Efficient FCTN Decomposition With Structural Sparsity for Noisy Tensor Completion

Wei-Jian Huang, Li Huang, *Member, IEEE*, Tai-Xiang Jiang, *Member, IEEE*, Yu-Bang Zheng, and Guisong Liu, *Member, IEEE*

***Abstract*—Recently, the fully-connected tensor network (FCTN) decomposition has shown a powerful capability of depicting intrinsic correlations between any pair of tensor modes. But there exists a challenging question in FCTN decomposition-based methods, i.e., the accurate determination of the complicated FCTN-rank, which contains $N(N-1)/2$ elements for Nth-order tensors. In this paper, we design a structural sparsity regularization for the FCTN decomposition, which estimates the complicated FCTN-rank by adaptively pruning near-zero groups in FCTN factor. Based on this regularization, we propose a noisy tensor completion (NTC) model, aiming at the recovery of a tensor from its partial and noisy observation. Besides, we design a proximal alternating minimization (PAM)-based algorithm to solve the model. In theorem, we prove a guarantee for the global convergence of the developed algorithm. To further accelerate our method for large-scale data sets, we customize the randomized block sampling strategy for general tensor network decomposition methods by updating factors from small samples. Experiments demonstrate that our strategy can accurately estimate the FCTN-rank and achieve better reconstruction performances, and our methods outperform the state-of-the-art methods in the reconstruction of different types of real-world tensors.**

***Index Terms*—Tensor decomposition, noisy tensor completion (NTC), structural sparsity, randomized block sampling.**

I. Introduction

THE tensor as a multidimensional array is a natural representation of high-dimensional real-world data, such as color images, color videos, hyper-spectral videos, audio data, and traffic data. The tensor decomposition (TD), an effective tool for mining high-dimensional tensor data, factorizes high-dimensional data into a series of latent factors through multi-linear operations. These concepts can be traced back to Hitchcock in the 1920s [1], [2] and Cattell in 1944 [3], [4]. By far, numerous efforts have been devoted into the TD and interests have expanded to various fields, such as signal processing [5], [6], computer vision [7], [8], machine learning [9], [10], [11], and neuroscience [12], [13].

In the past decades, two widely investigated TD schemes are the Tucker decomposition [14], [15] and the CANDECOMP/PARAFAC (CP) decomposition [16], [17]. The Tucker decomposition factorizes a target tensor into factor matrices and a factor tensor where each factor matrix is "connected" to the factor tensor via the tensor-matrix product. The CANDECOMP/PARAFAC (CP) decomposition decomposes a tensor into the sum of rank-one tensors constructed by outer products of vectors. Recently, the tensor singular value decomposition (t-SVD) framework [18], [19], [20], which is based on the tensor-tensor product between third-order tensors, newly emerges recently and operates third-order tensors integrally without any matricization or flattening. As above TDs are capable of characterizing the inner low-dimensional structure of real-world tensor data, and concepts of rank or the nuclear norm for tensors derived from them, are utilized for the tensor data recovery; See [21], [22], [23], [24], [25], [26].

In this work, we focus our attention on the popular tensor network (TN) [27], [28] decomposition, which represents a higher-order tensor as interconnected small-scale sub-tensors (also called TN factors). In the past ten years, many TN-based TDs have emerged and two of the most representative ones are the tensor train (TT) [29] decomposition and the tensor ring (TR) [30], [31] decomposition. The TT decomposition represents an Nth-order tensor as $N-2$ third-order sub-tensors and two matrices. Based on this foundation, the TR decomposition replaced two matrices with a cyclically connected third-order tensor resulting in a ring structure. As TT and TR have shown great abilities to characterize the higher-order tensor data's intrinsic inner correlation, which is essential in the area of higher-order tensor data recovery, they have been successfully applied for tensor recovery, e.g., tensor completion [32], [33], [34], [35] and the more challenging task of noisy tensor completion [36]. However, one imperfection of TT and TR is that they concentrate more on the local connections between adjacent modes of higher-order tensors, leading to limited descriptions of the global correlation of tensors. Very recently, Zheng et al. proposed a fully-connected tensor network (FCTN)

Received 12 January 2024; revised 9 September 2024; accepted 24 June 2025. Date of publication 10 July 2025; date of current version 13 November 2025. This work was supported in part by Sichuan Science and Technology Program under Grant 2024ZYD0147 and Grant 2024NSFSC0796, in part by Natural Science Foundation of Xinjiang Uygur Autonomous Region under Grant 2024D01A18, in part by the National Natural Science Foundation of China under Grant 62301456, and in part by the China Postdoctoral Researcher Funding Program under Grant GZB20230605 and Grant 2025T180066. Recommended for acceptance by N. Ramakrishnan. *(Corresponding author: Tai-Xiang Jiang.)*

Wei-Jian Huang, Li Huang, Tai-Xiang Jiang, and Guisong Liu are with the School of Computing and Artificial Intelligence, Southwestern University of Finance and Economics, Chengdu 610074, China, also with the Kash Institute of Electronics and Information Industry, Kash 844099, China, and also with Engineering Research Center of Intelligent Finance, Ministry of Education, Southwestern University of Finance and Economics, Chengdu 610074, China (e-mail: huangwj012@gmail.com; lihuang@swufe.edu.cn; taixiangjiang@gmail.com; gliu@swufe.edu.cn).

Yu-Bang Zheng is with the School of Information Science and Technology, Southwest Jiaotong University, Chengdu 611756, China (e-mail: zhengyubang@163.com).

Digital Object Identifier 10.1109/TBDATA.2025.3588029

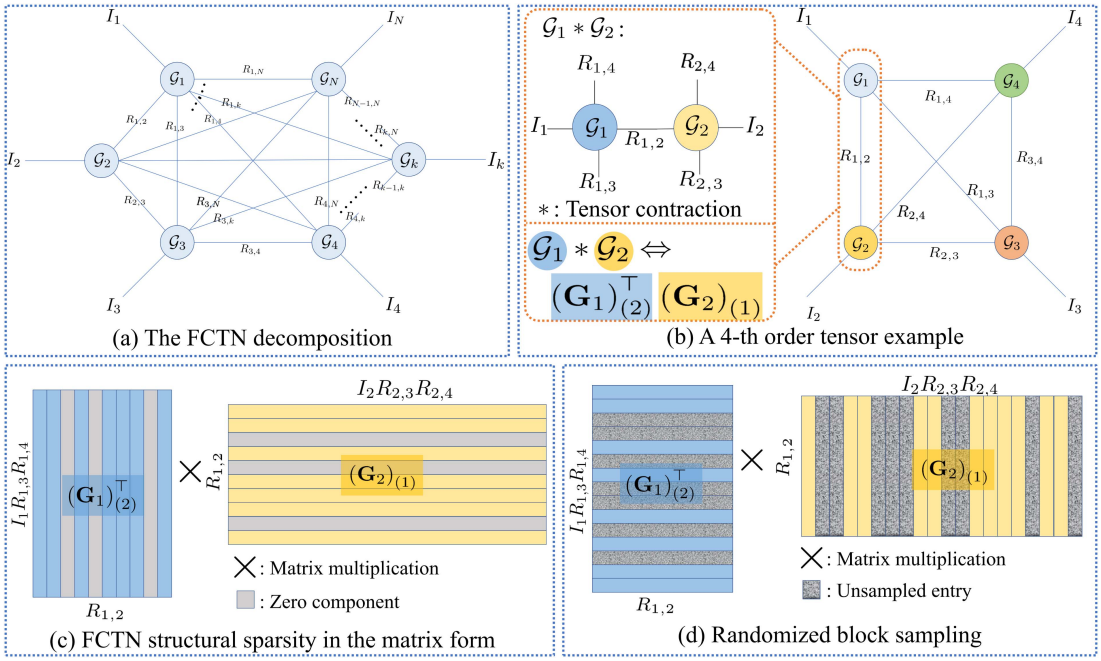

Fig. 1. Illustrations of the structural sparsity regularization and the randomized block sampling.

decomposition [37], which decomposes an Nth-order tensor into a set of Nth-order small-sized factor tensors. In the FCTN decomposition, exhibited in Fig. 1(a), each factor tensor is connected to all other factor tensors and thus the correlations between any two modes of the tensor are well encoded. However, there are two main limitations inherent in FCTN-based methods. First, FCTN decomposition is unable to adaptively determine the FCTN-rank, which contains $N(N-1)/2$ different elements for an Nth-order tensor. The accurate estimation of the FCTN-rank is difficult in itself, and the degradation of the tensor data in tensor recovery applications would further make it harder. Second, the computational complexity of TN-based decompositions is particularly high for high-order tensors, mainly due to the large-scale factor tensor computations involved. These limitations also apply for other TN-based decompositions, e.g., tensor wheel decomposition [38].

To address these two limitations, this paper proposes ideas aimed at addressing them. First, we design the structural sparsity regularization for the FCTN decomposition and apply it to the noisy tensor completion (NTC) problem. The proposed method can explore the low-rank structure and estimate the FCTN-rank automatically with high accuracy, and it can effectively address the performance degradation of FCTN decomposition. Second, we design a randomized block sampling strategy for tensor decompositions, which can significantly reduce the computational complexity. Specifically, our contributions include the following:

- We propose a structural sparsity regularization for the FCTN decomposition. We consider the equivalent matrix form of the tensor contraction and introduce the structural sparsity into the factor tensors within the framework of the FCTN decomposition. We illustrate that this regularization term can not only enhance the global low-rankness but also help to determine the complicated FCTN-rank values automatically with high accuracy. Then, we propose a structural sparsity regularized FCTN decomposition model for tensor completion from partial and noisy observations.
- A proximal alternating minimization (PAM) based algorithm is proposed to solve the proposed noisy tensor completion model. We prove that the sequence generated by our algorithm globally converges to a critical point. Besides, we customize the randomized block sampling method for the tensor neural network decomposition scheme to further accelerate our algorithm with acceptable performance impairments. Experiments carried out with synthetic and real-world data and results demonstrate the effectiveness of our method.

The rest of this paper is organized as follows. Section III presents basic notations and preliminaries. Section IV designs a structural sparsity regularization, and proposes a noisy tensor completion model with efficient solving algorithms. Section V proves a theoretical guarantee for the convergence. Sections VI and VII show the experimental results and conclusions, respectively.

II. Related Work

Tensor decompositions and low-rank tensor completion techniques have been extensively studied, with various methods developed to address different challenges in high-dimensional data processing. In this section, we review the most relevant work in the field, focusing on classical tensor decomposition methods and recent advancements in low-rank tensor completion.

A. Tensor Decomposition

a) CP decomposition: The CANDECOMP/PARAFAC (CP) decomposition [16], [17] is one of the earliest tensor decomposition techniques and has been applied to fields such as machine

learning [39], signal processing [9], and data mining [40]. CP decomposition factorizes a tensor into a sum of rank-one components, expressed as:

$$\mathcal{A} = \sum_{r=1}^{R} \mathbf{a}_1 \circ \mathbf{a}_2 \circ \cdots \circ \mathbf{a}_N,$$

where R is the minimum number of components required to express the tensor. Although CP decomposition is simple and interpretable, its primary challenge is in determining the rank [41] and the computational cost associated with fitting the model, especially for large-scale tensors.

b) Tucker decomposition: Tucker decomposition [14], [15] is another widely-used tensor decomposition technique preserving important interactions across different dimensions. The Tucker decomposition expresses of an Nth-order tensor as:

$$\mathcal{A} = \mathcal{B} \times_1 \mathbf{U}_1 \times_2 \mathbf{U}_2 \cdots \times_N \mathbf{U}_N,$$

where $\times_k$ represents mode-k multiplication [42]. Tucker decomposition is extensively applied in fields such as compressed sensing [43], high-order data analysis [44], and signal processing. The challenge with Tucker decomposition lies in selecting the appropriate ranks for each mode.

c) Tensor singular value decomposition (t-SVD): The t-SVD [19], [20], based on the t-product, extends the classical matrix SVD to third-order tensors. It decomposes a third-order tensor as:

$$\mathcal{A} = \mathcal{U} *_t \mathcal{S} *_t \mathcal{V}^H,$$

where $*_t$ denotes the tensor-tensor product proposed in [19]. T-SVD has proven effective in applications such as including image/video completion [25], [45], [46] and hyperspectral denoising [47] for its algebraic framework to handle the third-order tensor while preserving intrinsic structural information. However, t-SVD is primarily limited to third-order tensors and becomes less efficient for higher-order tensors.

d) Tensor network decomposition: Tensor network (TN)-based methods, such as tensor train (TT) [29] and tensor ring (TR) [30], have been widely studied in quantum physics, scientific computing, and machine learning [48]. These methods represent a high-order tensor as a sequence of lower-order tensors, reducing the computational complexity and storage requirements. FCTN [37] and tensor wheel (TW) [31] decompositions represent more recent advancements, where each factor tensor is connected to every other tensor or the core tensor, capturing all possible interactions between tensor modes.

B. Low-Rank Tensor Completion

The problem of low-rank tensor completion aims to recover missing entries by leveraging the low-rank structure of the tensor across multiple dimensions. Two common approaches are:

a) Rank minimization: Rank minimization methods seek to recover the missing entries by minimizing the rank of the tensor. This can be formulated as:

$$\min_{\mathcal{X}} \texttt{rank}(\mathcal{X}) \quad \text{s.t.} \quad \mathcal{X}_\Omega = \mathcal{M}_\Omega, \tag{1}$$

where $\texttt{rank}(\mathcal{X})$ refers to the rank of tensor $\mathcal{X}$, $\mathcal{M}$ is the observed tensor, and Ω is the set of observed indices. As minimizing the rank function is hard, convex/nonconvex surrogates such as the nuclear norm are typically used in practice [21]. While rank minimization approaches can recover low-rank tensors effectively, they often struggle with scalability when applied to large tensors due to the high computational cost involved in the computation of SVD [49].

b) Tensor decomposition: The second approach is tensor completion using decomposition. It is assumed that the target tensor has an inherent low-rank structure that can be captured using various tensor decomposition techniques. The problem can be formulated as:

$$\min_{\mathcal{X},\mathcal{G}_i} \frac{1}{2}||\mathcal{X} - \texttt{TD}(\mathcal{G}_i)||_F^2 \quad \text{s.t.} \mathcal{X}_\Omega = \mathcal{M}_\Omega, \tag{2}$$

where $\texttt{TD}(\mathcal{G}_i)$ denotes a specific tensor decomposition and $\mathcal{G}_i$s indicate the corresponding factor vectors/matrices/tensors. Tensor decomposition-based methods provide a scalable alternative to rank minimization but require the selection of an appropriate rank prior, which can be challenging in practical applications.

III. Notations and Preliminaries

We use $x, \mathbf{x}, \mathbf{X}$ and $\mathcal{X}$ to denote scalars, vectors, matrices, and tensors, respectively. Given an Nth-order tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times \cdots \times I_N}$, its $(i_1, i_2, \ldots, i_N)$-th element can be denoted by $\mathcal{X}(i_1, i_2, \ldots, i_N)$ or $\mathcal{X}_{i_1,i_2,\ldots,i_N}$. The vector $\mathbf{n} = [n_1, n_2, \ldots, n_N]$ based tensor permutation of the tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times \cdots \times I_N}$, denoted by $\mathcal{X}^{\mathbf{n}} \in \mathbb{R}^{I_{n_1} \times \cdots \times I_{n_N}}$, is obtained by rearranging modes of $\mathcal{X}$ in the order specified by $\mathbf{n}$. We use the symbol $\odot$ to denote the Hadamard (element-wise) product, and we use symbol $\circ$ to denote the outer product. The inner product between two same-sized tensors $\mathcal{A}$ and $\mathcal{B}$ is defined as $\langle \mathcal{A}, \mathcal{B} \rangle = \sum_{i_1,i_2,\ldots,i_N} \mathcal{A}_{i_1,i_2,\ldots,i_N} \mathcal{B}_{i_1,i_2,\ldots,i_N}$. The Frobenius norm of $\mathcal{X}$is defined as $\|\mathcal{X}\|_F = \sqrt{\langle \mathcal{X}, \mathcal{X} \rangle}$. For brevity of exposition, we use $\mathbf{x}_{1:d}$ to denoted a variable of vector $[x_1, x_2, \ldots, x_d]$. $\texttt{vec}(\cdot)$ is the vectorization operation reshaping all the entries into a column vector. Then, we restate those important and frequently used definitions.

Definition 1 (mode-n unfolding [42]): The mode-n unfolding of $\mathcal{A} \in \mathbb{R}^{I_1 \times \cdots \times I_N}$ is a matrix, which can be denoted as $\mathbf{A}_{(n)} \in \mathbb{R}^{I_n \times \prod_{m=1, m \neq n} I_m}$. The $(i_1, i_2, \ldots, i_N)$-th element of $\mathcal{A}$ is mapped to the (i_n, j)-th element of $\mathbf{A}_{(n)}$, where $j = 1 + \sum_{k=1,k\neq n}^{N} (i_k - 1) J_k$ with $J_k = \prod_{m=1,m\neq n}^{k-1} I_m$.

Intuitively, the lexicographic ordering of the mode-n unfolding of tensor $\mathcal{A}$ arranges the fibers along mode n. The columns of $\mathbf{A}_{(n)}$ can be denoted as $\mathcal{A}(i_1, i_2, \ldots, i_{n-1}, [1 : I_n], i_{n+1}, \ldots, i_N)$. It can also be easily accomplished with the Matlab command[1] `reshape` and we have $\mathbf{A}_{(n)} = \texttt{reshape}(\mathcal{A}^{\mathbf{b}}, [I_{\mathbf{b}_1}, \prod_{i=2}^{N} I_{\mathbf{b}_i}])$, where $\mathbf{b} = [n, 1, \ldots, n-1, n+1, \ldots, N]$. The inverse operation is denoted as mode-n folding.

Definition 2 (n-unfolding [30]): The n-unfolding of an Nth-order tensor $\mathcal{A} \in \mathbb{R}^{I_1 \times \cdots \times I_N}$ is a matrix, which can be denoted as $\mathbf{A}_{\langle n \rangle} \in \mathbb{R}^{I_1 \cdots I_n \times I_{n+1} \cdots I_N}$. And it can be implemented by

[1] https://www.mathworks.com/help/matlab/ref/reshape.html

Matlab command $\mathbf{A}_{\langle n\rangle} = \texttt{reshape}(\mathcal{A}, [\prod_{i=1}^{n} I_i, \prod_{i=n+1}^{N} I_i])$, where the first n indices enumerate the rows of $\mathbf{A}_{\langle n\rangle}$ and the last $N-n$ indices for its columns.

Definition 3 (tensor contraction [37]): $\mathcal{A} \in \mathbb{R}^{I_1\times I_2\times\cdots\times I_N}$ and $\mathcal{B} \in \mathbb{R}^{J_1\times J_2\times\cdots\times J_M}$ owns d common dimensions, i.e., $I_{n_i} = J_{m_i}$ with $i = 1, 2, \ldots, d$. $\mathbf{n}$ and $\mathbf{m}$ are permutations of vectors $[1, \ldots, N]$ and $[1, \ldots, M]$, respectively. The result of the contraction between $\mathcal{A}$ and $\mathcal{B}$ along their common dimensions is $\mathcal{C} \in \mathbb{R}^{I_{n_{d+1}}\times\cdots\times I_{n_N}\times J_{m_{d+1}}\cdots\times J_{m_N}}$, which is defined as $\mathcal{C} = \mathcal{A} \times_{\mathbf{n}_{1:d}}^{\mathbf{m}_{1:d}} \mathcal{B}$. Element-wisely, we have

$$\mathcal{C}(i_{n_{d+1}}, \ldots, i_{n_N}, j_{m_{d+1}}, \ldots, j_{m_N}) = \sum_{i_{n_1}=1}^{I_{n_1}} \sum_{i_{n_2}=1}^{I_{n_2}} \cdots \sum_{i_{n_d}=1}^{I_{n_d}} \mathcal{A}^{\mathbf{n}}(i_{n_1}, \ldots, i_{n_N}) \cdot \mathcal{B}^{\mathbf{m}}(i_{n_1}, \ldots, j_{m_M}).$$

Definition 4 (Fully-Connected Tensor Network Decomposition [37]): The FCTN decomposition of an Nth-order tensor $\mathcal{A} \in \mathbb{R}^{I_1\times\cdots\times I_N}$ is denoted as $\mathfrak{F}(\{\mathcal{G}_k\}_{k=1}^N)$, where $\{\mathcal{G}_k\}_{k=1}^N$ is the set of all FCTN factor tensors. Each $\mathcal{G}_k \in \mathbb{R}^{R_{1,k}\times\cdots\times R_{k-1,k}\times I_k\times R_{k,k+1}\times\cdots\times R_{k,N}}$ $(k = 1, 2, \ldots, N)$ is connected with all other factor tensors through the tensor contraction. Element-wisely, the FCTN decomposition can be formulated as

$$\mathcal{A}(i_1, i_2, \ldots, i_N) = \sum_{r_{1,2}=1}^{R_{1,2}} \cdots \sum_{r_{1,N}=1}^{R_{1,N}} \sum_{r_{2,3}=1}^{R_{2,3}} \cdots \sum_{r_{N-1,N}=1}^{R_{N-1,N}} \mathcal{G}_1(i_1, r_{1,2}, \ldots, r_{1,N}) \mathcal{G}_2(r_{1,2}, i_2, r_{2,3}, \ldots, r_{2,N}) \cdots \mathcal{G}_k(r_{1,k}, \ldots, i_k, r_{k,k+1}, \ldots, r_{k,N}) \cdots \mathcal{G}_N(r_{1,N}, r_{2,N}, \ldots, r_{N-1,N}, i_N).$$

The vector, constituted by the minimal positive values $R_{i,j}$s $(1 \le i < j \le N$ and $i, j \in \mathbb{N}^+)$ required to express $\mathcal{A}$, is correspondingly called the FCTN-rank.

IV. Noisy Tensor Completion Via Structural Sparsity Regularized FCTN

As the graphical representation of the FCTN decomposition illustrated in Fig. 1(a) illustrates and we can see that any two factor tensors are connected. The fully connected structure enables it to adequately characterize the intrinsic correlations between any two modes. However, this also leads to the complicated FCTN-rank of $O(N^2)$ positive values for an Nth-order tensor. Meanwhile, when applying FCTN decomposition for tensor data recovery, the noise or missing values would further burden the users for the accurate identification of the FCTN-rank. In this section, we would first tailor the structural sparsity regularization for the FCTN decomposition and then propose the NTC model and algorithms. The randomized block sampling strategy designed to accelerate our NTC method is subsequently introduced.

A. Structural Sparsity Regularization for FCTN

Although the structure of FCTN is complicated, the specific connection between any prescribed two factor tensors is clear and simple. We can see in the graphical representation of FCTN in Fig. 1(a) that there is only one edge between any two factor tensors. As we know that the tensor contraction operation would degenerate to the common matrix product when $d = 1$ in Definition 3. Thus, the interaction between any two factor tensors is indeed equivalent to the matrix multiplication between two matrices. To be specific, we take the fourth order tensor of size $I_1 \times I_2 \times I_3 \times I_4$ shown in Fig. 1(b) as an example. Six positive values, from $R_{1,1}$ to $R_{3,4}$, constitute its FCTN-rank. Let's focus on the first and second factor tensors, i.e., the orange dashed line circled $\mathcal{G}_1$ and $\mathcal{G}_2$ in Fig. 1(b). As mentioned above, there is only one same-sized dimension between $\mathcal{G}_1 \in \mathbb{R}^{I_1\times R_{1,2}\times R_{1,3}\times R_{1,4}}$ and $\mathcal{G}_2 \in \mathbb{R}^{R_{1,2}\times I_2\times R_{2,3}\times R_{2,4}}$ and their contraction[2] along their shared edge is equivalent to the matrix multiplication between their unfolding matrices, i.e., $(\mathbf{G}_1)_{(2)}^\top$ and $(\mathbf{G}_2)_{(1)}$.

With the above analysis, we can see that the elements in FCTN-rank are indeed the rank values of the multiplication results between unfolding matrices. Therefore, exploring the low FCTN-rank structure of a tensor is to find low-rank structures of a series of matrices. Similar insights can be found in [50]. However, different from [50], which considers the overall unfolding matrices of the tensor, we focus on specific low-rank structures corresponding to connections between any two factor tensors, resorting to the following theoretical result.

Theorem 1 (Lemma 3 in [51]): Given a matrix $\mathbf{X} \in \mathbb{R}^{m\times n}$ with $\text{rank}(\mathbf{X}) = r \le \min(m, n)$ and non-zero singular values σ_is $(i = 1, 2, \ldots, r)$, there exist two factor matrices $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_d] \in \mathbb{R}^{m\times d}$ and $\mathbf{B} = [\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_d]^\top \in \mathbb{R}^{n\times d}$ with $r \le d \le \min(m, n)$, such that the Schatten-$1/2$ norm (a quasi norm) of $\mathbf{X}$, denoted and defined as $\|\mathbf{X}\|_{S_{1/2}}^{1/2} \triangleq \sum_{i=1}^{r} \sigma_i^{1/2}$, satisfies

$$\|\mathbf{X}\|_{S_{1/2}}^{1/2} = \min_{\mathbf{X}=\mathbf{A}\mathbf{B}^\top} \sum_{k=1}^{d} \sqrt{\frac{\|\mathbf{a}_k\|_2^2 + \|\mathbf{b}_k\|_2^2}{2}}, \tag{3}$$

and the equation holds when $\mathbf{X}$ admits the singular value decomposition $\mathbf{X} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^\top$, $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}^{\frac{1}{2}}$, and $\mathbf{B} = \mathbf{V}\mathbf{\Sigma}^{\frac{1}{2}}$. Moreover, we can generalize the above to a tensor version, which can be defined as

$$||\mathbf{X}||_{S_{1/2}}^{1/2} = \min_{\mathcal{X}=\mathcal{A}*\mathcal{B}} \sum_{k=1}^{d} \sqrt{\frac{||\mathcal{A}_{:,\ldots,k,\ldots,:}||_F^2 + ||\mathcal{B}_{:,\ldots,k,\ldots,:}||_F^2}{2}}, \tag{4}$$

where $\mathcal{X} = \texttt{fold}(\mathbf{X})$, and $\texttt{fold}(.)$ denotes the inverse operation of n-folding.

Guided by Theorem 1, we propose the structural sparsity regularization for a tensor $\mathcal{X}$ with its FCTN decomposition, i.e.,

[2]We use the symbol $*$ to denote their contraction for simplicity.

$\mathcal{X} = \mathfrak{F}(\{\mathcal{G}_k\}_{k=1}^N) \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$, as

$$\mathfrak{S}\left(\{\mathcal{G}_k\}_{k=1}^N\right) = \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \sum_{r=1}^{R_{i,j}} \left(\left\| \mathcal{G}_i^{\{j,r\}} \right\|_F^2 + \left\| \mathcal{G}_j^{\{i,r\}} \right\|_F^2 \right)^{\frac{1}{2}}, \tag{5}$$

where $\mathcal{G}_i^{\{j,r\}}$ $(r = 1, 2, \ldots, R_{i,j})$ denotes the r-th subtensor of $\mathcal{G}_i \in \mathbb{R}^{R_{1,i} \times \cdots \times R_{i-1,i} \times I_i \times R_{i,i+1} \times \cdots \times R_{i,N}}$ along its j-th dimension, i.e., $\mathcal{G}_i^{\{j,r\}} = \mathcal{G}_i([1 : R_{1,i}], \ldots, [1 : R_{i,j-1}], r, [1 : R_{i,j+1}], \ldots, [1 : R_{i,N}])$, and the same applies to $\mathcal{G}_j^{\{i,r\}}$s. In (5), the first two summations are corresponding to $(N(N-1))/2$ FCTN-rank values, and the remaining $\sum_{r=1}^{R_{i,j}} (\|\mathcal{G}_i^{\{j,r\}}\|_F^2 + \|\mathcal{G}_j^{\{i,r\}}\|_F^2)^{1/2}$ reflects the structural sparsity of unfolding matrices of $\mathcal{G}_i$ and $\mathcal{G}_j$. As stated in Theorem 1, minimizing it would undoubtedly enhance the low-rankness of the contraction result of two factor tensors, yielding the entire low-dimensionality of the tensor data.

Intuitively, as shown in Fig. 1(c), regularizing the structural sparsity would result in a few columns or rows near to zero in the unfolding factor matrices. Thus, on the one hand, we can effectively exploit the low-rankness without the heavy burden of SVD computations. On the other hand, it is reasonable and connivent to prune those zero columns and rows (collectively called *groups*) for an adaptive and accurate determination of the FCTN-rank.

B. Proposed Noisy Tensor Completion Model

With the structural sparsity regularization for FCTN designed in (5), our noisy tensor completion model is formulated as

$$\min_{\{\mathcal{G}_k\}_{k=1}^N, \mathcal{X}} \left\| \mathcal{X} - \mathfrak{F}(\{\mathcal{G}_k\}_{k=1}^N) \right\|_F^2 + \beta \left\| \mathcal{W} \odot (\mathcal{X} - \mathcal{B}) \right\|_F^2 + \lambda \mathfrak{S}\left(\{\mathcal{G}_k\}_{k=1}^N\right), \tag{6}$$

where $\mathcal{B} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ is the observed tensor, β and λ are two nonnegative trade-off parameters,

$$\mathcal{W}_{i_1, \ldots, i_N} = \begin{cases} 1, & \text{if } (i_1, \ldots, i_N) \in \Omega, \\ 0, & \text{otherwise,} \end{cases}$$

and Ω denotes the index set of observed entries.

In (6), $\|\mathcal{X} - \mathfrak{F}(\{\mathcal{G}_k\}_{k=1}^N)\|_F^2$ explore the low-rankness by mining the FCTN decomposition structure of the resulting tensor $\mathcal{X}$. The weighted Frobenius norm, i.e., $\frac{\beta}{2}\|\mathcal{W} \odot (\mathcal{X} - B)\|_F^2$, accounts for the degradation process of missing entries and the Gaussian noise corruption. The structural sparsity regularization $\mathfrak{S}(\{\mathcal{G}_k\}_{k=1}^N)$ would be helpful to further exploit the precise low-rank structure and ensure the robustness of our model to disturbances, such as the noise and entry missing.

It can be found that direct minimizing the structural sparsity regularization $\mathfrak{S}(\{\mathcal{G}_k\}_{k=1}^N)$ is difficult. Considering the inequality of arithmetic and geometric means, we have

$$\left(\left\| \mathcal{G}_i^{\{j,r\}} \right\|_F^2 + \left\| \mathcal{G}_j^{\{i,r\}} \right\|_F^2 \right)^{\frac{1}{2}} \leq \frac{\frac{\left\| \mathcal{G}_i^{\{j,r\}} \right\|_F^2 + \left\| \mathcal{G}_j^{\{i,r\}} \right\|_F^2}{x} + x}{2}$$

for any $x > 0$ and the equality holds when $x = (\|\mathcal{G}_i^{\{j,r\}}\|_F^2 + \|\mathcal{G}_j^{\{i,r\}}\|_F^2)^{1/2}$. Thus, we adopt the iteratively reweighted least squares (IRLS) strategy [51], [52] and turn to solve the following model

$$\min_{\{\mathcal{G}_k\}_{k=1}^N, \mathcal{X}} \left\| \mathcal{X} - \mathfrak{F}(\{\mathcal{G}_k\}_{k=1}^N) \right\|_F^2 + \beta \left\| \mathcal{W} \odot (\mathcal{X} - \mathcal{B}) \right\|_F^2 \lambda \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \sum_{r=1}^{R_{i,j}} \eta_r^{(i,j)} \left(\|\mathcal{G}_i^{\{j,r\}}\|_F^2 + \|\mathcal{G}_j^{\{i,r\}}\|_F^2 \right), \tag{7}$$

where $\eta_r^{(i,j)} = \frac{1}{\sqrt{\|\mathcal{G}_i^{\{j,r\}}\|_F^2 + \|\mathcal{G}_j^{\{i,r\}}\|_F^2}}$.

C. Proposed PAM Algorithm

In this part, we employ the framework of PAM and give the solving algorithm for our noisy tensor completion model (7).

a) $\mathcal{G}_k$ subproblem ($k = 1, \ldots, N$): For $k = 1, \ldots, N$, with other variables fixed, we update $\mathcal{G}_k$ by solving the following subproblem

$$\min_{\mathcal{G}_k} \left\| \mathcal{X}^{(s)} - \mathfrak{F}(\mathcal{G}_{1:k-1}^{(s+1)}, \mathcal{G}_k, \mathcal{G}_{k+1:N}^{(s)}) \right\|_F^2 + \rho_1 \left\| \mathcal{G}_k - \mathcal{G}_k^{(s)} \right\|_F^2 + \lambda \sum_{i=1}^{k-1} \sum_{r=1}^{R_{i,k}} \eta_r^{(i,k)} \left\| \mathcal{G}_k^{\{i,r\}} \right\|_F^2 + \lambda \sum_{j=k+1}^{N} \sum_{r=1}^{R_{k,j}} \eta_r^{(k,j)} \left\| \mathcal{G}_k^{\{j,r\}} \right\|_F^2. \tag{8}$$

Moreover, $(\mathbf{G}_k)_{(k)}$ is the mode-k unfolding of $\mathcal{G}_k$, $\mathcal{M}_k$ is the incomplete FCTN structure consisting of all latent factors except $\mathcal{G}_k$, $(\mathbf{M}_k)_{\langle N-1 \rangle}$ is the $N-1$ unfolding of $\mathcal{M}_k$ (See Theorem 4 in [37]), ρ_1 is a nonnegative parameter, and $\mathcal{G}_k^{(s)}$ denote the value of $\mathcal{G}_k$ in the latest iteration. The summation term $\sum_{r=1}^{R_{i,k}} \eta_r^{(i,k)} \|\mathcal{G}_k^{\{i,r\}}\|_F^2$ can be viewed as the squared weighted Frobenius norm of $\mathcal{G}_k$ with assigning $\sqrt{\eta_r^{(i,k)}}$ to the r-th group of $\mathcal{G}_k$ along the i-th dimension, i.e., $\sum_{r=1}^{R_{i,k}} \eta_r^{(i,k)} \|\mathcal{G}_k^{\{i,r\}}\|_F^2 = \|\mathbf{W}^{(i,k)} \odot (\mathbf{G}_k)_{(i)}\|_F^2$, where columns of $\mathbf{W}^{(i,k)}$ are all the same as $[\sqrt{\eta_1^{(i,k)}}, \sqrt{\eta_2^{(i,k)}}, \ldots, \sqrt{\eta_{R_{i,k}}^{(i,k)}}]^\top$. The same applies to $\sum_{r=1}^{R_{k,j}} \eta_r^{(k,j)} \|\mathcal{G}_k^{\{j,r\}}\|_F^2$.

Thus, we unfold $\mathcal{G}_k$ in (8) and equivalently convert (8) into the matrix form as

$$\begin{aligned} \min_{\mathcal{G}_k} \quad & \left\| \mathbf{X}_{(k)}^{(s)} - (\mathbf{G}_k)_{(k)} (\mathbf{M}_k^{(s)})_{\langle N-1 \rangle} \right\|_F^2 \\ & + \lambda \sum_{i=1}^{k-1} \left\| \hat{\mathbf{W}}^{(i,k)} \odot (\mathbf{G}_k)_{(k)} \right\|_F^2 \\ & + \lambda \sum_{j=k+1}^{N} \left\| \hat{\mathbf{W}}^{(k,j)} \odot (\mathbf{G}_k)_{(k)} \right\|_F^2 \\ & + \rho_1 \left\| (\mathbf{G}_k)_{(k)} - (\mathbf{G}_k^{(s)})_{(k)} \right\|_F^2, \end{aligned} \tag{9}$$

where $\hat{\mathbf{W}}^{(i,k)}$ is reshaped from $\mathbf{W}^{(i,k)}$ according to the transition from $(\mathbf{G}_k)_{(i)}$ to $(\mathbf{G}_k)_{(k)}$, i.e., mode-i folding and mode-k unfolding, for $i = 1, 2, \ldots, k-1$. The same applies to $\hat{\mathbf{W}}^{(k,j)}$ for $j = k+1, \ldots, N$. As all columns of $\mathbf{W}^{(i,k)}$ are the same, it follows that all rows of $\hat{\mathbf{W}}^{(i,k)}$ are also the same. We use

Algorithm 1: FCTN-GS.

Input: The observed data $\mathcal{B} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$, the maximum value of the FCTN-rank $R_{\texttt{max}}$, the maximum iteration $s_{\texttt{max}} = 500$, the initial target tensor $\mathcal{X}^{(0)} = \mathcal{B}$, parameters λ, β, ρ_1, ρ_2, τ.
while Not converged and $s < s_{\max}$ **do**
  for $k = 1, 2, \ldots, N$ **do**
    Update $\mathcal{G}_k^{(s+1)}$ through (10)
    Update $\eta_r^{(i,k)}$s and $\eta_r^{(k,j)}$s via (11)
  end for
  Update $\mathcal{X}^{(s+1)}$ through (13) and (14)
  Update the FCTN-rank by pruning slices with the threshold τ
end while
Output: The reconstructed tensor $\mathcal{X}$ and the estimation of the FCTN-rank.

this common row vector to create a diagonal matrix $\Lambda^{(i,k)}$ with this common row placed along the diagonal. Thus, we have $\hat{\mathbf{W}}^{(i,k)} \odot (\mathbf{G}_k)_{(k)} = (\mathbf{G}_k)_{(k)}\Lambda^{(i,k)}$. Then, we consider the first order optimal condition and update $\mathcal{G}_k$ (or equivalently $(\mathbf{G}_k)_{(k)}$) by solving the following linear equation

$$(\mathbf{G}_k)_{(k)} = \left(\mathbf{X}_{(k)}^{(s)}(\mathbf{M}_k^{(s)})_{\langle N-1\rangle}^{\top} + \rho_1(\mathbf{G}_k^{(s)})_{(k)}\right)^{\dagger} \left(\lambda\left(\sum_{i=1}^{k-1}(\Lambda^{(i,k)})^2 + \sum_{j=k+1}^{N}(\Lambda^{(k,j)})^2\right) + \rho_1\mathbf{I}_R + (\mathbf{M}_k^{(s)})_{\langle N-1\rangle}^{\top}(\mathbf{M}_k^{(s)})_{\langle N-1\rangle}\right), \tag{10}$$

where † denotes pseudo-inverse. After updating $\mathcal{G}_k$, the weights are updated as

$$\eta_r^{(i,k)} = \frac{1}{\sqrt{\|\mathcal{G}_i^{\{k,r\}}\|_F^2 + \|\mathcal{G}_k^{\{i,r\}}\|_F^2}}, \tag{11}$$

where $r = 1, 2, \ldots, R_{i,k}$, and the same applies to $\eta_r^{(k,j)}$.

b) 2) $\mathcal{X}$ subproblem Then, the subproblem with respect to $\mathcal{X}$ is

$$\min_{\mathcal{X}} \left\|\mathcal{X} - \mathfrak{F}(\{\mathcal{G}_k^{(s+1)}\}_{k=1}^N)\right\|_F^2 + \beta\left\|\mathcal{W} \odot (\mathcal{X} - B)\right\|_F^2 + \rho_2\left\|\mathcal{X} - \mathcal{X}^{(s)}\right\|_F^2. \tag{12}$$

Considering the separability of the Frobenius norm, we have

$$\mathcal{X}_{i_1,\ldots,i_N}^{(s+1)} = \left(\frac{\mathfrak{F}(\{\mathcal{G}_k^{(s+1)}\}_{k=1}^N) + \rho_2\mathcal{X}^{(s)}}{1+\rho_2}\right)_{i_1,\ldots,i_N} \tag{13}$$

for $(i_1, \ldots, i_N) \notin \Omega$, and

$$\mathcal{X}_{i_1,\ldots,i_N}^{(s+1)} = \left(\frac{\mathfrak{F}(\{\mathcal{G}_k^{(s+1)}\}_{k=1}^N) + \beta\mathcal{B} + \rho_2\mathcal{X}^{(s)}}{1+\beta+\rho_2}\right)_{i_1,\ldots,i_N} \tag{14}$$

for $(i_1, \ldots, i_N) \in \Omega$. After updating $\mathcal{X}$, we compute the Frobenius norm of slices of different directions in all factor tensors. If the value is smaller than a preset positive value τ, this slice will be pruned and the FCTN-rank is then adjusted. Our algorithm to optimize (7) is summarized in Algorithm 1.

Algorithm 2: FCTN-GSR.

Input: The observed data $\mathcal{B} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$, the maximum value of the FCTN-rank $R_{\texttt{max}}$, the maximum iteration $s_{\max=500}$, the initial target tensor $\mathcal{X}^{(0)} = \mathcal{B}$, parameters λ, β, ρ_1, ρ_2, τ.
while not converged and $s < s_{\max}$ **do**
  for $k = 1, 2, \ldots, N$ **do**
    Obtain sampling size $(Z_1, \ldots, Z_N)$ through randomized block sampling strategy
    Conduct randomized block sampling to obtain $\mathcal{X}_z^{(s)}$ and $\mathcal{G}_k^{(s)Z_k}$s through sampling size
    Update $\eta_r^{(i,k)}$s and $\eta_r^{(k,j)}$s via (11)
    Update $\mathcal{G}_k^{(s+1)}$ through its related $\mathcal{X}_z^{(s)}$ and $\mathcal{G}_k^{(s)Z_k}$s
  end for
  Update $\mathcal{X}^{(s+1)}$ through (13) and (14)
  Update the FCTN-rank by pruning slices with the threshold τ
end while
Output: The reconstructed tensor $\mathcal{X}^{(s)}$ and the estimation of the FCTN-rank.

D. Randomized Block Sampling

Large-scale tensors are playing an increasingly important role in the real world as people become more capable of accessing data. So we introduce the randomized block sampling [53], [54], [55], [56] strategy to reduce the computational burden. As we will illustrate, the randomized block sampling naturally fits the TN decomposition. In the following, we show how to conduct randomized block sampling taking the FCTN as an example.

Given an Nth-order tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$, FCTN decomposition aims to find $\mathcal{G}_k$s via

$$\min_{\{\mathcal{G}_k\}_{k=1}^N, \mathcal{X}} \left\|\mathcal{X} - \mathfrak{F}(\{\mathcal{G}_k\}_{k=1}^N)\right\|_F^2, \tag{15}$$

we conduct random sampling along each dimension to obtain a sub-tensor $\mathcal{X}_Z \in \mathbb{R}^{Z_1 \times Z_2 \times \cdots \times Z_N}$, where $Z_k \leq I_k$ $(k = 1, 2, \ldots, N)$ is a positive integer representing the k-th sample size.

Specifically, in most TC problems with TN decomposition, we will encounter this form of optimization problem, such as

$$\min_{\mathcal{G}_k} ||\mathbf{X}_{(k)} - \mathbf{G}_{k(k)}\mathbf{M}_k||_F^2. \tag{16}$$

It can be easily observed that sampling along each dimension except k-th dimension will greatly reduce the computation of the $\mathbf{M}_k\mathbf{M}_k^T$ and $\mathbf{X}_{(k)}\mathbf{M}_k^T$. Given an Nth-order tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$, we conduct random sampling along each dimension to obtain a sub-tensor $\mathcal{X}_Z \in \mathbb{R}^{Z_1 \times Z_2 \times \cdots \times Z_N}$, where $Z_k \leq I_k$ $(k = 1, 2, \ldots, N)$ is a positive integer representing the k-th sample size. That is, for the k-th order of the length I_k, we only sample a length of Z_k uniformly at random. We can see from Fig. 1(d) that this sampling operation is equivalent to sampling the rows (or columns) of the factor tensor's unfolding matrix.

On the one hand, if the original tensor $\mathcal{X}$ maintains an FCTN decomposition structure, i.e., $\mathcal{X} = \mathfrak{F}(\{\mathcal{G}_k\}_{k=1}^N)$, its sub-tensor $\mathcal{X}_z$, of a much smaller size, can be expressed by the same FCTN decomposition structure as $\mathcal{X}_z = \mathfrak{F}(\{\mathcal{G}_k^{Z_k}\}_{k=1}^N)$, where $\mathcal{G}_k^{Z_k}$ is sampled from $\mathcal{G}_k$ and the sampling is along its k-th dimension. On the other hand, as we can see from Fig. 1(c) and (d), the sampling direction is orthogonal to the direction of employing the structural sparsity. Thus, when estimating the sub-tensor $\mathcal{X}_z$ and sub-factor tensor $\mathcal{G}_k^{Z_k}$s, our structural sparsity regularization can be simultaneously utilized.

The randomized block sampling preserves the FCTN structure and the dimension can be largely reduced. Thus, we turn to update the randomly sampled part of $\mathcal{G}_k$, i.e., $\mathcal{G}_k^{Z_k}$, instead of the entire $\mathcal{G}_k$ in Algorithm 1. For $k = 1, 2, \ldots, N$, the $\mathcal{G}_k^{Z_k}$ subproblem can be formulated as

$$\begin{aligned}\min_{\mathcal{G}_k^{Z_k}} & \left\| \mathcal{X}_z^{(s)} - \mathfrak{F}\left(\mathcal{G}_{1:k-1}^{(s+1)Z_{1:k-1}}, \mathcal{G}_k^{Z_k}, \mathcal{G}_{k+1:N}^{(s)Z_{k+1:N}}\right)\right\|_F^2 \\ & + \lambda \sum_{i=1}^{k-1}\sum_{r=1}^{R_{i,k}} \eta_r^{(i,k)} \|\mathcal{G}_k^{Z_k,\{i,r\}}\|_F^2 \\ & + \lambda \sum_{j=k+1}^{N}\sum_{r=1}^{R_{k,j}} \eta_r^{(k,j)} \|\mathcal{G}_k^{Z_k,\{j,r\}}\|_F^2 \\ & + \rho_1 \|\mathcal{G}_k^{Z_k} - \mathcal{G}_k^{(s)Z_k}\|_F^2. \end{aligned} \tag{17}$$

As (17) has the same form as (8), the optimal solution of the $\mathcal{G}_k^{Z_k}$ subproblem can be obtained by solving the equation similar to (10). For simplicity, we have summarized Algorithm 1 and extended it into Algorithm 2 by incorporating the randomized block sampling strategy. This strategy is applicable to other tensor network decomposition-based methods as their factors are similar to those used in FCTN. As a result, we developed accelerated versions of TW-TC and FCTN-TC, named TW-TCR and FCTN-TCR, respectively.

Additionally, to clearly illustrate the effectiveness of the randomized block sampling strategy, we present Tab. I, which lists the time complexities per iteration of methods that will be compared in the experimental part. In Table I, an Nth-order incomplete tensor $\mathcal{X} \in \mathcal{R}^{I\times I\times I\times\cdots\times I}$ is considered. The tensor rank values are all set to be R, and all the sampled data dimension sizes are assumed to be J. Moreover, d is the size of the columns in the dictionary matrix in [46], and k denotes the k-th unfolding.

V. Convergency Analysis

In this section, we establish the theoretical guarantee for the convergence of Algorithm 1, and its proof is provided in the following.

Theorem 2: Denoting that $f(\{\mathcal{G}_k\}_{k=1}^N, \mathcal{X})$ is the sum of $\Phi(\mathcal{X})$, $\mathfrak{S}(\{\mathcal{G}_k\}_{k=1}^N)$, and $h(\{\mathcal{G}_k\}_{k=1}^N, \mathcal{X})$. $\Phi(\mathcal{X})$ and $h(\{\mathcal{G}_k\}_{k=1}^N, \mathcal{X})$ are defined as $\beta\|\mathcal{W}\odot(\mathcal{X}-\mathcal{B})\|_F^2$ and $\|\mathcal{X} - \mathfrak{F}(\{\mathcal{G}_k\}_{k=1}^N)\|_F^2$, respectively. The sequence $\{\{\mathcal{G}_k^{(s)}\}_{k=1}^N, \mathcal{X}^{(s)}\}_{(s\in\mathbb{N})}$ obtained by the Algorithm 1 is bounded. And it converges to a critical point of $f(\{\mathcal{G}_k\}_{k=1}^N, \mathcal{X})$.

Algorithm 3: FCTN-TCR/TW-TCR.

Input: The observed data $\mathcal{B} \in \mathbb{R}^{I_1\times I_2\times\cdots\times I_N}$, the initial value of the rank R, the maximum value of the rank $R_{\max}$, the maximum iteration $s_{\max=500}$, the initial target tensor $\mathcal{X}^{(0)} = \mathcal{B}$, parameters β, ρ_1, ρ_2.
while not converged and $s < s_{\max}$ **do**
  for $k = 1, 2, \ldots, N$ **do**
    Obtain sampling size $(Z_1, \ldots, Z_N)$ through the randomized block sampling strategy
    Conduct randomized block sampling to obtain the sub-tensor $\mathcal{X}_z^{(s)}$ and sub-factor tensor $\mathcal{G}_k^{(s)Z_k}$
    Update each sub-factor tensor by solving a least square problem as in FCTN-TC [37] or TW-TC [31]
  end for
  Update $\mathcal{X}^{(s+1)}$ through (10) and (11)
  Update the rank R by increasing its value
end while
Output: The reconstructed tensor $\mathcal{X}^{(s)}$ and the estimation of the FCTN-rank.

As Algorithm 1 is factually a special instance of Algorithm 4 described in [58], the proof of Theorem 2 confirms Theorem 6.2 in [58] if the following 4 conditions are fulfilled:

i) $\{\mathcal{G}_k^{(s)}\}_{k=1}^N$ and $\mathcal{X}^{(s)}(s \in \mathbb{N})$ are bounded;
ii) $f(\{\mathcal{G}_k\}_{k=1}^N, \mathcal{X})$ is proper lower semi-continuous;
iii) the K-Łproperty of $f(\{\mathcal{G}_k\}_{k=1}^N, \mathcal{X})$ at each point;
iv) $f(\{\mathcal{G}_k\}_{k=1}^N, \mathcal{X})$ satisfies the relative error condition ((64) in [58]) and the sufficient decrease condition ((65) and (66) in [58]).

We first verify conditions i) to iii). As shown in Algorithm 1, the initial $\mathcal{G}_k^{(0)}(k = 1\cdots N)$ and $\mathcal{X}^{(0)}$ are apparently bounded. Therefore, we only need to verify that $\mathcal{G}_k^{(s+1)}$ and $\mathcal{X}^{(s+1)}$ are bounded when $\mathcal{G}_k^{(s)}$ and $\mathcal{X}^{(s)}$ are bounded. Denoting that $\|\mathcal{G}_k^{(s)}\|_F \le c$, $\|\mathcal{X}^{(s)}\|_F \le d$ and $\|\mathcal{A}_k^{(s)}\|_F \le e$, and according to (7)[3], we have

$$\begin{aligned}\left\|\mathcal{G}_1^{(s+1)}\right\| &\le \left(\left\|\mathcal{X}^{(s)}\right\|_F \left\|\mathcal{M}_1^{(s)}\right\|_F + \rho_1 \left\|\mathcal{G}_1^{(s)}\right\|_F\right) \\ &\quad \cdot \left\|\left(\mathbf{Q}_1^{(s)} + \lambda\mathbf{W}_1 + \rho_1\mathbf{I}_R\right)^{-1}\right\|_F \\ &\le \left(dc^{N-1} + \rho_1 c\right)\left(\sum_{i=1}^{j}\frac{1}{\nu_i + \lambda\eta_i + \rho_1}\right) \\ &\le \left(dc^{N-1} + \rho_1 c\right)\frac{\sqrt{j}}{\rho_1}, \end{aligned} \tag{18}$$

where $\mathbf{Q}_1^{(s)} = (\mathbf{M}_1^{(s)})_{\langle N-1\rangle}(\mathbf{M}_1^{(s)})_{\langle N-1\rangle}^T$, $\mathbf{W}_1 = \sum_{i=1}^{k-1}(\Lambda^{(i,k)})^2 + \sum_{j=k+1}^{N}(\Lambda^{(k,j)})^2$, ν_is $(i = 1, \ldots, j = I_1R)$ are the eigenvalues of $\mathbf{Q}_1^{(s)}$, η_is are the eigenvalues of $\mathbf{W}_1$. Thus, $\mathcal{G}_1^{(s+1)}$ is bounded. Similarly, we can obtain that $\mathcal{G}_2^{(s+1)}, \ldots, \mathcal{G}_N^{(s+1)}$

[3] $(\mathbf{G}_k)_{(k)} = (\mathbf{X}_{(k)}^{(s)}(\mathbf{M}_k^{(s)})_{\langle N-1\rangle}^\top + \rho_1(\mathbf{G}_k^{(s)})_{(k)})^\dagger \quad (\lambda(\sum_{i=1}^{k-1}(\Lambda^{(i,k)})^2 + \sum_{j=k+1}^{N}(\Lambda^{(k,j)})^2) + \rho_1\mathbf{I}_R + (\mathbf{M}_k^{(s)})_{\langle N-1\rangle}^\top(\mathbf{M}_k^{(s)})_{\langle N-1\rangle})$.

TABLE I
THE COMPLEXITIES PER ITERATION OF THE METHODS

Method	Vanilla version	Accelerated version
HaLRTC [21]	$O(NI^{2(N-1)})$	—
TNN [25]	$O(I^{2(N+1)})$	—
DTNN [46]	$O((d^2+d)I^N + dI^3)$	—
TRLRF [57]	$O(I^N R^2 + I^{N-1}R^4)$	$O(IJ^{N-1}R^2 + J^{N-1}R^4)$
TMacTT [32]	$O(3I^N R)$	$O(3I^k J^{N-k} R)$
HTNN-FFT [26]	$O(I^N \log(I^{N-2}) + I^{N+1})$	—
TW-TC [31]	$O(N\sum_{l=2}^{N-1} I^l R^{l+3} + NI^N R^3 + I^N R^{2N})$	$O(N\sum_{l=2}^{N-1} J^l R^{l+3} + NIJ^{N-1}R^3 + I^N R^{2N})$
FCTN-TC [47]	$O(N\sum_{l=2}^{N} I^l R^{lN-l^2+l-1} + NI^{N-1}R^{2N-2} + NR^{3N-3})$	$O(N\sum_{l=2}^{N} J^l R^{lN-l^2+l-1} + NJ^{N-1}R^{2N-2} + NR^{3N-3})$
FCTN-GS	$O(N\sum_{l=2}^{N} I^l R^{lN-l^2+l-1} + NI^{N-1}R^{2N-2} + NR^{3N-3})$	$O(N\sum_{l=2}^{N} J^l R^{lN-l^2+l-1} + NJ^{N-1}R^{2N-2} + NR^{3N-3})$

Tensor rank values are all set to be R with all data dimensions equaling to I. The sampled dimensions are set to be J.

are also bounded. Let $a \geq 0$ be the maximum of $\|\mathcal{G}_k^{(s+1)}\|_F$ for $k = 1, 2, \ldots, N$, we have

$$\|\mathcal{X}^{(s+1)}\|_F \leq \frac{(a^N + \rho_2 d)}{1+\beta+\rho_2} + \frac{\beta\|\mathcal{B}\|_F}{1+\beta+\rho_2}. \tag{19}$$

Therefore, $\mathcal{X}^{(s+1)}$ is also bounded and **condition i) holds**.

Second, it can be seen that $h(\{\mathcal{G}_k\}_{k=1}^N, \mathcal{X})$ and $\mathfrak{S}(\{\mathcal{G}_k\}_{k=1}^N)$ are C^1 functions whose gradients are Lipschitz continuous and $\Phi(\mathcal{X})$, which is in a Frobenius norm manner, is obviously a proper lower semi-continuous function. Therefore, as the sum of semi-algebraic functions is still semi-algebraic, **condition ii) holds**.

Third, since the semi-algebraic real-valued function satisfies the K-Łproperty [58], we have that the semi-algebraic real-valued function $f(\{\mathcal{G}_k\}_{k=1}^N, \mathcal{X})$ satisfies the K-Łproperty at each point. Therefore, **condition iii) is satisfied**.

The verification of condition iv) needs two additional lemmas as follows.

Lemma 1 (Sufficient Decrease): Denote that $\rho_1, \rho_2 > 0$ and let $\{\mathcal{G}_k^{(s)}\}_{k=1}^N, \mathcal{X}^{(s)}\}_{s\in\mathbb{N}}$ be generated by Algorithm 1. The following inequalities hold

$$\begin{aligned}
&f(\mathcal{G}_{1:k}^{(s+1)}, \mathcal{G}_{k+1:N}^{(s)}, \mathcal{X}^{(s)}) + \rho_1\|\mathcal{G}_k^{(s+1)} - \mathcal{G}_k^{(s)}\|_F^2 \\
&\leq f(\mathcal{G}_{1:k-1}^{(s+1)}, \mathcal{G}_{k:N}^{(s)}, \mathcal{X}^{(s)}), k = 1, 2, \ldots, N, \\
&f(\{\mathcal{G}_k^{(s+1)}\}_{k=1}^N, \mathcal{X}^{(s+1)}) + \rho_2\|\mathcal{X}^{(s+1)} - \mathcal{X}^{(s)}\|_F^2 \\
&\leq f(\{\mathcal{G}_k^{(s+1)}\}_{k=1}^N, \mathcal{X}^{(s)}).
\end{aligned}$$

Proof: Since the updating of $\mathcal{G}_k^{(s+1)}$ satisfies the first order optimal condition, for $k = 1, 2, \ldots, N$ we have

$$\begin{aligned}
&f(\mathcal{G}_{1:k}^{(s+1)}, \mathcal{G}_{k+1:N}^{(s)}, \mathcal{X}^{(s)}) + \rho_1\|\mathcal{G}_k^{(s+1)} - \mathcal{G}_k^{(s)}\|_F^2 \\
&\leq f(\mathcal{G}_{1:k-1}^{(s+1)}, \mathcal{G}_{k:N}^{(s)}, \mathcal{X}^{(s)}) + \rho_1\|\mathcal{G}_k^{(s)} - \mathcal{G}_k^{(s)}\|_F^2 \\
&= f(\mathcal{G}_{1:k-1}^{(s+1)}, \mathcal{G}_{k:N}^{(s)}, \mathcal{X}^{(s)}).
\end{aligned}$$

Similarly, since the updating of $\mathcal{X}$ is element-wisely optimal, we have

$$\begin{aligned}
&f(\{\mathcal{G}_k^{(s+1)}\}_{k=1}^N, \mathcal{X}^{(s+1)}) + \rho_2\|\mathcal{X}^{(s+1)} - \mathcal{X}^{(s)}\|_F^2 \\
&\leq f(\{\mathcal{G}_k^{(s+1)}\}_{k=1}^N, \mathcal{X}^{(s)}).
\end{aligned}$$

□

Lemma 2 (Relative Error): Let ρ_1 and ρ_2 be positive constants, and suppose that the sequence obtained by the Algorithm 1, denoted as $\{\{\mathcal{G}_k^{(s+1)}\}_{k=1}^N, \mathcal{X}^{(s)}\}_{s\in\mathbb{N}}$, is under consideration. There exist $\mathcal{A}_1^{(s+1)} \in \partial_{\mathcal{G}_k}\mathfrak{S}(\mathcal{G}_k^{(s+1)})$ and $\mathcal{A}_2^{(s+1)} \in \partial_{\mathcal{X}}\Phi(\mathcal{X}^{(s+1)})$, which satisfy

$$\begin{aligned}
&\|\mathcal{A}_1^{(s+1)} + \nabla_{\mathcal{G}_k} h(\mathcal{G}_{1:k}^{(s+1)}, \mathcal{G}_{k+1:N}^{(s)}, \mathcal{X}^{(s)})\|_F \\
&\qquad \leq \rho_1\|\mathcal{G}_k^{(s+1)} - \mathcal{G}_k^{(s)}\|_F, k = 1, 2, \ldots, N, \\
&\|\mathcal{A}_2^{(s+1)} + \nabla_{\mathcal{X}} h(\{\mathcal{G}_k^{(s+1)}\}_{k=1}^N, \mathcal{X}^{(s+1)})\|_F \\
&\qquad \leq \rho_2\|\mathcal{X}^{(s+1)} - \mathcal{X}^{(s)}\|_F,
\end{aligned}$$

where $h(\{\mathcal{G}_k^{(s)}\}_{k=1}^N, \mathcal{X}) = \|\mathcal{X} - \mathfrak{F}(\{\mathcal{G}_k^{(s)}\}_{k=1}^N)\|_F^2$.

Proof: As our updating of each sub-problem satisfies the first order optimal condition, we have

$$\begin{aligned}
0 &\in \nabla_{\mathcal{G}_k} h(\mathcal{G}_{1:k}^{(s+1)}, \mathcal{G}_{k+1:N}^{(s)}, \mathcal{X}^{(s)}) + \partial_{\mathcal{G}_k}\mathfrak{S}(\mathcal{G}_k) \\
&\quad + \rho_1(\mathcal{G}_k^{(s+1)} - \mathcal{G}_k^{(s)}), \\
0 &\in \nabla_{\mathcal{X}} h(\{\mathcal{G}_k^{(s+1)}\}_{k=1}^N, \mathcal{X}^{(s)}) + \partial_{\mathcal{X}}\Phi(\mathcal{X}) \\
&\quad + \rho_2(\mathcal{X}^{(s+1)} - \mathcal{X}^{(s)}).
\end{aligned}$$

Then, there exist

$$\begin{aligned}
\mathcal{A}_1^{(s+1)} &= -\nabla_{\mathcal{G}_k} h(\mathcal{G}_{1:k}^{(s+1)}, \mathcal{G}_{k+1:N}^{(s)}, \mathcal{X}^{(s)}) - \rho_1(\mathcal{G}_k^{(s+1)} - \mathcal{G}_k^{(s)}) \\
&\in \partial_{\mathcal{G}_k}\mathfrak{S}(\mathcal{G}_k^{(s+1)}) \\
\mathcal{A}_2^{(s+1)} &= -\nabla_{\mathcal{X}} h(\{\mathcal{G}_k^{(s+1)}\}_{k=1}^N, \mathcal{X}^{(s)}) - \rho_2(\mathcal{X}^{(s+1)} - \mathcal{X}^{(s)}) \\
&\in \partial_{\mathcal{X}}\Phi(\mathcal{X}).
\end{aligned}$$

□

Then, combining the above results, we can establish the proof of Theorem 2.

Proof of Theorem 2: From (18) and (19), we can see that the sequence generated by Algorithm 1 is bounded. Meanwhile, $f(\{\mathcal{G}_k\}_{k=1}^N, \mathcal{X})$ is a semi-algebraic real-valued function satisfying the K-Łproperty at each point. Then, from Lemma 1, the value of $f(\{\mathcal{G}_k\}_{k=1}^N, \mathcal{X})$ monotonically decreases, and we can obtain the relative error condition the sequence generated by Algorithm 1 from Lemma 2. Therefore, this proof conforms Theorem 6.2 in [58] under the above conditions. □

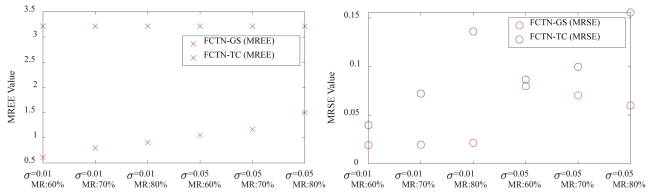

Fig. 2. The MREE and MRSE values of the results by our FCTN-GS and FCTN-TC with different missing rates (MRs) and noise levels. **Best viewed in** $\times 2$ **sized color pdf file.**

VI. Experiment

In this section, we compare our methods with many state-of-the-art methods on synthetic and real-world data sets. All the experiments are implemented on Matlab (R2020b) with an Intel(R) Xeon(R) Gold-5218 2.20 GHz CPU and 32 GB memory.

A. Synthetic Data

In this part, we conduct experiments on synthetic data to verify the FCTN-rank estimation ability of our method. The synthetic groundtruth (GT) tensor $\mathcal{X}^{\mathrm{GT}} = \mathfrak{F}(\{\mathcal{G}_k\}_{k=1}^N)$ with a given FCTN-rank, denoted by $\mathbf{r}^{\mathrm{GT}}$, is obtained by i) entries in each $\mathcal{G}_k$ being randomly generated following the uniform distribution in [0,1], and ii) contracting all $\mathcal{G}_k$s to obtain $\mathcal{X}$ and normalizing entries in $\mathcal{X}$ to the interval of [0,1]. Then, the observed data is generated by uniformly dropping 60% to 80% entries at random and adding a zero-mean Gaussian noise with standard deviations $\sigma = 0.01$ or 0.05. Finally, we run our structural sparsity regularized FCTN method in Algorithm 1, denoted as "FCTN-GS", and the method in [37][4], denoted as "FCTN-TC", to obtain the recovered tensor and estimated FCTN-rank. Denoting the estimated FCTN-rank by $\mathbf{r}^{\mathrm{Est}}$ and the recovered tensor by $\mathcal{X}^{\mathrm{Est}}$. Then we employ the rank estimation error (REE) [51] to measure the rank estimation ability. It is defined as

$$\mathrm{REE} = \left(\frac{2\sum_{k=1}^{\frac{N(N-1)}{2}} \left(\mathbf{r}^{\mathrm{GT}}(k) - \mathbf{r}^{\mathrm{Est}}(k)\right)^2}{N(N-1)} \right)^{1/2}. \tag{20}$$

Meanwhile, we employ the relative squared error (RSE) to measure the reconstruction quality. RSE is calculated by

$$\mathrm{RSE} = \frac{\|\mathcal{X}^{\mathrm{GT}} - \mathcal{X}^{\mathrm{Est}}\|_F^2}{\|\mathcal{X}^{\mathrm{GT}}\|_F^2} \tag{21}$$

For each degradation case, we randomly generate 10 tensors of the size $35\times20 \times 25\times20$ with different FCTN-ranks.

Fig. 2 exhibits the mean values of REE and RSE (denoted as MREE and MRSE, respectively). The MREE values for FTCN-TC are all the same. The main reason for this is that FTCN-TC adopts a rank-increasing strategy; however, this process of rank increase is imprecise, and as a result, its rank often increases to the maximum value we have set. From the left part, we can see that the FCTN-rank estimated by our FCTN-GS is much more accurate than that by FCTN-TC. Meanwhile, the MRSE values show that our recovery results are also better in all cases, illustrating that the structural sparsity regularization can better enhance the low-rankness with robustness.

[4]In [37], the authors only consider noise-free tensor completion. We remark here that it is easy to modify their methods for the noisy tensor completion problem by replacing the indicator function in their objective function with the weighted Frobenius norm. Meanwhile, the projection step in [37] should be correspondingly changed to be in the same manner as (13) and (14). This also applies to other tensor completion methods.

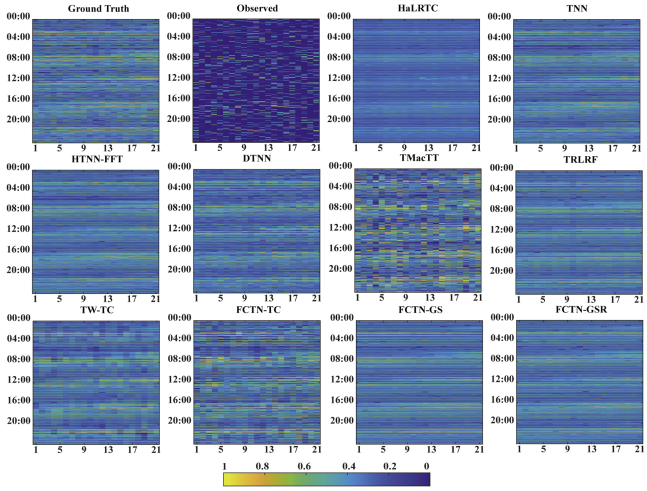

Fig. 3. The reconstructed slices (made up of the time interval and segment modes) on the traffic speed dataset with MR=80%. **Best viewed in** $\times 2$ **sized color pdf file.**

TABLE II
The RMSE Values and MAE Values (In Percentage) of Results by All Methods on the Traffic Speed Data

Method	MR=80%		MR=70%		MR=60%		Time
	RMSE	MAE	RMSE	MAE	RMSE	MAE	(s)
Observed	0.329	0.241	0.309	0.212	0.287	0.183	0
HaLRTC	0.098	0.069	0.075	0.052	0.062	0.042	0
TNN	0.064	0.042	0.055	0.034	0.048	0.027	1
DTNN	0.084	0.055	0.096	0.062	0.081	0.049	13
TRLRF	0.067	0.041	0.046	0.028	0.041	0.023	4
TMacTT	0.091	0.047	0.068	0.033	0.052	0.026	3
HTNN-FFT	0.055	0.036	0.049	0.030	0.043	0.024	1
TW-TC	0.064	0.037	0.051	0.029	0.039	0.022	44
FCTN-TC	0.061	0.037	0.051	0.029	0.044	0.023	19
FCTN-GS	0.046	0.029	0.041	0.025	0.037	0.021	12
FCTN-GSR	0.052	0.033	0.046	0.028	0.041	0.024	5

The red and blue colors stand for the best and second best values, respectively.

B. Real Data

In this section, we test our FCTN-GS and its randomized block sampling accelerated version (denoted as FCTN-GSR, the block sampling ratio is 70%) on real data. Compared methods consist of i) HaLRTC [21], ii) TNN [25], iii) HTNN-FFT [26], iv) TMacTT [32], v) TRLRF [57], vi) DTNN [46], vii) FCTN-TC [37], and viii) TW-TC [38]. Then we compute the peak signal-to-noise ratio (PSNR) and the structural similarity (SSIM) [59] values of the results by different methods on the face data and video data. The root mean square error (RMSE) and the mean absolute error (MAE) are employed to quantitatively measure the reconstruction quality of the traffic data, and MAE

TABLE III
THE PSNR VALUES AND SSIM VALUES OF RESULTS BY ALL METHODS ON THE HUMAN FACE DATA

Method	$\sigma = 0.05$						$\sigma = 0.07$						Time
	MR=90%		MR=80%		MR=70%		MR=90%		MR=80%		MR=70%		(s)
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	
Observed	10.643	0.015	11.094	0.040	11.618	0.071	10.597	0.014	11.019	0.037	11.512	0.065	0
HaLRTC	16.737	0.249	19.673	0.405	21.592	0.491	16.485	0.232	19.246	0.371	20.932	0.444	11
TNN	20.312	0.378	21.984	0.458	23.016	0.501	19.860	0.350	21.117	0.413	21.783	0.445	31
DTNN	20.463	0.372	20.264	0.394	21.715	0.465	19.345	0.319	18.288	0.312	19.368	0.373	1111
TRLRF	21.037	0.417	22.761	0.504	23.257	0.526	21.947	0.461	22.629	0.492	23.075	0.511	499
TMacTT	18.520	0.308	23.600	0.531	25.333	0.587	16.853	0.233	21.799	0.460	24.136	0.543	92
HTNN-FFT	20.799	0.400	22.156	0.464	23.171	0.504	20.281	0.367	21.283	0.416	21.990	0.448	40
TW-TC	22.968	0.511	23.429	0.528	23.945	0.550	22.831	0.501	23.298	0.518	23.706	0.534	795
FCTN-TC	23.786	0.538	24.852	0.579	25.352	0.595	23.363	0.519	24.478	0.561	24.956	0.576	324
FCTN-GS	24.179	0.562	25.102	0.595	25.632	0.612	23.860	0.543	24.753	0.576	25.215	0.589	369
FCTN-GSR	24.210	0.569	25.140	0.601	25.636	0.613	23.954	0.552	24.808	0.580	25.222	0.589	219

The red and blue colors stand for the best and second best values, respectively.

and RMSE are respectively defined as

$$\text{MAE} = \frac{\sum_{i_1,\ldots,i_N} \left| \mathcal{X}^{\text{GT}}_{i_1,\ldots,i_N} - \mathcal{X}^{\text{Est}}_{i_1,\ldots,i_N} \right|}{\prod_{k=1}^{k=N} I_k} \tag{22}$$

$$\text{RMSE} = \left(\frac{\|\mathcal{X}^{\text{GT}} - \mathcal{X}^{\text{Est}}\|_F^2}{\prod_{k=1}^{k=N} I_k} \right)^{1/2} \tag{23}$$

Higher values of PSNR and SSIM and lower values of MAE and RMSE indicate better recovery performances.

1) Traffic Data: In this part, we conduct experiments on the traffic speed dataset[5] collected from 21 road segments over a period of 30 days. The data point is sampled at 5-minute intervals. The size of traffic data is $12 \times 24 \times 21 \times 30$ (minute $\times$ hour $\times$ segment$\times$ day). The observed data is obtained by random sampling from noiseless[6] elements with MR $\in \{80\%, 70\%, 60\%\}$. Meanwhile, considering that the sensor failure could occur across time intervals, we also set some element-wise missing to simulate this situation. In Fig. 3, we show the reconstructed results on the 4th day. Table II presents the RMSE and MAE values. As observed, the results of our FCTN-GS and FCTN-GSR are closer to the ground truth with better index values.

2) Face Data: We employ the Extended Yale B face data set,[7] which includes 38 subjects under 64 illumination conditions. The size of the face image is 48×42. Thus, the size of face data is $48 \times 42 \times 64 \times 38$ (height $\times$ width $\times$ illumination $\times$ subjects). We test three MRs (90%, 80%, and 70%) at two noise levels $\sigma = 0.05$ and 0.07.

Table III shows the PSNR and SSIM values of the results by different methods. The correlations between different modes of the face data tensor are heterogeneous. The performances of the tensor network decomposition based methods are generally better. This illustrates the outstanding ability of the tensor network to characterize different types of inner correlations. Besides, Fig. 4 shows the reconstructed results of human faces. Although the quantitative metrics of FCTN-TC's results are better than those of the TW-TC method, TW-TC obtains a more clear visual result. We can easily find that TRLRF, TW-TC, FCTN-GS, and FCTN-GSR have fewer noisy pixels in their results.

3) Color and Hyperspectral Videos: In this part, three color videos[8] ("boat", "container", and "rhino"), and one hyperspectral video,[9] denoted as "HSV", are selected. The size of three color videos is $144 \times 176 \times 3 \times 50$ (height $\times$ width $\times$ channel $\times$ frame) and the size of the hyperspectral video is $60 \times 60 \times 20 \times 20$ (height$\times$ width $\times$ spectrum $\times$ frame). For each video, we test three missing rates (MRs): 80%, 90%, and 95%, and the zero-mean Gaussian noise with standard deviations $\sigma = 0.05$ and 0.07. Table IV shows the quantitative metrics of the results by different methods together with their running time in seconds (s). We can see that our FCTN-GS can largely outperform FCTN-TC in the MR cases and the superiority is more obvious for the high standard deviation value. This shows that our structural sparsity regularization enhances the robustness and indeed helps to exploit the accurate low-FCTN-rank structure with its accurate determination of the FCTN-rank values. The acceleration in FCTN-GSR is obvious however its performance is not quite stable. It occasionally obtains better results that those by the FCTN-GS and the reason might be that the randomized sampling acts as an implicit regularization. The performances of compared methods would vary for different data and our FCTN-GS and FCTN-GSR almost make a clean sweep of the best and second best metrics. In Fig. 5, we exhibit the results by different methods on the color videos "boat" and "containe" and the HSV. We can see that the completion from noisy and partial observations is difficult. Results by TW-TC, TRLRF, FCTN-GS, and FCTN-GSR are more clear and visually better than those by other methods. Their additional constraints or regularization help them to be robust to the noise while we can see more obvious noise patterns in the results of other methods.

[5]The data is available at https://pems.dot.ca.gov.

[6]We remark that we consider the noiseless completion for traffic data, as many compared methods are originally designed for this case. We think this would gravitate towards greater equity of the comparison and thus better show the superiority of our method.

[7]The data is available at http://vision.ucsd.edu/content/yale-face-database.

[8]The data is available at http://trace.eas.asu.edu/yuv.

[9]The data is available at https://openremotesensing.net.

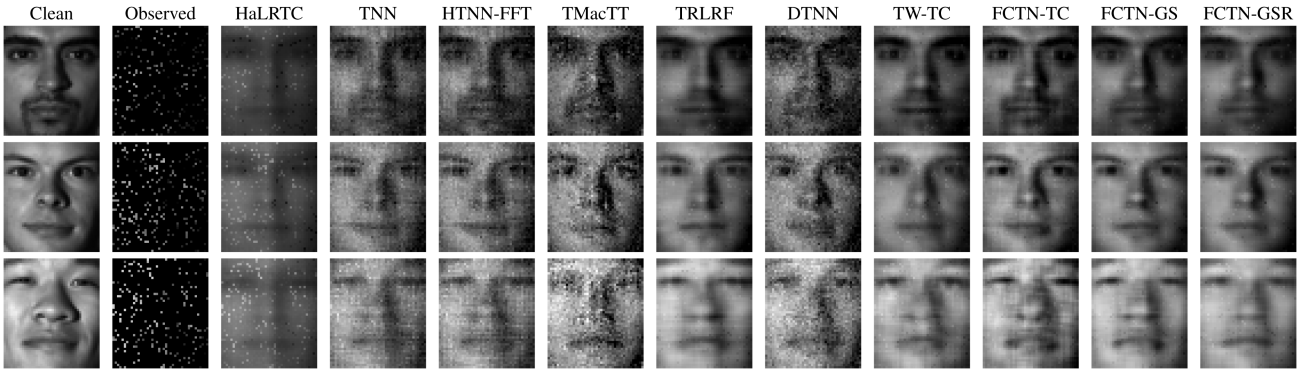

Fig. 4. The reconstructed results of the face data, which shows the 9-th, 13-th and 15-th face at the 15-th illumination with MR=90%, $\sigma = 0.05$.

TABLE IV
THE PSNR VALUES AND SSIM VALUES OF RESULTS BY ALL METHODS ON CVS AND THE HSV

Data	Method	$\sigma = 0.05$						$\sigma = 0.07$						Time (s)
		MR=95%		MR=90%		MR=80%		MR=95%		MR=90%		MR=80%		
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	
boat	Observed	5.876	0.002	6.108	0.006	6.613	0.016	5.874	0.002	6.104	0.006	6.603	0.015	0
	HaLRTC	17.957	0.154	21.210	0.253	24.138	0.367	17.390	0.143	20.599	0.231	23.255	0.323	14
	TNN	23.573	0.300	26.449	0.412	27.300	0.460	22.869	0.265	25.157	0.351	25.361	0.384	22
	DTNN	24.611	0.385	25.872	0.431	26.933	0.472	23.912	0.354	24.844	0.386	25.388	0.409	1188
	TRLRF	25.852	0.387	26.555	0.425	27.265	0.457	25.508	0.360	26.303	0.401	26.932	0.428	285
	TMacTT	26.196	0.376	26.968	0.418	27.601	0.454	24.998	0.326	26.686	0.399	27.361	0.433	25
	HTNN-FFT	25.654	0.369	26.708	0.422	27.107	0.453	24.670	0.317	25.195	0.356	25.108	0.376	32
	TW-TC	25.945	0.386	27.610	0.466	28.741	0.517	24.904	0.335	26.803	0.421	28.040	0.478	730
	FCTN-TC	24.728	0.343	27.367	0.460	29.252	0.537	23.360	0.286	25.943	0.397	28.029	0.485	150
	FCTN-GS	26.453	0.418	28.092	0.493	29.947	0.560	25.620	0.368	27.200	0.444	28.501	0.501	171
	FCTN-GSR	26.242	0.407	28.139	0.494	29.536	0.552	25.698	0.372	27.426	0.448	28.648	0.504	102
container	Observed	4.596	0.003	4.829	0.006	5.336	0.017	4.594	0.002	4.826	0.006	5.329	0.017	0
	HaLRTC	16.875	0.216	20.309	0.330	23.519	0.436	16.406	0.199	19.762	0.295	22.671	0.378	15
	TNN	23.085	0.362	27.581	0.464	28.570	0.479	22.383	0.325	25.888	0.410	26.186	0.421	22
	DTNN	25.565	0.476	27.355	0.515	28.637	0.524	24.664	0.430	25.944	0.461	26.534	0.464	1166
	TRLRF	24.882	0.431	26.417	0.489	27.141	0.506	24.682	0.407	26.154	0.461	26.802	0.473	298
	TMacTT	25.553	0.441	27.037	0.489	27.990	0.515	25.071	0.412	26.798	0.461	27.641	0.482	29
	HTNN-FFT	26.430	0.439	28.185	0.468	28.337	0.472	25.123	0.390	26.196	0.415	25.948	0.415	35
	TW-TC	27.055	0.482	29.408	0.545	30.715	0.569	25.843	0.431	28.360	0.499	29.732	0.525	796
	FCTN-TC	25.613	0.433	29.791	0.530	32.101	0.575	23.943	0.376	27.617	0.468	30.083	0.520	134
	FCTN-GS	27.300	0.497	30.437	0.565	32.465	0.588	26.272	0.446	29.409	0.512	30.707	0.531	162
	FCTN-GSR	26.657	0.494	29.101	0.545	31.469	0.584	26.382	0.458	28.730	0.509	30.310	0.537	107
rhino	Observed	8.292	0.006	8.523	0.016	9.023	0.041	8.288	0.006	8.515	0.015	9.006	0.040	0
	HaLRTC	13.459	0.090	16.201	0.202	19.064	0.394	13.256	0.086	15.972	0.190	18.662	0.359	15
	TNN	18.870	0.315	22.008	0.530	24.172	0.658	18.575	0.289	21.290	0.474	22.830	0.581	22
	DTNN	20.066	0.406	20.948	0.486	22.253	0.584	19.812	0.379	20.538	0.447	21.521	0.525	1160
	TRLRF	18.910	0.261	20.269	0.381	21.075	0.456	18.815	0.253	19.517	0.320	20.288	0.393	173
	TMacTT	16.070	0.199	18.445	0.402	24.289	0.691	16.016	0.175	15.404	0.240	20.878	0.537	68
	HTNN-FFT	20.814	0.431	22.719	0.573	24.293	0.662	20.296	0.386	21.784	0.507	22.824	0.580	33
	TW-TC	20.555	0.407	21.690	0.502	22.439	0.561	20.269	0.383	21.447	0.480	22.267	0.542	641
	FCTN-TC	20.482	0.424	22.412	0.561	23.624	0.638	19.860	0.380	21.983	0.530	23.258	0.613	152
	FCTN-GS	21.296	0.463	23.213	0.606	24.876	0.704	20.597	0.407	22.570	0.561	24.334	0.670	180
	FCTN-GSR	21.343	0.467	22.995	0.590	24.878	0.705	21.103	0.441	22.557	0.554	24.331	0.670	109
HSV	Observed	8.411	0.006	8.643	0.015	9.145	0.040	8.407	0.006	8.635	0.014	9.128	0.038	0
	HaLRTC	11.946	0.195	19.499	0.573	24.479	0.733	11.575	0.168	18.514	0.523	23.182	0.661	6
	TNN	25.039	0.666	28.910	0.739	29.115	0.727	24.013	0.617	26.859	0.671	26.548	0.649	9
	DTNN	29.688	0.773	30.026	0.758	28.910	0.719	27.805	0.713	27.653	0.691	26.404	0.646	171
	TRLRF	29.997	0.802	31.319	0.830	32.229	0.841	29.004	0.763	30.380	0.796	31.091	0.805	50
	TMacTT	30.841	0.789	33.636	0.847	34.433	0.859	25.635	0.625	29.186	0.728	31.116	0.777	14
	HTNN-FFT	28.304	0.735	29.083	0.734	28.818	0.717	26.632	0.675	26.850	0.664	26.293	0.639	13
	TW-TC	31.052	0.790	33.245	0.829	34.473	0.847	28.024	0.704	30.372	0.762	32.151	0.801	326
	FCTN-TC	25.805	0.631	28.902	0.716	31.682	0.784	23.196	0.541	25.897	0.629	28.615	0.707	164
	FCTN-GS	31.203	0.841	33.632	0.857	34.972	0.863	30.281	0.800	31.341	0.792	31.826	0.792	131
	FCTN-GSR	30.896	0.845	33.487	0.867	34.332	0.865	30.278	0.817	31.818	0.816	32.454	0.815	61

The red and blue colors stand for the best and second best values, respectively.

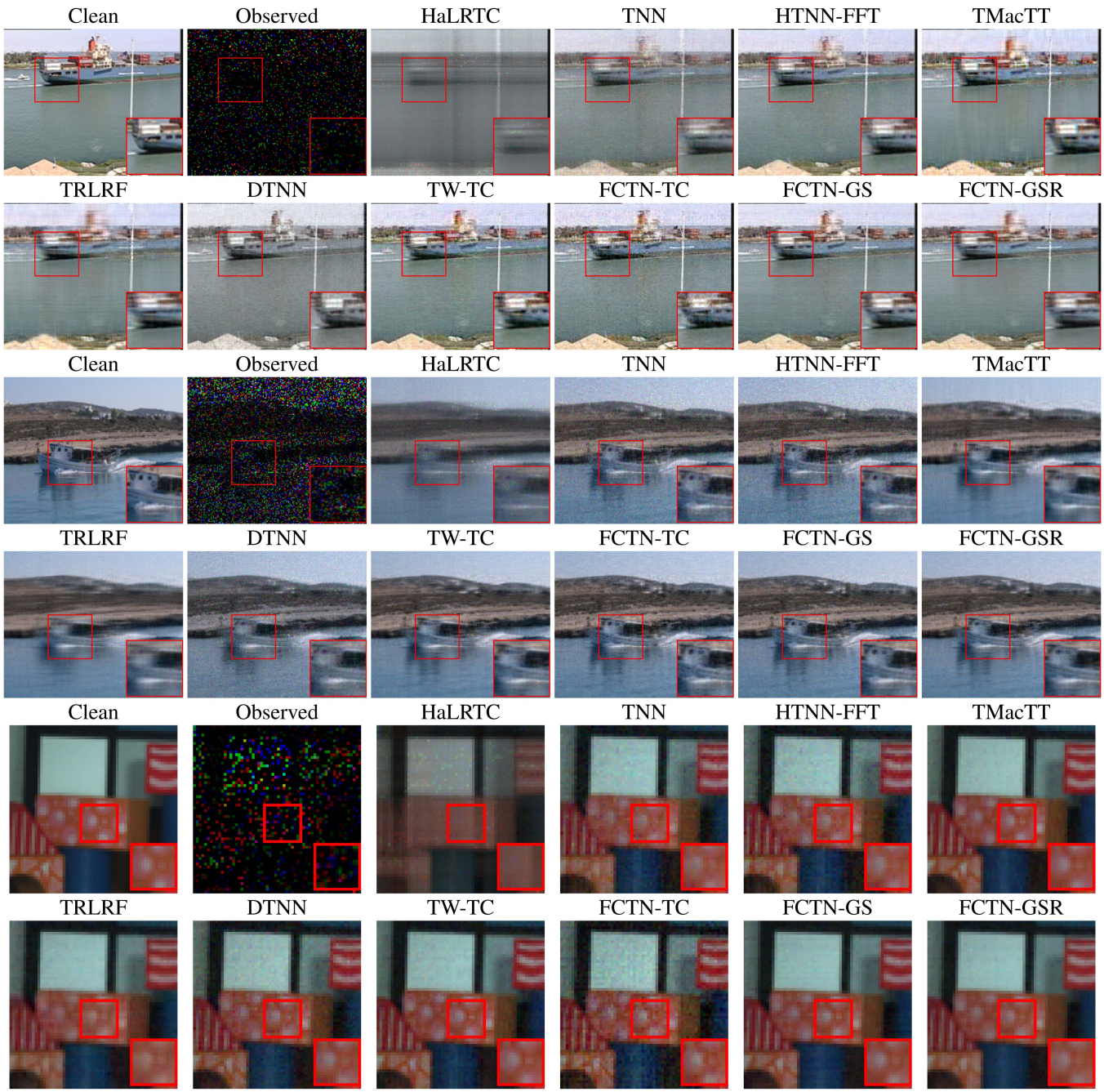

Fig. 5. Results on videos. The first two rows are reconstructed results of the 45-th frame of the color video "container" with MR=95%, $\sigma = 0.05$. The middle two rows are reconstructed results of the 35-th frame of the color video "boat" with MR=80%, $\sigma = 0.07$. The bottom two row are pseudo color images (R-16 G-8 B-4) of reconstructed results of the 15-th frame of the hyperspectral video "HSV" with MR=90% and $\sigma = 0.07$.

C. Discussions

In this part, we conduct more experiments to further discuss the settings that would affect the result.

a) Initial FCTN-rank values: In order to verify that our proposed method is less affected by the setting of the maximum values of the FCTN-rank R_{max}, we set $R_{max} \in \{[i, i, i, i, i]|i = 4, 5, 6, 7, 8\}$. Then, we compare our FCTN-GS with the FCTN-TC method on HSV with MR=90% and σ=0.05. The performance of FCTN-GS and FCTN-TC methods are presented in Fig. 6. In terms of PSNR values and SSIM values, it is easy to

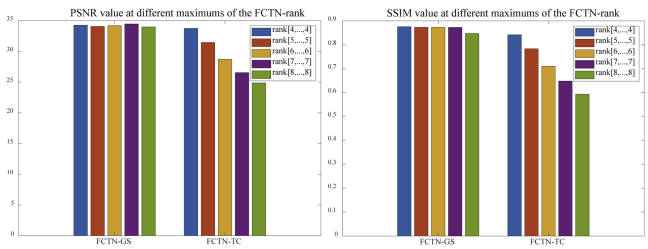

Fig. 6. The reconstruction performance of the FCTN-TC and FCTN-GS at different maximums of the FCTN-rank. **Best viewed in $\times 2$ sized color pdf file.**

TABLE V
THE RESULT OF THE RANDOMIZED BLOCK SAMPLING

Data	Method	$\sigma = 0.05$						$\sigma = 0.07$						Time
		MR=95%		MR=90%		MR=80%		MR=95%		MR=90%		MR=80%		
		PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	(s)
boat	Observed	5.88	0.00	6.11	0.01	6.61	0.02	5.87	0.00	6.10	0.01	6.60	0.02	0
	FCTN-TCR	25.37	0.37	27.60	0.47	28.25	0.49	23.92	0.31	26.26	0.41	26.97	0.44	**89**
	FCTN-TC	24.73	0.34	27.37	0.46	29.25	0.54	23.36	0.29	25.94	0.40	28.03	0.49	150
	TW-TCR	26.45	0.40	26.90	0.43	27.24	0.45	25.89	0.37	26.34	0.40	26.34	0.40	**144**
	TW-TC	25.95	0.39	27.61	0.47	28.74	0.52	24.90	0.34	26.80	0.42	28.04	0.49	730
face	Observed	10.64	0.02	11.09	0.04	11.62	0.07	10.60	0.01	11.02	0.04	11.51	0.07	0
	FCTN-TCR	22.50	0.49	22.60	0.49	22.81	0.50	22.30	0.48	22.24	0.47	22.37	0.47	**71**
	FCTN-TC	23.80	0.54	24.92	0.58	25.45	0.60	23.39	0.52	24.53	0.56	25.04	0.58	414
	TW-TCR	20.61	0.41	21.10	0.43	21.47	0.45	20.51	0.40	20.86	0.41	21.11	0.42	**170**
	TW-TC	22.95	0.51	23.50	0.53	23.89	0.55	22.81	0.50	23.36	0.52	23.66	0.53	1182

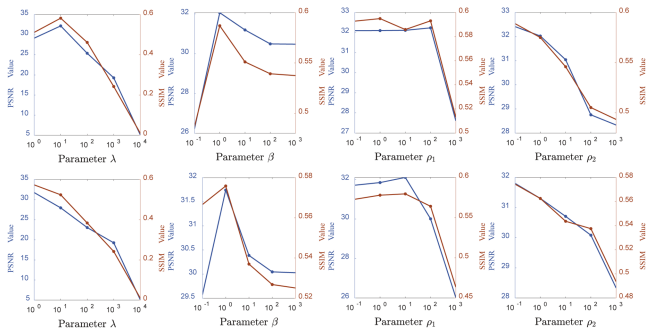

Fig. 7. The effect of the parameters $\rho_1, \rho_2, \beta, \lambda$ on the FCTN-GS algorithm and FCTN-GSR algorithm. The first row is the result of the FCTN-GS algorithm, and the second row is the result of the FCTN-GSR algorithm.

see that the results reconstructed by FCTN-GS are more stable than FCTN-TC at different maximums of the FCTN-rank.

b) Parameters: In our method, the main parameters are λ, ρ_1, ρ_2 and β. To test the effects from different values of them, we conduct 8 experiments on the color video "container" with MR $= 80\%$ and $\sigma = 0.05$. The PSNR values and SSIM values with respect to different parameters are represented in Fig. 7. When testing one parameter, other parameters are fixed as default values. The default values are set as follows: $\rho_1 = 0.1, \rho_2 = 0.1$, $\lambda = 10^1$, $\beta = 1$. From Fig. 7, we can see that the performance of FCTN-GS and FCTN-GSR is stable at $\lambda = \{10^0, 10^1\}$, $\rho_1 = \{10^{-1}, 10^0\}$ and $\rho_2 = \{10^{-1}, 10^0\}$ respectively. As observed, if λ or ρ_2 continues to increase, it will bring huge offsets to decrease the performance. Meanwhile, we can see the performance of our method is more sensitive to β and our method could obtain satisfactory results with a wide range of ρ_1.

c) The result of the randomized block sampling strategy: In recent years, some complicated tension network decompositions have emerged. Therefore, it is necessary to reduce the computation burden among them. TW decomposition and FCTN decomposition both are complicated, which is due to their factor tensor owning higher order. It can be naturally imagined that reducing the size of the dimensions of the core tensor can reduce computational burden.

To highlight the effectiveness of this approach, we compare TW-TCR and FCTN-TCR with their original counterparts on two datasets. Results are shown in Table V, demonstrating the improvements in speed and efficiency.

TABLE VI
THE PARAMETERS OF PROPOSED METHODS FOR REPRODUCIBILITY

Method	Data	Parameters
FCTN-GS	Color Video	$R_{\max} = [15, 3, 5, 3, 5, 3]$ $\lambda = 5$ $\beta = 1$ $\rho_1 = 0.01$ $\rho_2 = 0.001$
	Hyperspectral Video	$R_{\max} = [6, 6, 6, 6, 6, 6]$ $\lambda = 1.5$ $\beta = 1$ $\rho_1 = 0.01$ $\rho_2 = 0.001$
	Face Data	$R_{\max} = [6, 6, 6, 6, 6, 6]$ $\lambda = 1.5$ $\beta = 1$ $\rho_1 = 0.1$ $\rho_2 = 0.01$
	Traffic Data	$R_{\max} = [5, 5, 5, 5, 5, 5]$ $\lambda = 1$ $\beta = 10$ $\rho_1 = 0.1$ $\rho_2 = 0.01$
FCTN-GSR	Color Video	$R_{\max} = [15, 3, 5, 3, 5, 3]$ $\lambda = 2.5$ $\beta = 1$ $\rho_1 = 0.01$ $\rho_2 = 0.001$
	Hyperspectral Video	$R_{\max} = [6, 6, 6, 6, 6, 6]$ $\lambda = 0.5$ $\beta = 1$ $\rho_1 = 0.01$ $\rho_2 = 0.001$
	Face Data	$R_{\max} = [6, 6, 6, 6, 6, 6]$ $\lambda = 1$ $\beta = 1$ $\rho_1 = 0.1$ $\rho_2 = 0.01$
	Traffic Data	$R_{\max} = [5, 5, 5, 5, 5, 5]$ $\lambda = 0.5$ $\beta = 10$ $\rho_1 = 0.1$ $\rho_2 = 0.01$

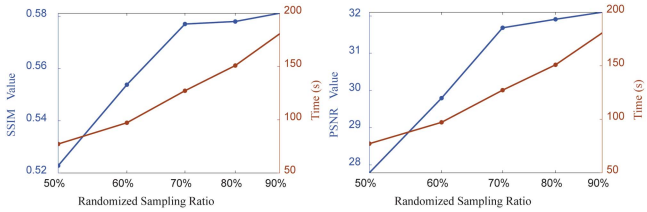

Fig. 8. The PSNR values and SSIM values related to the sampling ratios of the randomized block sampling.

d) Parameters settings for FCTN-GS and FCTN-GSR: For the experimental results presented in this paper, we provide the following table of parameters as a basis for reproducibility. From Table VI, we can summarize that the parameter requiring the most adjustment across different datasets is λ, which needs to be set smaller in FCTN-GSR compared to FCTN-GS. Otherwise, this paper aims to set $R_{\max}$ as close as possible to its estimation based on the results obtained from the experiments.

e) The effect from randomized block sampling: From real data experiments, we can see the acceleration brought in from the randomized block sampling. In this part, we test our FCTN-GSR also on the color video "container" with MR $= 80\%$ and $\sigma = 0.05$. We vary the sampling ratio of the randomized block sampling strategy from 50% to 90%. Then, the PSNR and

SSIM values are reported in Fig. 8. We also exhibit the running time. We can see that the performance is generally better as the sampling ratio grows. However, the running time would also increase.

VII. Conclusion

In this work, we propose a noisy tensor completion model based on the structural sparsity regularized FCTN decomposition. The structural sparsity regularization can lead to near-zero groups and thus the FCTN-rank can be adaptively adjusted by pruning them. We design a proximal alternating minimization algorithm and theoretically establish its global convergence. To further accelerate the proposed method, we customized a randomized block sampling strategy. Experiments on synthetic data show that our structural sparsity regularization helps to automatically and accurately determine complicated FCTN-rank in the presence of pixel missing and Gaussian noise corruptions. We also conduct abundant experiments on different types of real-world tensor data and results illustrate that our methods outperform compared state-of-the-art methods. It is imperative to emphasize that our randomized block sampling strategy, coupled with structural sparsity regularization, possesses the potential for broader applicability across various large-scale TNs. We have undertaken some experiments to substantiate the viability of this assertion.

Acknowledgment

The authors would like to thank the authors of [21], [25], [26], [32], [37], [38], [57] for their generous sharing of their codes or data.

References

[1] F. L. Hitchcock, "Multiple invariants and generalized rank of a p-way matrix or tensor," *J. Math. Phys.*, vol. 7, no. 1/4, pp. 39–79, 1928.
[2] F. L. Hitchcock, "The expression of a tensor or a polyadic as a sum of products," *J. Math. Phys.*, vol. 6, no. 1-4, pp. 164–189, 1927.
[3] R. B. Cattell, ""Parallel proportional profile," and other principles for determining the choice of factors by rotation," *Psychometrika*, vol. 9, no. 4, pp. 267–283, 1944.
[4] R. B. Cattell, "The three basic factor-analytic research designs—Their interrelations and derivatives," *Psychol. Bull.*, vol. 49, no. 5, 1952, Art. no. 499.
[5] N. D. Sidiropoulos, R. Bro, and G. B. Giannakis, "Parallel factor analysis in sensor array processing," *IEEE Trans. Signal Process.*, vol. 48, no. 8, pp. 2377–2388, Aug. 2000.
[6] A. Cichocki et al., "Tensor decompositions for signal processing applications: From two-way to multiway component analysis," *IEEE Signal Process. Mag.*, vol. 32, no. 2, pp. 145–163, Mar. 2015.
[7] A. Shashua and A. Levin, "Linear image coding for regression and classification using the tensor-rank principle," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 2001, pp. I–I.
[8] N. Hao, M. E. Kilmer, K. Braman, and R. C. Hoover, "Facial recognition using tensor-tensor decompositions," *SIAM J. Imag. Sci.*, vol. 6, no. 1, pp. 437–463, 2013.
[9] N. D. Sidiropoulos, L. De Lathauwer, X. Fu, K. Huang, E. E. Papalexakis, and C. Faloutsos, "Tensor decomposition for signal processing and machine learning," *IEEE Trans. Signal Process.*, vol. 65, no. 13, pp. 3551–3582, Jul. 2017.
[10] T.-X. Jiang, M. K. Ng, J. Pan, and G.-J. Song, "Nonnegative low rank tensor approximations with multidimensional image applications," *Numerische Mathematik*, vol. 153, no. 1, pp. 141–170, 2023.
[11] Y. Liu and M. K. Ng, "Deep neural network compression by Tucker decomposition with nonlinear response," *Knowl.-Based Syst.*, vol. 241, 2022, Art. no. 108171.
[12] C. F. Beckmann and S. M. Smith, "Tensorial extensions of independent component analysis for multisubject FMRI analysis," *Neuroimage*, vol. 25, no. 1, pp. 294–311, 2005.
[13] F. Cong, Q.-H. Lin, L.-D. Kuang, X.-F. Gong, P. Astikainen, and T. Ristaniemi, "Tensor decomposition of EEG signals: A brief review," *J. Neurosci. Methods*, vol. 248, pp. 59–69, 2015.
[14] L. R. Tucker, "Some mathematical notes on three-mode factor analysis," *Psychometrika*, vol. 31, no. 3, pp. 279–311, 1966.
[15] I. V. Oseledets, D. Savostianov, and E. E. Tyrtyshnikov, "Tucker dimensionality reduction of three-dimensional arrays in linear time," *SIAM J. Matrix Anal. Appl.*, vol. 30, no. 3, pp. 939–956, 2008.
[16] 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, no. 3, pp. 283–319, 1970.
[17] J. Douglas Carroll, S. Pruzansky, and J. B. Kruskal, "CANDELINC: A general approach to multidimensional analysis of many-way arrays with linear constraints on parameters," *Psychometrika*, vol. 45, no. 1, pp. 3–24, 1980.
[18] K. Braman, "Third-order tensors as linear operators on a space of matrices," *Linear Algebra Appl.*, vol. 433, no. 7, pp. 1241–1253, 2010.
[19] M. E. Kilmer and C. D. Martin, "Factorization strategies for third-order tensors," *Linear Algebra Appl.*, vol. 435, no. 3, pp. 641–658, 2011.
[20] M. E. Kilmer, K. Braman, N. Hao, and R. C. Hoover, "Third-order tensors as operators on matrices: A theoretical and computational framework with applications in imaging," *SIAM J. Matrix Anal. Appl.*, vol. 34, no. 1, pp. 148–172, 2013.
[21] 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, no. 1, pp. 208–220, Jan. 2013.
[22] Q. Zhao, L. Zhang, and A. Cichocki, "Bayesian CP factorization of incomplete tensors with automatic rank determination," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 9, pp. 1751–1763, Sep. 2015.
[23] S. Gandy, B. Recht, and I. Yamada, "Tensor completion and low-N-rank tensor recovery via convex optimization," *Inverse Problems*, vol. 27, no. 2, pp. 0 25010, 2011.
[24] C. Lu, J. Feng, Y. Chen, W. Liu, Z. Lin, and S. Yan, "Tensor robust principal component analysis with a new tensor nuclear norm," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 4, pp. 925–938, Apr. 2020.
[25] Z. Zhang and S. Aeron, "Exact tensor completion using t-SVD," *IEEE Trans. Signal Process.*, vol. 65, no. 6, pp. 1511–1526, Jun. 2017.
[26] W. Qin, H. Wang, F. Zhang, J. Wang, X. Luo, and T. Huang, "Low-rank high-order tensor completion with applications in visual data," *IEEE Trans. Image Process.*, vol. 31, pp. 2433–2448, 2022.
[27] A. Cichocki, N. Lee, I. Oseledets, A.-H. Phan, Q. Zhao, and D. P. Mandic, "Tensor networks for dimensionality reduction and large-scale optimization: Part 1 low-rank tensor decompositions," *Found. Trends Mach. Learn.*, vol. 9, no. 4/5, pp. 249–429, 2016.
[28] A. Cichocki et al., "Tensor networks for dimensionality reduction and large-scale optimization: Part 2 applications and future perspectives," *Found. Trends Mach. Learn.*, vol. 9, no. 6, pp. 431–673, 2017.
[29] I. V. Oseledets, "Tensor-train decomposition," *SIAM J. Sci. Comput.*, vol. 33, no. 5, pp. 2295–2317, 2011.
[30] Q. Zhao, G. Zhou, S. Xie, L. Zhang, and A. Cichocki, "Tensor ring decomposition," 2016, *arXiv:1606.05535*.
[31] P.-L. Wu, X.-L. Zhao, M. Ding, Y.-B. Zheng, L.-B. Cui, and T.-Z. Huang, "Tensor ring decomposition-based model with interpretable gradient factors regularization for tensor completion," *Knowl.-Based Syst.*, vol. 259, 2023, Art. no. 110094.
[32] J. A. Bengua, H. N. Phien, H. D. Tuan, and M. N. Do, "Efficient tensor completion for color image and video recovery: Low-rank tensor train," *IEEE Trans. Image Process.*, vol. 26, no. 5, pp. 2466–2479, May 2017.
[33] W. Wang, V. Aggarwal, and S. Aeron, "Efficient low rank tensor ring completion," in *Proc. IEEE Int. Conf. Comput. Vis.*, pp. 5697–5705, 2017.
[34] H. Huang, Y. Liu, J. Liu, and C. Zhu, "Provable tensor ring completion," *Signal Process.*, vol. 171, 2020, Art. no. 107486.
[35] Y. Li, D. Qiu, and X. Zhang, "Robust low transformed multi-rank tensor completion with deep prior regularization for multi-dimensional image recovery," *IEEE Trans. Big Data*, vol. 9, no. 5, pp. 1288–1301, Oct. 2023.
[36] Y. Qiu, G. Zhou, Q. Zhao, and S. Xie, "Noisy tensor completion via low-rank tensor ring," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 1, pp. 1127–1141, Jan. 2024.

[37] Y.-B. Zheng, T.-Z. Huang, X.-L. Zhao, Q. Zhao, and T.-X. Jiang, "Fully-connected tensor network decomposition and its application to higher-order tensor completion," in *Proc. AAAI Conf. Artif. Intell.*, 2021, pp. 11071–11078.
[38] Z.-C. Wu, T.-Z. Huang, L.-J. Deng, H.-X. Dou, and D. Meng, "Tensor wheel decomposition and its tensor completion application," in *Proc. Adv. Neural Inf. Process. Syst.*, 2022, pp. 27008–27020.
[39] S. Zhou, S. Erfani, and J. Bailey, "Online CP decomposition for sparse tensors," in *Proc. IEEE Int. Conf. Data Mining*, 2018, pp. 1458–1463.
[40] X. Luo, H. Wu, and Z. Li, "Neulft: A novel approach to nonlinear canonical polyadic decomposition on high-dimensional incomplete tensors," *IEEE Trans. Knowl. Data Eng.*, vol. 35, no. 6, pp. 6148–6166, Jun. 2023.
[41] C. J. Hillar and L.-H. Lim, "Most tensor problems are NP-hard," *J. ACM (JACM)*, vol. 60, no. 6, pp. 1–39, 2013.
[42] T. G. Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM Rev.*, vol. 51, no. 3, pp. 455–500, 2009.
[43] A. Karami, M. Yazdi, and G. Mercier, "Compression of hyperspectral images using discerete wavelet transform and Tucker decomposition," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 2, pp. 444–450, Apr. 2012.
[44] X. Li, M. K. Ng, G. Cong, Y. Ye, and Q. Wu, "MR-NTD: Manifold regularization nonnegative tucker decomposition for tensor data dimension reduction and representation," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 8, pp. 1787–1800, Aug. 2016.
[45] T.-X. Jiang, M. K. Ng, X.-L. Zhao, and T.-Z. Huang, "Framelet representation of tensor nuclear norm for third-order tensor completion," *IEEE Trans. Image Process.*, vol. 29, pp. 7233–7244, 2020.
[46] T.-X. Jiang, X.-L. Zhao, H. Zhang, and M. K. Ng, "Dictionary learning with low-rank coding coefficients for tensor completion," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 34, no. 2, pp. 932–946, Feb. 2023.
[47] Y.-B. Zheng, T.-Z. Huang, X.-L. Zhao, T.-X. Jiang, T.-H. Ma, and T.-Y. Ji, "Mixed noise removal in hyperspectral image via low-fibered-rank regularization," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 1, pp. 734–749, Jan. 2020.
[48] T. Huckle, K. Waldherr, and T. Schulte-Herbrüggen, "Computations in quantum tensor networks," *Linear Algebra Appl.*, vol. 438, no. 2, pp. 750–781, 2013.
[49] Y. Xu and W. Yin, "A block coordinate descent method for regularized multiconvex optimization with applications to nonnegative tensor factorization and completion," *SIAM J. Imag. Sci.*, vol. 3, pp. 1758–1789, Jan. 2013.
[50] Y.-Y. Liu, X.-L. Zhao, G.-J. Song, Y.-B. Zheng, M. K. Ng, and T.-Z. Huang, "Fully-connected tensor network decomposition for robust tensor completion problem," *Inverse Problems Imag.*, vol. 18, no. 1, pp. 208–238, 2024.
[51] J. Liu, C. Zhu, and Y. Liu, "Smooth compact tensor ring regression," *IEEE Trans. Knowl. Data Eng.*, vol. 34, no. 9, pp. 4439–4452, Sep. 2022.
[52] I. Daubechies, R. DeVore, M. Fornasier, and C. S. Güntürk, "Iteratively reweighted least squares minimization for sparse recovery," *Commun. Pure Appl. Math.: A J. Issued by Courant Inst. Math. Sci.*, vol. 63, no. 1, pp. 1–38, 2010.
[53] N. Vervliet and L. De Lathauwer, "A randomized block sampling approach to canonical polyadic decomposition of large-scale tensors," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 2, pp. 284–295, Mar. 2016.
[54] F. Zhang, L. Yang, J. Wang, and X. Luo, "Randomized sampling techniques based low-tubal-rank plus sparse tensor recovery," *Knowl.-Based Syst.*, vol. 261, 2023, Art. no. 110198.
[55] Y. Yuan, Q. He, X. Luo, and M. Shang, "A multilayered-and-randomized latent factor model for high-dimensional and sparse matrices," *IEEE Trans. Big Data*, vol. 8, no. 3, pp. 784–794, Jun. 2022.
[56] D. Teng, X. Zhang, L. Cheng, and D. Chu, "Least squares approximation via sparse subsampled randomized hadamard transform," *IEEE Trans. Big Data*, vol. 8, no. 2, pp. 446–457, Apr. 2022.
[57] L. Yuan, C. Li, D. Mandic, J. Cao, and Q. Zhao, "Tensor ring decomposition with rank minimization on latent space: An efficient approach for tensor completion," in *Proc. AAAI Conf. Artif. Intell.*, 2019, pp. 9151–9158.
[58] H. Attouch, J. Bolte, and B. F. Svaiter, "Convergence of descent methods for semi-algebraic and tame problems: Proximal algorithms, forward–backward splitting, and regularized Gauss–Seidel methods," *Math. Program.*, vol. 137, no. 1, pp. 91–129, 2013.
[59] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.