FULLY-CONNECTED TENSOR NETWORK DECOMPOSITION FOR ROBUST TENSOR COMPLETION PROBLEM

Yun-Yang Liu[✉][1], Xi-Le Zhao [✉][*][1], Guang-Jing Song[✉][2],
Yu-Bang Zheng[✉][3], Michael K. Ng[✉][4] and Ting-Zhu Huang[✉][1]

[1]School of Mathematical Sciences, University of Electronic Science and Technology of China,
Chengdu, 611731, China

[2]School of Mathematics and Information Sciences, Weifang University,
Weifang, 261061, China

[3]School of Information Science and Technology, Southwest Jiaotong University,
Chengdu, 611756, China

[4]Department of Mathematics, The University of Hong Kong, Pokfulam, 999077, Hong Kong

(Communicated by Hui Ji)

ABSTRACT. Motivated by the success of fully-connected tensor network (FCTN) decomposition, we suggest two FCTN-based models for the robust tensor completion (RTC) problem. Firstly, we propose an **FCTN**-based **r**obust **n**on**c**onvex optimization model (RNC-FCTN) directly based on FCTN decomposition for the RTC problem. Then, a proximal alternating minimization (PAM)-based algorithm is developed to solve the proposed RNC-FCTN. Meanwhile, we theoretically derive the convergence of the PAM-based algorithm. Although the nonconvex model has shown empirically excellent results, the exact recovery guarantee is still missing and $N(N-1)/2+1$ tuning parameters are difficult to choose for N-th order tensor. Therefore, we propose the FCTN nuclear norm as the convex surrogate function of the FCTN rank and suggest an **FCTN** nuclear norm-based **r**obust **c**onvex optimization model (RC-FCTN) for the RTC problem. For solving the constrained optimization model RC-FCTN, we develop an alternating direction method of multipliers (ADMM)-based algorithm, which enjoys the global convergence guarantee. To explore the exact recovery guarantee, we design a constructive singular value decomposition (SVD)-based FCTN decomposition, which is another crucial algorithm to obtain the factor tensors of FCTN decomposition. Accordingly, we rigorously establish the exact recovery guarantee for the RC-FCTN and suggest the theoretical optimal value for the only one parameter in the convex model. Comprehensive numerical experiments in several applications, such as video completion and video background subtraction, demonstrate that the suggested convex and nonconvex models have achieved state-of-the-art performance.

1. **Introduction.** Owing to various unpredictable or unavoidable reasons, the observed data is often contaminated by noise and suffers from missing information [39, 20, 32, 18], which significantly limits the accuracy of subsequent applications. The data recovery problems, such as image/video completion [42, 9, 10, 28] and

image/video noise removal [37, 8, 40, 24], are very significant. As the general-
ization of matrices, tensors can represent higher-dimensional data, most of which
is assumed to be low-rank in the recovery problems[27, 31, 21, 36]. As a typical
ill-posed inverse problem, this data recovery problems are modeled as a robust ten-
sor completion (RTC) problem, aiming to restructure a low-rank component and a
sparse component from the observed data.

Mathematically, the RTC problem can be expressed as:

$$\min_{\mathcal{X},\mathcal{E}} \; l(\mathcal{X}) + \lambda\|\mathcal{E}\|_1, \; \text{s.t.} \;\; \mathcal{P}_\Omega(\mathcal{X}+\mathcal{E}) = \mathcal{P}_\Omega(\mathcal{O}), \tag{1}$$

where $\mathcal{O}$ is the observed data, $\mathcal{X}$ and $\mathcal{E}$ are the low-rank component and the sparse
component, $l(\mathcal{X})$ is the regularization term, λ is regularization parameter, and $\mathcal{P}_\Omega$
is a projection that the entries in the set Ω are themselves while the other entries
are set to zeros. When Ω is the whole set, (1) transforms into sparse noise removal
problem. When $\lambda = 0$, (1) transforms into tensor completion problem.

High-dimensional images are usually globally correlated since they contain plenty
of highly similar information. Mathematically, we use the low-rankness to charac-
terize the global correlation. Unlike the matrix case, there exist different kinds
of tensor rank, such as Tucker rank [26], multi-rank and tubal rank [16], tensor
train (TT) rank [22], and tensor ring (TR) rank [41], which are derived from the
corresponding tensor decompositions. In general, the minimization of tensor rank
is NP-hard [12], the convex/nonconvex relaxation of tensor rank or the low-rank
tensor decomposition is usually used instead of the minimization of tensor rank.

Tucker decomposition decomposes an Nth-order tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ into
a small-sized Nth-order core tensor $\mathcal{G}$ multiplied by a matrix along each mode, i.e.,
$\mathcal{X} = \mathcal{G} \times_1 \mathbf{U}_1 \times_2 \mathbf{U}_2 \times_3 \cdots \times_N \mathbf{U}_N$. Tucker rank is a vector whose the kth entry is the
rank of the mode-k matricization of $\mathcal{X}$, i.e., $\text{rank}_{\text{Tu}}(\mathcal{X}) := (\text{rank}(\mathbf{X}_{(1)}), \text{rank}(\mathbf{X}_{(2)}),$
$\cdots, \text{rank}(\mathbf{X}_{(N)}))$, where $\mathbf{X}_{(k)} \in \mathbb{R}^{I_k \times (I_1 \cdots I_{k-1} I_{k+1} \cdots I_N)}$ is the mode-k matricization
of $\mathcal{X}$. Based on recent studies that the nuclear norm $(\|\cdot\|_*)$ was the convex relaxation
of the matrix rank, Liu et $al.$ [19] proposed the sum of nuclear norms (SNN) of all
unfolding matrices $\sum_{k=1}^N \alpha_k\|\mathbf{X}_{(k)}\|_*$ as the convex surrogate of tensor Tucker rank,
where $\alpha_k \geq 0$ and $\sum_{k=1}^N \alpha_k = 1$. Based on the SNN, Huang et $al.$ [13] studied the
robust low-Tucker-rank tensor completion problem, i.e.,

$$\min_{\mathcal{X},\mathcal{E}} \; \sum_{k=1}^N \alpha_k\|\mathbf{X}_{(k)}\|_* + \lambda\|\mathcal{E}\|_1, \; \text{s.t.} \;\; \mathcal{P}_\Omega(\mathcal{X}+\mathcal{E}) = \mathcal{P}_\Omega(\mathcal{O}), \tag{2}$$

and gave the theoretical guarantee of exact recovery under tensor certain incoher-
ence conditions. The framework of the SNN-based model is easier to calculate.
However, the SNN is not the tightest convex envelope of the sum of ranks of un-
folding matrices of a tensor. Moreover, unfolding a tensor into matrices along one
mode is an unbalanced matricization scheme. Therefore, Tucker rank cannot suit-
ably capture the global information of the tensor.

Tensor singular value decomposition (t-SVD) [17] decomposes a third-order ten-
sor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times I_3}$ into the tensor-product ($*$) of two orthogonal tensors $\mathcal{U}$, $\mathcal{V}$ and
an f-diagonal tensor $\mathcal{S}$, i.e., $\mathcal{X} = \mathcal{U} * \mathcal{S} * \mathcal{V}^{\text{T}}$. Tensor multi-rank [16] is a vector whose
entries are the rank of the frontal slice of $\hat{\mathcal{X}}$, $\text{rank}_{\text{mul}}(\mathcal{X}) := (\text{rank}(\hat{\mathbf{X}}^{(1)}), \text{rank}(\hat{\mathbf{X}}^{(2)}),$
$\cdots, \text{rank}(\hat{\mathbf{X}}^{(I_3)}))$, where $\hat{\mathcal{X}} = \text{fft}(\mathcal{X}, [], 3)$ denotes the tensor obtained by perform-
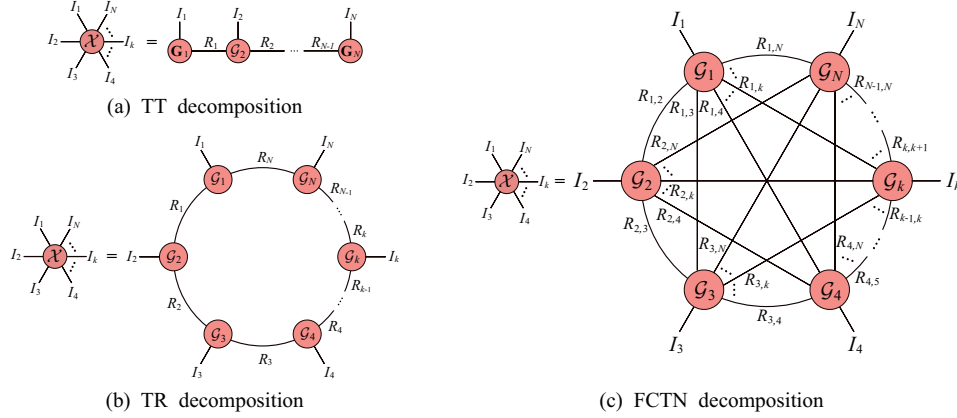ing one-dimensional Fourier transform on each tube of $\mathcal{X}$, and $\hat{\mathbf{X}}^{(k)}$ denotes the

FIGURE 1. The illustration of tensor network decomposition.

kth frontal slice of $\hat{\mathcal{X}}$. Tensor tubal rank [16] is the largest element of the tensor multi-rank. Semerci *et al.* [23] developed the tensor nuclear norm (TNN) $\sum_{k=1}^{I_3} \|\hat{\mathbf{X}}^{(k)}\|_*/I_3$ as the convex envelope of l_1 norm of the multi-rank. Utilizing the TNN, Jiang *et al.* [15] studied the RTC problem, i.e.,

$$\min_{\mathcal{X},\mathcal{E}} \frac{1}{I_3} \sum_{k=1}^{I_3} \|\hat{\mathbf{X}}^{(k)}\|_* + \lambda\|\mathcal{E}\|_1, \text{ s.t. } \mathcal{P}_\Omega(\mathcal{X} + \mathcal{E}) = \mathcal{P}_\Omega(\mathcal{O}), \tag{3}$$

and provided the theoretical guarantee for the exact recovery. Recently, Song *et al.* [25] generalized the t-SVD theory via multiplying by a unity matrix on all tubes instead of the fixed discrete Fourier transform matrix. A data-driven RTC model is also proposed, which has been shown to have obvious advantages in processing third-order tensors. However, the t-SVD focuses on third-order tensors. For high-order data, such as color videos and multi-temporal remote sensing images, the t-SVD and the corresponding TNN-based models may not effectively capture the low-dimensional structure of data.

Recently, TT and TR decompositions have emergend as powerful tools for dealing with higher-order tensors. TT decomposition (as shown in Fig. 1 (a)) [22] decomposes an Nth-order tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ into two matrices and $N-2$ third-order tensors, and the element-wise form is expressed as $\mathcal{X}(i_1, i_2, \cdots, i_N) = \sum_{r_1=1}^{R_1} \sum_{r_2=1}^{R_2} \cdots \sum_{r_{N-1}=1}^{R_{N-1}} \{\mathbf{G}_1(i_1, r_1)\mathcal{G}_2(r_1, i_2, r_2)\cdots\mathbf{G}_N(r_{N-1}, i_N)\}$. And TT rank is defined as an $(N-1)$-dimensional vector $(R_1, R_2, \cdots, R_{N-1})$. I. V. Oseledets [22] proved that there existed a TT decomposition such that $R_k \leq \text{rank}(\mathbf{X}_{[k]})$, $k = 1, 2, \cdots, N-1$, where $\mathbf{X}_{[k]} \in \mathbb{R}^{\Pi_{n=1}^k I_n \times \Pi_{n=k+1}^N I_n}$ is the k-mode matricization of $\mathcal{X}$. In [2], TT nuclear norm (TTNN) $\sum_{k=1}^{N-1} \alpha_k\|\mathbf{X}_{[k]}\|_*$ was proposed to use as a convex surrogate of TT rank for more convenient calculation and applied to low-rank tensor completion problem. Equiped with the TTNN, Chen *et al.* [9] studied the RTC problem, i.e.,

$$\min_{\mathcal{X},\mathcal{E}} \sum_{k=1}^{N-1} \alpha_k\|\mathbf{X}_{[k]}\|_* + \lambda\|\mathcal{E}\|_1, \text{ s.t. } \mathcal{P}_\Omega(\mathcal{X} + \mathcal{E}) = \mathcal{P}_\Omega(\mathcal{O}). \tag{4}$$

As an improvement of Tucker rank, TT rank help us to study the correlation between the first k modes (rather than one mode) and the rest modes. However, in real applications, the performance of TT decomposition is highly dependent on the permutations of the tensor dimensions. TR decomposition [41] (as shown in Fig. 1 (b)) decomposes an Nth-order tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ into a circular multilinear product of a list of third-order core tensors, and the element-wise form is expressed as $\mathcal{X}(i_1, i_2, \cdots, i_N) = \sum_{r_1=1}^{R_1} \sum_{r_2=1}^{R_2} \cdots \sum_{r_N=1}^{R_N} \{\mathcal{G}_1(r_1, i_1, r_2)\mathcal{G}_2(r_2, i_2, r_3) \cdots \mathcal{G}_N(r_N, i_N, r_1)\}$. And TR rank is defined as a N-dimensional vector $(R_1, R_2, \cdots, R_N)$. Inspired by the connection between TR rank and the rank of circularly unfolding matrices, TR nuclear norm minimization (TRNNM) [34] $\sum_{k=1}^{N} \alpha_k \|\mathbf{X}_{<k,L>}\|_*$ was suggested as a convex surrogate of TR rank for low-rank tensor completion problem, where $\mathbf{X}_{<k,L>} \in \mathbb{R}^{\Pi_{i=k}^{k+L-1} I_i \times \Pi_{i=k+L}^{k-1} I_i}$ is the tensor circular unfolding matrix. Then, Huang $et\ al.$ [14] studied the TRNNM-based RTC problem, i.e.,

$$\min_{\mathcal{X}, \mathcal{E}} \sum_{k=1}^{N} \alpha_k \|\mathbf{X}_{<k,L>}\|_* + \lambda \|\mathcal{E}\|_1, \ \text{s.t.} \ \ \mathcal{P}_\Omega(\mathcal{X} + \mathcal{E}) = \mathcal{P}_\Omega(\mathcal{O}). \tag{5}$$

TR decomposition has generalized representation abilities since it can be viewed as the linear combination of TT decomposition. However, TR decomposition only connects the adjacent two factors, which is highly sensitive to the order of tensor modes.

To compensate for the limitations of the TT and TR decompositions, Zheng $et\ al.$ [44] proposed the fully-connected tensor network (FCTN) decomposition (as shown in Fig. 1 (c)), which decomposes an Nth-order tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ into N small-sized Nth-order tensors, and each factor interacts with the others. Mathematically, the element-wise form is expressed as

$$\begin{aligned} \mathcal{X}(i_1, i_2, \cdots, i_N) = \\ \sum_{r_{1,2}=1}^{R_{1,2}} \sum_{r_{1,3}=1}^{R_{1,3}} \cdots \sum_{r_{1,N}=1}^{R_{1,N}} \sum_{r_{2,3}=1}^{R_{2,3}} \cdots \sum_{r_{2,N}=1}^{R_{2,N}} \cdots \sum_{r_{N-1,N}=1}^{R_{N-1,N}} \\ \{\mathcal{G}_1(i_1, r_{1,2}, r_{1,3}, \cdots, r_{1,N})\mathcal{G}_2(r_{1,2}, i_2, r_{2,3}, \cdots, r_{2,N}) \cdots \\ \mathcal{G}_k(r_{1,k}, r_{2,k}, \cdots, r_{k-1,k}, i_k, r_{k,k+1}, \cdots, r_{k,N}) \cdots \\ \mathcal{G}_N(r_{1,N}, r_{2,N}, \cdots, r_{N-1,N}, i_N)\}. \end{aligned} \tag{6}$$

FCTN rank is defined as vector $(R_{1,2}, R_{1,3}, \cdots, R_{1,N}, R_{2,3}, R_{2,4}, \cdots, R_{2,N}, \cdots, R_{N-1,N})$. Compared with other tensor decompositions, the FCTN decomposition obtains superior performance on the tensor completion problem. The reason is that it can flexibly characterize the correlation between arbitrary modes. In this paper, we leverage the strong expression ability of FCTN into the RTC problem. The contribution of this paper is threefold:

(i) To tackle the more general RTC problem, we propose an FCTN-based robust nonconvex optimization model (RNC-FCTN) and develop a proximal alternating minimization (PAM)-based algorithm to solve the proposed model. Moreover, we theoretically derive the convergence of the PAM-based algorithm. Although the nonconvex model RNC-FCTN has shown empirically excellent results, the exact recovery guarantee is still missing and $N(N-1)/2+1$ tuning parameters are difficult to choose for N-th order tensor.

(ii) To tackle the exact recovery theory and the problem of too many parameters, we firstly propose the FCTN nuclear norm as the convex surrogate of the

TABLE 1. The notations of the unfolding matrices of tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$.

Name	Form	Size	The corresponding tensor decomposition
The mode-k matricization	$\mathbf{X}_{(k)}$	$I_k \times \Pi_{i=1, i \neq k}^{N} I_i$	Tucker decomposition
The kth frontal slice	$\mathbf{X}^{(k)}$	$I_1 \times I_2$	t-SVD
The k-mode matricization	$\mathbf{X}_{[k]}$	$\Pi_{i=1}^{k} I_i \times \Pi_{i=k+1}^{N} I_i$	TT decomposition
The kth circularly unfolding matrice	$\mathbf{X}_{<k,L>}$	$\Pi_{i=k}^{k+L-1} I_i \times \Pi_{i=k+L}^{k-1} I_i$	TR decomposition
The generalized unfolding matrix	$\mathbf{X}_{[\mathbf{n}_{1:d}; \mathbf{n}_{d+1:N}]}$	$\Pi_{i=1}^{d} I_{\mathbf{n}_i} \times \Pi_{i=d+1}^{N} I_{\mathbf{n}_i}$	FCTN decomposition

FCTN rank. By applying the FCTN nuclear norm to the RTC problem, we suggest an FCTN nuclear norm-based robust convex optimization model (RC-FCTN). Secondly, we utilize the alternating direction method of multipliers (ADMM)-based algorithm to solve the proposed RC-FCTN. Finally, a constructive singular value decomposition (SVD)-based FCTN decomposition is designed, which is another crucial algorithm to obtain the factor tensors of FCTN decomposition. Accordingly, we theoretically **establish the exact recovery guarantee** and **derive the theoretical optimal value of the only one parameter** for RC-FCTN.

(iii) Extensive numerical experiments on several tasks, such as video completion and video background subtraction, demonstrate that the proposed convex and nonconvex models have achieved state-of-the-art performance.

The outline of the paper is as follows. We summarize necessary preliminaries throughout the paper in Section 2. We present the nonconvex model RNC-FCTN for the RTC problem and theoretically derive the convergence guarantee of the PAM-based algorithm in Section 3. We introduce the convex model RC-FCTN for the RTC problem and establish the exact recovery guarantee of the RC-FCTN model and the convergence guarantee of the ADMM-based algorithm in Section 4. We report extensive numerical experiments to verify the superior performance of the proposed methods in Section 5 and conclude this paper in Section 6.

2. **Preliminaries.** The material in this section mainly introduce the FCTN decomposition, which is essentially the same as in [44]. We use x, $\mathbf{x}$, $\mathbf{X}$, and $\mathcal{X}$ to denote scalars, vectors, matrices, and tensors, respectively. For tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$, we denote $\mathcal{X}(i_1, i_2, \cdots, i_N)$ as its $(i_1, i_2, \cdots, i_N)$th element and summarize its various unfolding matrices in Table 1. The **inner product** of two tensors $\mathcal{X}$ and $\mathcal{Y}$ with the same size is defined as the sum of the products of their entries, i.e., $\langle \mathcal{X}, \mathcal{Y} \rangle = \sum_{i_1, i_2, \cdots, i_N} \mathcal{X}(i_1, i_2, \cdots, i_N) \mathcal{Y}(i_1, i_2, \cdots, i_N)$. The l_1-**norm** and **Frobenius norm** of $\mathcal{X}$ are defined as $\|\mathcal{X}\|_1 = \sum_{i_1, i_2, \cdots, i_N} |\mathcal{X}(i_1, i_2, \cdots, i_N)|$ and $\|\mathcal{X}\|_F = \sqrt{\sum_{i_1, i_2, \cdots, i_N} |\mathcal{X}(i_1, i_2, \cdots, i_N)|^2}$, respectively. The **generalized transposition** of tensor $\mathcal{X}$ is $\vec{\mathcal{X}}^{\mathbf{n}} \in \mathbb{R}^{I_{n_1} \times I_{n_2} \times \cdots \times I_{n_N}}$, which is rearranging the modes of $\mathcal{X}$ by the specified vector $\mathbf{n}$. The corresponding operation and its inverse operation are denoted as $\vec{\mathcal{X}}^{\mathbf{n}} = \text{permute}[1](\mathcal{X}, \mathbf{n})$ and $\mathcal{X} = \text{ipermute}[1](\vec{\mathcal{X}}^{\mathbf{n}}, \mathbf{n})$, respectively. The **generalized unfolding** of $\mathcal{X}$ is a matrix defined as $\mathbf{X}_{[\mathbf{n}_{1:d}; \mathbf{n}_{d+1:N}]} = \text{reshape}[1](\vec{\mathcal{X}}^{\mathbf{n}}, \prod_{i=1}^{d} I_{\mathbf{n}_i},$

---

[1]Matlab commands

$\prod_{i=d+1}^{N} I_{\mathbf{n}_i}$), whose elements satisfy

$$\mathbf{X}_{[\mathbf{n}_{1:d};\mathbf{n}_{d+1:N}]}(j_1, j_2)$$

$$= \mathcal{X}(i_1, i_2, \cdots, i_N) \text{ with } \begin{cases} j_1 = i_{\mathbf{n}_1} + \sum_{k=2}^{d} \left((i_{\mathbf{n}_k}-1) \prod_{m=1}^{k-1} I_{\mathbf{n}_m} \right), \\ j_2 = i_{\mathbf{n}_{d+1}} + \sum_{k=d+2}^{N} \left((i_{\mathbf{n}_k}-1) \prod_{m=d+1}^{k-1} I_{\mathbf{n}_m} \right). \end{cases}$$

The corresponding inverse operation is defined as $\mathcal{X} = \Upsilon(\mathbf{X}_{[\mathbf{n}_{1:d};\mathbf{n}_{d+1:N}]})$.

Definition 2.1 (Tensor contraction). Suppose that $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ and $\mathcal{Y} \in \mathbb{R}^{J_1 \times J_2 \times \cdots \times J_M}$ have d modes of the same size ($I_{n_i} = J_{m_i}$ with $i = 1, 2, \cdots, d$), the **tensor contraction** along the $\mathbf{n}_{1:d}$th-modes of $\mathcal{X}$ and the $\mathbf{m}_{1:d}$th-modes of $\mathcal{Y}$ is an (N+M-2d)th-order tensor that satisfied

$$\mathcal{Z} = \mathcal{X} \times_{\mathbf{n}_{1:d}}^{\mathbf{m}_{1:d}} \mathcal{Y} \Leftrightarrow \mathbf{Z}_{[1:N-d;N-d+1:N+M-2d]} = \mathbf{X}_{[\mathbf{n}_{d+1:N};\mathbf{n}_{1:d}]} \mathbf{Y}_{[\mathbf{m}_{1:d};\mathbf{m}_{d+1:M}]}. \quad (7)$$

Definition 2.2 (FCTN decomposition). An Nth-order tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ can be decomposed into a series of Nth-order factor tensors $\mathcal{G}_k$ with size $R_{1,k} \times R_{2,k} \times \cdots \times R_{k-1,k} \times I_k \times R_{k,k+1} \times \cdots \times R_{k,N}, (k = 1, 2, \cdots, N)$, whose elements satisfied

$$\mathcal{X}(i_1, i_2, \cdots, i_N) =$$
$$\sum_{r_{1,2}=1}^{R_{1,2}} \sum_{r_{1,3}=1}^{R_{1,3}} \cdots \sum_{r_{1,N}=1}^{R_{1,N}} \sum_{r_{2,3}=1}^{R_{2,3}} \cdots \sum_{r_{2,N}=1}^{R_{2,N}} \cdots \sum_{r_{N-1,N}=1}^{R_{N-1,N}}$$
$$\{\mathcal{G}_1(i_1, r_{1,2}, r_{1,3}, \cdots, r_{1,N}) \mathcal{G}_2(r_{1,2}, i_2, r_{2,3}, \cdots, r_{2,N}) \cdots$$
$$\mathcal{G}_k(r_{1,k}, r_{2,k}, \cdots, r_{k-1,k}, i_k, r_{k,k+1}, \cdots, r_{k,N}) \cdots$$
$$\mathcal{G}_N(r_{1,N}, r_{2,N}, \cdots, r_{N-1,N}, i_N)\}. \quad (8)$$

This decomposition is defined as the FCTN decomposition. The factors $\mathcal{G}_1, \mathcal{G}_2, \cdots, \mathcal{G}_N$ are the core tensors of $\mathcal{X}$ and can be abbreviated as $\{\mathcal{G}\}_{1:N}$, then $\mathcal{X} = \text{FCTN}(\{\mathcal{G}\}_{1:N})$. The vector $(R_{1,2}, R_{1,3}, \cdots, R_{1,N}, R_{2,3}, R_{2,4}, \cdots, R_{2,N}, \cdots, R_{N-1,N})$ is defined as the FCTN-rank of the original tensor $\mathcal{X}$.

The introduction mentions several tensor decompositions and their corresponding tensor ranks. Now, we discuss the relationships between low-Tucker-rank (low-multi-tubal-rank, low-TT-rank, and low-TR-rank) tensor and low-FCTN-rank tensor.

Proposition 2.3. *The data with **low-Tucker-rank** is also **low-FCTN-rank**.*

Proof. Without loss of generality, we consider fourth-order data. The Tucker decomposition of a fourth-order tensor $\mathcal{X}$ is denoted as $\mathcal{X} = \mathcal{Z} \times_1 \mathbf{G}_1 \times_2 \mathbf{G}_2 \times_3 \mathbf{G}_3 \times_4 \mathbf{G}_4$. Since the tensor $\mathcal{X}$ is of low-Tucker-rank, that is, its Tucker-rank (R_1, R_2, R_3, R_4) is small. Then we perform FCTN decomposition on factor tensor $\mathcal{Z} \in \mathbb{R}^{R_1 \times R_2 \times R_3 \times R_4}$, the resulting FCTN-rank $(R_{1,2}, R_{1,3}, R_{1,4}, R_{2,3}, R_{2,4}, R_{3,4})$ will also be small since the FCTN-rank will not greater than the corresponding dimension size (R_1, R_2, R_3, R_4), see Fig. 2 (a). Finally, we can realize the FCTN decomposition of tensor $\mathcal{X}$ through the previously obtained factors, i.e., $\mathcal{X} = \text{FCTN}(\mathcal{G}_1 \times_1 \mathbf{G}_1, \mathcal{G}_2 \times_2 \mathbf{G}_2, \mathcal{G}_3 \times_3 \mathbf{G}_3, \mathcal{G}_4 \times_4 \mathbf{G}_4)$, and $(R_{1,2}, R_{1,3}, R_{1,4}, R_{2,3}, R_{2,4}, R_{3,4})$ becomes the FCTN-rank of tensor $\mathcal{X}$. It is obvious that the FCTN-rank of tensor $\mathcal{X}$ is small. □
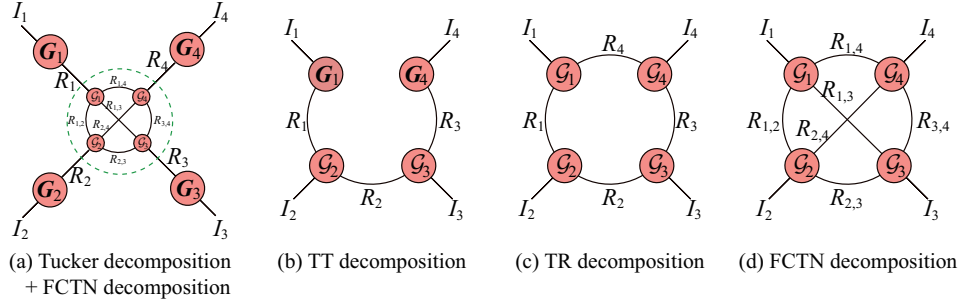
(a) Tucker decomposition + FCTN decomposition

(b) TT decomposition

(c) TR decomposition

(d) FCTN decomposition

FIGURE 2. The illustration of tensor network decompositions.

Proposition 2.4. *The data with **low-multi-tubal-rank** is also **low-FCTN-rank**.*

Proof. The tensor multi-tubal-rank is a vector (R_1, R_2, R_3) in which its elements characterize the low-rank structure of horizontal slices, lateral slices, and frontal slices of the transformed tensor along the first, second, and third mode, respectively. Actually, the tensor tubal-rank is R_3. In addition, Theorem 2.1 in [35] reveals that, in characterizing the low-rankness of data, multi-tubal-rank is more powerful than Tucker-rank. Equipped with Proposition 2.3, we can infer that when the data is of low-multi-tubal-rank, it is also low-FCTN-rank. $\qquad\square$

Proposition 2.5. *The data with **low-TT-rank** is also **low-FCTN-rank**.*

Proof. In fact, TT decomposition can be viewed as a special case of FCTN decomposition, see Fig. 2 (b). Since the tensor $\mathcal{X}$ is of low-TT-rank, that is, its TT-rank (R_1, R_2, R_3) is small. We can rewrite the TT-rank into the corresponding FCTN form, i.e., $(R_1, 1, 1, R_2, 1, R_3)$, which is obviously small. $\qquad\square$

Proposition 2.6. *The data with **low-TR-rank** is also **low-FCTN-rank**.*

Proof. Similarly, TR decomposition can also be viewed as a special case of FCTN decomposition, please see Fig. 2 (c). Since the tensor $\mathcal{X}$ is of low-TR-rank, that is, its TR-rank (R_1, R_2, R_3, R_4) is small. We can rewrite the TR-rank into the corresponding FCTN form, i.e., $(R_1, 1, R_4, R_2, 1, R_3)$, which is obviously small. $\quad\square$

3. **RNC-FCTN model.** In this section, inspired by the superiority of FCTN decomposition for tensor completion, we suggest an FCTN-based nonconvex optimization model for the general RTC problem as follows:

$$\min_{\mathcal{X},\mathcal{E},\{\mathcal{F}\}_{1:N}} \frac{1}{2}\|\mathcal{X} - \text{FCTN}(\{\mathcal{F}\}_{1:N})\|_{\text{F}}^2 + \lambda\|\mathcal{E}\|_1 \tag{9}$$
$$\text{s.t.} \quad \mathcal{P}_\Omega(\mathcal{X} + \mathcal{E}) = \mathcal{P}_\Omega(\mathcal{O}),$$

where λ is regularization parameter. By introducing auxiliary valuable $\mathcal{Y}$, we can equivalently rewrite the problem (9) as the following unconstraint problem

$$\min_{\mathcal{X},\mathcal{E},\{\mathcal{F}\}_{1:N},\mathcal{Y}} \frac{1}{2}\|\mathcal{X} - \text{FCTN}(\{\mathcal{F}\}_{1:N})\|_{\text{F}}^2 + \lambda\|\mathcal{E}\|_1 + \frac{\beta}{2}\|\mathcal{Y} - \mathcal{X} - \mathcal{E}\|_{\text{F}}^2 + \Phi(\mathcal{Y}), \tag{10}$$

where

$$\Phi(\mathcal{Y}) = \begin{cases} 0, & \mathcal{P}_\Omega(\mathcal{Y}) = \mathcal{P}_\Omega(\mathcal{O}) \\ \infty, & \text{otherwise} \end{cases}, \tag{11}$$

and β is a penalty parameter approaching positive infinity.

The nonconvex optimization problem (10) allows us to use the nonconvex solvers, such as PAM algorithm [4, 1] and BCD algorithm [29]. In this work, we consider to use the PAM algorithm to solve the suggested model. At each iteration, a single block of variables is optimized, while the remaining variables are fixed. Detailedly, the PAM-based algorithm is updated as the following iterative scheme:

$$
\begin{cases}
\mathcal{F}_k^{t+1} = \underset{\mathcal{F}_k}{\arg\min}\, f(\{\mathcal{F}\}_{1:k-1}^{t+1},\ \mathcal{F}_k,\ \{\mathcal{F}\}_{k+1:N}^t,\ \mathcal{X}^t,\ \mathcal{E}^t,\ \mathcal{Y}^t) + \frac{\rho}{2}\|\mathcal{F}_k - \mathcal{F}_k^t\|_F^2, \\[2mm]
\mathcal{X}^{t+1} = \underset{\mathcal{X}}{\arg\min}\, f(\{\mathcal{F}\}_{1:N}^{t+1},\ \mathcal{X},\ \mathcal{E}^t,\ \mathcal{Y}^t) + \frac{\rho}{2}\|\mathcal{X} - \mathcal{X}^t\|_F^2, \\[2mm]
\mathcal{E}^{t+1} = \underset{\mathcal{E}}{\arg\min}\, f(\{\mathcal{F}\}_{1:N}^{t+1},\ \mathcal{X}^{t+1},\ \mathcal{E},\ \mathcal{Y}^t) + \frac{\rho}{2}\|\mathcal{E} - \mathcal{E}^t\|_F^2, \\[2mm]
\mathcal{Y}^{t+1} = \underset{\mathcal{Y}}{\arg\min}\, f(\{\mathcal{F}\}_{1:N}^{t+1},\ \mathcal{X}^{t+1},\ \mathcal{E}^{t+1},\ \mathcal{Y}) + \frac{\rho}{2}\|\mathcal{Y} - \mathcal{Y}^t\|_F^2,
\end{cases}
\tag{12}
$$

where $f(\{\mathcal{F}\}_{1:N},\ \mathcal{X},\ \mathcal{E},\ \mathcal{Y})$ is the objective function of (10) and $\rho > 0$ is a proximal parameter.

The corresponding details are as follows.

1) Update $\mathcal{F}_k$: the $\mathcal{F}_k$ $(k = 1, 2, \cdots, N)$ subproblems can be equivalently rewritten as

$$
\begin{aligned}
\mathcal{F}_k^{t+1} &= \underset{\mathcal{F}_k}{\arg\min}\, \frac{1}{2}\|\mathcal{X}^t - \mathrm{FCTN}(\{\mathcal{F}\}_{1:k-1}^{t+1},\ \mathcal{F}_k,\ \{\mathcal{F}\}_{k+1:N}^t)\|_F^2 + \frac{\rho}{2}\|\mathcal{F}_k - \mathcal{F}_k^t\|_F^2 \\
&= \underset{\mathcal{F}_k}{\arg\min}\, \left\{ \frac{1}{2}\|\mathbf{X}_{(k)}^t - (\mathbf{F}_k)_{(k)}(\mathbf{M}_k^t)_{[\mathbf{m}_{1:N-1};\mathbf{n}_{1:N-1}]}\|_F^2 + \frac{\rho}{2}\|(\mathbf{F}_k)_{(k)} - (\mathbf{F}_k^t)_{(k)}\|_F^2 \right\},
\end{aligned}
\tag{13}
$$

where $\mathcal{M}_k^t = \mathrm{FCTN}(\{\mathcal{F}\}_{1:k-1}^{t+1},\ \{\mathcal{F}\}_{k+1:N}^t)$. The problem (13) can be directly solved as

$$
\begin{aligned}
(\mathbf{F}_k^{t+1})_{(k)} =& [\mathbf{X}_{(k)}^t(\mathbf{M}_k^t)_{[\mathbf{n}_{1:N-1};\mathbf{m}_{1:N-1}]} + \rho(\mathbf{F}_k^t)_{(k)}] \\
& [(\mathbf{M}_k^t)_{[\mathbf{m}_{1:N-1};\mathbf{n}_{1:N-1}]}(\mathbf{M}_k^t)_{[\mathbf{n}_{1:N-1};\mathbf{m}_{1:N-1}]} + \rho\mathbf{I}]^{-1}.
\end{aligned}
\tag{14}
$$

To facilitate the calculation of complexity, we simply set $I_1 = I_2 = \cdots = I_N = I$ and the FCTN rank R_{k_1,k_2} as the same value R. This step of calculation mainly includes tensor contraction, matrix multiplication, and matrix inversion. The computational complexity of updating $\mathbf{F}_k$ is $\mathcal{O}(N\sum_{i=2}^N I^i R^{i(N-i)+i-1} + NI^{N-1}R^{2(N-1)} + NR^{3(N-1)})$.

2) Update $\mathcal{X}$: the $\mathcal{X}$ subproblem can be simplified as

$$
\mathcal{X}^{t+1} = \underset{\mathcal{X}}{\arg\min}\, \frac{1}{2}\|\mathcal{X} - \mathrm{FCTN}(\{\mathcal{F}\}_{1:N}^{t+1})\|_F^2 + \frac{\beta}{2}\|\mathcal{Y}^t - \mathcal{X} - \mathcal{E}^t\|_F^2 + \frac{\rho}{2}\|\mathcal{X} - \mathcal{X}^t\|_F^2.
\tag{15}
$$

The optimization function (15) is differentiable and its closed-form solution can be obtained by solving the following Sylvester equation:

$$
\mathcal{X} + \beta\mathcal{X} + \rho\mathcal{X} = \mathrm{FCTN}(\{\mathcal{F}\}_{1:N}^{t+1}) + \beta(\mathcal{Y}^t - \mathcal{E}^t) + \rho\mathcal{X}^t,
\tag{16}
$$

Directly solving (16), we have

$$
\mathcal{X}^{t+1} = \frac{\mathrm{FCTN}(\{\mathcal{F}\}_{1:N}^{t+1}) + \beta(\mathcal{Y}^t - \mathcal{E}^t) + \rho\mathcal{X}^t}{1 + \beta + \rho}.
\tag{17}
$$

The computational complexity of updating $\mathcal{X}$ is $\mathcal{O}(\sum_{i=2}^N I^i R^{i(N-i)+i-1})$.

Algorithm 1 PAM-based algorithm for solving RNC-FCTN.

Input: The observed tensor $\mathcal{Y}$, the maximal FCTN rank $R^{\max}$, and parameter λ.
Initialization: $t = 0$, $\mathcal{X}^0$, $\mathcal{E}^0$, and $\mathcal{F}_k^0$, the original FCTN rank R, parameter β
   and ρ;
1: Update $\mathcal{F}_k^{t+1}$ via (14);
2: Update $\mathcal{X}^{t+1}$ via (17);
3: Update $\mathcal{E}^{t+1}$ via (19);
4: Update $\mathcal{Y}^{t+1}$ via (21);
5: If $\|\mathcal{X}^{t+1} - \mathcal{X}^t\|_F / \|\mathcal{X}^t\|_F \leq 10^{-2}$, set $R = \min\{R+1, R^{\max}\}$ and expand $\mathcal{F}_k^{t+1}$;
6: Check convergence criteria: $\|\mathcal{X}^{t+1} - \mathcal{X}^t\|_F / \|\mathcal{X}^t\|_F \leq \epsilon$;
7: If the convergence criteria is not meet, set $t := t + 1$ and go to Step 1.
Output: The low-rank part $\mathcal{X}$ and sparse part $\mathcal{E}$.

---

3) Update $\mathcal{E}$: the $\mathcal{E}$ subproblem can be equivalently rewritten as

$$
\begin{aligned}
\mathcal{E}^{t+1} &= \arg\min_{\mathcal{E}} \lambda\|\mathcal{E}\|_1 + \frac{\beta}{2}\|\mathcal{Y}^t - \mathcal{X}^{t+1} - \mathcal{E}\|_F^2 + \frac{\rho}{2}\|\mathcal{E} - \mathcal{E}^t\|_F^2 \\
&= \arg\min_{\mathcal{E}} \lambda\|\mathcal{E}\|_1 + \frac{\beta+\rho}{2}\left\|\mathcal{E} - \frac{\beta(\mathcal{Y}^t - \mathcal{X}^{t+1}) + \rho\mathcal{E}^t}{\beta+\rho}\right\|_F^2,
\end{aligned}
\tag{18}
$$

which can be updated as follows

$$
\mathcal{E}^{t+1} = \text{soft}\left(\frac{\beta(\mathcal{Y}^t - \mathcal{X}^{t+1}) + \rho\mathcal{E}^t}{\beta+\rho}, \frac{\lambda}{\beta+\rho}\right).
\tag{19}
$$

The computational complexity of updating $\mathcal{E}$ is $\mathcal{O}(I^N)$.

4) Update $\mathcal{Y}$: the $\mathcal{Y}$ subproblem can be rewritten as

$$
\begin{aligned}
\mathcal{Y}^{t+1} &= \arg\min_{\mathcal{Y}} \Phi(\mathcal{Y}) + \frac{\beta}{2}\|\mathcal{Y} - \mathcal{X}^{t+1} - \mathcal{E}^{t+1}\|_F^2 + \frac{\rho}{2}\|\mathcal{Y} - \mathcal{Y}^t\|_F^2 \\
&= \arg\min_{\mathcal{Y}} \Phi(\mathcal{Y}) + \frac{\beta+\rho}{2}\left\|\mathcal{Y} - \frac{\beta(\mathcal{X}^{t+1} + \mathcal{E}^{t+1}) + \rho\mathcal{Y}^t}{\beta+\rho}\right\|_F^2.
\end{aligned}
\tag{20}
$$

The $\mathcal{Y}$ subproblem has the following closed-form solution:

$$
\mathcal{Y}^{t+1} = \left(\frac{\beta(\mathcal{X}^{t+1} + \mathcal{E}^{t+1}) + \rho\mathcal{Y}^t}{\beta+\rho}\right)_{\Omega^C} + \mathcal{O}_\Omega.
\tag{21}
$$

where Ω^C is the complementary set of Ω. The computational complexity of updating $\mathcal{Y}$ is $\mathcal{O}(I^N)$.

Theorem 3.1. *The sequence $\{\{\mathcal{F}\}_{1:N}^t, \mathcal{X}^t, \mathcal{E}^t, \mathcal{Y}^t\}_{t\in\mathbb{N}}$ obtained by Algorithm 1 converges to a critical point of (9).*

The detailed proof for Theorem 3.1 is presented in Appendix.

Although the nonconvex model has shown empirically excellent results, the exact recovery guarantee is still missing and there are $N(N-1)/2+1$ tuning parameters for N-th order tensors.

4. **RC-FCTN model.** We first introduce a lemma in [44], which reveals the relationship between the rank of generalized tensor unfolding matrices and the FCTN rank.

Lemma 4.1. *An Nth-order tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ can be represented by (8), then we have*

$$\text{Rank}(\boldsymbol{X}_{[\boldsymbol{n}_{1:d};\boldsymbol{n}_{d+1:N}]}) \leq \prod_{i=1}^{d} \prod_{j=d+1}^{N} R_{\boldsymbol{n}_i,\boldsymbol{n}_j} (R_{\boldsymbol{n}_i,\boldsymbol{n}_j} = R_{\boldsymbol{n}_j,\boldsymbol{n}_i}, \text{if } \boldsymbol{n}_i > \boldsymbol{n}_j). \tag{22}$$

To tackle the exact recovery theory and parameters problem, we suggest a new FCTN nuclear norm as a convex surrogate of FCTN rank by utilizing the relationship between the rank of generalized tensor unfolding matrices and the FCTN rank.

Definition 4.2 (FCTN nuclear norm). For an Nth-order tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$, its **FCTN nuclear norm** is defined as

$$\sum_{k=1}^{\bar{N}} \alpha_k \|\mathbf{X}_{[\mathbf{n_1^k};\mathbf{n_2^k}]}\|_*, \tag{23}$$

where $\mathbf{n}^k$ is the k-th permutation of the vector $(1, 2, \cdots, N)$, $\mathbf{n_1^k} = \mathbf{n}_{1:\lfloor N/2 \rfloor}^k$, $\mathbf{n_2^k} = \mathbf{n}_{\lfloor N/2 \rfloor+1:N}^k$, $\alpha_k \geq 0$, $\sum_{k=1}^{\bar{N}} \alpha_k = 1$,

$$\bar{N} = \begin{cases} C_N^{\lfloor N/2 \rfloor}, & \text{if } N \text{ is odd}, \\ C_N^{\lfloor N/2 \rfloor}/2, & \text{if } N \text{ is even}, \end{cases} \tag{24}$$

and $\lfloor \cdot \rfloor$ denotes the floor function.

Based on the proposed FCTN nuclear norm, we suggest a robust convex optimization model RC-FCTN for the RTC problems as follows:

$$\begin{aligned} \min_{\mathcal{X},\mathcal{E}} \ & \sum_{k=1}^{\bar{N}} \alpha_k \|\mathbf{X}_{[\mathbf{n_1^k};\mathbf{n_2^k}]}\|_* + \lambda\|\mathcal{E}\|_1 \\ \text{s.t.} \ & \mathcal{P}_\Omega(\mathcal{X} + \mathcal{E}) = \mathcal{P}_\Omega(\mathcal{O}), \end{aligned} \tag{25}$$

where $\mathcal{X}$ and $\mathcal{E}$ are the low-rank component and the sparse component, λ is a regularization parameter, $\mathcal{O}$ is the observed data, and $\mathcal{P}_\Omega$ is a projection that the entries in the set Ω are themselves while the other entries are set to zeros.

4.1. **ADMM-based algorithm for solving RC-FCTN.** For the convex optimization problem (25), we develop an ADMM-based algorithm [5] to solve it. By introducing auxiliary variables $\mathcal{L}_k$ ($k = 1, 2, \cdots, \bar{N}$), the problem (25) can be rewritten as

$$\begin{aligned} \min_{\mathcal{X},\mathcal{E},\mathcal{L}^k} \ & \sum_{k=1}^{\bar{N}} \alpha_k \|\mathbf{L}_{k[\mathbf{n_1^k};\mathbf{n_2^k}]}\|_* + \lambda\|\mathcal{S}\|_1 + \Phi(\mathcal{Y}) \\ \text{s.t.} \ & \mathcal{Y} = \mathcal{X} + \mathcal{E}, \mathcal{S} = \mathcal{E}, \mathcal{L}_k = \mathcal{X}, k = 1, 2, \cdots, \bar{N}, \end{aligned} \tag{26}$$

where

$$\Phi(\mathcal{Y}) = \begin{cases} 0, & \mathcal{P}_\Omega(\mathcal{Y}) = \mathcal{P}_\Omega(\mathcal{O}) \\ \infty, & \text{otherwise} \end{cases}. \tag{27}$$

The augmented Lagrangian function of (26) is

$$L(\mathcal{X},\mathcal{E},\mathcal{S},\mathcal{Y},\mathcal{L}_k,\mathcal{Z}_k,\mathcal{P},\mathcal{Q}) = \sum_{k=1}^{\bar{N}}\left\{\alpha_k\|\mathbf{L}_{k[\mathbf{n}_1^k;\mathbf{n}_2^k]}\|_* + \langle\mathcal{Z}_k,\mathcal{L}_k - \mathcal{X}\rangle + \frac{\mu_k}{2}\|\mathcal{L}_k - \mathcal{X}\|_F^2\right\}$$
$$+ \lambda\|\mathcal{S}\|_1 + \langle\mathcal{P},\mathcal{Y} - \mathcal{X} - \mathcal{E}\rangle + \frac{\gamma}{2}\|\mathcal{Y} - \mathcal{X} - \mathcal{E}\|_F^2 + \langle\mathcal{Q},\mathcal{S} - \mathcal{E}\rangle + \frac{\sigma}{2}\|\mathcal{S} - \mathcal{E}\|_F^2 + \Phi(\mathcal{Y}),$$
$$(28)$$

where μ_k, γ, and σ are penalty parameters, and $\mathcal{Z}_k$, $\mathcal{P}$, and $\mathcal{Q}$ are Lagrangian multipliers. According to the ADMM framework [5, 38], $\mathcal{L}_k$, $\mathcal{S}$, $\mathcal{Y}$, $\mathcal{X}$, and $\mathcal{E}$ can be divided into two groups, and then the two groups of variables are updated alternately.

$$(\mathcal{L}_k^{t+1},\ \mathcal{S}^{t+1},\ \mathcal{Y}^{t+1}) = \underset{\mathcal{L}_k,\mathcal{S},\mathcal{Y}}{\arg\min}\, L(\mathcal{X}^t,\mathcal{E}^t,\mathcal{S},\mathcal{Y},\mathcal{L}_k,\mathcal{Z}_k^t,\mathcal{P}^t,\mathcal{Q}^t), \qquad (29)$$

and

$$(\mathcal{X}^{t+1},\ \mathcal{E}^{t+1}) = \underset{\mathcal{X},\mathcal{E}}{\arg\min}\, L(\mathcal{X},\mathcal{E},\mathcal{S}^{t+1},\mathcal{Y}^{t+1},\mathcal{L}_k^{t+1},\mathcal{Z}_k^t,\mathcal{P}^t,\mathcal{Q})^t. \qquad (30)$$

Now, we present more details of each subproblem.

1) Update $\mathcal{L}_k$: the $\mathcal{L}_k$ $(k = 1, 2, \cdots, \bar{N})$ subproblem can be easily transformed into its equivalent formulation:

$$\mathbf{L}_k^{t+1} = \underset{\mathcal{L}_k}{\arg\min}\, \alpha_k\|\mathbf{L}_{k[\mathbf{n}_1^k;\mathbf{n}_2^k]}\|_* + \langle\mathcal{Z}_k,\mathcal{L}_k - \mathcal{X}\rangle + \frac{\mu_k}{2}\|\mathcal{L}_k - \mathcal{X}\|_F^2$$
$$= \underset{\mathbf{L}_k}{\arg\min}\, \alpha_k\|\mathbf{L}_{k[\mathbf{n}_1^k;\mathbf{n}_2^k]}\|_* + \frac{\mu_k}{2}\|\mathbf{L}_{k[\mathbf{n}_1^k;\mathbf{n}_2^k]} - (\mathcal{X}^t - \frac{\mathcal{Z}_k^t}{\mu_k})_{[\mathbf{n}_1^k;\mathbf{n}_2^k]}\|_F^2,$$
$$(31)$$

which has the closed-form solution

$$\mathbf{L}_{k[\mathbf{n}_1^k;\mathbf{n}_2^k]}^{t+1} = \mathbf{U}\Sigma_{\alpha_k/\mu_k}\mathbf{V}^T, \qquad (32)$$

where $(\mathcal{X}^t - \frac{\mathcal{Z}_k^t}{\mu_k})_{[\mathbf{n}_1^k;\mathbf{n}_2^k]} = \mathbf{U}\Sigma\mathbf{V}^T$, $\Sigma_{\alpha_k/\mu_k} = \mathrm{diag}(\max(\Sigma_{r,r} - \alpha_k/\mu_k, 0))$, and $\Sigma_{r,r}$ is the rth singular value of Σ. $\mathcal{L}_k^{t+1} = \Upsilon(\mathbf{L}_k^{t+1})$. The computational complexity of updating $\mathbf{L}_k$ is $\mathcal{O}\big(\sum_{i=1}^{\bar{N}} p_i q_i \mathtt{min}\,(p_i, q_i)\big)$ $(p_i = \prod_{i=1}^{l} I_{\mathbf{n}_i}$ and $q_i = \prod_{i=l+1}^{N} I_{\mathbf{n}_i})$.

2) Update $\mathcal{S}$: the $\mathcal{S}$-subproblem is

$$\mathcal{S}^{t+1} = \underset{\mathcal{S}}{\arg\min}\, \lambda\|\mathcal{S}\|_1 + \langle\mathcal{Q}^t,\mathcal{S} - \mathcal{E}^t\rangle + \frac{\sigma}{2}\|\mathcal{S} - \mathcal{E}^t\|_F^2$$
$$= \underset{\mathcal{S}}{\arg\min}\, \lambda\|\mathcal{S}\|_1 + \frac{\sigma}{2}\|\mathcal{S} - \mathcal{E}^t + \frac{\mathcal{Q}^t}{\sigma}\|_F^2. \qquad (33)$$

It has the following closed-form solution:

$$\mathcal{S}^{t+1} = \mathrm{soft}_{\lambda/\sigma}(\mathcal{E}^t - \frac{\mathcal{Q}^t}{\sigma}), \qquad (34)$$

where $\mathrm{soft}_{\lambda/\sigma}(.)$ denotes the soft shrinkage operator with threshold value λ/σ. The computational complexity of updating $\mathcal{S}$ is $\mathcal{O}(\prod_{i=1}^{N} I_i)$.

3) Update $\mathcal{Y}$: the $\mathcal{Y}$ subproblem is

$$\mathcal{Y}^{t+1} = \underset{\mathcal{Y}}{\arg\min}\, \Phi(\mathcal{Y}) + \langle\mathcal{P}^t,\mathcal{Y} - \mathcal{X}^t - \mathcal{E}^t\rangle + \frac{\gamma}{2}\|\mathcal{Y} - \mathcal{X}^t - \mathcal{E}^t\|_F^2$$
$$= \underset{\mathcal{Y}}{\arg\min}\, \Phi(\mathcal{Y}) + \frac{\gamma}{2}\|\mathcal{Y} - \mathcal{X}^t - \mathcal{E}^t + \frac{\mathcal{P}^t}{\gamma}\|_F^2. \qquad (35)$$

Algorithm 2 ADMM-based algorithm for solving RC-FCTN.

Input: The observed tensor $\mathcal{Y}^0$, parameter λ.
Initialization: $t = 0$, $\mathcal{X}^0$, $\mathcal{E}^0$, $\mathcal{S}^0$, $\mathcal{L}_k^0$, Lagrangian multiplies $\mathcal{Z}_k^0$ $(k = 1, 2, \cdots, \bar{N})$,
    $\mathcal{P}$, $\mathcal{Q}$, parameters μ_k $(k = 1, 2, \cdots, \bar{N})$, γ and σ;
  1: Update $\mathcal{L}_k^{t+1}$ via (32);
  2: Update $\mathcal{S}^{t+1}$ via (34);
  3: Update $\mathcal{Y}^{t+1}$ via (36);
  4: Update $\mathcal{X}^{t+1}$ and $\mathcal{E}^{t+1}$ via (40) and (41);
  5: Update the multiplies via (42);
  6: Check convergence criteria: $\|\mathcal{X}^{t+1} - \mathcal{X}^t\|_F / \|\mathcal{X}^t\|_F \le \epsilon$;
  7: If the convergence criteria is not meet, set $t := t + 1$ and go to Step 1.
Output: The low-rank component $\mathcal{X}$ and sparse component $\mathcal{E}$.

Therefore, $\mathcal{Y}^{t+1}$ is updated via the following step:

$$\mathcal{Y}^{t+1} = \mathcal{O}_\Omega + (\mathcal{X}^t + \mathcal{E}^t - \frac{\mathcal{P}^t}{\gamma})_{\Omega^C}, \tag{36}$$

where Ω^C denotes the complementary set of Ω. The computational complexity of updating $\mathcal{Y}$ is $\mathcal{O}(\prod_{i=1}^N I_i)$.

4) Update $(\mathcal{X}, \mathcal{E})$: The $(\mathcal{X}, \mathcal{E})$ subproblem is a least squares problem

$$
\begin{aligned}
(\mathcal{X}^{t+1}, \mathcal{E}^{t+1}) &= \arg\min_{\mathcal{X},\mathcal{E}} \sum_{k=1}^{\bar{N}} \left\{ \langle \mathcal{Z}_k^t, \mathcal{L}_k^{t+1} - \mathcal{X} \rangle + \frac{\mu_k}{2} \|\mathcal{L}_k^{t+1} - \mathcal{X}\|_F^2 \right\} + \langle \mathcal{P}^t, \mathcal{Y}^{t+1} - \mathcal{X} - \mathcal{E} \rangle \\
&\quad + \frac{\gamma}{2} \|\mathcal{Y}^{t+1} - \mathcal{X} - \mathcal{E}\|_F^2 + \langle \mathcal{Q}^t, \mathcal{S}^{t+1} - \mathcal{E} \rangle + \frac{\sigma}{2} \|\mathcal{S}^{t+1} - \mathcal{E}\|_F^2 \\
&= \arg\min_{\mathcal{X},\mathcal{E}} \sum_{k=1}^{\bar{N}} \left\{ \frac{\mu_k}{2} \|\mathcal{L}_k^{t+1} - \mathcal{X} + \frac{\mathcal{Z}_k^t}{\mu_k}\|_F^2 \right\} + \frac{\gamma}{2} \|\mathcal{Y}^{t+1} - \mathcal{X} - \mathcal{E} + \frac{\mathcal{P}^t}{\gamma}\|_F^2 \\
&\quad + \frac{\sigma}{2} \|\mathcal{S}^{t+1} - \mathcal{E} + \frac{\mathcal{Q}^t}{\sigma}\|_F^2.
\end{aligned}
\tag{37}
$$

The objective function of (37) is represented by $F(\mathcal{X}, \mathcal{E})$. Taking $\partial F/\partial \mathcal{X} = 0$ and $\partial F/\partial \mathcal{E} = 0$, we have

$$\left(\sum_{k=1}^{\bar{N}} \mu_k + \gamma\right)\mathcal{X} + \gamma\mathcal{E} = \sum_{k=1}^{\bar{N}} \mu_k\left(\mathcal{L}_k^{t+1} + \frac{\mathcal{Z}_k^t}{\mu_k}\right) + \gamma\left(\mathcal{Y}^{t+1} + \frac{\mathcal{P}^t}{\gamma}\right) \tag{38}$$

and

$$\gamma\mathcal{X} + (\gamma + \sigma)\mathcal{E} = \gamma\left(\mathcal{Y}^{t+1} + \frac{\mathcal{P}^t}{\gamma}\right) + \sigma\left(\mathcal{S}^{t+1} + \frac{\mathcal{Q}^t}{\sigma}\right). \tag{39}$$

Based on the Cramer's Rule, $\mathcal{X}$ and $\mathcal{E}$ can be exactly obtained as follows:

$$\mathcal{X}^{t+1} = (\gamma\mathcal{N}^t - (\gamma + \sigma)\mathcal{M}^t)/(\gamma^2 - \left(\sum_{k=1}^{\bar{N}} \mu_k + \gamma\right)(\gamma + \sigma)) \tag{40}$$

and

$$\mathcal{E}^{t+1} = \left(\gamma\mathcal{M}^t - \left(\sum_{k=1}^{\bar{N}} \mu_k + \gamma\right)\mathcal{N}^t\right)/(\gamma^2 - \left(\sum_{k=1}^{\bar{N}} \mu_k + \gamma\right)(\gamma + \sigma)), \tag{41}$$

where $\mathcal{M}^t = \sum_{k=1}^{\bar{N}} \mu_k(\mathcal{L}_k^{t+1} + \mathcal{Z}_k^t/\mu_k) + \gamma(\mathcal{Y}^{t+1} + \mathcal{P}^t/\gamma)$ and $\mathcal{N}^t = \gamma(\mathcal{Y}^{t+1} + \mathcal{P}^t/\gamma) + \sigma(\mathcal{S}^{t+1} + \mathcal{Q}^t/\sigma)$. The computational complexity of updating $\mathcal{X}$ and $\mathcal{E}$ is $\mathcal{O}(\prod_{i=1}^N I_i)$.

Algorithm 3 SVD-based FCTN decomposition for tensor $\mathcal{X}$.

Input: Tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$, the FCTN rank.

Initialization: Core tensors $\mathcal{G}^k, k = 1, ..., N$ relate to the FCTN decomposition.

1: Choose one unfolding matrix $\mathbf{X}_{[\mathbf{n_1};\mathbf{n_2}]}$;

2: Compute the rank $\bar{R}_l$ truncated SVD of $\mathbf{X}_{[\mathbf{n_1};\mathbf{n_2}]}$: $\mathbf{X}_{[\mathbf{n_1};\mathbf{n_2}]} = \mathbf{U}_l \Sigma_l \mathbf{V}_l^T + \mathbf{E}$,
   where $d_l = l = \lfloor N/2 \rfloor, \bar{R}_l = \prod\limits_{i=1}^{l} \prod\limits_{j=l+1}^{N} R_{i,j}$, and $\mathbf{E}$ is error matrix;

3: Fold the matrix $\mathbf{U}_l$ and $\mathbf{V}_l$: $\mathcal{U}_l = \Upsilon(\mathbf{U}_l) \in \mathbb{C}^{I_1 I_2 \cdots I_l \bar{R}_l}$, $\mathcal{V}_l = \Upsilon(\mathbf{V}_l) \in$
   $\mathbb{C}^{I_{l+1} I_{l+2} \cdots I_N \bar{R}_l}$;

4: Let $d_{l-1} = \lfloor l/2 \rfloor$ and $d_{l+1} = \lfloor (N-l)/2 \rfloor$, unfold $\mathcal{U}_l$ and $\mathcal{V}_l$ with size
   $I_1 \prod\limits_{i=l+1}^{N} R_{1,i} \cdots I_{d_{l-1}} \prod\limits_{i=l+1}^{N} R_{d_{l-1},i} \times I_{d_{l-1}+1} \prod\limits_{i=l+1}^{N} R_{d_{l-1}+1,i} \cdots I_l \prod\limits_{i=l+1}^{N}, R_{l,i}$ and
   $\prod\limits_{i=1}^{l} R_{i,l+1} I_{l+1} \cdots \prod\limits_{i=1}^{l} R_{i,l+d_{l+1}} I_{l+d_{l+1}} \times \prod\limits_{i=1}^{l} R_{i,l+d_{l+1}+1} I_{l+d_{l+1}+1} \cdots \prod\limits_{i=1}^{l} R_{i,N} I_N$, re-
   spectively;

5: Compute the rank $\bar{R}_{l-1}$ and rank $\bar{R}_{l+1}$ truncated SVD of new matrix $\mathbf{U}_l$ and
   $\mathbf{V}_l$, respectively: $\mathbf{U}_l = \mathbf{U}_{l-1} \Sigma_{l-1} \mathbf{V}_{l-1}^T + \mathbf{E}$ and $\mathbf{V}_l = \mathbf{U}_{l+1} \Sigma_{l+1} \mathbf{V}_{l+1}^T + \mathbf{E}$, where
   $\bar{R}_{l-1} = \prod\limits_{i=1}^{d_{l-1}} \prod\limits_{j=d_{l-1}+1}^{l} R_{i,j}$ and $\bar{R}_{l+1} = \prod\limits_{i=l+1}^{l+d_{l+1}} \prod\limits_{j=l+d_{l+1}+1}^{N} R_{i,j}$;

6: Fold the matrix $\mathbf{U}_{l-1}$, $\mathbf{V}_{l-1}$, $\mathbf{U}_{l+1}$, and $\mathbf{V}_{l+1}$;
   $\cdots$

7: Form the corresponding factor tensors $\mathcal{G}_k$ with size $R_{1,k} \times R_{2,k} \times \cdots \times R_{k-1,k} \times I_k \times R_{k,k+1} \times \cdots \times R_{k,N}$, $(k = 1, 2, \cdots, N)$.

---

5) Update multipliers: the Lagrangian multipliers are updated as follows:

$$\begin{cases} \mathcal{Z}_k^{t+1} = \mathcal{Z}_k^t + \delta\mu_k(\mathcal{X}^{t+1} - \mathcal{L}_k^{t+1}), \ k = 1, 2, \cdots, \bar{N}, \\ \mathcal{P}^{t+1} = \mathcal{P}^t + \delta\gamma(\mathcal{Y} - \mathcal{X} - \mathcal{E}), \\ \mathcal{Q}^{t+1} = \mathcal{Q}^t + \delta\sigma(\mathcal{E} - \mathcal{S}). \end{cases} \tag{42}$$

where δ is the step length. The complexity of updating multipliers is $\mathcal{O}(\prod_{i=1}^{N} I_i)$.

The whole process of the ADMM-based algorithm for solving RC-FCTN is summarized in Algorithm 2.

Theorem 4.3. *Within the framework of ADMM, sequence $\{\mathcal{L}_k^t, \mathcal{S}^t, \mathcal{Y}^t, \mathcal{X}^t, \mathcal{E}^t\}_{t \in \mathbb{N}}$ obtained by Algorithm 2 converges to the global minimum point of the problem (25).*

The detailed proof for Theorem 4.3 is presented in Appendix.

4.2. **SVD-based FCTN decomposition.** To study the exact recovery theory of the proposed RC-FCTN, we propose a constructive SVD-based FCTN decomposition to compute the core tensors by sequential SVDs, which establishes the connection of the incoherence conditions between the proposed model and the robust matrix recovery model.

For an Nth-order tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$, its FCTN rank is

$$(R_{1,2}, R_{1,3}, \cdots, R_{1,N}, R_{2,3}, R_{2,4}, \cdots, R_{2,N}, \cdots, R_{N-1,N}).$$

Denote $\hat{R}_i = \Pi_{j=1,j\neq i}^N R_{ij}$, then it follows the results given in [33], the fully-connected graph based tensor network states can be categorized into three types:
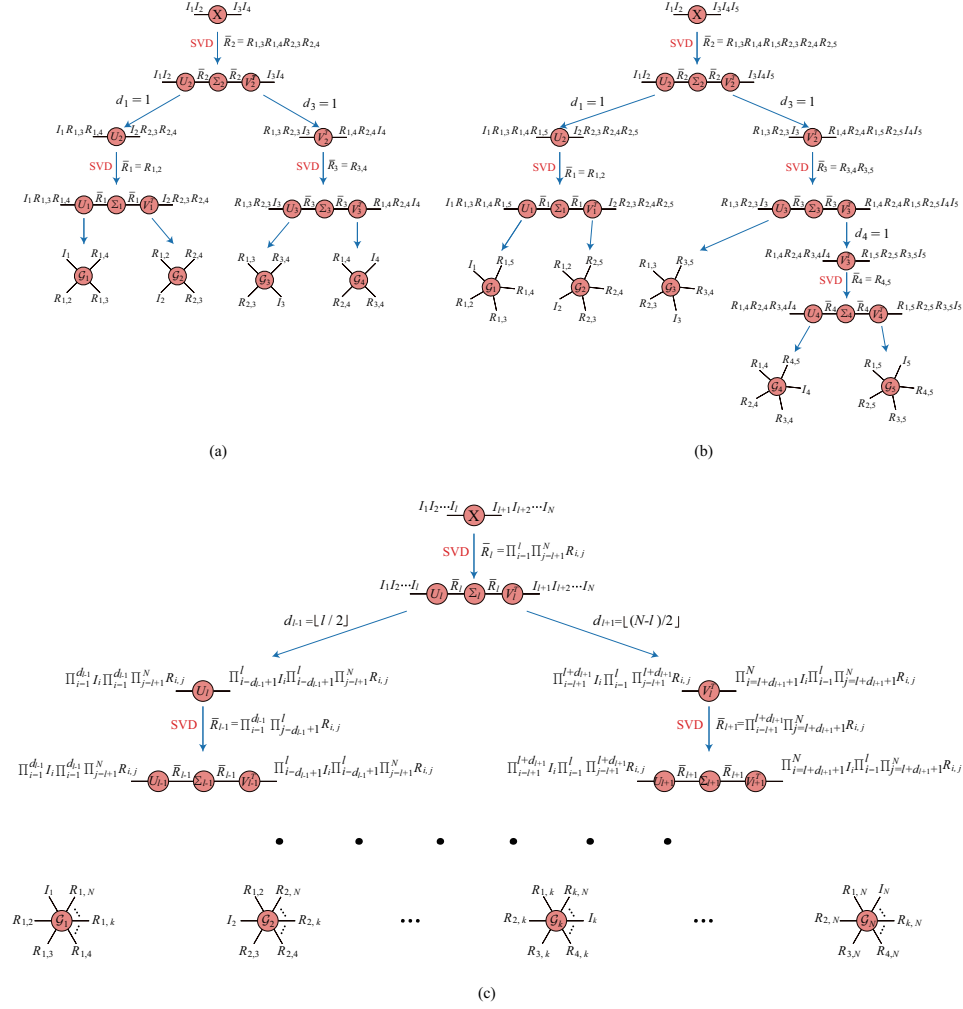
(a)

(b)

(c)

FIGURE 3. The graphical representation of the SVD-based FCTN decomposition. (a), (b), and (c) are the graphical representation of the SVD-based FCTN decomposition of a fourth-order tensor, a fifth-order tensor, and an Nth-order tensor, respectively.

subcritical, critical and supercritical. If $\hat{R}_i \leq I_i(\hat{R}_i \geq I_i), \forall i = 1, ..., N$, where at least one inequality is strict, then it is called subcritical (supercritical), if $\hat{R}_i = I_i, \forall i = 1, ..., N$, it is critical. In this paper we mainly focus on the subcritical case, since a supercritical FCTN case can be reduced to the subcritical case by a surjective birational map [33].

Let $d_l = l = \lfloor N/2 \rfloor$, $\mathbf{n}$ is the permutation of vector $(1, 2, \cdots, N)$, the whole process of SVD-based FCTN decomposition is summarized in Algorithm 3 [2]. To facilitate its understanding, we show the graphical representation of SVD-based FCTN decomposition of a fourth-order tensor, a fifth-order tensor, and an Nth-order tensor in Fig. 3.

---

[2] The SVD-based FCTN decomposition can be performed on arbitrary $\mathbf{n}$-based generalized tensor transposition. Here, for simplicity, we just take the vector $\mathbf{n} = (1, 2, \cdots, N)$ as an example.

The tensor SVD-based FCTN decomposition can be denoted as

$$\mathcal{X} = \text{FCTN-SVD}(\hat{\mathcal{G}}_1, \Sigma_1, \hat{\mathcal{G}}_2, \Sigma_2, \cdots, \Sigma_{N-1}, \hat{\mathcal{G}}_N). \tag{43}$$

When the Σ_i ($i = 1,\, 2,\, \cdots,\, N-1$) are absorbed into matrices $\mathbf{U}_i$ and $\mathbf{V}_i$ during the truncated SVD processed, we can get the FCTN decomposition $\mathcal{X} = \text{FCTN}(\{\mathcal{G}\}_{1:N})$.

4.3. FCTN incoherence conditions and exact recovery guarantee.
Based on the SVD-based FCTN decomposition, we firstly propose the FCTN incoherence conditions and then establish the exact recovery guarantee for the proposed RC-FCTN.

Theorem 4.4. *Suppose that $\mathcal{T} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ has a SVD-based FCTN decomposition*

$$\mathcal{T} = \text{FCTN-SVD}(\hat{\mathcal{G}}_1, \Sigma_1, \hat{\mathcal{G}}_2, \Sigma_2, ..., \Sigma_{N-1}, \hat{\mathcal{G}}_N), \tag{44}$$

with the FCTN-rank $(R_{1,2}, R_{1,3}, \cdots, R_{1,N}, R_{2,3}, R_{2,4}, \cdots, R_{2,N}, \cdots, R_{N-1,N})$. Then, its $\boldsymbol{n}$-based generalized tensor transposition $\vec{\mathcal{T}}^{\boldsymbol{n}}$ can be decomposes as

$$\vec{\mathcal{T}}^{\boldsymbol{n}} = \text{FCTN-SVD}(\vec{\mathcal{G}}^{\boldsymbol{n}}_{\boldsymbol{n}_1}, \vec{\Sigma}_1, \vec{\mathcal{G}}^{\boldsymbol{n}}_{\boldsymbol{n}_2}, \vec{\Sigma}_2, \cdots, \vec{\Sigma}_{N-1}, \vec{\mathcal{G}}^{\boldsymbol{n}}_{\boldsymbol{n}_N}), \tag{45}$$

and the FCTN-rank of $\vec{\mathcal{T}}^{\boldsymbol{n}}$ is $(R_{\boldsymbol{n}_1, \boldsymbol{n}_2}, R_{\boldsymbol{n}_1, \boldsymbol{n}_3}, \cdots, R_{\boldsymbol{n}_1, \boldsymbol{n}_N}, R_{\boldsymbol{n}_2, \boldsymbol{n}_3}, R_{\boldsymbol{n}_2, \boldsymbol{n}_4}, \cdots, R_{\boldsymbol{n}_2, \boldsymbol{n}_N}, \cdots, R_{\boldsymbol{n}_{N-1}, \boldsymbol{n}_N}), R_{\boldsymbol{n}_i, \boldsymbol{n}_j} = R_{\boldsymbol{n}_j, \boldsymbol{n}_i},$ if $\boldsymbol{n}_i > \boldsymbol{n}_j$.

Based on Theorem 4.4, when we suppose that $\mathcal{T}$ satisfies the following FCTN incoherent conditions for the vector $\mathbf{n} = (1, 2, \cdots, N)$, then for any vector $\mathbf{n}$, $\vec{\mathcal{T}}^{\mathbf{n}}$ satisfies the FCTN incoherent conditions. Therefore, we just consider the vector $\mathbf{n} = (1, 2, \cdots, N)$.

Definition 4.5 (FCTN Incoherence Conditions)**.** Let $\mathcal{T} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ be an Nth-order tensor represented by the SVD-based FCTN decomposition with FCTN-rank $(R_{1,2}, R_{1,3}, \cdots, R_{1,N}, R_{2,3}, R_{2,4}, \cdots, R_{2,N}, \cdots, R_{N-1,N})$, then $\mathcal{T}$ is said to satisfy the **FCTN incoherence conditions** if there exists parameters μ_i for any $\xi = \{\xi_1, \cdots, \xi_m\} \subseteq \{1, \cdots, i-1, i+1, \cdots, N\}$ ($m \leq \lfloor N/2 \rfloor$) such that

$$\max_{l=1,\cdots,L} \left\| \left(\hat{\boldsymbol{G}}_i\right)_{[\xi;\xi^c]} \cdot \boldsymbol{e}_{i,l} \right\|_F^2 \leq \frac{\mu_i \prod_{k\in\xi} R_{i,k}}{I_i \prod_{k\in\xi^c, k\neq i} R_{i,k}}, \quad i = 1, 2, \cdots, N, \tag{46}$$

where $L = I_i \prod_{k\in\xi^c, k\neq i} R_{i,k}$, $[e_{i,1}, e_{i,2}, \cdots, e_{i,L}]$ be an identity matrix of size $L \times L$, and ξ^c is the relative complement of ξ with respect to the set $\{1, 2, \cdots, N\}$.

If the generalized unfolding matrices of a tensor obey the matrix incoherent conditions, then the tensor will satisfy the FCTN incoherence conditions. Specifically, Candes and Tao listed some model matrices [7] that satisfies the matrix incoherent conditions, including uniformly bounded orthogonal model, low-rank low-coherence model, and random orthogonal model.

Based on the tensor network contraction operation given in Definition 2.1 and the tensor network incoherence conditions on core tensors given in Definition 4.5, we have the following results. For convenience, we denote $n^k_{(1)} = \Pi^l_{i=1} I_{\mathbf{n}^k_i}$, $n^k_{(2)} = \Pi^N_{i=l+1} I_{\mathbf{n}^k_i}$, $\bar{n}^k = \max(n^k_{(1)}, n^k_{(2)})$, $\hat{n}^k = \min(n^k_{(1)}, n^k_{(2)})$, and $r^k = \Pi^l_{i=1} \Pi^N_{j=l+1} R_{\mathbf{n}^k_i, \mathbf{n}^k_j}$ $(R_{\mathbf{n}^k_i, \mathbf{n}^k_j} = R_{\mathbf{n}^k_j, \mathbf{n}^k_i}, if\ \mathbf{n}^k_i > \mathbf{n}^k_j)$.

Theorem 4.6. *Suppose that $\mathcal{T}$ satisfies the FCTN incoherence conditions. Assume that the observation set Ω is uniformly distributed among all sets of cardinality $m = \rho \prod_{k=1}^{N} I_k$, where ρ is the sampling ratio. Also suppose that each observed entry is independently corrupted with probability γ. Then, there exist universal constants $c_{k1}, c_{k2} > 0$ such that with probability at least $1 - \sum_{k=1}^{\bar{N}} c_{k1}(\bar{n}^k)^{-c_{k2}}$, the recovery of underlying tensor $\mathcal{X}_0$ with $\hat{\lambda} = \sum_{k=1}^{\bar{N}} \alpha_k / \sqrt{\rho \bar{n}^k}$ is exact, provided that*

$$r^k \leq \frac{c_r \hat{n}^k}{\mu^k (\log(\bar{n}^k))^2} \quad \text{and} \quad \gamma \leq c_\gamma, k = 1, ..., \bar{N} \tag{47}$$

where $\mu^k = \max\{\Pi_{i=1}^{l}\mu_{\boldsymbol{n}_i^k}\Pi_{i=1}^{l-1}\Pi_{j=i+1}^{l}R_{\boldsymbol{n}_i^k,\boldsymbol{n}_j^k}^2, \Pi_{i=l+1}^{N}\mu_{\boldsymbol{n}_i^k}\Pi_{i=l+1}^{N-1}\Pi_{j=i+1}^{N}R_{\boldsymbol{n}_i^k,\boldsymbol{n}_j^k}^2\}$, c_r and c_γ are two positive constants.

Proof. The proof can be split into two steps. Firstly, we prove that the arbitrary unfolding matrix in (25) satisfies the matrix incoherence conditions in [6] when the original tensor satisfies the FCTN incoherence conditions given in (46). Secondly, we prove that the pair $(\mathcal{X}_0, \mathcal{E})$ derived from some convex optimization algorithms is the unique optimal solution to problem (25).

Firstly, we use a fourth-order tensor $\mathcal{T} \in \mathbb{R}^{I_1 \times I_2 \times I_3 \times I_4}$ as an example to show the process. For simplicity, set $\mathbf{n}^1 = (1, 2, 3, 4)$ and $l = \lfloor N/2 \rfloor = 2$. Then it follows the SVD-based FCTN decomposition algorithm we can get

$$\mathcal{T} = \text{FCTN-SVD}(\hat{\mathcal{G}}_1, \Sigma_1, \hat{\mathcal{G}}_2, \Sigma_2, \hat{\mathcal{G}}_3, \Sigma_3, \hat{\mathcal{G}}_4),$$

and

$$\mathbf{T}_{[1,2;3,4]} = \mathbf{U}_2 \Sigma_2 \mathbf{V}_2^H,$$

where

$$\mathbf{U}_2 = \mathbf{U}_{[1,2;3,4]} = \text{perm\&resh}((\hat{\mathbf{G}}_1)_{[1,3,4;2]}\Sigma_1(\hat{\mathbf{G}}_2)_{[1;2,3,4]}, \ I_1 I_2, \ R_{1,3}R_{1,4}R_{2,3}R_{2,4}),$$

$$\mathbf{V}_2 = \mathbf{V}_{[1,2;3,4]} = \text{perm\&resh}((\hat{\mathbf{G}}_3)_{[1,2,3;4]}\Sigma_3(\hat{\mathbf{G}}_4)_{[3;1,2,4]}, \ I_3 I_4, \ R_{1,3}R_{1,4}R_{2,3}R_{2,4}),$$

where the perm\&resh operation means that we first do the permute operation and then do the reshape operation with a specified dimension or size.

By the FCTN incoherence conditions given in (46), and choosing $i = 1, \xi = \{2\}$, we obtain

$$\max_{l=1,\cdots,I_1 R_{1,3}R_{1,4}} \|(\hat{\mathbf{G}}_1)_{[2;1,3,4]} \cdot \mathbf{e}_{1l}\|_F^2 \leq \frac{\mu_1 R_{1,2}}{I_1 R_{1,3}R_{1,4}}, \tag{48}$$

Choosing $i = 2, \xi = \{1\}$, we have

$$\max_{l=1,\cdots,I_2 R_{2,3}R_{2,4}} \|(\hat{\mathbf{G}}_2)_{[1;2,3,4]} \cdot \mathbf{e}_{2l}\|_F^2 \leq \frac{\mu_2 R_{1,2}}{I_2 R_{2,3}R_{2,4}}. \tag{49}$$

Moreover, following the SVD-based FCTN decomposition method, we have $\|\Sigma_1\| \leq \sqrt{R_{1,3}R_{1,4}R_{2,3}R_{2,4}}$. Combine above and note that

$$\|\mathbf{U}_{[1,2;3,4]}\|_\infty$$
$$= \max_{i=1,\cdots,I_1 R_{1,3}R_{1,4}, \ j=1,\cdots,I_2 R_{2,3}R_{2,4}} ((\hat{\mathbf{G}}_1)_{[2;1,3,4]})_i^T \Sigma_1 ((\hat{\mathbf{G}}_2)_{[1;2,3,4]})_j$$
$$\leq \max_{i=1,\cdots,I_1 R_{1,3}R_{1,4}, \ j=1,\cdots,I_2 R_{2,3}R_{2,4}} \|((\hat{\mathbf{G}}_1)_{[2;1,3,4]})_i^T\|\|\Sigma_1\|\|((\hat{\mathbf{G}}_2)_{[1;2,3,4]})_j\| \tag{50}$$
$$\leq \frac{\sqrt{\mu_1 \mu_2}R_{1,2}}{\sqrt{I_1 I_2}}$$

where $\mathbf{G}_i$ is the ith column of the matrix $\mathbf{G}$. Then

$$\max_{i=1,\cdots,I_1 I_2} \|\mathbf{U}_{[1,2;3,4]}^H \cdot \mathbf{e}_i\|_F^2 \leq \frac{\mu_1 \mu_2 R_{1,2}^2 R_{1,3} R_{1,4} R_{2,3} R_{2,4}}{I_1 I_2}. \tag{51}$$

Similarly, we can get

$$\max_{j=1,\cdots,I_3 I_4} \|\mathbf{V}_{[1,2;3,4]}^H \cdot \mathbf{e}_j\|_F^2 \leq \frac{\mu_3 \mu_4 R_{3,4}^2 R_{1,3} R_{1,4} R_{2,3} R_{2,4}}{I_3 I_4}. \tag{52}$$

Moreover,

$$\|\mathbf{U}_{[1,2;3,4]} \mathbf{V}_{[1,2;3,4]}^H\|_\infty \leq \sqrt{\frac{\mu_1 \mu_2 R_{1,2}^2 R_{1,3} R_{1,4} R_{2,3} R_{2,4}}{I_1 \cdot I_2}} \cdot \sqrt{\frac{\mu_3 \mu_4 R_{3,4}^2 R_{1,3} R_{2,3} R_{1,4} R_{2,4}}{I_3 \cdot I_4}}$$

$$= \frac{\sqrt{\mu_1 \mu_2 \mu_3 \mu_4} R_{1,2} R_{3,4} R_{1,3} R_{1,4} R_{2,3} R_{2,4}}{\sqrt{I_1 I_2 I_3 I_4}}. \tag{53}$$

Combine (51), (52), and (53), and set

$$\mu = \max\{\mu_1 \mu_2 R_{1,2}^2, \mu_3 \mu_4 R_{3,4}^2, \mu_1 \mu_2 \mu_3 \mu_4 R_{1,2}^2 R_{3,4}^2 R_{1,3} R_{1,4} R_{2,3} R_{2,4}\}$$

we can verify that the unfolding matrix $\mathbf{T}_{[1,2;3,4]}$ satisfies the corresponding incoherence conditions in [6]. Similarly, for any specified rearrangement $\mathbf{n}^k$, we can also verify that the generalized unfolding matrix $\mathbf{T}_{[\mathbf{n}_1^k;\mathbf{n}_2^k]}$ satisfies corresponding incoherence conditions in [6].

For an Nth-order tensor $\mathcal{T} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$, its SVD-based FCTN decomposition can be expressed as

$$\mathcal{T} = \text{FCTN-SVD}(\hat{\mathcal{G}}_1, \Sigma_1, \hat{\mathcal{G}}_2, \cdots, \Sigma_{l-1}, \hat{\mathcal{G}}_l, \Sigma_l, \hat{\mathcal{G}}_{l+1}, \cdots, \hat{\mathcal{G}}_N).$$

For simplicity, setting $\mathbf{n}^1 = (1, 2, \cdots, N)$, $l = \lfloor N/2 \rfloor$, $\mathbf{n}_1^1 = (1, \cdots, l)$, and $\mathbf{n}_2^1 = (l+1, \cdots, N)$, then,

$$\mathbf{T}_{[\mathbf{n}_1^1;\mathbf{n}_2^1]} = \mathbf{U}_l \Sigma_l \mathbf{V}_l^H.$$

For $\mathbf{U}_l$, we can get the following iteration expression

$$\mathbf{U}_2 = \text{perm\&resh}((\hat{\mathbf{G}}_1)_{[1,3,\cdots,N;2]} \Sigma_1 (\hat{\mathbf{G}}_2)_{[1;2,3,\cdots,N]}, \ I_1 I_2, \ \Pi_{i=3}^N R_{1,i} R_{2,i}),$$

$$\hat{\mathbf{U}}_2 = \text{perm\&resh}(\mathbf{U}_2, \ I_1 I_2 \Pi_{i=4}^N R_{1,i} R_{2,i}, \ R_{1,3} R_{2,3}),$$

$$\cdots,$$

$$\mathbf{U}_k = \text{perm\&resh}(\hat{\mathbf{U}}_{k-1} \Sigma_{k-1} (\hat{\mathbf{G}}_k)_{[1,\cdots,k-1;k,\cdots,N]}, \ \Pi_{i=1}^k I_i, \ \Pi_{i=1}^k \Pi_{j=k+1}^N R_{i,j}), \tag{54}$$

$$\hat{\mathbf{U}}_k = \text{perm\&resh}(\mathbf{U}_k, \ \Pi_{i=1}^k I_i \Pi_{i=1}^k \Pi_{j=k+2}^N R_{i,j}, \ \Pi_{i=1}^k R_{i,k+1}),$$

$$\cdots,$$

$$\mathbf{U}_l = \text{perm\&resh}(\hat{\mathbf{U}}_{l-1} \Sigma_{l-1} (\hat{\mathbf{G}}_l)_{[1,\cdots,l-1;l,\cdots,N]}, \ \Pi_{i=1}^l I_i, \ \Pi_{i=1}^l \Pi_{j=l+1}^N R_{i,j}).$$

Recall the incoherence conditions in (46), we have

$$\max_{l=1,\cdots,I_1 \Pi_{i=3}^N R_{1,i}} \|(\hat{\mathbf{G}}_1)_{[2;1,3,\cdots,N]} \cdot \mathbf{e}_{1l}\|_F^2 \leq \frac{\mu_1 R_{1,2}}{I_1 \Pi_{i=3}^N R_{1,i}} \tag{55}$$

and

$$\max_{l=1,\cdots,I_k \Pi_{i=k+1}^N R_{k,i}} \|(\hat{\mathbf{G}}_k)_{[1,\cdots,k-1;k,\cdots,N]} \cdot \mathbf{e}_{kl}\|_F^2 \leq \frac{\mu_k \Pi_{i=1}^{k-1} R_{1,i}}{I_k \Pi_{i=k+1}^N R_{k,i}}, k = 2, \cdots, l. \tag{56}$$

Moreover, following the SVD-based FCTN decomposition method, we have $\|\Sigma_1\| \le \sqrt{\Pi_{i=3}^{N} R_{1,i} R_{2,i}}$. Combine above and note that

$$\|\hat{\mathbf{U}}_2\|_\infty \le \frac{\sqrt{\mu_1 \mu_2} R_{1,2}}{\sqrt{I_1 I_2}},$$

$$\max_{i=1,\cdots,I_1 I_2} \|\mathbf{U}_2^H \cdot \mathbf{e}_i\|_F^2 \le \frac{\mu_1 \mu_2 R_{1,2}^2 \Pi_{i=3}^{N} R_{1,i} R_{2,i}}{I_1 I_2},$$

$$\|\hat{\mathbf{U}}_k\|_\infty \le \frac{\sqrt{\Pi_{i=1}^{k} \mu_i} \Pi_{i=1}^{k-1} \Pi_{j=i+1}^{k} R_{i,j}}{\sqrt{\Pi_{i=1}^{k} I_i}}, \ \ k = 3, 4, \cdots, l,$$

$$\max_{i=1,\cdots,\Pi_{i=1}^{k} I_i} \|\mathbf{U}_k^H \cdot \mathbf{e}_i\|_F^2 \le \frac{\Pi_{i=1}^{k} \mu_i \Pi_{i=1}^{k-1} \Pi_{j=i+1}^{k} R_{i,j}^2 \Pi_{i=1}^{k} \Pi_{j=k+1}^{N} R_{i,j}}{\Pi_{i=1}^{k} I_i}, \ \ k = 3, 4, \cdots, l.$$

$$(57)$$

Then, we have

$$\max_{i=1,\cdots,\Pi_{i=1}^{l} I_i} \|\mathbf{U}_l^H \cdot \mathbf{e}_i\|_F^2 \le \frac{\Pi_{i=1}^{l} \mu_i \Pi_{i=1}^{l-1} \Pi_{j=i+1}^{l} R_{i,j}^2 r}{\Pi_{i=1}^{l} I_i}, \tag{58}$$

$$\max_{j=1,\cdots,\Pi_{i=l+1}^{N} I_i} \|\mathbf{V}_l^H \cdot \mathbf{e}_j\|_F^2 \le \frac{\Pi_{i=l+1}^{N} \mu_i \Pi_{i=l+1}^{N-1} \Pi_{j=i+1}^{N} R_{i,j}^2 r}{\Pi_{i=l+1}^{N} I_i}. \tag{59}$$

Moreover,

$$\|\mathbf{U}_l \mathbf{V}_l^H\|_\infty \le \sqrt{\frac{\Pi_{i=1}^{N} \mu_i \Pi_{i=1}^{l-1} \Pi_{j=i+1}^{l} R_{i,j}^2 \Pi_{i=l+1}^{N-1} \Pi_{j=i+1}^{N} R_{i,j}^2 r}{\Pi_{i=1}^{N} I_i}},$$

where $r = \Pi_{i=1}^{l} \Pi_{j=l+1}^{N} R_{i,j}$. To sum up, the unfolding matrix $\mathbf{T}_{[\mathbf{n}_1^1; \mathbf{n}_2^1]}$ satisfies the corresponding incoherence conditions of robust matrix recovery problems in [6]. By using the same procedure as the proof of vector $(1, 2, \cdots, n)$, for any specified rearrangement $\mathbf{n}^k$, we can also verify that the generalized unfolding matrix $\mathbf{T}_{[\mathbf{n}_1^k; \mathbf{n}_2^k]}$ satisfies the corresponding incoherence conditions in [6].

Secondly, the convex model (25) can be regarded as a convex combination of $\bar{N}$ robust matrix completion models. Invoking the Theorem 1.1 in [6] and recalling that the incoherence conditions of the unfolding matrices that appeared in (25) are satisfied, we can prove the Theorem 4.6. □

This exact recovery guarantee ensures that one can recover a tensor of low-FCTN-rank exactly with overwhelming probability provided that its rank is sufficiently small and its corrupted entries are reasonably sparse.

5. **Numerical experiments.** In this section, we firstly conduct the RTC experiments on synthetic data in subsection 5.1, which further corroborates our theoretical results. In subsection 5.3-5.5, we conduct numerical experiments on fourth-order color videos and hyperspectral videos (HSV) to verify the effectiveness of the proposed RC-FCTN and RNC-FCTN. To adequately examine the recovery performance of RC-FCTN and RNC-FCTN, we compare the proposed methods with four representative RTC methods: Tucker rank based method [13] (denoted as "SNN"), tubal rank based method [20] (denoted as "TNN"), TT rank based method [32] (denoted as "TTNN"), and TR rank based method [14] (denoted as "RTRC"). For the fourth-order data $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times I_3 \times I_4}$, let $I = \prod_{i=1}^{4} I_i$, $R_1 = R_{1,2} R_{1,3} R_{1,4}$, $R_2 = R_{1,2} R_{2,3} R_{2,4}$, $R_3 = R_{1,3} R_{2,3} R_{3,4}$, and $R_4 = R_{1,4} R_{2,4} R_{3,4}$, we summarize the computational complexity of each iteration of the presented algorithms and the

TABLE 2. The comparison of the computational complexity of the proposed methods and compared methods.

Method	The computational complexity of each step
SNN	$\mathcal{O}\big(\sum_{k=1}^{4} p_k q_k \ \min\ (p_k, q_k)\big) \ \Big(p_k = I_k \text{ and } q_k = \prod_{i=1, i \neq k}^{4} I_i\Big)$
TNN	$\mathcal{O}\big(I_1 I_2 log(I_3 I_4)\big) + \min(I_1, I_2) \prod_{i=1}^{4} I_i\big)$
TTNN	$\mathcal{O}\big(\sum_{k=1}^{4} p_k q_k \ \min\ (p_k, q_k)\big) \ \Big(p_k = \prod_{i=1}^{k} I_i \text{ and } q_i = \prod_{i=k+1}^{4} I_i\Big)$
RTRC	$\mathcal{O}\big(\sum_{k=1}^{2} p_k q_k \ \min\ (p_k, q_k)\big) \ \Big(p_k = \prod_{i=k}^{k+1} I_i \text{ and } q_i = \prod_{i=k+2}^{k+3} I_i \text{ with } I_5 = I_1\Big)$
RC-FCTN	$\mathcal{O}\big(\sum_{k=1}^{3} p_k q_k \ \min\ (p_k, q_k)\big) \ \Big(p_k = \prod_{i=1}^{l} I_{\mathbf{n}_i} \text{ and } q_k = \prod_{i=l+1}^{4} I_{\mathbf{n}_i}\Big)$
RNC-FCTN	$\mathcal{O}(4 \sum_{k=2}^{4} I^k R^{k(4-k)+k-1} + 4I^3 R^6)$

compared ones in Table 2. For the tensor nuclear norm-based convex methods (including SNN, TTNN, RTRC, and RC-FCTN), their computational complexities are mainly related to the sizes of the corresponding unfolding matrices (e.g., $X_{(k)}$, $X_{[k]}$, $\mathbf{X}_{<k,L>}$, and $\mathbf{X}_{[\mathbf{n}_{1:d};\mathbf{n}_{d+1:N}]}$). Clearly, the computational complexity of RC-FCTN is higher than that of RTRC. For other methods, we cannot theoretically compare their computational complexities as the output of the min function is different for different dimension sizes combination.

For all methods in all experiments of subsection 5.3-5.5, we employ the following setup to make a fair comparison: (i) The data is normalized into [0, 1]; (ii) The relative error is set to 10^{-4}; (iii) We utilize a simple linear interpolation strategy [30] to obtain the $\mathcal{Y}^0$. We employ the mean of peak signal-to-noise rate (MPSNR) and the mean of structural similarity (MSSIM) as the initial tensor quantitative metric [43]. The parameters in compared methods are manually adjusted to the optimal performance, which refers to the discussion in their articles. Meanwhile, the best and the second-best results are highlighted by bold and underline, respectively.

5.1. **Synthetic tensor completion.** To verify the validity of exact recovery in Theorem 4.6, we firstly execute experiments on synthetic data. We simulate the tensor of size $I \times I \times I \times I$ by FCTN contraction with varying dimensions I=20 and 40. We generate the tensor $\mathcal{X}^0 = \text{FCTN}(\mathcal{G}_1, \mathcal{G}_2, \mathcal{G}_3, \mathcal{G}_4)$ with FCTN rank $R_{i,j}$ $(1 \leq i < j \leq 4, \text{ and } i, j \in \mathbb{N}^+)$ as the same value r. The core tensors are independently satisfied the uniform distribution $\mathcal{U}(0, 1)$. Then, the whole entries are corrupted by salt and pepper (SaP) noise with density s. And we choose the observation with sampling ratio (SR) ρ. We test the recovery ability of the proposed RC-FCTN on eight cases and show the results in Table 3. We design $r = 0.1I$ and $0.2I$, $\rho = 0.9$ and 0.8, and $s = 0.05$ and 0.1. As observed, our method obtains the negligible relative error $\|\mathcal{X} - \mathcal{X}^0\|_F / \|\mathcal{X}^0\|_F$. These numerical results of all cases validly corroborate the exact recovery in Theorem 4.6 well.

5.2. **Parameter setting.** In RNC-FCTN model, the important parameters are regularization parameter λ and FCTN-rank while in RC-FCTN model, the important parameter is regularization parameter λ. Now, we discuss the influence of parameters on the experiment results and provide the parameters selection strategy.

TABLE 3. Exact recovery on random synthetic data in different cases.

size I	rank r	ρ	s	$\|\mathcal{X} - \mathcal{X}^0\|_F/\|\mathcal{X}^0\|_F$	size I	rank r	ρ	s	$\|\mathcal{X} - \mathcal{X}^0\|_F/\|\mathcal{X}^0\|_F$
20	0.1I	1	0.05	1.28×10^{-4}	40	0.1I	1	0.05	1.06×10^{-4}
			0.1	1.56×10^{-4}				0.1	1.12×10^{-4}
		0.9	0.05	1.93×10^{-4}			0.9	0.05	1.72×10^{-4}
			0.1	1.96×10^{-4}				0.1	2.25×10^{-4}
	0.2I	1	0.05	4.74×10^{-4}		0.2I	1	0.05	1.77×10^{-4}
			0.1	7.83×10^{-4}				0.1	2.88×10^{-4}
		0.9	0.05	6.55×10^{-4}			0.9	0.05	2.66×10^{-4}
			0.1	8.35×10^{-4}				0.1	3.01×10^{-4}
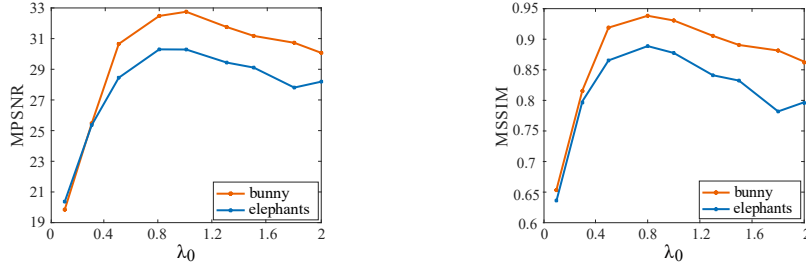
FIGURE 4. The MPSNR and MSSIM values recovered by RNC-FCTN with respect to λ_0 for two color videos.
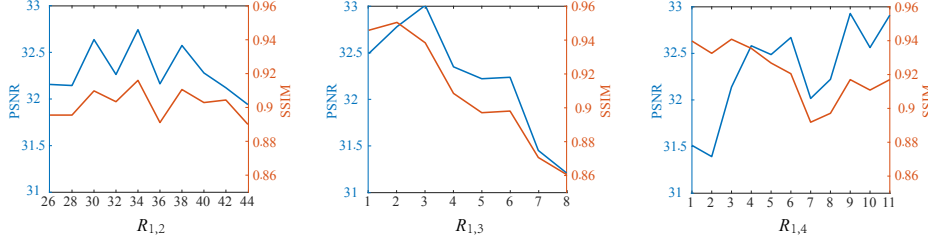
FIGURE 5. The MPSNR and MSSIM values recovered by RNC-FCTN with respect to the FCTN-rank for color video *bunny*.

1) **Influence of regularization parameter λ in RNC-FCTN:** The parameter λ actually balances the low-FCTN-rank term and the sparse term. By following [15, 25, 14], we set $\lambda = \lambda_0/\sqrt{\max(I_1, I_2)I_3I_4}$, and thus we discuss the influence of λ_0, which is proportionate to λ. We display the MPSNR and MSSIM values recovered by RNC-FCTN with respect to λ_0 for two color videos with SR=0.2 and SaP=0.1 in Fig. 4. As observed, the method RNC-FCTN achieves the optimal performance when λ_0 is around 1. Inspired by this observation, the parameter λ_0 can be selected from the candidate set $\{0.4, 0.6, 0.8, 1, 1.2, 1.4, 1.6\}$.

2) **Influence of FCTN-rank in RNC-FCTN:** By following [44], we just tune $R_{1,2}^{\max}$, $R_{1,3}^{\max}$, and $R_{1,4}^{\max}$, where $R_{2,4}^{\max}$ is set to be the same as $R_{1,4}^{\max}$, and $R_{2,3}^{\max}$ and
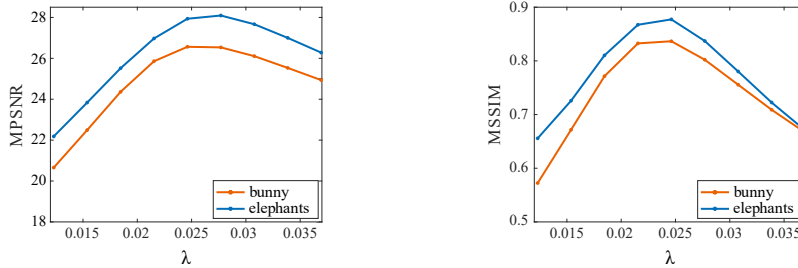
FIGURE 6. The MPSNR and MSSIM values recovered by RC-FCTN with respect to λ for two color videos.

$R_{3,4}^{\max}$ are set to be the same as $R_{1,3}^{\max}$. Fig. 5 shows the MPSNR and MSSIM values recovered by RNC-FCTN with respect to the FCTN-rank for color video *bunny* with SR=0.2 and SaP=0.1. When the FCTN-rank reaches the suitable value, we can achieve optimal performance. Motivated by this observation, parameters $R_{1,2}^{\max}$, $R_{1,3}^{\max}$, and $R_{1,4}^{\max}$ are suggested to select from the candidate sets $\{20, 25, 30, 35, 40, 45, 50\}$, $\{1, 3, 5, 7\}$, and $\{3, 6, 9, 12\}$, respectively.

3) **Influence of regularization parameter λ in RC-FCTN:** For the proposed method RC-FCTN, the theoretical optimal value for the regularization parameter λ is given in Theorem 4.6. In all real experiments, we finely tune the regularization parameter λ from the mesh grids (e.g., 0.024, 0.025, and 0.026) around the theoretical optimal value (e.g., 0.238 for the color videos with SR=0.2 and SaP=0.1) to obtain the best MPSNR and MSSIM results. We display the MPSNR and MSSIM values recovered by RC-FCTN with respect to λ for two color videos with SR=0.2 and SaP=0.1 in Fig. 6. As observed, it achieves the optimal performance when λ is around 0.025, which is slightly higher than the theoretical optimal value 0.0238.

5.3. **Color video completion.** To demonstrate the effectiveness of the proposed methods, we execute experiments on two color videos [3] (height $\times$ width $\times$ color channel $\times$ frames) including *bunny* and *elephants*. We consider RTC problem for the testing color videos with SaP=0.1 and different SRs $\{0.6, 0.4, 0.2\}$.

We report the MPSNR/MSSIM values, iteration number, and CPU time obtained by all compared RTC methods on the videos *bunny* and *elephants* in Table 4. Both convex method RC-FCTN and nonconvex method RNC-FCTN can achieve superior performance compared with the other methods although they need more CPU time. Moreover, the nonconvex RNC-FCTN requires more iterations than the convex RC-FCTN, which leads to a significant difference in CPU time. In RC-FCTN model, we consider the convex relaxation of the FCTN-rank, i.e., the FCTN nuclear norm, which allows us to apply the convex solver. However, the FCTN nuclear norm cannot exactly approximate the FCTN-rank. In RNC-FCTN model, we directly consider the FCTN-rank, which leads to a better recovery performance as compared to RC-FCTN.

Furthermore, Fig. 7 shows the visual results and their corresponding residual images (the mean of the absolute difference between three color channels of the recovered images and the ground truth) of the two color videos with SR=0.2 and SaP=0.1. From Fig. 7, it is easy to see that the proposed methods are markedly

---

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

TABLE 4. The MPSNR/MSSIM, iteration number, and CPU time (seconds) of the recovered color videos by different methods for different cases.

Data	Methods	SaP=0.1, SR=0.6				SaP=0.1, SR=0.4				SaP=0.1, SR=0.2			
		MPSNR	MSSIM	Iter	Time	MPSNR	MSSIM	Iter	Time	MPSNR	MSSIM	Iter	Time
bunny	Observed	9.357	0.095	-	-	8.035	0.062	-	-	7.024	0.033	-	-
	SNN	26.80	0.832	128	83.5	22.61	0.653	142	87.3	14.92	0.415	143	85.3
	TNN	34.61	0.955	**34**	**22.9**	30.59	0.925	**34**	**22.6**	25.67	0.774	**34**	**22.1**
	TTNN	32.86	0.966	<u>60</u>	27.8	29.80	0.937	<u>66</u>	29.3	23.90	0.764	74	33.7
	RTRC	34.05	0.962	66	<u>26.7</u>	30.87	0.924	67	<u>28.5</u>	<u>26.60</u>	0.835	<u>72</u>	<u>30.1</u>
	RC-FCTN	<u>34.92</u>	<u>0.975</u>	85	60.1	<u>31.86</u>	<u>0.946</u>	87	62.1	26.56	<u>0.837</u>	92	64.4
	RNC-FCTN	**39.72**	**0.983**	643	849.7	**37.11**	**0.977**	656	856.1	**32.85**	**0.941**	676	880.8
elephants	Observed	7.146	0.049	-	-	5.631	0.031	-	-	4.516	0.017	-	-
	SNN	28.80	0.863	141	88.7	24.21	0.759	143	89.9	14.45	0.516	144	85.7
	TNN	34.35	0.958	**32**	**22.2**	31.05	0.928	**33**	**22.5**	26.28	0.811	**32**	**21.8**
	TTNN	32.61	0.943	<u>62</u>	28.1	29.75	0.933	<u>68</u>	30.5	25.42	0.845	<u>67</u>	<u>29.4</u>
	RTRC	33.69	0.961	66	<u>27.2</u>	30.52	0.929	<u>68</u>	<u>28.5</u>	26.31	0.847	75	32.3
	RC-FCTN	<u>36.99</u>	<u>0.974</u>	83	62.3	<u>33.25</u>	<u>0.956</u>	87	62.1	<u>27.94</u>	<u>0.877</u>	94	66.4
	RNC-FCTN	**39.18**	**0.975**	651	841.7	**36.20**	**0.960**	663	850.0	**29.91**	**0.888**	669	889.9
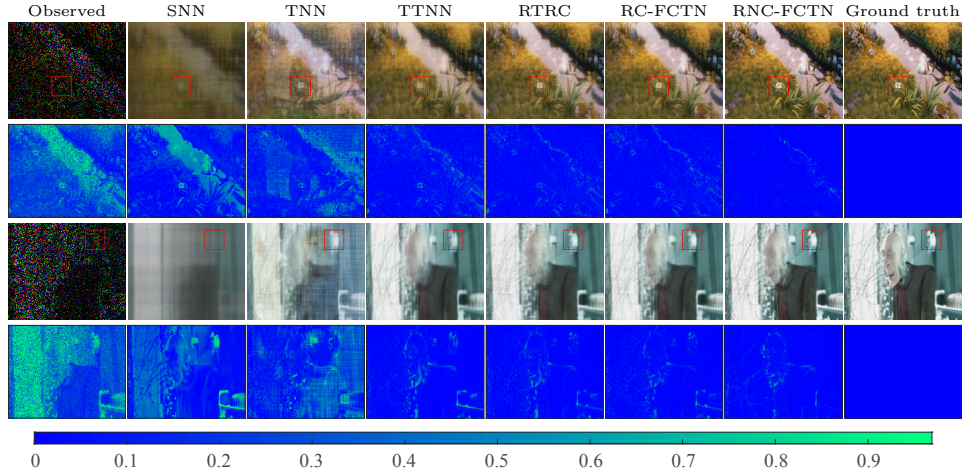
FIGURE 7. Recovered results on two color videos with SR=0.2 and SaP=0.1. The first row and third row are visual results at the 1st frame of *bunny* and the 50th frame of *elephants*, respectively. The second row and fourth row are the corresponding residual images.

superior in removing noise and preserving details than the compared ones, such as grasses in the *bunny* and construction in the *elephants*.

5.4. **Hyperspectral video completion.** To verify the effectiveness of the proposed methods, we conduct experiments on the hyperspectral video (HSV) [4] (height

---

[4]The data is available at http://openremotesensing.net/kb/data/.

TABLE 5. The MPSNR/MSSIM, iteration number, and CPU time (seconds) of the recovered HSV by different methods for different cases.

| Methods	SaP=0.1, SR=0.3				SaP=0.1, SR=0.2				SaP=0.1, SR=0.1			
	MPSNR	MSSIM	Iter	Time	MPSNR	MSSIM	Iter	Time	MPSNR	MSSIM	Iter	Time
Observed	9.367	0.071	-	-	8.936	0.048	-	-	8.545	0.026	-	-
SNN	26.89	0.872	94	24.9	24.12	0.790	97	25.6	19.23	0.628	98	25.8
TNN	43.23	0.993	**36**	**12.4**	37.77	0.983	**36**	**12.6**	27.72	0.893	**34**	**11.2**
TTNN	45.71	0.995	<u>89</u>	<u>28.7</u>	43.18	0.993	<u>94</u>	<u>30.7</u>	36.68	0.985	<u>94</u>	<u>29.2</u>
RTRC	45.73	<u>0.996</u>	93	33.0	42.70	<u>0.994</u>	95	32.1	36.56	0.983	99	36.4
RC-FCTN	<u>46.64</u>	**0.997**	109	67.2	<u>43.98</u>	**0.995**	110	67.7	<u>37.88</u>	**0.988**	112	66.4
RNC-FCTN	**46.91**	0.997	436	162.2	**44.38**	0.995	443	159.0	**39.20**	<u>0.986</u>	446	159.9
Methods	SaP=0.2, SR=0.3				SaP=0.2, SR=0.2				SaP=0.2, SR=0.1			
	MPSNR	MSSIM	Iter	Time	MPSNR	MSSIM	Iter	Time	MPSNR	MSSIM	Iter	Time
Observed	9.015	0.055	-	-	8.716	0.038	-	-	8.448	0.022	-	-
SNN	24.24	0.759	89	23.1	21.32	0.722	94	24.8	17.00	0.527	98	25.1
TNN	39.75	0.987	**36**	**12.4**	34.37	0.969	**36**	**12.2**	25.14	0.821	**34**	**11.0**
TTNN	43.43	0.994	<u>72</u>	<u>24.7</u>	<u>40.20</u>	**0.993**	<u>82</u>	<u>26.7</u>	34.38	0.971	<u>92</u>	<u>29.6</u>
RTRC	42.62	0.994	83	30.6	39.09	0.991	87	31.7	33.65	0.972	96	35.8
RC-FCTN	<u>43.79</u>	**0.996**	110	68.7	39.62	0.991	112	69.9	<u>34.49</u>	<u>0.974</u>	112	67.1
RNC-FCTN	**44.73**	<u>0.995</u>	421	149.7	**41.93**	<u>0.992</u>	423	155.2	**36.02**	**0.975**	426	157.8
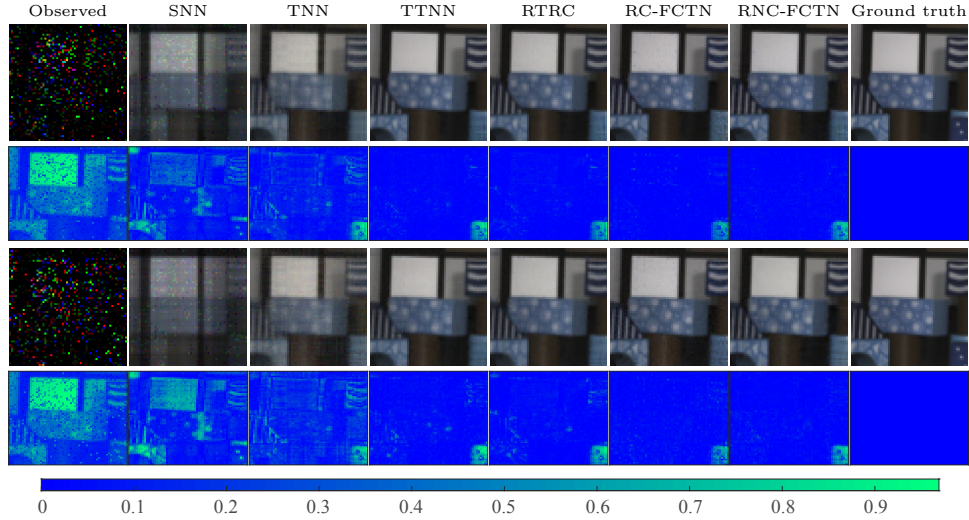
FIGURE 8. Recovered results on the HSV (band 8, 9, and 10 of first frame are picked as red, green, and blue channels). The first row and third row are visual results with SR=0.1, SaP=0.1 and 0.2, respectively. The second row and fourth row are the corresponding residual images.

$\times$ width $\times$ band $\times$ frames) of size $60 \times 60 \times 20 \times 20$. HSV contains a wealth of information, thus we consider more challenging situations. We conduct RTC problem for the HSV with different SRs $\{0.3, 0.2, 0.1\}$ and SaP $\{0.1, 0.2\}$.

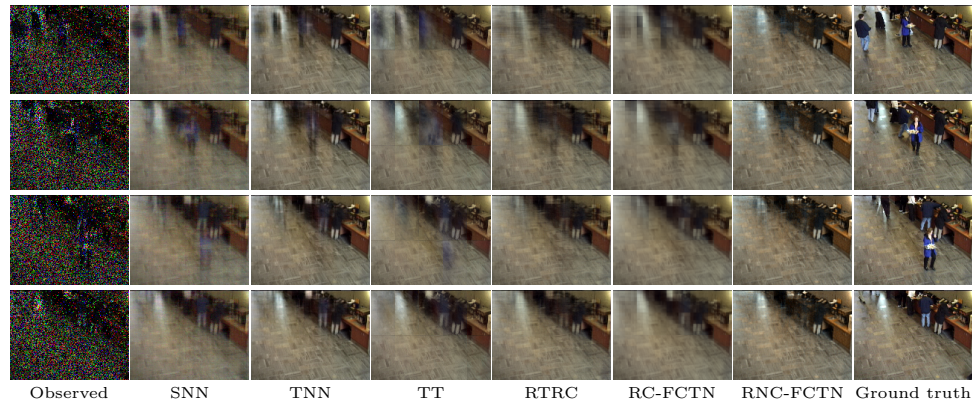|							
Observed	SNN	TNN	TT	RTRC	RC-FCTN	RNC-FCTN	Ground truth

FIGURE 9. The visual results of the 17th frame, the 27th frame, the 37th frame, and the 47th frame of six robust competition methods.

We show the MPSNR/MSSIM values, iteration number, and CPU time obtained by all compared restoration methods on the HSV in Table 5. The results indicates that the proposed methods obtain an overall better performance than the compared methods. As the SaP increases, the advantage of the proposed RNC-FCTN over the compared methods is more prominent. Furthermore, Fig. 8 shows the visual results and their corresponding residual images of the HSV with SR=0.1, SaP= 0.1 and 0.2, respectively. For visual effect, we have selected three bands to show pseudo-color images. From Fig. 8, the SaP in the reconstructed results of SNN and TNN is not removed well. The results reconstructed by the proposed methods are closer to the real image in terms of color than those of compared methods. The corresponding residual images can clearly confirm this phenomenon.

5.5. **Video background subtraction.** In this subsection, we apply the two proposed methods to color video background subtraction. We pick up consecutive 50 frames of *bootstrap* [5] which forms a $120 \times 160 \times 3 \times 50$ tensor. This video consists of a static architectural background and moving foreground, such as moving people. Since the background components of all frames are highly correlated, it can be regarded as a low-rank tensor. The foreground components occupy a few locations of the entire video, and it can be regarded as a sparse tensor.

We design a challenging task that executes a background subtraction from a corrupted video with SR=0.4 and SaP=0.1. Fig. 9 presents the visual comparison of the 17th frame, the 27th frame, the 37th frame, and the 47th frame of six robust competition methods. As observed, the proposed methods, especially RNC-FCTN, simultaneously extract background and preserve the global structure in completing the missing entries. The reason is that the FCTN rank can be flexibly adjusted to separate the low-rank background and sparse foreground well.

6. **Conclusion.** In this paper, we firstly proposed a robust nonconvex optimization model RNC-FCTN for the RTC problem. Then, we theoretically derive the convergence guarantee of the PAM-based algorithm. Moreover, we suggested an FCTN nuclear norm as a convex surrogate of FCTN rank. Based on the FCTN

---

[5]The data is available https://www.microsoft.com/en-us/download/details.aspx?id=54651.

nuclear norm, we propose a robust convex optimization model RC-FCTN for the RTC problem. Then, we theoretically establish the exact recovery conditions that one can recover a tensor of low-FCTN-rank exactly with overwhelming probability provided that its rank is sufficiently small and its corrupted entries are reasonably sparse. We develop an ADMM-based algorithm to solve the proposed RC-FCTN, which enjoys the global convergence guarantee. Experimental results demonstrate the usefulness of proposed methods with compared ones.

Appendix A. Proof of Theorem 3.1. Here, we exhibit the detailed proof of Theorem 3.1.

Proof. To prove it, we mainly demonstrate that the proposed RNC-FCTN satisfies the following three conditions:

(1) $f(\{\mathcal{F}\}_{1:N}, \mathcal{X}, \mathcal{E}, \mathcal{Y})$ is a proper lower semi-continuous function.

(2) $f(\{\mathcal{F}\}_{1:N}, \mathcal{X}, \mathcal{E}, \mathcal{Y})$ satisfies the K-L property [3] at each $\{\{\mathcal{F}\}_{1:N}^t, \mathcal{X}^t, \mathcal{E}^t, \mathcal{Y}^t\}$.

(3) The bounded sequence $\{\{\mathcal{F}\}_{1:N}^t, \mathcal{X}^t, \mathcal{E}^t, \mathcal{Y}^t\}_{t\in\mathbb{N}}$ satisfies the sufficient decrease and relative error conditions.

For convenience, we rewrite the objective function as

$$f(\{\mathcal{F}\}_{1:N}, \mathcal{X}, \mathcal{E}, \mathcal{Y}) = f_1(\{\mathcal{F}\}_{1:N}, \mathcal{X}, \mathcal{E}, \mathcal{Y}) + f_2(\mathcal{E}) + \Phi(\mathcal{Y}). \qquad (60)$$

where $f_1(\{\mathcal{F}\}_{1:N}, \mathcal{X}, \mathcal{E}, \mathcal{Y}) = \frac{1}{2}\|\mathcal{X} - \text{FCTN}(\{\mathcal{F}\}_{1:N})\|_F^2 + \frac{\beta}{2}\|\mathcal{Y} - \mathcal{X} - \mathcal{E}\|_F^2$ and $f_2(\mathcal{E}) = \lambda\|\mathcal{E}\|_1$.

Thus, PAM-based algorithm is updated as the following iterative scheme:

$$\begin{cases} \mathcal{F}_k^{t+1} = \underset{\mathcal{F}_k}{\arg\min}\, f(\{\mathcal{F}\}_{1:k-1}^{t+1}, \mathcal{F}_k, \{\mathcal{F}\}_{k+1:N}^t, \mathcal{X}^t, \mathcal{E}^t, \mathcal{Y}^t) + \frac{\rho}{2}\|\mathcal{F}_k - \mathcal{F}_k^t\|_F^2, \\ \mathcal{X}^{t+1} = \underset{\mathcal{X}}{\arg\min}\, f(\{\mathcal{F}\}_{1:N}^{t+1}, \mathcal{X}, \mathcal{E}^t, \mathcal{Y}^t) + \frac{\rho}{2}\|\mathcal{X} - \mathcal{X}^t\|_F^2, \\ \mathcal{E}^{t+1} = \underset{\mathcal{E}}{\arg\min}\, f(\{\mathcal{F}\}_{1:N}^{t+1}, \mathcal{X}^{t+1}, \mathcal{E}, \mathcal{Y}^t) + \frac{\rho}{2}\|\mathcal{E} - \mathcal{E}^t\|_F^2, \\ \mathcal{Y}^{t+1} = \underset{\mathcal{Y}}{\arg\min}\, f(\{\mathcal{F}\}_{1:N}^{t+1}, \mathcal{X}^{t+1}, \mathcal{E}^{t+1}, \mathcal{Y}) + \frac{\rho}{2}\|\mathcal{Y} - \mathcal{Y}^t\|_F^2. \end{cases}$$

$$(61)$$

Now, we prove that the three key conditions are holded respectively.

Firstly, it is easy to verify that f_1 is a C^1 function with locally Lipschitz continuous gradient, and f_2 and $\Phi(\mathcal{Y})$ are proper and lower semi-continuous functions. Therefore, $f(\{\mathcal{F}\}_{1:N}, \mathcal{X}, \mathcal{E}, \mathcal{Y})$ is a proper lower semi-continuous function.

Secondly, since the semi-algebraic real-valued function satisfies the K-L property [3], we only need to illustrate that $f(\{\mathcal{F}\}_{1:N}, \mathcal{X}, \mathcal{E}, \mathcal{Y})$ is a semi-algebraic function. $f_1(\{\mathcal{F}\}_{1:N}, \mathcal{X}, \mathcal{E}, \mathcal{Y})$, $f_2(\mathcal{E})$, and $\Phi(\mathcal{Y})$ are the sum of Frobenius norm, l_1-norm, and indicator function, respectively. It is easy to identity that they are semi-algebraic functions [4]. As the sum of three semi-algebraic functions, $f(\{\mathcal{F}\}_{1:N}, \mathcal{X}, \mathcal{E}, \mathcal{Y})$

is still a semi-algebraic function. Therefore, $f(\{\mathcal{F}\}_{1:N}, \mathcal{X}, \mathcal{E}, \mathcal{Y})$ satisfies the K-Ł property at each $\{\{\mathcal{F}\}_{1:N}^t, \mathcal{X}^t, \mathcal{E}^t, \mathcal{Y}^t\}$.

Thirdly, we prove that the bounded sequence $\{\{\mathcal{F}\}_{1:N}^t, \mathcal{X}^t, \mathcal{E}^t, \mathcal{Y}^t\}_{t\in\mathbb{N}}$ satisfies the sufficient decrease and relative error conditions, respectively.

Lemma 6.1. *(**Sufficient decrease**) Suppose $\{\{\mathcal{F}\}_{1:N}^t, \mathcal{X}^t, \mathcal{E}^t, \mathcal{Y}^t\}_{t\in\mathbb{N}}$ be the sequence obtained by Algorithm 1, then it satisfies*

$$
\begin{cases}
f(\{\mathcal{F}\}_{1:k}^{t+1}, \{\mathcal{F}\}_{k+1:N}^t, \mathcal{X}^t, \mathcal{E}^t, \mathcal{Y}^t) + \frac{\rho}{2}\|\mathcal{F}_k^{t+1} - \mathcal{F}_k^t\|_F^2 \leq f(\{\mathcal{F}\}_{1:k-1}^{t+1}, \{\mathcal{F}\}_{k:N}^t, \mathcal{X}^t, \mathcal{E}^t, \mathcal{Y}^t), \\
f(\{\mathcal{F}\}_{1:N}^{t+1}, \mathcal{X}^{t+1}, \mathcal{E}^t, \mathcal{Y}^t) + \frac{\rho}{2}\|\mathcal{X}^{t+1} - \mathcal{X}^t\|_F^2 \leq f(\{\mathcal{F}\}_{1:N}^{t+1}, \mathcal{X}^t, \mathcal{E}^t, \mathcal{Y}^t), \\
f(\{\mathcal{F}\}_{1:N}^{t+1}, \mathcal{X}^{t+1}, \mathcal{E}^{t+1}, \mathcal{Y}^t) + \frac{\rho}{2}\|\mathcal{E}^{t+1} - \mathcal{E}^t\|_F^2 \leq f(\{\mathcal{F}\}_{1:N}^{t+1}, \mathcal{X}^{t+1}, \mathcal{E}^t, \mathcal{Y}^t), \\
f(\{\mathcal{F}\}_{1:N}^{t+1}, \mathcal{X}^{t+1}, \mathcal{E}^{t+1}, \mathcal{Y}^{t+1}) + \frac{\rho}{2}\|\mathcal{Y}^{t+1} - \mathcal{Y}^t\|_F^2 \leq f(\{\mathcal{F}\}_{1:N}^{t+1}, \mathcal{X}^{t+1}, \mathcal{E}^{t+1}, \mathcal{Y}^t).
\end{cases}
\tag{62}
$$

We give the proof of Lemma 6.1. Let $\{\mathcal{F}\}_{1:N}^{t+1}, \mathcal{X}^{t+1}, \mathcal{E}^{t+1}$, and $\mathcal{Y}^{t+1}$ are the optimal solutions of $\mathcal{F}_k$-subproblem, $\mathcal{X}$-subproblem, $\mathcal{E}$-subproblem, and $\mathcal{Y}$-subproblem, we have

$$
\begin{aligned}
&f(\{\mathcal{F}\}_{1:k}^{t+1}, \{\mathcal{F}\}_{k+1:N}^t, \mathcal{X}^t, \mathcal{E}^t, \mathcal{Y}^t) + \frac{\rho}{2}\|\mathcal{F}_k^{t+1} - \mathcal{F}_k^t\|_F^2 \leq f(\{\mathcal{F}\}_{1:k-1}^{t+1}, \{\mathcal{F}\}_{k:N}^t, \mathcal{X}^t, \mathcal{E}^t, \mathcal{Y}^t), \\
&f(\{\mathcal{F}\}_{1:N}^{t+1}, \mathcal{X}^{t+1}, \mathcal{E}^t, \mathcal{Y}^t) + \frac{\rho}{2}\|\mathcal{X}^{t+1} - \mathcal{X}^t\|_F^2 \leq f(\{\mathcal{F}\}_{1:N}^{t+1}, \mathcal{X}^t, \mathcal{E}^t, \mathcal{Y}^t), \\
&f(\{\mathcal{F}\}_{1:N}^{t+1}, \mathcal{X}^{t+1}, \mathcal{E}^{t+1}, \mathcal{Y}^t) + \frac{\rho}{2}\|\mathcal{E}^{t+1} - \mathcal{E}^t\|_F^2 \leq f(\{\mathcal{F}\}_{1:N}^{t+1}, \mathcal{X}^{t+1}, \mathcal{E}^t, \mathcal{Y}^t), \\
&f(\{\mathcal{F}\}_{1:N}^{t+1}, \mathcal{X}^{t+1}, \mathcal{E}^{t+1}, \mathcal{Y}^{t+1}) + \frac{\rho}{2}\|\mathcal{Y}^{t+1} - \mathcal{Y}^t\|_F^2 \leq f(\{\mathcal{F}\}_{1:N}^{t+1}, \mathcal{X}^{t+1}, \mathcal{E}^{t+1}, \mathcal{Y}^t).
\end{aligned}
\tag{63}
$$

Lemma 6.2. *(**Relative error**) Suppose $\{\{\mathcal{F}\}_{1:N}^t, \mathcal{X}^t, \mathcal{E}^t, \mathcal{Y}^t\}_{t\in\mathbb{N}}$ be the sequence obtained by Algorithm 1, then there exists $\mathcal{U}_k^{t+1} \in \boldsymbol{0}$ ($k = 1, 2, \cdots, N$), $\mathcal{V}_1^{t+1} \in \boldsymbol{0}$, $\mathcal{V}_2^{t+1} \in \partial f_2(\mathcal{E}^{t+1})$, and $\mathcal{V}_3^{t+1} \in \partial \Phi(\mathcal{Y}^{t+1})$, such that*

$$
\begin{cases}
\|\mathcal{U}_k^{t+1} + \nabla_{\mathcal{F}_k} f_1(\{\mathcal{F}\}_{1:k}^{t+1}, \{\mathcal{F}\}_{k+1:N}^t, \mathcal{X}^t, \mathcal{E}^t, \mathcal{Y}^t)\|_F \leq \rho\|\mathcal{F}_k^{t+1} - \mathcal{F}_k^t\|_F, \ k = 1, 2, \cdots, N, \\
\|\mathcal{V}_1^{t+1} + \nabla_{\mathcal{X}} f_1(\{\mathcal{F}\}_{1:N}^{t+1}, \mathcal{X}^{t+1}, \mathcal{E}^t, \mathcal{Y}^t)\|_F \leq \rho\|\mathcal{X}^{t+1} - \mathcal{X}^t\|_F, \\
\|\mathcal{V}_2^{t+1} + \nabla_{\mathcal{E}} f_1(\{\mathcal{F}\}_{1:N}^{t+1}, \mathcal{X}^{t+1}, \mathcal{E}^{t+1}, \mathcal{Y}^t)\|_F \leq \rho\|\mathcal{E}^{t+1} - \mathcal{E}^t\|_F, \\
\|\mathcal{V}_3^{t+1} + \nabla_{\mathcal{Y}} f(\{\mathcal{F}\}_{1:N}^{t+1}, \mathcal{X}^{t+1}, \mathcal{E}^{t+1}, \mathcal{Y}^{t+1})\|_F \leq \rho\|\mathcal{Y}^{t+1} - \mathcal{Y}^t\|_F.
\end{cases}
\tag{64}
$$

To prove Lemma 6.2, we first show that the sequence is bounded. Since the initial tensors $\{\mathcal{F}\}_{1:N}^0, \mathcal{X}^0, \mathcal{E}^0$, and $\mathcal{Y}^0$ are apparently bounded, we prove that $\{\mathcal{F}\}_{1:N}^{t+1}$, $\mathcal{X}^{t+1}, \mathcal{E}^{t+1}$, and $\mathcal{Y}^{t+1}$ are bounded when $\{\mathcal{F}\}_{1:N}^t, \mathcal{X}^t, \mathcal{E}^t$, and $\mathcal{Y}^t$ are bounded.

(I) The sequence $\{\{\mathcal{F}\}_{1:N}^t\}_{t\in\mathbb{N}}$ are bounded: Supposing that $\|\mathcal{F}_k^t\|_F \leq a$ and $\|\mathcal{X}^t\|_F \leq b$, according to (14), we have

$$
\begin{aligned}
\|\mathcal{F}_1^{t+1}\|_F &\leq (\|\mathcal{X}^t\|_F\|\mathcal{M}_1^t\|_F + \rho\|\mathcal{F}_1^t\|_F)\|(\mathbf{N}_1^t + \rho\mathbf{I})^{-1}\|_F \\
&\leq (ba^{N-1} + \rho a)\sqrt{\sum_{i=1}^j (1/(\sigma_i + \rho))^2} \\
&\leq (ba^{N-1} + \rho a)\sqrt{j}/\rho,
\end{aligned}
\tag{65}
$$

where $\mathbf{N}_1^t = (\mathbf{M}_k^t)_{[\mathbf{m}_{1:N-1};\mathbf{n}_{1:N-1}]}(\mathbf{M}_k^t)_{[\mathbf{n}_{1:N-1};\mathbf{m}_{1:N-1}]}$ and σ_i is the eigenvalues of $\mathbf{N}_1^t$. It is clearly to see that $\mathcal{F}_1^{t+1}$ is bounded. Similarly, we can obtain that $\mathcal{F}_2^{t+1}, \mathcal{F}_3^{t+1}, \cdots, \mathcal{F}_N^{t+1}$ are bounded.

(II) The sequence $\{\mathcal{X}^t\}_{t\in\mathbb{N}}$ is bounded: Supposing that $\|\mathcal{F}_k^{t+1}\|_F \leq c$, $\|\mathcal{Y}^t\|_F \leq d$, and $\|\mathcal{E}^t\|_F \leq e$, according to (17)

$$
\begin{aligned}
\|\mathcal{X}^{t+1}\|_F &= \|\mathrm{FCTN}(\mathcal{F}_{1:N}^{t+1}) + \beta(\mathcal{Y}^t - \mathcal{E}^t) + \rho\mathcal{X}^t\|_F/(1+\beta+\rho) \\
&\leq (\|\mathrm{FCTN}(\mathcal{F}_{1:N}^{t+1})\|_F + \beta\|\mathcal{Y}^t\|_F + \beta\|\mathcal{E}^t\|_F + \rho\|\mathcal{X}^t\|_F)/(1+\beta+\rho) \quad (66) \\
&\leq (c^N + \beta d + \beta e + \rho b)/(1+\beta+\rho).
\end{aligned}
$$

Thus, $\mathcal{X}^{t+1}$ is bounded.

(III) The sequence $\{\mathcal{E}^t\}_{t\in\mathbb{N}}$ is bounded: Since $\mathcal{X}^{t+1}$, $\{\mathcal{F}\}_{1:N}^{t+1}$, $\mathcal{Y}^t$, and $\mathcal{E}^t$ are bounded, we suppose that $\|\mathcal{X}^{t+1} - \mathrm{FCTN}(\{\mathcal{F}\}_{1:N}^{t+1})\|_F \leq \mathrm{f}$, $\|\mathcal{Y}^t - \mathcal{X}^{t+1} - \mathcal{E}^t\|_F \leq g$, and $\|\mathcal{E}^t\|_1 \leq h$, according to (62), we have

$$
\begin{aligned}
f_2(\mathcal{E}^{t+1}) = \lambda\|\mathcal{E}^{t+1}\|_1 &\leq f(\{\mathcal{F}\}_{1:N}^{t+1}, \mathcal{X}^{t+1}, \mathcal{E}^t, \mathcal{Y}^t) \\
&= \frac{1}{2}\|\mathcal{X}^{t+1} - \mathrm{FCTN}(\{\mathcal{F}\}_{1:N}^{t+1})\|_F^2 + \frac{\beta}{2}\|\mathcal{Y}^t - \mathcal{X}^{t+1} - \mathcal{E}^t\|_F^2 + \lambda\|\mathcal{E}^t\|_1 + \Phi(\mathcal{Y}^t) \\
&\leq \frac{1}{2}f^2 + \frac{\beta}{2}g^2 + \lambda h.
\end{aligned}
$$
$$(67)$$

Therefore, $\mathcal{E}^{t+1}$ is bounded.

(IV) The sequence $\{\mathcal{Y}^t\}_{t\in\mathbb{N}}$ is bounded: Supposing that $\|\mathcal{X}^{t+1}\|_F \leq i$ and $\|\mathcal{E}^{t+1}\|_F \leq j$, according to (21), we have

$$
\begin{aligned}
\|\mathcal{Y}^{t+1}\|_F &\leq \|(\mathcal{Y}^{t+1/2})_{\Omega^C}\|_F + \|\mathcal{Y}_\Omega\|_F \\
&\leq \|\beta(\mathcal{X}^{t+1} + \mathcal{E}^{t+1}) + \rho\mathcal{Y}^t\|_F/(\beta+\rho) + \rho\|\mathcal{Y}^t\|_F \quad (68) \\
&\leq \beta(i+j)/(\beta+\rho) + \rho d/(\beta+\rho) + \rho d.
\end{aligned}
$$

Thus, $\mathcal{Y}^{t+1}$ is bounded.

In summary, the sequence $\{\{\mathcal{F}\}_{1:N}^t, \mathcal{X}^t, \mathcal{E}^t, \mathcal{Y}^t\}_{t\in\mathbb{N}}$ is bounded.

Let $\{\mathcal{F}\}_{1:N}^{t+1}$, $\mathcal{X}^{t+1}$, $\mathcal{E}^{t+1}$, and $\mathcal{Y}^{t+1}$ are the optimal solutions of $\mathcal{F}_k$-subproblem, $\mathcal{X}$-subproblem, $\mathcal{E}$-subproblem, and $\mathcal{Y}$-subproblem, we have

$$
\begin{cases}
\mathbf{0} \in \nabla_{\mathcal{F}_k} f_1(\{\mathcal{F}\}_{1:k-1}^{t+1}, \mathcal{F}_k, \{\mathcal{F}\}_{k+1:N}^t, \mathcal{X}^t, \mathcal{E}^t, \mathcal{Y}^t) + \rho(\mathcal{F}_k - \mathcal{F}_k^t), \\
\mathbf{0} \in \nabla_{\mathcal{X}} f_1(\{\mathcal{F}\}_{1:N}^{t+1}, \mathcal{X}, \mathcal{E}^t, \mathcal{Y}^t) + \rho(\mathcal{X} - \mathcal{X}^t), \\
\mathbf{0} \in \nabla_{\mathcal{E}} f_1(\{\mathcal{F}\}_{1:N}^{t+1}, \mathcal{X}^{t+1}, \mathcal{E}, \mathcal{Y}^t) + \partial f_2(\mathcal{E}) + \rho(\mathcal{E} - \mathcal{E}^t), \\
\mathbf{0} \in \nabla_{\mathcal{Y}} f_1(\{\mathcal{F}\}_{1:N}^{t+1}, \mathcal{X}^{t+1}, \mathcal{E}^{t+1}, \mathcal{Y}) + \partial\Phi(\mathcal{Y}) + \rho(\mathcal{Y} - \mathcal{Y}^t).
\end{cases}
\quad (69)
$$

Then, we define $\mathcal{U}_k^{t+1}$, $\mathcal{V}_1^{t+1}$, $\mathcal{V}_2^{t+1}$, and $\mathcal{V}_3^{t+1}$ as

$$
\begin{cases}
\mathcal{U}_k^{t+1} = -\nabla_{\mathcal{F}_k} f_1(\{\mathcal{F}\}_{1:k}^{t+1}, \{\mathcal{F}\}_{k+1:N}^t, \mathcal{X}^t, \mathcal{E}^t, \mathcal{Y}^t) - \rho(\mathcal{F}_k^{t+1} - \mathcal{F}_k^t) \in \mathbf{0}, \\
\mathcal{V}_1^{t+1} = -\nabla_{\mathcal{X}} f_1(\{\mathcal{F}\}_{1:N}^{t+1}, \mathcal{X}^{t+1}, \mathcal{E}^t, \mathcal{Y}^t) - \rho(\mathcal{X}^{t+1} - \mathcal{X}^t) \in \mathbf{0}, \\
\mathcal{V}_2^{t+1} = -\nabla_{\mathcal{E}} f_1(\{\mathcal{F}\}_{1:N}^{t+1}, \mathcal{X}^{t+1}, \mathcal{E}^{t+1}, \mathcal{Y}^t) - \rho(\mathcal{E}^{t+1} - \mathcal{E}^t) \in \partial f_2(\mathcal{E}^{t+1}), \\
\mathcal{V}_3^{t+1} = -\nabla_{\mathcal{Y}} f_1(\{\mathcal{F}\}_{1:N}^{t+1}, \mathcal{X}^{t+1}, \mathcal{E}^{t+1}, \mathcal{Y}^{t+1}) - \rho(\mathcal{Y}^{t+1} - \mathcal{Y}^t) \in \partial\Phi(\mathcal{Y}^{t+1}).
\end{cases}
\quad (70)
$$

Since the sequence $\{\{\mathcal{F}\}_{1:N}^t, \mathcal{X}^t, \mathcal{E}^t, \mathcal{Y}^t\}_{t\in\mathbb{N}}$ is bounded, and ∇f_1 is Lipschitz continuous on any bounded set. Then there exists $\mathcal{U}_k^{t+1} \in \mathbf{0}$ ($k = 1, 2, \cdots, N$),

$\mathcal{V}_1^{t+1} \in \mathbf{0}$, $\mathcal{V}_2^{t+1} \in \partial f_2(\mathcal{E}^{t+1})$, and $\mathcal{V}_3^{t+1} \in \partial\Phi(\mathcal{Y}^{t+1})$, such that

$$
\begin{aligned}
\|\mathcal{U}_k^{t+1} + \nabla_{\mathcal{F}_k} f_1(\{\mathcal{F}\}_{1:k}^{t+1}, \ \{\mathcal{F}\}_{k+1:N}^t, \ \mathcal{X}^t, \ \mathcal{E}^t, \ \mathcal{Y}^t)\|_F &\leq \rho\|\mathcal{F}_k^{t+1} - \mathcal{F}_k^t\|_F, \\
\|\mathcal{V}_1^{t+1} + \nabla_{\mathcal{X}} f_1(\{\mathcal{F}\}_{1:N}^{t+1}, \ \mathcal{X}^{t+1}, \ \mathcal{E}^t, \ \mathcal{Y}^t)\|_F &\leq \rho\|\mathcal{X}^{t+1} - \mathcal{X}^t\|_F, \\
\|\mathcal{V}_2^{t+1} + \nabla_{\mathcal{E}} f_1(\{\mathcal{F}\}_{1:N}^{t+1}, \ \mathcal{X}^{t+1}, \ \mathcal{E}^{t+1}, \ \mathcal{Y}^t)\|_F &\leq \rho\|\mathcal{E}^{t+1} - \mathcal{E}^t\|_F, \\
\|\mathcal{V}_3^{t+1} + \nabla_{\mathcal{Y}} f(\{\mathcal{F}\}_{1:N}^{t+1}, \ \mathcal{X}^{t+1}, \ \mathcal{E}^{t+1}, \ \mathcal{Y}^{t+1})\|_F &\leq \rho\|\mathcal{Y}^{t+1} - \mathcal{Y}^t\|_F.
\end{aligned}
\tag{71}
$$

Combining these conditions, the proposed algorithm conforms to Theorem 6.2 in [1], the bounded sequence $\{\{\mathcal{F}\}_{1:N}^t, \ \mathcal{X}^t, \ \mathcal{E}^t, \ \mathcal{Y}^t\}_{t\in\mathbb{N}}$ converges to the critical point of $f(\{\mathcal{F}\}_{1:N}, \ \mathcal{X}, \ \mathcal{E}, \ \mathcal{Y})$. $\qquad\square$

Appendix B. Proof of Theorem 4.3. Here, we exhibit the detailed proof of Theorem 4.3.

Proof. We first recall the general convergence result of the ADMM [11]. Consider the following two-block convex problems:

$$
\begin{aligned}
\arg\min_{\mathbf{x},\mathbf{y}} \ & f(\mathbf{x}) + g(\mathbf{y}), \\
\text{s.t. } & \mathbf{Ax} + \mathbf{By} = \mathbf{b}, \ \mathbf{x} \in X, \ \mathbf{y} \in Y,
\end{aligned}
\tag{72}
$$

where $f : \mathbb{R}^m \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}$ are closed proper convex functions, $X \subseteq \mathbb{R}^m$ and $Y \subseteq \mathbb{R}^n$ are closed convex sets, $\mathbf{A} \in \mathbb{R}^{l\times m}$ and $\mathbf{B} \in \mathbb{R}^{l\times n}$ are matrices, and $\mathbf{b} \in \mathbb{R}^l$ is a given vector. The augmented Lagrangian function of (72) is

$$
L(\mathbf{x}, \mathbf{y}, \mathbf{z}) = f(\mathbf{x}) + g(\mathbf{y}) + \langle \mathbf{z}, \mathbf{Ax} + \mathbf{By} - \mathbf{b}\rangle + \frac{\beta}{2}\|\mathbf{Ax} + \mathbf{By} - \mathbf{b}\|_2^2,
\tag{73}
$$

where $\mathbf{z}$ is the Lagrangian multiplier and β is a penalty parameter. The ADMM algorithm iterates as

$$
\begin{cases}
\mathbf{x}^{t+1} = \arg\min_{\mathbf{x}} f(\mathbf{x}) + \langle\mathbf{z}^t, \mathbf{Ax}\rangle + \frac{\beta}{2}\|\mathbf{Ax} + \mathbf{By}^t - \mathbf{b}\|_2^2, \\
\mathbf{y}^{t+1} = \arg\min_{\mathbf{y}} g(\mathbf{y}) + \langle\mathbf{z}^t, \mathbf{By}\rangle + \frac{\beta}{2}\|\mathbf{Ax}^{t+1} + \mathbf{By} - \mathbf{b}\|_2^2 \\
\mathbf{z}^{t+1} = \mathbf{z}^t + \tau\beta\left(\mathbf{Ax}^{t+1} + \mathbf{By}^{t+1} - \mathbf{b}\right),
\end{cases}
$$

where τ is the step length and the superscript t refers to the iteration index. The following theorem establishes the convergence of ADMM.

Lemma 6.3. *(Theorem B.1 in [11]) Assume that the solution set of (72) is nonempty and there exists $(\boldsymbol{x}^0, \boldsymbol{y}^0) \in ri(\mathrm{dom}f \times \mathrm{dom}g) \cap P$, where P is the constraint set in (72). Assume also that both $\boldsymbol{A}^T\boldsymbol{A}$ and $\boldsymbol{B}^T\boldsymbol{B}$ are positive definite. Let $\{(\boldsymbol{x}^t, \boldsymbol{y}^t, \boldsymbol{z}^t)\}$ be generated from the ADMM algorithm. If the step length $\tau \in (0, (1 + \sqrt{5})/2)$, then the sequence $\{(\boldsymbol{x}^t, \boldsymbol{y}^t)\}$ converges to an optimal solution to (72) and $\{\boldsymbol{z}^t\}$ converges to an optimal solution to the dual problem of (72). Therefore, the sequence $\{(\boldsymbol{x}^t, \boldsymbol{y}^t, \boldsymbol{z}^t)\}$ generated from the ADMM algorithm is convergent.*

Then we equivalently rewrite the proposed convex model which fits the form (72). By introducing auxiliary variables $\mathcal{L}_k$ $(k = 1, 2, \cdots, \bar{N})$ and $\mathcal{S}$, the optimization

problem (25) can be rewritten as

$$\min_{\mathcal{X},\mathcal{E}} \sum_{k=1}^{\bar{N}} \alpha_k \|\mathbf{L}_{k[\mathbf{n}_1^k;\mathbf{n}_2^k]}\|_* + \lambda\|\mathcal{S}\|_1 + \Phi(\mathcal{Y}) \tag{74}$$
$$\text{s.t.} \quad \mathcal{Y} = \mathcal{X} + \mathcal{E}, \mathcal{S} = \mathcal{E}, \mathcal{L}_k = \mathcal{X}, k = 1, 2, \cdots, \bar{N},$$

The linear constraints can be reformulated as the following matrix-vector multiplication form:

$$\begin{pmatrix} \mathbf{I} & \mathbf{I} \\ \mathbf{0} & \mathbf{I} \\ \mathbf{I} & \mathbf{0} \\ \vdots & \vdots \\ \mathbf{I} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ \mathbf{e} \end{pmatrix} + \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ -\mathbf{I} & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & -\mathbf{I} & \mathbf{0} & \cdots & \mathbf{0} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \cdots & -\mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{s} \\ \boldsymbol{l}_1 \\ \boldsymbol{l}_2 \\ \vdots \\ \boldsymbol{l}_{\bar{N}} \end{pmatrix} = \begin{pmatrix} \mathbf{y} \\ \mathbf{0} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{pmatrix}, \tag{75}$$

where $\mathbf{I}$ and $\mathbf{0}$ respectively denote the identify matrix and zero matrix, $\mathbf{x}$, $\mathbf{e}$, $\mathbf{s}$, $\{\boldsymbol{l}_k\}_{k=1}^{\bar{N}}$, and $\mathbf{y}$ denote the vectorization of $\mathcal{X}$, $\mathcal{E}$, $\mathcal{S}$, $\{\mathcal{L}_k\}_{k=1}^{\bar{N}}$, and $\mathcal{Y}$, respectively. We separate all the variables into two groups $(\mathcal{X},\mathcal{E})$ and $(\{\mathcal{L}_k\}_{k=1}^{\bar{N}},\mathcal{S})$, and decompose the objective function as $f + g$ with $f = \Phi(\mathcal{X}+\mathcal{E})$ $(\Phi(\mathcal{Y}) = \Phi(\mathcal{X}+\mathcal{E}))$ and $g = \sum_{k=1}^{\bar{N}} \alpha_k\|\mathbf{L}_{k[\mathbf{n}_1^k;\mathbf{n}_2^k]}\|_* + \lambda\|\mathcal{S}\|_1)$. Then the minimization problem (74) fits the framework of ADMM (72).

Now, we examine the conditions in Lemma 6.3. The verification is divided into three parts.

First, it is clear that the objective function of (74), denote by $F(\mathcal{X},\mathcal{E},\mathcal{L}_k,\mathcal{S})$, is a proper convex function.

Second, the P in our model is an affine space, thus there exists $(\mathbf{x}^0, \mathbf{y}^0) \in ri(domf \times domg) \cap P$.

Third, both $\mathbf{A}^T\mathbf{A}$ and $\mathbf{B}^T\mathbf{B}$ are positive definite, since $\mathbf{A}$ and $\mathbf{B}$ in (75) are full column rank, where $\mathbf{A}$ and $\mathbf{B}$ denote the coefficient matrices of the variables $(\mathbf{x}^T, \mathbf{e}^T)^T$ and $(\mathbf{s}^T, \boldsymbol{l}_1^T, \cdots, \boldsymbol{l}_{\bar{N}}^T)^T$, respectively.

We set the step length $\tau = 1.1$, then according to the Lemma 6.3, the sequence $\{\mathcal{L}_k^t, \mathcal{S}^t, \mathcal{X}^t, \mathcal{E}^t\}_{t\in\mathbb{N}}$ obtained by ADMM is convergent. $\square$

REFERENCES

[1] 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. Prog.*, **137** (2013), 91-129.

[2] 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.*, **26** (2017), 2466-2479.

[3] J. Bolte, A. Daniilidis, A. Lewis and M. Shiota, Clarke subgradients of stratifiable functions, *SIAM J. Optim.*, **18** (2007), 556-572.

[4] J. Bolte, S. Sabach and M. Teboulle, Proximal alternating linearized minimization for non-convex and nonsmooth problems, *Math. Prog.*, **146** (2014), 459-494.

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

[6] E. J. Candès, X. Li, Y. Ma and J. Wright, Robust principal component analysis?, *J. ACM*, **58** (2011), 1-37.

[7] E. J. Candes and T. Tao, The power of convex relaxation: Near-optimal matrix completion, *IEEE Trans. Inf. Theory*, **56** (2010), 2053-2080.

[8] C. Chen, X. Li, M. K. Ng and X. Yuan, Total variation based tensor decomposition for multi-dimensional data with time dimension, *Numer. Linear Algeb. Appl.*, **22** (2015), 999-1019.

[9] C. Chen, Z.-B. Wu, Z.-T. Chen, Z.-B. Zheng and X.-J. Zhang, Auto-weighted robust low-rank tensor completion via tensor-train, *Inform. Sci.*, **567** (2021), 100-115.

[10] M. Ding, T.-Z. Huang, X.-L. Zhao, M. K. Ng and T.-H. Ma, Tensor train rank minimization with nonlocal self-similarity for tensor completion, *Inverse Probl. Imag.*, **15** (2021), 475-498.

[11] 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.*, **34** (2013), 946-977.

[12] C. J. Hillar and L.-H. Lim, Most tensor problems are NP-hard, *J. ACM*, **60** (2013), 1-39.

[13] B. Huang, C. Mu, D. Goldfarb and J. Wrigh, Provable models for robust low-rank tensor completion, *Pac. J. Optim.*, **11** (2015), 339-364.

[14] H. Huang, Y. Liu, Z. Long and C. Zhu, Robust low-rank tensor ring completion, *IEEE Trans. Comput. Imaging*, **6** (2020), 1117-1126.

[15] Q. Jiang and M. Ng, Robust low-tubal-rank tensor completion via convex optimization, in *Proceedings of the IJCAI*, 2649-2655.

[16] 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.*, **34** (2013), 148-172.

[17] M. E. Kilmer and C. D. Martin, Factorization strategies for third-order tensors, *Linear Algeb. Appl.*, **435** (2011), 641-658.

[18] B.-Z. Li, X.-L. Zhao, T.-Y. Ji, X.-J. Zhang and T.-Z. Huang, Nonlinear transform induced tensor nuclear norm for tensor completion, *J. Sci. Comput.*, **92** (2022), Paper No. 83, 30 pp.

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

[20] 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.*, **42** (2020), 925-938.

[21] C.-Y. Lyu, X.-L. Zhao, B.-Z. Li, H. Zhang and T.-Z. Huang, Multi-dimensional image recovery via fully-connected tensor network decomposition under the learnable transforms, *J. Sci. Comput.*, **93** (2022), Paper No. 49, 24 pp.

[22] I. V. Oseledets, Tensor-train decomposition, *SIAM J. Sci. Comput.*, **33** (2011), 2295-2317.

[23] O. Semerci, N. Hao, M. E. Kilmer and E. L. Miller, Tensor-based formulation and nuclear norm regularization for multienergy computed tomography, *IEEE Trans. Image Process.*, **23** (2014), 1678-1693.

[24] Y. Shi, Z. Liu, X. Wang and J. Zhang, Edge detection with mixed noise based on maximum a posteriori approach, *Inverse Probl. Imag.*, **15** (2021), 1223-1245.

[25] G. Song, M. K. Ng and X. Zhang, Robust tensor completion using transformed tensor singular value decomposition, *Numer. Linear Algeb. Appl.*, **27** (2020), e2299, 27 pp.

[26] L. R. Tucker, Some mathematical notes on three-mode factor analysis, *Psychometrika*, **31** (1966), 279-311.

[27] K. Wei, J.-F. Cai, T. F. Chan and S. Leung, Guarantees of riemannian optimization for low rank matrix completion, *Inverse Probl. Imag.*, **14** (2020), 233-265.

[28] Y. Xu, R. Hao, W. Yin and Z. Su, Parallel matrix factorization for low-rank tensor completion, *Inverse Probl. Imag.*, **9** (2015), 601-624.

[29] Y. Xu and W. Yin, A block coordinate descent method for regularized multiconvex optimization with applications to nonnegative tensor factorization and completion, *SIAM J. Imaging Sci.*, **6** (2013), 1758-1789.

[30] N. Yair and T. Michaeli, Multi-scale weighted nuclear norm image restoration, in *Proceedings of the CVPR*, 3165-3174.

[31] H. Yang, X. Ding, R. Chan, H. Hu, Y. Peng and T. Zeng, A new initialization method based on normed statistical spaces in deep networks, *Inverse Probl. Imag.*, **15** (2021), 147-158.

[32] J.-H. Yang, X.-L. Zhao, T.-Y. Ji, T.-H. Ma and T.-Z. Huang, Low-rank tensor train for tensor robust principal component analysis, *Appl. Math. Comput.*, **367** (2020), 124783, 15 pp.

[33] K. Ye and L.-H. Lim, Tensor network ranks, `arXiv:1801.02662`.

[34] J. Yu, C. Li, Q. Zhao and G. Zhou, Tensor-ring nuclear norm minimization and application for visual data completion, in *Proceedings of the ICASSP*, (2019), 3142-3146.

[35] Q. Yu, X. Zhang and Z.-H. Huang, Multi-tubal rank of third order tensor and related low rank tensor completion problem, arXiv preprint, `arXiv:2012.05065`.

[36] H. Zhang, X.-L. Zhao, T.-X. Jiang, M. K. Ng and T.-Z. Huang, Multiscale feature tensor train rank minimization for multidimensional image recovery, *IEEE Trans. Cybern.*, **52** (2021), 13395-13410.

[37] J. Zhang, Y. Duan, Y. Lu, M. K. Ng and H. Chang, Bilinear constraint based admm for mixed poisson-gaussian noise removal, *Inverse Probl. Imag.*, **15** (2021), 339-366.

[38] X. Zhang, A nonconvex relaxation approach to low-rank tensor completion, *IEEE Trans. Neural Netws. Learn. Syst.*, **30** (2019), 1659-1671.

[39] X. Zhang and M. K. Ng, Robust tensor train component analysis, *Numer. Linear Algeb. Appl.*, **29** (2022), e2403.

[40] M. Zhao, Y.-W. Wen, M. Ng and H. Li, A nonlocal low rank model for poisson noise removal, *Inverse Probl. Imag.*, **15** (2021), 519-537.

[41] Q. Zhao, G. Zhou, S. Xie, L. Zhang and A. Cichocki, Tensor ring decomposition, arXiv preprint, `arXiv:1606.05535`.

[42] X. Zhao, M. Bai and M. K. Ng, Nonconvex optimization for robust tensor completion from grossly sparse observations, *J. Sci. Comput.*, **85** (2020), Paper No. 46, 32 pp.

[43] 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.*, **58** (2020), 734-749.

[44] 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 *Proceedings of the AAAI Conf. Artifi. Intell.*, **35** (2021).