Hankel Matrix Nuclear Norm Regularized Tensor Completion for N-dimensional Exponential Signals

Jiaxi Ying, Hengfa Lu, Qingtao Wei, Jian-Feng Cai, Di Guo, Jihui Wu, Zhong Chen, Xiaobo Qu*

Abstract—**Signals are generally modeled as a superposition of exponential functions in spectroscopy of chemistry, biology and medical imaging. For fast data acquisition or other inevitable reasons, however, only a small amount of samples may be acquired and thus how to recover the full signal becomes an active research topic. But existing approaches can not efficiently recover N-dimensional exponential signals with $N \geq 3$. In this paper, we study the problem of recovering N-dimensional (particularly $N \geq 3$) exponential signals from partial observations, and formulate this problem as a low-rank tensor completion problem with exponential factor vectors. The full signal is reconstructed by simultaneously exploiting the CANDECOMP/PARAFAC tensor structure and the exponential structure of the associated factor vectors. The latter is promoted by minimizing an objective function involving the nuclear norm of Hankel matrices. Experimental results on simulated and real magnetic resonance spectroscopy data show that the proposed approach can successfully recover full signals from very limited samples and is robust to the estimated tensor rank.**

Index Terms—**tensor completion, exponential signal, Hankel matrix, low rank, spectroscopy, NMR.**

I. INTRODUCTION

S IGNAL reconstruction from its sampled measurements is recognized as a fundamental theme of signal processing. Under some circumstances, these acquired measurements are incomplete due to costly experiments, hardware limitation, or other inevitable reasons. For example, for the purpose of fast data acquisition, nonuniform sampling is used to obtain partial entries of the time-domain signal in nuclear magnetic resonance (NMR) spectroscopy [1]–[3], which has been widely used in chemistry and biology. Recovering the full signal is essential for the next step of data analyses in these applications.

Jiaxi Ying, Hengfa Lu, Zhong Chen and Xiaobo Qu are with the Department of Electronic Science, Fujian Provincial Key Laboratory of Plasma and Magnetic Resonance, Xiamen University, Xiamen, China (e-mail: quxiaobo@xmu.edu.cn)

Qingtao Wei and Jihui Wu are with School of Life Sciences, University of Science and Technology of China, Hefei, Anhui, China.

Jian-Feng Cai is with Department of Mathematics, Hong Kong University of Science and Technology, Hong Kong SAR, China.

Di Guo is with the School of Computer and Information Engineering, Fujian Provincial University Key Laboratory of Internet of Things Application Technology, Xiamen University of Technology, Xiamen, China.

In this paper, the signal of interest $f(t_1, t_2, \ldots, t_N)$ can be modeled or approximated by a superposition of limited R N-dimensional (N-D) exponentials, i.e.,

$$f(t_1, t_2, \ldots, t_N) = \sum_{r=1}^{R} d_r \phi_{1,r}(t_1) \phi_{2,r}(t_2) \cdots \phi_{N,r}(t_N), \quad (1)$$

where f denotes a scalar function of N variables; $\phi_{n,r}(t_n)$ the exponential function for any n ($1 \leq n \leq N$) and r ($1 \leq r \leq R$), and d_r the complex amplitudes of the associated coefficient. This formulation implies that each component N-D exponential in (1) is an exponential function with respect to any variable. In particular, when $\phi_{n,r}(t_n) = e^{j2\pi f_{n,r}t_n}$, f denotes a spectrally sparse signal [4], [5], i.e., a superposition of R N-D undamped complex sinusoids. When $\phi_{n,r}(t_n) = e^{(-1/\tau_{n,r}+j2\pi f_{n,r})t_n}$ with the damping factor $\tau_{n,r} \in \mathbb{R}_+$, f denotes a sum of R N-D damped complex sinusoids. These two types of signals arise in various applications such as multiple-input multiple-output (MIMO) radars [6], harmonic analysis [7]–[9], analog-to-digital conversion [10], magnetic resonance imaging [11]–[15], and NMR spectroscopy [1], [16]. In this paper, we study the problem of reconstructing N-D exponential signals from a small amount of measurements.

Recently, recovering a spectrally sparse signal becomes of great interest in signal processing community [4], [5], [9], [17]–[20]. Among emerging approaches, the compressed sensing [21] suggests to reconstruct a signal from its partial observations if it enjoys a sparse representation in some transform domain and the observation operator satisfies some incoherence conditions. The spectrally sparse signal can be sparse in the discrete Fourier transform domain if the frequencies are aligned well with the discrete frequencies and the number of exponentials is small. In this case, signals can be recovered from very few measurements by enforcing the sparsity in the discrete Fourier domain [21]. However, true frequencies in practical applications generally take values on a continuous domain, and the resultant basis mismatch between the true frequencies and the discretized grid [22] leads to the loss of sparsity and hence worsens the performance of compressed sensing. For the purpose of addressing this problem, total variation or atomic norm [23] minimization methods were proposed to deal with signal recovery with continuous-valued frequencies [4], [5], [9], [17]. However, these methods [4], [5], [9], [17] work only for undamped complex sinusoids but not for generic exponential signals, such as damped complex sinusoids. Furthermore, being computationally expensive, they search the solution in a space whose dimension is the *square*
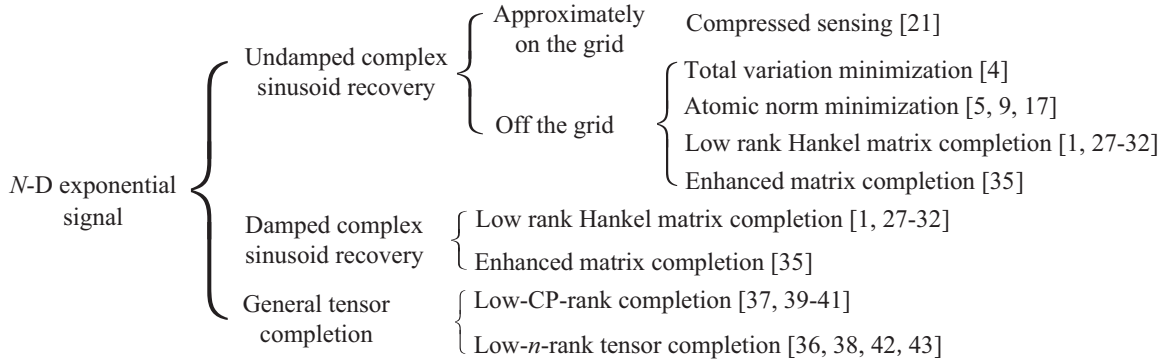
Fig. 1: Summary of methods in recovering N-dimensional exponential signals.

of the dimension of the underlying signal, which is intractable for large scale problems (e.g., $N \geq 3$).

More recently, inspired by matrix pencil method [24] and matrix completion [25], [26], low rank Hankel matrix (LRHM) reconstruction [1], [27]–[32] has been proposed to recover generic exponential signals. The LRHM was proposed in missing data recovery of non-uniformly sampling in realistic protein NMR spectroscopy [1], showing that broad peaks can be recovered much more reliably than minimizing the l_1 norm on the spectrum [2], [3], [33], [34]. The minimal number of samples for LRHM method is theoretically predicted when the measurements are taken with random Gaussian sampling [27]. However, LRHM is limited to recovering only 1-D signals. The N-D exponential signal is possibly recovered by enhanced matrix completion (EMaC) [35] with theoretical guarantee of the stable recovery. Nonetheless, EMaC invokes minimization problems with a huge unknown matrix, prohibiting its applicability to N-D ($N \geq 3$) exponential signals. Note that the structured matrix based methods [17], [35] become prohibitive even in the 3-order tensor completion, because they solve the tensor completion problem by lifting to a search space whose dimension is on the order of the square of the dimension of the tensor. For example, for a signal of size $50 \times 50 \times 50$, EMaC [35] needs to solve an optimization problem that involves a matrix of size 15625×17576. Therefore, how to recover N-D (particularly $N \geq 3$) exponential signals still remains challenging.

It can be easily checked from (1) that the signal of interest, viewed as an N-D tensor in the discrete domain, enjoys a low CANDECOMP/PARAFAC (CP) rank and a low n-rank if R is sufficiently small. Therefore, the exponential signal recovery can be formulated as a low-rank tensor completion. The low-rank tensor completion has been successfully applied to a large class of real-world problems, such as data analyses in computer vision [36], [37], remote sensing [38], and electroencephalogram [39]. Since the rank of a tensor can be defined differently, there exist several types of low-rank tensor completion approaches, e.g., the low-CP-rank tensor completion [37], [39]–[41] and the low-n-rank tensor completion [36], [38], [42], [43]. Although both low-CP-rank and low-n-rank tensor completions can be applied to recover N-D ($N \geq 3$) signals, they ignore the specific exponential structure, e.g. time domain signal in NMR spectroscopy, probably causing the requirement

of an unnecessarily large number of observed entries for a stable recovery.

In this paper, we propose an approach to reconstruct the N-D exponential signal satisfying (1) from a small amount of samples. The proposed approach simultaneously exploits the low-CP-rank structure and the exponential structure of the associated factor vectors (See the definition in Section III). To enforce the former structure, we represent the signal in the CP decomposition form and imposed a least square fitting to the sampled data. To promote the latter structure, we penalize the nuclear norm of the Hankel matrices formed by factor vectors. We will verify, with comprehensive numerical experiments on simulated and real-world data, the effectiveness of the new method by comparisons with state-of-the-art tensor completion methods.

The rest of this paper is organized as follows. In Section II, we will introduce the notations and related backgrounds. Section III converts the N-D exponential signal reconstruction to a low-rank tensor completion problem and reviews existing low-rank tensor completion methods. In Section IV, we will propose our Hankel matrix nuclear norm regularized tensor completion approach, including the algorithm, convergence and complexity analysis. Section V presents the experimental results on simulated and real-world data. Section VI discusses higher-dimensional experiments, comparisons with other state-of-the-art methods and some parameters. Section VII concludes this work and discusses the future work.

II. NOTATIONS AND BACKGROUNDS

Notations and backgrounds of tensors are given below.

A. Notations

Notations and nomenclatures of tensors are introduced following the review paper [44] and a summary of these notations is shown in Table I.

Tensor is the generalization of matrix to high dimensions. For a tensor $\mathcal{X} \in \mathbb{C}^{I_1 \times \cdots \times I_N}$, the number of dimensions N is called the order, also known as way or mode. Throughout the paper, the i-th entry of a vector $\mathbf{x}$ is denoted by x_i, the (i, j)-element of a matrix $\mathbf{X}$ is denoted by $x_{i,j}$, and the $(i_1, i_2, \ldots, i_N)$-element of an order-N tensor $\mathcal{X}$ is denoted by $x_{i_1, i_2, \ldots, i_N}$.

TABLE I: Basic notations

Symbols	Notations
$x, \mathbf{x}, \mathbf{X}, \mathcal{X}$	scalar, vector, matrix, tensor
$\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_R]$	matrix $\mathbf{X}$ with column vectors $\mathbf{x}_r$
$\mathbf{X}_{(n)}$	mode-n matricization of tensor $\mathcal{X}$
$\mathbf{X}^T, \mathbf{X}^H$	transpose, hermitian transpose
$\mathbf{x}_{i_1,\ldots,i_{n-1},:,i_{n+1},\ldots,i_N}$	mode-n fiber of tensor $\mathcal{X}$
$\mathbf{X}_{:,:,i_3,\ldots,i_N}$	matrix slice of tensor $\mathcal{X}$

a) Fibers and Slices: Fibers are the higher-order analogue of matrix rows and columns. The fiber is a collection of entries of the tensor by fixing all but one indices. The mode-n fibers are all vectors $\mathbf{x}_{i_1,\ldots,i_{n-1},:,i_{n+1},\ldots,i_N}$ that are obtained by fixing the indices $\{i_1, \ldots, i_N\} \setminus i_n$. Slices are 2-D sections of a tensor, defined by fixing all but two indices.

b) Kronecker product: The Kronecker product of matrices $\mathbf{A} \in \mathbb{C}^{m \times n}$ and $\mathbf{B} \in \mathbb{C}^{p \times q}$, denoted by $\mathbf{A} \otimes \mathbf{B}$, is defined by

$$\mathbf{A} \otimes \mathbf{B} := [a_{ij}\mathbf{B}]_{mp \times nq} \qquad (2)$$

c) Khatri-Rao product: The Khatri-Rao product is the columnwise Kronecker product. Given two matrices $\mathbf{A} \in \mathbb{C}^{p \times n}$ and $\mathbf{B} \in \mathbb{C}^{q \times n}$, their Khatri-Rao product $\mathbf{A} \odot \mathbf{B}$ satisfies

$$\mathbf{A} \odot \mathbf{B} = \begin{bmatrix} \mathbf{a}_1 \otimes \mathbf{b}_1 & \mathbf{a}_2 \otimes \mathbf{b}_2 & \cdots & \mathbf{a}_n \otimes \mathbf{b}_n \end{bmatrix} \qquad (3)$$

where $\otimes$ denotes Kronecker product.

d) Frobenius norm: The Frobenius norm of a tensor is the square root of the sum of the squares of the absolute value of each element, i.e.,

$$\|\mathcal{X}\|_F = \left(\sum_{i_1=1}^{I_1} \sum_{i_2=1}^{I_2} \cdots \sum_{i_N=1}^{I_N} |x_{i_1,\cdots,i_N}|^2 \right)^{\frac{1}{2}}. \qquad (4)$$

B. CP decomposition and tensor CP-rank

a) factor and rank-1 tensor: An N-order tensor $\mathcal{X}$ is rank-1 if it can be written as the outer product of vectors as follows

$$\mathcal{X} = \mathbf{u}^{(1)} \circ \mathbf{u}^{(2)} \circ \cdots \circ \mathbf{u}^{(N)},$$

where the symbol $\circ$ denotes the vector outer product and the vector $\mathbf{u}^{(n)}$, for all $1 \leq n \leq N$, is called a *factor*. This means each element of the tensor is the product of the corresponding vector elements:

$$x_{i_1,\cdots,i_N} = \prod_{n=1}^{N} u_{i_n}^{(n)}$$

for all $1 \leq i_n \leq I_n$ and $1 \leq n \leq N$.

b) CP decomposition and tensor CP-rank: The CP decomposition [45] factorizes a tensor into a linear combination of rank-1 tensors. A tensor $\mathcal{X} \in \mathbb{C}^{I_1 \times \cdots \times I_N}$ is represented with CP decomposition as

$$\mathcal{X} = \sum_{r=1}^{R} d_r \mathbf{u}_r^{(1)} \circ \mathbf{u}_r^{(2)} \circ \cdots \circ \mathbf{u}_r^{(N)}, \qquad (5)$$

where R is a positive integer and $d_r \in \mathbb{C}$. The tensor CP-rank is defined as the smallest number of rank-1 tensors that composes the $\mathcal{X}$ in (5).

Following [46], we combine the factors in CP decomposition (5) to form *factor matrices*

$$\mathbf{U}^{(n)} = \left[\mathbf{u}_1^{(n)} \mathbf{u}_2^{(n)} \ldots \mathbf{u}_R^{(n)} \right]$$

for $1 \leq n \leq N$, and the CP decomposition (5) is concisely expressed as

$$\mathcal{X} = \llbracket \mathbf{d}; \mathbf{U}^{(1)}, \mathbf{U}^{(2)}, \ldots, \mathbf{U}^{(N)} \rrbracket$$

where $\mathbf{d} = [d_1, \ldots, d_R]^T \in \mathbb{C}^R$ and $\llbracket \ \rrbracket$ is called the Tucker operator [46]. Note that we can always rescale the factor matrices so that all the entries in $\mathbf{d}$ are 1's, and for simplicity, we will drop $\mathbf{d}$, i.e.,

$$\llbracket \mathbf{U}^{(1)}, \mathbf{U}^{(2)}, \ldots, \mathbf{U}^{(N)} \rrbracket \equiv \sum_{r=1}^{R} \mathbf{u}_r^{(1)} \circ \mathbf{u}_r^{(2)} \circ \cdots \circ \mathbf{u}_r^{(N)}.$$

C. Tensor matricization and Tensor n-rank

There are some other definitions of tensor rank. The n-rank is defined via tensor *matricization* (also known as *unfolding*), reordering the elements of a tensor into a matrix. The mode-n matricization of a tensor $\mathcal{X} \in \mathbb{C}^{I_1 \times \cdots \times I_N}$ is denoted by $\mathbf{X}_{(n)}$, and the tensor element $x_{i_1,\cdots,i_N}$ is mapped to the matrix element (i_n, j_n), where

$$j_n = 1 + \sum_{k=1,k\neq n}^{N} (i_k - 1) J_k \ \text{with} \ J_k = \prod_{m=1,m\neq n}^{k-1} I_m. \qquad (6)$$

Therefore, $\mathbf{X}_{(n)} \in \mathbb{C}^{I_n \times K_n}$ with $K_n = \prod_{k=1,k\neq n}^{N} I_k$. With the help of Khatri-Rao product, the mode-i matricization can be expressed as

$$\mathbf{X}_{(i)} = \mathbf{U}^{(i)} \left(\mathbf{U}^{(N)} \odot \cdots \odot \mathbf{U}^{(i+1)} \odot \mathbf{U}^{(i-1)} \odot \cdots \odot \mathbf{U}^{(1)} \right)^T.$$

The n-rank of an N-D tensor $\mathcal{X}$ is the tuple of the ranks of the mode-n unfoldings:

$$n\text{-rank}(\mathcal{X}) = \left(\text{rank}\left(\mathbf{X}_{(1)}\right), \text{rank}\left(\mathbf{X}_{(2)}\right), \ldots, \text{rank}\left(\mathbf{X}_{(N)}\right) \right).$$

III. N-D EXPONENTIAL SIGNAL RECONSTRUCTION AS A TENSOR COMPLETION

Without loss of generality, the frequencies in (1) can be normalized with respect to the Nyquist frequency and hence the measurements are sampled at integer values. Therefore, by sampling the signal (1) on a uniform grid, we can obtain an N-order tensor $\mathcal{Y} \in \mathbb{C}^{I_1 \times \cdots \times I_N}$, and each entry $y_{i_1,\cdots,i_N}$ can be expressed as

$$y_{i_1,\cdots,i_N} = \sum_{r=1}^{R} \left(d_r \prod_{n=1}^{N} z_{n,r}^{i_n-1} \right), \qquad (7)$$

for $i_n = 0, \ldots, I_n - 1$, $n = 1, \ldots, N$, with $d_r \in \mathbb{C}$, and $z_{n,r} \in \mathbb{C}$. Fig. 2 shows a graphical illustration of (7) when $N = 3$, implying that each component, rank-1 tensor in 3-D exponential signal (7), can be written as the outer product of vectors that are of exponential structure. Here we assume $R \ll \prod_{n=1}^{N} I_n$.

In this paper, we aim to recover $\mathcal{Y}$ from its small subset of entries $\mathcal{P}_\Omega(\mathcal{Y})$,

$$[\mathcal{P}_\Omega(\mathcal{Y})]_{i_1,\ldots,i_N} = \begin{cases} y_{i_1,\ldots,i_N} & \text{if } (i_1,\ldots,i_N) \in \Omega, \\ 0 & \text{otherwise.} \end{cases} \quad (8)$$

One can check that $\mathcal{Y}$ is low-rank in the sense of both CP-rank and n-rank if R is sufficiently small. Therefore, the problem of reconstructing the signal $\mathcal{Y}$ can be recast as a low-rank *tensor completion*. Depending on how the tensor rank is defined, we have a couple of possible N-D exponential signal reconstruction approaches using available low-rank tensor completion methods.

A. Low-CP-rank tensor completion

The signal $\mathcal{Y}$ enjoys a low CP-rank. Actually, (7) implies the following CP-decomposition

$$\mathcal{Y} = [\![\mathbf{d}; \mathbf{A}^{(1)}, \mathbf{A}^{(2)}, \ldots, \mathbf{A}^{(N)}]\!], \quad (9)$$

where $\mathbf{d} = [d_1, \ldots, d_R]^T$ and the factor matrices $\mathbf{A}^{(n)} \in \mathbb{C}^{I_n \times R}$ are

$$\mathbf{A}^{(n)} = [\mathbf{a}_1^{(n)}, \ldots, \mathbf{a}_R^{(n)}] \quad (10)$$

with

$$\mathbf{a}_r^{(n)} = [1, z_{n,r}, z_{n,r}^2, \ldots, z_{n,r}^{I_n-1}]^T. \quad (11)$$

These equations imply that $\mathcal{Y}$ has a CP-rank at most R. Once R is relatively small, $\mathcal{Y}$ is low-CP-rank.
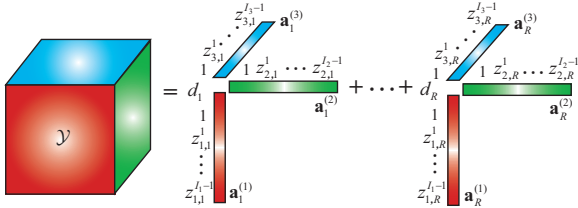

Fig. 2: CP decomposition for the signal of interest in 3-D.

Therefore, reconstructing $\mathcal{Y}$ from its partial entries can be formulated as a low-CP-rank tensor completion problem. There exist several methods available in the literature to solve this problem and generally they can be categorized into non-convex and convex methods. The weighted CP (WCP) decomposition method [39], [40] is a typical non-convex method, and it recovers tensors by solving the following non-convex optimization

$$\min_{\mathbf{U}^{(1)},\cdots,\mathbf{U}^{(N)}} \left\| \mathcal{P}_\Omega \left(\mathcal{Y} - [\![\mathbf{U}^{(1)}, \mathbf{U}^{(2)}, \cdots, \mathbf{U}^{(N)}]\!] \right) \right\|_F^2, \quad (12)$$

where $\mathbf{U}^{(n)} \in \mathbb{C}^{I_n \times \hat{R}}$ are the factor matrices, in which $\hat{R}$ is an estimated rank. Regarding convex methods, the following optimization was proposed [41]

$$\min_{\mathcal{X}} \|\mathcal{X}\|_* \quad \text{s.t.} \quad \mathcal{P}_\Omega \mathcal{X} = \mathcal{P}_\Omega \mathcal{Y}, \quad (13)$$

where $\|\mathcal{X}\|_*$ is the tensor nuclear norm defined by

$$\|\mathcal{X}\|_* = \sup_{\|\mathbf{a}_n\|_2=1,\ 1\le n\le N} \langle \mathcal{X}, \mathbf{a}_1 \circ \mathbf{a}_2 \circ \ldots \circ \mathbf{a}_N \rangle.$$

It was shown that the solution (13) will be $\mathcal{Y}$ if the number of known entries exceeds a certain amount [41]. Unfortunately, (13) is NP-hard and computationally intractable. Recently some computational methods [47], [48] are proposed to approximate or equally achieve the tensor nuclear norm.

B. Low-n-rank tensor completion

The signal $\mathcal{Y}$ enjoys low n-rank as long as R is small compared with $\min\{I_1, I_2, \ldots, I_N\}$, since the rank of the matricization $\mathbf{Y}_{(n)}$ of $\mathcal{Y}$ is at most R if $R \le I_n$. Thus, recovering $\mathcal{Y}$ from its partial entries can be viewed as a low-n-rank tensor completion problem. One finds a tensor that has a small n-rank [36], [38]

$$\min_{\mathcal{X}} \sum_{n=1}^{N} \alpha_n \text{rank}\left(\mathbf{X}_{(n)}\right), \quad \text{s.t.} \quad \mathcal{P}_\Omega\left(\mathcal{X}\right) = \mathcal{P}_\Omega\left(\mathcal{Y}\right), \quad (14)$$

where α_n is the weight and $\mathbf{X}_{(n)}$ denotes the mode-n matricization.

One may approximate the non-convex objective in (14) by a convex one. The only non-convexity in (14) is the rank function. It is well known [25], [49] that the best convex approximation of the rank function is the nuclear norm, i.e., the sum of singular values. Also, to handle data with noise, the linear constraint (14) may be replaced by a least square fitting term. Altogether, one gets a convex optimization for low-n-rank tensor completion as follows

$$\min_{\mathcal{X}} \sum_{n=1}^{N} \alpha_n \|\mathbf{X}_{(n)}\|_* + \frac{\lambda}{2} \|\mathcal{P}_\Omega\left(\mathcal{Y}\right) - \mathcal{P}_\Omega\left(\mathcal{X}\right)\|_F^2, \quad (15)$$

where $\|\cdot\|_*$ denotes the nuclear norm of a matrix and λ is the regularization parameter. When $\alpha_n = 1$ for all $n = 1, \ldots, N$, (15) is the convex model proposed in [38]. Low-n-rank tensor completion has been successfully applied in computer vision [36] and remote sensing data analyses [38].

IV. THE PROPOSED METHOD

Though generic low-rank tensor completion methods discussed in Section III are applicable to our N-D exponential signal reconstruction, they ignore the specific exponential structure of the factor vectors. From (9), it is observed that each factor matrix $\mathbf{A}^{(n)}$ defined in (10) is *Vandermonde matrix* and each factor vector in (11) is an exponential function. As a consequence, they will need unnecessarily large number of measurements for a stable reconstruction of $\mathcal{Y}$. Take $N = 3$ and $I_1 = I_2 = I_3 = I$ as an example. As discussed in [41], (13) needs $O\left(R^{1/2}(I \log I)^{3/2}\right)$ known entries to give a robust recovery for low-CP-rank tensor completion, and matrix completion theory suggests that $O(RI^2 \log^2(I))$ observed entries are necessary for a reliable recovery for low-n-rank tensor completion (15). However, there are only $O(R)$ degree of freedoms in the signal model (7). Therefore, it is expected that, if we explore the exponential structure of the factor vectors, we can design an N-D exponential signal reconstruction method that requires much fewer measurements than generic low-rank tensor completion methods.

In the following, we propose an approach that utilizes the exponential structure of the factor vectors, in addition to the low-CP-rank structure of the signal.

A. The proposed model

We will use the low-CP-rank structure of the tensor $\mathcal{Y}$, although the low-n-rank property is applicable as well. To promote the exponential structure of the factor vector, we enforce the Hankel matrix of each factor vector to be low-rank by nuclear norm. We propose the following reconstruction model,

$$\min_{\substack{\mathbf{u}_r^{(n)} \\ r=1,\ldots,\hat{R} \\ n=1,\ldots,N}} \sum_{r=1}^{\hat{R}} \sum_{n=1}^{N} \|\mathcal{R}\mathbf{u}_r^{(n)}\|_*$$
$$+ \frac{\lambda}{2} \left\| \mathcal{P}_\Omega \Big(\sum_{r=1}^{\hat{R}} \mathbf{u}_r^{(1)} \circ \mathbf{u}_r^{(2)} \circ \cdots \circ \mathbf{u}_r^{(N)}\Big) - \mathcal{P}_\Omega(\mathcal{Y}) \right\|_F^2, \quad (16)$$

where λ is a regularization parameter that trades off the nuclear norm against the data consistency and $\mathcal{R}$ is a linear operator defined as $\mathcal{R} : \mathbb{C}^{I_n} \to \mathbb{C}^{S_1^{(n)} \times S_2^{(n)}}$, for some integers $S_1^{(n)}$ and $S_2^{(n)}$ satisfying $S_1^{(n)} + S_2^{(n)} = I_n + 1$, as follows

$$[\mathcal{R}\mathbf{a}_r^{(n)}]_{k,l} = [\mathbf{a}_r^{(n)}]_{k+l-1}, \quad \forall \ 1 \le k \le S_1^{(n)}, \ 1 \le l \le S_2^{(n)}.$$

The $S_1^{(n)}$ is chosen to make the Hankel matrix square or approximately square to minimize the reconstruction error [35]. We call the proposed model (16) Hankel Matrix nuclear norm Regularized low-CP-rank Tensor Completion (HMRTC).

The proposed HMRTC simultaneously exploits the low-CP-rank structure and the exponential structure of the associated factor vectors, which makes it superior to existing methods. 1) HMRTC utilizes the exponential structure of factor vectors, being demonstrated in Section V that HMRTC can significantly reduce the number of necessary samples, compared with generic low-rank tensor completion methods that do not consider the structure of the factor matrices. 2) HMRTC represents the unknown tensor in a CP decomposition form, significantly reducing the size of variables in numerical algorithms, compared with atomic norm minimization (e.g. [17]) and other low-rank structured matrix methods (e.g. EMaC [35]). In particular, as we will see in Section IV-B, the numerical algorithm of HMRTC invokes $O(\hat{R}N)$ matrices of size $O(I^2)$ if $I_1 = \ldots = I_N = I$. As a comparison, the standard alternating direction method of multipliers (ADMM) algorithm for solving atomic norm minimization and EMaC needs $O(1)$ huge matrices of size $O(I^{2N})$. Consequently, our proposed HMRTC can easily reconstruct 3-D exponential signals of size $50 \times 50 \times 50$, which, however, is a prohibitive task for EMaC or atomic norm minimization.

We conclude this subsection by a toy example to demonstrate the great potential of the proposed HMRTC. It is well known that neither the low-n-rank tensor completion nor WCP can recover missing slices [25], [39] since the information is unknown in the missing part. On the contrary, HMRTC is able to estimate these missing data because the exponential structure of one single factor is exploited. Fig. 3 shows that HMRTC can perform well even when half of the slices are missing. Thus, HMRTC will be useful for those applications requiring recover the truncated data in the end of signals to improve the frequency resolution or increase the signal-to-noise ratio [1], [50].

B. Numerical algorithm

By combining the unknown factors $\mathbf{u}_n^{(r)}$ into factor matrices $\mathbf{U}^{(n)} = \left[\mathbf{u}_1^{(n)} \ldots \mathbf{u}_R^{(n)}\right]$ for $n = 1, \ldots, N$, Eq. (16) is rewritten concisely as

$$\min_{\substack{\mathbf{U}^{(n)} \\ n=1,\ldots,N}} \sum_{r=1}^{\hat{R}} \sum_{n=1}^{N} \|\mathcal{R}\mathcal{Q}_r \mathbf{U}^{(n)}\|_*$$
$$+ \frac{\lambda}{2} \big\| \mathcal{P}_\Omega(\mathcal{Y}) - \mathcal{P}_\Omega([\![\mathbf{U}^{(1)}, \mathbf{U}^{(2)}, \cdots, \mathbf{U}^{(N)}]\!]) \big\|_F^2, \quad (17)$$

where $\mathcal{Q}_r$ extracts the r-th column from $\mathbf{U}^{(n)}$ for $r = 1, \ldots, R$ and $n = 1, \ldots, N$.

Recently, it has been shown in [37] and [51] that the ADMM is very efficient for some convex or nonconvex problems in various applications, such as medical imaging [52]. To solve (17), we also propose an algorithm based on ADMM. Some auxiliary variables, $\mathbf{Z}_r^{(n)} = \mathcal{R}Q_r\mathbf{U}^{(n)}$, $n = 1, \ldots, N$ and $r = 1, \ldots, \hat{R}$ are introduced and then (17) is reformulated into the following equivalent form:

$$\min_{\substack{\mathbf{U}^{(n)} \\ n=1,\ldots,N \\ r=1,\ldots,\hat{R}}} \sum_{r=1}^{\hat{R}} \sum_{n=1}^{N} \|\mathbf{Z}_r^{(n)}\|_*$$
$$+ \frac{\lambda}{2} \big\| \mathcal{P}_\Omega(\mathcal{Y}) - \mathcal{P}_\Omega([\![\mathbf{U}^{(1)}, \mathbf{U}^{(2)}, \ldots, \mathbf{U}^{(N)}]\!]) \big\|_F^2 \quad (18)$$
$$\text{s.t.} \ \mathbf{Z}_r^{(n)} = \mathcal{R}Q_r\mathbf{U}^{(n)}, \forall \ 1 \le r \le \hat{R}, \ 1 \le n \le N.$$

For ease of presentation, we define

$$\mathcal{U} = \big\{\mathbf{U}^{(1)}, \ldots, \mathbf{U}^{(N)}\big\},$$

$$\mathcal{Z} = \big\{\mathbf{Z}_1^{(1)}, \ldots, \mathbf{Z}_1^{(N)}, \ldots, \mathbf{Z}_{\hat{R}}^{(1)}, \ldots, \mathbf{Z}_{\hat{R}}^{(N)}\big\},$$

and

$$\mathcal{D} = \big\{\mathbf{D}_1^{(1)}, \ldots, \mathbf{D}_1^{(N)}, \ldots, \mathbf{D}_{\hat{R}}^{(1)}, \ldots, \mathbf{D}_{\hat{R}}^{(N)}\big\}.$$

The augmented Lagrangian function of (18) is

$$\mathcal{L}_\beta(\mathcal{U}, \mathcal{Z}, \mathcal{D}) = \sum_{r=1}^{\hat{R}} \sum_{n=1}^{N} \big(\langle \mathbf{D}_r^{(n)}, \mathcal{R}Q_r\mathbf{U}^{(n)} - \mathbf{Z}_r^{(n)} \rangle$$
$$+ \|\mathbf{Z}_r^{(n)}\|_* + \frac{\beta}{2}\|\mathcal{R}Q_r\mathbf{U}^{(n)} - \mathbf{Z}_r^{(n)}\|_F^2\big) \quad (19)$$
$$+ \frac{\lambda}{2}\big\|\mathcal{P}_\Omega(\mathcal{Y}) - \mathcal{P}_\Omega([\![\mathbf{U}^{(1)}, \mathbf{U}^{(2)}, \ldots, \mathbf{U}^{(N)}]\!])\big\|_F^2,$$

where $\mathbf{D}_r^{(n)}$ is the matrix of Lagrange multipliers for $n = 1, \ldots, N$ and $r = 1, \ldots, \hat{R}$.

The ADMM is an iterative algorithm. Given $\mathcal{U}_k$, $\mathcal{Z}_k$, and $\mathcal{D}_k$ at step k, it updates $\mathcal{U}$, $\mathcal{Z}$, and $\mathcal{D}$ as follows.

a) : The variable $\mathcal{U}$ is updated by solving the following optimization,

$$\min_{\mathcal{U}} \mathcal{L}_{\beta_k}(\mathcal{U}, \mathcal{Z}_k, \mathcal{D}_k). \quad (20)$$

Due to the multi-linearity of the CP decomposition, it is not easy to solve (20) exactly. Here, we employ an alternating minimization procedure to solve (20) approximately. Fixing
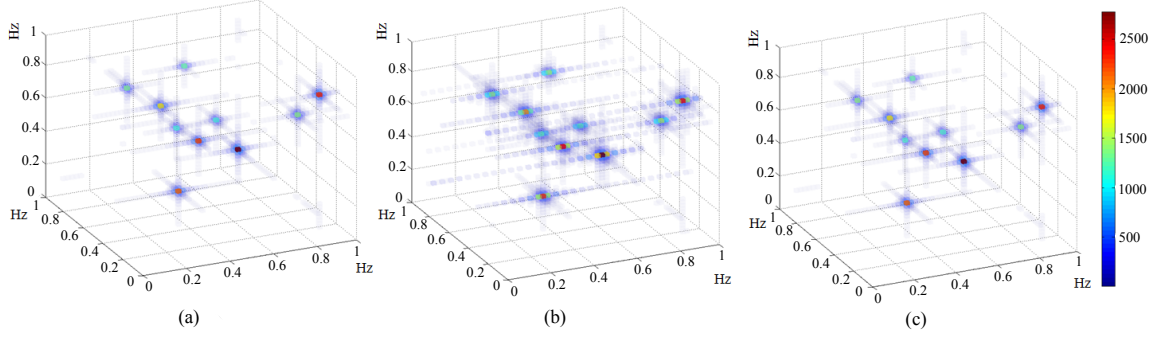
Fig. 3: The 3-D spectra recovered by the HMRTC from a tensor with half of slices missing. (a) The ground truth spectrum; (b) Reconstructed spectrum by filling zeroes into missing data points in the time domain and then performing Fourier transform on the full time domain signal; (c) Reconstructed spectrum by HMRTC. Note: We simulate a tensor with 10 damped complex sinusoids and discard half of slices. We draw 3-D spectra of tensors by using the tensor toolbox [53]. The color stands for magnitude of spectra.

$\mathbf{U}^{(1)}, \ldots, \mathbf{U}^{(n-1)}, \mathbf{U}^{(n+1)}, \ldots, \mathbf{U}^{(N)}$, we solve (20) with respect to $\mathbf{U}^{(n)}$, which is a convex optimization as follows:

$$
\min_{\mathbf{U}^{(n)}} \sum_{r=1}^{\hat{R}} \left(\frac{\beta_k}{2} \left\| \mathcal{R}Q_r \mathbf{U}^{(n)} - \mathbf{Z}_{r;k}^{(n)} + (\beta_k)^{-1} \mathbf{D}_{r;k}^{(n)} \right\|_F^2 \right) \tag{21}
$$
$$
+ \frac{\lambda}{2} \left\| \mathcal{P}_{\Omega^{(n)}}(\mathbf{Y}_{(n)}) - \mathcal{P}_{\Omega^{(n)}}(\mathbf{U}^{(n)} \mathbf{G}_k^{(n)}) \right\|_F^2,
$$

where $\mathbf{G}_k^{(n)} = (\mathbf{U}_k^{(N)} \odot \cdots \odot \mathbf{U}_k^{(n+1)} \odot \mathbf{U}_{k+1}^{(n-1)} \odot \cdots \odot \mathbf{U}_{k+1}^{(1)})^T$; $\mathbf{D}_{r;k}^{(n)}$ and $\mathbf{Z}_{r;k}^{(n)}$ are the k-th update of $\mathbf{D}_r^{(n)}$ and $\mathbf{Z}_r^{(n)}$; $\mathbf{Y}_{(n)}$ and $\mathcal{P}_{\Omega^{(n)}}(\mathbf{Y}_{(n)})$ are the mode-n matricization of the tensors $\mathcal{Y}$ and $\mathcal{P}_\Omega(\mathcal{Y})$ respectively. It is obvious (21) is a least squares problem, and therefore its solution is a solution of linear system. In particular, if we define $\mathcal{A}_k^{(n)}(\mathbf{X}) \triangleq \mathcal{P}_{\Omega^{(n)}}\left(\mathbf{X}\mathbf{G}_k^{(n)}\right)$, then (21) is rewritten as

$$
\min_{\mathbf{U}^{(n)}} \sum_{r=1}^{\hat{R}} \left(\frac{\beta_k}{2} \left\| \mathcal{R}Q_r \mathbf{U}^{(n)} - \mathbf{Z}_{r;k}^{(n)} + (\mu^k)^{-1} \mathbf{D}_{r;k}^{(n)} \right\|_F^2 \right) \tag{22}
$$
$$
+ \frac{\lambda}{2} \left\| \mathcal{P}_{\Omega^{(n)}}(\mathbf{Y}_{(n)}) - \mathcal{A}_k^{(n)}\left(\mathbf{U}^{(n)}\right) \right\|_F^2,
$$

whose solution satisfies

$$
\lambda \mathcal{A}_k^{(n)*} \mathcal{A}_k^{(n)} \mathbf{U}^{(n)} + \beta_k \sum_{r=1}^{\hat{R}} \mathcal{Q}_r \mathcal{R}\mathcal{R} \mathcal{Q}_r \mathbf{U}^{(n)}
$$
$$
= \lambda \mathcal{A}_k^{(n)*}\left(\mathcal{P}_{\Omega^{(n)}}(\mathbf{Y}_{(n)})\right) \tag{23}
$$
$$
+ \beta_k \sum_{r=1}^{\hat{R}} \mathcal{Q}_r^* \mathcal{R}^* \left(\mathbf{Z}_{r;k}^{(n)} - (\beta_k)^{-1} \mathbf{D}_{r;k}^{(n)} \right),
$$

The closed-form of $\mathbf{U}^{(n)}$ are derived in Appendix A.

b) : The variable $\mathcal{Z}$ is updated via the following optimization

$$
\min_{\mathcal{Z}} \mathcal{L}_{\beta_k}(\mathcal{U}_{k+1}, \mathcal{Z}, \mathcal{D}_k). \tag{24}
$$

which is equivalent to $N\hat{R}$ independent sub-problems

$$
\min_{\mathbf{Z}_r^{(n)}} \left\| \mathbf{Z}_r^{(n)} \right\|_* + \frac{\beta_k}{2} \left\| \mathcal{R}Q_r \mathbf{U}_{k+1}^{(n)} - \mathbf{Z}_r^{(n)} + (\beta_k)^{-1} \mathbf{D}_{r;k}^{(n)} \right\|_F^2. \tag{25}
$$

for $n = 1, \ldots, N$ and $r = 1, \ldots, \hat{R}$. Following [49], the closed-form solution of (25) is

$$
\mathbf{Z}_r^{(n)} = S_{1/\beta_k}\left(\mathcal{R}\mathcal{Q}_r \mathbf{U}_{k+1}^{(n)} + (\beta_k)^{-1} \mathbf{D}_{r;k}^{(n)} \right), \tag{26}
$$

where S_{1/β_k} is the soft singular value thresholding operator with parameter $1/\beta_k$.

c) : Update $\mathbf{D}_r^{(n)}$ by

$$
\mathbf{D}_{r;k+1}^{(n)} = \mathbf{D}_{r;k}^{(n)} + \beta_k \left(\mathcal{R}Q_r \mathbf{U}_{k+1}^{(n)} - \mathbf{Z}_{r;k+1}^{(n)} \right). \tag{27}
$$

The full algorithm is described in Algorithm 1. This algorithm can also be accelerated by setting $\beta_{k+1} = \rho\beta_k$, where $\rho \in (1.0, 1.1)$ [36], [37], [54].

Algorithm 1 The pseudo code of the proposed algorithm

Input: The tensor $\mathcal{Y} \in \mathbb{C}^{I_1 \times \cdots \times I_N}$, the set Ω, and parameters $\hat{R}, \lambda$;

Output: The reconstruction $\mathcal{X} = [\![\mathbf{U}^{(1)}, \mathbf{U}^{(2)}, \cdots, \mathbf{U}^{(N)}]\!]$ of $\mathcal{Y}$;

1: Initialize $\mathbf{U}^{(1)}, \ldots, \mathbf{U}^{(N)}$, $k = \Delta x = 1$, $\beta_0 = 0.1$ and $\rho = 1.05$.

2: **while** $\Delta x < 10^{-4}$ and $k < 10^3$ **do**

3:     Update $\mathbf{U}^{(1)}, \mathbf{U}^{(2)}, \cdots, \mathbf{U}^{(N)}$ by solving (23);

4:     Update $\mathbf{Z}_r^{(n)}$ by solving (26) for $r = 1, \ldots, \hat{R}$, $n = 1, \ldots, N$;

5:     Update $\mathbf{D}_r^{(n)}$ by solving (27) for $r = 1, \ldots, \hat{R}$, $n = 1, \ldots, N$;

6:     Update β by $\beta_{k+1} = \rho\beta_k$;

7:     $\mathcal{X} = [\![\mathbf{U}^{(1)}, \mathbf{U}^{(2)}, \cdots, \mathbf{U}^{(N)}]\!]$;

8:     $\Delta x = \|\mathcal{X} - \mathcal{X}_{\text{last}}\|_F / \|\mathcal{X}_{\text{last}}\|_F$; $\mathcal{X}_{\text{last}} \leftarrow \mathcal{X}$;

9:     $k = k + 1$;

10: **end while**

C. Convergence analysis

We provide the convergence of Algorithm 1 stated in **Theorem 1** and **Theorem 2**. From the theorems, we know that the sequence $\{\mathcal{U}_k\}$ generated by Algorithm 1 converges. Furthermore, if we further impose some condition on the

Lagrange multipliers $\{\mathcal{D}_k\}$, then the limit is a critical point of (17).

Theorem 1 *The sequence $\{\mathcal{U}_k\}$ generated by Algorithm 1 is a Cauchy sequence.*

Theorem 2 *If $\lim_{k\to\infty}\left\|\mathbf{D}_{r;k+1}^{(n)} - \mathbf{D}_{r;k}^{(n)}\right\|_F = 0$, for all $r \in \{1,\ldots,\hat{R}\}$, $n \in \{1,\ldots,N\}$, then the limit of $\{\mathcal{U}_k\}$ satisfies the KKT condition for (17).*

Proofs of **Theorem 1** and **Theorem 2** are added in Appendix C.

D. Complexity analysis

The computational complexity of HMRTC is analyzed here. Besides the typical tensor operation such as Tucker operation, the running time of Algorithm 1 is dominated by the singular value decomposition (SVD) for the singular value thresholding operator in (26). Consider to recover a tensor $\mathcal{X} \in \mathbb{C}^{I_1 \times \cdots \times I_N}$ with $I_1 = \ldots = I_N = I$. The SVD of $\mathbf{Z}_r^{(n)}$, which is of a small size $0.5I \times 0.5I$, can be done in $O(I^3)$ operations. Since we have $N\hat{R}$ SVDs to compute at each iteration, the total computational complexity for SVD in each iteration is $O(N\hat{R}I^3)$, which is only sub-linear with the tensor size I^N when $N \geq 3$. Furthermore, the SVDs in a single iteration can be computed in parallel, since each $\mathbf{Z}_r^{(n)}$, $r = 1,\ldots,\hat{R}$, $n = 1,\ldots,N$, is updated independently. Therefore, HMRTC has the potential to be applied in N-D ($N \geq 3$) exponential signal recovery shown in Section VI-A. In addition, from the spatial complexity point of view, the storage $O(N\hat{R}I)$ of HMRTC can be significantly smaller than that of the original tensor when the estimated tensor rank $\hat{R}$ is much smaller than I.

V. EXPERIMENTAL RESULTS

In this section, we will evaluate the proposed HMRTC on simulated exponential signals, including undamped and damped complex sinusoids, and real NMR spectroscopy data. Two state-of-the-art algorithms of low-rank tensor completion, the alternating direction method-based tensor recovery (ADM-TR) [38] and WCP [39], are compared with HMRTC.

The parameters of HMRTC are listed in Algorithm 1. For ADM-TR and WCP, the maximum number of iterations is 10^3. In ADM-TR, the parameters are $c_\lambda = c_\beta = 1$ and $\beta = 1$. The alternating least squares method is applied to solve WCP and the algorithm is terminated when $\mathrm{abs}\,(f_k - f_{k+1})/(1 + f_k) \leq 10^{-6}$, where f_k is the objective function value in (12) after iteration k.

A. Experiment setup for simulated data

The proposed algorithm is tested on 3-D simulated signals $\mathcal{Y} \in \mathbb{C}^{50 \times 50 \times 50}$ as follows. The frequency $f_{n,r}$ is uniform randomly drawn from the interval $[0, 1)$ for all $1 \leq r \leq R$ and $1 \leq n \leq 3$, where R is the number of exponentials. The coefficient d_r is generated by $d_r = 1 + 10^{0.5m_r}$ and the damping factor $\tau_{n,r}$ is generated by $\tau_{n,r} = 10 + 30g_{n,r}$,

where m_r and $g_{n,r}$ follow the uniformly random distribution on $[0, 1]$. The undamped complex sinusoid is synthesized as

$$y_{i_1,i_2,i_3} = \sum_{r=1}^{R} d_r \left(e^{j2\pi f_{1,r}}\right)^{i_1} \left(e^{j2\pi f_{2,r}}\right)^{i_2} \left(e^{j2\pi f_{3,r}}\right)^{i_3}, \quad (28)$$

and damped complex sinusoid is simulated as

$$y_{i_1,i_2,i_3} = \sum_{r=1}^{R} d_r \left(e^{-1/\tau_{1,r}+2\pi j f_{1,r}}\right)^{i_1} \\ \left(e^{-1/\tau_{2,r}+2\pi j f_{2,r}}\right)^{i_2} \left(e^{-1/\tau_{3,r}+2\pi j f_{3,r}}\right)^{i_3}, \quad (29)$$

for $i_n = 0, \cdots, 49$ and n=1, 2, 3.

In each experiment, the simulated ground-truth tensor $\mathcal{Y}$ is normalized by dividing the maximum magnitude of its entries. Gaussian white noise with standard deviation σ is added on both real and imaginary parts of $\mathcal{Y}$.

We denote by $\mathcal{X}$ the reconstruction of HMRTC, i.e., the output of Algorithm 1, and the relative least normalized error (RLNE) is defined as

$$\mathrm{RLNE} = \|\mathcal{X} - \mathcal{Y}\|_F / \|\mathcal{Y}\|_F. \quad (30)$$

The average RLNE is calculated by averaging the RLNEs over 10 Monte Carlo trials as conducted in [36], [55]. To facilitate comparison in color map, the average RLNE is set to be 1 if it is larger than 1. In each trial, the observed entries are sampled in the uniformly random fashion and the sampling ratio (SR) is denoted as the proportion of available data to the full data.

The numerical experiments are conducted on a Dell PC running Windows 7 operating system with Intel Core i7 2600 CPU and 12-GB RAM. The average computation time for ADM-TR, WCP and HMRTC to recover simulated signals with $R = 50$ and $SR = 0.5$ is 185s, 7395s and 675s, respectively. The average memory requirements for ADM-TR, WCP and HMRTC is 0.69GB, 0.65GB and 0.66GB, respectively.

B. Robustness to the estimated tensor rank

We first evaluate the robustness of WCP and HMRTC to the estimated tensor rank. ADM-TR is not compared since it does not require an estimation of the tensor rank.

The comparison is presented in Fig. 4. It is observed that the HMRTC can always achieve low reconstruction errors (RLNE$\leq$0.1) as $\hat{R}$ increases from $2R$ to $10R$, while WCP fails to recover if R is over-estimated too much. In particular, when the ground truth $R = 60$, as shown in Fig. 4(b), WCP can work only when $\hat{R} = 2R$ but HMRTC is robust to any $\hat{R}$ chosen from $2R$ to $10R$. Hence, HMRTC will be more useful for those applications where it is difficult to estimate the true number of exponentials. In the following simulated experiments, we set $\hat{R} = 2R$ in WCP and $\hat{R} = 100$ in HMRTC.

C. Recovery of simulated complex sinusoids

In this subsection, undamped and damped complex sinusoids are simulated to evaluate the construction performance. The regularization parameter λ in ADM-TR is set be 10, and be
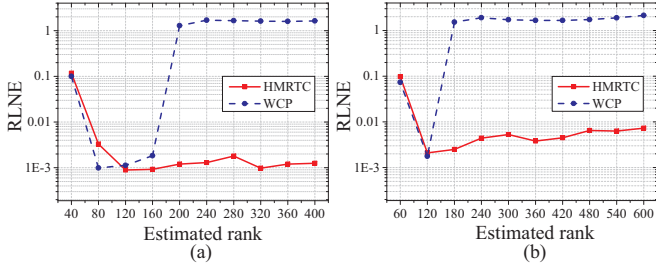
Fig. 4: RLNE versus the estimated tensor rank. (a) and (b) illustrate reconstruction errors of WCP and HMRTC with different estimated ranks $\hat{R}$ chosen from $\{R, 2R, \dots, 10R\}$. The sampling ratio SR is 0.3 and the number of exponentials R in (a) and (b) is 40 and 60, respectively. Here $\rho = 1.02$.

10^2 and 10^3 in HMRTC for data with noise level 0.01 and 0.005, respectively.

Fig. 6 shows that HMRTC yields an average RLNE that is much smaller than ADM-TR and WCP for all R's and SR's, no matter in recovering undamped or damped sinusoids. Furthermore, for a fixed R, HMRTC needs a much smaller SR than ADM-TR and WCP to achieve an average RLNE within noise level, implying that HMRTC requires a much smaller number of sampled entries for the stable recovery than ADM-TR and WCP.

Fig. 5 further examines the stability of the proposed algorithm at different noise levels. It is shown that HMRTC produces reconstructions with the lowest RLNE, as well as the smallest variance, among the three methods, implying that HMRTC is the most robust to noise. The spectra of selected fibers in the reconstruction in Fig. 5(c) and Fig. 5(d) indicate that HMRTC leads to most consistent spectra to the ground-truth.

D. Recovery of real NMR spectroscopy data

NMR spectroscopy has been an indispensable tool in the study of structure, dynamics, and interactions of biopolymers in chemistry and biology. The duration of an N-D NMR spectroscopy experiment is proportional to the number of measured data points, and the nonuniform sampling of time domain signal can dramatically reduce measurement time [1], [56]–[58].Here we apply the proposed HMRTC to recover full spectrum in fast NMR, since the time domain signal of NMR is generally modeled as damped complex sinusoids [1].

A 3-D HNCO spectrum is tested and its sample is the U-[^{15}N, ^{13}C] RNA recognition motifs domain of protein RNA binding motif protein 5 [59], which is a component of the spliceosome A-complex. The ADM-TR, WCP and HMRTC are compared in recovering this 3-D spectrum with the size of $64 \times 128 \times 512$ from a 3-D Poisson-gap [56] nonuniformly sampled time-domain data. All the spectra are processed in NMRPipe [60] using a routine processing manner.

Fig. 7 shows that HMRTC leads to the most faithful recovery (Fig. 7(d)) of the ground truth (Fig. 7(a)) than ADM-TR (Fig. 7(b)) and WCP (Fig. 7(c)). Moreover, Fig. 7(d) indicates that HMRTC can achieve high quality of reconstruction even with the sampling rate 10% and hence allow a significant reduction in measurement time. Thus, the proposed HMRTC algorithm
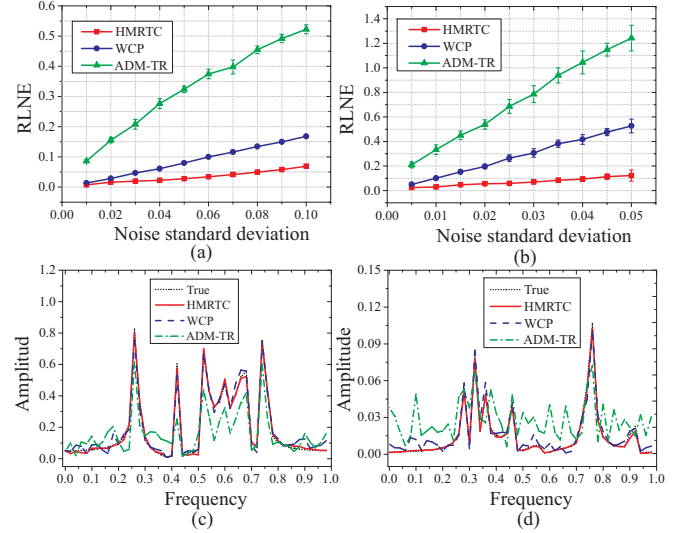

Fig. 5: Reconstruction results of ADM-TR, WCP and HMRTC from noisy data: (a) and (b) The average RLNE versus noise standard deviation σ in recovering undamped and damped complex sinusoid with $R = 10$, respectively. The sampling ratio SR is set to be 0.3 and 0.5 in (a) and (b), respectively. (c) and (d) The spectra of chosen fibers in reconstruction in the undamped case with $\sigma = 0.1$ and damped case with $\sigma = 0.025$, respectively.

may serve as a versatility method for studying biological or chemical molecules using N-D NMR spectroscopy.

VI. DISCUSSIONS

In the following, a RLNE smaller than 0.1 is considered as a low reconstruction error since the corresponding energy loss is less than 1%.

A. Higher-dimensional experiments

A 5-D experiment is conducted to explore the capability of HMRTC in higher-dimensional exponential signal recovery. Fig. 8(a) indicates that HMRTC achieves a low reconstruction error although R is as large as around 6 times of ambient dimensions. The reconstruction error will increase dramatically when R further increases ($R \geq 250$). Therefore, HMRTC has the potential to reconstruct the signal with a quite larger R than ambient dimensions. However, reconstructing a signal with larger R needs further development.

B. Comparison with other state-of-the-art methods

Another three state-of-the-art methods, ADMM-R [48], TNN [47] and FaLRTC [36], are compared with the proposed HMRTC on undamped complex sinusoid recovery. Since tensor nuclear norm is computationally intractable, the approximate solution was proposed recently in [47] by providing sub-optimality guarantees. More recently, another computational method to equally achieve tensor nuclear norm was proposed in a preprint paper [48]. FaLRTC [36] is another typical method in tensor completion which minimizes low n-rank.
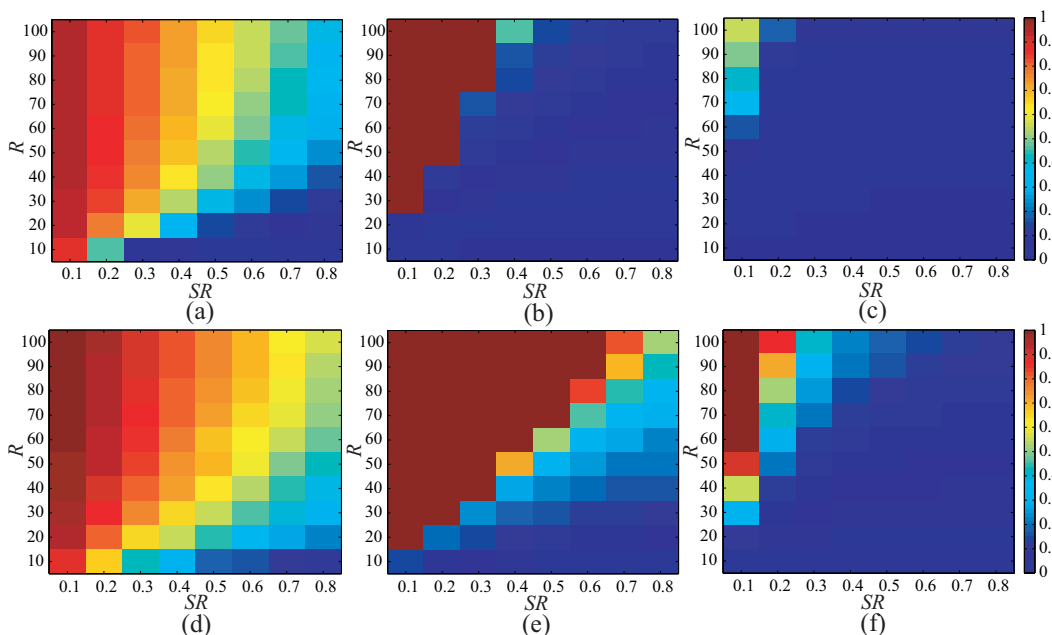
Fig. 6: Reconstruction errors of ADM-TR, WCP and HMRTC. (a), (b) and (c) reflect the average RLNEs by ADM-TR, WCP and HMRTC, respectively, in recovering undamped complex sinusoids with noise level $\sigma = 0.01$. (d), (e) and (f) indicate the average RLNEs by ADM-TR, WCP and HMRTC, respectively, in recovering damped complex sinusoids with noise level $\sigma = 0.005$.
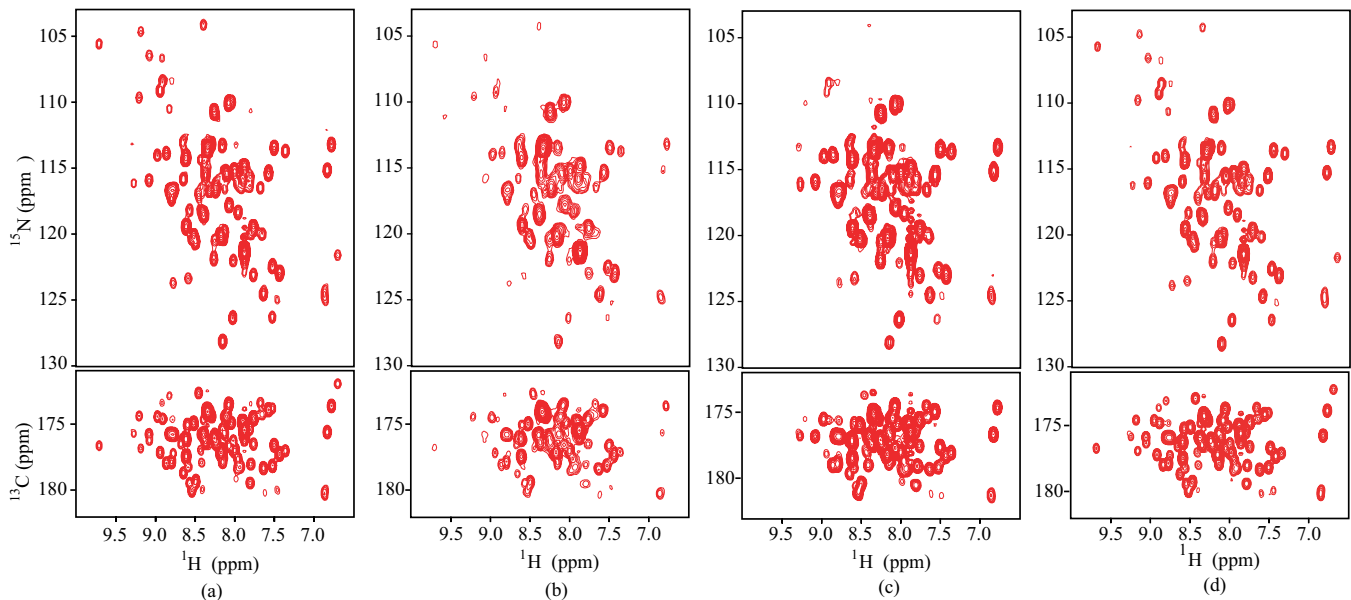

Fig. 7: The ^{1}H-^{15}N and ^{1}H-^{13}C skyline projection spectra in the 3-D HNCO experiment. (a) the uniformly-sampled spectrum; (b), (c) and (d) are the ADM-TR, WCP and HMRTC reconstruction using 10% sampled data, respectively. The skyline projection spectra [61], plotted by summing all the frequency points along the rest dimensions, are used here to simplify the description of spectra. The experiment is carried out on a Bruker AVANCE III 600 MHz spectrometer equipped with a cryogenic probe at 293K. The estimated rank in WCP and HMRTC is 200 and 500, respectively. The λ in ADM-TR and HRMTC is 10 and 10^3, respectively. The ppm denotes parts per million, the unit of chemical shift.

Fig. 9 shows that HMRTC holds advantage over ADMM-R [48], TNN [47] and FaLRTC [36] on achieving much lower reconstruction errors. Note that the estimated tensor rank $\hat{R}$ is set to be exactly ground-truth number of exponentials in ADMM-R [48], which may be unknown in practice.

C. Success rate of factor vector recovery

The low-CP-rank completion method may lead the factor vector to be consisted of single or multiple exponentials in real applications [63]. Considering the proposed model also explores the low-CP-rank structure, one may wonder the
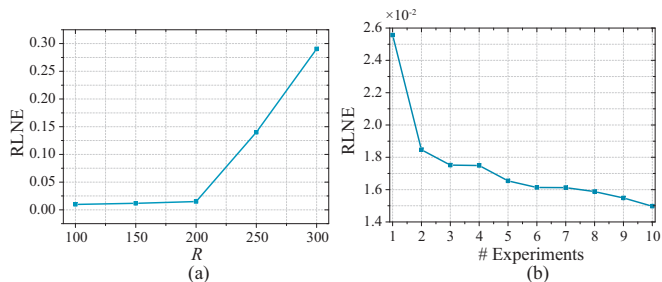
Fig. 8: (a) The reconstruction error by HMRTC on 5-D data with the size 30^5 and sampling rate $SR = 0.5$. (b) The effect of different initializations on the 3-D data in Fig. 3 with sampling ratio $SR = 0.3$. We sort the RLNE for better visualization though these RLNEs are randomly distributed in experiments.

possibility of the reconstructed factor vectors being a single exponential as it was introduced in (1). Here we discuss this possibility based on numerical experiments.

Following [64], the factor vector recovery is declared successful if

$$\text{sim}\,(r) = \frac{\left|\mathbf{a}_r^H \hat{\mathbf{a}}_r\right|}{\|\mathbf{a}_r\|\,\|\hat{\mathbf{a}}_r\|} \times \frac{\left|\mathbf{b}_r^H \hat{\mathbf{b}}_r\right|}{\|\mathbf{b}_r\|\,\|\hat{\mathbf{b}}_r\|} \times \frac{\left|\mathbf{c}_r^H \hat{\mathbf{c}}_r\right|}{\|\mathbf{c}_r\|\,\|\hat{\mathbf{c}}_r\|} > 0.99^3$$

for all $r \in \{1, \ldots, R\}$, where $\mathbf{a}_r$ is the recovery of the ground truth $\mathbf{a}$. Here the empirical success rate is calculated by averaging over 10 Monte Carlo trials.

Fig. 10 indicates that HMRTC has a high probability to achieve accurate factor vectors, as long as the sampling ratio is sufficiently large, though it is not theoretically guaranteed.

D. Frequency estimation

Although the focus of this work lies in reconstructing a full N-D exponential signal from partial measurements, it is still important to estimate frequency in some applications. Here, we adopt the method in [62] to estimate the frequency and root mean square error (RMSE) to quantify the accuracy of estimation. The RMSE is defined as

$$\text{RMSE} = \sqrt{\frac{1}{R}\sum_{r=1}^{R}\left(f_r - \hat{f}_r\right)^2},$$

where f_r denotes the estimated frequency of the fully sampled signal, $\hat{f}_r$ the estimated frequency of the reconstructed signal, and R the number of exponentials. Table II indicates that the proposed HMRTC obtains the best frequency estimation, comparing with ADM-TR and WCP.

E. Regularization parameter

The optimal λ in HMRTC, producing the lowest reconstruction error, generally decreases as the noise level increases. For example, as shown in Fig. 11(a), the optimal λ is 50, 20 and 10 when the noise level is 0.01, 0.02 and 0.04, respectively. This trend means that a smaller λ should be set for a higher noise level.
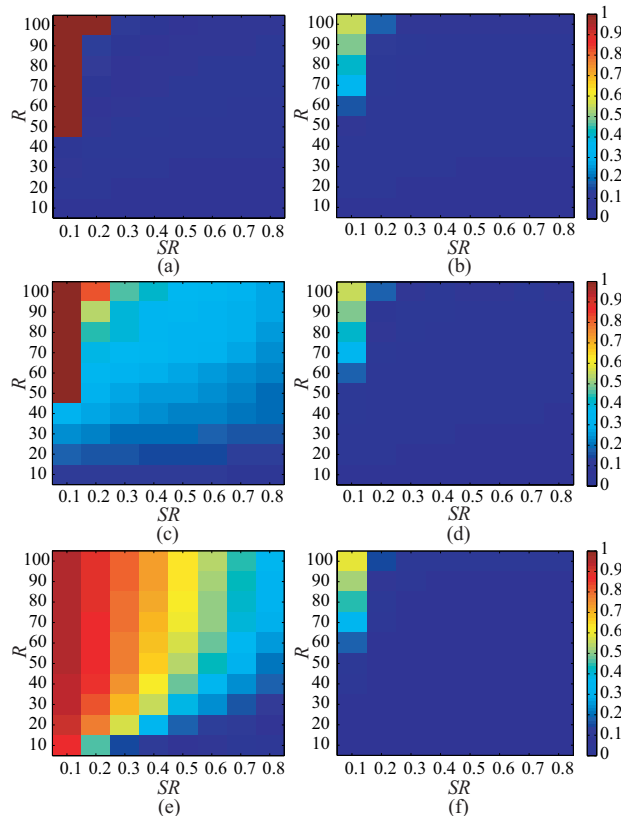

Fig. 9: Comparison between HMRTC with other state-of-the-art methods. (a) and (b) show average RLNEs by TNN [47] and HMRTC for real exponential signal recovery, respectively; (c) and (d) indicate average RLNEs by ADMM-R [48] and HMRTC for complex exponential signal recovery, respectively; (e) and (f) show average RLNEs by FaLRTC [36] and HMRTC for noiseless complex exponential signal recovery. The experiment settings are based on the consideration that TNN [47] cannot support complex number and FaLRTC [36] aims to recover noiseless tensors.

Fig. 11(a) indicates that there exists an available range ($10 \leq \lambda \leq 10^2$) for λ that leads to low reconstruction errors (RLNE ≤ 0.1), while a smaller or larger λ will bring high reconstruction errors (RLNE > 0.1). It is primarily because that a smaller λ results in missing some spectral peaks (Fig. 12(b)) while a larger λ introduces more noise into the spectrum (Fig. 12(c)). In addition, this available range turns narrowed as the noise level increases (Fig. 11(a)). Fig. 11(b) further indicates that, the available range of λ turns narrowed if the number of exponentials increases, implying that it becomes more difficult to find an optimal λ.

F. Effect of local minima

Due to the CP decomposition, HMRTC is non-convex and thus local minima may be encountered. Fig. 8(b) presents that, under different initializations, all reconstruction errors RLNEs are very small (in the order of 10^{-2}) though each reconstruction error varies. We further quantitatively compare the frequency estimation errors of reconstructions with the

TABLE II: Frequency estimation from different reconstructions.

# Peaks	Estimated frequencies from noiseless full data	Frequency errors (10^{-5})		
		ADM-TR	WCP	Proposed
1	(0.12201, 0.45297, 0.31506)	(Failed,108, 338)	(9, 1, 5)	(3,2,2)
2	(0.23503, 0.54308, 0.38204)	(Failed, 2913, 579)	(1, 5, 20)	(1,1,1)
3	(0.26608, 0.57491, 0.44104)	(Failed, 7207, 3499)	(2, 2, 1)	(2,0,1)
4	(0.31721, 0.64294, 0.47386)	(Failed, 8213 ,7507)	(6, 14, 2)	(1,0,0)
5	(0.38207, 0.72565, 0.55160)	(252, 1471, 3559)	(5, 0, 1)	(2,2,3)
6	(0.41720, 0.74228, 0.58575)	(Failed, 7016, 418)	(2, 4, 0)	(2,1,0)
7	(0.44421, 0.81311, 0.62796)	(Failed, 3299, 6558)	(0, 4, 3)	(1,1,0)
8	(0.56599, 0.84617, 0.69391)	(Failed, 5922, 4492)	(3, 0, 3)	(0,1,2)
9	(0.62214, 0.90486, 0.73689)	(235, 3533, 6835)	(6, 1,3)	(0,0,1)
10	(0.69392, 0.94303, 0.81405)	(51, 105, 388)	(2, 0, 3)	(1,0,1)

Note: The signal reconstruction error, RLNE, of ADM-TR, WCP and the proposed method are 0.9321, 0.0373 and 0.0105, respectively. The *Failed* means that too many pseudo peaks are presented thus the frequency cannot be estimated. The noisy 3-D signal with known frequencies in Fig. 3 is used for simulation. The number of peaks R is 10 and the sampling rate is 6%.
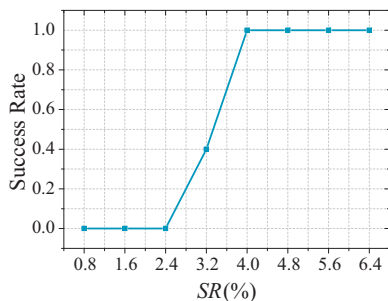

Fig. 10: The empirical success rate of the factor vector recovery versus sampling ratio. The simulated signals contains 10 exponentials. We set $\hat{R} = 10$ and $\rho = 1.02$, and measurements are randomly sampled.
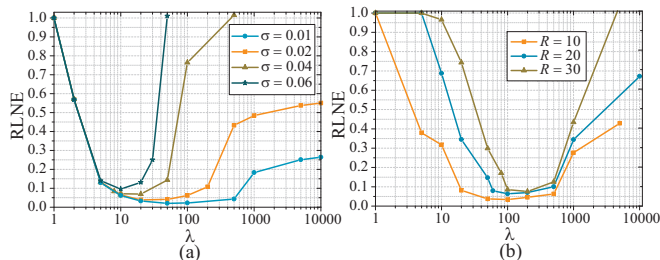

Fig. 11: Reconstruction errors versus regularization parameter λ. (a) denotes RLNE via λ under different noise level σ, (b) denotes RLNE via λ under different R. Note: The source 3-D data in (a) is Fig. 3 and in (b) is generated with noise level 0.01. The sampling ratio in (a) and (b) is 0.3.

Fig. 12: A 2-D plane of the reconstructed 3-D spectrum with different regularization parameters λ. (a) is the noiseless ground-truth spectrum, (b)-(d) are reconstructed spectra with $\lambda = 2$, 1000 and 50, respectively. Note: The 3-D data in Fig. 3 is used for simulation. The sampling ratio $SR = 0.3$, the number of exponentials $R = 10$ and the noise level $\sigma = 0.02$.

maximum and minimum RLNE in Fig. 8(b). Table III shows that each frequency estimation error is in the order of 10^{-5}, implying that the initializations has no significant impact on the reconstruction.

VII. CONCLUSIONS AND FUTURE WORK

A Hankel matrix nuclear norm regularized CANDECOMP/PARAFAC(CP)-based tensor completion method is proposed to recover missing data of high-dimensional exponential signals. The full signal is reconstructed by simultaneously exploring the low-CP-rank 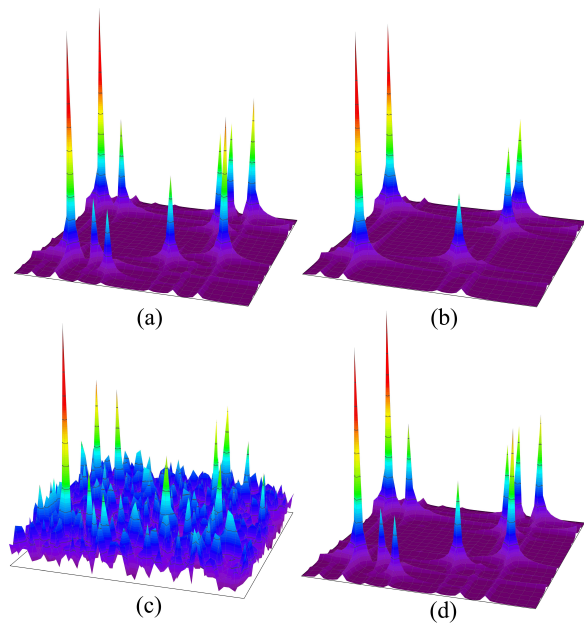structure of the signal and the exponential structure of the associated factor vectors. Numerical experiments on simulated undamped and damped complex sinusoids, as well as the real NMR protein spectroscopy data, demonstrate that the proposed method can successfully recover full signals from very limited samples and is robust to the estimated tensor rank. The proposed method is particularly useful for fast sampling of high-dimensional spectroscopy in chemistry, biology and imaging sciences [13]–[15], especially those applications where the number of spectrum peaks is hard to estimate.

The proposed algorithm may face challenge in computation when the number of dimensions and ambient dimensions are large and thus developing a more efficient numerical algorithm to accommodate larger datasets will be considered

TABLE III: Comparisons on frequency estimation error of reconstructions using different initializations.

# Peaks	Frequencies estimation from noiseless full data	Frequency estimation errors (10^{-5})	
		Maximum RLNE	Minimum RLNE
1	(0.69999, 0.29999, 0.30005)	(1, 0, 2)	(1, 13, 2)
2	(0.19994, 0.40000, 0.70007)	(2, 2, 2)	(4, 1, 0)
3	(0.29997, 0.7999, 0.59999)	(2, 7, 2)	(1, 9, 3)
4	(0.80004, 0.80007, 0.79944)	(1, 2, 1)	(1, 3, 1)
5	(0.09992, 0.90001, 0.20001)	(1, 3, 0)	(1, 4, 1)
6	(0.20178, 0.69992, 0.09993)	(3, 2, 2)	(11, 2, 3)
7	(0.19992, 0.30194, 0.69963)	(0, 4, 1)	(1, 26, 3)
8	(0.30004, 0.70162, 0.59979)	(0, 0, 0)	(1, 2, 1)
9	(0.89893, 0.20247, 0.29986)	(7, 0, 1)	(6, 0,0)
10	(0.29743, 0.10002, 0.39997)	(4, 5, 0)	(4, 4, 0)

Note: The frequency estimation are conducted on reconstructions with the maximum and minimum RLNE in Fig. 8(b).

in the further work. For example, it is meaningful to consider more computationally appealing characterization of the nuclear norm [65]–[67] to accelerate the algorithm. In addition, it is worthy of deriving the recovery conditions of this method, including the possible frequency separation conditions. Besides, how to extend these Hankel matrix reconstruction methods into fast biomedical magnetic resonance imaging [13]–[15] will be further developed. The code of this paper was shared at http://www.quxiaobo.org/project/LowRank_Hankel_NMR/Toolbox_Tensor_Hankel_NMR.zip.

VIII. ACKNOWLEDGMENTS

APPENDIX

A. Derivation of closed form solution of $\mathbf{U}^{(\mathbf{n})}$

1) Operator Definition

Before deriving the closed-form of $\mathbf{U}^{(n)}$, we define two linear operators and their adjoint operators. We rewrite the definition of Hankel operator $\mathcal{R}$ as follows

$$\mathcal{R}: \mathbf{x} \in \mathbb{C}^{s_1+s_2-1} \mapsto \mathcal{R}\mathbf{x} \in \mathbb{C}^{s_1 \times s_2}, [\mathcal{R}\mathbf{x}]_{i,j} = x_{i+j},$$
$$\forall i \in \{0,\ldots,s_1-1\}, j \in \{0,\ldots,s_2-1\}, \quad (31)$$

Meanwhile, the adjoint $\mathcal{R}^*$ of $\mathcal{R}$ is given by

$$\mathcal{R}^*: \mathbf{X} \in \mathbb{C}^{s_1 \times s_2} \mapsto \mathcal{R}^*\mathbf{X} \in \mathbb{C}^{s_1+s_2-1},$$
$$[\mathcal{R}^*\mathbf{X}]_k = \sum_{i+j=k} x_{i,j}, \quad (32)$$
$$\forall i \in \{0,\ldots,s_1-1\}, j \in \{0,\ldots,s_2-1\},$$

where $[\cdot]_k$ and $[\cdot]_{i,j}$ denote the k-th entry of a vector and the (i,j)-element of a matrix, respectively.

We denote $\mathcal{D} = \mathcal{R}^*\mathcal{R}$ and it is a diagonal operator from a vector to a vector of the form. In other words, for any $\mathbf{x} \in \mathbb{C}^{s_1+s_2-1}$, we have

$$\mathcal{D}\mathbf{x} = \mathbf{w} \circledast \mathbf{x}, \quad (33)$$

where $\mathbf{w}$ is a vector and its k-th element w_k is the number of elements in k-th anti-diagonal of a matrix with size $s_1 \times s_2$; $\circledast$ denotes the Hadamard product.

We also define another linear operator $\mathcal{Q}_r$ which aims to extract the rth column from $\mathbf{X}$. For a matrix $\mathbf{X} \in \mathbb{C}^{n_1 \times n_2}$, specifically, we define $\mathcal{Q}_r$ by

$$\mathcal{Q}_r: \mathbf{X} \in \mathbb{C}^{n_1 \times n_2} \mapsto \mathcal{Q}_r\mathbf{X} = \mathbf{X}_{(:,r)} \in \mathbb{C}^{n_1 \times 1},$$
$$\forall r \in \{0,\ldots,n_2-1\}. \quad (34)$$

Then the adjoint $\mathcal{Q}_r^*$ of $\mathcal{Q}_r$ is given by

$$\mathcal{Q}_r^*: \mathbf{x} \in \mathbb{C}^{n_1 \times 1} \mapsto \mathcal{Q}_r^*\mathbf{x} \in \mathbb{C}^{n_1 \times n_2},$$
$$[\mathcal{Q}_r^*\mathbf{x}]_{(:,k)} = \begin{cases} \mathbf{x} & k = r, \\ 0 & k \neq r. \end{cases} \quad \forall r \in \{0,\ldots,n_2-1\}. \quad (35)$$

where $[\cdot]_{(:,k)}$ denotes the k-th column of a matrix, $k = 0,\ldots,s_2-1$. Thus we have

$$[\mathcal{Q}_r^*\mathcal{Q}_r\mathbf{X}]_{(:,k)} = \begin{cases} \mathbf{X}_{(:,r)} & k = r, \\ 0 & k \neq r. \end{cases} \quad (36)$$

According to (33) and (36), for any matrix $\mathbf{X} \in \mathbb{C}^{n_1 \times n_2}$, we have

$$\sum_{r=1}^{n_2} \mathcal{Q}_r\mathcal{R}\mathcal{R}\mathcal{Q}_r\mathbf{X} = \mathbf{C} \circledast \mathbf{X}, \quad (37)$$

where $\mathbf{C} \in \mathbb{C}^{n_1 \times n_2}$ and each column equals to $\mathbf{w}$ defined in (33). Note that here $n_1 = s_1 + s_2 - 1$, i.e., Hankel operator $\mathcal{R}$ maps a vector with size n_1 to a Hankel matrix with size $s_1 \times s_2$.

2) Derivation of closed form solution of $\mathbf{U}^{(\mathbf{n})}$

For simplicity, we replace the terms in the right of (23) with a constant matrix $\mathbf{E} = \lambda\mathcal{A}_k^{(n)*}\left(\mathcal{P}_{\Omega^{(n)}}\left(\mathbf{Y}_{(n)}\right)\right) + \beta_k \sum_{r=1}^{\hat{R}} \mathcal{Q}_r^*\mathcal{R}^*\left(\mathbf{Z}_{r;k}^{(n)} - (\beta_k)^{-1}\mathbf{D}_{r;k}^{(n)}\right)$, and by (37) one has

$$\lambda\mathcal{A}_k^{(n)*}\mathcal{A}_k^{(n)}\mathbf{U}^{(n)} + \beta_k\mathbf{C} \circledast \mathbf{U}^{(n)} = \mathbf{E}. \quad (38)$$

Since $\mathcal{A}_k^{(n)}$ can be rewritten as

$$\mathcal{A}_k^{(n)}\left(\mathbf{U}^{(n)}\right) = \mathcal{P}_{\Omega^{(n)}}\left(\mathbf{U}^{(n)}\mathbf{G}_k^{(n)}\right)$$
$$= \begin{bmatrix} \mathcal{P}_{\Omega_1^{(n)}}\left(\mathbf{v}_1^{(n)}\mathbf{G}_k^{(n)}\right) \\ \vdots \\ \mathcal{P}_{\Omega_{I_n}^{(n)}}\left(\mathbf{v}_{I_n}^{(n)}\mathbf{G}_k^{(n)}\right) \end{bmatrix} = \begin{bmatrix} \mathcal{A}_{1;k}^{(n)}\left(\mathbf{v}_1^{(n)}\right) \\ \vdots \\ \mathcal{A}_{I_n;k}^{(n)}\left(\mathbf{v}_{I_n}^{(n)}\right) \end{bmatrix}, \quad (39)$$

where $\mathbf{v}_i^{(n)} \in \mathbb{C}^{1 \times \hat{R}}$ denotes the i-th row of $\mathbf{U}^{(n)}$ and $\Omega_i^{(n)}$ is the subset of $\Omega^{(n)}$ corresponding the i-th row. Here $\mathcal{A}_{i;k}^{(n)}(\mathbf{x}) \triangleq \mathcal{P}_{\Omega_i^{(n)}}\left(\mathbf{x}\mathbf{G}_k^{(n)}\right)$. Therefore, to obtain the closed-form of $\mathbf{U}^{(n)}$, we can first derive the closed-form of each row of $\mathbf{U}^{(n)}$, i.e., $\mathbf{v}_i^{(n)}$. Then (38) can be written as

$$\lambda\mathcal{A}_{i;k}^{(n)*}\mathcal{A}_{i;k}^{(n)}(\mathbf{v}_i^{(n)}) + \beta\mathbf{c}_i \circledast \mathbf{v}_i^{(n)} = \mathbf{e}_i$$
$$\text{for } i \in \{1,\cdots,I_n\}, \quad (40)$$

where $\mathbf{e}_i$ and $\mathbf{c}_i$ denote the i-th row of $\mathbf{E}$ and $\mathbf{C}$, respectively.

To get the closed form of $\mathbf{v}_i^{(n)}$, we simplify the operator $\mathcal{A}_{i;k}^{(n)}$. According to the definition $\mathcal{A}_{i;k}^{(n)}(\mathbf{x}) = \mathcal{P}_{\Omega_i^{(n)}}\left(\mathbf{x}\mathbf{G}_k^{(n)}\right)$ and $\mathbf{x}\mathbf{G}_k^{(n)} \in \mathbb{C}^{1 \times K_n}$ with $K_n = \prod_{k=1,k\neq n}^{N} I_k$, $\mathcal{A}_{i;k}^{(n)}$ can be

rewritten as $\mathcal{A}_{i;k}^{(n)}(\mathbf{x}) = \mathbf{x}\mathbf{G}_k^{(n)}\mathbf{P}_i$, where $\mathbf{P}_i \in \mathbb{C}^{K_n \times K_n}$ is a diagonal matrix and the (k,k)-th entry in the main diagonal is

$$[\mathbf{P}_i]_{k,k} = \begin{cases} 1 & \text{if } (k) \in \Omega_i^{(n)}, \\ 0 & \text{otherwise.} \end{cases} \quad (41)$$

for all $k \in \{1, \ldots, K_n\}$, where $[\cdot]_{k,k}$ denotes the (k,k)-th entry of a matrix. Hence (40) can be written as

$$\lambda \mathbf{v}_i^{(n)} \mathbf{G}_k^{(n)} \mathbf{P}_i \mathbf{P}_i^* \mathbf{G}_k^{(n)*} + \beta_k \mathbf{v}_i^{(n)} D(\mathbf{c}_i) = \mathbf{e}_i, \quad (42)$$

where $D(\mathbf{c}_i)$ denotes a diagonal matrix and the (k,k)-th entry in $D(\mathbf{c}_i)$ equals to the k-th entry in $\mathbf{c}_i$.

Therefore the closed form of $\mathbf{v}_i^{(n)}$ is

$$\mathbf{v}_i^{(n)} = \mathbf{e}_i \left(\lambda \mathbf{G}_k^{(n)} \mathbf{P}_i \mathbf{P}_i^* \mathbf{G}_k^{(n)*} + \beta_k D(\mathbf{c}_i) \right)^{-1}. \quad (43)$$

Then the closed form solution of $\mathbf{U}^{(\mathbf{n})}$ can be obtained by $\mathbf{U}^{(n)} = [\mathbf{v}_1^{(n)T} \ldots \mathbf{v}_{I_n}^{(n)T}]^T$

B. Lemmas

To prove **Theorem 1** and **2**, we first prove the boundedness of multipliers and some variables generated by Algorithm 1, and then we analyze the convergence of the algorithm.

Lemma 1 *[54] Let $\mathcal{T}$ be Hilbert space endowed with an inner product $\langle \cdot, \cdot \rangle$ and a corresponding norm $\|\cdot\|$, and $\mathbf{y} \in \partial \|\mathbf{x}\|$, where $\partial f(\mathbf{x})$ is the subgradient of a convex $f(\mathbf{x})$. Then $\|\mathbf{y}\|^* = 1$ if $\mathbf{x} \neq \mathbf{0}$, and $\|\mathbf{y}\|^* \leq 1$ if $\mathbf{x} = \mathbf{0}$, where $\|\cdot\|^*$ is the dual norm of $\|\cdot\|$.*

Lemma 2 *The sequences $\{\mathbf{D}_{r;k}^{(n)}\}$, $r = 1, \ldots, \hat{R}$ and $n = 1, \ldots, N$, are bounded.*

Proof: By the optimality of $\mathbf{Z}_{r;k+1}^{(n)}$, we have that:

$$\mathbf{0} \in \partial \|\mathbf{Z}_{r;k+1}^{(n)}\|_* - \mathbf{D}_{r;k}^{(n)} - \beta_k \left(\mathcal{R}\mathcal{Q}_r \mathbf{U}_k^{(n)} - \mathbf{Z}_{r;k+1}^{(n)} \right).$$

This together with

$$\mathbf{D}_{r;k+1}^{(n)} = \mathbf{D}_{r;k}^{(n)} + \beta_k \left(\mathcal{R}\mathcal{Q}_r \mathbf{U}_k^{(n)} - \mathbf{Z}_{r;k+1}^{(n)} \right)$$

implies

$$\mathbf{D}_{r;k+1}^{(n)} \in \partial \|\mathbf{Z}_{r;k+1}^{(n)}\|_*.$$

From **lemma 1** and [49], every element of the subgradient of the nuclear norm is bounded by 1 in spectral norm. Therefore, $\|\mathbf{D}_{r;k}^{(n)}\|_2 \leq 1$ and hence the sequence $\{\mathbf{D}_{r;k}^{(n)}\}$ is bounded for all r and n.

Lemma 3 *The sequences $\{\mathbf{U}_k^{(n)}\}$ and $\{\mathbf{Z}_{r;k}^{(n)}\}$, $r = 1, \ldots, \hat{R}$, $n = 1, \ldots, N$, produced by algorithm 1 are bounded.*

Proof: From the iteration procedure, we have that

$$\mathcal{L}_{\beta_k}(\mathcal{U}_{k+1}, \mathcal{Z}_{k+1}, \mathcal{D}_k) \leq \mathcal{L}_{\beta_k}(\mathcal{U}_k, \mathcal{Z}_k, \mathcal{D}_k)$$

$$= \sum_{n=1}^{N} \sum_{r=1}^{\hat{R}} \left(\langle \mathbf{D}_{r;k}^{(n)} - \mathbf{D}_{r;k-1}^{(n)}, \mathcal{R}\mathcal{Q}_r \mathbf{U}_k^{(n)} - \mathbf{Z}_{r;k}^{(n)} \rangle + \right.$$

$$\left. \frac{\beta_k - \beta_{k-1}}{2} \|\mathcal{R}\mathcal{Q}_r \mathbf{U}_k^{(n)} - \mathbf{Z}_{r;k}^{(n)}\|_F^2 \right) + \mathcal{L}_{\beta_{k-1}}(\mathcal{U}_k, \mathcal{Z}_k, \mathcal{D}_{k-1})$$

$$= \sum_{n=1}^{N} \sum_{r=1}^{\hat{R}} \left(\frac{1}{\beta_{k-1}} \|\mathbf{D}_{r;k}^{(n)} - \mathbf{D}_{r;k-1}^{(n)}\|_F^2 \right.$$

$$\left. + \frac{\beta_k - \beta_{k-1}}{2\beta_{k-1}^2} \|\mathbf{D}_{r;k}^{(n)} - \mathbf{D}_{r;k-1}^{(n)}\|_F^2 \right) + \mathcal{L}_{\beta_{k-1}}(\mathcal{U}_k, \mathcal{Z}_k, \mathcal{D}_{k-1})$$

$$= \frac{\beta_k + \beta_{k-1}}{2\beta_{k-1}^2} \sum_{n=1}^{N} \sum_{r=1}^{\hat{R}} \|\mathbf{D}_{r;k}^{(n)} - \mathbf{D}_{r;k-1}^{(n)}\|_F^2$$

$$+ \mathcal{L}_{\beta_{k-1}}(\mathcal{U}_k, \mathcal{Z}_k, \mathcal{D}_{k-1}).$$

Therefore $\mathcal{L}_{\beta_{k-1}}(\mathcal{U}_k, \mathcal{Z}_k, \mathcal{D}_{k-1})$ is upper bounded thanks to the boundedness of $\{\mathbf{D}_{r;k}^{(n)}\}$ and

$$\sum_{k=1}^{\infty} \frac{\beta_k + \beta_{k-1}}{2(\beta_{k-1})^2} = \frac{\rho(\rho+1)}{2\beta_0(\rho-1)} < \infty,$$

where $\rho \in (1.0, 1.1]$. Then we have

$$\sum_{r=1}^{\hat{R}} \sum_{n=1}^{N} \|\mathbf{Z}_{r;k}^{(n)}\|_*$$

$$+ \frac{\lambda}{2} \|\mathcal{P}_\Omega(\mathcal{Y}) - \mathcal{P}_\Omega([\![\mathbf{U}_k^{(1)}, \mathbf{U}_k^{(2)}, \ldots, \mathbf{U}_k^{(N)}]\!])\|_F^2$$

$$= \mathcal{L}_{\beta_{k-1}}(\mathcal{U}_k, \mathcal{Z}_k, \mathcal{D}_{k-1}) - \frac{1}{2\beta_{k-1}} \sum_{n=1}^{N} \sum_{r=1}^{\hat{R}} \left(\|\mathbf{D}_{r;k}^{(n)} - \mathbf{D}_{r;k-1}^{(n)}\|_F^2 \right.$$

$$\left. + 2\langle \mathbf{D}_{r;k-1}^{(n)}, \mathbf{D}_{r;k}^{(n)} - \mathbf{D}_{r;k-1}^{(n)} \rangle \right),$$

is upper bounded, which implies $\{\mathbf{Z}_{r;k}^{(n)}\}$ is bounded for all $r \in \{1, \ldots, \hat{R}\}$ and $n \in \{1, \ldots, N\}$. Furthermore, $\{\mathbf{U}_k^{(n)}\}$ is bounded for $n = 1, \ldots, N$, due to the equation $\mathcal{R}\mathcal{Q}_r \mathbf{U}_{k+1}^{(n)} = \mathbf{Z}_{r;k+1}^{(n)} + (\mathbf{D}_{r;k+1}^{(n)} - \mathbf{D}_{r;k}^{(n)})/\beta_k$ for $r = 1, \ldots, \hat{R}$.

C. Proof of convergence

Proof of Theorem 1: By $\mathbf{D}_{r;k+1}^{(n)} = \mathbf{D}_{r;k}^{(n)} + \beta_k(\mathcal{R}\mathcal{Q}_r \mathbf{U}_{k+1}^{(n)} - \mathbf{Z}_{r;k+1}^{(n)})$, the boundedness of $\{\mathbf{D}_{r;k}^{(n)}\}$ and $\lim_{k\to\infty} \beta_k = \infty$, we have

$$\lim_{k\to\infty} \left\| \mathcal{R}\mathcal{Q}_r \mathbf{U}_{k+1}^{(n)} - \mathbf{Z}_{r;k+1}^{(n)} \right\|_F = 0, \quad (44)$$

for $r = 1, \ldots, \hat{R}$ and $n = 1, \ldots, N$. Therefore, $\{\mathcal{Z}_k, \mathcal{U}_k\}$ approaches to a feasible solution.

Moreover, $\mathbf{D}_{r;k}^{(n)} = \mathbf{D}_{r;k-1}^{(n)} + \beta_{k-1}(\mathcal{R}\mathcal{Q}_r\mathbf{U}_k^{(n)} - \mathbf{Z}_{r;k}^{(n)})$ and (23) imply

$$
\begin{aligned}
&\beta_k \sum_{r=1}^{\hat{R}} \mathcal{Q}_r^*\mathcal{R}^*\mathcal{R}\mathcal{Q}_r\big(\mathbf{U}_{k+1}^{(n)} - \mathbf{U}_k^{(n)}\big)\\
&=\lambda\mathcal{A}_k^{(n)*}\big(\mathcal{P}_{\Omega^{(n)}}(\mathbf{Y}_{(n)})\big) - \lambda\mathcal{A}_k^{(n)*}\mathcal{A}_k^{(n)}\mathbf{U}_{k+1}^{(n)}\\
&\quad + \beta_k \sum_{r=1}^{\hat{R}} \mathcal{Q}_r^*\mathcal{R}^*\big(\mathbf{Z}_{r;k}^{(n)} - (\beta_k)^{-1}\mathbf{D}_{r;k}^{(n)} - \mathcal{R}\mathcal{Q}_r\mathbf{U}_k^{(n)}\big)\\
&=\lambda\mathcal{A}_k^{(n)*}\big(\mathcal{P}_{\Omega^{(n)}}(\mathbf{Y}_{(n)})\big) - \lambda\mathcal{A}_k^{(n)*}\mathcal{A}_k^{(n)}\mathbf{U}_{k+1}^{(n)}\\
&\quad - \sum_{r=1}^{\hat{R}} \mathcal{Q}_r^*\mathcal{R}^*\big((\rho+1)\mathbf{D}_{r;k}^{(n)} - \rho\mathbf{D}_{r;k-1}^{(n)}\big).
\end{aligned}
\tag{45}
$$

According to (37), we have

$$
\begin{aligned}
&\mathbf{C} \circledast \big(\mathbf{U}_{k+1}^{(n)} - \mathbf{U}_k^{(n)}\big)\\
&=\frac{1}{\beta_k}\big(\lambda\mathcal{A}_k^{(n)*}\big(\mathcal{P}_{\Omega^{(n)}}(\mathbf{Y}_{(n)})\big) - \lambda\mathcal{A}_k^{(n)*}\mathcal{A}_k^{(n)}\mathbf{U}_{k+1}^{(n)}\\
&\quad - \sum_{r=1}^{\hat{R}} \mathcal{Q}_r^*\mathcal{R}^*\big((\rho+1)\mathbf{D}_{r;k}^{(n)} - \rho\mathbf{D}_{r;k-1}^{(n)}\big)\big) = \frac{\mathbf{L}_k}{\beta_k},
\end{aligned}
\tag{46}
$$

where $\mathbf{C}$ is a constant matrix and $\circledast$ denotes Hadamard product. Since $\mathbf{L}_k$ is upper bounded, $\big\|\mathbf{U}_{k+1}^{(n)} - \mathbf{U}_k^{(n)}\big\|_F = O\big((\beta_k)^{-1}\big)$. We further have

$$
\begin{aligned}
&\big\|\mathbf{C} \circledast \big(\mathbf{U}_m^{(n)} - \mathbf{U}_k^{(n)}\big)\big\|_F \leq \big\|\mathbf{C} \circledast \big(\mathbf{U}_m^{(n)} - \mathbf{U}_{m-1}^{(n)}\big)\big\|_F\\
&\quad + \big\|\mathbf{C} \circledast \big(\mathbf{U}_{m-1}^{(n)} - \mathbf{U}_{m-2}^{(n)}\big)\big\|_F + \cdots +\\
&\quad \big\|\mathbf{C} \circledast \big(\mathbf{U}_{k+1}^{(n)} - \mathbf{U}_k^{(n)}\big)\big\|_F\\
&=\frac{\|\mathbf{L}_{m-1}\|_F}{\beta_{m-1}} + \frac{\|\mathbf{L}_{m-2}\|_F}{\beta_{m-2}} + \cdots + \frac{\|\mathbf{L}_k\|_F}{\beta_k} \leq\\
&\frac{\delta}{\beta_k}\Big(\frac{1}{\rho_{m-k-1}} + \frac{1}{\rho_{m-k-2}} + \cdots + 1\Big) < \frac{\delta\rho}{\beta_k(1-\rho)},
\end{aligned}
\tag{47}
$$

where $\delta = \max\{\|\mathbf{L}_{m-1}\|_F, \|\mathbf{L}_{m-2}\|_F, \ldots, \|\mathbf{L}_k\|_F\}$. Since $\frac{\delta\rho}{\beta_k(1-\rho)} \to 0$, $\{\mathbf{U}_k^{(n)}\}$ is a Cauchy sequence for all $n \in \{1, \ldots, N\}$, i.e., $\{\mathcal{U}_k\}$ is a Cauchy sequence.

Similarly, $\{\mathbf{Z}_{r;k}^{(n)}\}$, $r = 1, \ldots, \hat{R}$, $n = 1, \ldots, N$, also is a Cauchy sequence.

Proof of Theorem 2: Let $(\tilde{\mathbf{U}}_k^{(1)} \cdots, \tilde{\mathbf{U}}_k^{(N)})$ be a stationary point of (17), then the KKT conditions of (17) are

$$
\begin{aligned}
\mathbf{0} \in &\sum_{r=1}^{\hat{R}} \mathcal{Q}_r^*\mathcal{R}^*\partial\big\|\mathcal{R}\mathcal{Q}_r\tilde{\mathbf{U}}^{(n)}\big\|_*\\
&+ \lambda\tilde{\mathcal{A}}^{(n)*}\tilde{\mathcal{A}}^{(n)}\tilde{\mathbf{U}}^{(n)} - \lambda\tilde{\mathcal{A}}^{(n)*}\big(\mathcal{P}_{\Omega^{(n)}}(\mathbf{Y}_{(n)})\big),
\end{aligned}
\tag{48}
$$

for all $n \in \{1, \ldots, N\}$, where $\tilde{\mathcal{A}}^{(n)}(\mathbf{X}) \triangleq \mathcal{P}_{\Omega^{(n)}}\big(\mathbf{X}\tilde{\mathbf{G}}^{(n)}\big)$ and $\tilde{\mathbf{G}}^{(n)} = (\tilde{\mathbf{U}}^{(N)} \odot \cdots \odot \tilde{\mathbf{U}}^{(n+1)} \odot \tilde{\mathbf{U}}^{(n-1)} \odot \cdots \odot \tilde{\mathbf{U}}^{(1)})^T$.

According to Algorithm 1, the first-order optimal condition of (21) at the $(k+1)$-th iteration is

$$
\begin{aligned}
\mathbf{0} = &\lambda\mathcal{A}_k^{(n)*}\mathcal{A}_k^{(n)}\mathbf{U}_{k+1}^{(n)} - \lambda\mathcal{A}_k^{(n)*}\big(\mathcal{P}_{\Omega^{(n)}}(\mathbf{Y}_{(n)})\big)\\
&+ \beta_k \sum_{r=1}^{\hat{R}} \mathcal{Q}_r^*\mathcal{R}^*(\mathcal{R}\mathcal{Q}_r\mathbf{U}_{k+1}^{(\mathbf{n})} - \mathbf{Z}_{r;k}^{(n)} + (\beta_k)^{-1}\mathbf{D}_{r;k}^{(n)}).
\end{aligned}
\tag{49}
$$

The first-order optimal condition of the problem (25) is

$$
\mathbf{0} \in \partial\big\|\mathbf{Z}_{r;k+1}^{(n)}\big\|_* + \beta_k(\mathbf{Z}_{r;k+1}^{(n)} - \mathcal{R}\mathcal{Q}_r\mathbf{U}_{k+1}^{(n)}) - \mathbf{D}_{r;k}^{(n)},
$$

implying that

$$
\mathbf{0} \in \partial\big\|\mathbf{Z}_{r;k+1}^{(n)}\big\|_* - \mathbf{D}_{r;k+1}^{(n)}.
\tag{50}
$$

Since $\{\mathbf{U}_k^{(n)}\}$ and $\{\mathbf{Z}_{r;k}^{(n)}\}$ are all Cauchy sequences, $\{\mathbf{U}_\infty^{(n)}\}$ and $\{\mathbf{Z}_{r;\infty}^{(n)}\}$ are their limit points, respectively. We also have $\mathcal{R}\mathcal{Q}_r\mathbf{U}_\infty^{(n)} = \mathbf{Z}_{r;\infty}^{(n)}$ according to (44). By (49) and (50), if $\lim_{k\to\infty}\big\|\mathbf{D}_{r;k+1}^{(n)} - \mathbf{D}_{r;k}^{(n)}\big\|_F = 0$ for all $r = 1, \ldots, \hat{R}$ and $n = 1, \ldots, N$, we have

$$
\begin{aligned}
\mathbf{0} \in &\sum_{r=1}^{\hat{R}} \mathcal{Q}_r^*\mathcal{R}^*\partial\big\|\mathcal{R}\mathcal{Q}_r\mathbf{U}_\infty^{(n)}\big\|_*\\
&+ \lambda\mathcal{A}_\infty^{(n)*}\mathcal{A}_\infty^{(n)}\mathbf{U}_\infty^{(n)} - \lambda\mathcal{A}_\infty^{(n)*}\big(\mathcal{P}_{\Omega^{(n)}}(\mathbf{Y}_{(n)})\big).
\end{aligned}
\tag{51}
$$

where $\mathcal{A}_\infty^{(n)}(\mathbf{X}) \triangleq \mathcal{P}_{\Omega^{(n)}}\big(\mathbf{X}\mathbf{G}_\infty^{(n)}\big)$ and $\mathbf{G}_\infty^{(n)} = (\mathbf{U}_\infty^{(N)} \odot \cdots \odot \mathbf{U}_\infty^{(n+1)} \odot \mathbf{U}_\infty^{(n-1)} \odot \cdots \odot \mathbf{U}_\infty^{(1)})^T$

Therefore the limit of the sequence $\{\mathcal{U}_k\}$ satisfies the KKT conditions of the problem (17).

REFERENCES

[1] X. Qu, M. Mayzel, J.-F. Cai, Z. Chen, and V. Orekhov, "Accelerated NMR spectroscopy with low-rank reconstruction," *Angew. Chem. Int. Ed.*, vol. 54, pp. 852-854, 2015.

[2] X. Qu, X. Cao, D. Guo, and Z. Chen. "Compressed sensing for sparse magnetic resonance spectroscopy," *Int. Society for Magn. Reson. in Med. 18th Sci. Meeting*, Stockholm, Sweden, pp. 3371, 2010.

[3] X. Qu, D. Guo, X. Cao, S. Cai, and Z. Chen, "Reconstruction of self-sparse 2D NMR spectra from undersampled data in the indirect dimension," *Sensors*, vol. 11, pp. 8888-8909, 2011.

[4] E. J. Candès and C. Fernandez-Granda, "Towards a mathematical theory of super-resolution," *Commun. Pure Appl. Math.*, vol. 67, pp. 906-956, 2014.

[5] G. Tang, B. N. Bhaskar, P. Shah, and B. Recht, "Compressed sensing off the grid," *IEEE Trans. Inf. Theory*, vol. 59, pp. 7465-7490, 2013.

[6] D. Nion and N. D. Sidiropoulos, "Tensor algebra and multidimensional harmonic retrieval in signal processing for MIMO radar," *IEEE Trans. Signal Process.*, vol. 58, pp. 5693-5705, 2010.

[7] F. Wen and H. C. So, "Robust multi-dimensional harmonic retrieval using iteratively reweighted HOSVD," *IEEE Signal Process. Lett.*, vol. 22, pp. 2464-2468, 2015.

[8] M. Haardt and J. A. Nossek, "Simultaneous Schur decomposition of several nonsymmetric matrices to achieve automatic pairing in multidimensional harmonic retrieval problems," *IEEE Trans. Signal Process.*, vol. 46, pp. 161-169, 1998.

[9] Y. Chi and Y. Chen, "Compressive two-dimensional harmonic retrieval via atomic norm minimization," *IEEE Trans. Signal Process.*, vol. 63, pp. 1030-1042, 2015.

[10] J. A. Tropp, J. N. Laska, M. F. Duarte, J. K. Romberg, and R. G. Baraniuk, "Beyond nyquist: Efficient sampling of sparse bandlimited signals," *IEEE Trans. Inf. Theory*, vol. 56, pp. 520-544, 2010.

[11] M. Lustig, D. Donoho, and J. M. Pauly, "Sparse MRI: The application of compressed sensing for rapid MR imaging," *Magn. Reson. Med*, vol. 58, pp. 1182-1195, 2007.

[12] F. Lam and Z.-P. Liang, "A subspace approach to high-resolution spectroscopic imaging" *Magn. Reson. Med*, vol. 71, pp. 1349-1357, 2014.

[13] X. Qu, D. Guo, B. Ning, Y. Hou, Y. Lin, S. Cai, Z. Chen, "Undersampled MRI reconstruction with patch-based directional wavelets," *Magn. Reson. Imag.*, vol. 30, pp. 964977, 2012

[14] X. Qu, Y. Hou, F. Lam, D. Guo, J. Zhong, and Z. Chen, "Magnetic resonance image reconstruction from undersampled measurements using a patch-based nonlocal operator," *Med. Image Anal.*, vol. 18, pp. 843856, 2014.

[15] Y. Liu, Z. Zhan, J.-F. Cai, D. Guo, Z. Chen, X. Qu, "Projected iterative soft-thresholding algorithm for tight frames in compressed sensing magnetic resonance imaging," *IEEE Trans. Med. Imag.*, vol. 35, 2130-2140, 2016.

[16] Y. Li, J. Razavilar, and K. J. R. Liu, "A high-resolution technique for multidimensional NMR spectroscopy," *IEEE Trans. Biomed. Eng.*, vol. 45, pp. 78-86, 1998.

[17] W. Xu, J.-F. Cai, K. V. Mishra, C. Myung, and A. Kruger, "Precise semidefinite programming formulation of atomic norm minimization for recovering d-dimensional ($d \geq 2$) off-the-grid frequencies," in *Proc. IEEE Inf. Theory Appl. Workshop (ITA)*, 2014, pp. 1-4.

[18] K. V. Mishra, M. Cho, A. Kruger, and W. Xu, "Spectral super-resolution with prior knowledge," *IEEE Trans. Signal Process.*, vol. 63, pp. 5342-5357, 2015.

[19] M. Cho, K. V. Mishra, J. F. Cai, and W. Xu, "Block iterative reweighted algorithms for super-resolution of spectrally sparse signals," *IEEE Signal Process. Lett.*, vol. 22, pp. 2319-2313, 2015.

[20] Z. Tan, Y. C. Eldar, and A. Nehorai, "Direction of arrival estimation using co-prime arrays: A super resolution viewpoint," *IEEE Trans. Signal Process.*, vol. 62, pp. 5565-5576, 2014.

[21] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, pp. 489-509, 2006.

[22] Y. Chi, L. L. Scharf, A. Pezeshki, and A. R. Calderbank, "Sensitivity to basis mismatch in compressed sensing," *IEEE Trans. Signal Process.*, vol. 59, pp. 2182-2195, 2011.

[23] V. Chandrasekaran, B. Recht, P. Parrilo, and A. Willsky, "The convex geometry of linear inverse problems," *Found. Comut. Math.*, vol. 12, pp. 805-849, 2012.

[24] Y. Hua and T. K. Sarkar, "Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise," *IEEE Trans. Acoust., Speech and Signal Process.*, vol. 38, pp. 814-824, 1990.

[25] E. J. Candès and B. Recht, "Exact matrix completion via convex optimization," *Found. Comut. Math.*, vol. 9, pp. 717-772, 2009.

[26] R. H. Keshavan, A. Montanari, and O. Sewoong, "Matrix completion from a few entries," *IEEE Trans. Inf. Theory*, vol. 56, pp. 2980-2998, 2010.

[27] J.-F. Cai, X. Qu, W. Xu, and G.-B. Ye, "Robust recovery of complex exponential signals from random Gaussian projections via low rank Hankel matrix reconstruction," *Appl. Comput. Harmon. Anal.*, vol. 41, pp. 470-490, 2016.

[28] J.-F. Cai, S. Liu, and W. Xu, "A fast algorithm for reconstruction of spectrally sparse signals in super-resolution," *Proc. of SPIE, Wavelets and Sparsity XVI*, vol. 9597, 2015.

[29] F. Andersson, M. Carlsson, J. Y. Tourneret, and H. Wendt, "A new frequency estimation method for equally and unequally spaced data," *IEEE Trans. Signal Process.*, vol. 62, pp. 5761-5774, 2014.

[30] M. Fazel, T. K. Pong, D. Sun, and P. Tseng, "Hankel matrix rank minimization with applications to system identification and realization," *SIAM J. Matrix Anal. Appl.*, vol. 34, pp. 946-977, 2013.

[31] I. Markovsky and K. Usevich, "Structured low-rank approximation with missing data," *SIAM J. Matrix Anal. Appl.*, vol. 34, pp. 814-830, 2013.

[32] K. Usevich and P. Comon, "Hankel low-rank matrix completion: Performance of the nuclear norm relaxation," *IEEE J. Sel. Top. Signal Process.*, vol. 10, pp. 637-646, 2016.

[33] K. Kazimierczuk and V. Orekhov, "Accelerated NMR spectroscopy by using compressed sensing," *Angew. Chem. Int. Ed.*, vol. 123, pp. 5670-5673, 2011.

[34] D. J. Holland, M. J. Bostock, L. F. Gladden, and D. Nietlispach, "Fast multidimensional NMR spectroscopy using compressed sensing," *Angew. Chem. Int. Ed.*, vol. 50, pp. 6548-6551, 2011.

[35] Y. Chen and Y. Chi, "Robust spectral compressed sensing via structured matrix completion," *IEEE Trans. Inf. Theory*, vol. 60, pp. 6576-6601, 2014.

[36] J. Liu, P. Musialski, P. Wonka, and J. Ye, "Tensor completion for estimating missing values in visual data," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, pp. 208-220, 2013.

[37] Y. Liu, F. Shang, L. Jiao, J. Cheng, and H. Cheng, "Trace norm regularized CANDECOMP/PARAFAC decomposition with missing data," *IEEE Trans. Cybern.*, vol. 45, pp. 2437-2448, 2015.

[38] S. Gandy, B. Recht, and I. Yamada, "Tensor completion and low-n-rank tensor recovery via convex optimization," *Inverse Probl.*, vol. 27, pp. 025010, 2011.

[39] E. Acar, D. M. Dunlavy, T. G. Kolda, and M. Mørup, "Scalable tensor factorizations for incomplete data," *Chemometr. Intell. Lab. Syst.*, vol. 106, pp. 41-56, 2011.

[40] G. Tomasi and R. Bro, "PARAFAC and missing values," *Chemometr. Intell. Lab. Syst.*, vol. 75, pp. 163-180, 2005.

[41] M. Yuan and C.-H. Zhang, "On tensor completion via nuclear norm minimization". *Found. Comut. Math.*, vol. 7, pp. 1-38, 2014.

[42] M. Signoretto, L. D. Lathauwer, and J. A. K. Suykens, "Nuclear norms for tensors and their use for convex multilinear estimation," Tech. Rep. 10-186, K.U.Leuven, 2010.

[43] L. Yang, Z. H. Huang, and X. Shi, "A fixed point iterative method for low n-rank tensor pursuit," *IEEE Trans. Signal Process.*, vol. 61, pp. 2952–2962, 2013.

[44] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM Rev.*, vol. 51, pp. 455-500, 2009.

[45] J. D. Carroll and J.-J. Chang, "Analysis of individual differences in multidimensional scaling via an n-way generalization of Eckart-Young decomposition," *Psychometrika*, vol. 35, pp. 283-319, 1970.

[46] T. G. Kolda, "Multilinear operators for higher-order decompositions," Tech. Rep. SAND2006-2081, Sandia National Laboratories, 2006.

[47] Y. Yu, H. Cheng and X. Zhang, "Approximate low-rank tensor learning", *32nd NIPS Workshop on Optimization for Machine Learning, JMLR Workshop Conf. Proc. 37*, 2014.

[48] Q. Li, A. Prater, L. Shen and G. Tang, "Overcomplete tensor decomposition via convex optimization," *Arxiv: 1602.08614*, 2016.

[49] J.-F. Cai, E. J. Candès, and Z. Shen, "A singular value thresholding algorithm for matrix completion," *SIAM J. Optim.*, vol. 20, pp. 1956-1982, 2010.

[50] A. S. Stern, D. L. Donoho, and J. C. Hoch, "NMR data processing using iterative thresholding and minimum l_1-norm reconstruction," *J. Magn. Reson.*, vol. 188, pp. 295-300, 2007.

[51] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, pp. 1-122, 2011.

[52] Z. Zhan, J.-F. Cai, D. Guo, Y. Liu, Z. Chen and X. Qu, "Fast multiclass dictionaries learning with geometrical directions in MRI reconstruction," *IEEE Trans. Biomed. Eng.*, vol. 63, pp. 1850-1861, 2016.

[53] M. V. B. Laurent Sorber and Lieven De Lathauwer, "Tensorlab v2.0," 2014.

[54] Z. Lin, M. Chen, and Y. Ma, "The augmented lagrange multiplier method for exact recovery of corrupted low-rank matrices," *UIUC Technical Report UILU-ENG-09-2215*.

[55] Y.-L. Chen, C.-T. Hsu, and H.-Y. Liao, "Simultaneous tensor decomposition and completion using factor priors," *IEEE Trans. Pattern Anal. Mach.Intell.*, vol. 36, pp. 577-591, 2014.

[56] S. G. Hyberts, K. Takeuchi, and G. Wagner, "Poisson-Gap sampling and forward maximum entropy reconstruction for enhancing the resolution and sensitivity of protein NMR data," *J. Am. Chem. Soc.*, vol. 132, pp. 2145-2147, 2010.

[57] M. Mobli, "Reducing seed dependent variability of non-uniformly sampled multidimensional NMR data," *J. Magn. Reson.* , vol. 256, pp. 60-69, 2015 .

[58] V. Orekhov and V. A. Jaravine, "Analysis of non-uniformly sampled spectra with multi-dimensional decomposition," *Prog. Nucl. Magn. Reson. Spectrosc.*, vol. 59, pp. 271-292, 2011.

[59] Z. Song, P. Wu, P. Ji, J. Zhang, Q. Gong, J. Wu, and Y. Shi, "Solution structure of the second RRM domain of RBM5 and its unusual binding characters for different RNA targets," *Biochemistry*, vol. 51, pp. 6667–78, 2012.

[60] F. Delaglio, S. Grzesiek, G.W. Vuister, G. Zhu, J. Pfeifer, and A. Bax, "NMRPipe: A multidimensional spectral processing system based on Unix pipes," *J. Biomol. NMR*, vol. 6, pp. 277-293, 1995.

[61] B. Blümich and D. Ziessow, "Skyline projections in two-dimensional NMR spectroscopy," *J. Magn. Reson.*, vol. 49, pp. 151-154, 1982.

[62] J. Swärd, S. I. Adalbjörnsson and A. Jakobsson, "High resolution sparse estimation of exponentially decaying N-dimensional signals," *Signal Process.*, vol. 128, pp. 309–317, 2016.

[63] V. Y. Orekhov, I. V. Ibraghimov, and M. Billeter, "MUNIN: A new approach to multi-dimensional NMR spectra interpretation," *J. Biomol. NMR*, vol. 20, pp. 49–60, 2001.

[64] E. Acar, D. M. Dunlavy, T. G. Kolda, and M. Mørup, "Scalable tensor factorizations with missing data," in *Proc. SIAM International Conference on Data Mining*, 2010, pp. 701-712.

[65] S. Burer and R. D. Monteiro, "Local minima and convergence in low rank semidefinite programming," *Math. Program.*, vol. 103, pp. 427-444, 2005.

[66] M. Mardani, G. Mateos, and G. B. Giannakis, "Subspace learning and imputation for streaming big data matrices and tensors," *IEEE Trans. Signal Process.*, vol. 63, pp. 2663–2677, 2015.

[67] N. Srebro and A. Shraibman, "Rank, trace-norm and max-norm," in *Proc. of Learning Theory*, pp. 545-560, 2005.