 \right\},$$

where

$$\boldsymbol{c}(\tau) = \frac{1}{\sqrt{4M+1}} \left[e^{-j2\pi(-2M)\tau}, \ldots, 1, \ldots, e^{-j2\pi(2M)\tau} \right]^T$$

represents a complex sinusoid with the frequency $\tau \in [0, 1)$. The atomic norm of a matrix $\boldsymbol{Z} \in \mathbb{C}^{(4M+1) \times L}$ is then [8]

$$\begin{aligned} \|\boldsymbol{Z}\|_{\mathcal{A}} &= \inf \{ t > 0 : \ \boldsymbol{Z} \in t \operatorname{conv}(\mathcal{A}) \} \\ &= \inf \left\{ \sum_k c_k \middle| \boldsymbol{Z} = \sum_k c_k \boldsymbol{A}(\tau_k, \boldsymbol{u}_k), c_k \geq 0 \right\}, \end{aligned} \quad (7)$$

where $\operatorname{conv}(\mathcal{A})$ is the convex hull of $\mathcal{A}$. Moreover, $\|\boldsymbol{Z}\|_{\mathcal{A}}$ admits an equivalent semidefinite programming (SDP) characterization [8] which can be computed efficiently using off-the-shelf solvers:

$$\|\boldsymbol{Z}\|_{\mathcal{A}} = \inf_{\boldsymbol{u} \in \mathbb{C}^n, \boldsymbol{W} \in \mathbb{C}^{L \times L}} \left\{ \frac{1}{2} \operatorname{Tr}(\operatorname{toep}(\boldsymbol{u})) + \frac{1}{2} \operatorname{Tr}(\boldsymbol{W}) \middle| \begin{bmatrix} \operatorname{toep}(\boldsymbol{u}) & \boldsymbol{Z} \\ \boldsymbol{Z}^H & \boldsymbol{W} \end{bmatrix} \succeq \boldsymbol{0} \right\},$$

where $\operatorname{toep}(\boldsymbol{u})$ is the Toeplitz matrix with $\boldsymbol{u}$ as the first column. We then propose the following *AtomicLift* algorithm to motivate the joint spectral sparsity $\boldsymbol{Z}$ obeying the measurements:

$$\min \ \|\boldsymbol{Z}\|_{\mathcal{A}} \quad \text{s. t.} \quad \boldsymbol{y} = \mathcal{X}(\boldsymbol{Z}). \quad (8)$$

A. Spike Localization via Dual Polynomials

Define $\langle \boldsymbol{Y}, \boldsymbol{X} \rangle_{\mathbb{R}} = \operatorname{Re}(\langle \boldsymbol{Y}, \boldsymbol{X} \rangle)$. The dual norm of $\|\cdot\|_{\mathcal{A}}$ can be defined as

$$\begin{aligned} \|\boldsymbol{Y}\|_{\mathcal{A}}^\star &= \sup_{\|\boldsymbol{X}\|_{\mathcal{A}} \leq 1} \langle \boldsymbol{Y}, \boldsymbol{X} \rangle_{\mathbb{R}} = \sup_{\tau \in [0,1), \|\boldsymbol{u}\|_2 = 1} \langle \boldsymbol{Y}, \boldsymbol{c}(\tau) \boldsymbol{u}^H \rangle_{\mathbb{R}} \\ &= \sup_{\tau \in [0,1)} \left\| \boldsymbol{Y}^H \boldsymbol{c}(\tau) \right\|_2. \end{aligned}$$

Then the dual problem of (8) can thus be written as

$$\max_{\boldsymbol{p} \in \mathbb{C}^{4M+1}} \ \langle \boldsymbol{p}, \boldsymbol{y} \rangle_{\mathbb{R}} \quad \text{s.t.} \quad \|\mathcal{X}^*(\boldsymbol{p})\|_{\mathcal{A}}^\star \leq 1, \quad (9)$$

where $\mathcal{X}^*(\boldsymbol{p}) = \sum_{n=-2M}^{2M} p_n \boldsymbol{e}_n \boldsymbol{b}_n^H$. Let $(\boldsymbol{Z}, \boldsymbol{p})$ be primal-dual feasible to (8) and (9), we have $\langle \mathcal{X}^*(\boldsymbol{p}), \boldsymbol{Z} \rangle_{\mathbb{R}} = \langle \mathcal{X}^*(\boldsymbol{p}), \boldsymbol{Z}^\star \rangle_{\mathbb{R}}$. The solution of the dual problem (9) can be used to localize the spikes. Write the vector-valued dual polynomial $\boldsymbol{Q}(\tau) \in \mathbb{C}^L$ as

$$\boldsymbol{Q}(\tau) = (\mathcal{X}^*(\boldsymbol{p}))^H \boldsymbol{c}(\tau) = \frac{1}{\sqrt{4M+1}} \sum_{n=-2M}^{2M} e^{-j2\pi\tau n} p_n \boldsymbol{b}_n,$$

then the spikes can be localized as the peak of $\|\boldsymbol{Q}(\tau)\|_2$:

$$\hat{\mathcal{T}} = \{ \tau \in [0, 1) | \, \|\boldsymbol{Q}(\tau)\|_2 = 1 \}.$$

B. Optimality Condition

We provide a sufficient condition for optimality of the proposed algorithm (8).

Proposition 1. *The solution to* (8) *is unique if there exists a vector* $\boldsymbol{q} \in \mathbb{C}^{4M+1}$ *such that the vector-valued dual polynomial* $\boldsymbol{Q}(\tau) = (\mathcal{X}^*(\boldsymbol{q}))^H \boldsymbol{c}(\tau) = \boldsymbol{B}^T \operatorname{diag}(\boldsymbol{q})\boldsymbol{c}(\tau) = \frac{1}{\sqrt{4M+1}} \sum_{n=-2M}^{2M} e^{-j2\pi\tau n} q_n \boldsymbol{b}_n \in \mathbb{C}^L$ *satisfies*

$$\boldsymbol{Q}(\tau_k) = \frac{1}{\|\boldsymbol{h}\|_2^2} sign(a_k^*)\boldsymbol{h}^* \quad \forall \tau_k \in \mathcal{T}, \quad and \tag{10}$$

$$\|\boldsymbol{Q}(\tau)\|_2 < 1, \quad \forall \tau \in [0,1] \backslash \mathcal{T}, \tag{11}$$

where $\boldsymbol{h}^*$ *is the conjugate of* $\boldsymbol{h}$*, and the sign*$(\cdot)$ *is the complex sign function.*

Proof: First, any $\boldsymbol{q}$ satisfying (10) and (11) is dual feasible. We have

$$\begin{aligned}
\|\boldsymbol{Z}^\star\|_{\mathcal{A}} &\geq \|\boldsymbol{Z}^\star\|_{\mathcal{A}} \|\mathcal{X}^*(\boldsymbol{q})\|_{\mathcal{A}}^\star \geq \langle \mathcal{X}^*(\boldsymbol{q}), \boldsymbol{Z}^\star \rangle_{\mathbb{R}} \\
&= \left\langle \mathcal{X}^*(\boldsymbol{q}), \sum_{k=1}^K a_k \boldsymbol{c}(\tau_k)\boldsymbol{h}^T \right\rangle_{\mathbb{R}} \\
&= \sum_{k=1}^K \operatorname{Re}\left(a_k^* \left\langle \mathcal{X}^*(\boldsymbol{q}), \boldsymbol{c}(\tau_k)\boldsymbol{h}^T \right\rangle\right) \\
&= \sum_{k=1}^K \operatorname{Re}\left(a_k^* \langle \boldsymbol{Q}^H(\tau_k), \boldsymbol{h}^T \rangle\right) \\
&= \sum_{k=1}^K \operatorname{Re}\left(a_k^* \operatorname{sign}(a_k)\right) = \sum_{k=1}^K |a_k| \geq \|\boldsymbol{Z}^\star\|_{\mathcal{A}}.
\end{aligned}$$

Hence $\langle \mathcal{X}^*(\boldsymbol{q}), \boldsymbol{Z}^\star \rangle_{\mathbb{R}} = \|\boldsymbol{Z}^\star\|_{\mathcal{A}}$. By strong duality we have $\boldsymbol{Z}^\star$ is primal optimal and $\boldsymbol{q}$ is dual optimal.

For uniqueness, suppose $\hat{\boldsymbol{Z}}$ is another optimal solution. If $\hat{\boldsymbol{Z}}$ and $\boldsymbol{Z}^\star$ have the same support set $\mathcal{T}$, they must coincide since the set of atoms in $\mathcal{T}$ is independent. Let $\hat{\boldsymbol{Z}} = \sum_k \hat{a}_k \boldsymbol{c}(\hat{\tau}_k)\hat{\boldsymbol{h}}_k^T$ be its atomic decomposition where $\hat{a}_k > 0$, with some support $\hat{\tau}_k \notin \mathcal{T}$. We then have

$$\begin{aligned}
&\langle \mathcal{X}^*(\boldsymbol{q}), \hat{\boldsymbol{Z}} \rangle_{\mathbb{R}} \\
&= \sum_{\hat{\tau}_k \in \mathcal{T}} \operatorname{Re}\left(\hat{a}_k^* \langle \boldsymbol{Q}^H(\hat{\tau}_k), \hat{\boldsymbol{h}}_k^T \rangle\right) + \sum_{\hat{\tau}_l \notin \mathcal{T}} \operatorname{Re}\left(\hat{a}_l^* \langle \boldsymbol{Q}^H(\hat{\tau}_l), \hat{\boldsymbol{h}}_l^T \rangle\right) \\
&\leq \sum_{\hat{\tau}_k \in \mathcal{T}} \hat{a}_k \|\boldsymbol{Q}(\hat{\tau}_k)\|_2 \|\hat{\boldsymbol{h}}_k\|_2 + \sum_{\hat{\tau}_l \notin \mathcal{T}} \hat{a}_l \|\boldsymbol{Q}(\hat{\tau}_l)\|_2 \|\hat{\boldsymbol{h}}_l\|_2 \\
&< \sum_{\hat{\tau}_k \in \mathcal{T}} \hat{a}_k \|\hat{\boldsymbol{h}}_k\|_2 + \sum_{\hat{\tau}_l \notin \mathcal{T}} \hat{a}_l \|\hat{\boldsymbol{h}}_l\|_2 = \|\hat{\boldsymbol{Z}}\|_{\mathcal{A}},
\end{aligned}$$

which contradicts strong duality. Therefore the optimal solution of (8) is unique. ∎

C. Performance Guarantees

If we assume each row of the subspace $\boldsymbol{B}$ are drawn from some distribution with incoherence properties, together with a mild separation condition for the spike signal, the proposed AtomicLift algorithm provably recovers the spike signal with high probability as long as M is large enough. Specifically, we assume each row of $\boldsymbol{B}$ is sampled independently and identically from a population F, i.e. $\boldsymbol{b}_n \sim F$. Furthermore, we require F satisfies the following properties:

- Isometry property: We say F satisfies the isometry property if $\mathbb{E}\boldsymbol{b}\boldsymbol{b}^H = \boldsymbol{I}_L$, $\boldsymbol{b} \sim F$.
- Incoherence property: for $\boldsymbol{b} = [b_1, \ldots, b_L]^T \sim F$, define the coherence parameter μ of F as the smallest number that $\max_{1 \leq i \leq L} |b_i|^2 \leq \mu$ holds.

The above conditions are adopted from [12] which are employed in the development of a RIPless theory of compressed sensing. In particular, the incoherence parameter μ is a deterministic bound on the maximum entry of $\boldsymbol{b}$. Furthermore, define the separation of the spike signal as $\Delta = \min_{1 \leq i < j \leq K} |\tau_i - \tau_j|$, where the difference is taken as the wrap-around distance on the unit circle. The performance guarantee is presented in Theorem 1.

Theorem 1. *Let* $M \geq 4$*. Assume* $\boldsymbol{g}$ *lies in a random subspace* $\boldsymbol{B}$ *whose rows are sampled i.i.d. from a population* F *satisfying the isometry property and the incoherence property, with the coherence parameter* μ*. If* $\Delta \geq 1/M$*, then there exists a numerical constant* C *such that*

$$M \geq C\mu K^2 L \log^2\left(\frac{M}{\delta}\right)$$

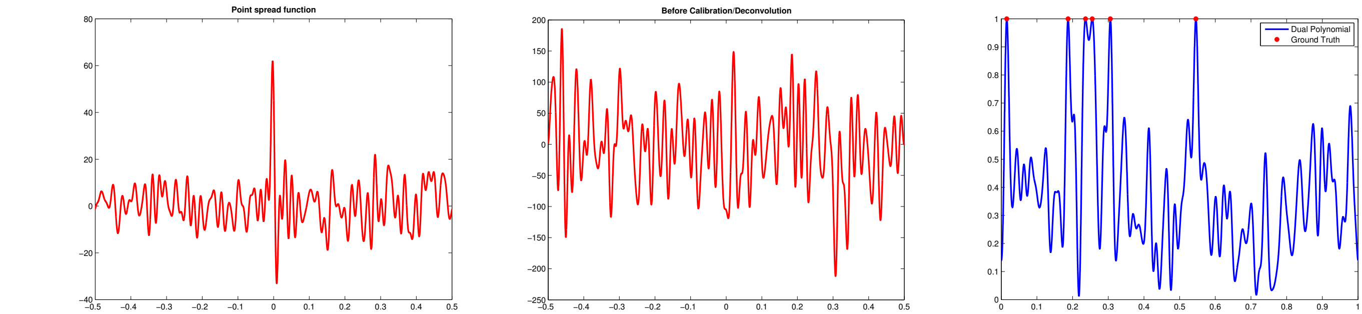
is sufficient to guarantee that we can recover $\boldsymbol{Z}^\star = \boldsymbol{x}\boldsymbol{h}^T$ *via the AtomicLift algorithm with probability at least* $1 - \delta$*.*

Theorem 1 allows $\boldsymbol{g}$ to have arbitrary orientation in the subspace $\boldsymbol{B}$, and applies to any *deterministic* spike signal as long as it satisfies the separation condition $\Delta \geq 1/M$. This is the same as required by Candès and Fernandez-Granda [13] for spikes deconvolution using total variation minimization even when the PSF is known perfectly. Therein they established that $N = O(K)$ measurements are sufficient to exactly recover the spike signal. In comparison, our performance guarantee is probabilistic that holds with high probability.

Theorem 1 suggests that as long as M is on the order of $O(K^2L)$ up to logarithmic factors, AtomicLift provably recovers the spike signal with high probability, as long as the conditions in Theorem 1 are satisfied. If L is a small constant independent of K and M, our bound simplifies to $M/\log^2 M \gtrsim O(K^2)$, which suggests blind spike deconvolution is possible at a cost of more measurements. Due to space limits, the proof of this theorem is presented elsewhere [14].

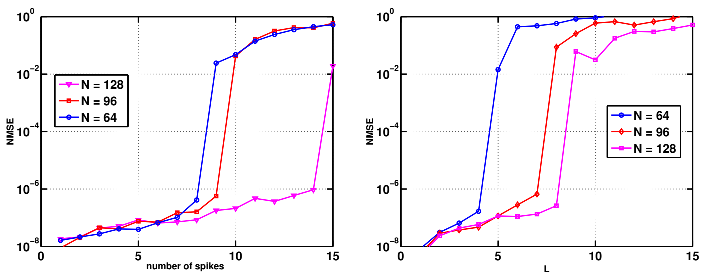
IV. Numerical Experiments

We perform a series of numerical experiments to validate the performance of the proposed algorithm which is implemented with MOSEK in the CVX toolbox [15]. Consider a sensor array of size N. We first randomly generate the spike locations uniformly at random, respecting the minimum separation $\Delta \geq 1/N$, whose coefficients are generated with a dynamic range of 10dB and random phase selected with a uniform distribution from $[0, 2\pi]$. The subspace of the PSF $\boldsymbol{B} \in \mathbb{R}^{N \times L}$ is generated with i.i.d. standard Gaussian entries, and the coefficient vector



(a) PSF (b) Observation (d) Localization

Fig. 2: Blind spikes deconvolution: (a) PSF; (b) convolution between the PSF in (a) and a sparse spike signal; (c) perfect localization.



(a) (b)

Fig. 1: The NMSE with respect to (a) the number of spikes K for a fixed $L = 3$ and (b) the subspace dimension L for a fixed $K = 5$ for various $N = 64, 96, 128$.

$\boldsymbol{h} \in \mathbb{R}^L$ is also generated with i.i.d. standard Gaussian entries. Fig. 1 shows the normalized mean squared error (NMSE) $\|\hat{\boldsymbol{Z}} - \boldsymbol{Z}^*\|_{\mathrm{F}}^2 / \|\boldsymbol{Z}^*\|_{\mathrm{F}}^2$, where $\hat{\boldsymbol{Z}}$ is the returned estimate of the lifted matrix $\boldsymbol{Z}^* = \boldsymbol{x}\boldsymbol{h}^T$. with respect to (a) the number of spikes K for a fixed $L = 3$ and (b) the subspace dimension L for a fixed $K = 5$ for various $N = 64, 96, 128$.

Fig. 2 (a) shows an example PSF in the time domain, where $\boldsymbol{g}$ is randomly generated from the subspace model, where the rows are given as $\boldsymbol{b}_i = [1, e^{j2\pi\alpha}, \cdots, e^{j2\pi(L-1)\alpha}]^T$, $\alpha \in [0, 1]$ uniformly generated at random. Its convolution with a spike signal containing $K = 6$ spikes is shown in Fig. 2 (b). Fig. 2 (c) demonstrates that the dual polynomial of the proposed algorithm can be used to perfectly localize the spikes.

V. Conclusions

This paper proposes a convex optimization framework based on minimization of the atomic norm for jointly spectrally-sparse ensembles to simultaneously estimate the point spread function as well as the spike signal with provable performance guarantees, by mildly constraining the point spread function lies in a known low-dimensional subspace with an unknown orientation. Compared with the sample complexity requirement $M \gtrsim O(K)$ when the PSF is perfectly known [2], our bounds require $M/\log^2 M \gtrsim O(K^2)$ samples. It remains an interesting future research direction to further reduce the sample complexity of blind super-resolution.

References

[1] R. Roy and T. Kailath, "ESPRIT-estimation of signal parameters via rotational invariance techniques," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 37, no. 7, pp. 984 –995, Jul 1989.

[2] 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.

[3] B. Friedlander and A. Weiss, "Eigenstructure methods for direction finding with sensor gain and phase uncertainties," in *Acoustics, Speech, and Signal Processing, 1988. ICASSP-88., 1988 International Conference on*, Apr 1988, pp. 2681–2684 vol.5.

[4] M. Viberg and A. Swindlehurst, "A bayesian approach to auto-calibration for parametric array signal processing," *Signal Processing, IEEE Transactions on*, vol. 42, no. 12, pp. 3495–3507, Dec 1994.

[5] A. Ahmed, B. Recht, and J. Romberg, "Blind deconvolution using convex programming," *Information Theory, IEEE Transactions on*, vol. 60, no. 3, pp. 1711–1732, 2014.

[6] S. Ling and T. Strohmer, "Self-calibration and biconvex compressive sensing," *arXiv preprint arXiv:1501.06864*, 2015.

[7] E. J. Candes, T. Strohmer, and V. Voroninski, "Phaselift: Exact and stable signal recovery from magnitude measurements via convex programming," *Communications on Pure and Applied Mathematics*, 2012.

[8] Y. Chi, "Joint sparsity recovery for spectral compressed sensing," in *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*. IEEE, 2014, pp. 3938–3942.

[9] Y. Li and Y. Chi, "Off-the-grid line spectrum denoising and estimation with multiple measurement vectors," *arXiv preprint arXiv:1408.2242*, 2014.

[10] K. Lee, Y. Li, M. Junge, and Y. Bresler, "Stability in blind deconvolution of sparse signals and reconstruction by alternating minimization," in *Int. Conf. Sampling Theory and Applications (SampTA)*, 2015.

[11] Y. Chi, L. Scharf, A. Pezeshki, and A. Calderbank, "Sensitivity to basis mismatch in compressed sensing," *IEEE Transactions on Signal Processing*, vol. 59, no. 5, pp. 2182–2195, May 2011.

[12] E. Candes and Y. Plan, "A probabilistic and RIPless theory of compressed sensing," *IEEE Transactions on Information Theory*, vol. 57, no. 11, pp. 7235–7254, 2011.

[13] 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.

[14] Y. Chi, "Guaranteed blind sparse spikes deconvolution via lifting and convex optimization," *arXiv preprint arXiv:1506.02751*, 2015.

[15] M. Grant, S. Boyd, and Y. Ye, "Cvx: Matlab software for disciplined convex programming," *Online accessiable: http://stanford. edu/~ boyd/cvx*, 2008.