3D Poissonian image deblurring via patch-based tensor logarithmic Schatten-*p* minimization

To cite this article before publication: Jian Lu *et al* 2024 *Inverse Problems* in press https://doi.org/10.1088/1361-6420/ad40c9

3D Poissonian Image Deblurring via Patch-based Tensor Logarithmic Schatten-p Minimization

Jian Lu*‡, Lin Huang†‡, Xiaoxia Liu§‖, Ning Xie¶‖, Qingtang Jiang**, Yuru Zou†

Abstract

In medical and biological image processing, multi-dimensional images are often corrupted by blur and Poisson noise. In this paper, we first propose a new tensor logarithmic Schatten-p (t-log-S_p) low-rank measure and a tensor iteratively reweighted Schatten-p minimization (t-IRSpM) algorithm for minimizing such measure. Furthermore, we adopt this low-rank measure to regularize the non-local tensors formed by similar 3D image patches and develop a patch-based non-local low-rank model. The data fidelity term of the model characterizes the Poisson noise distribution and blur operator. The optimization model is further solved by an alternating minimization technique combined with variable splitting. Experimental results tested on 3D fluorescence microscope images show that the proposed patch-based tensor logarithmic Schatten-p minimization (TLSpM) method outperforms state-of-the-art methods in terms of image evaluation metrics and visual quality.

Keywords: tensor low-rank measure, non-local low-rank regularization, Poisson noise, deblurring.

1 Introduction

Image degradation by blur and Poisson noise is inevitable in electronic microscopy [1], astronomical imaging [2], single particle emission computed tomography (SPECT) [3,4], positron emission tomography (PET) [5], and so on. On one hand, images are convoluted by a point spread function (PSF) of the imaging device or body movement caused by the respiratory shake of the patient. On the other hand, due to the low photon count [6], images such as X-ray tomography [7], fluorescence microscopes [1], astronomy [2], mammography [8], and tomosynthesis [9], are often affected by Poisson noise.

For deconvoluting Poissonian images, a popular method is the Richardson-Lucy (RL) algorithm [10], which calculated a Poisson maximum likelihood estimate. The Ameliorated Richardson-Lucy (ARL) algorithm [11] accelerated the deblurring procedure of the RL algorithm. But the RL and ARL algorithms may amplify the noise after several iterations. To efficiently restore blurry Poissonian images, various optimization models with regularization terms were developed and further solved by efficient algorithms. The most commonly used regularization is the total variation (TV) regularization [12–19]. Dey et al. [12] enhanced the RL algorithm by the TV regularization; Harmany et al. [13] solved the TV regularized

*Shenzhen Key Laboratory of Advanced Machine Learning and Applications, College of Mathematics and Statistics, Shenzhen University, Shenzhen 518060, China; National Center for Applied Mathematics Shenzhen (NCAMS), Shenzhen 518055, China.

†Shenzhen Key Laboratory of Advanced Machine Learning and Applications, College of Mathematics and Statistics, Shenzhen University, Shenzhen 518060, China.

‡J. Lu and L. Huang contributed equally to this work.

§Department of Applied Mathematics, The Hong Kong Polytechnic University, Hong Kong, China.

¶Guangdong Key Laboratory of Intelligent Information Processing, College of Information Engineering, Shenzhen University, Shenzhen 518060, China.

‖Authors to whom any correspondence should be addressed. Emails: xiaoxia.liu@polyu.edu.hk and ningxie@szu.edu.cn.

**Department of Mathematics and Statistics, University of Missouri-St. Louis, St. Louis, MO 63121, United States of America.

model by sequential quadratic approximations; Bonettini et al. [14] combined a Poisson log-likelihood data fidelity term with the TV regularization term and used an alternating extragradient algorithm to solve the model; Figueiredo et al. [15] solved the model by the alternating direction method of multipliers; Ma et al. [16] proposed a dictionary learning model in addition to the TV regularization for Poissonian image restoration. Other regularizations such as wavelet based regularizations [20–25] and Hessian Schatten norm regularization [26] were also proposed. However, those regularization techniques are primarily designed for 2D images and cannot be easily extended to 3D images.

Recent approaches for 3D Poissonian image deblurring converted Poisson noise into Gaussian noise through some transformations and then restored the image via denoising tools for Gaussian noise. Dupe et al. [27] utilized the Anscombe variance stable transformation (VST) [28] leading to Gaussian noise and denoised the blurry Gaussian image by a convex optimization model; Azzari et al. [29] deconvolved the blurry image by a linear regularized inverse filter and then adopted VST and block matching 3D (BM3D) [30] or BM4D [31] to remove Poisson noise. Besides these approaches, the methods based on the Poisson unbiased risk estimate (PURE) also achieved great performance. The PURE-LET method that characterized the deconvolution process as a linear combination of elementary functions (LET) was proposed in [32] for 2D images and in [33] for 3D images. Each LET function contains a Wiener filtering and wavelet-domain thresholding and the PURE is used to estimate the coefficients of the linear combination.

In this paper, we propose a patch-based approach for 3D Poissonian image deblurring. First, a new tensor low-rank measure called the t-log-S_p low-rank measure is proposed, and an efficient algorithm with convergence results is also proposed for minimizing such measure. Second, according to the image non-local self-similarity, we use the proposed tensor low-rank measure to regularize the low-rankness of the tensors formed by similar 3D patches extracted from the 3D image. Then we further propose a non-local low-rank model with a data fidelity term for Poissonian deblurring and solve it by an alternating minimization algorithm with a proximal term. Lastly, we demonstrate the proposed method outperforms the state-of-the-art methods in removing Poisson noise and deblurring of fluorescence microscope images.

The main contributions of this paper are as follows:

- We propose a matrix logarithmic Schatten-p (log-S_p) low-rank measure for 2D images, which can reveal the weighting strategy used in the weighted Schatten p-norm minimization [34]. Then we further extend the log-S_p low-rank measure to tensor log-S_p (t-log-S_p) low-rank measure for 3D images. It can be demonstrated in this paper that the t-log-S_p measure is efficient and suitable for applications in 3D image restoration such as 3D Poissonian image deblurring.

- For the proposed log-S_p and t-log-S_p measures, we introduce some properties and develop reliable solvers for their minimization problems. In particular, we develop an iteratively reweighted S_p minimization (IRSpM) algorithm for the log-S_p minimization and a tensor IRSpM (t-IRSpM) algorithm for the t-log-S_p minimization. A convergence analysis of each algorithm is provided in detail, showing any accumulation point generated by the algorithm is a stationary point of the problem.

- We build a new patch-based non-local low-rank model using the proposed t-log-S_p measure for 3D Poissonian image deblurring. This approach can achieve state-of-the-art performance for 3D Poissonian image deblurring.

This paper is organized as follows. In section 2, we provide some tensor notations and definitions, then introduce matrix and tensor logarithmic Schatten-p (log-S_p) low-rank measures and their properties. To solve the matrix and tensor log-S_p minimization problems, in section 3 we propose matrix and tensor iteratively reweighted Schatten-p minimization (IRSpM) algorithms, respectively, along with convergence analysis. We further develop our model for 3D Poissonian image deblurring in section 4. Experimental results tested on 3D fluorescence microscope images are presented in section 5. Section 6 concludes this paper.

2 Tensor Logarithmic Schatten-p Low-rank Measure

In this section, we first introduce the definitions and notations of tensors including tensor singular value decomposition (t-SVD). Then we propose a t-log-S_p low-rank measure and present its properties.

2.1 Preliminaries on tensors

Tensors are represented by bold calligraphy letters, e.g., $\boldsymbol{\mathcal{X}}$; matrices are represented by bold capital letters, e.g., $\boldsymbol{X}$; vectors are represented by bold lowercase letters, e.g., $\boldsymbol{x}$; and scalars are represented by lowercase letters, e.g., x. For an N-order tensor $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_N}$, the vectorization of $\boldsymbol{\mathcal{X}}$ is denoted as $\boldsymbol{x} = \mathrm{vec}(\boldsymbol{\mathcal{X}}) \in \mathbb{R}^{n_1 n_2 \ldots n_N}$, and the j-th element of $\boldsymbol{x}$ is equal to the $(i_1, i_2, \ldots, i_N)$-th element of $\boldsymbol{\mathcal{X}}$ with $j = i_1 + \sum_{s=2}^{N} \left((i_s - 1) \prod_{m=1}^{s-1} n_m\right)$. The mode-$k$ tensor matricization of $\boldsymbol{\mathcal{X}}$ is denoted as $\boldsymbol{X}_{(k)} \in \mathbb{R}^{n_k \times \prod_{s \neq k} n_s}$, and the (i_k, j)-th element of $\boldsymbol{X}_{(k)}$ is equal to the $(i_1, i_2, \ldots, i_N)$-th element of $\boldsymbol{\mathcal{X}}$, where $j = 1 + \sum_{s=1, s\neq k}^{N} (i_s - 1) J_s$ with $J_s = \prod_{m=1, m\neq k}^{s-1} n_m$. And the operator "unfold" and its inverse operator "fold" are defined by $\boldsymbol{X}_{(k)} = \mathrm{unfold}_{(k)}(\boldsymbol{\mathcal{X}})$ and $\boldsymbol{\mathcal{X}} = \mathrm{fold}_{(k)}(\boldsymbol{X}_{(k)})$, respectively.

For a 3-order tensor $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, x_{ijk} denotes the (i, j, k)-th entry of $\boldsymbol{\mathcal{X}}$, $\boldsymbol{X}^{(k)}$ denotes the k-th frontal slice $\boldsymbol{\mathcal{X}}(:, :, k)$, and $\overline{\boldsymbol{\mathcal{X}}}$ denotes the discrete Fourier transform (DFT) of $\boldsymbol{\mathcal{X}}$ along the 3-rd dimension, i.e., $\overline{\boldsymbol{\mathcal{X}}} = \mathrm{fft}(\boldsymbol{\mathcal{X}}, [\,], 3)$. This also implies $\boldsymbol{\mathcal{X}} = \mathrm{ifft}(\overline{\boldsymbol{\mathcal{X}}}, [\,], 3)$. $\overline{\boldsymbol{X}}^{(i)}$ denotes the i-th frontal slice of $\overline{\boldsymbol{\mathcal{X}}}$. The block diagonal matrix of $\overline{\boldsymbol{\mathcal{X}}}$ is defined as

$$\mathrm{bdiag}(\overline{\boldsymbol{\mathcal{X}}}) = \begin{bmatrix} \overline{\boldsymbol{X}}^{(1)} & & & \\ & \overline{\boldsymbol{X}}^{(2)} & & \\ & & \ddots & \\ & & & \overline{\boldsymbol{X}}^{(n_3)} \end{bmatrix},$$

and the block circulant matrix of $\boldsymbol{\mathcal{X}}$ is defined as a matrix of size $n_1 n_3 \times n_2 n_3$ having the following form:

$$\mathrm{bcirc}(\boldsymbol{\mathcal{X}}) = \begin{bmatrix} \boldsymbol{X}^{(1)} & \boldsymbol{X}^{(n_3)} & \ldots & \boldsymbol{X}^{(2)} \\ \boldsymbol{X}^{(2)} & \boldsymbol{X}^{(1)} & \ldots & \boldsymbol{X}^{(3)} \\ \vdots & \vdots & \ddots & \vdots \\ \boldsymbol{X}^{(n_3)} & \boldsymbol{X}^{(n_3-1)} & \ldots & \boldsymbol{X}^{(1)} \end{bmatrix}.$$

As for block unfolding $\boldsymbol{\mathcal{X}}$ and its inverse operation, the operations are defined as follows:

$$\mathrm{bvec}(\boldsymbol{\mathcal{X}}) = \begin{pmatrix} \boldsymbol{X}^{(1)} \\ \boldsymbol{X}^{(2)} \\ \vdots \\ \boldsymbol{X}^{(n_3)} \end{pmatrix}, \quad \mathrm{bfold}(\mathrm{bvec}(\boldsymbol{\mathcal{X}})) = \boldsymbol{\mathcal{X}}.$$

For a 3-order tensor $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, the Frobenius norm of $\boldsymbol{\mathcal{X}}$ is $\|\boldsymbol{\mathcal{X}}\|_F = \sqrt{\sum_{ijk} |x_{ijk}|^2}$ and the tensor transpose of $\boldsymbol{\mathcal{X}}$ is $\boldsymbol{\mathcal{X}}^T \in \mathbb{R}^{n_2 \times n_1 \times n_3}$ defined as

$$\boldsymbol{\mathcal{X}}^T = \mathrm{bfold}\left(\left[\boldsymbol{X}^{(1)}, \boldsymbol{X}^{(n_3)}, \boldsymbol{X}^{(n_3-1)}, \ldots, \boldsymbol{X}^{(2)}\right]^T\right).$$

Using the tensor notations above, we present the definition of a tensor product.

Definition 2.1 [35] (t-product). *Let $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{n_1\times n_2\times n_3}$ and $\boldsymbol{\mathcal{Y}} \in \mathbb{R}^{n_2\times l\times n_3}$. Then the t-product of $\boldsymbol{\mathcal{X}}$ and $\boldsymbol{\mathcal{Y}}$ is $\boldsymbol{\mathcal{Z}} \in \mathbb{R}^{n_1\times l\times n_3}$ defined as:*

$$\boldsymbol{\mathcal{Z}} = \boldsymbol{\mathcal{X}} * \boldsymbol{\mathcal{Y}} = \mathrm{bfold}(\mathrm{bcirc}(\boldsymbol{\mathcal{X}}) \cdot \mathrm{bvec}(\boldsymbol{\mathcal{Y}})).$$

Note that if $n_3 = 1$, the operator $*$ reduces to matrix multiplication.

In fact, the t-product can also be calculated via the following equivalence under the DFT:

$$\overline{\boldsymbol{Z}}^{(i)} = \overline{\boldsymbol{X}}^{(i)}\overline{\boldsymbol{Y}}^{(i)}, \tag{1}$$

that is, the i-th frontal slice of the DFT of the t-product is equal to the matrix product of the i-th frontal slices of the DFT of $\boldsymbol{\mathcal{X}}$ and $\boldsymbol{\mathcal{Y}}$.

Definition 2.2 [35] (Identity tensor). *The identity tensor $\boldsymbol{\mathcal{I}} \in \mathbb{R}^{n\times n\times n_3}$ is the tensor whose first frontal slice is the $n \times n$ identity matrix, and other frontal slices are all zeros.*

Definition 2.3 [35] (Orthogonal tensor). *A tensor $\boldsymbol{\mathcal{Q}} \in \mathbb{R}^{n\times n\times n_3}$ is orthogonal if it satisfies $\boldsymbol{\mathcal{Q}}^T\boldsymbol{\mathcal{Q}} = \boldsymbol{\mathcal{Q}}\boldsymbol{\mathcal{Q}}^T = \boldsymbol{\mathcal{I}}$.*

Definition 2.4 [35] (F-diagonal tensor). *A tensor is F-diagonal if each frontal slice is diagonal.*

Note that each frontal slice of $\overline{\boldsymbol{\mathcal{I}}}$ is the identity matrix and each frontal slide of $\overline{\boldsymbol{\mathcal{Q}}}$, where $\boldsymbol{\mathcal{Q}}$ is orthogonal, is an orthogonal matrix. Next, we present the definition of a t-SVD and several tensor low-rank measures.

Theorem 2.5 [35] (t-SVD). *Let $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{n_1\times n_2\times n_3}$. Then there exist $\boldsymbol{\mathcal{U}} \in \mathbb{R}^{n_1\times n_1\times n_3}$, $\boldsymbol{\mathcal{S}} \in \mathbb{R}^{n_1\times n_2\times n_3}$ and $\boldsymbol{\mathcal{V}} \in \mathbb{R}^{n_2\times n_2\times n_3}$ such that:*

$$\boldsymbol{\mathcal{X}} = \boldsymbol{\mathcal{U}} * \boldsymbol{\mathcal{S}} * \boldsymbol{\mathcal{V}}^T, \tag{2}$$

where $\boldsymbol{\mathcal{U}}$ and $\boldsymbol{\mathcal{V}}$ are orthogonal, and $\boldsymbol{\mathcal{S}}$ is a frontal-slice-diagonal tensor.

Definition 2.6 [36] (Tensor tubal rank). *For a 3D tensor $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{n_1\times n_2\times n_3}$, the tensor tubal rank of $\boldsymbol{\mathcal{X}}$, denoted as $\mathrm{rank}_t(\boldsymbol{\mathcal{X}})$, is defined as the number of non-zero tubes of $\boldsymbol{\mathcal{S}}$ where $\boldsymbol{\mathcal{S}}$ is from the t-SVD of $\boldsymbol{\mathcal{X}} = \boldsymbol{\mathcal{U}} * \boldsymbol{\mathcal{S}} * \boldsymbol{\mathcal{V}}^T$. That is,*

$$\mathrm{rank}_t(\boldsymbol{\mathcal{X}}) = \#\{i : \boldsymbol{\mathcal{S}}(i, i, :) \neq \mathbf{0}\}.$$

The tensor tubal rank is a tensor low-rank measure based on t-SVD, which counts the number of non-zero tubes in t-SVD. In fact, the tensor tubal rank only depends on the first frontal slice of $\boldsymbol{\mathcal{S}}$, that is, $\mathrm{rank}_t(\boldsymbol{\mathcal{X}}) = \#\{i : \boldsymbol{\mathcal{S}}(i, i, 1) \neq 0\}$. Since the tensor tubal rank minimization is NP-hard, several tensor low-rank measures were proposed to approximate the tensor tubal rank.

Definition 2.7 [37] (Tensor nuclear norm). *Let $\boldsymbol{\mathcal{X}} = \boldsymbol{\mathcal{U}} * \boldsymbol{\mathcal{S}} * \boldsymbol{\mathcal{V}}^T$ be the t-SVD of $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{n_1\times n_2\times n_3}$. Then the tensor nuclear norm is defined as*

$$\|\boldsymbol{\mathcal{X}}\|_* = \sum_{i=1}^{n_3} \boldsymbol{\mathcal{S}}(i, i, 1).$$

The tensor nuclear norm can also be computed via the frontal slices of $\overline{\boldsymbol{\mathcal{X}}}$, that is,

$$\|\boldsymbol{\mathcal{X}}\|_* = \frac{1}{n_3}\sum_{i=1}^{n_3} \|\overline{\boldsymbol{X}}^{(i)}\|_*.$$

As the frontal slices $\overline{\boldsymbol{X}}^{(i)}$ are matrices, the matrix nuclear norm can be replaced by other non-convex surrogates of the matrix rank. For example, the tensor p-shrinkage nuclear norm [38] replaces the nuclear norm by the Schatten-p quasi-norm; the tensor weighted Schatten-p norm [39] uses the weighted Schatten-p norm, and the log-based tensor nuclear norm [40] uses the log-det function.

2.2 Tensor logarithmic Schatten-p low-rank measure and its properties

Before we propose the tensor log-S_p low-rank measure, we first define a new matrix log-S_p low-rank measure as follows.

Definition 2.8 (log-S_p)**.** *Given $\boldsymbol{X} \in \mathbb{R}^{m\times n}$, the matrix logarithmic Schatten-p (log-S_p) low-rank measure of $\boldsymbol{X}$ is defined as*

$$\mathcal{M}_{\log,S_p}(\boldsymbol{X}) = \sum_{j=1}^{\min\{m,n\}} \log\left(\sigma_j^p(\boldsymbol{X}) + \varepsilon\right), \tag{3}$$

where $\varepsilon > 0$, $0 < p \leq 1$, and $\sigma_j(\boldsymbol{X})$ represents the j-th largest singular value of $\boldsymbol{X}$.

If $p = 1$, then this log-S_p low-rank measure $\mathcal{M}_{\log,S_p}(\cdot)$ reduces to the log-det function [41], which is a non-convex surrogate of the matrix rank. If $0 < p < 1$, due to the non-convexity of the ℓ_p norm, $\mathcal{M}_{\log,S_p}(\cdot)$ is also a non-convex relaxation of the matrix rank, and in fact it can achieve a better approximation than the S_p quasi-norm [42] or log-det function.

Next, we propose a new tensor low-rank measure called the t-log-S_p low-rank measure, which adopts the matrix log-S_p low-rank measure to characterize the low-rankness of the frontal slices $\overline{\boldsymbol{X}}^{(i)}$. The t-log-$S_p$ low-rank measure also denoted as $\mathcal{M}_{\log,S_p}(\cdot)$ is defined as follows.

Definition 2.9 (t-log-S_p)**.** *Given $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{n_1\times n_2\times n_3}$, the tensor logarithm Schatten-p (t-log-S_p) low-rank measure of $\boldsymbol{\mathcal{X}}$ is defined as*

$$\mathcal{M}_{\log,S_p}(\boldsymbol{\mathcal{X}}) = \frac{1}{n_3}\sum_{i=1}^{n_3} \mathcal{M}_{\log,S_p}\left(\overline{\boldsymbol{X}}^{(i)}\right) = \frac{1}{n_3}\sum_{i=1}^{n_3}\sum_{j=1}^{\min\{n_1,n_2\}} \log\left(\sigma_j^p\left(\overline{\boldsymbol{X}}^{(i)}\right) + \varepsilon\right), \tag{4}$$

where $\varepsilon > 0$ and $0 < p \leq 1$. When $n_3 = 1$, the t-log-S_p low-rank measure reduces to the matrix log-S_p low-rank measure.

Proposition 2.10 (Orthogonal invariance)**.** *The following assertions hold:*

(i) For a given matrix $\boldsymbol{X} \in \mathbb{R}^{m\times n}$, if $\boldsymbol{U} \in \mathbb{R}^{m\times m}$ and $\boldsymbol{V} \in \mathbb{R}^{n\times n}$ are orthogonal matrices, then

$$\mathcal{M}_{\log,S_p}(\boldsymbol{X}) = \mathcal{M}_{\log,S_p}(\boldsymbol{U}\boldsymbol{X}\boldsymbol{V}^T).$$

(ii) For a given tensor $\boldsymbol{\mathcal{X}} \in \mathbb{R}^{n_1\times n_2\times n_3}$, if $\boldsymbol{\mathcal{U}} \in \mathbb{R}^{n_1\times n_1\times n_3}$ and $\boldsymbol{V} \in \mathbb{R}^{n_2\times n_2\times n_3}$ are orthogonal tensors, then

$$\mathcal{M}_{\log,S_p}(\boldsymbol{\mathcal{X}}) = \mathcal{M}_{\log,S_p}(\boldsymbol{\mathcal{U}} * \boldsymbol{\mathcal{X}} * \boldsymbol{\mathcal{V}}^T).$$

Proof. (i) It immediately follows from $\sigma_j(\boldsymbol{X}) = \sigma_j(\boldsymbol{U}\boldsymbol{X}\boldsymbol{V}^T)$, $j = 1, 2, \ldots, \min\{m, n\}$.

(ii) Let $\boldsymbol{\mathcal{Z}} = \boldsymbol{\mathcal{U}} * \boldsymbol{\mathcal{X}} * \boldsymbol{\mathcal{V}}^T$. By equation (1), we have $\overline{\boldsymbol{Z}}^{(i)} = \overline{\boldsymbol{U}}^{(i)}\overline{\boldsymbol{X}}^{(i)}(\overline{\boldsymbol{V}}^{(i)})^T$, where $\overline{\boldsymbol{U}}^{(i)}$ and $\overline{\boldsymbol{V}}^{(i)}$ are orthogonal matrices. According to (i), we have (ii) holds. □

As shown in Proposition 2.1, both the matrix and tensor log-S_p low-rank measures $\mathcal{M}_{\log,S_p}(\cdot)$ satisfy the orthogonal invariance property. This property is useful when we minimize these measures together with another function that also has an orthogonal invariance property. For example, the log-S_p minimization problem that will be discussed in the next section may be reduced to a minimization problem only in terms of the singular values using this orthogonal invariance property.

3 Tensor Iteratively Reweighted S_p Minimization Algorithm for the t-log-S_p Minimization

For a 3D tensor $\boldsymbol{\mathcal{Y}} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, the t-log-$S_p$ minimization problem is written as

$$\min_{\boldsymbol{\mathcal{X}}} \frac{1}{2}\|\boldsymbol{\mathcal{X}} - \boldsymbol{\mathcal{Y}}\|_F^2 + \tau \mathcal{M}_{\log,S_p}(\boldsymbol{\mathcal{X}}), \tag{5}$$

where $\tau > 0$. By the definition of the t-log-S_p low-rank measure, $\mathcal{M}_{\log,S_p}(\cdot)$ is separable in terms of the frontal slices of $\overline{\boldsymbol{\mathcal{X}}}$. Then solving the t-log-S_p minimization problem (5) is equivalent to solving for each frontal slice $\overline{\boldsymbol{X}}^{(i)}$ via the following problem

$$\min_{\overline{\boldsymbol{X}}^{(i)}} \frac{1}{2}\|\overline{\boldsymbol{X}}^{(i)} - \overline{\boldsymbol{Y}}^{(i)}\|_F^2 + \tau \mathcal{M}_{\log,S_p}(\overline{\boldsymbol{X}}^{(i)}). \tag{6}$$

In subsection 3.1 we will propose an IRSpM algorithm for the log-S_p minimization problem as in (6), and conduct in subsection 3.2 a convergence analysis for IRSpM algorithm. Then in subsection 3.3, we summarize the tensor IRSpM (t-IRSpM) algorithm and its convergence analysis for solving the t-log-S_p minimization problem (5).

3.1 Iteratively reweighted S_p minimization algorithm for the log-S_p minimization

We consider the log-S_p minimization problem as follows

$$\min_{\boldsymbol{X} \in \mathbb{R}^{m \times n}} \frac{1}{2}\|\boldsymbol{X} - \boldsymbol{Y}\|_F^2 + \tau \mathcal{M}_{\log,S_p}(\boldsymbol{X}), \tag{7}$$

where $\boldsymbol{Y} \in \mathbb{R}^{m \times n}$ is the given data, $\boldsymbol{X} \in \mathbb{R}^{m \times n}$ is the unknown to be computed, and $\tau > 0$. Note that $\boldsymbol{X}$ and $\boldsymbol{Y}$ can represent $\overline{\boldsymbol{X}}^{(i)}$ and $\overline{\boldsymbol{Y}}^{(i)}$, respectively. By definition, the log-S_p low-rank measure can be written as

$$\mathcal{M}_{\log,S_p}(\boldsymbol{X}) = \sum_{j=1}^{\min\{m,n\}} g\left(\sigma_j^p(\boldsymbol{X})\right),$$

where $g : [0, \infty) \to \mathbb{R}$ is defined by $g(t) = \log(t + \varepsilon)$. The function g is monotonically increasing, concave, and continuously differentiable. Also, g has a Lipschitz continuous gradient with constant $L_g > 0$, i.e.,

$$\left|g'(s) - g'(t)\right| \leq L_g|s - t|, \quad \forall s, t \in [0, \infty).$$

To solve the log-S_p minimization as in (7), we propose an *iteratively reweighted S_p minimization (IRSpM) algorithm* as follows

$$\boldsymbol{X}^{k+1} = \operatorname*{argmin}_{\boldsymbol{X} \in \mathbb{R}^{m \times n}} \frac{\mu}{2} \left\|\boldsymbol{X} - \left[\boldsymbol{X}^k - \frac{1}{\mu}\left(\boldsymbol{X}^k - \boldsymbol{Y}\right)\right]\right\|_F^2 + \tau \sum_{j=1}^{l} \omega_j^k \sigma_j^p(\boldsymbol{X}), \tag{8}$$

where $w_j^k = g'\left(\sigma_j^p(\boldsymbol{X}^k)\right) = \frac{1}{\sigma_j^p(\boldsymbol{X}^k)+\varepsilon}$, $l = \min\{m, n\}$ and $\mu > 1$.

Before solving equation (8), we recall some notations on the singular value decomposition (SVD). Given a vector $\boldsymbol{x} \in \mathbb{R}^l$, let $\mathrm{Diag}(\boldsymbol{x})$ denote the $l \times l$ diagonal matrix with the j-th diagonal element as x_j. Given a matrix $\boldsymbol{X} \in \mathbb{R}^{m \times n}$, the SVD of $\boldsymbol{X}$ is computed as $\boldsymbol{X} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T$, where $\boldsymbol{U} \in \mathbb{R}^{m \times l}$ and $\boldsymbol{V} \in \mathbb{R}^{n \times l}$ are orthogonal matrices with $\boldsymbol{U}^T\boldsymbol{U} = \boldsymbol{V}^T\boldsymbol{V} = \boldsymbol{I}$, and $\boldsymbol{\Sigma} \in \mathbb{R}^{l \times l}$ is a diagonal matrix, $l = \min\{m, n\}$. In particular, $\boldsymbol{\Sigma} = \mathrm{Diag}(\boldsymbol{\sigma}(\boldsymbol{X}))$, where $\boldsymbol{\sigma}(\boldsymbol{X}) := [\sigma_1(\boldsymbol{X}), \sigma_2(\boldsymbol{X}), \cdots, \sigma_l(\boldsymbol{X})]^T$ and $\sigma_j(\boldsymbol{X})$ is the j-th largest singular value of $\boldsymbol{X}$.

Since equation (8) can be viewed as a weighted S_p minimization problem, we recall some preliminary results in [34].

Lemma 3.1 [34]. *For the following optimization problem:*

$$\min_{\delta \geq 0} f(\delta) = \frac{1}{2}(\delta - \sigma)^2 + w\delta^p \tag{9}$$

with $w \geq 0$ and $0 < p \leq 1$, there exists a specific threshold:

$$\tau_p^{GST}(w) = (2w(1-p))^{\frac{1}{2-p}} + wp(2w(1-p))^{\frac{p-1}{2-p}},$$

and we have the following conclusions.

(i) When $|\sigma| \leq \tau_p^{GST}(w)$, f has an optimal solution $T_p^{GST}(\sigma, w) = 0$;

(ii) When $|\sigma| > \tau_p^{GST}(w)$, f has one unique optimal solution $T_p^{GST}(\sigma, w) = \mathrm{sign}(\sigma)S_p^{GST}(|\sigma|, w)$ and $S_p^{GST}(|\sigma|, w)$ can be obtain by solving

$$S_p^{GST}(|\sigma|, w) - |\sigma| + wp\left(S_p^{GST}(|\sigma|, w)\right)^{p-1} = 0. \tag{10}$$

The generalized soft-thresholding (GST) algorithm proposed in [43] for finding an optimal solution $T_p^{GST}(\sigma, w)$ of problem (9) is summarized in Algorithm 1.

Algorithm 1 Generalized soft-thresholding (GST) [43]

Input: σ, w, p, J
$\tau_p^{GST}(w) = (2w(1-p))^{\frac{1}{2-p}} + wp(2w(1-p))^{\frac{p-1}{2-p}}$;
if $|\sigma| \leq \tau_p^{GST}(w)$ **then**
  $T_p^{GST}(\sigma, w) = 0$;
else
  $k = 0$, $\delta^k = |\sigma|$;
  for $k = 0, 1, \ldots, J$ **do**
    $\delta^{k+1} = |\sigma| - wp(\delta^k)^{p-1}$;
    $k \leftarrow k + 1$;
  end for
  $T_p^{GST}(\sigma, w) = \mathrm{sign}(\sigma)\delta^k$.
end if
Output: $T_p^{GST}(\sigma, w)$

Theorem 3.2 [34]. *Let $\boldsymbol{Y} \in \mathbb{R}^{m \times n}$ and $\tau > 0$. And let $\boldsymbol{w} = [w_1, \ldots, w_l]^T \in \mathbb{R}^l$ such that $0 \leq w_1 \leq w_2 \leq \cdots \leq w_l$, $l = \min\{m, n\}$. Then a global optimal solution for the following problem*

$$\min_{\boldsymbol{X} \in \mathbb{R}^{m \times n}} \frac{1}{2}\|\boldsymbol{X} - \boldsymbol{Y}\|_F^2 + \tau \sum_{j=1}^{l} w_j \sigma_j^p(\boldsymbol{X})$$

is given by

$$\Gamma_{\tau \boldsymbol{w}}(\boldsymbol{Y}) = \boldsymbol{U}\,\mathrm{Diag}(\boldsymbol{\gamma})\boldsymbol{V}^T,$$

where $\boldsymbol{Y} = \boldsymbol{U\Sigma V}^T$ is the SVD of $\boldsymbol{Y}$, $\boldsymbol{\Sigma} = \mathrm{Diag}(\boldsymbol{\sigma}(\boldsymbol{Y}))$, and $\boldsymbol{\gamma} = [\gamma_1, \gamma_2, \ldots, \gamma_l]^T \in \mathbb{R}^l$ satisfies $\gamma_j = T_p^{GST}(\sigma_j(\boldsymbol{Y}), \tau w_j)$, $i = 1, 2, \ldots, l$. In particular, $\boldsymbol{\gamma}$ also satisfies $\gamma_1 \geq \gamma_2 \geq \cdots \geq \gamma_l \geq 0$.

For the IRSpM algorithm given in equation (8), it can be easily verified that $\boldsymbol{w}^k = [w_1^k, w_2^k, \dots, w_l^k]^T \in \mathbb{R}^l$ satisfies $0 \le w_1^k \le w_2^k \le \dots \le w_l^k$. Then a global optimal solution of (8) can be efficiently solved according to Theorem 3.2 as follows

$$\boldsymbol{X}^{k+1} = \Gamma_{\frac{\tau}{\mu}\boldsymbol{w}^k}\left(\boldsymbol{X}^k - \frac{1}{\mu}(\boldsymbol{X}^k - \boldsymbol{Y})\right). \tag{11}$$

We summarize the IRSpM algorithm in Algorithm 2.

Algorithm 2 IRSpM algorithm for solving the log-S_p minimization problem (7)

Input: $\boldsymbol{Y}$ and parameter τ
1: **Initialize** $\boldsymbol{X}^0$
2: **Set** $k=0$, $\mu > 1$ and $\boldsymbol{w}_j^0 = \frac{1}{\sigma_j(\boldsymbol{X}^0)^p + \varepsilon}$
3: **while** stopping criterion is not satisfied **do**
4: **Compute the SVD of** $\boldsymbol{X}^k - \frac{1}{\mu}(\boldsymbol{X}^k - \boldsymbol{Y})$, i.e., $\boldsymbol{X}^k - \frac{1}{\mu}(\boldsymbol{X}^k - \boldsymbol{Y}) = \boldsymbol{U}^{k+1}\boldsymbol{\Sigma}^{k+1}(\boldsymbol{V}^{k+1})^T$
5: **for** $j = 1, 2, \dots, l$ **do**
6: $\gamma_j^{k+1} = T_p^{GST}\left(\boldsymbol{\Sigma}_{jj}^{k+1}, \frac{\tau}{\mu} w_j^k\right)$
7: $w_j^{k+1} = \frac{1}{(\gamma_j^{k+1})^p + \varepsilon}$
8: **end for**
9: $\boldsymbol{X}^{k+1} = \boldsymbol{U}^{k+1}\,\mathrm{Diag}(\boldsymbol{\gamma}^{k+1})(\boldsymbol{V}^{k+1})^T$
10: $k \leftarrow k+1$.
11: **end while**
Output: $\boldsymbol{X}^k$

Remark 3.3. *If we initialize $\boldsymbol{X}^0$ by $\boldsymbol{X}^0 = \boldsymbol{Y}$, the IRSpM algorithm can be simplified and only requires one SVD operation. Suppose $\boldsymbol{Y} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T$ is the SVD of $\boldsymbol{Y}$, where $\boldsymbol{\Sigma} = \mathrm{Diag}(\boldsymbol{\sigma}(\boldsymbol{Y}))$. Then the sequence $\{\boldsymbol{X}^k\}$ generated by the IRSpM algorithm in (11) can be computed by*

$$\boldsymbol{X}^{k+1} = \boldsymbol{U}\,\mathrm{Diag}(\boldsymbol{\gamma}^{k+1})\boldsymbol{V}^T,$$

where $\boldsymbol{\gamma}^{k+1} = [\gamma_1^{k+1}, \gamma_2^{k+1}, \dots, \gamma_l^{k+1}]^T \in \mathbb{R}^l$, $l = \min\{m, n\}$, satisfies that $\boldsymbol{\gamma}^0 = \boldsymbol{\sigma}(\boldsymbol{Y})$ and for each j

$$\gamma_j^{k+1} = T_p^{GST}\left(\gamma_j^k - \frac{1}{\mu}(\gamma_j^k - \sigma_j(\boldsymbol{Y})), \frac{\tau}{\mu} w_j^k\right), \quad k = 0, 1, \dots.$$

3.2 Convergence analysis of the IRSpM algorithm

We can prove that any accumulation point of the sequence $\{\boldsymbol{X}^k\}$ generated by Algorithm 2 is a stationary point of the objective function of the log-S_p minimization as in (7).

First, we recall some definitions of subdifferentials and some results on computing the subdifferential of singular value functions introduced in [44].

Definition 3.4 Subdifferentials**.** *Let $f : \mathbb{R}^d \to (-\infty, +\infty]$ be a proper and lower semicontinuous function.*

(1) For a given $\boldsymbol{x} \in \mathrm{dom}\,\partial f := \{\boldsymbol{x} \in \mathbb{R}^d : \partial f(\boldsymbol{x}) \neq \emptyset\}$, the Fréchet subdifferential of f at $\boldsymbol{x}$, written $\hat{\partial} f(\boldsymbol{x})$, is the set of all vectors $\boldsymbol{u} \in \mathbb{R}^d$ which satisfy

$$\liminf_{\substack{\boldsymbol{y} \neq \boldsymbol{x} \\ \boldsymbol{y} \to \boldsymbol{x}}} \frac{f(\boldsymbol{y}) - f(\boldsymbol{x}) - \langle \boldsymbol{u}, \boldsymbol{y} - \boldsymbol{x} \rangle}{\|\boldsymbol{y} - \boldsymbol{x}\|} \ge 0.$$

When $\boldsymbol{x} \notin \mathrm{dom}\, f$, we set $\hat{\partial} f(\boldsymbol{x}) = \emptyset$.

(2) The subdifferential of f at $\boldsymbol{x} \in \mathbb{R}^d$, written $\partial f(\boldsymbol{x})$, is defined through the following closure process

$$\partial f(\boldsymbol{x}) := \{\boldsymbol{u} \in \mathbb{R}^d : \exists \boldsymbol{x}_k \to \boldsymbol{x}, f(\boldsymbol{x}_k) \to f(\boldsymbol{x}) \text{ and } \boldsymbol{u}_k \in \hat{\partial} f(\boldsymbol{x}_k) \to \boldsymbol{u} \text{ as } k \to \infty\}.$$

Definition 3.5. *A function $f : \mathbb{R}^n \to \mathbb{R}$ is absolutely symmetric if*

$$f(x_1, x_2, \dots, x_n) = f(|x_{\pi(1)}|, |x_{\pi(2)}|, \dots, |x_{\pi(n)}|),$$

for any permutation π.

Definition 3.6. *A function $F : \mathbb{R}^{m\times n} \to \mathbb{R}$ is a singular value function if $F(\boldsymbol{X}) = (f \circ \boldsymbol{\sigma})(\boldsymbol{X})$, where $f : \mathbb{R}^l \to \mathbb{R}$ is an absolutely symmetric function, $l = \min\{m, n\}$.*

Lemma 3.7 [44]**.** *Let f be an absolutely symmetric function, then the subdifferential of the corresponding singular value function $f \circ \boldsymbol{\sigma}$ at a matrix $\boldsymbol{X}$ is given by the formula*

$$\partial(f \circ \boldsymbol{\sigma})(\boldsymbol{X}) = \boldsymbol{U}\,\mathrm{Diag}\left(\partial f\left[\boldsymbol{\sigma}(\boldsymbol{X})\right]\right)\boldsymbol{V}^T$$

with $\boldsymbol{X} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^T$ being the SVD of $\boldsymbol{X}$.

The log-S_p low-rank measure can be viewed as a singular value function and its subdifferential can be computed by Lemma 3.7. However, it is still challenging to find an explicit expression for the subdifferential of the log-S_p low-rank measure due to the non-smoothness of the S_p quasi-norm.

Second, motivated by the class of first-order stationary points for ℓ_p regularized low-rank approximation problems introduced in [42], we define a class of first-order stationary points for the log-S_p minimization problem (7) using

$$\widetilde{\mathcal{O}}(\boldsymbol{X}) := \left\{(\tilde{\boldsymbol{U}}, \tilde{\boldsymbol{V}}) \in \mathbb{R}^{m\times r} \times \mathbb{R}^{n\times r} : \tilde{\boldsymbol{U}}^T\tilde{\boldsymbol{U}} = \tilde{\boldsymbol{V}}^T\tilde{\boldsymbol{V}} = \boldsymbol{I} \text{ and } \boldsymbol{X} = \tilde{\boldsymbol{U}}\,\mathrm{Diag}(\tilde{\boldsymbol{\sigma}}(\boldsymbol{X}))\tilde{\boldsymbol{V}}^T\right\},$$

where $\tilde{\boldsymbol{\sigma}}(\boldsymbol{X}) := [\sigma_1(\boldsymbol{X}), \sigma_2(\boldsymbol{X}), \dots, \sigma_r(\boldsymbol{X})]^T$ and $r = \mathrm{rank}(\boldsymbol{X})$. Note that $\widetilde{\mathcal{O}}(\boldsymbol{X})$ is the set of all such pairs $(\tilde{\boldsymbol{U}}, \tilde{\boldsymbol{V}})$ of the rank reduced SVD of $\boldsymbol{X}$.

Definition 3.8. *A point $\boldsymbol{X}^*$ is a first-order stationary point of problem* (7) *if*

$$\boldsymbol{0} \in \left\{\tilde{\boldsymbol{U}}^T\left(\boldsymbol{X}^* - \boldsymbol{Y}\right)\tilde{\boldsymbol{V}} + \tau p\,\mathrm{Diag}(\boldsymbol{d}) : (\tilde{\boldsymbol{U}}, \tilde{\boldsymbol{V}}) \in \widetilde{\mathcal{O}}\left(\boldsymbol{X}^*\right) \text{ and } d_j = \sigma_j^{p-1}(\boldsymbol{X}^*)(\sigma_j^p(\boldsymbol{X}^*) + \varepsilon)^{-1}\right\}. \tag{12}$$

The next theorem shows that a local minimizer of problem (7) is a first-order stationary point.

Theorem 3.9. *Suppose that $\boldsymbol{X}^*$ is a local minimizer of problem* (7)*. Then $\boldsymbol{X}^*$ is a first-order stationary point of problem* (7)*, that is,* (12) *holds at $\boldsymbol{X}^*$.*

Proof. Let $\boldsymbol{X}^* = \boldsymbol{U}\,\mathrm{Diag}(\tilde{\boldsymbol{\sigma}}(\boldsymbol{X}^*))\boldsymbol{V}^T$ for some $(\boldsymbol{U}, \boldsymbol{V}) \in \widetilde{\mathcal{O}}(\boldsymbol{X}^*)$ and $r = \mathrm{rank}(\boldsymbol{X}^*)$. Define $\varphi : \mathbb{R}^{r\times r} \to \mathbb{R}$ as

$$\begin{aligned}\varphi(\boldsymbol{Z}) &= \frac{1}{2}\left\|\boldsymbol{X}^* + \boldsymbol{U}\boldsymbol{Z}\boldsymbol{V}^T - \boldsymbol{Y}\right\|_F^2 + \tau\mathcal{M}_{\log,S_p}(\boldsymbol{X}^* + \boldsymbol{U}\boldsymbol{Z}\boldsymbol{V}^T)\\ &= \frac{1}{2}\left\|\boldsymbol{X}^* + \boldsymbol{U}\boldsymbol{Z}\boldsymbol{V}^T - \boldsymbol{Y}\right\|_F^2 + \tau\mathcal{M}_{\log,S_p}(\mathrm{Diag}(\tilde{\boldsymbol{\sigma}}(\boldsymbol{X}^*)) + \boldsymbol{Z}).\end{aligned}$$

By Theorem 7.1 in [44] and the definition of $\widetilde{\mathcal{O}}(\cdot)$, the subdifferential of $\varphi(\cdot)$ at $\boldsymbol{Z} = \boldsymbol{0}$ is given by

$$\begin{aligned}\partial\varphi(\boldsymbol{0}) = \big\{&\boldsymbol{U}^T(\boldsymbol{X}^* - \boldsymbol{Y})\boldsymbol{V} + \tau p\hat{\boldsymbol{U}}\,\mathrm{Diag}(\boldsymbol{d})\hat{\boldsymbol{V}}^T :\\ &(\hat{\boldsymbol{U}}, \hat{\boldsymbol{V}}) \in \widetilde{\mathcal{O}}\left(\mathrm{Diag}(\tilde{\boldsymbol{\sigma}}(\boldsymbol{X}^*))\right) \text{ and } d_j = \sigma_j^{p-1}(\boldsymbol{X}^*)(\sigma_j^p(\boldsymbol{X}^*) + \varepsilon)^{-1}\big\}.\end{aligned}$$

Since $\mathbf{0}$ is a local minimizer of $\varphi(\cdot)$, we have $\mathbf{0} \in \partial\varphi(\mathbf{0})$. Hence, there exists some $(\hat{\boldsymbol{U}}, \hat{\boldsymbol{V}}) \in \widetilde{\mathcal{O}}\,(\mathrm{Diag}(\tilde{\boldsymbol{\sigma}}(\boldsymbol{X}^*)))$ such that

$$\boldsymbol{U}^T(\boldsymbol{X}^* - \boldsymbol{Y})\boldsymbol{V} + \tau p \hat{\boldsymbol{U}}\,\mathrm{Diag}(\boldsymbol{d})\hat{\boldsymbol{V}}^T = \mathbf{0},$$

where $d_j = \sigma_j^{p-1}(\boldsymbol{X}^*)(\sigma_j^p(\boldsymbol{X}^*) + \varepsilon)^{-1}$, $j = 1, 2, \ldots, r$. Upon pre- and post-multiplying the above equation by $\hat{\boldsymbol{U}}^T$ and $\hat{\boldsymbol{V}}$, and using $\hat{\boldsymbol{U}}^T\hat{\boldsymbol{U}} = \hat{\boldsymbol{V}}^T\hat{\boldsymbol{V}} = \boldsymbol{I}$, we obtain

$$\tilde{\boldsymbol{U}}^T\,(\boldsymbol{X}^* - \boldsymbol{Y})\,\tilde{\boldsymbol{V}} + \tau p\,\mathrm{Diag}(\boldsymbol{d}) = \mathbf{0},$$

where $\tilde{\boldsymbol{U}} = \boldsymbol{U}\hat{\boldsymbol{U}}$ and $\tilde{\boldsymbol{V}} = \boldsymbol{V}\hat{\boldsymbol{V}}$. Since $(\boldsymbol{U}, \boldsymbol{V}) \in \widetilde{\mathcal{O}}(\boldsymbol{X}^*)$ and $(\hat{\boldsymbol{U}}, \hat{\boldsymbol{V}}) \in \widetilde{\mathcal{O}}\,(\mathrm{Diag}(\tilde{\boldsymbol{\sigma}}(\boldsymbol{X}^*)))$, then we have

$$\tilde{\boldsymbol{U}}\,\mathrm{Diag}(\tilde{\boldsymbol{\sigma}}(\boldsymbol{X}^*))\tilde{\boldsymbol{V}}^T = \boldsymbol{U}\left(\hat{\boldsymbol{U}}\,\mathrm{Diag}(\tilde{\boldsymbol{\sigma}}(\boldsymbol{X}^*))\hat{\boldsymbol{V}}^T\right)\boldsymbol{V}^T = \boldsymbol{U}\,\mathrm{Diag}(\tilde{\boldsymbol{\sigma}}(\boldsymbol{X}^*))\boldsymbol{V}^T = \boldsymbol{X}^*.$$

Hence, $(\tilde{\boldsymbol{U}}, \tilde{\boldsymbol{V}}) \in \widetilde{\mathcal{O}}(\boldsymbol{X}^*)$ and (12) holds. □

Third, we show some convergence results on the sequence $\{\boldsymbol{X}^k\}$ generated by the proposed IRSpM algorithm in Algorithm 2. The objective function of (7) evaluated at the sequence $\{\boldsymbol{X}^k\}$ is strictly decreasing and any accumulation point of $\{\boldsymbol{X}^k\}$ is a stationary point.

Proposition 3.10. *Let Ψ denote the objective function of the log-S_p minimization problem (7). Suppose that $\{\boldsymbol{X}^k\}$ is a sequence generated by Algorithm 2 and $\mu > 1$. Then we have*

$$\Psi(\boldsymbol{X}^k) - \Psi(\boldsymbol{X}^{k+1}) \geq \frac{\mu - 1}{2}\|\boldsymbol{X}^{k+1} - \boldsymbol{X}^k\|_F^2. \tag{13}$$

Proof. By the descent lemma [45] and the concavity of the function g, we obtain

$$\begin{aligned}
\Psi(\boldsymbol{X}^k) - \Psi(\boldsymbol{X}^{k+1}) &= \frac{1}{2}\|\boldsymbol{X}^k - \boldsymbol{Y}\|_F^2 - \frac{1}{2}\|\boldsymbol{X}^{k+1} - \boldsymbol{Y}\|_F^2 + \tau\sum_{j=1}^{l}\left[g\left(\sigma_j^p(\boldsymbol{X}^k)\right) - g\left(\sigma_j^p(\boldsymbol{X}^{k+1})\right)\right] \\
&= \langle \boldsymbol{X}^k - \boldsymbol{Y}, \boldsymbol{X}^k - \boldsymbol{X}^{k+1}\rangle - \frac{1}{2}\|\boldsymbol{X}^{k+1} - \boldsymbol{X}^k\|_F^2 + \tau\sum_{j=1}^{l}\left[g\left(\sigma_j^p(\boldsymbol{X}^k)\right) - g\left(\sigma_j^p(\boldsymbol{X}^{k+1})\right)\right] \\
&\geq \langle \boldsymbol{X}^k - \boldsymbol{Y}, \boldsymbol{X}^k - \boldsymbol{X}^{k+1}\rangle - \frac{1}{2}\|\boldsymbol{X}^{k+1} - \boldsymbol{X}^k\|_F^2 + \tau\sum_{j=1}^{l} w_j^k\left(\sigma_j^p(\boldsymbol{X}^k) - \sigma_j^p(\boldsymbol{X}^{k+1})\right),
\end{aligned} \tag{14}$$

where $w_j^k = g'\left(\sigma_j^p(\boldsymbol{X}^k)\right)$ and $l = \min\{m, n\}$. Note that X^{k+1} is a minimizer of (8), and thus we have

$$\begin{aligned}
&\langle \boldsymbol{X}^k - \boldsymbol{Y}, \boldsymbol{X}^{k+1} - \boldsymbol{X}^k\rangle + \frac{\mu}{2}\|\boldsymbol{X}^{k+1} - \boldsymbol{X}^k\|_F^2 + \tau\sum_{j=1}^{l} w_j^k\sigma_j^p(\boldsymbol{X}^{k+1}) \\
&\leq \langle \boldsymbol{X}^k - \boldsymbol{Y}, \boldsymbol{X}^k - \boldsymbol{X}^k\rangle + \frac{\mu}{2}\|\boldsymbol{X}^k - \boldsymbol{X}^k\|_F^2 + \tau\sum_{j=1}^{l} w_j^k\sigma_j^p(\boldsymbol{X}^k) \\
&= \tau\sum_{j=1}^{l} w_j^k\sigma_j^p(\boldsymbol{X}^k).
\end{aligned}$$

That is,

$$\langle \boldsymbol{X}^k - \boldsymbol{Y}, \boldsymbol{X}^k - \boldsymbol{X}^{k+1}\rangle + \tau\sum_{j} w_j^k\left(\sigma_j^p(\boldsymbol{X}^k) - \sigma_j^p(\boldsymbol{X}^{k+1})\right) \geq \frac{\mu}{2}\left\|\boldsymbol{X}^{k+1} - \boldsymbol{X}^k\right\|_F^2. \tag{15}$$

Then substituting (15) into (14) yields (13). □

Theorem 3.11. *Let Ψ denote the objective function of the log-S_p minimization problem (7). Suppose that $\{\boldsymbol{X}^k\}$ is a sequence generated by Algorithm 2 and $\mu > 1$. Then the following assertions hold:*

(i) The sequence $\{\boldsymbol{X}^k\}$ is bounded.

(ii) $\lim_{k\to\infty} \|\boldsymbol{X}^{k+1} - \boldsymbol{X}^k\|_F = 0$.

(iii) Any accumulation point of $\{\boldsymbol{X}^k\}$ is a stationary point of Ψ.

Proof. (i) It follows from Proposition 3.10 that the decreasing sequence $\{\Psi(\boldsymbol{X}^k)\}$ is bounded above by $\Psi(\boldsymbol{X}^0)$. Also, $\Psi_{\inf} = \inf_{\boldsymbol{X}} \Psi(\boldsymbol{X}) > -\infty$. Then assertion (i) holds since Ψ is coercive.

(ii) Summing (13) from $k = 0$ to $k = K$, we have

$$\sum_{k=0}^{K} \|\boldsymbol{X}^{k+1} - \boldsymbol{X}^k\|_F^2 \le \frac{2}{\mu - 1}\left(\Psi(\boldsymbol{X}^0) - \Psi(\boldsymbol{X}^{K+1})\right) \le \frac{2}{\mu - 1}\left(\Psi(\boldsymbol{X}^0) - \Psi_{\inf}\right) < +\infty.$$

Taking $K \to \infty$, we have

$$\sum_{k=0}^{\infty} \|\boldsymbol{X}^{k+1} - \boldsymbol{X}^k\|_F^2 < +\infty.$$

This yields assertion (ii).

(iii) Let $\boldsymbol{X}^*$ be an accumulation point of the sequence $\{\boldsymbol{X}^k\}$ and let $\boldsymbol{\gamma}^* \in \mathbb{R}^l$ be a vector such that $\boldsymbol{\gamma}^* = \boldsymbol{\sigma}(\boldsymbol{X}^*)$. Assume that a subsequence $\{\boldsymbol{X}^{k_i}\}$ of $\{\boldsymbol{X}^k\}$ converges to $\boldsymbol{X}^*$ as $i \to \infty$. Due to assertion (ii), we also have $\boldsymbol{X}^{k_i+1} \to \boldsymbol{X}^*$ as $i \to \infty$. Then $\boldsymbol{\sigma}(\boldsymbol{X}^{k_i}) \to \boldsymbol{\sigma}(\boldsymbol{X}^*)$ and $\boldsymbol{\sigma}(\boldsymbol{X}^{k_i+1}) \to \boldsymbol{\sigma}(\boldsymbol{X}^*)$, i.e., $\boldsymbol{\gamma}^{k_i} \to \boldsymbol{\gamma}^*$ and $\boldsymbol{\gamma}^{k_i+1} \to \boldsymbol{\gamma}^*$, as $i \to \infty$.

Let $r = \operatorname{rank}(\boldsymbol{X}^*)$. Then there exists some $I_0 > 0$ such that $\gamma_j^{k_i+1} > 0$ for all $j \le r$ and $i > I_0$. And $\gamma_j^{k_i+1} > 0$ implies that $\gamma_j^{k_i} - \frac{1}{\mu}(\gamma_j^{k_i} - \sigma_j(\boldsymbol{Y})) > \tau_p^{GST}(\frac{\tau}{\mu} w_j^{k_i})$ and

$$\gamma_j^{k_i+1} = T_p^{GST}\left(\boldsymbol{\Sigma}_{jj}^{k_i+1}, \frac{\tau}{\mu} w_j^{k_i}\right) = S_p^{GST}\left(\boldsymbol{\Sigma}_{jj}^{k_i+1}, \frac{\tau}{\mu} w_j^{k_i}\right).$$

By Lemma 3.1, the following equation holds for all $j \le r$ and $i > I_0$,

$$\gamma_j^{k_i+1} - \boldsymbol{\Sigma}_{jj}^{k_i+1} + \frac{\tau}{\mu} w_j^{k_i} p (\gamma_j^{k_i+1})^{p-1} = 0. \tag{16}$$

Denote $\tilde{\boldsymbol{\gamma}}^{k_i+1} := [\gamma_1^{k_i+1}, \gamma_2^{k_i+1}, \ldots, \gamma_r^{k_i+1}]^T$ and denote $\boldsymbol{d}^{k_i+1} := [d_1^{k_i+1}, d_2^{k_i+1}, \ldots, d_r^{k_i+1}]^T$ with $d_j^{k_i+1} = w_j^{k_i}(\gamma_j^{k_i+1})^{p-1} = ((\gamma_j^{k_i})^p + \varepsilon)^{-1}(\gamma_j^{k_i+1})^{p-1}$. Let $\tilde{\boldsymbol{\Sigma}}^{k_i+1}$ denote the $r \times r$ matrix formed by the first r rows and first r columns of $\boldsymbol{\Sigma}^{k_i+1}$, let $\tilde{\boldsymbol{U}}^{k_i+1}$ denote the $m \times r$ matrix formed by the first r columns of $\boldsymbol{U}^{k_i+1}$ and let $\tilde{\boldsymbol{V}}^{k_i+1}$ denote the $n \times r$ matrix formed by the first r columns of $\boldsymbol{V}^{k_i+1}$. Then from equation (16), we obtain

$$\tilde{\boldsymbol{U}}^{k_i+1} \operatorname{Diag}(\tilde{\boldsymbol{\gamma}}^{k_i+1})(\tilde{\boldsymbol{V}}^{k_i+1})^T - \tilde{\boldsymbol{U}}^{k_i+1}\tilde{\boldsymbol{\Sigma}}^{k_i+1}(\tilde{\boldsymbol{V}}^{k_i+1})^T + \frac{\tau p}{\mu}\tilde{\boldsymbol{U}}^{k_i+1} \operatorname{Diag}(\boldsymbol{d}^{k_i+1})(\tilde{\boldsymbol{V}}^{k_i+1})^T = \boldsymbol{0}.$$

We observe that $\tilde{\boldsymbol{U}}^{k_i+1} \operatorname{Diag}(\tilde{\boldsymbol{\gamma}}^{k_i+1})(\tilde{\boldsymbol{V}}^{k_i+1})^T = \boldsymbol{U}^{k_i+1} \operatorname{Diag}(\boldsymbol{\gamma}^{k_i+1})(\boldsymbol{V}^{k_i+1})^T - \sum_{j=r+1}^{l} \gamma_j^{k_i+1} \boldsymbol{U}_j^{k_i+1}(\boldsymbol{V}_j^{k_i+1})^T = \boldsymbol{X}^{k_i+1} - \sum_{j=r+1}^{l} \gamma_j^{k_i+1}\boldsymbol{U}_j^{k_i+1}(\boldsymbol{V}_j^{k_i+1})^T$, where $\boldsymbol{U}_j^{k_i+1}$ and $\boldsymbol{V}_j^{k_i+1}$ denote the j-th column of $\boldsymbol{U}^{k_i+1}$ and $\boldsymbol{V}^{k_i+1}$, respectively. Also, $\tilde{\boldsymbol{U}}^{k_i+1}\tilde{\boldsymbol{\Sigma}}^{k_i+1}(\tilde{\boldsymbol{V}}^{k_i+1})^T = \boldsymbol{U}^{k_i+1}\boldsymbol{\Sigma}^{k_i+1}(\boldsymbol{V}^{k_i+1})^T - \sum_{j=r+1}^{l} \boldsymbol{\Sigma}_{jj}^{k_i+1}\boldsymbol{U}_j^{k_i+1}(\boldsymbol{V}_j^{k_i+1})^T = \boldsymbol{X}^{k_i} - \frac{1}{\mu}(\boldsymbol{X}^{k_i} - \boldsymbol{Y}) - \sum_{j=r+1}^{l} \boldsymbol{\Sigma}_{jj}^{k_i+1}\boldsymbol{U}_j^{k_i+1}(\boldsymbol{V}_j^{k_i+1})^T$. These imply that

$$\mu(\boldsymbol{X}^{k_i+1} - \boldsymbol{X}^{k_i}) + (\boldsymbol{X}^{k_i} - \boldsymbol{Y}) + \mu \sum_{j=r+1}^{l} (\boldsymbol{\Sigma}_{jj}^{k_i+1} - \gamma_j^{k_i+1})\boldsymbol{U}_j^{k_i+1}(\boldsymbol{V}_j^{k_i+1})^T + \tau p \tilde{\boldsymbol{U}}^{k_i+1} \operatorname{Diag}(\boldsymbol{d}^{k_i+1})(\tilde{\boldsymbol{V}}^{k_i+1})^T = \boldsymbol{0}.$$

Upon pre- and post-multiplying the equation above by $(\tilde{\boldsymbol{U}}^{k_i+1})^T$ and $\tilde{\boldsymbol{V}}^{k_i+1}$ and using $(\tilde{\boldsymbol{U}}^{k_i+1})^T\tilde{\boldsymbol{U}}^{k_i+1} = \boldsymbol{I}$ and $(\tilde{\boldsymbol{V}}^{k_i+1})^T\tilde{\boldsymbol{V}}^{k_i+1} = \boldsymbol{I}$, we obtain for all $i > I_0$

$$\mu(\tilde{\boldsymbol{U}}^{k_i+1})^T(\boldsymbol{X}^{k_i+1} - \boldsymbol{X}^{k_i})\tilde{\boldsymbol{V}}^{k_i+1} + (\tilde{\boldsymbol{U}}^{k_i+1})^T(\boldsymbol{X}^{k_i} - \boldsymbol{Y})\tilde{\boldsymbol{V}}^{k_i+1} + \tau p\,\mathrm{Diag}(\boldsymbol{d}^{k_i+1}) = \boldsymbol{0}. \tag{17}$$

Next, it can be easily verified that $\{\tilde{\boldsymbol{U}}^{k_i+1}\}$ and $\{\tilde{\boldsymbol{V}}^{k_i+1}\}$ are bounded. Considering a convergent subsequence if necessary, without loss of generality we assume that $\tilde{\boldsymbol{U}}^{k_i+1} \to \tilde{\boldsymbol{U}}^*$ and $\tilde{\boldsymbol{V}}^{k_i+1} \to \tilde{\boldsymbol{V}}^*$. Then taking the limit of both sides of equation (17) as $i \to \infty$ and using assertion (ii), we have

$$(\tilde{\boldsymbol{U}}^*)^T(\boldsymbol{X}^* - \boldsymbol{Y})\tilde{\boldsymbol{V}}^* + \tau p\,\mathrm{Diag}(\boldsymbol{d}^*) = \boldsymbol{0},$$

where $\boldsymbol{d}^* = [d_1^*, d_2^*, \ldots, d_r^*]^T \in \mathbb{R}^r$ such that $d_j^* = (\gamma_j^*)^{p-1}((\gamma_j^*)^p + \varepsilon)^{-1} = \sigma_j^{p-1}(\boldsymbol{X}^*)(\sigma_j^p(\boldsymbol{X}^*) + \varepsilon)^{-1}$. Since $(\tilde{\boldsymbol{U}}^{k_i+1})^T\tilde{\boldsymbol{U}}^{k_i+1} = \boldsymbol{I}$ and $(\tilde{\boldsymbol{V}}^{k_i+1})^T\tilde{\boldsymbol{V}}^{k_i+1} = \boldsymbol{I}$, we have $(\tilde{\boldsymbol{U}}^*)^T\tilde{\boldsymbol{U}}^* = \boldsymbol{I}$ and $(\tilde{\boldsymbol{V}}^*)^T\tilde{\boldsymbol{V}}^* = \boldsymbol{I}$. Since $\gamma_j^* = 0$ for all $j > r$, we have $\boldsymbol{X}^* = \tilde{\boldsymbol{U}}^*\,\mathrm{Diag}(\tilde{\gamma}^*)(\tilde{\boldsymbol{V}}^*)^T$, that is, $(\tilde{\boldsymbol{U}}^*, \tilde{\boldsymbol{V}}^*) \in \widetilde{\mathcal{O}}(\boldsymbol{X}^*)$. Therefore, $\boldsymbol{X}^*$ is a stationary point of Ψ.

□

3.3 The tensor IRSpM algorithm and its convergence analysis for the t-log-S_p minimization

As mentioned in the beginning of section 3, solving the t-log-S_p minimization (5) is equivalent to solving a log-S_p minimization (6) for each frontal slice $\overline{\boldsymbol{X}}^{(i)}$. Then a tensor IRSpM (t-IRSpM) algorithm can be proposed for the t-log-S_p minimization using the IRSpM algorithm for the log-S_p minimization. We summarize the t-IRSpM algorithm in Algorithm 3 and its convergence results in Theorem 3.13.

Algorithm 3 The t-IRSpM Algorithm for solving the t-log-S_p minimization problem (5)

Input: $\boldsymbol{\mathcal{Y}}$ and parameter τ
1: **Set initialization** denoted as $\boldsymbol{\mathcal{Z}}$
2: $\overline{\boldsymbol{\mathcal{Y}}} = \mathrm{fft}(\boldsymbol{\mathcal{Y}}, [\,], 3)$
3: $\overline{\boldsymbol{\mathcal{Z}}} = \mathrm{fft}(\boldsymbol{\mathcal{Z}}, [\,], 3)$
4: **for** $i = 1, 2, \ldots, n_3$ **do**
5: $\quad\overline{\boldsymbol{X}}^{(i)} = \mathrm{IRSpM}(\overline{\boldsymbol{Y}}^{(i)}, \tau)$ initialized with $\overline{\boldsymbol{Z}}^{(i)}$.
6: **end for**
7: $\boldsymbol{\mathcal{X}} = \mathrm{ifft}(\overline{\boldsymbol{\mathcal{X}}}, [\,], 3)$
Output: $\boldsymbol{\mathcal{X}}$

Definition 3.12. *A point $\boldsymbol{\mathcal{X}}_*$ is a first-order stationary point of problem (5) if for $i = 1, 2, \ldots, n_3$,*

$$\boldsymbol{0} \in \{\tilde{\boldsymbol{U}}^T\left(\overline{\boldsymbol{X}_*}^{(i)} - \overline{\boldsymbol{Y}}^{(i)}\right)\tilde{\boldsymbol{V}} + \tau p\,\mathrm{Diag}(\boldsymbol{d}) :(\tilde{\boldsymbol{U}}, \tilde{\boldsymbol{V}}) \in \widetilde{\mathcal{O}}\left(\overline{\boldsymbol{X}_*}^{(i)}\right) \\ \text{and } d_j = \sigma_j^{p-1}(\overline{\boldsymbol{X}_*}^{(i)})(\sigma_j^p(\overline{\boldsymbol{X}_*}^{(i)}) + \varepsilon)^{-1}\}. \tag{18}$$

To present the convergence results of the t-IRSpM algorithm, we denote $\{\overline{\boldsymbol{X}}_k^{(i)}\}$ as the sequence generated by IRSpM in the fifth line in Algorithm 3 and denote $\boldsymbol{\mathcal{X}}_k = \mathrm{ifft}(\overline{\boldsymbol{\mathcal{X}}_k}, [\,], 3)$, where the i-frontal slice of $\overline{\boldsymbol{\mathcal{X}}_k}$ is $\overline{\boldsymbol{X}}_k^{(i)}$.

Theorem 3.13. *Let Φ denote the objective function of the log-S_p minimization problem (7). Suppose that $\{\boldsymbol{\mathcal{X}}_k\}$ is a sequence generated by Algorithm 3. Then the following assertions hold:*

(i) $\Phi(\boldsymbol{\mathcal{X}}_k) - \Phi(\boldsymbol{\mathcal{X}}_{k+1}) \geq \frac{\mu-1}{2}\|\boldsymbol{\mathcal{X}}_{k+1} - \boldsymbol{\mathcal{X}}_k\|_F^2$, *where* $\mu > 1$ *is an IRSpM algorithm parameter.*

(ii) *The sequence* $\{\boldsymbol{\mathcal{X}}_k\}$ *is bounded.*

(iii) $\lim\limits_{k\to\infty} \|\boldsymbol{\mathcal{X}}_{k+1} - \boldsymbol{\mathcal{X}}_k\|_F = 0.$

(iv) *Any accumulation point of* $\{\boldsymbol{\mathcal{X}}_k\}$ *is a stationary point of* Φ.

Proof. Let $\Psi_i(\overline{\boldsymbol{X}}^{(i)})$ denote the objective function of problem (6). Note $\Phi(\boldsymbol{\mathcal{X}}) = \frac{1}{n_3}\sum_{i=1}^{n_3}\Psi_i(\overline{\boldsymbol{X}}^{(i)})$ and $\|\boldsymbol{\mathcal{X}}_{k+1} - \boldsymbol{\mathcal{X}}_k\|_F = \frac{1}{n_3}\sum_{i=1}^{n_3}\|\overline{\boldsymbol{X}_{k+1}}^{(i)} - \overline{\boldsymbol{X}_k}^{(i)}\|_F^2$. All the assertions immediately hold. □

4 Patch-based Approach for 3D Poissonian Image Deblurring

In this section, we propose a non-local low-rank model for 3D Poissonian image deblurring by exploiting low-rank priors of the non-local similar patch groups extracted from the observed images.

4.1 Problem statement

For a 3D image $\boldsymbol{x} \in \mathbb{R}^N$, the image degradation model under Poisson noise can be written as

$$\boldsymbol{y} = P(\boldsymbol{H}\boldsymbol{x}), \tag{19}$$

where $\boldsymbol{x}$ denotes an image that is not degraded, $\boldsymbol{H} \in \mathbb{R}^{n\times n}$ denotes a matrix operation of the convolution of a PSF, $P(\cdot)$ denotes a process in which the image is contaminated with Poisson noise, and $\boldsymbol{y}$ denotes a degraded image. If $\boldsymbol{H}$ is an identity matrix, the model becomes a simple denoising model. In this paper, we consider periodic boundary conditions and then the blurring operator $\boldsymbol{H}$ keeps the block-cyclic structure.

Since the variance of the Poisson noise is proportional to the intensity of the signal in each pixel, more precisely, assuming that the observed value of image f at position i is independent, we can write

$$P(\boldsymbol{y} \mid \boldsymbol{H}\boldsymbol{x}) = \prod_i \frac{e^{-(\boldsymbol{H}\boldsymbol{x})_i}\,((\boldsymbol{H}\boldsymbol{x})_i)^{y_i}}{y_i!},$$

where y_i denotes the pixel value of the observed image at each position i, and $\boldsymbol{x}$ denotes the original clear image. Using the Bayesian framework, Le et al. [46] proposed a minimization model as follows for 2D Poissonian image deblurring

$$\min_{\boldsymbol{x}} \tau\langle \boldsymbol{H}\boldsymbol{x} - \boldsymbol{y}\log \boldsymbol{H}\boldsymbol{x}, \boldsymbol{1}\rangle + \|\nabla \boldsymbol{x}\|_1, \tag{20}$$

where $\boldsymbol{1}$ denotes the vector whose entries are all ones, the logarithm and multiplication with $\boldsymbol{y}$ are component-wise operations, and $\tau > 0$ is a parameter. The first term of model (20) is the data fidelity term derived from the log-likelihood function of the Poisson distribution, and the second term is the classical discrete TV regularization [47] defined as the composition of the l_1 norm and the first-order difference operator ∇.

For 3D Poissonian image deblurring, the data fidelity term of model (20) for 2D Poissonian image deblurring can also be used. However, due to the ill-posedness of the problem, TV regularization-based methods have some limitations in preserving the image textures, especially for 3D images. Therefore, we propose a non-local low-rank model based on the t-log-S_p low-rank measure.

4.2 Non-local low-rank model for 3D images

Non-local self-similarity for 2D images indicates that for each patch of the 2D image, similar patches can be found in the image and grouped to obtain a low-rank matrix. And using this property, non-local low-rank models have been developed for various applications in image restoration [48–54]. For example, weighted nuclear norm minimization [52] has been applied to image denoising [52], image deblurring [53], Rician noise removal [54] and phase retrieval [51]. For 3D images, the non-local self-similarity property also exists. The non-local low-rank regularization for 3D images can be imposed by using matrix low-rank measures [55] or tensor low-rank measures [56, 57]. For example, Kronecker-Basis-Representation (KBR) tensor sparsity regularization [58] has been applied to multispectral image denoising [58] and low-dose dynamic cerebral perfusion CT reconstruction [59] and low-dose CT sinogram recovery [60]. In the following, we adopt our t-log-S_p tensor low-rank measure proposed in section 2, and develop a non-local low-rank model for 3D Poissonian image deblurring.

First, we group non-local 3D patches, also called cubes, with similarity together by cube matching and form a non-local similar patch tensor. Given a 3D image $\boldsymbol{x}$, suppose it can be divided into L overlapping cubes of size $\sqrt{n_1}\times\sqrt{n_1}\times n_3$, denoted as $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_L\}$. For each reference cube $\boldsymbol{x}_i$ of the image, a total number of n_2 non-local self-similar cubes $\{\boldsymbol{x}_{i,1}, \boldsymbol{x}_{i,2}, \cdots, \boldsymbol{x}_{i,n_2}\}$ can be found by cube matching. Here, the cubes are grouped using Euclidean distances, and the tensor $\mathcal{R}_i(\boldsymbol{x})$ is generated for the reference cube $\boldsymbol{x}_i$ by stacking the grouped unfolding cubes in the ascending order of Euclidean distance in the second dimension, see Definition 4.1.

Definition 4.1. *Given a vectorized 3D image $\boldsymbol{x} \in \mathbb{R}^N$ and a reference vectorized cube $\boldsymbol{x}_i \in \mathbb{R}^{n_1 n_3}$, the non-local similar patch matrix $\boldsymbol{R}_{i,j} \in \mathbb{R}^{n_1 n_3 \times N}$ is a binary matrix (whose terms are 1 or 0), $i = 1, 2. \dots, L, j = 1, 2. \dots .n_2$, such that $\boldsymbol{R}_{i,j}\boldsymbol{x}$ is the j-th vectorized cube in the i-th non-local similar group $\boldsymbol{x}_{i,j}$, that is, $\boldsymbol{R}_{i,j}\boldsymbol{x} = \boldsymbol{x}_{i,j}$. Let $\mathcal{R}_i : \mathbb{R}^N \to \mathbb{R}^{n_1\times n_2\times n_3}$ be the extraction operator for the i-th non-local self-similar tensor defined as*

$$\mathcal{R}_i(\boldsymbol{x}) = \mathrm{fold}_{(2)}([\boldsymbol{R}_{i,1}\boldsymbol{x}, \dots, \boldsymbol{R}_{i,n_2}\boldsymbol{x}]^T).$$

Here, $\mathcal{R}_i(\boldsymbol{x})$ is the constructed tensor for the i-th reference cube. And this tensor describes the spatial correlation along the first dimension, presents the repeated patterns of similar cubes along the second dimension, and keeps the mode-3 correlation of the 3D image along the third dimension. Note that the order of the modes can be switched. And $\mathcal{R}_i(\boldsymbol{x})$ should be a low-rank tensor according to non-local self-similarity if $\boldsymbol{x}$ is a clean image.

Second, we adopt the t-log-S_p low-rank measure to regularize the low-rank properties of these non-local similar patch tensors. By combining the low-rank tensor regularization using t-log-S_p low-rank measure defined as in (4) with the tensor Poissonian image deblurring model (20), a non-local low-rank tensor model for image restoration is as follows

$$\min_{\boldsymbol{x}} \tau\langle \boldsymbol{Hx} - \boldsymbol{y}\log \boldsymbol{Hx}, \boldsymbol{1}\rangle_{\boldsymbol{W}} + \sum_{i=1}^{L} \eta_i \mathcal{M}_{\log,S_p}(\mathcal{R}_i(\boldsymbol{x})), \tag{21}$$

where $\mathcal{R}_i(\boldsymbol{x})$ represents the constructed tensor for each reference cube, $\boldsymbol{W} = \sum_{i=1}^{L} \mathcal{R}_i^T \circ \mathcal{R}_i = \sum_{i=1}^{L} \sum_{l=1}^{n_2} \boldsymbol{R}_{i,l}^T \boldsymbol{R}_{i,l}$ is a diagonal matrix whose main diagonal entries indicate the counts for each pixel, $\langle \cdot, \cdot \rangle$ is the $\boldsymbol{W}$-weighted inner product and $\eta_i > 0$. This model preserves the structural correlation of the constructed tensors, thus obtaining better denoising results.

Lastly, we use variable splitting to reformulate the model. By introducing relaxation variables, and problem (21) can be rewritten as a constrained problem:

$$\min_{\boldsymbol{x}} \tau\langle \boldsymbol{h} - \boldsymbol{y}\log \boldsymbol{h}, \boldsymbol{1}\rangle_{\boldsymbol{W}} + \sum_{i=1}^{L} \eta_i \mathcal{M}_{\log,S_p}(\boldsymbol{\mathcal{L}}_i), \quad s.t. \quad \boldsymbol{Hx} = \boldsymbol{h}, \quad \boldsymbol{\mathcal{L}}_i = \mathcal{R}_i(\boldsymbol{x}).$$

Then by relaxing these equalities of the splitting variables, the constrained problem can be relaxed to an unconstrained problem as follows

$$\min_{\boldsymbol{x},\boldsymbol{h},\boldsymbol{\mathcal{L}}_i} \tau\langle \boldsymbol{h}-\boldsymbol{y}\log\boldsymbol{h},\boldsymbol{1}\rangle_{\boldsymbol{W}} + \frac{\alpha}{2}\|\boldsymbol{h}-\boldsymbol{H}\boldsymbol{x}\|_{\boldsymbol{W}}^2 + \sum_{i=1}^{L}\left[\frac{1}{2}\|\boldsymbol{\mathcal{L}}_i-\mathcal{R}_i(\boldsymbol{x})\|_F^2 + \eta_i\mathcal{M}_{\log,S_p}(\boldsymbol{\mathcal{L}}_i)\right], \tag{22}$$

where $\tau > 0$, $\alpha > 0$ and $\eta_i > 0$. We call this model as the non-local low-rank tensor model for 3D Poissonian image deblurring.

4.3 The full algorithm for 3D Poissonian image deblurring

To solve the proposed model (22) for 3D Poissonian image deblurring, we perform an alternating minimization algorithm with a proximal term as follows.

- **Update of $\boldsymbol{\mathcal{L}}_i$:** given $\boldsymbol{x}=\boldsymbol{x}^k$, we update $\boldsymbol{\mathcal{L}}_i^{k+1}$ by solving the following subproblem

$$\min_{\boldsymbol{\mathcal{L}}_i} \frac{1}{2}\left\|\boldsymbol{\mathcal{L}}_i-\mathcal{R}_i(\boldsymbol{x}^k)\right\|_F^2 + \eta_i\mathcal{M}_{\log,S_p}(\boldsymbol{\mathcal{L}}_i). \tag{23}$$

  We solve this t-log-S_p minimization problem by the t-IRSpM algorithm given in Algorithm 3 using $\boldsymbol{\mathcal{L}}_i^k$ as an initial solution.

- **Update of $\boldsymbol{h}$:** given $\boldsymbol{x}=\boldsymbol{x}^k$, we update $\boldsymbol{h}^{k+1}$ by minimizing problem (22) with respect to $\boldsymbol{h}$ as follows

$$\boldsymbol{h}^{k+1} = \operatorname*{argmin}_{\boldsymbol{h}} \tau\langle \boldsymbol{h}-\boldsymbol{y}\log\boldsymbol{h},\boldsymbol{1}\rangle_{\boldsymbol{W}} + \frac{\alpha}{2}\|\boldsymbol{h}-\boldsymbol{H}\boldsymbol{x}^k\|_{\boldsymbol{W}}^2.$$

  This is a least squares problem. Its closed-form solution is

$$\boldsymbol{h}^{k+1} = \frac{1}{2}\left(\boldsymbol{H}\boldsymbol{x}^k-\frac{\tau}{\alpha}\right) + \sqrt{\frac{1}{4}\left(\boldsymbol{H}\boldsymbol{x}^k-\frac{\tau}{\alpha}\right)^2+\frac{\tau\boldsymbol{y}}{\alpha}}. \tag{24}$$

- **Update of $\boldsymbol{x}$:** given $\boldsymbol{h}=\boldsymbol{h}^{k+1}$ and $\boldsymbol{\mathcal{L}}_i={\boldsymbol{\mathcal{L}}_i}^{k+1}$, the update of the estimated image $\boldsymbol{x}^{k+1}$ at the $(k+1)$-th step is computed by minimizing problem (22) together with a proximal term as follows

$$\boldsymbol{x}^{k+1} = \operatorname*{argmin}_{\boldsymbol{x}} \frac{\alpha}{2}\left\|\boldsymbol{h}^{k+1}-\boldsymbol{H}\boldsymbol{x}\right\|_{\boldsymbol{W}}^2 + \sum_{i=1}^{L}\frac{1}{2}\left\|\boldsymbol{\mathcal{L}}_i^{k+1}-\mathcal{R}_i(\boldsymbol{x})\right\|_F^2 + \frac{\beta}{2}\left\|\boldsymbol{x}-\boldsymbol{x}^k\right\|_{\boldsymbol{W}}^2,$$

  where $\beta > 0$. The update of $\boldsymbol{x}$ has a closed-form solution as follows

$$\boldsymbol{x}^{k+1} = \left[\alpha\boldsymbol{H}^T\boldsymbol{H}+(\beta+1)\boldsymbol{I}\right]^{-1}\left(\alpha\boldsymbol{H}^T\boldsymbol{h}^{k+1}+\sum_{i=1}^{L}\boldsymbol{W}^{-1}\mathcal{R}_i^T\left(\boldsymbol{\mathcal{L}}_i^{k+1}\right)+\beta\boldsymbol{x}^k\right), \tag{25}$$

  where $\mathcal{R}_i^T : \mathbb{R}^{n_1\times n_2\times n_3}\to\mathbb{R}^N$ is an inverse process of $\mathcal{R}_i$ defined as $\mathcal{R}_i^T(\boldsymbol{\mathcal{X}}) = \sum_{j=1}^{n_2}\boldsymbol{R}_{i,j}^T\,\mathrm{vec}(\boldsymbol{\mathcal{X}}(:,j,:))$.

Since the update of $\boldsymbol{h}$ has a closed-form solution, the algorithm can be viewed as alternatively updating variables $\boldsymbol{\mathcal{L}}_i$ and $\boldsymbol{x}$. We call this algorithm as the patch-based tensor logarithmic S_p minimization (TL-SpM) algorithm and summarize it in Algorithm 4. The convergence analysis of this algorithm is presented in the next subsection.

Algorithm 4 Patch-based TLSpM algorithm for 3D Poissonian deblurring

Input: $\boldsymbol{y}$, and parameters τ, α, η_i, and β.

1: **Initialize** $\boldsymbol{x}^0 = \boldsymbol{y}$, $k = 0$
2: Set extraction $\mathcal{R}_i$ by cube matching
3: **repeat**
4: Update $\boldsymbol{\mathcal{L}}_i^{k+1}$ by t-IRSpM$(\mathcal{R}_i(\boldsymbol{x}^k), \eta_i)$ initialized with $\boldsymbol{\mathcal{L}}_i^k$, $i = 1, 2, ..., L$;
5: Update $\boldsymbol{h}^{k+1}$ by eq. (24);
6: Update $\boldsymbol{x}^{k+1}$ by eq. (25);
7: $k \leftarrow k + 1$.
8: **until** convergence

Output: $\boldsymbol{x}^k$

4.4 Convergence analysis of the Patch-based TLSpM algorithm

We can prove that any accumulation point of the sequence $\{(\boldsymbol{x}^*, \boldsymbol{h}^*, \{\boldsymbol{\mathcal{L}}_i{}^*\})\}$, where $\{\boldsymbol{\mathcal{L}}_i{}^*\}$ denotes $\{\boldsymbol{\mathcal{L}}_1{}^*, \boldsymbol{\mathcal{L}}_2{}^*, \dots, \boldsymbol{\mathcal{L}}_L{}^*\}$, generated by Patch-based TLSpM algorithm in Algorithm 4 is a stationary point of the objective function of the proposed model in (22).

For the sake of proving convergence results for Algorithm 4, we assume with loss of generality that the t-IRSpM algorithm in line 4 performs q inner iterations. And we denote the inner updates of $\boldsymbol{\mathcal{L}}_i$ from the initial $\boldsymbol{\mathcal{L}}_i^{qk}$ to the output $\boldsymbol{\mathcal{L}}_i^{q(k+1)}$, which are corresponding to the initial $\boldsymbol{\mathcal{L}}_i^k$ to the output $\boldsymbol{\mathcal{L}}_i^{k+1}$ in line 4.

Definition 4.2. *A point $(\boldsymbol{x}^*, \boldsymbol{h}^*, \{\boldsymbol{\mathcal{L}}_i{}^*\})$ is a first-order stationary point of problem* (22) *if*

$$\begin{aligned}\mathbf{0} &= \tau(\boldsymbol{1} - \frac{\boldsymbol{y}}{\boldsymbol{h}^*}) + \alpha(\boldsymbol{h}^* - \boldsymbol{H}\boldsymbol{x}^*)\\ \mathbf{0} &= \alpha \boldsymbol{H}^T(\boldsymbol{H}\boldsymbol{x}^* - \boldsymbol{h}^*) + \boldsymbol{x}^* - \sum_i \boldsymbol{W}^{-1}\mathcal{R}_i^T(\boldsymbol{\mathcal{L}}_i^*),\end{aligned}$$

where the division of $\boldsymbol{y}$ is a component-wise operation and for $s = 1, 2, \dots, n_3$, $i = 1, 2, \dots, L$,

$$\begin{aligned}\mathbf{0} \in \Big\{\tilde{\boldsymbol{U}}^T \left(\overline{\boldsymbol{L}_i^*}^{(s)} - \overline{\mathcal{R}_i(\boldsymbol{x}^*)}^{(s)}\right) \tilde{\boldsymbol{V}} + \tau p \,\mathrm{Diag}(\boldsymbol{d}) : &(\tilde{\boldsymbol{U}}, \tilde{\boldsymbol{V}}) \in \widetilde{\mathcal{O}}\left(\overline{\boldsymbol{L}_i^*}^{(s)}\right)\\ &\text{and } d_j = \sigma_j^{p-1}(\overline{\boldsymbol{L}_i^*}^{(s)})(\sigma_j^p(\overline{\boldsymbol{L}_i^*}^{(s)}) + \varepsilon)^{-1}\Big\}.\end{aligned}$$

Proposition 4.3. *Let Φ denote the objective function of model* (22)*. Suppose that $\{(\boldsymbol{x}^k, \boldsymbol{h}^k, \{\boldsymbol{\mathcal{L}}_i{}^{qk}\})\}$ is a sequence generated by Algorithm 4. Then the following assertions hold:*

(i) The following inequality holds for $k = 1, 2, \dots$

$$\Phi(\boldsymbol{x}^k, \boldsymbol{h}^k, \{\boldsymbol{\mathcal{L}}_i{}^{qk}\}) - \Phi(\boldsymbol{x}^{k+1}, \boldsymbol{h}^{k+1}, \{\boldsymbol{\mathcal{L}}_i{}^{q(k+1)}\}) \geq \frac{\mu - 1}{2} \sum_{i=1}^{L} \sum_{j=1}^{q} \|\boldsymbol{\mathcal{L}}_i^{qk+j} - \boldsymbol{\mathcal{L}}_i^{qk+j-1}\|_F^2 + \frac{\beta}{2}\|\boldsymbol{x}^{k+1} - \boldsymbol{x}^k\|_{\boldsymbol{W}}^2, \tag{26}$$

where $\mu > 1$ is an IRSpM algorithm parameter.

(ii) The sequence $\{(\boldsymbol{x}^k, \boldsymbol{h}^k, \{\boldsymbol{\mathcal{L}}_i{}^{qk}\})\}$ is bounded.

(iii) $\lim_{k\to\infty} \|\boldsymbol{x}^{k+1} - \boldsymbol{x}^k\|_{\boldsymbol{W}} = 0$, $\lim_{k\to\infty} \|\boldsymbol{h}^{k+1} - \boldsymbol{h}^k\|_F = 0$ and $\lim_{k\to\infty} \|\boldsymbol{\mathcal{L}}_i^{q(k+1)} - \boldsymbol{\mathcal{L}}_i^{qk}\|_F = 0$

(iv) Any accumulation point of $\{(\boldsymbol{x}^k, \boldsymbol{h}^k, \{\boldsymbol{\mathcal{L}}_i{}^{qk}\})\}$ is a stationary point of Ψ.

Proof. (i) By the update of $\boldsymbol{\mathcal{L}}_i^{q(k+1)}$ via t-IRSpM algorithm, it follows from Theorem 3.13 that

$$\Phi(\boldsymbol{x}^k, \boldsymbol{h}^k, \{\boldsymbol{\mathcal{L}}_i{}^{qk}\}) - \Phi(\boldsymbol{x}^k, \boldsymbol{h}^k, \{\boldsymbol{\mathcal{L}}_i{}^{q(k+1)}\}) \geq \frac{\mu - 1}{2} \sum_{i=1}^{L} \sum_{j=1}^{q} \|\boldsymbol{\mathcal{L}}_i^{qk+j} - \boldsymbol{\mathcal{L}}_i^{qk+j-1}\|_F^2.$$

By the updates of $\boldsymbol{h}^{k+1}$ and $\boldsymbol{x}^{k+1}$, we have $\Phi(\boldsymbol{x}^k, \boldsymbol{h}^{k+1}, \{\boldsymbol{\mathcal{L}}_i{}^{q(k+1)}\}) - \Phi(\boldsymbol{x}^k, \boldsymbol{h}^k, \{\boldsymbol{\mathcal{L}}_i{}^{q(k+1)}\}) \geq 0$ and $\Phi(\boldsymbol{x}^{k+1}, \boldsymbol{h}^{k+1}, \{\boldsymbol{\mathcal{L}}_i{}^{q(k+1)}\}) - \Phi(\boldsymbol{x}^k, \boldsymbol{h}^{k+1}, \{\boldsymbol{\mathcal{L}}_i{}^{q(k+1)}\}) \geq \frac{\beta}{2}\|\boldsymbol{x}^{k+1} - \boldsymbol{x}^k\|_{\boldsymbol{W}}^2$. It follows immediately from these inequalities that equation (26) holds.

(ii) Since Φ is bounded below and coercive assertion (ii) holds.

(iii) Summing (26) from $k = 0$ to $k = K$, we have

$$\frac{\mu - 1}{2} \sum_{i=1}^{L} \sum_{j=1}^{q} \|\boldsymbol{\mathcal{L}}_i^{qk+j} - \boldsymbol{\mathcal{L}}_i^{qk+j-1}\|_F^2 + \frac{\beta}{2}\|\boldsymbol{x}^{k+1} - \boldsymbol{x}^k\|_{\boldsymbol{W}}^2 \leq \Phi(\boldsymbol{x}^0, \boldsymbol{h}^0, \{\boldsymbol{\mathcal{L}}_i{}^0\}) - \Phi(\boldsymbol{x}^K, \boldsymbol{h}^K, \{\boldsymbol{\mathcal{L}}_i{}^{qK}\}) < +\infty.$$

Taking $K \to \infty$, we have

$$\sum_{i=1}^{L} \sum_{j=1}^{q} \|\boldsymbol{\mathcal{L}}_i^{qk+j} - \boldsymbol{\mathcal{L}}_i^{qk+j-1}\|_F^2 < +\infty \quad \text{and} \quad \|\boldsymbol{x}^{k+1} - \boldsymbol{x}^k\|_{\boldsymbol{W}}^2 < +\infty.$$

These together with (24) yield assertion (iii).

(iv) Let $(\boldsymbol{x}^*, \boldsymbol{h}^*, \{\boldsymbol{\mathcal{L}}_i{}^*\})$ be an accumulation point of the sequence $\{(\boldsymbol{x}^k, \boldsymbol{h}^k, \{\boldsymbol{\mathcal{L}}_i{}^{qk}\})\}$. Assume that a subsequence $\{(\boldsymbol{x}^k, \boldsymbol{h}^k, \{\boldsymbol{\mathcal{L}}_i{}^{qk}\})\}_{\mathcal{K}}$ converges to $(\boldsymbol{x}^*, \boldsymbol{h}^*, \{\boldsymbol{\mathcal{L}}_i{}^*\})$ as $k \to \infty$.

According to the t-IRSpM algorithm and Theorem 3.13, we have for $s = 1, 2, \ldots, n_3$, $i = 1, 2, \ldots, L$,

$$\mathbf{0} \in \Big\{\tilde{\boldsymbol{U}}^T \left(\overline{\boldsymbol{L}_i^{q(k+1)}}^{(s)} - \overline{\mathcal{R}_i(\boldsymbol{x}^k)}^{(s)}\right) \tilde{\boldsymbol{V}} + \tau p \operatorname{Diag}(\boldsymbol{d}) : (\tilde{\boldsymbol{U}}, \tilde{\boldsymbol{V}}) \in \widetilde{\mathcal{O}}\left(\overline{\boldsymbol{L}_i^{q(k+1)}}^{(s)}\right)$$
$$\text{and } d_j = \sigma_j^{p-1}(\overline{\boldsymbol{L}_i^{q(k+1)}}^{(s)})(\sigma_j^p(\overline{\boldsymbol{L}_i^{q(k+1)}}^{(s)}) + \varepsilon)^{-1}\Big\}.$$

According to the updates of $\boldsymbol{h}^{k+1}$ and $\boldsymbol{x}^{k+1}$, we have

$$\mathbf{0} = \tau(\mathbf{1} - \frac{\boldsymbol{y}}{\boldsymbol{h}^{k+1}}) + \alpha(\boldsymbol{h}^{k+1} - \boldsymbol{H}\boldsymbol{x}^k)$$
$$\mathbf{0} = \alpha\boldsymbol{H}^T(\boldsymbol{H}\boldsymbol{x}^{k+1} - \boldsymbol{h}^{k+1}) + \boldsymbol{x}^{k+1} - \sum_i \boldsymbol{W}^{-1}\mathcal{R}_i^T(\boldsymbol{\mathcal{L}}_i^{q(k+1)}) + \beta(\boldsymbol{x}^{k+1} - \boldsymbol{x}^k),$$

Taking $k \in \mathcal{K}$ approaches ∞ and using assertion (iii), we can obtain the assertion (iv). $\square$

5 Experimental Results

In this section, we demonstrate the performance of the patch-based TLSpM algorithm in Algorithm 4 for 3D Poissonian image deblurring. We compare this algorithm with other Poisson deblurring algorithms including RL [10], ARL [11], VST-BM3D [29] and PURE-LET [33] algorithms. Also, we test the KBR-denoising [61] for 3D Poissonian image deblurring by using our proposed model and algorithm scheme. For example, in the KBR-PoisDebl algorithm, the model is (22) where $\mathcal{M}_{\log,S_p}$ is replaced by KBR. The experiments were implemented in MATLAB 2016b running a 64-bit Ubuntu 18.04 system and executed on an eight-core Intel Xeon E5-2640v3 128GB CPU at 2.6 GHz. The proposed algorithm was accelerated using parallel computing, as the estimation of each patch tensor can be computed in parallel.

5.1 Experiments on fluorescence microscope images

Poisson noise and blur degradation often occur simultaneously in fluorescence microscope images. Fluorescence microscopy is widely used in biological studies to analyze cell and tissue structures. Its resolution is affected by two factors. One is the ambiguity caused by the Abbe diffraction limit, and the other is the noise that strongly depends on the signal. We use 3D fluorescence microscope images for testing. Three test images are "Spherical-beads"[1] ($128 \times 128 \times 64$), "Micro-tubules"[1] ($128 \times 128 \times 64$), and "Pollen"[2] ($256 \times 256 \times 32$). The 10th frontal slice of each original image is shown in Figure 1. To simulate blurry Poissonian images, we adopt the procedure in [62]. First, the original image is scaled by Peak$/I_{\max}$, where $I_{\max}$ is the maximum value of the original image and Peak is the peak value set as 255. Then the image is further convolved with three different 3D blur kernels obtained by a microscope PSF generator[3], including one 3D Gibson & Lanni blur (G&L) [63] and two different 3D Gaussian blur (G1 and G2). Lastly, Poisson noise is added to the blurry image.

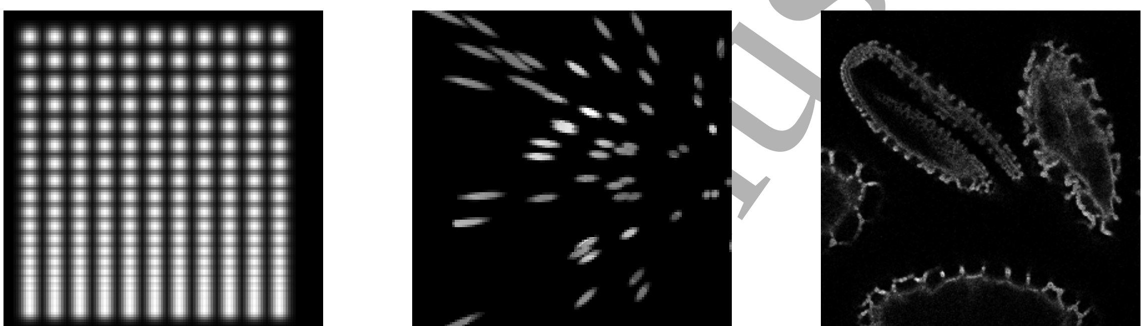


Figure 1: The 10th frontal slices of fluorescence microscope images of "Spherical-beads", "Micro-tubules" and "Pollen", respectively.

For the proposed patch-based TLSpM algorithm, we first set the search window as 35×35 and the number of non-local patches for each group as 60. The cube size is $7 \times 7 \times 7$ for the G&L blur kernel and $6 \times 6 \times 14$ for the G1 and G2 blur kernels. Also, the parameters $p = 0.95$, $\beta = 0.0001$ and $\mu = 1.0001$ and the rest are shown in Table 1. And to achieve better performance, the cube matching $\mathcal{R}_i$ is also updated for certain iterations and then remains unchanged afterward.

Table 1: Parameter settings for patch-based TLSpM algorithm

Image	Spherical-beads			Micro-tubules			Pollen		
PSF	G&L	G1	G2	G&L	G1	G2	G&L	G1	G2
α	20	20	2	20	20	2	20	20	2
τ	150	200	80	150	200	80	240	200	60
η_i	5000	5000	900	5000	5000	900	5000	35000	700

In the experiment, the peak signal-to-noise ratio (PSNR) [64] and structural similarity index measure (SSIM) [65] are used to measure the quality of the restored images. In particular, the PSNR value is defined as

$$\mathrm{PSNR} = 10 \log_{10} \frac{\mathrm{Peak}^2}{\|\boldsymbol{x}^* - \boldsymbol{x}\|_2^2},$$

where $\boldsymbol{x}^*$ is the restored image and $\boldsymbol{x}$ is the original image. And the SSIM value is defined in [65].

[1] The "Spherical-beads" and "Micro-tubules" images are collected from http://bigwww.epfl.ch/deconvolution/index.html#data.
[2] The "Pollen" image is collected from http://www. cellimagelibrary.org/images/35532.
[3] The software package is downloaded from http://bigwww.epfl.ch/algorithms/psfgenerator/.

Table 2: PSNR and SSIM comparison among different algorithms under different blur kernels

Image	Metric	Spherical-beads			Micro-tubules			Pollen		
PSF		G&L	G1	G2	G&L	G1	G2	G&L	G1	G2
Noisy	PSNR	14.53	13.82	14.67	19.66	19.64	19.66	23.77	23.76	23.64
	SSIM	0.291	0.197	0.315	0.220	0.221	0.224	0.429	0.462	0.450
RL	PSNR	18.80	17.78	18.60	20.85	20.95	20.92	26.46	25.17	25.82
	SSIM	0.782	0.674	0.766	0.325	0.337	0.324	0.626	0.475	0.581
ARL	PSNR	17.44	15.64	17.44	19.79	19.81	19.74	24.72	23.23	23.90
	SSIM	0.639	0.451	0.686	0.305	0.286	0.280	0.557	0.417	0.501
VST-BM4D	PSNR	19.09	19.13	19.24	22.01	22.32	22.18	27.25	25.60	26.51
	SSIM	0.756	0.746	0.766	0.358	0.381	0.365	0.642	0.546	0.594
PURE-LET	PSNR	19.41	18.97	19.28	22.39	22.72	22.69	28.23	26.82	27.59
	SSIM	0.779	0.734	0.787	0.357	0.356	0.371	0.741	0.595	0.627
PoisDebl-KBR	PSNR	**19.70**	19.21	19.56	23.15	23.55	23.67	28.40	27.34	28.10
	SSIM	**0.820**	0.787	0.812	0.607	0.574	0.628	0.717	0.600	0.658
Patch-based	PSNR	19.67	**19.42**	**20.97**	**23.58**	**23.65**	**23.87**	**28.49**	**27.35**	**28.40**
TLSpM (ours)	SSIM	0.788	**0.795**	**0.846**	**0.634**	**0.602**	**0.634**	**0.720**	**0.689**	**0.718**

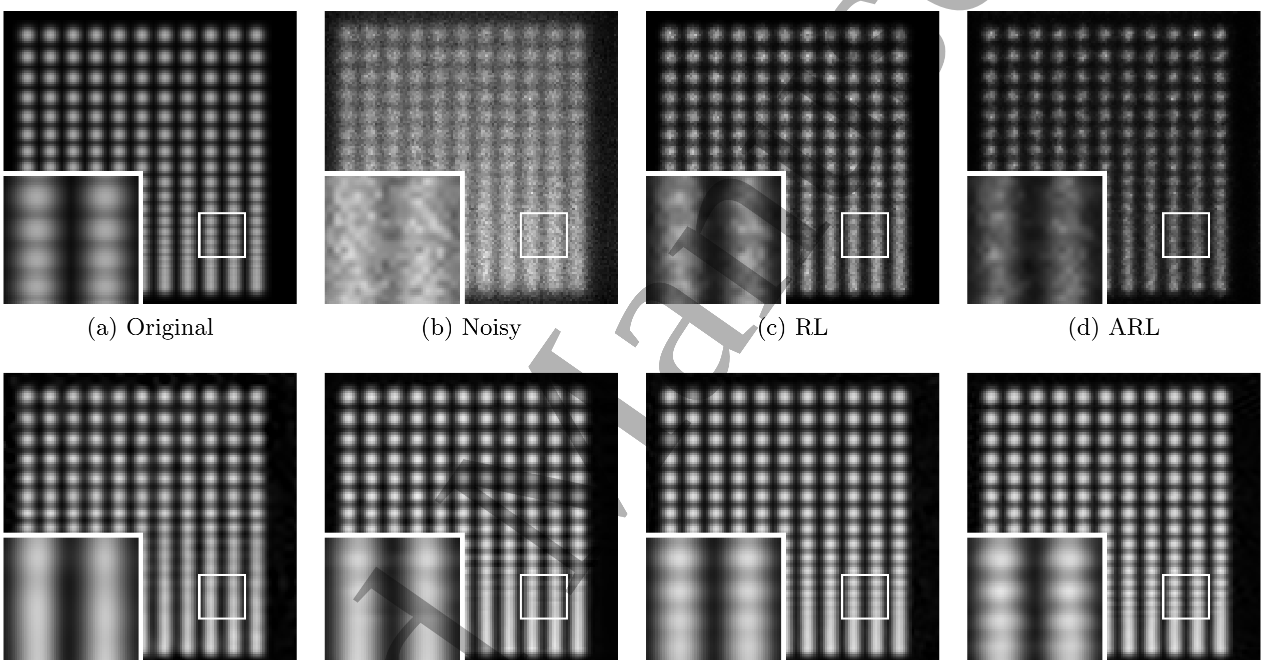
The PSNR and SSIM values of the restored images obtained by different algorithms are shown in Table 2. It shows that the proposed patch-based TLSpM algorithm achieves the best numerical values for most of the testing cases. For example, for "Micro-tublules" image with the G2 blur kernel, the PSNR value of the proposed algorithm exceeds the state-of-the-art PURE-LET algorithm by 1.18 dB. The PoisDebl-KBR method that is modified from our proposed model performs very competitive numerical results, achieving only 0.17 dB in average less than our Patch-based TLSpM in terms of PSNR values.

To evaluate the visual quality of the restored images obtained by different algorithms, we compare several selected slices of the restored 3D images in Figures 2-4. In Figure 2, for the "Spherical-beads" image with the G&L blur kernel, the proposed patch-based TLSpM algorithm obtains the best performance in preserving the spherical structure of beads and separating distinct beads. In contrast, The RL and ARL algorithms were not able to remove Poisson noise and restore the shape of the beads; the VST-BM4D and PURE-LET fail to separate the distinct beads if they are too close to each other; and the PoisDebl-KBR method can separate the beads but the gaps between the beads are not as clear as our proposed method, even though the PSNR value of PoisDebl-KBR method exceeds our proposed method by 0.03dB. In Figure 3, the lateral slice of the original "Micro-tubules" image contains many luminous points. The ARL and RL algorithms cannot recognize luminous points, while the VST-BM4D, PURE-LET, PoisDebl-KBR and proposed algorithms can identify most of the luminous points. In fact, the proposed algorithm can restore images with higher accuracy and fewer artifacts, compared to the VST-BM4D, PURE-LET and PoisDebl-KBR algorithms, as shown in the zoomed-in image of Figure 3. Lastly, in Figure 4, for the "Pollen" image with the G2 blur kernel, the proposed and PoisDebl-KBR algorithms can recover the pattern of the cell wall, while the RL and ARL algorithms fail to remove the noise on the cell wall and the state-of-the-art VST-BM4D and PURE-LET algorithms restore blurry cell walls without details.

All in all, the proposed patch-based TLSpM algorithm outperforms the competing algorithms in removing Poisson noise and retrieving details from blurry images.

5.2 Analysis on the parameter p

The denoising performance of the proposed patch-based TLSpM method is related to the parameter p, which is used in the t-log-S_p low-rank measure. We conduct a sensitivity analysis on parameter p using the "Spherical-beads" image with a G2 blur kernel. Figure 5(a) presents the plot of the PSNR value for $p \in (0, 1)$ vs the number of iterations. We can observe that when $p \in (0, 0.65)$, the PSNR value decreases as p decreases; when $p \in [0.65, 1)$, the differences in PSNR are not very significant. To analyze



(a) Original (b) Noisy (c) RL (d) ARL

(e) VST-BM4D (f) PURE-LET (g) PoisDebl-KBR (h) Patch-based TLSpM (ours)

Figure 2: The 20th frontal slices of the images restored by different algorithms from the noisy "Spherical-beads" image with the G&L blur kernel. The PSNR values of the restored images are: (b) noisy image (14.53 dB); (c) RL (18.80 dB); (d) ARL (17.44 dB); (e) VST-BM4D (19.09 dB); (f) PURE-LET (19.41 dB); (g) PoisDebl-KBR (19.70 dB); (h) Patch-based TLSpM (ours) (19.67 dB).

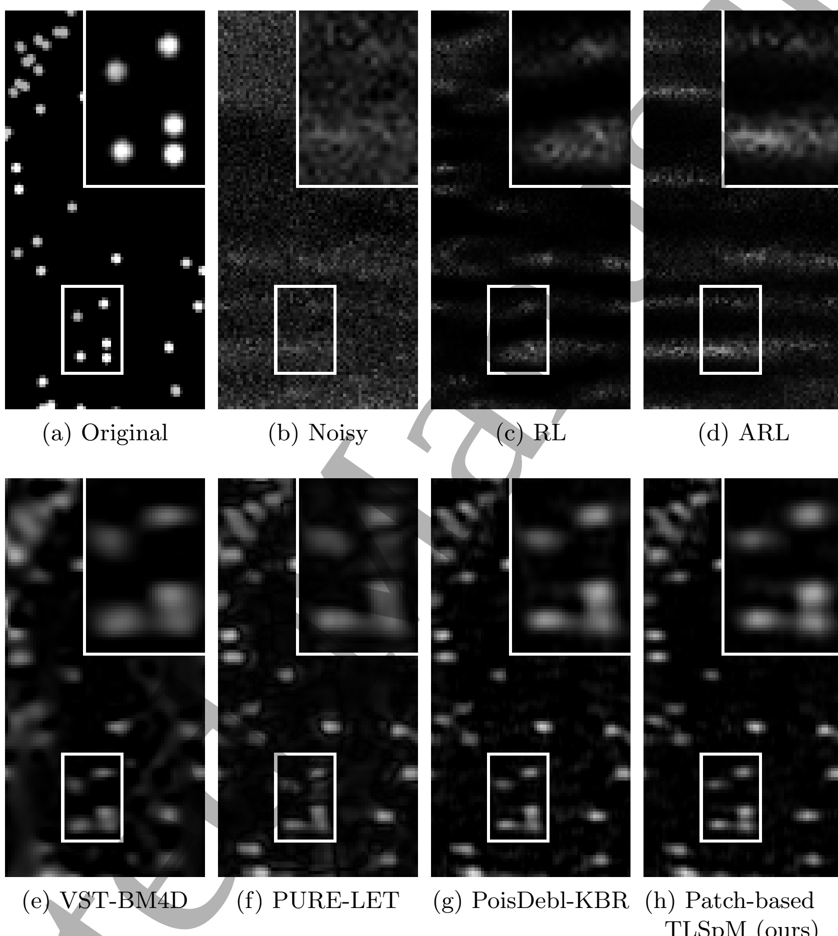


Figure 3: The 33rd lateral slices of the images restored by different algorithms from the noisy "Microtubules" image with the G1 blur kernel. The PSNR values of the restored images are: (b) noisy image (19.64 dB); (c) RL (20.95 dB); (d) ARL (19.81 dB); (e) VST-BM4D (22.32 dB); (f) PURE-LET (22.72 dB); (g) PoisDebl-KBR (23.55 dB); (h) Patch-based TLSpM (ours) (23.65 dB).

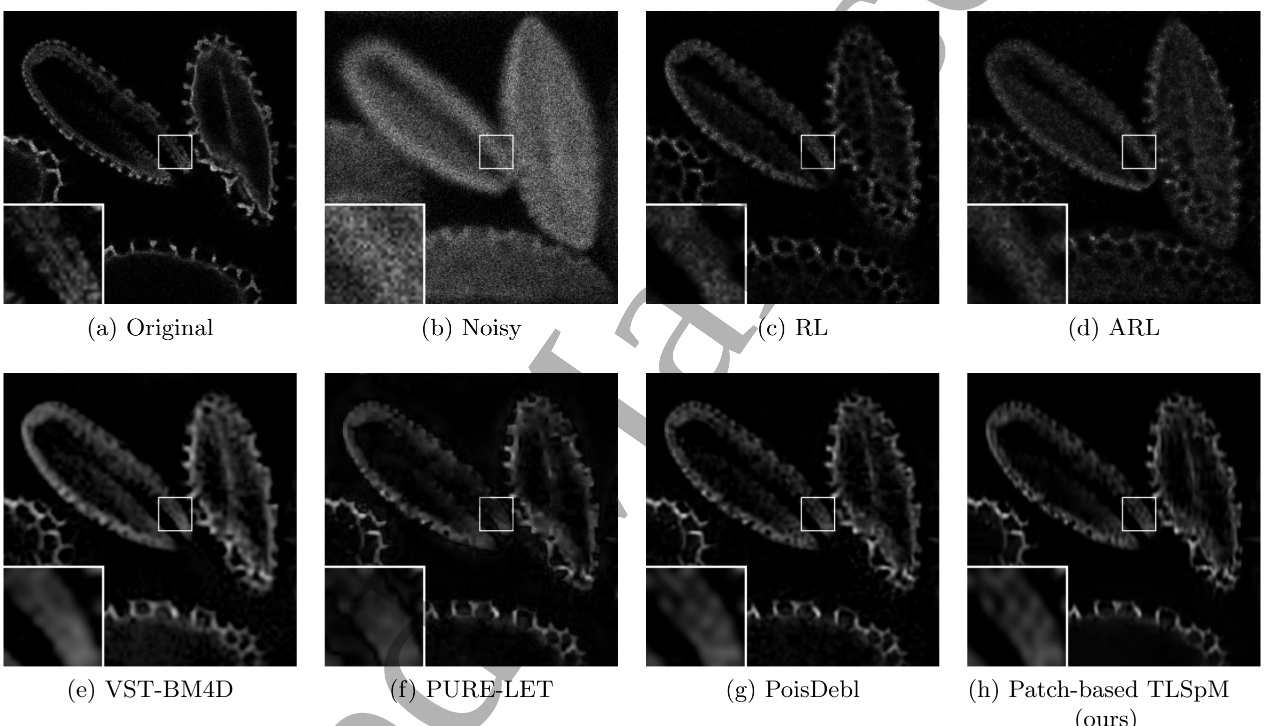
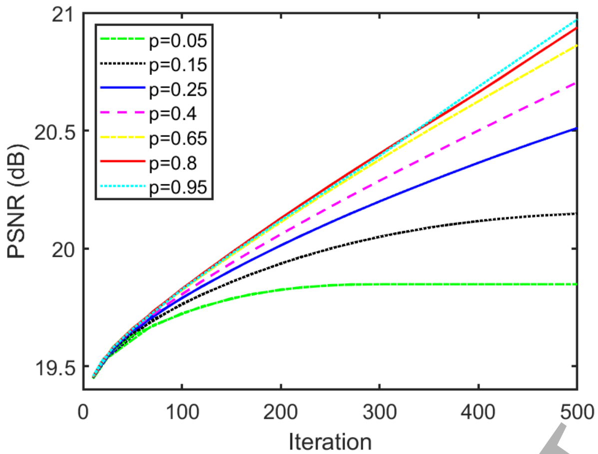
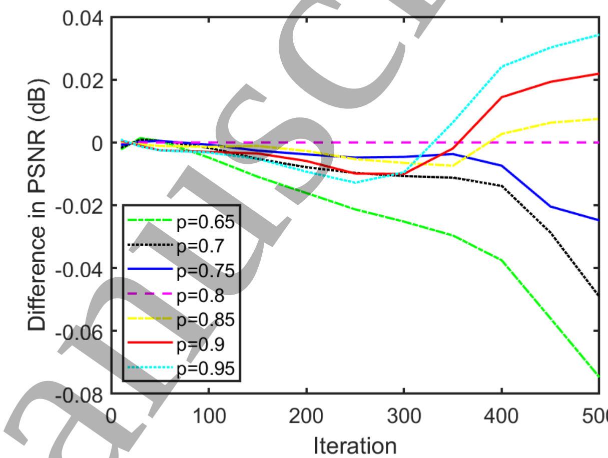


Figure 4: The 12th frontal slices of the images restored by different algorithms from the noisy "Pollen" image with the G2 blur kernel. The PSNR values of the restored images are: (b) noisy image (23.64 dB); (c) RL (25.82 dB); (d) ARL (23.90 dB); (e) VST-BM4D (25.51 dB); (f) PURE-LET (27.59 dB); (g) PoisDebl-KBR (28.10 dB); (h) Patch-based TLSpM (ours) (28.40 dB).

the sensitivity of parameter $p \in [0.65, 1)$, we use the PSNR value for $p = 0.8$ as a reference and compute the difference between the PSNR value for each p and the reference PSNR. Figure 5(b) presents the plot of the PSNR difference vs the number of iterations. When fewer than 100 iterations are performed, the PSNR differences are not significant; when 100-300 iterations are performed, $p = 0.8$ performs the best; when 500 iterations are performed, $p = 0.95$ performs the best. In summary, the proposed method can achieve satisfactory performance by choosing $p \in [0.65, 1)$. When early stop is preferred, one may choose $p = 0.8$; otherwise, one may choose $p = 0.95$.



(a) Plot of PSNR values.



(b) Plot of differences in PSNR values.

Figure 5: Sensitivity analysis of parameter p. (a) Plot of the PSNR value for $p \in (0, 1)$ vs the number of iterations; (b) The PSNR difference for $p \in [0.65, 1)$ vs the number of iterations.

6 Conclusion

In this paper, we first define a new t-log-S_p low-rank measure for tensors. Then we propose a patch-based non-local low-rank approach, called patched-based TLSpM, for removing blur and Poisson noise. The experimental results show that this algorithm is effective in improving the image quality of 3D fluorescence microscopes, and it is superior to the existing methods in terms of visual quality and quantitative quality measures.

Acknowledge

The authors would like to thank the anonymous reviewers and Editorial Board Member for providing valuable comments and suggestions which led to significant improvements in our article. This work is supported in part by the National Natural Science Foundation of China under grants U21A20455, 61972265, and 11871348, by the Natural Science Foundation of Guangdong Province of China under grants 2020B1515310008 and 2023A1515011691, by the Educational Commission of Guangdong Province of China under grant 2019KZDZX1007, and by the PolyU internal grant P0040271.

Data availability statement

The data that support the findings of this study are available upon reasonable request from the authors.

References

[1] Yide Zhang, Yinhao Zhu, Evan Nichols, Qingfei Wang, Siyuan Zhang, Cody Smith, and Scott Howard. A Poisson-Gaussian denoising dataset with real fluorescence microscopy images. In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 11702–11710, 2019.

[2] Elena Anisimova, Jan Bednar, and Petr Pata. Astronomical image denoising using curvelet and starlet transform. In *2013 23rd International Conference Radioelektronika (RADIOELEKTRONIKA)*, pages 255–260, 2013.

[3] Vladimir. Y. Panin, Gengsheng L. Zeng, and Grant T. Gullberg. Total variation regulated EM algorithm [SPECT reconstruction]. *IEEE Transactions on Nuclear Science*, 46(6):2202–2210, 1999.

[4] Wei Zheng, Si Li, Andrzej Krol, C. Ross Schmidtlein, Xueying Zeng, and Yuesheng Xu. Sparsity promoting regularization for effective noise suppression in SPECT image reconstruction. *Inverse Problems*, 35(11):115011, Oct 2019.

[5] Lawrence A. Shepp and Yehuda Vardi. Maximum likelihood reconstruction for emission tomography. *IEEE Transactions on Medical Imaging*, 1(2):113–122, 1982.

[6] Neil Savage. Medical imagers lower the dose. *IEEE Spectrum*, 47(3):14–16, 2010.

[7] Renliang Gu and Aleksandar Dogandžić. Blind X-ray CT image reconstruction from polychromatic Poisson measurements. *IEEE Transactions on Computational Imaging*, 2(2):150–165, 2016.

[8] Ghada Hamed, Mohammed Marey, Safaa El-Sayed Amin, and Mohamed F. Tolba. A proposed model for denoising breast mammogram images. In *2018 13th International Conference on Computer Engineering and Systems (ICCES)*, pages 652–657, 2018.

[9] Pei Chen and K. Barner. Maximum likelihood reconstruction for tomosynthesis. In *2003 IEEE 29th Annual Proceedings of Bioengineering Conference*, pages 59–60, 2003.

[10] William H. Richardson. Bayesian-based iterative method of image restoration. *Journal of the Optical Society of America*, 62(1):55–59, 1972.

[11] Zohair Al-Ameen. Faster deblurring for digital images using an ameliorated Richardson-Lucy algorithm. *IEIE Transactions on Smart Processing & Computing*, 7(4):289–295, 2018.

[12] Nicolas Dey, Laure Blanc-Feraud, Christophe Zimmer, Pascal Roux, Zvi Kam, Jean-Christophe Olivo-Marin, and Josiane Zerubia. Richardson-Lucy algorithm with total variation regularization for 3D confocal microscope deconvolution. *Microscopy Research and Technique*, 69(4):260–266, APR 2006.

[13] Zachary T. Harmany, Roummel F. Marcia, and Rebecca M. Willett. This is spiral-tap: Sparse Poisson intensity reconstruction algorithms-theory and practice. *IEEE Transactions on Image Processing*, 21(3):1084–1096, MAR 2012.

[14] S. Bonettini and V. Ruggiero. An alternating extragradient method for total variation-based image restoration from Poisson data. *Inverse Problems*, 27(9):095001, Jul 2011.

[15] Mário A. T. Figueiredo and José M. Bioucas-Dias. Restoration of Poissonian images using alternating direction optimization. *IEEE Transactions on Image Processing*, 19(12):3133–3145, 2010.

[16] Liyan Ma, Lionel Moisan, Jian Yu, and Tieyong Zeng. A dictionary learning approach for Poisson image deblurring. *IEEE Transactions on Medical Imaging*, 32(7):1277–1289, 2013.

[17] Dai-Qiang Chen and Li-Zhi Cheng. Spatially adapted regularization parameter selection based on the local discrepancy function for Poissonian image deblurring. *Inverse Problems*, 28(1):015004, Dec 2011.

[18] Tingting Zhang, Jie Chen, Caiying Wu, Zhifei He, Tieyong Zeng, and Qiyu Jin. Edge adaptive hybrid regularization model for image deblurring. *Inverse Problems*, 38(6):065010, May 2022.

[19] Simon Setzer, Gabriele Steidl, and Tanja Teuber. Deblurring Poissonian images by split Bregman techniques. *Journal of Visual Communication and Image Representation*, 21(3):193–199, APR 2010.

[20] Dai-Qiang Chen. Regularized generalized inverse accelerating linearized alternating minimization algorithm for frame-based Poissonian image deblurring. *SIAM Journal on Imaging Sciences*, 7(2):716–739, 2014.

[21] CÉdric Vonesch and Michael Unser. A fast multilevel algorithm for wavelet-regularized image restoration. *IEEE Transactions on Image Processing*, 18(3):509–523, 2009.

[22] Mikael Carlavan and Laure Blanc-Feraud. Sparse Poisson noisy image deblurring. *IEEE Transactions on Image Processing*, 21(4):1834–1846, 2012.

[23] Haimiao Zhang, Yichuan Dong, and Qibin Fan. Wavelet frame based Poisson noise removal and image deblurring. *Signal Processing*, 137:363–372, 2017.

[24] T. Jeong, H. Woo, and S. Yun. Frame-based Poisson image restoration using a proximal linearized alternating direction method. *Inverse Problems*, 29(7):075007, Jun 2013.

[25] Jingjing Liu, Yifei Lou, Guoxi Ni, and Tieyong Zeng. An image sharpening operator combined with framelet for image deblurring. *Inverse Problems*, 36(4):045015, Mar 2020.

[26] Stamatios Lefkimmiatis and Michael Unser. Poisson image reconstruction with Hessian Schatten-norm regularization. *IEEE Transactions on Image Processing*, 22(11):4314–4327, 2013.

[27] FranÇois-Xavier Dupe, Jalal M. Fadili, and Jean-Luc Starck. A proximal iteration for deconvolving Poisson noisy images using sparse representations. *IEEE Transactions on Image Processing*, 18(2):310–321, 2009.

[28] F. J. Anscombe. The transformation of Poisson, binomial and negative-binomial data. *Biometrika*, 35(3-4), 1957.

[29] Lucio Azzari and Alessandro Foi. Variance stabilization in Poisson image deblurring. In *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*, pages 728–731, 2017.

[30] Kostadin Dabov, Alessandro Foi, Vladimir Katkovnik, and Karen Egiazarian. Image denoising with block-matching and 3D filtering. In *Image Processing: Algorithms and Systems, Neural Networks, and Machine Learning*, pages 354–365, 2006.

[31] Matteo Maggioni, Vladimir Katkovnik, Karen Egiazarian, and Alessandro Foi. Nonlocal transform-domain filter for volumetric data denoising and reconstruction. *IEEE Transactions on Image Processing*, 2013.

[32] Jizhou Li, Florian Luisier, and Thierry Blu. Pure-let image deconvolution. *IEEE Transactions on Image Processing*, 27(1):92–105, 2018.

[33] Jizhou Li, Florian Luisier, and Thierry Blu. Pure-let deconvolution of 3D fluorescence microscopy images. In *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*, pages 723–727, 2017.

[34] Yuan Xie, Shuhang Gu, Yan Liu, Wangmeng Zuo, Wensheng Zhang, and Lei Zhang. Weighted Schatten p-norm minimization for image denoising and background subtraction. *IEEE Transactions on Image Processing*, 25(10):4842–4857, 2016.

[35] Misha E Kilmer and Carla D Martin. Factorization strategies for third-order tensors. *Linear Algebra and its Applications*, 435(3):641–658, 2011.

[36] Misha E Kilmer, Karen Braman, Ning Hao, and Randy C Hoover. Third-order tensors as operators on matrices: A theoretical and computational framework with applications in imaging. *SIAM Journal on Matrix Analysis and Applications*, 34(1):148–172, 2013.

[37] Canyi Lu, Jiashi Feng, Yudong Chen, Wei Liu, Zhouchen Lin, and Shuicheng Yan. Tensor robust principal component analysis with a new tensor nuclear norm. *IEEE transactions on pattern analysis and machine intelligence*, 42(4):925–938, 2019.

[38] Chunsheng Liu, Hong Shan, and Chunlei Chen. Tensor p-shrinkage nuclear norm for low-rank tensor completion. *Neurocomputing*, 387:255–267, 2020.

[39] Quanxue Gao, Pu Zhang, Wei Xia, Deyan Xie, Xinbo Gao, and Dacheng Tao. Enhanced tensor RPCA and its application. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(6):2133–2140, 2021.

[40] Yu-Bang Zheng, Ting-Zhu Huang, Xi-Le Zhao, Tai-Xiang Jiang, Tian-Hui Ma, and Teng-Yu Ji. Mixed noise removal in hyperspectral image via low-fibered-rank regularization. *IEEE Transactions on Geoscience and Remote Sensing*, 58(1):734–749, 2020.

[41] Maryam Fazel, Haitham Hindi, and Stephen P. Boyd. Log-det heuristic for matrix rank minimization with applications to Hankel and Euclidean distance matrices. In *Proceedings of American Control Conference*, volume 3, pages 2156 – 2162, 2003.

[42] Zhaosong Lu, Yong Zhang, and Jian Lu. ℓ_p Regularized low-rank approximation via iterative reweighted singular value minimization. *Computational Optimization and Applications*, 68(3):619–642, 2017.

[43] Wangmeng Zuo, Deyu Meng, Lei Zhang, Xiangchu Feng, and David Zhang. A generalized iterated shrinkage algorithm for non-convex sparse coding. In *2013 IEEE International Conference on Computer Vision*, pages 217–224, 2013.

[44] Adrian S. Lewis and Hristo S. Sendov. Nonsmooth analysis of singular values. part I: Theory. *Set-Valued Analysis*, 13(3):213–241, 2005.

[45] James M. Ortega and Werner C. Rheinboldt. *Iterative solution of nonlinear equations in several variables.* Academic Press, New York, 1970.

[46] Triet Le, Rick Chartrand, and Thomas J. Asaki. A variational approach to reconstructing images corrupted by Poisson noise. *Journal of Mathematical Imaging and Vision*, 27(3):257–263, 2007.

[47] Leonid I. Rudin, Stanley Osher, and Emad Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 60(1):259–268, 1992.

[48] Chen Xu, Xiaoxia Liu, Jian Zheng, Lixin Shen, Qingtang Jiang, and Jian Lu. Nonlocal low-rank regularized two-phase approach for mixed noise removal. *Inverse Problems*, 37(8):085001, Jul 2021.

[49] Xiaoxia Liu, Jian Lu, Lixin Shen, Chen Xu, and Yuesheng Xu. Multiplicative noise removal: Nonlocal low-rank model and its proximal alternating reweighted minimization algorithm. *SIAM J. Imag. Sci.*, 13(3):1595–1629, 2020.

[50] Shanzhou Niu, Gaohang Yu, Jianhua Ma, and Jing Wang. Nonlocal low-rank and sparse matrix decomposition for spectral CT reconstruction. *Inverse Problems*, 34(2):024003, Jan 2018.

[51] Zhi Li, Ming Yan, Tieyong Zeng, and Guixu Zhang. Phase retrieval from incomplete data via weighted nuclear norm minimization. *Pattern Recognition*, 125:108537–, 2022.

[52] Shuhang Gu, Lei Zhang, Wangmeng Zuo, and Xiangchu Feng. Weighted nuclear norm minimization with application to image denoising. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2862–2869, 2014.

[53] Liyan Ma, Li Xu, and Tieyong Zeng. Low rank prior and total variation regularization for image deblurring. *Journal of Scientific Computing*, 2017.

[54] Jian Lu, Jiapeng Tian, Qingtang Jiang, Xiaoxia Liu, and Yuru Zou. Rician noise removal via weighted nuclear norm penalization. *Applied and Computational Harmonic Analysis*, 53(4), 2021.

[55] Jian Lu, Chen Xu, Zhenwei Hu, Xiaoxia Liu, Qingtang Jiang, Deyu Meng, and Zhouchen Lin. A new nonlocal low-rank regularization method with applications to magnetic resonance image denoising. *Inverse Problems*, 38(6):065012, May 2022.

[56] Duo Qiu, Minru Bai, Michael K. Ng, and Xiongjun Zhang. Nonlocal robust tensor recovery with nonconvex regularization. *Inverse Problems*, 37(3):035001, Jan 2021.

[57] Jize Xue, Yongqiang Zhao, Wenzhi Liao, and Jonathan Cheung-Wai Chan. Nonlocal low-rank regularized tensor decomposition for hyperspectral image denoising. *IEEE Transactions on Geoscience and Remote Sensing*, 57(7):5174–5189, 2019.

[58] Qi Xie, Qian Zhao, Deyu Meng, and Zongben Xu. Kronecker-basis-representation based tensor sparsity and its applications to tensor recovery. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 1–1, 2017.

[59] Dong Zeng, Qi Xie, Wenfei Cao, Jiahui Lin, Hao Zhang, Shanli Zhang, Jing Huang, Zhaoying Bian, Deyu Meng, and Zongben Xu. Low-dose dynamic cerebral perfusion computed tomography reconstruction via Kronecker-basis-representation tensor sparsity regularization. *IEEE Transactions on Medical Imaging*, PP(12):1–1, 2017.

[60] Jian Lu, Huaxuan Hu, Yuru Zou, Zhaosong Lu, Xiaoxia Liu, Keke Zu, and Lin Li. A nonlocal Kronecker-basis-representation method for low-dose CT sinogram recovery. *Journal of Computational Mathematics*, 2023.

[61] Qi Xie, Qian Zhao, Deyu Meng, and Zongben Xu. Kronecker-basis-representation based tensor sparsity and its applications to tensor recovery. *IEEE transactions on pattern analysis and machine intelligence*, 40(8):1888–1902, 2017.

[62] Liyan Ma, Lionel Moisan, Jian Yu, and Tieyong Zeng. A dictionary learning approach for Poisson image deblurring. *IEEE Transactions on medical imaging*, 32(7):1277–1289, 2013.

[63] Sarah Frisken Gibson and Frederick Lanni. Experimental test of an analytical model of aberration in an oil-immersion objective lens used in three-dimensional light microscopy. *JOSA A*, 9(1):154–166, 1992.

[64] Alan C. Bovik. *Handbook of image and video processing.* Academic press, 2010.

[65] Zhou Wang, Alan C. Bovik, Hamid R. Sheikh, and Eero P. Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612, 2004.