NON-STATIONARY BLIND SUPER-RESOLUTION

Dehui Yang, Gongguo Tang and Michael B. Wakin

Dept. of Electrical Engineering and Computer Science
Colorado School of Mines

ABSTRACT

In this paper, we propose a new framework for parameter estimation of complex exponentials from their modulations with unknown waveforms via convex programming. Our model generalizes the recently developed blind sparse spike deconvolution framework by Y. Chi [1] to the non-stationary scenario and encompasses a wide spectrum of applications. Under the assumption that the unknown waveforms live in a common random subspace, we recast the problem into an atomic norm minimization framework by a lifting trick, and this problem can be solved using computationally efficient semidefinite programming. We show that the number of measurements for exact recovery is proportional to the number of degrees of freedom in the problem, up to polylogarithmic factors. Numerical experiments support our theoretical findings.

Index Terms— Atomic norm, super-resolution, semidefinite programming, blind deconvolution

1. INTRODUCTION

1.1. Motivation

Many real-world problems involve the recovery of unknown complex exponentials from their modulations with unknown waveforms. Mathematically, we consider the following model:

$$\boldsymbol{y}(n) = \sum_{j=1}^{J} c_j e^{-i2\pi n\tau_j} \boldsymbol{g}_j(n), \tag{1}$$

where $\{\boldsymbol{y}(n)\}$ are samples of a continuous-time output, $\{c_j\} \subset \mathbb{C}$ and $\{\tau_j\} \subset [0,1)$ are unknown parameters of the complex exponentials that we aim to recover, and $\boldsymbol{g}_j(n)$ are samples of the unknown waveforms, whose forms may vary with the index j. This model arises in a variety of applications, ranging from single-molecule imaging in biology [2, 3] , to radar signal processing [4, 5], to nuclear magnetic resonance spectroscopy [6]. Here we list three stylized applications, where the problems can be modeled using our general mathematical framework (1).

This work was supported by NSF CAREER Grant CCF-1149225, NSF Grant CCF-1464205, NSF Grant CCF-1409258, and NSF Grant CCF-1409261.

Super-resolution with unknown point spread functions: In applications like single-molecule imaging, we are interested in resolving the unknown locations of point sources from their convolutions with unknown point spread functions, whose forms may depend on the locations of the point sources. This is the case in three-dimensional single-molecule imaging [3], where the point spread function depends on the depth (z-axis) of the target. Other applications include non-stationary deconvolution of seismic data [7], computational photography [8], astronomy [9], etc. Mathematically speaking, we hope to determine the unknown delays $\{\tau_j\}$ and amplitudes $\{c_j\}$ from

$$y(t) = \sum_{j=1}^{J} c_j \delta(t-\tau_j) * g_{\tau_j}(t). \tag{2}$$

with g_{τ_j} being the unknown point spread functions. By taking the Fourier transform of both sides of (2), we have

$$\widehat{y}(f) = \sum_{j=1}^{J} c_j e^{-i2\pi f\tau_j} \widehat{g}_{\tau_j}(f). \tag{3}$$

It is easy to see that (1) is a sampled version of (3). With the samples $\boldsymbol{y}(n)$, we hope to simultaneously recover $\{c_j, \tau_j\}$ and samples of the point spread functions $g_{\tau_j}(f)$.

Parameter estimation in radar imaging: In radar imaging, given the input probing signal $x(t)$ transmitted by the radar, the output $y(t)$ can be characterized by the following relationship [4]

$$y(t) = \sum_{j=1}^{J} c_j e^{i2\pi\nu_j t} x(t-\mu_j), \tag{4}$$

where (μ_j, ν_j) are delay-Doppler pairs, which capture the distances and velocities of the J point targets relative to the radar. It is easy to see that (1) can also be obtained by sampling (4).

Frequency estimation with damping: In applications such as nuclear magnetic resonance spectroscopy [6], the signal is the superposition of complex exponentials with unknown frequencies and unknown damping factors. More precisely, the signal has the form:

$$y(t) = \sum_{j=1}^{J} c_j e^{i2\pi f_j t} e^{-\xi_j t} \tag{5}$$

with parameters $\{f_j, \xi_j\}$. In this situation, resolving the f_j's becomes difficult since the energy of the signal $y(t)$ attenuates as time moves forward. By sampling the continuous variable t, one again obtains an instance of (1).

1.2. Contributions and related work

In this work, we study the recovery of the unknown parameters ($\{c_j\}$ and $\{\tau_j\}$) associated with the complex exponentials and the recovery of the samples of the unknown waveforms, i.e., $\{\boldsymbol{g}_j\}$, from $\boldsymbol{y}$ by assuming that $\{\boldsymbol{g}_j\}$ lie in a known subspace with dimension K. Under randomness and incoherence assumptions on the subspace and minimum separation conditions of τ_j's, we show that $O(JK)$ measurements suffice to retrieve $\{\tau_j\}$ exactly and recover $\{c_j\}$ and $\{\boldsymbol{g}_j\}$ up to a multiplicative factor via a computationally efficient convex program. Note that our sample complexity matches the number of degrees of freedom in the problem and, thus, cannot be further improved.

Much of the recent work on super-resolution has involved convex optimization since the foundational work [10, 11]. Our proposed model (1) is a natural generalization of that by Y. Chi [1], where the author considers the super-resolution of point sources with an unknown point spread function that lives in a known low-dimensional subspace. Our model allows for non-stationary blind super-resolution, where $\boldsymbol{g}_j$ may change over the index of j within the subspace. We show that non-stationary blind super-resolution is possible when the number of measurements exceeds $O(JK)$. Furthermore, as a special case of our model, we improve the bound in [1] to $O(JK)$ when the unknown point spread function does not vary with j.

The rest of the paper is organized as follows. Section 2 introduces our problem setup, its connection to the atomic norm minimization framework via a lifting trick, and an exact semidefinite programming (SDP) reformulation of the atomic norm minimization [11]. Section 3 presents the main theorem and discusses its implications. Section 4 performs numerical experiments to support our theoretical findings. Conclusion is given in Section 5.

2. PROBLEM SETUP AND LIFTING

2.1. Problem formulation via atomic norm minimization

Consider the model

$$\boldsymbol{y}(n) = \sum_{j=1}^{J} c_j e^{-i2\pi n\tau_j} \boldsymbol{g}_j(n) \quad n = -2M, \cdots, 2M, \quad (6)$$

where $\boldsymbol{y}(n)$ are samples and $\{c_j, \tau_j\}$ and $\{\boldsymbol{g}_j(n)\}$ are unknowns. Without loss of generality, we assume that $\tau_j \in [0,1),\ j = 1, \cdots, J$. Note that this problem is severely ill-posed without additional constraints on $\boldsymbol{g}_j$ since the number of samples in (6) is $N := 4M + 1$, while the number of unknowns in (6) is $JN + 2J$. To make the problem well-posed, we assume that each $\boldsymbol{g}_j$ lives in a common low-dimensional subspace spanned by the columns of a known $N \times K$ matrix $\boldsymbol{B} = \begin{bmatrix} \boldsymbol{b}_{-2M} & \boldsymbol{b}_{-2M+1} & \cdots & \boldsymbol{b}_{2M-1} & \boldsymbol{b}_{2M} \end{bmatrix}^H$, where each $\boldsymbol{b}_n \in \mathbb{C}^{K\times 1}$. In other words, $\boldsymbol{g}_j = \boldsymbol{B}\boldsymbol{h}_j$, where each $\boldsymbol{h}_j \in \mathbb{C}^{K\times 1}$ is unknown. Thus, we can rewrite (6) equivalently as

$$\boldsymbol{y}(n) = \sum_j c_j e^{-i2\pi n\tau_j} \boldsymbol{b}_n^H \boldsymbol{h}_j,$$

which only involves $JK + 2J$ unknowns.
Denoting

$$\boldsymbol{a}(\tau) = \begin{bmatrix} e^{i2\pi(-2M)\tau} & \cdots & 1 & \cdots & e^{i2\pi(2M)\tau} \end{bmatrix}^T,$$

we have

$$\begin{aligned} \boldsymbol{y}(n) &= \sum_j c_j \boldsymbol{b}_n^H \boldsymbol{h}_j \boldsymbol{a}(\tau_j)^H \boldsymbol{e}_n \\ &= \mathbf{Tr}\left(\sum_j c_j \boldsymbol{b}_n^H \boldsymbol{h}_j \boldsymbol{a}(\tau_j)^H \boldsymbol{e}_n \right) \\ &= \mathbf{Tr}\left(\boldsymbol{e}_n \boldsymbol{b}_n^H \sum_j c_j \boldsymbol{h}_j \boldsymbol{a}(\tau_j)^H \right) \end{aligned}$$

where $\boldsymbol{e}_n, -2M \le n \le 2M$, is the $(n + 2M + 1)$th column of the identity matrix $\boldsymbol{I}_N$, and $\mathbf{Tr}(\cdot)$ and $(\cdot)^H$ denote the trace and Hermitian operations, respectively. Then, using the lifting trick [12, 13], we obtain

$$\boldsymbol{y}(n) = \left\langle \sum_j c_j \boldsymbol{h}_j \boldsymbol{a}(\tau_j)^H, \boldsymbol{b}_n \boldsymbol{e}_n^H \right\rangle \quad (7)$$

where we have defined $\langle \boldsymbol{X}, \boldsymbol{Y} \rangle = \mathbf{Tr}(\boldsymbol{Y}^H \boldsymbol{X})$. Note that (7) leads to a parameterized rank-J matrix sensing problem, which we write as $\boldsymbol{y} = \mathcal{B}(\boldsymbol{X}_o), \boldsymbol{X}_o = \sum_j c_j \boldsymbol{h}_j \boldsymbol{a}(\tau_j)^H$, with the linear operator $\mathcal{B} : \mathbb{C}^{K\times N} \to \mathbb{C}^N$ defined as $[\mathcal{B}(\boldsymbol{X}_o)]_n = \left\langle \boldsymbol{X}_o, \boldsymbol{b}_n \boldsymbol{e}_n^H \right\rangle$.

Define the atomic norm [14] associated with the following set of atoms

$$\mathcal{A} = \left\{ \boldsymbol{h}\boldsymbol{a}(\tau)^H : \tau \in [0,1), \|\boldsymbol{h}\|_2 = 1, \boldsymbol{h} \in \mathbb{C}^{K\times 1} \right\}$$

as

$$\begin{aligned} \|\boldsymbol{X}\|_{\mathcal{A}} &= \inf\{t > 0 : \boldsymbol{X} \in t\mathbf{conv}(\mathcal{A})\} \\ &= \inf_{c_k, \tau_k, \|\boldsymbol{h}_k\|_2 = 1} \left\{ \sum_k |c_k| : \boldsymbol{X} = \sum_k c_k \boldsymbol{h}_k \boldsymbol{a}(\tau_k)^H \right\}. \end{aligned}$$

To enforce the sparsity of the atomic representation, we solve

$$\begin{aligned} &\mathbf{minimize} \ \ \|\boldsymbol{X}\|_{\mathcal{A}} \\ &\mathbf{subject\ to} \ \ \boldsymbol{y}(n) = \langle \boldsymbol{X}, \boldsymbol{b}_n \boldsymbol{e}_n^H \rangle, n = -2M, \cdots, 2M. \end{aligned} \quad (8)$$

Standard Lagrangian analysis shows that the dual of (8) is given by

$$\textbf{maximize } \langle \boldsymbol{\lambda}, \boldsymbol{y} \rangle_{\mathbb{R}} \quad \textbf{subject to } \|\mathcal{B}^*(\boldsymbol{\lambda})\|_{\mathcal{A}}^* \leq 1 \quad (9)$$

where $\mathcal{B}^*$ denotes the adjoint operator of $\mathcal{B}$ and $\|\cdot\|_{\mathcal{A}}^*$ is the dual norm of the atomic norm.

2.2. SDP characterization and optimality condition

Since the convex hull of the set of atoms $\mathcal{A}$ can be characterized by semidefinite programming, $\|\boldsymbol{X}\|_{\mathcal{A}}$ admits an equivalent SDP representation.

Lemma 2.1 *[15] For any $\boldsymbol{X} \in \mathbb{C}^{K \times N}$,*

$$\|\boldsymbol{X}\|_{\mathcal{A}} = \inf_{\boldsymbol{u}, \boldsymbol{T}} \left\{ \frac{1}{2N} \mathbf{Tr}\left(\mathrm{Toep}(\boldsymbol{u})\right) + \frac{1}{2} \mathbf{Tr}(\boldsymbol{T}) : \begin{bmatrix} \mathrm{Toep}(\boldsymbol{u}) & \boldsymbol{X}^H \\ \boldsymbol{X} & \boldsymbol{T} \end{bmatrix} \succeq 0 \right\}$$

where $\mathrm{Toep}(\boldsymbol{u})$ *denotes the Hermitian Toeplitz matrix whose first column is* $\boldsymbol{u}$*.*

Hence, (8) can be solved efficiently using available off-the-shelf solvers such as [16]. The following proposition characterizes the optimality condition of (8).

Proposition 2.2 *Suppose that the atomic set $\mathcal{A}$ is composed of atoms of the form $\boldsymbol{h}_j \boldsymbol{a}(\tau_j)^H$ with $\|\boldsymbol{h}_j\|_2 = 1, \tau_j \in [0, 1)$. Define the set $\mathbb{D} = \{\tau_j, 1 \leq j \leq J\}$. Let $\widehat{\boldsymbol{X}}$ be the optimal solution to (8). Then $\widehat{\boldsymbol{X}} = \boldsymbol{X}_o$ is the unique optimal solution if the following two conditions are satisfied:*
1) There exists a dual polynomial

$$\begin{aligned} \boldsymbol{q}(\tau) &= \mathcal{B}^*(\boldsymbol{\lambda}) \boldsymbol{a}(\tau) \\ &= \sum_{n=-2M}^{2M} \boldsymbol{\lambda}(n) e^{i2\pi n \tau} \boldsymbol{b}_n \end{aligned}$$

such that

$$\boldsymbol{q}(\tau_j) = \mathrm{sign}(c_j) \boldsymbol{h}_j, \quad \forall\, \tau_j \in \mathbb{D} \quad (10)$$

$$\|\boldsymbol{q}(\tau)\|_2 < 1, \quad \forall\, \tau \notin \mathbb{D}. \quad (11)$$

Here $\boldsymbol{\lambda}$ is a dual optimal solution and $\mathrm{sign}(c_j) := \frac{c_j}{|c_j|}$.
2) $\left\{ \begin{bmatrix} \vdots \\ \boldsymbol{a}(\tau_j)^H \boldsymbol{e}_n \boldsymbol{b}_n^H \\ \vdots \end{bmatrix}, j = 1, \cdots, J \right\}$ *is a linearly independent set.*

Note that the vector polynomial $\boldsymbol{q}(\tau)$ serves as a dual certificate to certify the optimality of $\boldsymbol{X}_o$ in the primal problem (8). We omit the proof of Proposition 2.2 here due to limited space.

3. MAIN THEOREM AND DISCUSSIONS

Before showing the main results of our work, we discuss the assumptions that are used in the main theorem. The assumptions can be grouped into three parts: (a) randomness and incoherence of the subspace spanned by the columns of $\boldsymbol{B}$, (b) minimum separation of $\{\tau_j\}$, and (c) uniform distribution of $\boldsymbol{h}_j$ on the complex unit sphere $\mathbb{CS}^{K-1}$.

We assume that the rows of the matrix $\boldsymbol{B}$, namely, $\boldsymbol{b}_n, -2M \leq n \leq 2M$, are independently sampled from a population $\mathcal{F}$ with the following properties [1, 17]:

- **Isotropy property:** We assume that the distribution $\mathcal{F}$ obeys the isotropy property in that
$$\mathbb{E}\boldsymbol{b}\boldsymbol{b}^H = \boldsymbol{I}_K, \quad \boldsymbol{b} \sim \mathcal{F}.$$
- **Incoherence property:** We assume that $\mathcal{F}$ satisfies the incoherence property with coherence $\mu(\mathcal{F})$ in that
$$\max_{1 \leq p \leq K} |\boldsymbol{b}(p)|^2 \leq \mu(\mathcal{F}), \quad \boldsymbol{b} \in \mathcal{F},$$
where $\boldsymbol{b}(p)$ is the pth element of $\boldsymbol{b}$.

Furthermore, we require the following conditions on the parameters of the complex exponentials.

- **Minimum separation:** We assume that
$$\Delta_\tau = \min_{k \neq j} |\tau_k - \tau_j| \geq \frac{1}{M}$$
where the distance $|\tau_k - \tau_j|$ is understood as the wrap-around distance on $[0, 1)$.
- **Random sign:** We assume that the coefficient vectors $\boldsymbol{h}_j$ are drawn i.i.d. from the uniform distribution on the complex unit sphere $\mathbb{CS}^{K-1}$.

Theorem 3.1 *Assume that the minimum separation condition $\Delta_\tau \geq \frac{1}{M}$ is satisfied and that $M \geq 64$. Also, assume that $\boldsymbol{g}_j = \boldsymbol{B}\boldsymbol{h}_j$ with the columns of $\boldsymbol{B}^H$, namely, $\boldsymbol{b}_n$, being i.i.d. samples from a distribution $\mathcal{F}$ that satisfies the isotropy and incoherence properties with coherence parameter $\mu(\mathcal{F})$. Additionally, assume that $\boldsymbol{h}_j$ are drawn i.i.d. from the uniform distribution on the complex unit sphere $\mathbb{CS}^{K-1}$. Then, there exists a numerical constant C such that*

$$M \geq C\mu(\mathcal{F}) JK \log\left(\frac{MJK}{\delta}\right) \log^2\left(\frac{MK}{\delta}\right)$$

is sufficient to guarantee that we can recover $\boldsymbol{X}_o$ and localize the τ_j's via solving (8) with probability at least $1 - \delta$.

Remarks:

- According to (10), $\{\tau_j\}$ can be localized by checking the ℓ_2 norm of the dual polynomial $\boldsymbol{q}(\tau)$. Once τ_j's are identified, one can find $\boldsymbol{h}_j$'s by solving a least squares problem, up to phase ambiguities.

- Our bound on M suggests that when $\mu(\mathcal{F})$ is a constant (e.g., when the rows of $\boldsymbol{B}$ are drawn from a sub-Gaussian distribution, $\mu(\mathcal{F})$ can be bounded by a constant times $\log K$ with high probability [17]), $M = O(JK)$ is sufficient for exact recovery and this matches the number of degrees of freedom in the problem, up to polylogarithmic factors. Thus, our sample complexity bound is tight and there is little room for further improvement. Furthermore, when the dimension of the subspace is bounded by a constant, $M = O(J)$ (up to a polylogarithmic factor) is sufficient for exact recovery. This bound matches the one in the deterministic super-resolution framework [10], where $N = O(J)$ suffices to exactly localize the unknown spikes under the same minimum separation condition used here.
- Our bound improves the one derived by Y. Chi [1] even when $\boldsymbol{g}_j = \boldsymbol{B}\boldsymbol{h}$, i.e., when $\boldsymbol{g}_j$ has no dependence on j. We note that the number of degrees of freedom in their problem is $O(J + K)$. It would be interesting to see if further improvement upon our bound is possible in this scenario.

The proof of our main theorem involves the construction of a dual polynomial $\boldsymbol{q}(\tau)$ that satisfies conditions (10) and (11) in Proposition 2.2. The form of $\boldsymbol{q}(\tau)$ motivates us to use the the random Fejér kernel

$$\boldsymbol{K}_M(\tau) = \frac{1}{M} \sum_{n=-2M}^{2M} g_M(n) \boldsymbol{b}_n \boldsymbol{b}_n^H e^{i2\pi n\tau}$$

as a building block. Specifically, we explicitly construct

$$\boldsymbol{q}(\tau) = \sum_{j=1}^{J} \boldsymbol{K}_M(\tau - \tau_j)\boldsymbol{\alpha}_j + \sum_{j=1}^{J} \boldsymbol{K}'_M(\tau - \tau_j)\boldsymbol{\beta}_j. \quad (12)$$

First of all, we aim to find $\boldsymbol{\alpha}_j$ and $\boldsymbol{\beta}_j$ such that $\boldsymbol{q}(\tau)$ satisfies condition (10) in Proposition 2.2. This can be done by solving the following $2JK \times 2JK$ linear system of equations

$$\underbrace{\begin{bmatrix} \boldsymbol{D}_0 & \frac{1}{\sigma}\boldsymbol{D}_1 \\ -\frac{1}{\sigma}\boldsymbol{D}_1 & -\frac{1}{\sigma^2}\boldsymbol{D}_2 \end{bmatrix}}_{\boldsymbol{D}} \begin{bmatrix} \boldsymbol{\alpha} \\ \sigma\boldsymbol{\beta} \end{bmatrix} = \begin{bmatrix} \boldsymbol{h} \\ \boldsymbol{0} \end{bmatrix}$$

where $\boldsymbol{\alpha} = \begin{bmatrix} \boldsymbol{\alpha}_1^H & \cdots & \boldsymbol{\alpha}_J^H \end{bmatrix}^H$, $\boldsymbol{\beta} = \begin{bmatrix} \boldsymbol{\beta}_1^H & \cdots & \boldsymbol{\beta}_J^H \end{bmatrix}^H$, $\boldsymbol{h} = \begin{bmatrix} (\mathrm{sign}(c_1)\boldsymbol{h}_1)^H & \cdots & (\mathrm{sign}(c_J)\boldsymbol{h}_J)^H \end{bmatrix}^H$, $\sigma = \sqrt{\frac{4\pi^2(M^2-1)}{3}}$, and $[\boldsymbol{D}_\ell]_{sj} = \boldsymbol{K}_M^\ell(\tau_s - \tau_j)$, where ℓ denotes the ℓth derivative of the entries of $\boldsymbol{K}_M(\cdot)$. The rest of the proof is omitted for space, but outlined as follows:

- Showing that the matrix $\boldsymbol{D}$ is invertible with high probability under the minimum separation condition and the properties of $\boldsymbol{b}_n$.
- Showing that $\|\boldsymbol{q}(\tau)\|_2 < 1, \ \forall\, \tau \notin \mathbb{D}$.

We refer the interested readers to [18] for the detailed proof of Theorem 3.1.

4. NUMERICAL SIMULATIONS

In this part, we provide synthetic numerical simulations to support our theoretical findings. We characterize the phase transition of atomic norm minimization (8) for non-stationary blind super-resolution. We solve the atomic norm minimization problem (8) using CVX [16]. We fix $N = 64$ and vary the values of J and K. We generate $\{\tau_j\}$ uniformly at random between 0 and 1 under the minimum separation condition $\Delta_\tau = \frac{1}{N}$, which is slightly smaller than $\frac{1}{M}$ required by our theorem. $\{c_j\}$ are generated with dynamic range of 10. Entries of $\boldsymbol{B}$ are generated randomly from the standard Gaussian distribution. Each $\boldsymbol{h}_j$ is also generated using i.i.d. real Gaussian random variables. We run 40 trials for each pair of J and K. For each trial we declare success if the relative reconstruction Frobenius norm error of X_o is less than 10^{-4}. Figure 1 shows the phase transition of our proposed method.

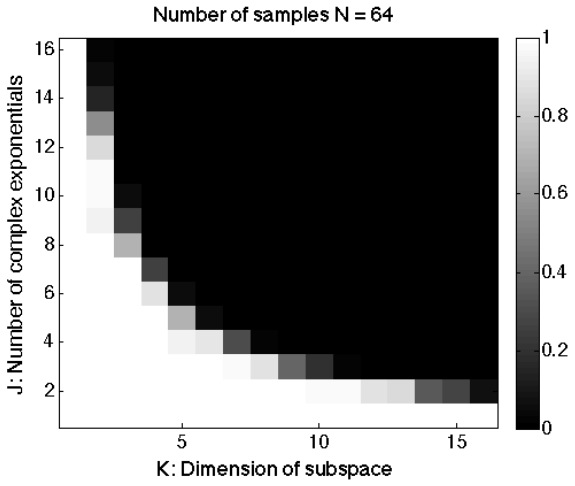

Fig. 1. The phase transition of non-stationary blind super-resolution using atomic norm minimization. This figure shows the empirical success rate of (8) for various numbers of complex exponentials J and different dimensions of subspace K when the number of samples $N = 64$ over 40 trials.

5. CONCLUSION

We developed a new model for non-stationary blind super-resolution. Using the lifting trick, we were able to formulate the problem as a convex program under the subspace assumption of the unknown waveforms. A sample complexity bound that is proportional to the number of degrees of freedom in the problem has been derived for exact recovery under the minimum separation condition. Numerical simulations were provided to validate our proposed approach.

6. REFERENCES

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

[2] B. Huang, W. Wang, M. Bates, and X. Zhuang, "Three-dimensional super-resolution imaging by stochastic optical reconstruction microscopy," *Science*, vol. 319, no. 5864, pp. 810–813, 2008.

[3] S. Quirin, S. Pavani, and R. Piestun, "Optimal 3d single-molecule localization for superresolution microscopy with aberrations and engineered point spread functions," *Proceedings of the National Academy of Sciences*, vol. 109, no. 3, pp. 675–679, 2012.

[4] R. Heckel, V. Morgenshtern, and M. Soltanolkotabi, "Super-resolution radar," *arXiv preprint arXiv:1411.6272*, 2014.

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

[6] J. Cai, X. Qu, W. Xu, and G. Ye, "Robust recovery of complex exponential signals from random gaussian projections via low rank hankel matrix reconstruction," *arXiv preprint arXiv:1503.02893*, 2015.

[7] G. Margrave, M. Lamoureux, and D. Henley, "Gabor deconvolution: Estimating reflectivity by nonstationary deconvolution of seismic data," *Geophysics*, vol. 76, no. 3, pp. W15–W30, 2011.

[8] R. Fergus, B. Singh, A. Hertzmann, S. Roweis, and W. Freeman, "Removing camera shake from a single photograph," *ACM Trans. Graphics (TOG)*, vol. 25, no. 3, pp. 787–794, 2006.

[9] J. Starck, E. Pantin, and F. Murtagh, "Deconvolution in astronomy: A review," *Publications of the Astronomical Society of the Pacific*, vol. 114, no. 800, pp. 1051–1069, 2002.

[10] E. 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.

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

[12] E. Candès, Y. Eldar, T. Strohmer, and V. Voroninski, "Phase retrieval via matrix completion," *SIAM Review*, vol. 57, no. 2, pp. 225–251, 2015.

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

[14] V. Chandrasekaran, B. Recht, P. Parrilo, and A. Willsky, "The convex geometry of linear inverse problems," *Foundations of Computational Mathematics*, vol. 12, no. 6, pp. 805–849, 2012.

[15] Y. Chi, "Joint sparsity recovery for spectral compressed sensing," in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Processing (ICASSP),*, pp. 3938–3942, 2014.

[16] M. Grant, S. Boyd, and Y. Ye, "Cvx: Matlab software for disciplined convex programming," 2008.

[17] E. Candès and Y. Plan, "A probabilistic and ripless theory of compressed sensing," *IEEE Trans. Information Theory,*, vol. 57, no. 11, pp. 7235–7254, 2011.

[18] D. Yang, G. Tang, and M. Wakin, "Super-resolution of complex exponentials from modulations with unknown waveforms," *arXiv preprint arXiv:1601.03712*, 2016.