Spectral Compressed Sensing via Structured Matrix Completion

Yuxin Chen   YXCHEN@STANFORD.EDU

Department of Electrical Engineering, Stanford University, Stanford, CA 94305, USA

Yuejie Chi   CHI@ECE.OSU.EDU

Electrical and Computer Engineering, The Ohio State University, Columbus, OH 43210, USA

Abstract

The paper studies the problem of recovering a spectrally sparse object from a small number of time domain samples. Specifically, the object of interest with ambient dimension n is assumed to be a mixture of r complex multi-dimensional sinusoids, while the underlying frequencies can assume any value in the unit disk. Conventional compressed sensing paradigms suffer from the *basis mismatch* issue when imposing a discrete dictionary on the Fourier representation. To address this problem, we develop a novel nonparametric algorithm, called enhanced matrix completion (EMaC), based on structured matrix completion. The algorithm starts by arranging the data into a low-rank enhanced form with multi-fold Hankel structure, then attempts recovery via nuclear norm minimization. Under mild incoherence conditions, EMaC allows perfect recovery as soon as the number of samples exceeds the order of $\mathcal{O}(r \log^2 n)$. We also show that, in many instances, accurate completion of a low-rank multi-fold Hankel matrix is possible when the number of observed entries is proportional to the information theoretical limits (except for a logarithmic gap). The robustness of EMaC against bounded noise and its applicability to super resolution are further demonstrated by numerical experiments.

1. Introduction

A large class of practical applications features high-dimensional objects that can be modeled or approximated by a superposition of spikes in the spectral (resp. time) domain, and involves estimation of the object from its time (resp. frequency) domain samples. A partial list includes medical imaging (Lustig et al., 2007), radar systems (Potter et al., 2010), seismic imaging (Borcea et al., 2002), microscopy (Schermelleh et al., 2010), etc. The data acquisition devices, however, are often limited by physical and hardware constraints, precluding sampling with the desired resolution. It is thus of paramount interest to reduce the sensing complexity while retaining recovery resolution.

Fortunately, in many instances, it is possible to recover an object even when the number of samples is far below the ambient dimension, provided that the object has a parsimonious representation in the transform domain. In particular, recent advances in compressed sensing (CS) (Candes et al., 2006) popularize the non-parametric methods based on convex surrogates. Such tractable methods do not require prior information on the model order, and are often robust against noise.

Nevertheless, the success of CS relies on sparse representation or approximation of the object of interest in a finite discrete dictionary, while the true parameters in many application are actually specified in a *continuous* dictionary. For concreteness, consider an object $x(\boldsymbol{t})$ that is a weighted sum of K-dimensional sinusoids at r distinct frequencies $\{\boldsymbol{f}_i \in [0,1]^K : 1 \le i \le r\}$. Conventional CS paradigms operate under the assumptions that these frequencies lie on a pre-determined grid on the unit disk. However, cautions need to be taken when imposing a discrete dictionary on continuous frequencies, since nature never poses the frequencies on the pre-determined grid, no matter how fine the grid is (Chi et al., 2011; Duarte & Baraniuk, 2012). This issue, known as *basis mismatch* between the true frequencies and the discretized grid, results in loss of sparsity due to spectral leakage along the Dirichlet kernel, and hence degeneration in the performance of CS algorithms. While one might impose finer gridding to mitigate this weakness, this approach often leads

to numerical instability and high correlation between dictionary elements, which significantly weakens the advantage of these CS approaches (Tang et al., 2012).

In this paper, we explore the above *spectral compressed sensing* problem, which aims to recover a spectrally sparse object from a small set of time-domain samples. The underlying (possibly *multi-dimensional*) frequencies can assume any value in the unit disk, and need to be recovered with infinite precision. To address this problem, we develop a *nonparametric* algorithm, called enhanced matrix completion (EMaC), based on structured matrix completion. Specifically, EMaC starts by converting the data samples into an enhanced matrix with (K-fold) Hankel structures, and then solves a nuclear-norm minimization program to complete the enhanced matrix. We show that, under mild incoherence conditions, EMaC admits exact recovery from $\mathcal{O}(r \log^2 n)$ random samples, where r and n denote respectively the spectral sparsity and the ambient dimension. Additionally, we provide theoretical guarantee for the low-rank Hankel matrix completion problem, which is of great importance in control, natural language processing, computer vision, etc. To the best of our knowledge, our results provide the first theoretical bounds that are close to the information theoretic limit. Furthermore, numerical experiments demonstrate that our algorithm is robust against noise and is applicable to the problem of super resolution.

2. Connection to Prior Art

The spectral compressed sensing problem is closely related to *harmonic retrieval*, which seeks to extract the underlying frequencies of an object from a collection of its time-domain samples. This spans many signal processing applications including radar localization systems (Nion & Sidiropoulos, 2010), array imaging systems (Borcea et al., 2002), wireless channel sensing (Sayeed & Aazhang, 1999; Gedalyahui et al., 2011), etc. In fact, if the time-domain representation of an object can be estimated accurately, then its underlying frequencies can be identified using harmonic super-resolution methods.

Conventional approaches for these problems, such as ESPRIT (Roy & Kailath, 1989) and the matrix pencil method (Hua, 1992), are based on the eigenvalue decomposition of covariance matrices constructed from *equi-spaced* samples, which can accommodate infinite frequency precision. One weakness of these techniques lies in that they require prior information on the model order, that is, the number of underlying frequency spikes or, at least, an estimate of it. Besides, their performance largely depends on the knowledge of noise spectra; some of them are unstable in the presence of

noise and outliers (Dragotti et al., 2007).

Nonparametric algorithms based on convex optimization differ from the above parametric techniques in that the model order does not need to be specified *a priori*. Recently, Candès and Fernandez-Granda (2013) proposed a total-variation minimization algorithm to super-resolve a sparse object from frequency samples at the low end of its spectrum. This algorithm allows accurate super-resolution when the point sources are appropriately separated, and is stable against noise (Candes & Fernandez-Granda, 2012). Inspired by this approach, Tang et. al. (2012) developed an atomic norm minimization algorithm for line spectral estimation from $\mathcal{O}(r \log r \log n)$ time domain samples. This work is limited to 1-D frequency models and assumes randomness in the data model.

In contrast, our approach can accommodate multi-dimensional frequencies, and only assumes randomness in the observation basis. The algorithm is inspired by the recent advances of matrix completion (MC) problem, which aims at recovering a low-rank matrix from partial entries. It has been shown (Candes & Recht, 2009; Gross, 2011) that exact recovery is possible via nuclear norm minimization, as soon as the number of observed entries is on the order of the information theoretic limit. Encouragingly, this line of algorithms is also robust against noise and outliers (Negahban & Wainwright, 2012). Nevertheless, the theoretical guarantees of these algorithms do not apply to the more structured observation models associated with Hankel structure. Consequently, direct application of existing MC results yields pessimistic bounds on the number of samples, which is far beyond the degrees of freedom underlying the sparse object.

The rest of this paper is organized as follows. Section 3 describes the data model and the EMaC algorithm. Section 4 presents the theoretical guarantee of EMaC, with a proof outlined in the Appendix. Numerical validation of EMaC is given in Section 5. We then discuss the extension to low-rank Hankel matrix completion in Section 6, and conclude the paper in Section 7.

3. Model and Algorithm

Assume that the object of interest $x(\boldsymbol{t})$ can be modeled as a weighted sum of K-dimensional sinusoids at r distinct frequencies $\boldsymbol{f}_i \in [0,1]^K$ ($1 \leq i \leq r$), i.e.

$$x(\boldsymbol{t}) = \sum_{i=1}^{r} d_i e^{j2\pi \langle \boldsymbol{t}, \boldsymbol{f}_i \rangle} \tag{1}$$

where d_i's denote the complex amplitude. For concreteness, our discussion is mainly devoted to a 2-dimensional (2-D) frequency model. This subsumes

1-D line spectral estimation as a special case, and indicates how to address multi-dimensional models.

3.1. 2-D Frequency Model and Problem Setup

Consider a data matrix $\boldsymbol{X} = (x_{k,l})_{0 \leq k < n_1, 0 \leq l < n_2}$ of size $n_1 \times n_2$. Suppose each entry can be expressed as

$$x_{k,l} = \sum_{i=1}^{r} d_i y_i^k z_i^l,$$

where $y_i = \exp(j2\pi f_{1i})$ and $z_i = \exp(j2\pi f_{2i})$ for some set of frequency pairs $\{(f_{1i}, f_{2i}) \mid 1 \leq i \leq r\}$ (normalized by the Nyquist rate). We can then write $\boldsymbol{X}$ in a matrix form as follows:

$$\boldsymbol{X} = \boldsymbol{YDZ}^T, \qquad (2)$$

where $\boldsymbol{D} := \mathrm{diag}[d_1, d_2, \cdots, d_r]$, and $\boldsymbol{Y}$ and $\boldsymbol{Z}$ are defined as

$$\boldsymbol{Y} := \begin{bmatrix} 1 & 1 & \cdots & 1 \\ y_1 & y_2 & \cdots & y_r \\ \vdots & \vdots & \vdots & \vdots \\ y_1^{n_1-1} & y_2^{n_1-1} & \cdots & y_r^{n_1-1} \end{bmatrix}, \qquad (3)$$

$$\boldsymbol{Z} := \begin{bmatrix} 1 & 1 & \cdots & 1 \\ z_1 & z_2 & \cdots & z_r \\ \vdots & \vdots & \vdots & \vdots \\ z_1^{n_2-1} & z_2^{n_2-1} & \cdots & z_r^{n_2-1} \end{bmatrix}. \qquad (4)$$

Suppose that there exists a location set Ω of size m such that $x_{k,l}$ is observed iff $(k,l) \in \Omega$, and assume that Ω is sampled uniformly at random. We are interested in recovering $\boldsymbol{X}$ from its partial observation on the location set Ω.

Before continuing, we introduce a few notations that will be used throughout. The spectral norm (operator norm), the Frobenius norm, and the nuclear norm (sum of singular values) of a matrix $\boldsymbol{X}$ are denoted by $\|\boldsymbol{X}\|$, $\|\boldsymbol{X}\|_{\mathrm{F}}$ and $\|\boldsymbol{X}\|_*$, respectively. The inner product between two matrices is defined by $\langle \boldsymbol{B}, \boldsymbol{C} \rangle = \mathrm{trace}(\boldsymbol{B}^* \boldsymbol{C})$. We let $\mathcal{P}_\Omega$ be the orthogonal projection onto the subspace of matrices that vanishes outside Ω.

3.2. Enhanced Matrix Completion (EMaC)

One might naturally attempt recovery by applying the matrix completion algorithms (Candes & Recht, 2009) on the original data matrix, arguing that when r is small, perfect recovery of $\boldsymbol{X}$ is possible from partial measurements. Specifically, this corresponds to the following optimization program

$$\underset{\boldsymbol{M} \in \mathbb{C}^{n_1 \times n_2}}{\mathrm{minimize}} \quad \|\boldsymbol{M}\|_* \qquad (5)$$

$$\text{subject to} \quad \mathcal{P}_\Omega(\boldsymbol{M}) = \mathcal{P}_\Omega(\boldsymbol{X}),$$

which is the convex relaxation of the rank minimization problem. However, generic matrix completion algorithms (Gross, 2011) require at least the order of $r \max(n_1, n_2)$ samples, which far exceeds the degrees of freedom (which is $\mathcal{O}(r \log n)$ due to a coupon collector's effect) in our problem. What is worse, when $r > \min(n_1, n_2)$ (which is possible since r can be as large as $n_1 n_2$), $\boldsymbol{X}$ is no longer low-rank. This motivates us to construct other forms (e.g. Hua (1992)) that better capture the harmonic structure.

Specifically, we adopt an effective enhanced form of $\boldsymbol{X}$ based on 2-fold Hankel structure as follows. The enhanced matrix $\boldsymbol{X}_{\mathrm{e}}$ with respect to $\boldsymbol{X}$ is defined as a $k_1 \times (n_1 - k_1 + 1)$ block Hankel matrix

$$\boldsymbol{X}_{\mathrm{e}} := \begin{bmatrix} \boldsymbol{X}_0 & \boldsymbol{X}_1 & \cdots & \boldsymbol{X}_{n_1-k_1} \\ \boldsymbol{X}_1 & \boldsymbol{X}_2 & \cdots & \boldsymbol{X}_{n_1-k_1+1} \\ \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{X}_{k_1-1} & \boldsymbol{X}_{k_1} & \cdots & \boldsymbol{X}_{n_1-1} \end{bmatrix}, \qquad (6)$$

where each block is a $k_2 \times (n_2 - k_2 + 1)$ Hankel matrix defined such that for all $0 \leq l < n_1$:

$$\boldsymbol{X}_l := \begin{bmatrix} x_{l,0} & x_{l,1} & \cdots & x_{l,n_2-k_2} \\ x_{l,1} & x_{l,2} & \cdots & x_{l,n_2-k_2+1} \\ \vdots & \vdots & \vdots & \vdots \\ x_{l,k_2-1} & x_{l,k_2} & \cdots & x_{l,n_2-1} \end{bmatrix}.$$

This form allows us to derive, through algebraic manipulation or a tensor product approach, the Vandermonde decomposition of each row of $\boldsymbol{X}$ as follows

$$\forall l \; (0 \leq l < n_1): \quad \boldsymbol{X}_l = \boldsymbol{Z}_{\mathrm{L}} \boldsymbol{Y}_{\mathrm{d}}^l \boldsymbol{D} \boldsymbol{Z}_{\mathrm{R}}, \qquad (7)$$

where $\boldsymbol{Z}_{\mathrm{L}}$, $\boldsymbol{Z}_{\mathrm{R}}$ and $\boldsymbol{Y}_{\mathrm{d}}$ are defined as

$$\boldsymbol{Z}_{\mathrm{L}} := \begin{bmatrix} 1 & 1 & \cdots & 1 \\ z_1 & z_2 & \cdots & z_r \\ \vdots & \vdots & \vdots & \vdots \\ z_1^{k_2-1} & z_2^{k_2-1} & \cdots & z_r^{k_2-1} \end{bmatrix},$$

$$\boldsymbol{Z}_{\mathrm{R}} := \begin{bmatrix} 1 & z_1 & \cdots & z_1^{n_2-k_2} \\ 1 & z_2 & \cdots & z_2^{n_2-k_2} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & z_r & \cdots & z_r^{n_2-k_2} \end{bmatrix},$$

and $\boldsymbol{Y}_{\mathrm{d}} := \mathrm{diag}[y_1, y_2, \cdots, y_r]$. Plugging (7) into (6) yields the following:

$$\boldsymbol{X}_{\mathrm{e}} = \underbrace{\begin{bmatrix} \boldsymbol{Z}_{\mathrm{L}} \\ \boldsymbol{Z}_{\mathrm{L}} \boldsymbol{Y}_{\mathrm{d}} \\ \vdots \\ \boldsymbol{Z}_{\mathrm{L}} \boldsymbol{Y}_{\mathrm{d}}^{k_1-1} \end{bmatrix}}_{\sqrt{k_1 k_2} \boldsymbol{E}_{\mathrm{L}}} \boldsymbol{D} \underbrace{\left[\boldsymbol{Z}_{\mathrm{R}}, \boldsymbol{Y}_{\mathrm{d}} \boldsymbol{Z}_{\mathrm{R}}, \cdots, \boldsymbol{Y}_{\mathrm{d}}^{n_1-k_1} \boldsymbol{Z}_{\mathrm{R}} \right]}_{\sqrt{(n_1-k_1+1)(n_2-k_2+1)} \boldsymbol{E}_{\mathrm{R}}},$$

where $\boldsymbol{E}_{\mathrm{L}}$ and $\boldsymbol{E}_{\mathrm{R}}$ characterize the column and row space of $\boldsymbol{X}_{\mathrm{e}}$, respectively. The effectiveness of the enhanced form relies on the shift-invariance of harmonic structures (any k consecutive samples lies in the same subspace), which promotes the use of Hankel matrices.

One can now see that $\boldsymbol{X}_{\mathrm{e}}$ is low-rank or, $\mathrm{rank}\left(\boldsymbol{X}_{\mathrm{e}}\right) \leq r$. We then attempt recovery via the following Enhanced Matrix Completion (EMaC) algorithm:

$$(\text{EMaC}) : \quad \underset{\boldsymbol{M} \in \mathbb{C}^{n_1 \times n_2}}{\text{minimize}} \quad \|\boldsymbol{M}_{\mathrm{e}}\|_* \qquad (8)$$
$$\text{subject to} \quad \mathcal{P}_\Omega\left(\boldsymbol{M}\right) = \mathcal{P}_\Omega\left(\boldsymbol{X}\right),$$

which minimizes the nuclear norm of the enhanced form over the constraint set. This convex program can be solved using off-the-shelf semidefinite program solvers in a tractable manner.

3.3. Higher-Dimensional Frequency Model

The EMaC method extends to higher dimensional frequency models without difficulty. For K-dimensional frequency models, one can convert the original data to a K-fold Hankel matrix of rank at most r. For instance, consider a 3 dimensional (3-D) model such that $x_{l_1,l_2,l_3} = \sum_{i=1}^r d_i y_i^{l_1} z_i^{l_2} w_i^{l_3}$. An enhanced form can be defined as a 3-fold Hankel matrix such that

$$\boldsymbol{X}_{\mathrm{e}} := \begin{bmatrix} \boldsymbol{X}_{0,\mathrm{e}} & \boldsymbol{X}_{1,\mathrm{e}} & \cdots & \boldsymbol{X}_{n_3-k_3,\mathrm{e}} \\ \boldsymbol{X}_{1,\mathrm{e}} & \boldsymbol{X}_{2,\mathrm{e}} & \cdots & \boldsymbol{X}_{n_3-k_3+1,\mathrm{e}} \\ \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{X}_{k_3-1,\mathrm{e}} & \boldsymbol{X}_{k_1,\mathrm{e}} & \cdots & \boldsymbol{X}_{n_3-1,\mathrm{e}} \end{bmatrix},$$

where $\boldsymbol{X}_{i,\mathrm{e}}$ denotes the 2-D enhanced form of the matrix consisting of all entries x_{l_1,l_2,l_3} obeying $l_3 = i$. One can verify that $\boldsymbol{X}_{\mathrm{e}}$ is of rank at most r, and can thus apply EMaC on the 3-D enhanced form $\boldsymbol{X}_{\mathrm{e}}$. To summarize, for K-dimensional frequency models, EMaC minimizes the nuclear norm over all K-fold Hankel matrices consistent with the observed entries.

3.4. Noisy Data

In practice, measurements are always corrupted by a certain amount of noise. To make our model and algorithm more practically applicable, we can replace our measurements by $\boldsymbol{X}_{il}^{\mathrm{o}}$ through the following noisy model

$$\forall (i,l) \in \Omega : \quad \boldsymbol{X}_{il}^{\mathrm{o}} = \boldsymbol{X}_{il} + \boldsymbol{N}_{il},$$

where $\boldsymbol{X}_{il}^{\mathrm{o}}$ is the observed (i,l)-th entry, and $\boldsymbol{N}_{il}$ denotes the noise. Suppose that the noise satisfies $\|\mathcal{P}_\Omega\left(\boldsymbol{N}\right)\|_{\mathrm{F}} \leq \delta$, then EMaC can be modified as:

$$(\text{EMaC-Noisy}) : \quad \underset{\boldsymbol{M} \in \mathbb{C}^{n_1 \times n_2}}{\text{minimize}} \quad \|\boldsymbol{M}_{\mathrm{e}}\|_* \qquad (9)$$
$$\text{subject to} \quad \|\mathcal{P}_\Omega\left(\boldsymbol{M} - \boldsymbol{X}^{\mathrm{o}}\right)\|_{\mathrm{F}} \leq \delta.$$

4. Main Results

Encouragingly, under certain incoherence conditions, the simple EMaC enables accurate recovery of the true data matrix from a small number of noiseless time-domain samples, and is stable against bounded noise.

For convenience of presentation, we denote $\Omega_{\mathrm{e}}(i,l)$ as the set of locations in the enhanced matrix $\boldsymbol{X}_{\mathrm{e}}$ containing copies of $x_{i,l}$, and let $\omega_{i,l} := |\Omega_{\mathrm{e}}\left(i,l\right)|$. For each (i,l), we use $\boldsymbol{A}_{(i,l)}$ to denote a *basis* matrix that extracts the average of all entries in $\Omega_{\mathrm{e}}\left(i,l\right)$, i.e.

$$\left(\boldsymbol{A}_{(i,l)}\right)_{\alpha,\beta} := \begin{cases} \frac{1}{\sqrt{|\Omega_{\mathrm{e}}(i,l)|}}, & \text{if } (\alpha,\beta) \in \Omega_{\mathrm{e}}(i,l), \\ 0, & \text{else.} \end{cases}$$

4.1. Incoherence Measures

In general, matrix completion from a few entries is hopeless unless the underlying structure is uncorrelated with the observation basis. This inspires us to define certain incoherence measures. Let $\boldsymbol{G}_{\mathrm{L}}$ and $\boldsymbol{G}_{\mathrm{R}}$ be $r \times r$ correlation matrices such that for any $i \neq j$,

$$\left(\boldsymbol{G}_{\mathrm{L}}\right)_{ij} := \frac{1}{k_1 k_2} \frac{1 - (y_i^* y_j)^{k_1}}{1 - y_i^* y_j} \frac{1 - (z_i^* z_j)^{k_2}}{1 - z_i^* z_j},$$

$$(\boldsymbol{G}_{\mathrm{R}})_{ij} := \frac{1 - (y_i^* y_j)^{n_1-k_1+1}}{(n_1 - k_1 + 1)(1 - y_i^* y_j)} \frac{1 - (z_i^* z_j)^{n_2-k_2+1}}{(n_2 - k_2 + 1)(1 - z_i^* z_j)},$$

with the convention $\left(\boldsymbol{G}_{\mathrm{L}}\right)_{ii} = \left(\boldsymbol{G}_{\mathrm{R}}\right)_{ii} = 1$. Note that $\boldsymbol{G}_{\mathrm{L}}$ and $\boldsymbol{G}_{\mathrm{R}}$ can be obtained by sampling the 2-D Dirichlet kernel, which is frequently used in Fourier analysis. Our incoherence measure is defined as follows.

Definition 1 (Incoherence) *Let $\boldsymbol{X}_{\mathrm{e}}$ denote the enhanced matrix associated with $\boldsymbol{X}$, and suppose the SVD of $\boldsymbol{X}_{\mathrm{e}}$ is given by $\boldsymbol{X}_{\mathrm{e}} = \boldsymbol{U}\boldsymbol{\Lambda}\boldsymbol{V}^*$. Then $\boldsymbol{X}$ is said to have incoherence (μ_1, μ_2, μ_3) if they are respectively the smallest quantities such that*

$$\sigma_{\min}\left(\boldsymbol{G}_{\mathrm{L}}\right) \geq 1/\mu_1, \quad \sigma_{\min}\left(\boldsymbol{G}_{\mathrm{R}}\right) \geq 1/\mu_1; \qquad (10)$$

$$\max_{(i,l) \in [n_1] \times [n_2]} \frac{\left| \sum_{(\alpha,\beta) \in \Omega_{\mathrm{e}}(i,l)} \left(\boldsymbol{U}\boldsymbol{V}^*\right)_{\alpha,\beta} \right|^2}{|\Omega_{\mathrm{e}}(i,l)|^2} \leq \frac{\mu_2 r}{n_1^2 n_2^2}; \qquad (11)$$

and

$$\sum_{\boldsymbol{a} \in [n_1] \times [n_2]} \left|\langle \boldsymbol{U}\boldsymbol{U}^* \boldsymbol{A}_b \boldsymbol{V}\boldsymbol{V}^*, \sqrt{\omega_a} \boldsymbol{A}_a \rangle\right|^2 \leq \frac{\mu_3 r}{n_1 n_2} \omega_{\boldsymbol{b}} \qquad (12)$$

holds for all $\boldsymbol{b} \in [n_1] \times [n_2]$.

Some brief interpretations of the above incoherence conditions are in order:

Condition (10) specifies certain incoherence among the locations of frequency pairs, which does not coincide

with and is not subsumed by the separation condition required in (Candes & Fernandez-Granda, 2013; Tang et al., 2012). The frequency pairs can be spread out (e.g. when their locations are generated in some random fashion), or minimally separated (e.g. when they are small perturbation of a fine grid).

Condition (11) can be satisfied when the total energy of each skew diagonal of UV^* is proportional to the dimension of this skew diagonal. This is indeed a weaker condition than the one introduced in (Candes & Recht, 2009) for matrix completion, which requires uniform energy distribution over all entries of UV^*. For instance, an ideal μ_2 can often be obtained when the complex phase of all frequencies are generated in some random fashion.

Condition (12) is an incoherence measure based on the (K-fold) Hankel structures. For example, one can reason that a desired μ_3 can be obtained if the magnitude of all entries of $UU^*A_aVV^*$ is mostly even. Condition (10) and (12) depend only on the frequency pairs. In fact, one can verify that $UU^* = E_\mathrm{L}\left(E_\mathrm{L}^*E_\mathrm{L}\right)^{-1}E_\mathrm{L}^*$ and $VV^* = E_\mathrm{R}^*\left(E_\mathrm{R}E_\mathrm{R}^*\right)^{-1}E_\mathrm{R}$, which depend only on the locations of the frequency pairs. Condition (11), however, might also rely on the amplitudes D.

Finally, the incoherence measures (μ_1, μ_2, μ_3) are mutually correlated, which is supplied as follows.

Lemma 1 *The incoherence measures (μ_1, μ_2, μ_3) of X_e satisfy*

$$\mu_2 \le \mu_1^2 c_s^2 r, \quad and \quad \mu_3 \le \mu_1^2 c_s^2 r. \tag{13}$$

4.2. Theoretical Guarantees

With the above incoherence measures, the main theoretical guarantee is supplied in the following theorem.

Theorem 1 *Let X be a data matrix with matrix form (2), and Ω the random location set of size m. Define $c_\mathrm{s} := \max\left(\frac{n_1 n_2}{k_1 k_2}, \frac{n_1 n_2}{(n_1-k_1+1)(n_2-k_2+1)}\right)$. If all measurements are noiseless, then there exists a constant $c_1 > 0$ such that under either of the following conditions:*

- *Conditions (10), (11) and (12) hold and*

$$m > c_1 \max\left(\mu_1 c_\mathrm{s}, \mu_2, \mu_3 c_\mathrm{s}\right) r \log^2\left(n_1 n_2\right); \tag{14}$$

- *Condition (10) holds and*

$$m > c_1 \mu_1^2 c_\mathrm{s}^2 r^2 \log^2(n_1 n_2); \tag{15}$$

X is the unique solution of EMaC with probability exceeding $1 - (n_1 n_2)^{-2}$.

Note that (15) is an immediate consequence of (14) by Lemma 1. Theorem 1 states the following: (1) under *strong* incoherence condition (i.e. given that (μ_1, μ_2, μ_3) are all constants), prefect recovery is possible as soon as the number of measurements exceeds the order of $r \log^2(n_1 n_2)$; (2) under *weak* incoherence condition (i.e. given only that μ_1 is a constant), perfect recovery is possible from $\mathcal{O}\left(r^2 \log(n_1 n_2)\right)$ samples. Since there are at least $\Theta(r)$ degrees of freedom in total, this establishes the near optimality of EMaC under strong incoherence condition.

We would like to note that while we assume random observation models, the conditions imposed on the data model are deterministic. This is different from (Tang et al., 2012), where randomness are assumed for both the observation model and the data model.

On the other hand, our method enables stable recovery even when the time-domain samples are noisy copies of the true data. Here, we say the recovery is stable if the solution of EMaC-Noisy is "close" to the ground truth. To this end, we establish the following theorem, which is a counterpart of Theorem 1 in the noisy setting.

Theorem 2 *Suppose X° is a noisy copy of X that satisfies $\|\mathcal{P}_\Omega(X - X^\circ)\|_\mathrm{F} \le \delta$. Under the conditions of Theorem 1, the solution to EMaC-Noisy in (9) satisfies*

$$\|\hat{X}_\mathrm{e} - X_\mathrm{e}\|_\mathrm{F} \le \left\{2\sqrt{n_1 n_2} + 8n_1 n_2 + \frac{8\sqrt{2}n_1^2 n_2^2}{m}\right\}\delta$$

with probability exceeding $1 - (n_1 n_2)^{-2}$.

Theorem 2 basically implies that the recovered enhanced matrix (which contains $\mathcal{O}(n_1^2 n_2^2)$ entries) is close to the true enhanced matrix at high signal-to-noise ratio. In particular, the average entry inaccuracy is bounded above by $\mathcal{O}(\frac{n_1 n_2}{m}\delta)$. We note that in practice, EMaC-Noisy usually yields better estimate, possibly by a polynomial factor. The practical applicability will be illustrated in Section 5 through numerical examples.

5. Numerical Experiments

5.1. Phase Transition for Exact Recovery

We examine phase transition of the EMaC algorithm to evaluate its practical ability. A square enhanced form was adopted with $n_1 = n_2$, which corresponds to the smallest c_s. For each (r, m) pair, we generated a spectrally sparse data matrix X by randomly generating r frequency spikes in $[0, 1] \times [0, 1]$, and sampled a subset Ω of size m entries uniformly at random. The EMaC algorithm was conducted using CVX with SDPT3. Each trial is declared successful if the
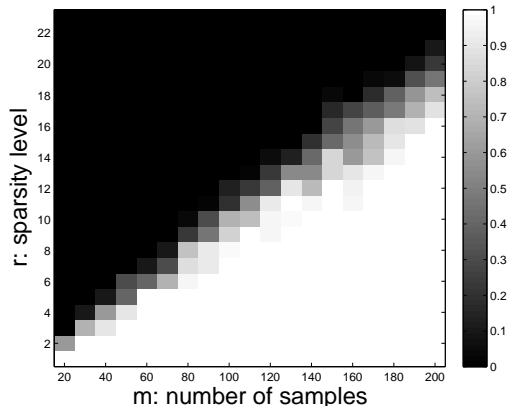
Figure 1. Phase transition diagrams where spike locations are randomly generated. The results are shown for the case where $n_1 = n_2 = 15$.

normalized mean squared error $\|\hat{X} - X\|_{\mathrm{F}}/\|X\|_{\mathrm{F}} \leq 10^{-4}$, where $\hat{X}$ denotes the estimate obtained through EMaC. The empirical success rate is calculated by averaging over 100 Monte Carlo trials.

Fig. 1 illustrates the results of these Monte Carlo experiments when X is a 15×15 matrix. The empirical success rate is reflected by the color of each cell. It can be seen that the number of samples m grows approximately linearly with respect to the spectral sparsity r, in line with our theoretical guarantee in Theorem 1. This phase transition diagram validates the practical applicability of our algorithm in the noiseless setting.

5.2. Stable Recovery via Singular Value Thresholding

The above experiments were conducted using the advanced semidefinite programming solver SDPT3. This and many other popular solvers (like SeDuMi) are based on interior point methods, which are typically inapplicable to large-scale data. In fact, SDPT3 fails to handle an $n \times n$ data matrix when n exceeds 19, which corresponds to a 100×100 enhanced matrix.

One solution for large-scale data is the first-order algorithms tailored for MC problems, e.g. the singular value thresholding (SVT) algorithm developed in (Cai et al., 2010). We propose a modified SVT algorithm in Algorithm 1 to exploit the Hankel structure.

In particular, $\mathcal{D}_{\tau_t}(\cdot)$ in Algorithm 1 denotes singular value shrinkage operator. Specifically, if the SVD of X is given by $X = U \Sigma V^*$ with $\Sigma = \mathrm{diag}\left(\{\sigma_i\}\right)$, then

$$\mathcal{D}_{\tau_t}(X) := U \mathrm{diag}\left(\left\{(\sigma_i - \tau_t)_+\right\}\right) V^*.$$

Algorithm 1 Singular Value Thresholding for EMaC

Input: The observed data matrix X° on the location set Ω.

initialize let X_{e}° denote the enhanced form of $\mathcal{P}_\Omega(X^\circ)$; set $M_0 = X_{\mathrm{e}}^\circ$ and $t = 0$.

repeat
    1) $Q_t \leftarrow \mathcal{D}_{\tau_t}(M_t)$
    2) $M_t \leftarrow \mathcal{H}_{X^\circ}(Q_t)$
    3) $t \leftarrow t + 1$
until convergence
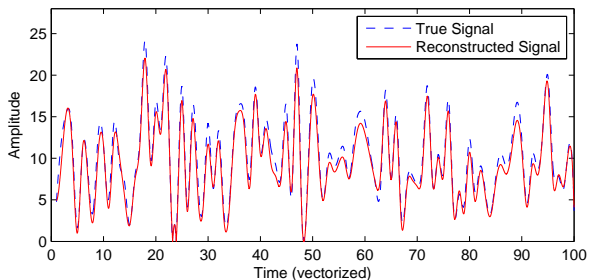**output** $\hat{X}$ as the matrix with enhanced form M_t.

Figure 2. The performance of SVT for a 101×101 data matrix that contains 30 random frequency spikes. 5.8% of all entries ($m = 600$) are observed with signal-to-noise amplitude ratio 10. Here, $\tau_t = 0.1\sigma_{\max}(M_t) / \lceil \frac{t}{10} \rceil$ empirically. The reconstruction against the true data for the first 100 time instances (after vectorization) are plotted.

where $\tau_t > 0$ is the soft-thresholding level[1]. Besides, in the K-dimensional frequency model, $\mathcal{H}_{X^\circ}(Q_t)$ denotes the projection of Q_t onto the subspace of enhanced matrices (i.e. K-fold Hankel matrices) that are consistent with the observed entries. Consequently, at each iteration, a pair (Q_t, M_t) is produced by first performing singular value shrinkage and then projecting the outcome onto the space of K-fold Hankel matrices that are consistent with observed entries.

Fig. 2 illustrates the performance of Algorithm 1. We generated a 101×101 data matrix X with a superposition of 30 random complex sinusoids, and revealed 5.8% of the total entries (i.e. $m = 600$) uniformly at random. The noise was i.i.d. Gaussian-distributed giving a signal-to-noise amplitude ratio of 10. The reconstructed signal is superimposed on the ground truth in Fig. 2. The normalized reconstruction error was $\|\hat{X} - X\|_{\mathrm{F}}/\|X\|_{\mathrm{F}} = 0.1098$, validating the stability of EMaC in the presence of noise.

---

[1] A good soft thresholding level is hard to pick, and needs to be selected by cross validation. Hence, the phase transitions of SVT and EMaC do not necessarily coincide.

5.3. Super Resolution

The proposed EMaC algorithm works beyond the random observation model in Theorem 1. Fig. 3 considers a synthetic super resolution example, where the ground truth in Fig. 3 (a) contains 6 point sources with constant amplitude. The low-resolution observation in Fig. 3 (b) is obtained by measuring low-frequency components $[-f_{\text{lo}}, f_{\text{lo}}]$ of the ground truth. Due to the large width of the associated point-spread function, both the locations and amplitudes of the point sources are distorted in the low-resolution image.

We apply EMaC to extrapolate high-frequency components up to $[-f_{\text{hi}}, f_{\text{hi}}]$, where $f_{\text{hi}}/f_{\text{lo}} = 2$. The reconstruction in Fig. 3 (c) is obtained via applying directly inverse Fourier transform of the spectrum to avoid parameter estimation such as the number of modes. The resolution is greatly enhanced from Fig. 3 (b), suggesting that EMaC is a promising approach for super resolution tasks.

6. Structured Matrix Completion

One problem closely related to our method is completion of K-fold Hankel matrices from a small number of entries. Consider for instance the 2-D model. While each spectrally sparse signal can be mapped to a low-rank 2-fold Hankel matrix, it is not clear whether all 2-fold Hankel matrices of rank r can be written as the enhanced form of an object with spectral sparsity r. Therefore, one can think of recovery of K-fold Hankel matrices as a more general problem than the spectral compressed sensing problem. Indeed, Hankel matrix completion has found numerous applications in system identification (Fazel et al., 2011), natural language processing (Balle & Mohri, 2012), computer vision (Sankaranarayanan et al., 2010), medical imaging (Shin et al., 2012), etc.

There has been several work concerning algorithms and numerical experiments for Hankel matrix completions (Fazel et al., 2003; 2011; Markovsky, 2008). However, to the best of our knowledge, there has been little theoretical guarantee that addresses directly Hankel matrix completion. Our analysis framework in Theorem 1 can be straightforwardly adapted to the general K-fold Hankel matrix completions. Notice that μ_2 and μ_3 are defined using the SVD of $\boldsymbol{X}_e$ in (11) and (12), and we only need to modify the definition of μ_1, as stated in the following theorem.

Theorem 3 *Consider a K-fold Hankel matrix $\boldsymbol{X}_e$ of rank r. The bounds in Theorem 1 and Theorem 2 continue to hold, if the incoherence μ_1 is defined as the smallest quantity that satisfies: $\forall (i,l) \in [n_1] \times [n_2]$,*

$$\max \left\{ \|\boldsymbol{U}\boldsymbol{U}^* \boldsymbol{A}_{(i,l)}\|_F^2, \|\boldsymbol{A}_{(i,l)} \boldsymbol{V}\boldsymbol{V}^*\|_F^2 \right\} \leq \frac{\mu_1 c_s r}{n_1 n_2}. \quad (16)$$

Condition (16) requires that the left and right singular vectors are sufficiently uncorrelated with the observation basis. In fact, condition (16) is a much weaker assumption than (10).

It is worth mentioning that low-rank Hankel matrices can often be converted to low-rank Toeplitz counterparts. Both Hankel and Toeplitz matrices are important forms that capture the underlying harmonic structures. Our results and analysis framework easily extend to the Toeplitz matrix completion problem.

7. Conclusions

We present an efficient nonparametric algorithm to estimate a spectrally sparse object from its partial time-domain samples, which poses spectral compressed sensing as a low-rank Hankel structured matrix completion problem. Under mild conditions, our algorithm enables recovery of the multi-dimensional unknown frequencies with infinite precision, which remedies the basis mismatch issue that arises in conventional CS paradigms. To the best of our knowledge, our result on Hankel matrix completion is also the first theoretical guarantee that is close to the information-theoretical limit (up to a logarithmic factor).

Appendix: Proof Outline of Theorem 1

The EMaC method has similar spirit as the well-known matrix completion algorithms (Candes & Recht, 2009; Gross, 2011), however the additional Hankel or block-Hankel structures on the matrices make existing theoretical results inapplicable in our framework. Nevertheless, the golfing scheme introduced in (Gross, 2011) lays the foundation of our analysis. We provide here a sketch of the proof, with detailed derivation deferred to (Chen & Chi, 2013) and the supplemental material.

Denote by T the tangent space with respect to $\boldsymbol{U}$ and $\boldsymbol{V}$. Let $\mathcal{P}_T$ be the orthogonal projection onto T, with $\mathcal{P}_{T^\perp} := \mathcal{I} - \mathcal{P}_T$ denoting its orthogonal complement. Denote by $\mathcal{A}_{(i,l)}$ the orthogonal projection onto the subspace spanned by $\boldsymbol{A}_{(i,l)}$. The projection onto the space spanned by all $\boldsymbol{A}_{(i,l)}$'s and its complement are defined as

$$\mathcal{A} := \sum_{(i,l) \in [n_1] \times [n_2]} \mathcal{A}_{(i,l)}, \quad \text{and} \quad \mathcal{A}^\perp = \mathcal{I} - \mathcal{A}. \quad (17)$$

Suppose that Ω is obtained by sampling with replacement, i.e. Ω contains m indices $\{z_i\}_{i=1}^m$ that are i.i.d.

(a) Ground truth   (b) Low-resolution observation   (c) High-resolution reconstruction
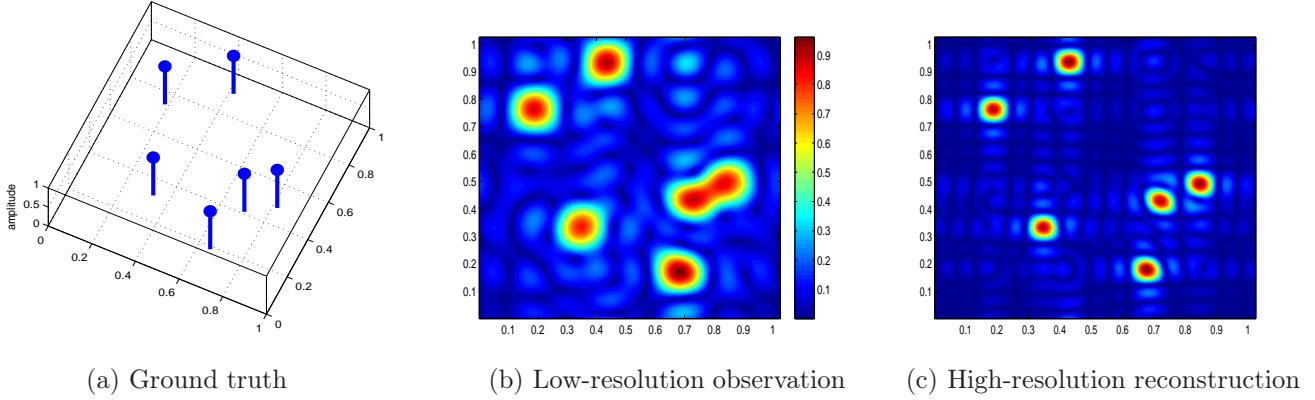
Figure 3. A synthetic super resolution example, where the observation (b) is taken from low-frequency components of the ground truth in (a), and the reconstruction (c) is done via inverse Fourier transform of extrapolated high-frequency components.

generated from $[n_1] \times [n_2]$. We define the associated projection operators as $\mathcal{A}_\Omega := \sum_{i=1}^m \mathcal{A}_{z_i}$, and introduce another projection operator $\mathcal{A}'_\Omega$ similar to $\mathcal{A}_\Omega$ but with the summation only over distinct samples.

To prove exact recovery of EMaC, it is sufficient to produce a dual certificate as follows.

Lemma 2 *For a location set Ω of size m, suppose that the sampling operator $\mathcal{P}_\Omega$ obeys*

$$\left\| \mathcal{P}_T \mathcal{A} \mathcal{P}_T - \frac{n_1 n_2}{m} \mathcal{P}_T \mathcal{A}_\Omega \mathcal{P}_T \right\| \leq \frac{1}{2}. \quad (18)$$

If there exists a matrix $\boldsymbol{W}$ that obeys

$$\begin{cases} (\mathcal{A} - \mathcal{A}'_\Omega)(\boldsymbol{U V}^* + \boldsymbol{W}) = 0, \\ \|\mathcal{P}_T(\boldsymbol{W})\|_{\mathrm{F}} \leq \frac{1}{2n_1^2 n_2^2}, \\ \|\mathcal{P}_{T^\perp}(\boldsymbol{W})\| \leq \frac{1}{2}, \end{cases} \quad (19)$$

then $\boldsymbol{X}_e$ is the unique optimizer of EMaC.

The dual certificate is constructed via the golfing scheme introduced in (Gross, 2011). Specifically, we generate j_0 independent random location sets $\Omega_j \sim$ Bernoulli(q), which is sampled with replacement. Set q such that $1 - \rho = (1-q)^{j_0}$, where $\rho := \frac{m}{n_1 n_2}$, then the distribution of Ω is the same as $\Omega_1 \cup \Omega_2 \cup \cdots \cup \Omega_{j_0}$. For small ρ, we have $q = \mathcal{O}(\rho/j_0)$ through the Taylor expansion. Consider a small constant $\epsilon < \frac{1}{3}$. The construction of a dual certificate $\boldsymbol{W}$ proceeds as follows:

1. Set $\boldsymbol{B}_0 = 0$, and $j_0 := 3 \log_{1/\epsilon}(n_1 n_2)$.

2. For $1 \leq i \leq j_0$, let

$$\boldsymbol{B}_i = \boldsymbol{B}_{i-1} + \left(q^{-1} \mathcal{A}_{\Omega_i} + \mathcal{A}^\perp\right) \mathcal{P}_T(\boldsymbol{U V}^* - \boldsymbol{B}_{i-1}).$$

3. Set $\boldsymbol{W} := -(\boldsymbol{U V}^* - \boldsymbol{B}_{j_0})$.

We will verify that $\boldsymbol{W}$ is a valid dual certificate satisfying the conditions (19) in Lemma 2. Our construction immediately yields $\left(\mathcal{A}'_\Omega + \mathcal{A}^\perp\right)(\boldsymbol{B}_i) = \boldsymbol{B}_i$ for $1 \leq i \leq j_0$, which gives

$$(\mathcal{A} - \mathcal{A}'_\Omega)(\boldsymbol{U V}^* + \boldsymbol{W}) = 0.$$

Define $\boldsymbol{F}_i := \boldsymbol{U V}^* - \boldsymbol{B}_i$, and hence $\boldsymbol{W} = \boldsymbol{F}_{j_0}$, one has

$$\mathcal{P}_T(\boldsymbol{F}_i) = \left(\mathcal{P}_T \mathcal{A} \mathcal{P}_T - q^{-1} \mathcal{P}_T \mathcal{A}_{\Omega_i} \mathcal{P}_T\right)(\boldsymbol{F}_{i-1}).$$

To proceed, we present Lemma 3 which shows that $\mathcal{A}_\Omega$ is sufficiently incoherent with respect to T.

Lemma 3 *There exists a constant $c_1 > 0$ such that if $m > c_1 \mu_1 c_s r \log(n_1 n_2)$, then*

$$\left\| \frac{n_1 n_2}{m} \mathcal{P}_T \mathcal{A}_\Omega \mathcal{P}_T - \mathcal{P}_T \mathcal{A} \mathcal{P}_T \right\| \leq \epsilon \quad (20)$$

with probability exceeding $1 - (n_1 n_2)^{-4}$.

Under the assumptions of Lemma 3, we have

$$\|\mathcal{P}_T(\boldsymbol{W})\|_{\mathrm{F}} = \|\mathcal{P}_T(\boldsymbol{F}_{j_0})\|_{\mathrm{F}} \leq \epsilon^{j_0} \sqrt{r} < n_1^{-2} n_2^{-2}.$$

It remains to show $\|\mathcal{P}_{T^\perp}(\boldsymbol{W})\| \leq \frac{1}{2}$ as follows.

Lemma 4 *There exist constants $c_8, c_9 > 0$ such that if $m > c_9 \max(\mu_1 c_s, \mu_2, \mu_3) r \log^2(n_1 n_2)$, then*

$$\forall 1 \leq i \leq j_0: \ \left\| \mathcal{P}_{T^\perp}\left(q^{-1} \mathcal{A}_{\Omega_i} + \mathcal{A}^\perp\right) \mathcal{P}_T(\boldsymbol{F}_i) \right\| < 2^{-i-2}$$

with probability at least $1 - c_8(n_1 n_2)^{-3}$.

Under the assumptions of Lemma 4, we have

$$\|\mathcal{P}_{T^\perp}(\boldsymbol{W})\| \leq \sum_{i=0}^{j_0} \left\| \mathcal{P}_{T^\perp}\left(q^{-1} \mathcal{A}_\Omega + \mathcal{A}^\perp\right) \mathcal{P}_T(\boldsymbol{F}_i) \right\| < 1/2.$$

So far, we have successfully shown that $\boldsymbol{W}$ is a valid dual certificate with high probability, and hence EMaC allows exact recovery with high probability.

References

Balle, B. and Mohri, M. Spectral learning of general weighted automata via constrained matrix completion. *Advances in Neural Information Processing Systems (NIPS)*, pp. 2168–2176, 2012.

Borcea, L., Papanicolaou, G., Tsogka, C., and Berryman, J. Imaging and time reversal in random media. *Inverse Problems*, 18(5):1247, 2002.

Cai, J. F., Candes, E. J., and Shen, Z. A singular value thresholding algorithm for matrix completion. *SIAM Journal on Optimization*, 20(4):1956–1982, 2010.

Candes, E. J. and Fernandez-Granda, C. Super-resolution from noisy data. *Arxiv 1211.0290*, November 2012.

Candes, E. J. and Fernandez-Granda, C. Towards a mathematical theory of super-resolution. *to appear in Communications on Pure and Applied Mathematics*, 2013.

Candes, E. J. and Recht, B. Exact matrix completion via convex optimization. *Foundations of Computational Mathematics*, 9(6):717–772, April 2009.

Candes, E. J., Romberg, J., and Tao, T. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory*, 52(2): 489–509, Feb. 2006.

Chen, Y. and Chi, Y. Robust spectral compressed sensing via structured matrix completion. *Arxiv: 1304.8126*, May 2013.

Chi, Y., Scharf, L.L., Pezeshki, A., and Calderbank, A.R. Sensitivity to basis mismatch in compressed sensing. *IEEE Transactions on Signal Processing*, 59(5):2182–2195, May 2011.

Dragotti, P. L., Vetterli, M., and Blu, T. Sampling moments and reconstructing signals of finite rate of innovation: Shannon meets strang-fix. *IEEE Trans on Signal Processing*, 55(5):1741 –1757, May 2007.

Duarte, M.F. and Baraniuk, R.G. Spectral compressive sensing. *Applied and Computational Harmonic Analysis*, 2012.

Fazel, M., Hindi, H., and Boyd, S. P. Log-det heuristic for matrix rank minimization with applications to Hankel and Euclidean distance matrices. *American Control Conference*, 3:2156 – 2162 vol.3, June 2003.

Fazel, M., Pong, T. K., Sun, D., and Tseng, P. Hankel matrix rank minimization with applications in system identification and realization, 2011.

Gedalyahui, K., Tur, R., and Eldar, Y. C. Multichannel sampling of pulse streams at the rate of innovation. *IEEE Trans on Sig. Proc*, 59:1491–1504, 2011.

Gross, D. Recovering low-rank matrices from few coefficients in any basis. *IEEE Transactions on Information Theory*, 57(3):1548–1566, March 2011.

Hua, Y. Estimating two-dimensional frequencies by matrix enhancement and matrix pencil. *IEEE Trans on Sig. Proc.*, 40(9):2267 –2280, Sep 1992.

Lustig, M., Donoho, D., and Pauly, J. M. Sparse MRI: The application of compressed sensing for rapid MR imaging. *Magnetic Resonance in Medicine*, 58(6): 1182–1195, 2007.

Markovsky, I. Structured low-rank approximation and its applications. *Automatica*, 44(4):891–909, 2008.

Negahban, S. and Wainwright, M.J. Restricted strong convexity and weighted matrix completion: Optimal bounds with noise. *The Journal of Machine Learning Research*, 98888:1665–1697, May 2012.

Nion, D. and Sidiropoulos, N. D. Tensor algebra and multidimensional harmonic retrieval in signal processing for MIMO Radar. *IEEE Transactions on Signal Processing*, 58(11):5693 –5705, Nov. 2010.

Potter, L.C., Ertin, E., Parker, J.T., and Cetin, M. Sparsity and compressed sensing in radar imaging. *Proceedings of the IEEE*, 98(6):1006–1020, 2010.

Roy, R. and Kailath, T. Esprit-estimation of signal parameters via rotational invariance techniques. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 37(7):984 –995, Jul 1989.

Sankaranarayanan, A., Turaga, P., Baraniuk, R., and Chellappa, R. Compressive acquisition of dynamic scenes. *ECCV 2010*, pp. 129–142, 2010.

Sayeed, A. M. and Aazhang, B. Joint multipath-Doppler diversity in mobile wireless communications. *IEEE Transactions on Communications*, 47 (1):123 –132, Jan 1999.

Schermelleh, L., Heintzmann, R., and Leonhardt, H. A guide to super-resolution fluorescence microscopy. *The Journal of cell biology*, 190(2):165–175, 2010.

Shin, P., Larson, P., Ohliger, M., Elad, M., Pauly, J., Vigneron, D., and Lustig, M. Calibrationless parallel imaging reconstruction based on structured low-rank matrix completion. *submitted to Magnetic Resonance in Medicine*, 2012.

Tang, G., Bhaskar, B. N., Shah, P., and Recht, B. Compressed sensing off the grid. *Arxiv 1207.6053*, July 2012.