Numer. Math. Theor. Meth. Appl.
doi: 10.4208/nmtma.OA-2023-0135

Vol. **xx**, No. x, pp. 1-25
x 2024

Non-Local and Fully Connected Tensor Network Decomposition for Remote Sensing Image Denoising

Zhihui Tu[1], Shunda Chen[1], Jian Lu[1,2,3,*], Lin Li[4] and Qingtang Jiang[5]

[1] *Shenzhen Key Laboratory of Advanced Machine Learning and Applications, School of Mathematical Sciences, Shenzhen University, Shenzhen 518060, China*
[2] *National Center for Applied Mathematics Shenzhen (NCAMS), Shenzhen 518055, China*
[3] *Pazhou Lab, Guangzhou 510320, China*
[4] *School of Electronic Engineering, Xidian University, Xi'an 710071, China*
[5] *Department of Mathematics and Statistics, University of Missouri-St. Louis, St. Louis, MO 63121, USA*

Received 25 October 2023; Accepted (in revised version) 8 January 2024

Abstract. Remote sensing images (RSIs) encompass abundant spatial and spectral/temporal information, finding wide applications in various domains. However, during image acquisition and transmission, RSI often encounter noise interference, which adversely affects the accuracy of subsequent applications. To address this issue, this paper proposes a novel non-local fully connected tensor network (NLFCTN) decomposition algorithm for denoising RSI, aiming to fully exploit their global correlation and non-local self-similarity (NSS) characteristics. FCTN, as a recently developed tensor decomposition technique, exhibits remarkable capability in capturing global correlations and minimizing information loss. In addition, we introduce an efficient algorithm based on proximal alternating minimization (PAM) to efficiently solve the model and prove the convergence. The effectiveness of the proposed method is validated through denoising experiments on both simulated and real RSI data, employing objective evaluation metrics and subjective visual assessments. The results of the experiment show that the proposed method outperforms other RSI denoising techniques in terms of denoising performance.

AMS subject classifications: 65F22, 65F35, 94A08, 46N10
Key words: Remote sensing images, denoising, tensor decomposition, non-local self-similarity.

1. Introduction

Recently, remote sensing images (RSIs) [10] have gained widespread utilization across various domains. However, in practical applications, RSIs inevitably suffer from

*Corresponding author. *Email address:* jianlu@szu.edu.cn (J. Lu)

noise contamination [17]. Noise presence markedly degrades image quality, consequently impacting subsequent tasks, including classification [26], sparse unmixing, target segmentation [25], and detection. Consequently, denoising emerges as an inevitable challenge during post-analysis and preprocessing of RSIs.

RSI exhibit rich spatial, spectral, or temporal features, which can be utilized to aid in RSI denoising. In general, RSI denoising involves leveraging the global correlations [23], non-local self-similarity (NSS) [5, 8], piecewise smoothness [7], and deep priors [22, 32] of the data. Due to the vast amount of data, RSIs often contain redundant information, namely global correlation. For instance, in hyperspectral images (HSIs), there is notable correlation among their spectral bands, resulting in the spectral vector residing in a lower-dimensional subspace [3]. Mathematically, we can characterize this global correlation using a low-rank representation, representing high-dimensional data based on learned lower-dimensional bases [34]. In the past two decades, low-rank matrix approximation methods have received significant attention, providing a theoretical foundation for RSI denoising and achieving promising results [15, 16, 29]. For instance, Zhang *et al.* [30] proposed a low-rank matrix recovery (LRMR) model, which achieved promising results by unfolding the RSI into a matrix. He *et al.* [12] introduced the total variation (TV) and proposed a low-rank matrix decomposition with total variation regularization (LRTV) method. However, matrix-based approaches often disrupt the intrinsic structure of high-order RSI data when unfolding it into matrices.

In recent years, tensor-based models have been proposed to recover RSI, drawing inspiration from the successful application of tensors in image processing inverse problems. These models enable effective exploration of the intrinsic properties of high-dimensional data. RSI data can be considered as tensors, where HSIs and multispectral images (MSIs) correspond to third-order tensors, while multi-temporal RSI are represented as fourth-order tensors. Liu *et al.* [14] used the PARAFAC model and statistical performance analysis to efficiently reconstruct the noisy RSI. Guo *et al.* [11] recovered RSI based on rank-1 tensor decomposition. Zhao *et al.* [35] proposed the constrained tube rank and sparsity model (CTSD) for addressing the problem of mixed noise removal in RSI. This model incorporates a low tube rank constraint, as well as ℓ_0 and ℓ_1 norm constraints, to effectively characterize the underlying clean RSI and sparse noise components. Zhuang *et al.* [39] proposed a global local factorization (GLF) method combining global matrix decomposition and local tensor decomposition, and achieved good results. Chen *et al.* [6] proposed a non-local group sparsifying transform learning (TLNLGS) method for HSI denoising. Nonetheless, t-SVD-based methods are limited to third-order tensors and lack flexibility in handling tensor correlation across different modes. Based on Tucker decomposition, Renard *et al.* [21] proposed a low-rank tensor approximation (LRTA) method, which effectively mines the low-rank attributes of different modes of data. However, the LRTA method only considers the global correlation of data, and cannot reconstruct the details and edge information of RSI well. To this end, Wang *et al.* [27] proposed a new method combining low-rank tensor decomposition and total variation (LRTDTV), which can better restore the RSI damaged

by mixed noise. Tu *et al.* [24] introduced a non-convex low-rank tensor approximation model (NLRTA) based on the γ-function, effectively addressing the issue of mixed noise in RSI. However, the Tucker decomposition expands the tensor along a single mode, making it impossible to mine correlations between different modes.

Recently, tensor network decomposition [18, 20, 33, 37, 38] has received considerable attention and has shown better performance in reconstructing high-dimensional data. Tensor train (TT) decomposition [20] decomposes the N-order tensor into two matrices and $N-2$ third-order tensors, which can more accurately consider the correlation between adjacent factors. Gong *et al.* [9] proposed the TT minimum description length (TT-MDL) method, where they employed the minimum description length principle to estimate the TT rank. The TT-MDL method theoretically reveals the potential advantage of the TT model in characterizing latent low tensor ranks. Tensor ring (TR) decomposition [33] decomposes the N-order tensor into N third-order tensors. Zhang *et al.* [31] introduced the T-RSTR model, which combines reconcile sparsity and low TR rank prior in the transformed domain, for RSI denoising. Chen *et al.* [5] proposed a non-local tensor ring (NLTR) approximation method that utilizes the TR decomposition to simultaneously explore non-local self-similarity and global spectral correlations for RSI denoising. However, both the TT and TR decompositions fail to effectively explore the correlations between any two factors.

To address these concerns, we took inspiration from the highly successful fully connected tensor network (FCTN) decomposition method as presented in [38]. This paper presents a novel approach known as the Non-Local Fully Connected Tensor Network Decomposition (NLFCTN) model for the denoising of RSI. Initially, we employ FCTN decomposition to initialize the observed data, thereby unveiling global correlations within the RSI dataset. Subsequently, we perform block-wise processing on the initialized data, where similar blocks are superimposed and denoised using FCTN, resulting in clear similar groups. Finally, we split and aggregate the clear similar groups to obtain denoised data. Moreover, for the introduced RSI denoising model, we propose an efficient iterative solution algorithm based on the Proximal Alternating Minimization (PAM) method, and its convergence is proved. By conducting comprehensive experiments on simulated and real datasets, we validate the efficacy of the NLFCTN approach. In summary, the contributions of this paper are as follows:

- We introduce a novel denoising model, NLFCTN, which considers correlations among various modes of RSI data. This model effectively leverages global similarity and prior NSS information within high-dimensional data, enhancing denoising performance.
- We develop an iterative algorithm for NLFCTN model based on PAM method and prove the convergence. Experiments on simulated and real datasets, whether assessed quantitatively or qualitatively, illustrate the effectiveness of our proposed method in reconstructing RSI.

Specific instructions for the remainder of this paper follow. In Section 2, we provide a concise introduction to the symbolic notation of tensors and offer a detailed

explanation of the FCTN decomposition. Subsequently, in Section 3, we introduce the NLFCTN model and present an iterative solution algorithm based on the PAM method. In addition, Section 5 presents the numerical experimental results of our method and comparative methods, which quantitatively and qualitatively demonstrate the effectiveness of our method. In Section 6, we analyze the effects of parameters on the model and algorithm. Finally, Section 7 summarizes the work.

2. Tensor notations and preliminaries

In this section, we first briefly introduce the framework of RSI denoising problem. On this basis, we introduce the NLFCTN denoising model in detail.

2.1. Basic notations

In this paper, our symbol definitions can be referred to [38]. We adopt the following notation: a scalar, vector, matrix, and tensor are represented as $h, \mathbf{h}, \mathbf{H}$, and $\mathcal{H}$, respectively. For the tensor $\mathcal{H} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$, its $(i_1, i_2, \cdots, i_N)$ element is defined as $\mathcal{H}(i_1, i_2, \cdots, i_N)$, and its Frobenius norm is defined as $\|\mathcal{H}\|_F = \sqrt{\sum |\mathcal{H}(i_1, i_2, \cdots, i_N)|^2}$. To simplify notation, we use $\mathcal{H}_{1:d}$ to refer to the subset $(\mathcal{H}_1, \mathcal{H}_2, \cdots, \mathcal{H}_d)$.

2.2. Basic definitions

Before introducing the FCTN decomposition, we first present three relevant definitions. Among them, mode-k matricization and generalized tensor unfolding are two ways of tensor unfolding, while tensor contraction defines the operation relating two tensors.

Definition 2.1 (cf. Kolda & Bader [13]). *The mode-k matricization of $\mathcal{A} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ is $\mathbf{A}_{(k)} \in \mathbb{R}^{I_k \times (I_1 \cdots I_{k-1} I_{k+1} \cdots I_N)}$. The corresponding operation and its inverse operation are denoted as $\mathbf{A}_{(k)} = \mathrm{unfold}_k(\mathcal{A})$ and $\mathcal{A} = \mathrm{fold}_k(\mathbf{A}_{(k)})$, respectively.*

Definition 2.2 (cf. Zheng *et al.* [38]). *Assume that $\mathbf{n}$ is a permutation of $(1, 2, \ldots, N)$. For an N-th order tensor $\mathcal{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$, the generalized tensor unfolding operation yields a matrix $\mathbf{X}_{[n_{1:d}; n_{d+1:N}]} \in \mathbb{R}^{\prod_{i=1}^{d} I_{n_i} \times \prod_{i=d+1}^{N} I_{n_i}}$, defined as*

$$\mathbf{X}_{[n_{1:d}; n_{d+1:N}]} = \mathrm{reshape}\left(\mathcal{X}^{\mathbf{n}}, \prod_{i=1}^{d} I_{n_i}, \prod_{i=d+1}^{N} I_{n_i}\right),$$

where $\mathcal{X}^{\mathbf{n}}$ represents the $\mathbf{n}$-based generalized tensor transposition of $\mathcal{X}$, with dimensions $\mathbb{R}^{I_{n_1} \times I_{n_2} \times \cdots \times I_{n_N}}$ (for detailed information, refer to [38]). The operations of generalized tensor unfolding and its inverse are denoted as

$$\mathbf{X}_{[n_{1:d}; n_{d+1:N}]} = \mathrm{GenUnfold}(\mathcal{X}, n_{1:d}; n_{d+1:N}),$$

and

$$\mathcal{X} = \mathrm{GenFold}(\mathbf{X}_{[n_{1:d};n_{d+1:N}]}, n_{1:d}; n_{d+1:N}),$$

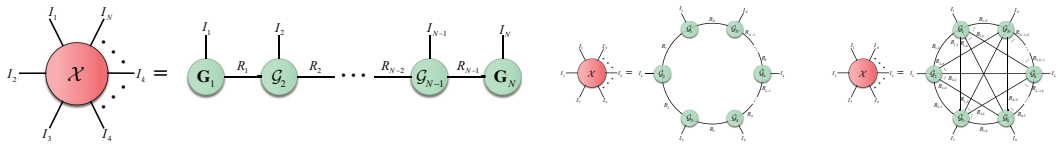
respectively.

Definition 2.3 (cf. Zheng *et al.* [38]). *Let* $\mathbf{n}$ *and* $\mathbf{m}$ *be permutations of* $(1,2,\ldots,N)$ *and* $(1,2,\ldots,M)$*, respectively. Given* $\mathcal{X} \in \mathbb{R}^{I_1\times I_2\times\cdots\times I_N}$ *and* $\mathcal{Y} \in \mathbb{R}^{J_1\times J_2\times\cdots\times J_M}$*, where* $I_{n_i} = J_{m_i}$ *for* $i = 1,2,\ldots,d$*, the tensor contraction along the* $\mathbf{n}_{1:d}$ *modes of* $\mathcal{X}$ *and the* $\mathbf{m}_{1:d}$ *modes of* $\mathcal{Y}$ *yields an* $(N+M-2d)$*-th order tensor defined as*

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

2.3. FCTN decomposition

The FCTN decomposition [38] is a generalization of the TT and TR decompositions. As shown in Fig. 1, FCTN decomposes an N-th order target tensor $\mathcal{X} \in \mathbb{R}^{I_1\times I_2\times\cdots\times I_N}$ into N factor tensors $\mathcal{G}_{1:N}$. Each factor tensor $\mathcal{G}_k$ $(k = 1,2,\ldots,N)$ has dimensions $\mathcal{G}_k \in \mathbb{R}^{R_{1,k}\times R_{2,k}\times\cdots\times R_{k-1,k}\times I_k\times R_{k,k+1}\times\cdots\times R_{k,N}}$. In contrast to TT and TR decompositions, the FCTN decomposition exhibits transposition invariance and the ability to account for correlations between each factor.



TT decomposition TR decomposition FCTN decomposition

Figure 1: A visual illustration of tensor network decomposition.

The elementwise form of the FCTN decomposition is expressed as

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

The tensor product form of the FCTN decomposition can be expressed as

$$\begin{aligned}\mathcal{X} = &\sum_{r_{1,2}=1}^{R_{1,2}}\sum_{r_{1,3}=1}^{R_{1,3}}\cdots\sum_{r_{1,N}=1}^{R_{1,N}}\sum_{r_{2,3}=1}^{R_{2,3}}\cdots\sum_{r_{2,N}=1}^{R_{2,N}}\cdots\sum_{r_{N-1,N}}^{R_{N-1,N}}\\ &\Big\{\mathcal{G}_1(:,r_{1,2},r_{1,3},\cdots,r_{1,N})\circ\mathcal{G}_2(r_{1,2},:,r_{2,3},\cdots,r_{2,N})\circ\cdots\circ\end{aligned}$$

$$\mathcal{G}_k(r_{1,k}, r_{2,k}, \cdots, r_{k-1,k}, :, r_{k,k+1}, \cdots, r_{k,N}) \circ \cdots \circ$$
$$\mathcal{G}_N(r_{1,N}, r_{2,N}, \cdots, r_{N-1,N}, :)\Big\},$$

where "$\circ$" represents the outer product of tensor fibers. Additionally, the FCTN decomposition is defined as $\mathcal{X} = \mathrm{FCTN}(\mathcal{G}_{1:N}) = \mathrm{FCTN}\ (\mathcal{G}_1, \mathcal{G}_2, \cdots, \mathcal{G}_N)$. The rank of FCTN is denoted by a vector composed of R_{k_1,k_2} $(1 \leq k_1 < k_2 \leq N)$.

3. Proposed model and solving algorithm

In this section, we first briefly introduce the framework of RSI denoising problem. On this basis, we introduce the NLFCTN denoising model in detail.

3.1. Problem formulation

Due to the inevitable noise pollution in RSI, the degradation model can be expressed as

$$\mathcal{Y} = \mathcal{X} + \mathcal{N},$$

where $\mathcal{Y}$ represents the observed image, $\mathcal{X}$ represents the clean image, and $\mathcal{N}$ corresponds to Gaussian noise. Our objective is to reconstruct clean data from observational data. Nevertheless, the estimation of a clean image $\mathcal{X}$ from observational data $\mathcal{Y}$ poses a challenging ill-posed inverse problem. Employing regularization methods to address this inverse problem proves to be an effective solution strategy. Typically, the model is expressed as follows:

$$\min_{\mathcal{X},\mathcal{N}} \varphi_1(\mathcal{X}) + \lambda_1 \varphi_2(\mathcal{N})$$
$$\text{s.t. } \mathcal{Y} = \mathcal{X} + \mathcal{N},$$

where λ_1 represents a regularization parameter that balances two regularization terms, $\varphi_1(\mathcal{X})$ and $\varphi_2(\mathcal{N})$, which capture the global correlation of $\mathcal{X}$ and the Gaussian distribution of $\mathcal{N}$, respectively. In the subsequent discussion, we will delve into the details of $\varphi_1(\mathcal{X})$ and $\varphi_2(\mathcal{N})$.

3.2. Proposed model and solving algorithm

Fig. 2 illustrates the flowchart of the proposed NLFCTN model. To leverage the global correlation inherent in RSI data, we employ FCTN decomposition as an initial step to process the observed image. Subsequently, we partition the initialized image into small blocks exhibiting similar characteristics, which are then stacked to form nonlocal similar groups. This approach enables us to fully harness the remarkable global correlation representation capabilities of FCTN, while intelligently capitalizing on the

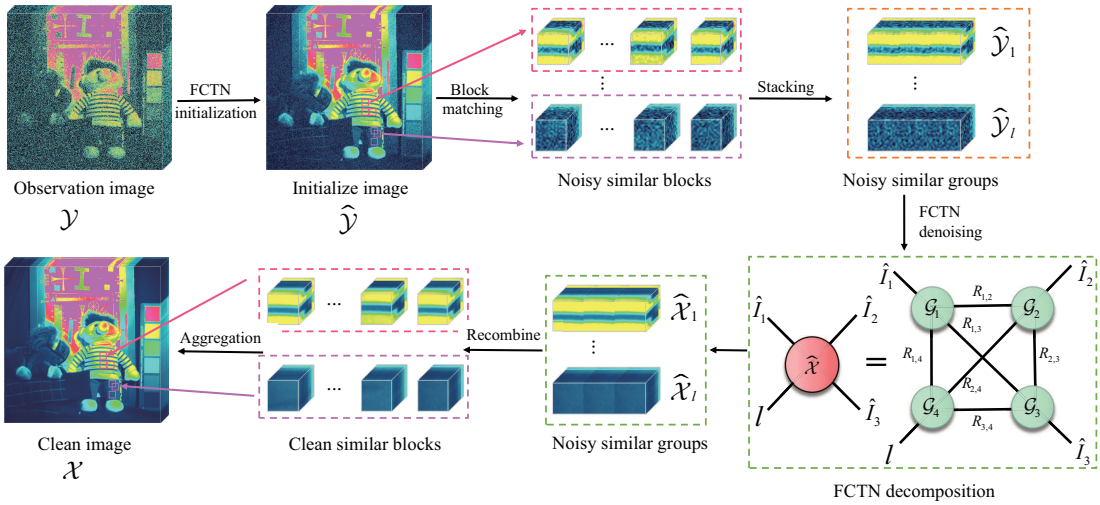


Figure 2: Flowchart of the NLFCTN method. Our method first uses the FCTN method for initialization, then blocks the initialized image, and combines similar blocks to obtain noise similarity groups. Here it can be seen that the noise similarity group is one order higher than the original observed image. For example, the third-order observation image will get the fourth-order noise similarity group. Next, we employ FCTN for decomposition to better extract the inherent features of the original data, resulting in the reconstructed clean image.

benefits offered by NSS-based tensor-order incremental operations for handling high-order tensors [36]. Finally, we disassemble and aggregate the recovered clean non-local similarity groups to their original positions, completing the denoising process.

For observation data $\mathcal{Y}$ of order $\mathcal{N}$, we examine its noise similarity group denoted as $\widehat{\mathcal{Y}}_l$. As depicted in Fig. 2, it is evident that $\widehat{\mathcal{Y}}_l$ possesses an order of $\mathcal{N}+1$. To address the denoising problem of the noise similarity group $\widehat{\mathcal{Y}}_l$, we employ FCTN and establish the NLFCTN model

$$\begin{aligned} &\min_{\widehat{\mathcal{X}}_l,\mathcal{G}_{l,k},\widehat{\mathcal{N}}_l} \frac{1}{2}\|\widehat{\mathcal{X}}_l - \mathrm{FCTN}(\mathcal{G}_{l,1},\mathcal{G}_{l,2},\ldots,\mathcal{G}_{l,N})\|_{\mathrm{F}}^2 + \frac{\lambda}{2}\|\widehat{\mathcal{N}}_l\|_F^2 \\ &\text{s.t. } \widehat{\mathcal{Y}}_l = \widehat{\mathcal{X}}_l + \widehat{\mathcal{N}}_l, \end{aligned} \tag{3.1}$$

where $\mathrm{FCTN}(\mathcal{G}_{l,1},\mathcal{G}_{l,2},\ldots,\mathcal{G}_{l,N})$ represents the FCTN decomposition, and $\mathcal{G}_{l,k}$ represents the k-th factor of the $\widehat{\mathcal{X}}_l$ FCTN decomposition. Additionally, $\widehat{\mathcal{Y}}_l$, $\widehat{\mathcal{X}}_l$, and $\widehat{\mathcal{N}}_l$ represent the l-th NSS groups of $\mathcal{Y}$, $\mathcal{X}$, and $\mathcal{N}$, respectively.

We consider the following penalty function of problem (3.1):

$$\min_{\widehat{\mathcal{X}}_l,\mathcal{G}_{l,k},\widehat{\mathcal{N}}_l} \frac{1}{2}\|\widehat{\mathcal{X}}_l - \mathrm{FCTN}(\mathcal{G}_{l,1},\mathcal{G}_{l,2},\ldots,\mathcal{G}_{l,N})\|_F^2 + \frac{\lambda}{2}\|\widehat{\mathcal{N}}_l\|_F^2 + \frac{\beta}{2}\|\widehat{\mathcal{Y}}_l - \widehat{\mathcal{X}}_l - \widehat{\mathcal{N}}_l\|_F^2, \tag{3.2}$$

where β is penalty parameter. Given the interdependence of all optimization variables, we employ the PAM framework [1] to solve the aforementioned equation. The solution

is obtained through iterative updates of $\widehat{\mathcal{X}}_l$, $\mathcal{G}_{l,k}$, and $\widehat{\mathcal{N}}_l$, as illustrated below

$$\begin{cases} \widehat{\mathcal{X}}_l^{(t+1)} = \underset{\widehat{\mathcal{X}}_l}{\operatorname{argmin}} f\big(\widehat{\mathcal{X}}_l, \mathcal{G}_{l,k}^{(t)}, \widehat{\mathcal{N}}_l^{(t)}\big) + \frac{\rho}{2}\big\|\widehat{\mathcal{X}}_l - \widehat{\mathcal{X}}_l^{(t)}\big\|_F^2, \\ \mathcal{G}_{l,k}^{(t+1)} = \underset{\mathcal{G}_{l,k}}{\operatorname{argmin}} f\big(\widehat{\mathcal{X}}_l^{(t+1)}, \mathcal{G}_{l,1:k-1}^{(t+1)}, \mathcal{G}_{l,k}, \mathcal{G}_{l,k+1:N}^{(t)}, \widehat{\mathcal{N}}_l^{(t)}\big) + \frac{\rho}{2}\big\|\mathcal{G}_{l,k} - \mathcal{G}_{l,k}^{(t)}\big\|_F^2, \\ \widehat{\mathcal{N}}_l^{(t+1)} = \underset{\widehat{\mathcal{N}}_l}{\operatorname{argmin}} f\big(\widehat{\mathcal{X}}_l^{(t+1)}, \mathcal{G}_{l,k}^{(t+1)}, \widehat{\mathcal{N}}_l\big) + \frac{\rho}{2}\big\|\widehat{\mathcal{N}}_l - \widehat{\mathcal{N}}_l^{(t)}\big\|_F^2. \end{cases} \tag{3.3}$$

1. Update $\widehat{\mathcal{X}}_l^{(t+1)}$.

   The subproblem of $\widehat{\mathcal{X}}_l^{(t+1)}$ can be simplified to

$$\begin{aligned} \widehat{\mathcal{X}}_l^{(t+1)} &= \arg\min_{\widehat{\mathcal{X}}_l} \frac{1}{2}\big\|\widehat{\mathcal{X}}_l - \mathrm{FCTN}(\mathcal{G}_{l,1}^{(t)}, \mathcal{G}_{l,2}^{(t)}, \ldots, \mathcal{G}_{l,N}^{(t)})\big\|_{\mathrm{F}}^2 \\ &\quad + \frac{\beta}{2}\big\|\widehat{\mathcal{Y}}_l - \widehat{\mathcal{X}}_l - \widehat{\mathcal{N}}_l^{(t)}\big\|_F^2 + \frac{\rho}{2}\big\|\widehat{\mathcal{X}}_l - \widehat{\mathcal{X}}_l^{(t)}\big\|_F^2 \\ &= \arg\min_{\widehat{\mathcal{X}}_l} \frac{1+\beta+\rho}{2}\left\|\widehat{\mathcal{X}}_l - \frac{\mathrm{FCTN}(\{\mathcal{G}_{l,k}\}_{1:N}^{(t)}) + \beta(\widehat{\mathcal{Y}}_l - \widehat{\mathcal{N}}_l^{(t)}) + \rho\widehat{\mathcal{X}}_l^{(t)}}{1+\beta+\rho}\right\|_F^2. \end{aligned}$$

   This is a least squares problem, and the following closed solution can be obtained by deriving $\widehat{\mathcal{X}}_l$ and setting the derivative equal to 0:

$$\widehat{\mathcal{X}}_l^{(t+1)} = \frac{\mathrm{FCTN}(\{\mathcal{G}_{l,k}\}_{1:N}^{(t)}) + \beta(\widehat{\mathcal{Y}}_l - \widehat{\mathcal{N}}_l^{(t)}) + \rho\widehat{\mathcal{X}}_l^{(t)}}{1+\beta+\rho}. \tag{3.4}$$

2. Update $\mathcal{G}_{l,k}^{(t+1)}$.

   As stated in [38] (refer to Theorem 4), the subproblem of $\mathcal{G}_{l,k}^{(t+1)}$ can be reformulated as follows:

$$\begin{aligned} \mathcal{G}_{l,k}^{(t+1)} = \underset{\mathcal{G}_{l,k}}{\operatorname{argmin}} &\frac{1}{2}\big\|(\widehat{\mathbf{X}}_{l,k})_{(k)}^{(t+1)} - (\mathbf{G}_{l,k})_{(k)}(\mathbf{M}_{l,k})_{[m_{1:N-1};n_{1:N-1}]}\big\|_F^2 \\ &+ \frac{\beta}{2}\big\|(\mathbf{G}_{l,k})_{(k)} - (\mathbf{G}_{l,k})_{(k)}^{(t)}\big\|_F^2, \end{aligned}$$

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

$$\begin{aligned} \big(\mathbf{G}_{l,k}^{(t+1)}\big)_{(k)} = &\big[\mathbf{X}_{l,k}^{(t+1)}(\mathbf{M}_{l,k}^{(t)})_{[n_{1:N-1};m_{1:N-1}]} + \beta(\mathbf{G}_{l,k})_{(k)}^{(t)}\big] \\ &\big[(\mathbf{M}_{l,k}^{(t)})_{[m_{1:N-1};n_{1:N-1}]}(\mathbf{M}_{l,k}^{(t)})_{[n_{1:N-1};;m_{1:N-1}]} + \beta\mathbf{I}\big]^{-1}, \end{aligned} \tag{3.5}$$

   and $\mathcal{G}_{l,k}^{(t+1)} = \mathrm{GenFold}((\mathbf{G}_{l,k}^{(t+1)})_{(k)}, k; 1, \cdots, k-1, k+1, \cdots, N)$.

3. Update $\widehat{\mathcal{N}}_l$.

   The subproblem of $\widehat{\mathcal{N}}_l$ can be reformulated as follows:

$$\widehat{\mathcal{N}}_l^{(t+1)} = \arg\min_{\widehat{\mathcal{N}}_l} \frac{\lambda}{2}\|\widehat{\mathcal{N}}_l\|_F^2 + \frac{\beta}{2}\|\widehat{\mathcal{Y}}_l - \widehat{\mathcal{X}}_l^{(t+1)} - \widehat{\mathcal{N}}_l\|_F^2 + \frac{\rho}{2}\|\widehat{\mathcal{N}}_l - \widehat{\mathcal{N}}_l^{(t)}\|_F^2. \quad (3.6)$$

   Solving the subproblem $\widehat{\mathcal{N}}_l$ is also a least squares problem, which can be solved similarly to $\widehat{\mathcal{X}}_l$

$$\widehat{\mathcal{N}}_l^{(t+1)} = \frac{\beta(\widehat{\mathcal{Y}}_l - \widehat{\mathcal{X}}_l^{(t+1)}) + \rho\widehat{\mathcal{N}}_l^{(t)}}{\lambda + \beta + \rho}. \quad (3.7)$$

See Algorithm 3.1 for the algorithm flow of NLFCTN model.

Algorithm 3.1 Optimization process of NLFCTN solution

Input: Observation data $\mathcal{Y}$, regularization parameter λ, penalty parameter β and penalty parameter ρ.
Use FCTN decomposition to get initialized $\widehat{\mathcal{Y}}$.
while not converged **do**
    Divide $\widehat{\mathcal{Y}}$ into blocks, match similar blocks, and then superimpose to obtain a similar group $\widehat{\mathcal{Y}}_l$;
    Update $\widehat{\mathcal{X}}_l$ by Eq. (3.4);
    Update $\mathcal{G}_{l,k}$ by Eq. (3.5);
    Update $\widehat{\mathcal{N}}_l$ by Eq. (3.7);
    Check the convergence condition $\|\widehat{\mathcal{X}}_l^{(t+1)} - \widehat{\mathcal{X}}_l^{(t)}\|_F / \|\widehat{\mathcal{X}}_l^{(t)}\|_F < 10^{-4}$.
end while
Reorganize the reconstructed clean similarity group $\widehat{\mathcal{X}}_l$ to obtain clean similar blocks, and aggregate these clean similar blocks to obtain a clean image $\mathcal{X}$.
Output: The clean image $\mathcal{X}$.

3.3. Computational complexity analysis

In this section, we analyze the computational complexity of the proposed NLFCTN method. For an N-order noise similar group $\widehat{\mathcal{Y}}_l$, the FCTN rank R_{k_1,k_2} $(1 \le k_1 \le k_2 \le N)$ is assumed to be R. The computational complexity of NLFCTN can be divided into three parts, namely 1) update $\widehat{\mathcal{X}}_l$, 2) update $\widehat{\mathcal{G}}_{l,k}$ $(k = 1, 2, \ldots, N)$ and 3) update $\widehat{\mathcal{N}}_l$. In Eq. (3.4), it costs $\mathcal{O}(\sum_{k=2}^{N} \widehat{I}^k R^{k(N-k)+k-1})$ to update the subproblem of $\widehat{\mathcal{X}}_l$. In Eq. (3.5), it costs $\mathcal{O}(N\sum_{k=2}^{N} \widehat{I}^k\ R^{k(N-k)+k-1} + N\widehat{I}^{N-1}R^{2(N-1)} + NR^{3(N-1)})$ to update the subproblem of $\widehat{\mathcal{G}}_{l,k}$ $(k = 1, 2, \ldots, N)$. In Eq. (3.7), updating the subproblem of $\widehat{\mathcal{N}}_l$ requires an $\mathcal{O}(\widehat{I}^k)$ cost. Thus, the cost of fully calculating a noise similarity group is $\mathcal{O}(\sum_{k=2}^{N} \widehat{I}^k R^{k(N-k)+k-1} + N\sum_{k=2}^{N} \widehat{I}^k R^{k(N-k)+k-1} + N\widehat{I}^{N-1}R^{2(N-1)} + NR^{3(N-1)} + \widehat{I}^k)$.

4. Convergence analysis of NLFCTN algorithm

In this section, we prove the global convergence of NLFCTN method under PAM framework. Let $\mathcal{H}_l = \{\widehat{\mathcal{X}}_l, \mathcal{G}_{l,k}, \widehat{\mathcal{N}}_l\}$, we define the following function:

$$F(\mathcal{H}_l) = F\big(\widehat{\mathcal{X}}_l, \mathcal{G}_{l,k}, \widehat{\mathcal{N}}_l\big) = \frac{1}{2}\big\|\widehat{\mathcal{X}}_l - \mathrm{FCTN}(\mathcal{G}_{l,1}, \mathcal{G}_{l,2}, \ldots, \mathcal{G}_{l,N})\big\|_F^2 + \frac{\lambda}{2}\big\|\widehat{\mathcal{N}}_l\big\|_F^2 + \frac{\beta}{2}\big\|\widehat{\mathcal{Y}}_l - \widehat{\mathcal{X}}_l - \widehat{\mathcal{N}}_l\big\|_F^2. \tag{4.1}$$

In addition, we introduce some definitions and lemmas to prove convergence.

Definition 4.1 (cf. Attouch *et al.* [2]). *The function $F(x) : \mathbb{R}^n \to \mathbb{R} \cup +\infty$ is said to have the Kurdyka-Łojasiewicz (K-Ł) property at $x^* \in \mathrm{dom}(\partial F(x))$ if there exist $\eta \in (0, +\infty]$, a neighborhood U of x^*, and a continuous concave function $\psi(x) : [0, \eta) \to \mathbb{R}_+$ such that*

- $\psi(0) = 0$,
- $\psi(x)$ *is* C^1 *on* $(0, \eta]$,
- $\psi'(x) > 0$ *for any* $x \in (0, \eta)$,
- *for any x in $U \cap [F(x^*) < F(x) < F(x^*) + \eta]$, the following K-Ł inequality holds:*

$$\psi'\big(F(x) - F(x^*)\big)\,\mathrm{dist}\,\big(0, \partial F(x)\big) \geq 1,$$

where $\partial F(x)$ denotes the subdifferential of $F(x)$. Furthermore, proper lower semi-continuous functions are called K-Ł functions when they satisfy the K-Ł property at every point within the domain of their subdifferential $\partial F(x)$.

Definition 4.2 (cf. Attouch *et al.* [2]). *If there exists a series of real polynomial functions m_{ij} and n_{ij} satisfying $S = \cap_j \cup_i \{x \in \mathbb{R}^n : m_{ij}(x) = 0, n_{ij}(x) < 0\}$, then the subset $S \in \mathbb{R}$ is a semi-algebraic set. If the graph $\{(x, y) \in \mathbb{R}^n \times \mathbb{R}, F(x) = y\}$ of the function F is a semi-algebraic set, then F is a semi-algebraic function.*

Remark 4.1. A semi-algebraic real valued function F satisfies K-Ł property at each $x \in \mathrm{dom}(F)$, i.e., F is a K-Ł function.

Lemma 4.1 (Sufficient Decrease Lemma). *Assume that $\mathcal{H}_l^{(t)} \triangleq \{\widehat{\mathcal{X}}_l^{(t)}, \mathcal{G}_{l,k}^{(t)}, \widehat{\mathcal{N}}_l^{(t)}\}$ generated by Eq.* (3.3). *Then, $\mathcal{H}_l^{(t)}$ satisfies the following formula:*

$$F\big(\mathcal{H}_l^{(t+1)}\big) + \frac{\rho}{2}\big\|\mathcal{H}_l^{(t+1)} - \mathcal{H}_l^{(t)}\big\|_F^2 \leq F\big(\mathcal{H}_l^{(t)}\big).$$

Proof. Suppose $\widehat{\mathcal{X}}_l^{(t+1)}$, $\mathcal{G}_{l,k}^{(t+1)}$ and $\widehat{\mathcal{N}}_l^{(t+1)}$ represent the optimal solutions of the respective subproblems in Eq. (3.3). Then, the following holds:

$$\begin{cases} F\big(\widehat{\mathcal{X}}_l^{(t+1)}, \mathcal{G}_{l,k}^{(t)}, \widehat{\mathcal{N}}_l^{(t)}\big) + \frac{\rho}{2}\big\|\widehat{\mathcal{X}}_l^{(t+1)} - \widehat{\mathcal{X}}_l^{(t)}\big\|_F^2 \leq F\big(\widehat{\mathcal{X}}_l^{(t)}, \mathcal{G}_{l,k}^{(t)}, \widehat{\mathcal{N}}_l^{(t)}\big), \\ F\big(\widehat{\mathcal{X}}_l^{(t+1)}, \mathcal{G}_{l,k}^{(t+1)}, \widehat{\mathcal{N}}_l^{(t)}\big) + \frac{\rho}{2}\big\|\mathcal{G}_{l,k}^{(t+1)} - \mathcal{G}_{l,k}^{(t)}\big\|_F^2 \leq F\big(\widehat{\mathcal{X}}_l^{(t+1)}, \mathcal{G}_{l,k}^{(t)}, \widehat{\mathcal{N}}_l^{(t)}\big), \quad k = 1, 2, \ldots, N, \\ F\big(\widehat{\mathcal{X}}_l^{(t+1)}, \mathcal{G}_{l,k}^{(t+1)}, \widehat{\mathcal{N}}_l^{(t+1)}\big) + \frac{\rho}{2}\big\|\widehat{\mathcal{N}}_l^{(t+1)} - \widehat{\mathcal{N}}_l^{(t)}\big\|_F^2 \leq F\big(\widehat{\mathcal{X}}_l^{(t+1)}, \mathcal{G}_{l,k}^{(t+1)}, \widehat{\mathcal{N}}_l^{(t)}\big). \end{cases}$$

By combining the above inequalities, we have

$$\begin{aligned} &F\big(\widehat{\mathcal{X}}_l^{(t+1)}, \mathcal{G}_{l,k}^{(t+1)}, \widehat{\mathcal{N}}_l^{(t+1)}\big) + \frac{\rho}{2}\big\|\widehat{\mathcal{X}}_l^{(t+1)} - \widehat{\mathcal{X}}_l^{(t)}\big\|_F^2 \\ &\quad + \frac{\rho}{2}\big\|\mathcal{G}_{l,k}^{(t+1)} - \mathcal{G}_{l,k}^{(t)}\big\|_F^2 + \frac{\rho}{2}\big\|\widehat{\mathcal{N}}_l^{(t+1)} - \widehat{\mathcal{N}}_l^{(t)}\big\|_F^2 \\ &\leq F(\widehat{\mathcal{X}}_l^{(t)}, \mathcal{G}_{l,k}^{(t)}, \widehat{\mathcal{N}}_l^{(t)}), \end{aligned}$$

i.e.,

$$F\big(\mathcal{H}_l^{(t+1)}\big) + \frac{\rho}{2}\big\|\mathcal{H}_l^{(t+1)} - \mathcal{H}_l^{(t)}\big\|_F^2 \leq F\big(\mathcal{H}_l^{(t)}\big).$$

The proof of the sufficient decrease lemma is complete. □

Lemma 4.2 (Relative Error Lemma)**.** *Assume that* $\mathcal{H}_l^{(t)} \triangleq \{\widehat{\mathcal{X}}_l^{(t)}, \mathcal{G}_{l,k}^{(t)}, \widehat{\mathcal{N}}_l^{(t)}\}$ *is generated by Eq.* (3.3)*. Then,* $\{\mathcal{H}_l^{(t)}\}$ *is bounded, and there exists* $d^{k+1} \in \partial F(\mathcal{H}_l^{(t+1)})$*, satisfies the following formula:*

$$\|d^{(t+1)}\|_F \leq \rho\big\|\mathcal{H}_l^{(t+1)} - \mathcal{H}_l^{(t)}\big\|_F.$$

Proof. First, we prove $\mathcal{H}_l^{(t)}$ is bounded. Since

$$\begin{cases} \lim\limits_{\|\widehat{\mathcal{X}}_l\|_F \to +\infty} \dfrac{\beta}{2}\big\|\widehat{\mathcal{Y}}_l - \widehat{\mathcal{X}}_l - \widehat{\mathcal{N}}_l\big\|_F = +\infty, \\ \lim\limits_{\|\mathcal{G}_{l,k}\|_F \to +\infty} \dfrac{1}{2}\big\|\widehat{\mathcal{X}}_l - \mathrm{FCTN}(\mathcal{G}_{l,1}, \mathcal{G}_{l,2}, \ldots, \mathcal{G}_{l,N})\big\|_F = +\infty, \\ \lim\limits_{\|\widehat{\mathcal{N}}_l\|_F \to +\infty} \dfrac{\lambda}{2}\big\|\widehat{\mathcal{N}}_l\big\|_F = +\infty, \end{cases}$$

we can respectively obtain

$$\begin{cases} \lim\limits_{\|\widehat{\mathcal{X}}_l\|_F \to +\infty} F\big(\widehat{\mathcal{X}}_l, \mathcal{G}_{l,k}, \widehat{\mathcal{N}}_l\big) = +\infty, \\ \lim\limits_{\|\mathcal{G}_{l,k}\|_F \to +\infty} F\big(\widehat{\mathcal{X}}_l, \mathcal{G}_{l,k}, \widehat{\mathcal{N}}_l\big) = +\infty, \\ \lim\limits_{\|\widehat{\mathcal{N}}_l\|_F \to +\infty} F\big(\widehat{\mathcal{X}}_l, \mathcal{G}_{l,k}, \widehat{\mathcal{N}}_l\big) = +\infty. \end{cases}$$

Therefore, we can conclude that $F(\mathcal{H}_l)$ tends to infinity if the sequence $\mathcal{H}_l^{(t)}$ is unbounded. In other words, if $F(\mathcal{H}_l^{(t+1)})$ is finite, then the sequence $\mathcal{J}^k$ is bounded. Consequently, we establish the finiteness of $F(\mathcal{H}_l^{(t+1)})$ as demonstrated below. Referring to Lemma 4.1, we obtain the following:

$$\begin{aligned} F(\mathcal{H}_l^{(t+1)}) &\leq F\big(\mathcal{H}_l^{(t+1)}\big) + \frac{\rho}{2}\big\|\mathcal{H}_l^{(t+1)} - \mathcal{H}_l^{(t)}\big\|_F^2 \\ &\leq F\big(\mathcal{H}_l^{(t)}\big) \leq F\big(\mathcal{H}_l^{(t)}\big) + \frac{\rho}{2}\big\|\mathcal{H}_l^{(t)} - \mathcal{H}_l^{(t-1)}\big\|_F^2 \\ &\leq \cdots \leq F\big(\mathcal{H}_l^{(0)}\big), \end{aligned}$$

then $F(\mathcal{H}_l^{(t+1)})$ is finite. Consequently, we can infer that $\mathcal{H}^{(t)}$ is bounded. Next, consider the optimal solutions of problem Eq. (3.3), denoted as $\widehat{\mathcal{X}}_l^{(t+1)}$, $\mathcal{G}_{l,k}^{(t+1)}$ and $\widehat{\mathcal{N}}_l^{(t+1)}$. By applying the Karush-Kuhn-Tucker (KKT) conditions, the following formulas are obtained:

$$\begin{cases} 0 = \nabla_{\widehat{\mathcal{X}}_l} F\big(\widehat{\mathcal{X}}_l^{(t+1)}, \mathcal{G}_{l,k}^{(t)}, \widehat{\mathcal{N}}_l^{(t)}\big) + \rho\big(\widehat{\mathcal{X}}_l^{(t+1)} - \widehat{\mathcal{X}}_l^{(t)}\big), \\ 0 = \nabla_{\mathcal{G}_{l,k}} F\big(\widehat{\mathcal{X}}_l^{(t+1)}, \mathcal{G}_{l,k}^{(t+1)}, \widehat{\mathcal{N}}_l^{(t)}\big) + \rho\big(\mathcal{G}_{l,k}^{(t+1)} - \mathcal{G}_{l,k}^{(t)}\big), \quad k = 1, 2, \ldots, N, \\ 0 = \nabla_{\widehat{\mathcal{N}}_l} F\big(\widehat{\mathcal{X}}_l^{(t+1)}, \mathcal{G}_{l,k}^{(t+1)}, \widehat{\mathcal{N}}_l^{(t+1)}\big) + \rho\big(\widehat{\mathcal{N}}_l^{(t+1)} - \widehat{\mathcal{N}}_l^{(t)}\big). \end{cases} \tag{4.2}$$

Then, we define $\mathcal{A}^{(t+1)}$, $\mathcal{B}_k^{(t+1)}$, and $\mathcal{C}^{(t+1)}$ as

$$\begin{cases} \mathcal{A}^{(t+1)} = -\nabla_{\widehat{\mathcal{X}}_l} F\big(\widehat{\mathcal{X}}_l^{(t+1)}, \mathcal{G}_{l,k}^{(t)}, \widehat{\mathcal{N}}_l^{(t)}\big) - \rho\big(\widehat{\mathcal{X}}_l^{(t+1)} - \widehat{\mathcal{X}}_l^{(t)}\big), \\ \mathcal{B}_k^{(t+1)} = -\nabla_{\mathcal{G}_{l,k}} F\big(\widehat{\mathcal{X}}_l^{(t+1)}, \mathcal{G}_{l,k}^{(t+1)}, \widehat{\mathcal{N}}_l^{(t)}\big) - \rho\big(\mathcal{G}_{l,k}^{(t+1)} - \mathcal{G}_{l,k}^{(t)}\big), \quad k = 1, 2, \ldots, N, \\ \mathcal{C}^{(t+1)} = -\nabla_{\widehat{\mathcal{N}}_l} F\big(\widehat{\mathcal{X}}_l^{(t+1)}, \mathcal{G}_{l,k}^{(t+1)}, \widehat{\mathcal{N}}_l^{(t+1)}\big) - \rho\big(\widehat{\mathcal{N}}_l^{(t+1)} - \widehat{\mathcal{N}}_l^{(t)}\big). \end{cases}$$

By Eq. (4.2), we have $\mathcal{A}^{(t+1)} = 0$, $\mathcal{B}_k^{(t+1)} = 0$ and $\mathcal{C}^{(t+1)} = 0$. We define

$$\begin{cases} d_1^{(t+1)} = \nabla_{\widehat{\mathcal{X}}_l} F\big(\widehat{\mathcal{X}}_l^{(t+1)}, \mathcal{G}_{l,k}^{(t+1)}, \widehat{\mathcal{N}}_l^{(t+1)}\big) + \mathcal{A}^{(t+1)}, \\ d_{2k}^{(t+1)} = \nabla_{\mathcal{G}_{l,k}} F\big(\widehat{\mathcal{X}}_l^{(t+1)}, \mathcal{G}_{l,k}^{(t+1)}, \widehat{\mathcal{N}}_l^{(t+1)}\big) + \mathcal{B}_k^{(t+1)}, \quad k = 1, 2, \ldots, N, \\ d_3^{(t+1)} = \nabla_{\widehat{\mathcal{N}}_l} F\big(\widehat{\mathcal{X}}_l^{(t+1)}, \mathcal{G}_{l,k}^{(t+1)}, \widehat{\mathcal{N}}_l^{(t+1)}\big) + \mathcal{C}^{(t+1)}. \end{cases}$$

It is easy to check that $d^{k+1} \triangleq \{d_1^{(t+1)}, d_{2k}^{(t+1)}, d_3^{(t+1)}\} \in \partial F(\mathcal{H}_l^{(t+1)})$. Since ∇F is Lipschitz continuous on any bounded set. By the boundedness of $\mathcal{H}_l^{(t)}$ and Lipschitz continuity of ∇F, we have

$$\|d^{k+1}\|_F \leq \rho \big\|\mathcal{H}_l^{(t+1)} - \mathcal{H}_l^{(t)}\big\|_F.$$

The proof of the relative error lemma is complete. □

Subsequently, we provide the theoretical convergence guarantee of Algorithm 3.1.

Theorem 4.1. *The bounded sequence $\{\widehat{\mathcal{X}}_l, \mathcal{G}_{l,k}, \widehat{\mathcal{N}}_l\}$ generated by Algorithm 3.1 globally converges to a critical point of Eq.* (3.2).

Proof. In order to establish the global convergence of $\{\widehat{\mathcal{X}}_l, \mathcal{G}_{l,k}, \widehat{\mathcal{N}}_l\}$ to a critical point of $F(\widehat{\mathcal{X}}_l^{(t)}, \mathcal{G}_{l,k}^{(t)}, \widehat{\mathcal{N}}_l^{(t)})$, we need to satisfy the following three key conditions:

1. $F(\widehat{\mathcal{X}}_l^{(t)}, \mathcal{G}_{l,k}^{(t)}, \widehat{\mathcal{N}}_l^{(t)})$ is a proper lower semi-continuous function.
2. $F(\widehat{\mathcal{X}}_l^{(t)}, \mathcal{G}_{l,k}^{(t)}, \widehat{\mathcal{N}}_l^{(t)})$ satisfies the K-Ł property at each $\{\widehat{\mathcal{X}}_l, \mathcal{G}_{l,k}, \widehat{\mathcal{N}}_l\} \in \mathrm{dom}(F)$.

3. The sequence $\{\widehat{\mathcal{X}}_l^{(t)}, \mathcal{G}_{l,k}^{(t)}, \widehat{\mathcal{N}}_l^{(t)}\}$ satisfies the sufficient decrease and relative error conditions.

First, F is composed of five Frobenius norm functions that its gradient is Lipschitz continuous. Therefore, $F(\widehat{\mathcal{X}}_l^{(t)}, \mathcal{G}_{l,k}^{(t)}, \widehat{\mathcal{N}}_l^{(t)})$ is a proper and lower semi-continuous function.

Second, the Frobenius norm function $\frac{1}{2}\|\widehat{\mathcal{X}}_l - \mathrm{FCTN}(\mathcal{G}_{l,1}, \mathcal{G}_{l,2}, \ldots, \mathcal{G}_{l,N})\|_F^2$, $\frac{\lambda}{2}\|\widehat{\mathcal{N}}_l\|_F^2$ and $\frac{\beta}{2}\|\widehat{\mathcal{Y}}_l - \widehat{\mathcal{X}}_l - \widehat{\mathcal{N}}_l\|_F^2$ are semi-algebraic [4] functions. According to Remark 4.1, they are K-Ł functions.

Thirdly, by referring to Lemmas 4.1 and 4.2, the sequence $F(\widehat{\mathcal{X}}_l^{(t)}, \mathcal{G}_{l,k}^{(t)}, \widehat{\mathcal{N}}_l^{(t)})$ satisfies both the sufficient decrease and relative error conditions.

In conclusion, combining these three key conditions, Algorithm 3.1 satisfies [2, Theorem 6.2]. Therefore, we can conclude that the sequence $\{\widehat{\mathcal{X}}_l^{(t)}, \mathcal{G}_{l,k}^{(t)}, \widehat{\mathcal{N}}_l^{(t)}\}$ generated by Algorithm 3.1 converges to a critical point of F.

5. Experiments

In this section, we conducted distinct simulation experiments and real experiments to validate the effectiveness of our RSI denoising method. Additionally, we compared it with five state-of-the-art methods, namely BM4D [19], LRTA [21], LRTDTV [27], NLRTA [24], and FCTN [38]. These are RSI reconstruction methods based on matrix approximation or tensor approximation, respectively. The source code for the aforementioned comparative methods was obtained from the authors' respective websites, and all parameters in the code were set as per the provided settings or manually adjusted to achieve the best results following reference recommendations. All experiments were conducted on a desktop workstation running the Windows 10 operating system and MATLAB R2021b. The workstation was equipped with a Core i7 3.20 GHz CPU and 64 GB of memory.

5.1. Evaluation indexes

The assessment of image recovery quality involves the use of essential metrics, namely the peak signal-to-noise ratio (PSNR) and the structural similarity index (SSIM) [28]. These metrics are widely employed in the field of image processing to quantitatively evaluate the fidelity. In the context of image restoration, a higher PSNR value indicates a more faithful and less noisy reconstruction, with the optimal value being infinity, representing a perfect restoration devoid of any noise distortion.

Similarly, the SSIM index gauges the structural similarity between two images by considering luminance, contrast, and structure. Its value ranges from -1 to 1, with 1 indicating a perfect similarity and -1 representing complete dissimilarity. A higher SSIM value signifies a closer resemblance between the reconstructed and original images, signifying better preservation of image structures and details. It is important to note

that the combination of PSNR and SSIM evaluations provides a robust and balanced assessment of the restoration results.

5.2. Simulation experiments on multitemporal data

In this subsection, we conducted simulation experiments using two datasets, namely France[†] and Jiangsu[†], to evaluate the robustness of the NLFCTN method on multi-temporal RSI datasets. Unlike HSI datasets, multi-temporal RSI datasets not only have spatial and spectral properties, but also have temporal properties.

Benchmark DataSet: The France dataset was obtained through Landsat-8 imaging over the airspace of France. It comprises four time points, each consisting of four spectral bands (B1, B2, B3, and B4). The dataset has a spatial resolution of 30 meters and a size of $400 \times 400 \times 4 \times 4$, with 10m/pixel spatial resolution. In addition, the Jiangsu dataset was acquired using Sentinel-2 imaging over the Jiangsu region in China. Similar to the France dataset, it consists of four time points, with each time point encompassing four spectral bands (B2, B3, B4, and B8). The dataset has a spatial resolution of 10 meters and a size of $400 \times 400 \times 4 \times 4$, with 10m/pixel spatial resolution. These datasets were specifically chosen to simulate realistic multi-temporal scenarios and provide a comprehensive evaluation of the NLFCTN method's performance across varying spatial and spectral resolutions.

Experimental results: We added Gaussian noise with different mean values to the multitemporal RSI dataset to evaluate the effectiveness of our NLFCTN method. This step is intended to simulate different noise levels that may be encountered in real-world application scenarios, so as to more comprehensively assess the practicality and adaptability of the NLFCTN method. Specifically, we generated three noise datasets with different variances corresponding to $\sigma = 0.06$, $\sigma = 0.08$ and $\sigma = 0.10$. By introducing these different levels of noise data, we were able to test and analyze the NLFCTN method in more complex and diverse noise environments to fully understand how its denoising effect performs under different noise conditions.

For the France dataset, we have compiled the quantitative results of various methods under different variances of Gaussian noise in Table 1. The best-performing outcomes from the comparison experiment are denoted in bold, enabling easy identification. Upon thorough examination, it becomes evident that the reconstruction results of the NLFCTN method exhibit higher MPSNR and MSSIM values. This significant finding suggests that the RSI reconstructed by our proposed NLFCTN method achieves a closer resemblance to the clean RSI in quantitative evaluation. In Fig. 3, we showcase the recovery results achieved on the France dataset, illustrating a compelling demonstration of the superiority of our NLFCTN method over the alternative approaches. Our method

[†]`https://theia.cnes.fr/atdistrib/rocket/#/home`

Table 1: Numerical results of France and Jiangsu datasets. The larger the value of MPSNR and MSSIM, the better the reconstruction effect, which is emphasized in bold. The smaller the running time (seconds), the faster the running speed.

Dataset	σ	Index	Noisy	BM4D	LRTA	LRTDTV	NLRTA	FCTN	NLFCTN
France	0.06	MPSNR	24.429	36.004	34.470	35.297	24.440	34.114	**36.509**
		MSSIM	0.3345	0.8683	0.8227	0.8770	0.3349	0.8091	**0.8965**
		Time	-	49.066	**2.1632**	33.597	45.597	388.34	799.50
	0.08	MPSNR	21.936	33.119	33.337	34.063	21.928	31.398	**34.133**
		MSSIM	0.2277	0.7680	0.7829	0.8051	0.2274	0.6954	**0.8391**
		Time	-	49.131	**1.7451**	33.673	34.844	381.18	785.46
	0.1	MPSNR	20.005	32.038	32.421	32.429	20.000	29.229	**33.001**
		MSSIM	0.1624	0.7046	0.7488	0.7395	0.1623	0.5824	**0.7939**
		Time	-	49.107	**1.5671**	35.372	32.213	381.70	788.81
Jiangsu	0.06	MPSNR	24.435	31.789	30.363	30.508	24.438	31.485	**32.273**
		MSSIM	0.5572	0.8045	0.8107	0.8305	0.5573	0.8575	**0.8840**
		Time	-	49.216	**3.8802**	34.287	37.522	384.98	808.71
	0.08	MPSNR	21.939	30.410	28.812	29.946	21.930	30.047	**31.202**
		MSSIM	0.4344	0.8057	0.7499	0.8099	0.4340	0.8081	**0.8518**
		Time	-	49.245	**3.1621**	34.782	32.706	382.59	797.34
	0.1	MPSNR	20.002	29.341	27.944	29.259	19.996	28.578	**30.074**
		MSSIM	0.3432	0.8092	0.7084	0.7814	0.3429	0.7485	**0.8125**
		Time	-	49.318	**2.7361**	34.546	32.546	412.61	989.87

excels in preserving crucial image structures and intricate details, as clearly observed in the magnified regions.

Similarly, Table 1 provides comprehensive MPSNR, MSSIM, and runtime results for all competing methods on the Jiangsu dataset. Once again, the highest MPSNR and MSSIM values, along with the lowest runtime values, are marked in bold for easy reference. Remarkably, our proposed NLFCTN method consistently outperforms the alternative approaches in terms of MPSNR and MSSIM metrics. To further validate the effectiveness of our method in preserving image quality, Fig. 4 presents the recovery results obtained on the Jiangsu dataset. The visual comparison convincingly supports the superiority of our NLFCTN method, as it faithfully retains image quality and exhibits superior performance compared to other methods.

5.3. Simulation experiments on HSI data

In this subsection, we perform simulation experiments using two datasets, Pavia Centre[‡] (PaC) and Washington DC Mall[‡] (WDC), to evaluate the robustness of the NLFCTN method on HSI.

[‡]https://rslab.ut.ac.ir/data

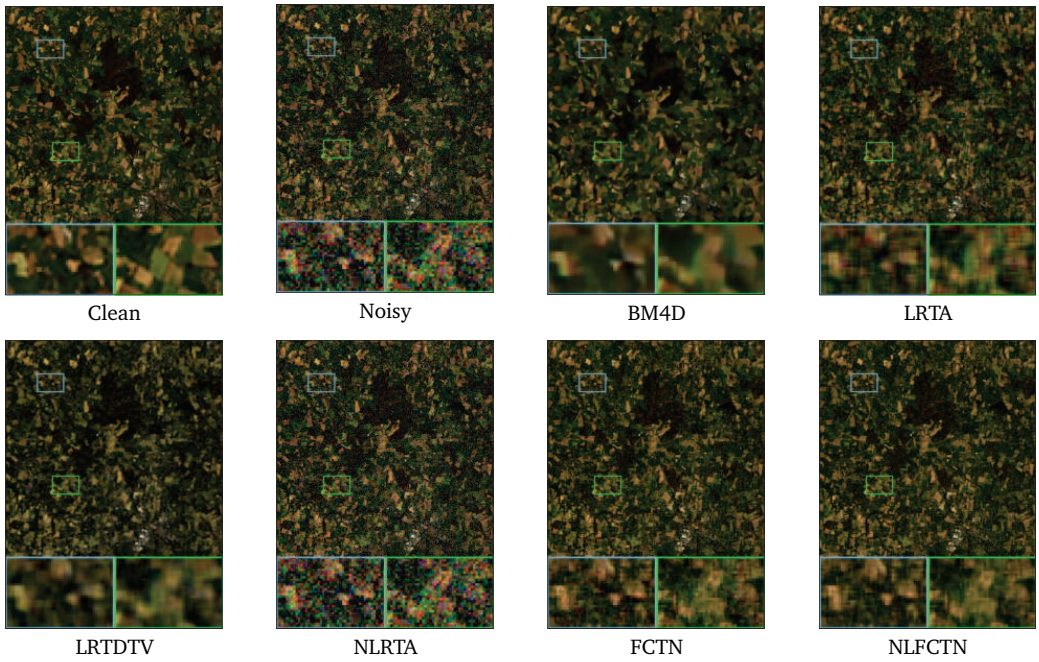


Figure 3: Experimental results on the France dataset. Gaussian noise with variance $\sigma = 0.06$ was added to the simulation experiment. For a better view of the visualization, we performed a double local magnification, where the blue and green boxes correspond to the enlarged area.

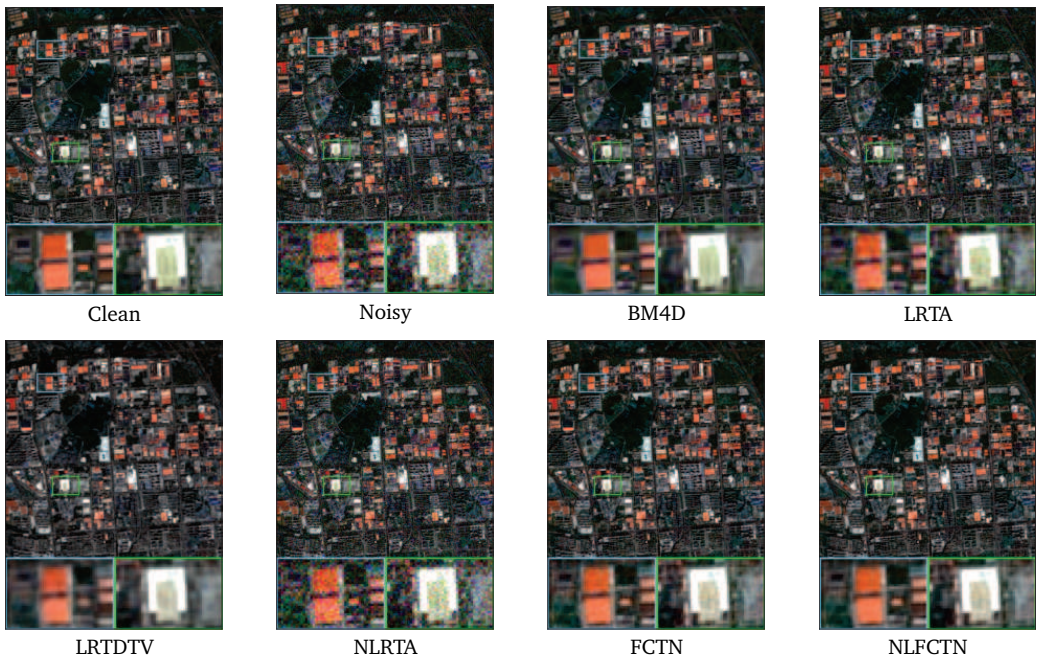


Figure 4: Experimental results on the Jiangsu dataset. Gaussian noise with variance $\sigma = 0.06$ was added to the simulation experiment. For a better view of the visualization, we performed a double local magnification, where the blue and green boxes correspond to the enlarged area.

Benchmark DataSet: In our experiments, we select a subimage with dimensions $200 \times 200 \times 80$ from the PaC dataset, where these 80 bands are noise-free bands, with 1.3m/pixel spatial resolution. In addition, we select a subimage of size $256 \times 256 \times 191$ from the WDC dataset, where these 191 bands are noise-free bands, with 2m/pixel spatial resolution. Both PaC and WDC datasets are HSI datasets, which contain rich spatial and spectral information.

Experimental results: Table 2 presents the results for MPSNR, MSSIM, and runtime of all comparative methods on the PaC dataset, enabling a quantitative assessment. The highest values for MPSNR and MSSIM, as well as the lowest runtime values, are highlighted in bold. Remarkably, the NLFCTN method proposed in this study exhibits

significantly superior MPSNR and MSSIM scores compared to the other competing methods. Specifically, at the noise level of $\sigma = 0.08$, NLFCTN demonstrates an approximate 2.79dB improvement in MPSNR compared to the second-best LRTDTV method. This improvement can be attributed to the ability of FCTN decomposition to preserve more intrinsic tensor structure, resulting in enhanced denoising performance for NLFCTN. To conduct a visual quality assessment, Fig. 5 presents the pseudo-color images illustrating the noisy result at a noise level of $\sigma = 0.10$. To facilitate a compre-

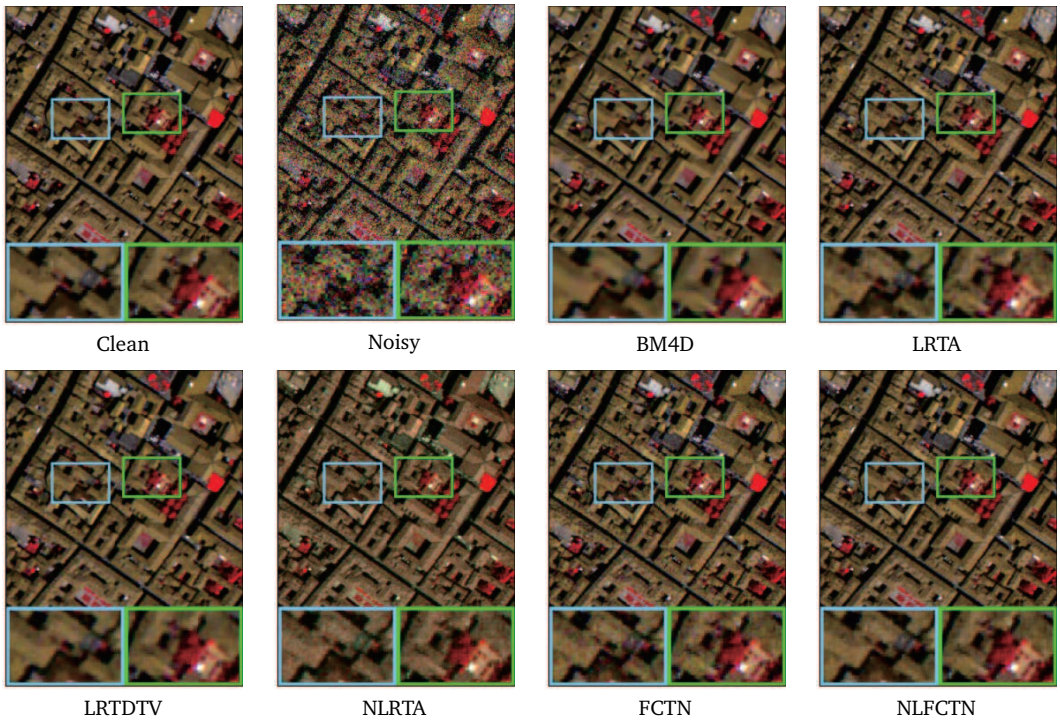


Figure 5: Experimental results on the PaC dataset. Gaussian noise with variance $\sigma = 0.1$ was added to the simulation experiment. For a better view of the visualization, we performed a double local magnification, where the blue and green boxes correspond to the enlarged area.

Table 2: Numerical results of PaC and WDC datasets. The larger the value of MPSNR and MSSIM, the better the reconstruction effect, which is emphasized in bold. The smaller the running time (seconds), the faster the running speed.

Dataset	σ	Index	Noisy	BM4D	LRTA	LRTDTV	NLRTA	FCTN	NLFCTN
PaC	0.06	MPSNR	24.436	36.780	35.940	35.903	37.671	36.681	**39.575**
		MSSIM	0.6530	0.9660	0.9581	0.9601	0.9690	0.9630	**0.9821**
		Time	-	68.418	**1.0871**	30.529	33.615	156.64	1189.9
	0.08	MPSNR	21.933	35.016	34.284	35.006	34.186	34.116	**37.790**
		MSSIM	0.5330	0.9500	0.9411	0.9512	0.9460	0.9361	**0.9730**
		Time	-	68.461	**1.2840**	30.044	30.055	156.25	1187.9
	0.1	MPSNR	20.006	33.680	33.192	34.094	32.840	32.031	**36.255**
		MSSIM	0.4370	0.9340	0.9271	0.9400	0.9271	0.9030	**0.9622**
		Time	-	68.320	**2.0371**	30.252	30.119	157.91	1205.9
WDC	0.06	MPSNR	24.439	36.197	37.256	33.160	**38.993**	36.62	37.521
		MSSIM	0.6771	0.9651	0.9700	0.9340	**0.9780**	0.9680	0.9761
		Time	-	345.63	**2.9821**	98.633	127.41	400.83	4814.7
	0.08	MPSNR	21.939	34.343	35.142	33.004	35.931	35.703	**37.168**
		MSSIM	0.5661	0.9471	0.9542	0.9311	0.9565	0.9602	**0.9741**
		Time	-	340.62	**3.1151**	98.032	147.01	389.33	4604.0
	0.1	MPSNR	20.000	32.924	33.792	32.822	33.267	34.609	**36.624**
		MSSIM	0.4741	0.9271	0.9405	0.9291	0.9232	0.9505	**0.9711**
		Time	-	344.61	**2.5891**	99.766	199.25	393.18	4616.5

hensive comparison, we employed the double local magnification technique to select two representative regions of the dataset for magnification comparison. The same areas in each subgraph are highlighted with blue and green boxes. As depicted in Fig. 5, the NLFCTN method effectively eliminates Gaussian noise, yielding clearer and more detailed results.

Table 2 shows the results of comparison experiments on WDC simulation dataset. The NLFCTN method proposed in this paper demonstrates superior results in terms of MPSNR and MSSIM measurement metrics, particularly under high noise levels. Compared with matrix method, NLFCTN has better effect under MPSNR. This demonstrates that tensors are more effective in capturing the global data correlation compared to matrices. By observing Fig. 6, it can be seen that NLFCTN is more effective in reconstructing the contaminated spectral curve on PaC dataset than other methods. Moreover, our proposed NLFCTN method outperforms the alternative method, particularly at high noise levels, showcasing its enhanced robustness for RSI denoising. Fig. 7 presents the visualization results of the reconstruction obtained by our method alongside the comparison method, providing visual evidence of the efficacy of our approach. By observing Fig. 8, it can be seen that the NLFCTN method effectively reconstructs the spectral curve polluted by noise.

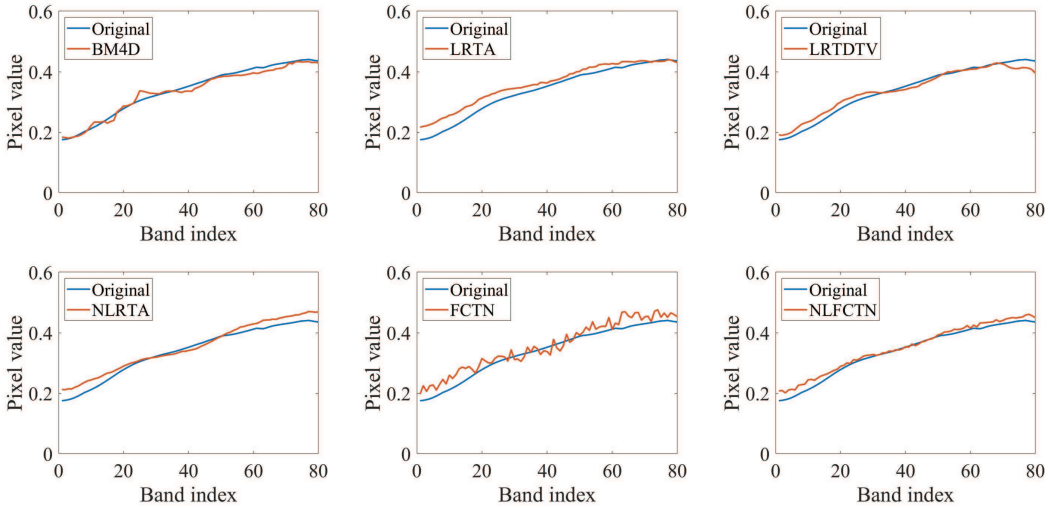


Figure 6: Spectral curves reconstructed by the NLFCTN method and other comparison methods on the PaC dataset.

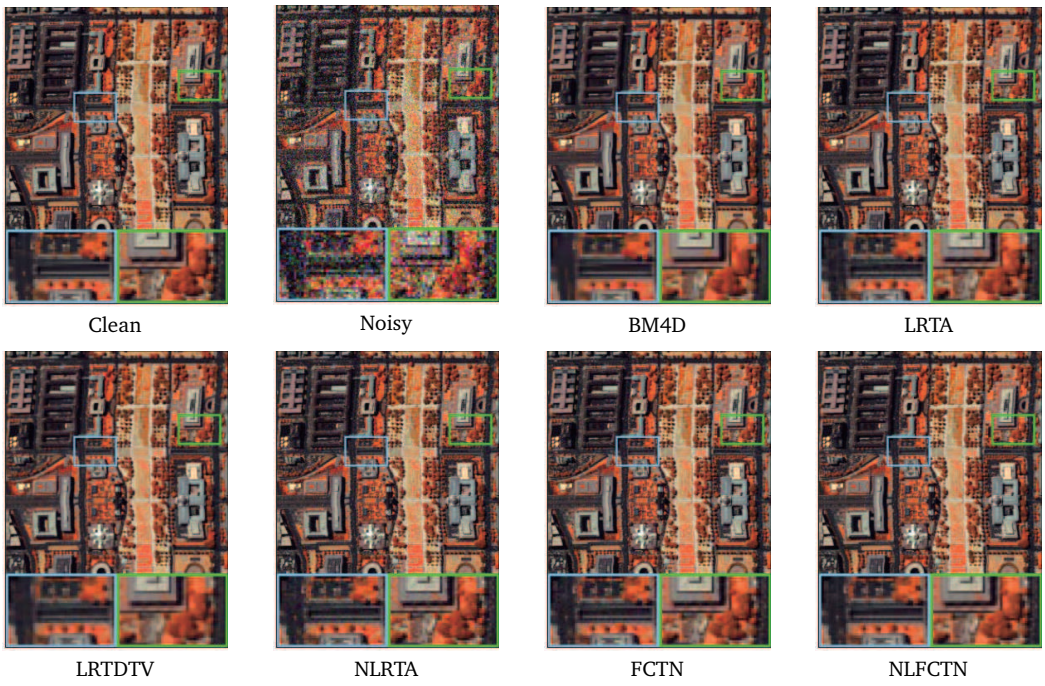


Figure 7: Experimental results on the WDC dataset. Gaussian noise with variance $\sigma = 0.1$ was added to the simulation experiment. For a better view of the visualization, we performed a double local magnification, where the blue and green boxes correspond to the enlarged area.

5.4. Real experiments on HSI data

In order to ascertain the efficacy of our proposed method, we conducted a comparative analysis against existing techniques using two real-noise HSIs. Since real-noise

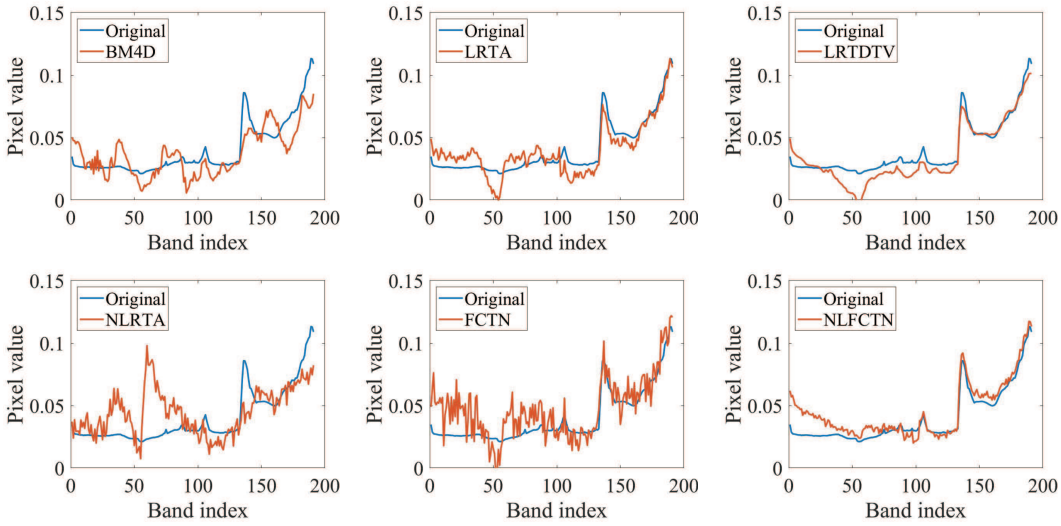


Figure 8: Spectral curves reconstructed by the NLFCTN method and other comparison methods on the WDC dataset.

HSIs typically lack corresponding noise-free reference images, we are constrained to solely conduct qualitative assessments of the denoising performance using real data. We assess the performance of our proposed NLFCTN method on real datasets by utilizing the Indian Pines[§] and Urban[§] datasets.

Benchmark DataSet: The Indian Pines dataset was acquired using AVIRIS and consists of 145×145 pixels and 224 bands. In our experiments, we utilized sub-images with a size of $145 \times 145 \times 200$, with 20m/pixel spatial resolution. The Urban dataset was collected using HYDICE and comprises 307×307 pixels and 210 spectral bands, with 2m/pixel spatial resolution. In our experiments, we employed the entire image of the Urban dataset.

Experimental results: Fig. 9 depicts the pseudo-color images of the original Indian Pines dataset and the corresponding denoised results. It can be seen from Fig. 9 that the noise reduction effect of LRTA method is poor, the image reconstructed by BM4D method is too smooth, while the noise reduction effect of LRTDTV, NLRTA, FCTN and our proposed NLFCTN method is better. For a detailed comparison, we employed the double local magnification technique to magnify the same region in each sub-map. The denoising results of the LRTA show edge distortion or blurring, while the BM4D, LRTDTV, NLRTA, FCTN, and NLFCTN retain sharp edge details. Fig. 10 presents the pseudo-color images of the Urban dataset. It can be observed from Fig. 10 that the original image is contaminated with Gaussian noise and streaks. The results from LRTA and NLRTA retained some traces of blue noise and streaks, while the results from BM4D

[§]https://rslab.ut.ac.ir/data

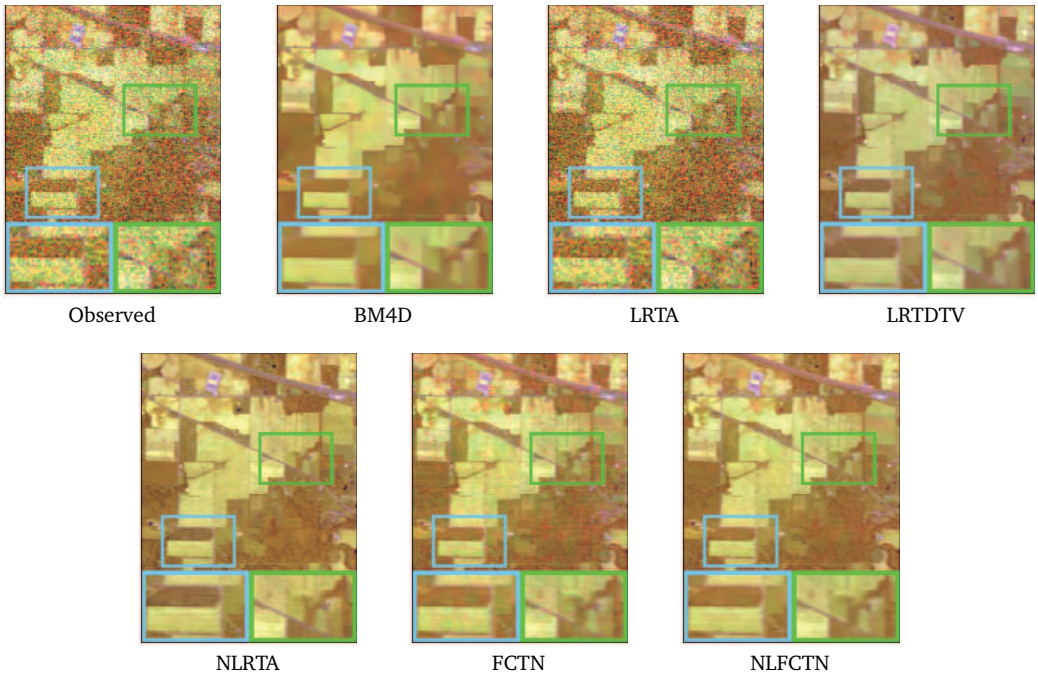


Figure 9: Experimental results on the Indian Pines dataset. For a better view of the visualization, we performed a double local magnification, where the blue and green boxes correspond to the enlarged area.

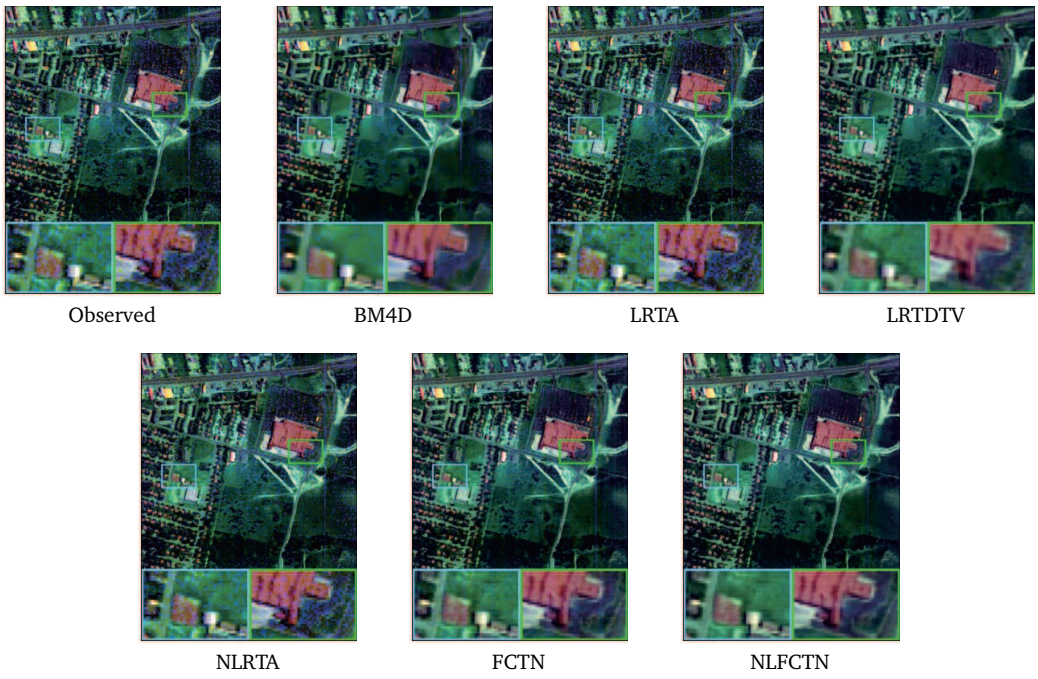


Figure 10: Experimental results on the Urban dataset. For a better view of the visualization, we performed a double local magnification, where the blue and green boxes correspond to the enlarged area.

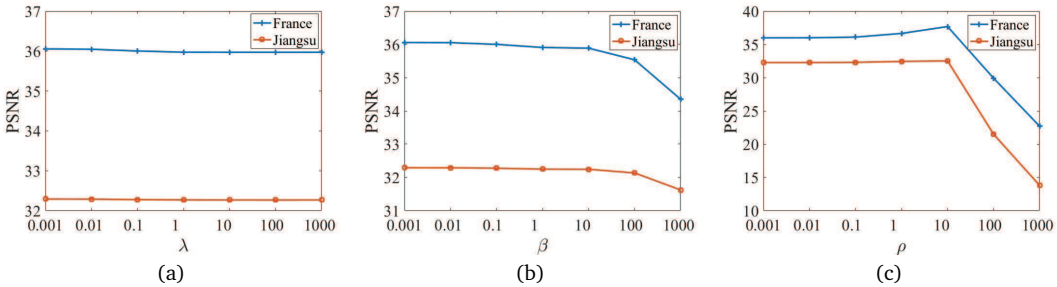
and LRTDTV smoothed the image excessively. On the other hand, FCTN and our proposed NLFCTN method successfully remove Gaussian noise and blue stripes.

6. Discussion

In this section, we quantitatively analyze the influence of parameters on the NLFCTN method. There are two types of parameters in our model, one is the regularization parameter λ, and the other is the penalty parameter β and ρ. Next, we analyze the influence of these two types of parameters on the model respectively.

Regularization parameter λ: This parameter is used to adjust the weight of the regular term $\|\widehat{\mathcal{N}}_l\|_F^2$. The larger the regularization parameter λ, the greater the weight of the regular term $\|\widehat{\mathcal{N}}_l\|_F^2$. We conducted experiments on the France and Jiangsu datasets. The Gaussian noise variance $\sigma = 0.06$ added to the experiment, and the value of λ was selected from $\{0.001, 0.01, 0.1, 1, 10, 100, 1000\}$. By observing Fig. 11(a), it can be seen that the model is not very sensitive to the regularization parameter λ. As the value of λ changes, the PSNR value of the simulated experiment fluctuates in a stable interval. Therefore, we suggest that λ can be selected in the range of $\{0.001, 0.01, 0.1, 1, 10, 100, 1000\}$.

Penalty parameter β and ρ: The penalty parameters β and ρ are used to balance the increased penalty term. Similar to the regularization parameter λ analysis, the penalty parameters β and ρ were also subjected to parameter sensitivity experiments on the France and Jiangsu datasets, and the experimental settings and parameter selection ranges were consistent. By observing Fig. 11(b), we can find that the performance of the NLFCTN method decreases when the parameter β is large. Especially when $\beta = 1000$, our proposed method has limited recovery performance on both datasets. Therefore, we recommend β to be selected between $\{0.001, 0.01, 0.1, 1, 10\}$, and the RSI reconstruction effect is better. In addition, Fig. 11(c) shows the effect of parameter ρ on the model. It can be seen that the PSNR values of the two data sets have a process



(a) (b) (c)

Figure 11: Parameter sensitivity analysis on France and Jiangsu datasets, Gaussian noise level $\sigma = 0.06$. (a) The PSNR value varies with the regularization parameter λ. (b) The PSNR value changes with the penalty parameter β. (c) The PSNR value changes with the penalty parameter ρ.

of rising first and then falling with the parameter ρ. For this reason, we recommend that the parameter ρ be selected between $\{0.1, 1, 10\}$.

7. Conclusions

Compared with conventional images, RSIs contain more spectral information and spatial features with a higher number of bands, and they play an increasingly important role in the fields of feature identification and military operations. In this paper, we first apply the FCTN algorithm to the global image for initial denoising, and then perform NSS image block finding on the initial denoised image, and recombine the image blocks with NSS to form NSS image blocks, which improves the dimensionality of the input data of the subsequent FCTN algorithm. The FCTN algorithm is continued for subsequent NSS image blocks, which both exploits the similar image information of NSS image blocks and the ability of FCTN to not destroy tensor potential information on higher-order tensors. Finally, each group of noise reduced NSS image blocks is reassembled to form the final denoised global image. In this process we utilize the PAM algorithm for model solving, which ensures the convergence of the results. The final experimental results show that the proposed NLFCTN denoising algorithm can achieve significant improvement in MPSNR and MSSIM data indexes for remote sensing images with high noise pollution.

Acknowledgement

This work is supported in part by the National Natural Science Foundation of China (Grant Nos. U21A20455, 61972265, 12326619, 62071349), by the Natural Science Foundation of Guangdong Province of China (Grant Nos. 2020B1515310008, 2023A1515011691), by the Educational Commission of Guangdong Province of China (Grant No. 2019KZDZX1007), and by the Key Research and Development Program of Shaanxi (Program No. 2023-YBGY-223).

References

[1] H. ATTOUCH, J. BOLTE, P. REDONT, AND A. SOUBEYRAN, *Proximal alternating minimization and projection methods for nonconvex problems: An approach based on the Kurdyka-Łojasiewicz inequality*, Math. Oper. Res. 35(2) (2010), 438–457.

[2] H. ATTOUCH, J. BOLTE, AND B. F. SVAITER, *Convergence of descent methods for semi-algebraic and tame problems: Proximal algorithms, forward-backward splitting, and regularized Gauss-Seidel methods*, Math. Program. 137(1-2) (2013), 91–129.

[3] J. M. BIOUCAS-DIAS AND J. M. P. NASCIMENTO, *Hyperspectral subspace identification*, IEEE Trans. Geosci. Remote. Sens. 46(8) (2008), 2435–2445.

[4] J. BOLTE, S. SABACH, AND M. TEBOULLE, *Proximal alternating linearized minimization for nonconvex and nonsmooth problems*, Math. Program. 146(1-2) (2014), 459–494.

[5] Y. CHEN, W. HE, N. YOKOYA, T.-Z. HUANG, AND X.-L. ZHAO, *Nonlocal tensor-ring decomposition for hyperspectral image denoising*, IEEE Trans. Geosci. Remote. Sens. 58(2) (2019), 1348–1362.
[6] Y. CHEN, W. HE, X.-L. ZHAO, T.-Z. HUANG, J. ZENG, AND H. LIN, *Exploring nonlocal group sparsity under transform learning for hyperspectral image denoising*, IEEE Trans. Geosci. Remote. Sens. 60 (2022), 1–18.
[7] Y. CHEN, T.-Z. HUANG, X.-L. ZHAO, L.-J. DENG, AND J. HUANG, *Stripe noise removal of remote sensing images by total variation regularization and group sparsity constraint*, Remote Sens. 9(6) (2017), 559.
[8] K. DABOV, A. FOI, V. KATKOVNIK, K. EGIAZARIAN, *Image denoising by sparse 3-d transform-domain collaborative filtering*, IEEE Trans. Image Process. 16(8) (2007), 2080–2095.
[9] X. GONG, W. CHEN, J. CHEN, AND B. AI, *Tensor denoising using low-rank tensor train decomposition*, IEEE Signal Process. Lett. 27 (2020), 1685–1689.
[10] J. GUO, Y. GUO, Q. JIN, M. KWOK-PO NG, AND S. WANG, *Gaussian patch mixture model guided low-rank covariance matrix minimization for image denoising*, SIAM J. Imaging Sci. 15(4) (2022), 1601–1622.
[11] X. GUO, X. HUANG, L. ZHANG, AND L. ZHANG, *Hyperspectral image noise reduction based on rank-1 tensor decomposition*, ISPRS journal of Photogrammetry and Remote Sensing 83 (2013), 50–63.
[12] W. HE, N. YOKOYA, L. YUAN, AND Q. ZHAO, *Remote sensing image reconstruction using tensor ring completion and total variation*, IEEE Trans. Geosci. Remote. Sens. 57(11) (2019), 8998–9009.
[13] T. G. KOLDA AND B. W. BADER, *Tensor decompositions and applications*, SIAM Rev. 51(3) (2009), 455–500.
[14] X. LIU, S. BOURENNANE, AND C. FOSSATI, *Denoising of hyperspectral images using the parafac model and statistical performance analysis*, IEEE Trans. Geosci. Remote. Sens. 50(10) (2012), 3717–3724.
[15] J. LU, J. TIAN, Q. JIANG, X. LIU, Z. HU, AND Y. ZOU, *Rician noise removal via weighted nuclear norm penalization*, Appl. Comput. Harmon. Anal. 53 (2021), 180–198.
[16] J. LU, C. XU, Z. HU, X. LIU, Q. JIANG, D. MENG, AND Z. LIN, *A new nonlocal low-rank regularization method with applications to magnetic resonance image denoising*, Inverse Problems 38(6) (2022), 065012.
[17] J. LU, Y. YE, Y. DONG, X. LIU, AND Y. ZOU, *Impulse noise removal by L1 weighted nuclear norm minimization*, J. Comput. Math. 41(6) (2023), 1171–1191.
[18] 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(2) (2022), 49.
[19] M. MAGGIONI, V. KATKOVNIK, K. EGIAZARIAN, AND A. FOI, *Nonlocal transform-domain filter for volumetric data denoising and reconstruction*, IEEE Trans. Image Process. 22(1) (2012), 119–133.
[20] I. V. OSELEDETS, *Tensor-train decomposition*, SIAM J. Sci. Comput. 33(5) (2011), 2295–2317.
[21] N. RENARD, S. BOURENNANE, AND J. BLANC-TALON, *Denoising and dimensionality reduction using multilinear tools for hyperspectral images*, IEEE Geosci. Remote. Sens. Lett. 5(2) (2008), 138–142.
[22] J. SUI, X. MA, X. ZHANG, AND M.-O. PUN, *GCRDN: Global context-driven residual dense network for remote sensing image super-resolution*, IEEE J. Sel. Top. Appl. Earth Obs. Remote. Sens. 16 (2023), 4457–4468.

[23] Z. TU, L. GUO, H. PAN, J. LU, C. XU, AND Y. ZOU, *Multitemporal image cloud removal using group sparsity and nonconvex low-rank approximation*, J. Nonlinear Var. Anal. 7(4) (2023), 527–548.

[24] Z. TU, J. LU, H. ZHU, H. PAN, W. HU, Q. JIANG, AND Z. LU, *A new nonconvex low-rank tensor approximation method with applications to hyperspectral images denoising*, Inverse Problems 39(6) (2023), 065003.

[25] W. WANG, N. TIAN, AND C. WU, *Two-phase image segmentation by nonconvex nonsmooth models with convergent alternating minimization algorithms*, J. Comput. Math. 41(4) (2023), 588–622.

[26] X. WANG AND G. XU, *A fast classification method for single-particle projections with a translation and rotation invariant*, J. Comput. Math. (2013), 137–153.

[27] Y. WANG, J. PENG, Q. ZHAO, Y. LEUNG, X.-L. ZHAO, AND D. MENG, *Hyperspectral image restoration via total variation regularized low-rank tensor decomposition*, IEEE J. Sel. Top. Appl. Earth Obs. Remote. Sens. 11(4) (2017), 1227–1243.

[28] Z. WANG, A. C. BOVIK, H. R. SHEIKH, AND E. P. SIMONCELLI, *Image quality assessment: from error visibility to structural similarity*, IEEE Trans. Image Process. 13(4) (2004), 600–612.

[29] C. XU, X. LIU, J. ZHENG, L. SHEN, Q. JIANG, AND J. LU, *Nonlocal low-rank regularized two-phase approach for mixed noise removal*, Inverse Problems 37(8) (2021), 085001.

[30] H. ZHANG, W. HE, L. ZHANG, H. SHEN, AND Q. YUAN, *Hyperspectral image restoration using low-rank matrix recovery*, IEEE Trans. Geosci. Remote Sens 52(8) (2013), 4729–4743.

[31] H. ZHANG, T.-Z. HUANG, X.-L. ZHAO, W. HE, J. K. CHOI, AND Y.-B. ZHENG, *Hyperspectral image denoising: Reconciling sparse and low-tensor-ring-rank priors in the transformed domain*, IEEE Trans. Geosci. Remote Sens. 61 (2023), 1-13.

[32] X. ZHANG, W. YU, AND M.-O. PUN, *Multilevel deformable attention-aggregated networks for change detection in bitemporal remote sensing imagery*, IEEE Trans. Geosci. Remote Sens. 60 (2022), 1–18.

[33] Q. ZHAO, G. ZHOU, S. XIE, L. ZHANG, AND A. CICHOCKI, *Tensor ring decomposition*, arXiv:1606.05535 (2016).

[34] X.-L. ZHAO, J.-H. YANG, T.-H. MA, T.-X. JIANG, M. K. NG, AND T.-Z. HUANG, *Tensor completion via complementary global, local, and nonlocal priors*, IEEE Trans. Image Process. 31 (2021), 984–999.

[35] X.-L. ZHAO, H. ZHANG, T.-X. JIANG, M. K. NG, AND X.-J. ZHANG, *Fast algorithm with theoretical guarantees for constrained low-tubal-rank tensor recovery in hyperspectral images denoising*, Neurocomputing 413 (2020), 397–409.

[36] W.-J. ZHENG, X.-L. ZHAO, Y.-B. ZHENG, AND Z.-F. PANG, *Nonlocal patch-based fully connected tensor network decomposition for multispectral image inpainting*, IEEE Geosci. Remote Sens. Lett. 19 (2021), 1–5.

[37] Y.-B. ZHENG, T.-Z. HUANG, X.-L. ZHAO, AND Q. ZHAO, *Tensor completion via fully-connected tensor network decomposition with regularized factors*, J. Sci. Comput. 92(1) (2022), 8.

[38] 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 Conference On Artificial Intelligence 35, (2021), 11071–11078.

[39] L. ZHUANG, X. FU, M. K. NG, AND J. M. BIOUCAS-DIAS, *Hyperspectral image denoising based on global and nonlocal low-rank factorizations*, IEEE Trans. Geosci. Remote Sens. 59(12) (2021), 10438–10454.