

PAPER • OPEN ACCESS

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


To cite this article: Jian Lu *et al* 2024 *Inverse Problems* **40** 065010

View the [article online](#) for updates and enhancements.

You may also like

- [Thaddäus Derfflinger's Sunspot Observations during 1802–1824: A Primary Reference to Understand the Dalton Minimum](#)
Hisashi Hayakawa, Bruno P. Besser, Tomoya Iju et al.
- [Multipliers for continuous frames in Hilbert spaces](#)
P Balazs, D Bayer and A Rahimi
- [Rényi entropy and subsystem distances in finite size and thermal states in critical XY chains](#)
Raúl Arias and Jiaju Zhang

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

Jian Lu^{1,2,6} , Lin Huang^{1,6}, Xiaoxia Liu^{3,*} , Ning Xie^{4,*},
Qingtang Jiang⁵ and Yuru Zou¹

¹ Shenzhen Key Laboratory of Advanced Machine Learning and Applications, School of Mathematical Sciences, Shenzhen University, Shenzhen 518060, People's Republic of China

² National Center for Applied Mathematics Shenzhen (NCAMS), Shenzhen 518055, People's Republic of China

³ Department of Applied Mathematics, The Hong Kong Polytechnic University, Hong Kong Special Administrative Region of China, People's Republic of China

⁴ Guangdong Key Laboratory of Intelligent Information Processing, College of Information Engineering, Shenzhen University, Shenzhen 518060, People's Republic of China

⁵ Department of Mathematics and Statistics, University of Missouri-St. Louis, St. Louis, MO 63121, United States of America

E-mail: xiaoxia.liu@polyu.edu.hk and ningxie@szu.edu.cn

Received 21 July 2023; revised 28 March 2024

Accepted for publication 19 April 2024

Published 9 May 2024



CrossMark

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 ($t\text{-log-}S_p$) low-rank measure and a tensor iteratively reweighted Schatten- p minimization 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

⁶ J Lu and L Huang contributed equally to this work.

* Authors to whom any correspondence should be addressed.



Original Content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](https://creativecommons.org/licenses/by/4.0/). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

show that the proposed patch-based tensor logarithmic Schatten- p minimization 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 RL (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 model by sequential quadratic approximations; Bonettini and Ruggiero [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 and Bioucas-Dias [15] solved the model by the alternating direction method of multipliers; Liyan *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 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. \mathcal{X} ; matrices are represented by bold capital letters, e.g. \mathbf{X} ; vectors are represented by bold lowercase letters, e.g. \mathbf{x} ; and scalars are represented by lowercase letters, e.g. x . For an N -order tensor $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times \dots \times n_N}$, the vectorization of \mathcal{X} is denoted as $\mathbf{x} = \text{vec}(\mathcal{X}) \in \mathbb{R}^{n_1 n_2 \dots n_N}$, and the j th element of \mathbf{x} is equal to the (i_1, i_2, \dots, i_N) th element of \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 \mathcal{X} is denoted as $\mathbf{X}_{(k)} \in \mathbb{R}^{n_k \times \prod_{s \neq k} n_s}$, and the (i_k, j) th element of $\mathbf{X}_{(k)}$ is equal to the (i_1, i_2, \dots, i_N) th element of \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 $\mathbf{X}_{(k)} = \text{unfold}_{(k)}(\mathcal{X})$ and $\mathcal{X} = \text{fold}_{(k)}(\mathbf{X}_{(k)})$, respectively.

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

$$\text{bdiag}(\bar{\mathcal{X}}) = \begin{bmatrix} \bar{\mathbf{X}}^{(1)} & & & \\ & \bar{\mathbf{X}}^{(2)} & & \\ & & \ddots & \\ & & & \bar{\mathbf{X}}^{(n_3)} \end{bmatrix},$$

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

$$\text{bcirc}(\mathcal{X}) = \begin{bmatrix} \mathbf{X}^{(1)} & \mathbf{X}^{(n_3)} & \dots & \mathbf{X}^{(2)} \\ \mathbf{X}^{(2)} & \mathbf{X}^{(1)} & \dots & \mathbf{X}^{(3)} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{X}^{(n_3)} & \mathbf{X}^{(n_3-1)} & \dots & \mathbf{X}^{(1)} \end{bmatrix}.$$

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

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

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

$$\mathcal{X}^T = \text{bfold} \left(\left[\mathbf{X}^{(1)}, \mathbf{X}^{(n_3)}, \mathbf{X}^{(n_3-1)}, \dots, \mathbf{X}^{(2)} \right]^T \right).$$

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

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

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

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:

$$\bar{\mathcal{Z}}^{(i)} = \bar{\mathcal{X}}^{(i)} \bar{\mathcal{Y}}^{(i)}, \tag{1}$$

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 \mathcal{X} and \mathcal{Y} .

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

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

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

Note that each frontal slice of $\overline{\mathcal{I}}$ is the identity matrix and each frontal slice of $\overline{\mathcal{Q}}$, where \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 (t-SVD, [35]). Let $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$. Then there exist $\mathcal{U} \in \mathbb{R}^{n_1 \times n_1 \times n_3}$, $\mathcal{S} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ and $\mathcal{V} \in \mathbb{R}^{n_2 \times n_2 \times n_3}$ such that:

$$\mathcal{X} = \mathcal{U} * \mathcal{S} * \mathcal{V}^T, \quad (2)$$

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

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

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

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 \mathcal{S} , that is, $\text{rank}_t(\mathcal{X}) = \#\{i : \mathcal{S}(i, i, 1) \neq 0\}$. 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 (tensor nuclear norm, [37]). Let $\mathcal{X} = \mathcal{U} * \mathcal{S} * \mathcal{V}^T$ be the t-SVD of $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$. Then the tensor nuclear norm is defined as

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

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

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

As the frontal slices $\overline{\mathcal{X}}^{(i)}$ 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 $\mathbf{X} \in \mathbb{R}^{m \times n}$, the matrix logarithmic Schatten- p (log- S_p) low-rank measure of \mathbf{X} is defined as

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

where $\varepsilon > 0$, $0 < p \leq 1$, and $\sigma_j(\mathbf{X})$ represents the j th largest singular value of \mathbf{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 $\bar{\mathbf{X}}^{(i)}$. 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 $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, the tensor logarithm Schatten- p (t-log- S_p) low-rank measure of \mathcal{X} is defined as

$$\mathcal{M}_{\log, S_p}(\mathcal{X}) = \frac{1}{n_3} \sum_{i=1}^{n_3} \mathcal{M}_{\log, S_p}(\bar{\mathbf{X}}^{(i)}) = \frac{1}{n_3} \sum_{i=1}^{n_3} \sum_{j=1}^{\min\{n_1, n_2\}} \log(\sigma_j^p(\bar{\mathbf{X}}^{(i)}) + \varepsilon), \quad (4)$$

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 $\mathbf{X} \in \mathbb{R}^{m \times n}$, if $\mathbf{U} \in \mathbb{R}^{m \times m}$ and $\mathbf{V} \in \mathbb{R}^{n \times n}$ are orthogonal matrices, then

$$\mathcal{M}_{\log, S_p}(\mathbf{X}) = \mathcal{M}_{\log, S_p}(\mathbf{UXV}^T).$$

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

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

Proof. (i) It immediately follows from $\sigma_j(\mathbf{X}) = \sigma_j(\mathbf{UXV}^T)$, $j = 1, 2, \dots, \min\{m, n\}$.

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

As shown in proposition 2.10, 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 $\mathcal{Y} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$, the t-log- S_p minimization problem is written as

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

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 $\bar{\mathcal{X}}$. Then solving the t-log- S_p minimization problem (5) is equivalent to solving for each frontal slice $\bar{\mathbf{X}}^{(i)}$ via the following problem

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

In section 3.1 we will propose an IRSpM algorithm for the log- S_p minimization problem as in (6), and conduct in section 3.2 a convergence analysis for IRSpM algorithm. Then in section 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

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

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

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

where $g: [0, \infty) \rightarrow \mathbb{R}$ 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.

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

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

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

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

Before solving equation (8), we recall some notations on the singular value decomposition (SVD). Given a vector $\mathbf{x} \in \mathbb{R}^l$, let $\operatorname{Diag}(\mathbf{x})$ denote the $l \times l$ diagonal matrix with

the j th diagonal element as x_j . Given a matrix $\mathbf{X} \in \mathbb{R}^{m \times n}$, the SVD of \mathbf{X} is computed as $\mathbf{X} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$, where $\mathbf{U} \in \mathbb{R}^{m \times l}$ and $\mathbf{V} \in \mathbb{R}^{n \times l}$ are orthogonal matrices with $\mathbf{U}^T\mathbf{U} = \mathbf{V}^T\mathbf{V} = \mathbf{I}$, and $\mathbf{\Sigma} \in \mathbb{R}^{l \times l}$ is a diagonal matrix, $l = \min\{m, n\}$. In particular, $\mathbf{\Sigma} = \text{Diag}(\boldsymbol{\sigma}(\mathbf{X}))$, where $\boldsymbol{\sigma}(\mathbf{X}) := [\sigma_1(\mathbf{X}), \sigma_2(\mathbf{X}), \dots, \sigma_l(\mathbf{X})]^T$ and $\sigma_j(\mathbf{X})$ is the j th largest singular value of \mathbf{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:

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

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 $T_p^{GST}(\sigma, w) = \text{sign}(\sigma)S_p^{GST}(|\sigma|, w)$ and $S_p^{GST}(|\sigma|, w)$ can be obtain by solving

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

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

- 1: $\tau_p^{GST}(w) = (2w(1-p))^{\frac{1}{2-p}} + wp(2w(1-p))^{\frac{p-1}{2-p}}$;
- 2: **if** $|\sigma| \leq \tau_p^{GST}(w)$ **then**
- 3: $T_p^{GST}(\sigma, w) = 0$;
- 4: **else**
- 5: $k = 0, \delta^k = |\sigma|$;
- 6: **for** $k = 0, 1, \dots, J$ **do**
- 7: $\delta^{k+1} = |\sigma| - wp(\delta^k)^{p-1}$;
- 8: $k \leftarrow k + 1$;
- 9: **end for**
- 10: $T_p^{GST}(\sigma, w) = \text{sign}(\sigma)\delta^k$.
- 11: **end if**

Output: $T_p^{GST}(\sigma, w)$

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

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

is given by

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

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

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

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

We summarize the IRSpM algorithm in algorithm 2.

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

Input: \mathbf{Y} and parameter τ

- 1: **Initialize** \mathbf{X}^0
 - 2: **Set** $k = 0$, $\mu > 1$ and $\mathbf{w}_j^0 = \frac{1}{\sigma_j(\mathbf{X}^0)^p + \varepsilon}$
 - 3: **while** stopping criterion is not satisfied **do**
 - 4: **Compute the SVD of** $\mathbf{X}^k - \frac{1}{\mu} (\mathbf{X}^k - \mathbf{Y})$, i.e. $\mathbf{X}^k - \frac{1}{\mu} (\mathbf{X}^k - \mathbf{Y}) = \mathbf{U}^{k+1} \boldsymbol{\Sigma}^{k+1} (\mathbf{V}^{k+1})^T$
 - 5: **for** $j = 1, 2, \dots, l$
 - 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: $\mathbf{X}^{k+1} = \mathbf{U}^{k+1} \text{Diag}(\boldsymbol{\gamma}^{k+1}) (\mathbf{V}^{k+1})^T$
 - 10: $k \leftarrow k + 1$.
 - 11: **end while**
- Output:** \mathbf{X}^k
-

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

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

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 $\boldsymbol{\gamma}^0 = \boldsymbol{\sigma}(\mathbf{Y})$ and for each j

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

3.2. Convergence analysis of the IRSpM algorithm

We can prove that any accumulation point of the sequence $\{\mathbf{X}^k\}$ 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 $f: \mathbb{R}^d \rightarrow (-\infty, +\infty]$ be a proper and lower semicontinuous function.

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

$$\liminf_{\mathbf{y} \neq \mathbf{x}, \mathbf{y} \rightarrow \mathbf{x}} \frac{f(\mathbf{y}) - f(\mathbf{x}) - \langle \mathbf{u}, \mathbf{y} - \mathbf{x} \rangle}{\|\mathbf{y} - \mathbf{x}\|} \geq 0.$$

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

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

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

Definition 3.5. A function $f: \mathbb{R}^n \rightarrow \mathbb{R}$ 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 $F: \mathbb{R}^{m \times n} \rightarrow \mathbb{R}$ is a singular value function if $F(\mathbf{X}) = (f \circ \sigma)(\mathbf{X})$, where $f: \mathbb{R}^l \rightarrow \mathbb{R}$ 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 $f \circ \sigma$ at a matrix \mathbf{X} is given by the formula

$$\partial(f \circ \sigma)(\mathbf{X}) = \mathbf{U} \text{Diag}(\partial f[\sigma(\mathbf{X})]) \mathbf{V}^T$$

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

The $\log\text{-}S_p$ 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 $\log\text{-}S_p$ 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 $\log\text{-}S_p$ minimization problem (7) using

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

That is,

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

Then substituting (15) into (14) yields (13). \square

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

- (i) The sequence $\{\mathbf{X}^k\}$ is bounded.
- (ii) $\lim_{k \rightarrow \infty} \|\mathbf{X}^{k+1} - \mathbf{X}^k\|_F = 0$.
- (iii) Any accumulation point of $\{\mathbf{X}^k\}$ is a stationary point of Ψ .

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

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

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

Taking $K \rightarrow \infty$, we have

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

This yields assertion (ii).

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

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

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

Denote $\tilde{\boldsymbol{\gamma}}^{k_i+1} := [\gamma_1^{k_i+1}, \gamma_2^{k_i+1}, \dots, \gamma_r^{k_i+1}]^T$ and denote $\mathbf{d}^{k_i+1} := [d_1^{k_i+1}, d_2^{k_i+1}, \dots, d_r^{k_i+1}]^T$ 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 $\tilde{\mathbf{U}}^{k_i+1}$ denote the $m \times r$ matrix formed by the first r columns of \mathbf{U}^{k_i+1} and let $\tilde{\mathbf{V}}^{k_i+1}$ denote the $n \times r$ matrix formed by the first r columns of \mathbf{V}^{k_i+1} . Then from equation (16), we obtain

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

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

$$\begin{aligned} & \mu(\mathbf{X}^{k_i+1} - \mathbf{X}^{k_i}) + (\mathbf{X}^{k_i} - \mathbf{Y}) + \mu \sum_{j=r+1}^l \left(\boldsymbol{\Sigma}_{jj}^{k_i+1} - \gamma_j^{k_i+1}\right) \mathbf{U}_j^{k_i+1} \left(\mathbf{V}_j^{k_i+1}\right)^T \\ & + \tau p \tilde{\mathbf{U}}^{k_i+1} \text{Diag}(\mathbf{d}^{k_i+1})\left(\tilde{\mathbf{V}}^{k_i+1}\right)^T = \mathbf{0}. \end{aligned}$$

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

$$\mu \left(\tilde{\mathbf{U}}^{k_i+1}\right)^T (\mathbf{X}^{k_i+1} - \mathbf{X}^{k_i}) \tilde{\mathbf{V}}^{k_i+1} + \left(\tilde{\mathbf{U}}^{k_i+1}\right)^T (\mathbf{X}^{k_i} - \mathbf{Y}) \tilde{\mathbf{V}}^{k_i+1} + \tau p \text{Diag}(\mathbf{d}^{k_i+1}) = \mathbf{0}. \quad (17)$$

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

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

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

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 $\bar{\mathbf{X}}^{(i)}$. 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: \mathcal{Y} and parameter τ

1: **Set initialization** denoted as \mathcal{Z}

2: $\bar{\mathcal{Y}} = \text{fft}(\mathcal{Y}, [], 3)$

3: $\bar{\mathcal{Z}} = \text{fft}(\mathcal{Z}, [], 3)$

4: **for** $i = 1, 2, \dots, n_3$ **do**

5: $\bar{\mathbf{X}}^{(i)} = \text{IRSpM}(\bar{\mathbf{Y}}^{(i)}, \tau)$ initialized with $\bar{\mathbf{Z}}^{(i)}$.

6: **end for**

7: $\mathcal{X} = \text{ifft}(\bar{\mathcal{X}}, [], 3)$

Output: \mathcal{X}

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

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

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

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

- (i) $\Phi(\mathcal{X}_k) - \Phi(\mathcal{X}_{k+1}) \geq \frac{\mu-1}{2} \|\mathcal{X}_{k+1} - \mathcal{X}_k\|_F^2$, where $\mu > 1$ is an IRS p M algorithm parameter.
- (ii) The sequence $\{\mathcal{X}_k\}$ is bounded.
- (iii) $\lim_{k \rightarrow \infty} \|\mathcal{X}_{k+1} - \mathcal{X}_k\|_F = 0$.
- (iv) Any accumulation point of $\{\mathcal{X}_k\}$ is a stationary point of Φ .

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

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 $\mathbf{x} \in \mathbb{R}^N$, the image degradation model under Poisson noise can be written as

$$\mathbf{y} = P(\mathbf{H}\mathbf{x}), \quad (19)$$

where \mathbf{x} denotes an image that is not degraded, $\mathbf{H} \in \mathbb{R}^{n \times n}$ 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 \mathbf{y} denotes a degraded image. If \mathbf{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 \mathbf{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

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

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

$$\min_{\mathbf{x}} \tau \langle \mathbf{H}\mathbf{x} - \mathbf{y} \log \mathbf{H}\mathbf{x}, \mathbf{1} \rangle + \|\nabla \mathbf{x}\|_1, \quad (20)$$

where $\mathbf{1}$ denotes the vector whose entries are all ones, the logarithm and multiplication with \mathbf{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 \mathbf{x} , suppose it can be divided into L overlapping cubes of size $\sqrt{n_1} \times \sqrt{n_1} \times n_3$, denoted as $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_L\}$. For each reference cube \mathbf{x}_i of the image, a total number of n_2 non-local self-similar cubes $\{\mathbf{x}_{i,1}, \mathbf{x}_{i,2}, \dots, \mathbf{x}_{i,n_2}\}$ can be found by cube matching. Here, the cubes are grouped using Euclidean distances, and the tensor $\mathcal{R}_i(\mathbf{x})$ is generated for the reference cube \mathbf{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 $\mathbf{x} \in \mathbb{R}^N$ and a reference vectorized cube $\mathbf{x}_i \in \mathbb{R}^{n_1 n_3}$, the non-local similar patch matrix $\mathbf{R}_{i,j} \in \mathbb{R}^{n_1 n_3 \times N}$ is a binary matrix (whose terms are 1 or 0), $i = 1, 2, \dots, L, j = 1, 2, \dots, n_2$, such that $\mathbf{R}_{i,j} \mathbf{x}$ is the j th vectorized cube in the i th non-local similar group $\mathbf{x}_{i,j}$, that is, $\mathbf{R}_{i,j} \mathbf{x} = \mathbf{x}_{i,j}$. Let $\mathcal{R}_i : \mathbb{R}^N \rightarrow \mathbb{R}^{n_1 \times n_2 \times n_3}$ be the extraction operator for the i th non-local self-similar tensor defined as

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

Here, $\mathcal{R}_i(\mathbf{x})$ 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 $\mathcal{R}_i(\mathbf{x})$ should be a low-rank tensor according to non-local self-similarity if \mathbf{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

$$\min_{\mathbf{x}} \tau \langle \mathbf{H}\mathbf{x} - \mathbf{y} \log \mathbf{H}\mathbf{x}, \mathbf{1} \rangle_W + \sum_{i=1}^L \eta_i \mathcal{M}_{\log, S_p}(\mathcal{R}_i(\mathbf{x})), \quad (21)$$

where $\mathcal{R}_i(\mathbf{x})$ represents the constructed tensor for each reference cube, $\mathbf{W} = \sum_{i=1}^L \mathcal{R}_i^T \circ \mathcal{R}_i = \sum_{i=1}^L \sum_{l=1}^{n_2} \mathbf{R}_{i,l}^T \mathbf{R}_{i,l}$ is a diagonal matrix whose main diagonal entries indicate the counts for each pixel, $\langle \cdot, \cdot \rangle$ is the \mathbf{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:

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Since the update of \mathbf{h} has a closed-form solution, the algorithm can be viewed as alternatively updating variables \mathcal{L}_i and \mathbf{x} . We call this algorithm as the patch-based tensor logarithmic S_p minimization (TLSpM) 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: \mathbf{y} , and parameters τ, α, η_i , and β .

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

4.4. Convergence analysis of the Patch-based TLSpM algorithm

We can prove that any accumulation point of the sequence $\{(\mathbf{x}^*, \mathbf{h}^*, \{\mathcal{L}_i^*\})\}$, where $\{\mathcal{L}_i^*\}$ denotes $\{\mathcal{L}_1^*, \mathcal{L}_2^*, \dots, \mathcal{L}_L^*\}$, 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 \mathcal{L}_i from the initial \mathcal{L}_i^{qk} to the output $\mathcal{L}_i^{q(k+1)}$, which are corresponding to the initial \mathcal{L}_i^k to the output \mathcal{L}_i^{k+1} in line 4.

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

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

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

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

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

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

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

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

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

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

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

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

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

By the updates of \mathbf{h}^{k+1} and \mathbf{x}^{k+1} , we have $\Phi(\mathbf{x}^k, \mathbf{h}^{k+1}, \{\mathcal{L}_i^{q(k+1)}\}) - \Phi(\mathbf{x}^k, \mathbf{h}^k, \{\mathcal{L}_i^{q(k+1)}\}) \geq 0$ and $\Phi(\mathbf{x}^{k+1}, \mathbf{h}^{k+1}, \{\mathcal{L}_i^{q(k+1)}\}) - \Phi(\mathbf{x}^k, \mathbf{h}^{k+1}, \{\mathcal{L}_i^{q(k+1)}\}) \geq \frac{\beta}{2} \|\mathbf{x}^{k+1} - \mathbf{x}^k\|_{\mathbf{W}}^2$. 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

$$\begin{aligned} &\frac{\mu-1}{2} \sum_{i=1}^L \sum_{j=1}^q \|\mathcal{L}_i^{qk+j} - \mathcal{L}_i^{qk+j-1}\|_F^2 + \frac{\beta}{2} \|\mathbf{x}^{k+1} - \mathbf{x}^k\|_{\mathbf{W}}^2 \\ &\leq \Phi(\mathbf{x}^0, \mathbf{h}^0, \{\mathcal{L}_i^{q0}\}) - \Phi(\mathbf{x}^K, \mathbf{h}^K, \{\mathcal{L}_i^{qK}\}) < +\infty. \end{aligned}$$

Taking $K \rightarrow \infty$, we have

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

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

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

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

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

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

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

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’⁷ ($128 \times 128 \times 64$), ‘Micro-tubules’¹ ($128 \times 128 \times 64$), and ‘Pollen’⁸ ($256 \times 256 \times 32$). The 10th frontal slice of each original image is shown in figure 1. To simulate blurry Poissonian images, we adopt the procedure in [16]. 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 generator⁹, including one 3D Gibson & Lanni blur (G&L) [62] and two different 3D Gaussian blur (G1 and G2). Lastly, Poisson noise is added to the blurry image.

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.

⁷ The ‘Spherical-beads’ and ‘Micro-tubules’ images are collected from <http://bigwww.epfl.ch/deconvolution/index.html#data>.

⁸ The ‘Pollen’ image is collected from www.cellimagelibrary.org/images/35532.

⁹ The software package is downloaded from <http://bigwww.epfl.ch/algorithms/psfgenerator/>.



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

Table 1. Parameter settings for patch-based TLSpM algorithm.

Image	Spherical-beads			Micro-tubules			Pollen		
	G&L	G1	G2	G&L	G1	G2	G&L	G1	G2
PSF									
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	35 000	700

In the experiment, the peak signal-to-noise ratio (PSNR) [63] and structural similarity index measure (SSIM) [64] are used to measure the quality of the restored images. In particular, the PSNR value is defined as

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

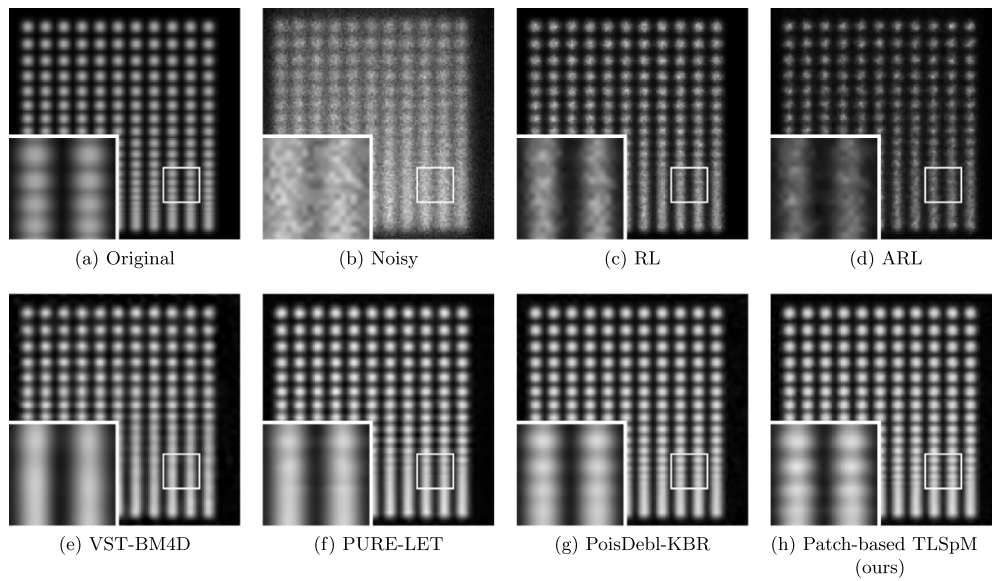
where \mathbf{x}^* is the restored image and \mathbf{x} is the original image. And the SSIM value is defined in [64].

The PSNR and SSIM values of the restored images obtained by different algorithms are shown in table 2. The best PSNR and SSIM values for each case are marked in bold. It shows that the proposed patch-based TLSpM algorithm achieves the best numerical values for most of the testing cases. For example, for ‘Micro-tubules’ 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.

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

Image	Metric	Spherical-beads			Micro-tubules			Pollen		
		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 TLSpM (ours)	PSNR	19.67	19.42	20.97	23.58	23.65	23.87	28.49	27.35	28.40
	SSIM	0.788	0.795	0.846	0.634	0.602	0.634	0.720	0.689	0.718

**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).

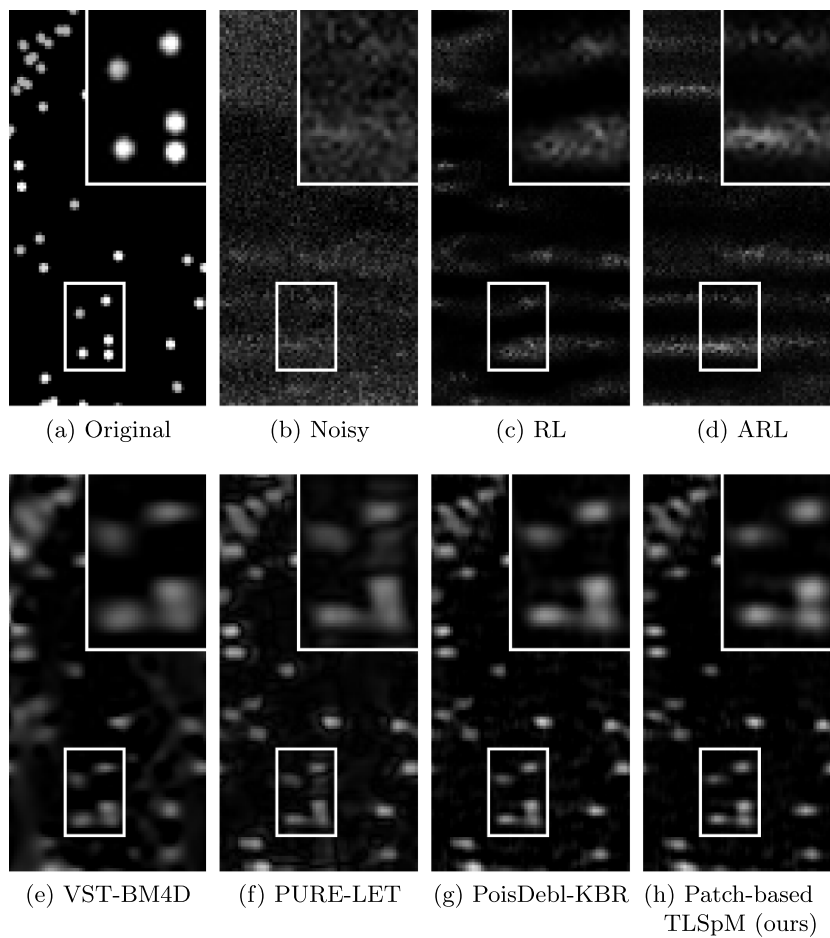


Figure 3. The 33rd lateral slices of the images restored by different algorithms from the noisy ‘Micro-tubules’ image with the G1 blur kernel. The PSNR values of the restored images are: (b) noisy image (19.64 dB); (c) RL (20.95 dB); (d) ARL (19.81 dB); (e) VST-BM4D (22.32 dB); (f) PURE-LET (22.72 dB);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

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.

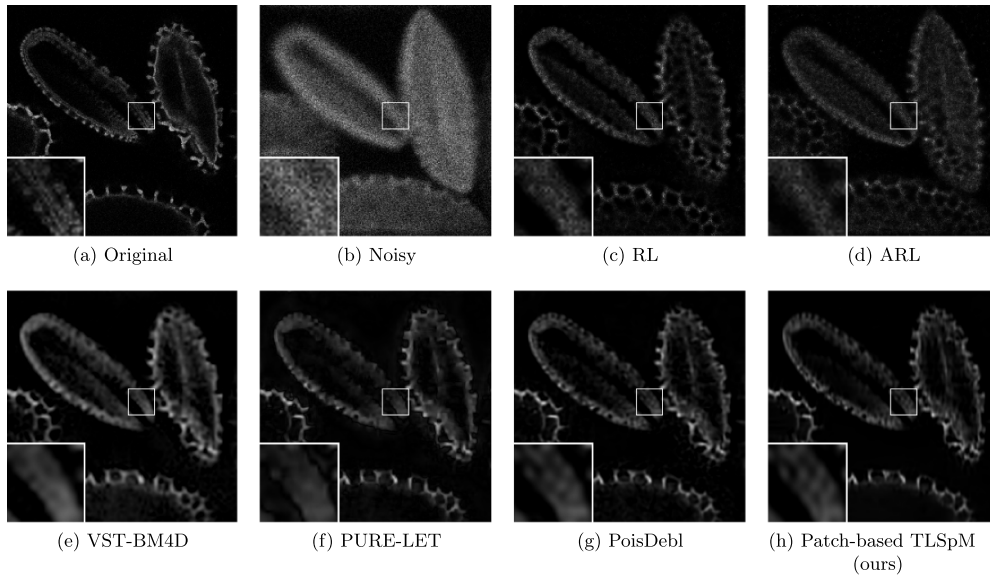


Figure 4. The 12th frontal slices of the images restored by different algorithms from the noisy ‘Pollen’ image with the G2 blur kernel. The PSNR values of the restored images are: (b) noisy image (23.64 dB); (c) RL (25.82 dB); (d) ARL (23.90 dB); (e) VST-BM4D (25.51 dB); (f) PURE-LET (27.59 dB); (g)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

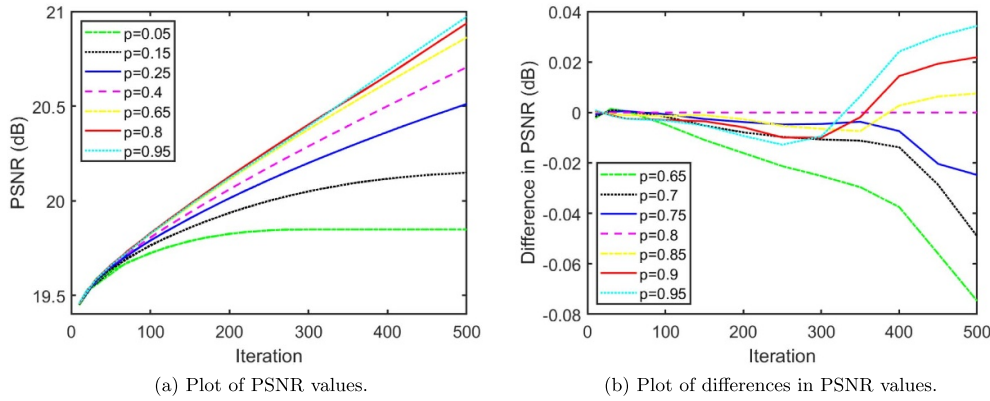


Figure 5. Sensitivity analysis of parameter p . (a) Plot of the PSNR value for $p \in (0, 1)$ vs the number of iterations; (b) The PSNR difference for $p \in [0.65, 1)$ vs the number of iterations.

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 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 $p = 0.95$.

6. Conclusion

In this paper, we first define a new $t\text{-log-}S_p$ 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.

Data availability statement

The data cannot be made publicly available upon publication because they are owned by a third party and the terms of use prevent public distribution. The data that support the findings of this study are available upon reasonable request from the authors.

Acknowledgments

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 12326619, by the Natural Science Foundation of Guangdong Province of China under Grants 2020B1515310008, 2023A1515011691 and 2024A1515011913, by the Educational Commission of Guangdong Province of China under Grant 2019KZDZX1007, by the Shenzhen Basis Research Project under Grant JCYJ20180305125521534, and by the PolyU internal Grant P0040271.

ORCID iDs

Jian Lu  <https://orcid.org/0000-0003-4599-7281>

Xiaoxia Liu  <https://orcid.org/0000-0002-2289-7520>

References

- [1] Zhang Y, Zhu Y, Nichols E, Wang Q, Zhang S, Smith C and Howard S 2019 A Poisson-Gaussian denoising dataset with real fluorescence microscopy images *2019 IEEE/CVF Conf. on Computer Vision and Pattern Recognition (CVPR)* pp 11702–10
- [2] Anisimova E, Bednar J and Pata P 2013 Astronomical image denoising using curvelet and starlet transform *2013 23rd Int. Conf. Radioelektronika (RADIOELEKTRONIKA)* pp 255–60
- [3] Panin V Y, Zeng G L and Gullberg G T 1999 Total variation regulated EM algorithm [SPECT reconstruction] *IEEE Trans. Nucl. Sci.* **46** 2202–10

- [4] Zheng W, Li S, Andrzej Krol C R S, Zeng X and Yuesheng X 2019 Sparsity promoting regularization for effective noise suppression in SPECT image reconstruction *Inverse Problems* **35** 115011
- [5] Shepp L A and Vardi Y 1982 Maximum likelihood reconstruction for emission tomography *IEEE Trans. Med. Imaging* **1** 113–22
- [6] Savage N 2010 Medical imagers lower the dose *IEEE Spectr.* **47** 14–16
- [7] Renliang G and Dogandžić A 2016 Blind x-ray CT image reconstruction from polychromatic Poisson measurements *IEEE Trans. Comput. Imaging* **2** 150–65
- [8] Hamed G, Marey M, El-Sayed Amin S and Tolba M F 2018 A proposed model for denoising breast mammogram images *2018 13th Int. Conf. on Computer Engineering and Systems (ICCES)* pp 652–7
- [9] Pei C and Barner. K 2003 Maximum likelihood reconstruction for tomosynthesis *2003 IEEE 29th Annual Proc. Bioengineering Conf.* pp 59–60
- [10] William H R 1972 Bayesian-based iterative method of image restoration *J. Opt. Soc. Am.* **62** 55–59
- [11] Al-Ameen Z 2018 Faster deblurring for digital images using an ameliorated Richardson-Lucy algorithm *IEIE Trans. Smart Process. Comput.* **7** 289–95
- [12] Dey N, Blanc-Feraud L, Zimmer C, Roux P, Kam Z, Olivo-Marin J-C and Zerubia J 2006 Richardson-Lucy algorithm with total variation regularization for 3D confocal microscope deconvolution *Microscopy Res. Tech.* **69** 260–6
- [13] Harmany Z T, Marcia R F and Willett R M 2012 This is spiral-tap: sparse Poisson intensity reconstruction algorithms-theory and practice *IEEE Trans. Image Process.* **21** 1084–96
- [14] Bonettini S and Ruggiero V 2011 An alternating extragradient method for total variation-based image restoration from Poisson data *Inverse Problems* **27** 095001
- [15] Figueiredo M A T and Bioucas-Dias J M 2010 Restoration of Poissonian images using alternating direction optimization *IEEE Trans. Image Process.* **19** 3133–45
- [16] Liyan M, Moisan L, Jian Y and Zeng T 2013 A dictionary learning approach for Poisson image deblurring *IEEE Trans. Med. Imaging* **32** 1277–89
- [17] Chen D-Q and Cheng Li-Z 2011 Spatially adapted regularization parameter selection based on the local discrepancy function for Poissonian image deblurring *Inverse Problems* **28** 015004
- [18] Zhang T, Chen J, Caiying W, Zhifei H, Zeng T and Jin Q 2022 Edge adaptive hybrid regularization model for image deblurring *Inverse Problems* **38** 065010
- [19] Setzer S, Steidl G and Teuber T 2010 Deblurring Poissonian images by split Bregman techniques *J. Visual Commun. Image Represent.* **21** 193–9
- [20] Chen D-Q 2014 Regularized generalized inverse accelerating linearized alternating minimization algorithm for frame-based Poissonian image deblurring *SIAM J. Imaging Sci.* **7** 716–39
- [21] Vonesch C and Unser M 2009 A fast multilevel algorithm for wavelet-regularized image restoration *IEEE Trans. Image Process.* **18** 509–23
- [22] Carlván M and Blanc-Feraud L 2012 Sparse Poisson noisy image deblurring *IEEE Trans. Image Process.* **21** 1834–46
- [23] Zhang H, Dong Y and Fan Q 2017 Wavelet frame based Poisson noise removal and image deblurring *Signal Process.* **137** 363–72
- [24] Jeong T, Woo H and Yun S 2013 Frame-based Poisson image restoration using a proximal linearized alternating direction method *Inverse Problems* **29** 075007
- [25] Liu J, Lou Y, Guoxi N and Zeng T 2020 An image sharpening operator combined with framelet for image deblurring *Inverse Problems* **36** 045015
- [26] Lefkimmiatis S and Unser M 2013 Poisson image reconstruction with Hessian Schatten-norm regularization *IEEE Trans. Image Process.* **22** 4314–27
- [27] Dupe F-X, Fadili J M and Starck J-L 2009 A proximal iteration for deconvolving Poisson noisy images using sparse representations *IEEE Trans. Image Process.* **18** 310–21
- [28] Anscombe F J 1957 The transformation of Poisson, binomial and negative-binomial data *Biometrika* **35** 246–54
- [29] Azzari L and Foi A 2017 Variance stabilization in Poisson image deblurring *2017 IEEE 14th Int. Symp. on Biomedical Imaging (ISBI 2017)* pp 728–31
- [30] Dabov K, Foi A, Katkovnik V and Egiazarian K 2006 Image denoising with block-matching and 3D filtering *Image Processing: Algorithms and Systems, Neural Networks and Machine Learning* pp 354–65
- [31] Maggioni M, Katkovnik V, Egiazarian K and Foi A 2013 Nonlocal transform-domain filter for volumetric data denoising and reconstruction *IEEE Trans. Image Process.*

- [32] Jizhou Li, Luisier F and Blu T 2018 Pure-let image deconvolution *IEEE Trans. Image Process.* **27** 92–105
- [33] Jizhou Li, Luisier F and Blu T 2017 Pure-let deconvolution of 3D fluorescence microscopy images *2017 IEEE 14th Int. Symp. on Biomedical Imaging (ISBI 2017)* pp 723–7
- [34] Xie Y, Shuhang G, Liu Y, Zuo W, Zhang W and Zhang L 2016 Weighted Schatten p -norm minimization for image denoising and background subtraction *IEEE Trans. Image Process.* **25** 4842–57
- [35] Kilmer M E and Martin C D 2011 Factorization strategies for third-order tensors *Linear Algebr. Appl.* **435** 641–58
- [36] Kilmer M E, Braman K, Hao N and Hoover R C 2013 Third-order tensors as operators on matrices: a theoretical and computational framework with applications in imaging *SIAM J. Matrix Anal. Appl.* **34** 148–72
- [37] Canyi L, Feng J, Chen Y, Liu W, Lin Z and Yan S 2019 Tensor robust principal component analysis with a new tensor nuclear norm *IEEE Trans. Pattern Anal. Mach. Intell.* **42** 925–38
- [38] Liu C, Shan H and Chen C 2020 Tensor p -shrinkage nuclear norm for low-rank tensor completion *Neurocomputing* **387** 255–67
- [39] Gao Q, Zhang P, Xia W, Xie D 2021 Enhanced tensor RPCA and its application *IEEE Trans. Pattern Anal. Mach. Intell.* **43** 2133–40
- [40] Zheng Y-B, Huang T-Z, Zhao Xi-L, Jiang T-X, Tian-Hui M and Teng-Yu J 2020 Mixed noise removal in hyperspectral image via low-fibered-rank regularization *IEEE Trans. Geosci. Remote Sens.* **58** 734–49
- [41] Fazel M, Hindi H and Boyd S P 2003 Log-det heuristic for matrix rank minimization with applications to Hankel and Euclidean distance matrices *Proc. American Control Conf.* vol 3 pp 2156–62
- [42] Zhaosong L, Zhang Y and Jian L 2017 ℓ_p Regularized low-rank approximation via iterative reweighted singular value minimization *Comput. Optim. Appl.* **68** 619–42
- [43] Zuo W, Meng D, Zhang L, Feng X and Zhang D 2013 A generalized iterated shrinkage algorithm for non-convex sparse coding *2013 IEEE Int. Conf. on Computer Vision* pp 217–24
- [44] Lewis A S and Sendov H S 2005 Nonsmooth analysis of singular values. part I: theory *Set-Valued Anal.* **13** 213–41
- [45] Ortega J M and Rheinboldt W C 1970 *Iterative Solution of Nonlinear Equations in Several Variables* (Academic)
- [46] Triet L, Chartrand R and Asaki T J 2007 A variational approach to reconstructing images corrupted by Poisson noise *J. Math. Imaging Vis.* **27** 257–63
- [47] Rudin L I, Osher S and Fatemi E 1992 Nonlinear total variation based noise removal algorithms *Physica D* **60** 259–68
- [48] Chen X, Liu X, Zheng J, Shen L, Jiang Q and Jian L 2021 Nonlocal low-rank regularized two-phase approach for mixed noise removal *Inverse Problems* **37** 085001
- [49] Liu X, Jian L, Shen L, Chen X and Yuesheng X 2020 Multiplicative noise removal: Nonlocal low-rank model and its proximal alternating reweighted minimization algorithm *SIAM J. Imag. Sci.* **13** 1595–629
- [50] Niu S, Gaochang Y, Jianhua M and Wang J 2018 Nonlocal low-rank and sparse matrix decomposition for spectral CT reconstruction *Inverse Problems* **34** 024003
- [51] Zhi Li, Yan M, Zeng T and Zhang G 2022 Phase retrieval from incomplete data via weighted nuclear norm minimization *Pattern Recognit.* **125** 108537
- [52] Shuhang G, Zhang L, Zuo W and Feng X 2014 Weighted nuclear norm minimization with application to image denoising *Proc. IEEE Conf. on Computer Vision and Pattern Recognition* pp 2862–9
- [53] Liyan M, Xu Li and Zeng T 2017 Low rank prior and total variation regularization for image deblurring *J. Sci. Comput.* **70** 1336–57
- [54] Jian L, Tian J, Jiang Q, Liu X and Zou Y 2021 Rician noise removal via weighted nuclear norm penalization *Appl. Comput. Harmon. Anal.* **53** 180–98
- [55] Jian L, Chen X, Zhenwei H, Liu X, Jiang Q, Meng D and Lin Z 2022 A new nonlocal low-rank regularization method with applications to magnetic resonance image denoising *Inverse Problems* **38** 065012
- [56] Qiu D, Bai M, Michael K N and Zhang X 2021 Nonlocal robust tensor recovery with nonconvex regularization *Inverse Problems* **37** 035001
- [57] Xue J, Zhao Y, Liao W and Cheung-Wai Chan J 2019 Nonlocal low-rank regularized tensor decomposition for hyperspectral image denoising *IEEE Trans. Geosci. Remote Sens.* **57** 5174–89

- [58] Xie Q, Zhao Q, Meng D and Zongben X 2017 Kronecker-basis-representation based tensor sparsity and its applications to tensor recovery *IEEE Trans. Pattern Anal. Mach. Intell.* **40** 1888–902
- [59] Zeng D, Xie Q, Cao W, Lin J, Zhang H, Zhang S, Huang J, Bian Z, Meng D and Zongben X 2017 Low-dose dynamic cerebral perfusion computed tomography reconstruction via Kronecker-basis-representation tensor sparsity regularization *IEEE Trans. Med. Imaging* **36** 2546–56
- [60] Jian L, Huaxuan H, Zou Y, Zhaosong L, Liu X, Keke Z and Lin Li 2024 A nonlocal Kronecker-basis-representation method for low-dose CT sinogram recovery *J. Comput. Math.* **42** 1080–108
- [61] Xie Q, Zhao Q, Meng D and Zongben X 2017 Kronecker-basis-representation based tensor sparsity and its applications to tensor recovery *IEEE Trans. Pattern Anal. Mach. Intell.* **40** 1888–902
- [62] Frisken Gibson S and Lanni F 1992 Experimental test of an analytical model of aberration in an oil-immersion objective lens used in three-dimensional light microscopy *J. Opt. Soc. Am. A* **9** 154–66
- [63] Alan C B 2010 *Handbook of Image and Video Processing* (Academic)
- [64] Wang Z, Bovik A C, Sheikh H R and Simoncelli E P 2004 Image quality assessment: from error visibility to structural similarity *IEEE Trans. Image Process.* **13** 600–12