



OPEN ACCESS

EDITED BY

Yannan Chen,
South China Normal University, China

REVIEWED BY

Douglas Soares Goncalves,
Federal University of Santa Catarina, Brazil
Can Li,
Honghe University, China
Xueli Bai,
Guangdong University of Foreign Studies,
China

*CORRESPONDENCE

Jinjie Liu
✉ jinjie.liu@cqnu.edu.cn

RECEIVED 04 March 2025

ACCEPTED 07 April 2025

PUBLISHED 06 May 2025

CITATION

Duan Y, Ling C, Liu J and Yang X (2025)
Internet traffic data recovery via a low-rank
spatio-temporal regularized optimization
approach without d -th order T-SVD.
Front. Appl. Math. Stat. 11:1587681.
doi: 10.3389/fams.2025.1587681

COPYRIGHT

© 2025 Duan, Ling, Liu and Yang. This is an
open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](#). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic practice.
No use, distribution or reproduction is
permitted which does not comply with these
terms.

Internet traffic data recovery via a low-rank spatio-temporal regularized optimization approach without d -th order T-SVD

Yuxuan Duan¹, Chen Ling², Jinjie Liu^{1*} and Xinmin Yang¹

¹National Center for Applied Mathematics in Chongqing, Chongqing Normal University, Chongqing, China, ²Department of Mathematics, Hangzhou Dianzi University, Hangzhou, China

Accurate recovery of Internet traffic data can mitigate the adverse impact of incomplete data on network task processes. In this study, we propose a low-rank recovery model for incomplete Internet traffic data with a fourth-order tensor structure, incorporating spatio-temporal regularization while avoiding the use of d -th order T-SVD. Based on d -th order tensor product, we first establish the equivalence between d -th order tensor nuclear norm and the minimum sum of the squared Frobenius norms of two factor tensors under the unitary transformation domain. This equivalence allows us to leave aside the d -th order T-SVD, significantly reducing the computational complexity of solving the problem. In addition, we integrate the alternating direction method of multipliers (ADMM) to design an efficient and stable algorithm for precise model solving. Finally, we validate the proposed approach by simulating scenarios with random and structured missing data on two real-world Internet traffic datasets. Experimental results demonstrate that our method exhibits significant advantages in data recovery performance compared to existing methods.

KEYWORDS

Internet traffic data recovery, d -th order TNN, spatio-temporal regularization, ADMM algorithm, tensor completion

1 Introduction

Internet traffic data, characterized by distinct spatio-temporal attributes, are crucial for documenting information transmission volumes over specified time periods. It plays a vital role in network design and management [1–3]. However, in practice, uncontrollable factors often result in incomplete or corrupted traffic data, hindering its effective use. Therefore, accurately recovering original data from incomplete traffic data is of significant importance.

Since Zhang et al. [4] introduced the sparse regularized matrix factorization (SRMF) model for data recovery, while incorporating the unique characteristics of Internet traffic data, research in the field of Internet traffic data recovery has continued to advance. As higher-order generalizations of matrices, tensors are more effective at capturing the structural characteristics within the data. Consequently, many researchers have proposed various effective methods for Internet traffic data recovery based on different tensor decomposition techniques, such as the CANDECOMP/PARAFAC (CP) decomposition

[5, 6], the tensor-train (TT) decomposition [7], and the third-order T-product factorization [8]. Given that low-rankness is a typical characteristic of incomplete data in scenarios with high levels of missing samples, low-rank modeling serves as another effective strategy beyond tensor decomposition. Candès et al. [9] were the pioneers in establishing a low-rank matrix data recovery model. To address its inherent non-convexity and NP hardness, they employed the nuclear norm of a matrix to convexly relax the model, enabling an effective solution. Unlike the definition of matrix rank, which is well-defined, the definition of tensor rank relies on specific decomposition method.

In 2014, based on the tensor singular value decomposition (T-SVD), Zhang et al. [10] proposed third-order tensor nuclear norm (TNN) and employed it to convexly relax the tubal rank of tensors, and they established a model for image restoration. On this basis, Li et al. [11] developed the SRTNN model for the recovery of Internet traffic data by incorporating spatial-temporal regularization with the nuclear norm of third-order tensors, which demonstrates significant advantages in terms of data recovery effectiveness. Meanwhile, to mitigate the computational burden associated with SVD, He et al. [12] proposed the equivalence between the so-called transformed tubal nuclear norm for a third-order tensor and the minimum of the sum of two factor tensors' squared Frobenius norms under a general invertible linear transform.

Currently, the recovery of Internet traffic data primarily relies on model construction based on third-order tensors. While these studies have achieved certain advancements, the effectiveness in restoring data with structural missingness remains limited. Addressing this challenge requires data recovery models to more thoroughly account for the intrinsic structural characteristics of the data. In the case of Internet traffic data, considering its complexity and multidimensionality, it can be organized in the form of a fourth-order tensor ($\mathcal{X} \in \mathbb{R}^{D \times T \times N \times N}$, where N signifies network nodes, T is the daily recording times, and D indicates the recording days) to capture the internal structural features of the data more comprehensively. Therefore, exploring a low-rank fourth-order tensor model for Internet traffic data recovery is essential.

Notably, Martin et al. [13] have extended the third order T-product to d -th order T-product, which is explained in a recursive way but for computational speed is implemented directly using the fast Fourier transform. Inspired by this, Qin et al. [14] have generalized the multiplication in the discrete Fourier transform (DFT) domain to tensor product in the domain of general invertible transforms and introduced the d -th order tensor nuclear norm. They employed the d -th order TNN to perform convex relaxation on the fourth order tensor low-rank model, establishing the HTNN method for recovering visual data. However, for solving optimization problems involving nuclear norms, the thresholding operator is often employed, which involves performing SVD on a large number of matrices during the process, significantly increasing the computational time [11, 14]. Hence, a natural question arises about how to build a low-rank recovery model for a d -th order tensor without the need for SVD computation.

In this study, by arranging the observed Internet traffic data as a tensor $\mathcal{M} \in \mathbb{R}^{D \times T \times N \times N}$, we enable unitary transformations on mode-3 and mode-4 to fully integrate the internal data of these modes. Meanwhile, a spatio-temporal regularization term is applied

to modes-1 and-2, aiming to preserve the periodicity and similarity of the recovered data as much as possible, thereby enhancing the accuracy of data recovery. We develop a new low-rank model for the recovery of fourth-order Internet traffic data as follows:

$$\begin{aligned} \min_{\mathcal{X}} \quad & \|\mathcal{X}\|_{*,L} + \frac{\alpha_1}{2} \|H \cdot \text{Mat}_1(\mathcal{X})\|_F^2 + \frac{\alpha_2}{2} \|K \cdot \text{Mat}_2(\mathcal{X})\|_F^2 + \gamma \|\mathcal{X}\|_F^2 \\ \text{s.t.} \quad & P_{\Omega}(\mathcal{X}) = P_{\Omega}(\mathcal{M}), \end{aligned}$$

where $\|\cdot\|_{*,L}$ means the tensor nuclear norm under the unitary transform L , and $\text{Mat}_1(\mathcal{X})$ and $\text{Mat}_2(\mathcal{X})$ are consisted by flow data of consecutive days and time points, respectively. Through the equivalence relationship between the d -th order tensor nuclear norm and the minimum sum of two factor tensors' squared Frobenius norms within the unitary transformation domain and under the d -th order T-product which is first established by us, we reformulate this model into the following version without relying on d -th order T-SVD:

$$\begin{aligned} \min_{\mathcal{X}, \mathcal{U}, \mathcal{V}} \quad & \frac{1}{2} (\|\mathcal{U}\|_F^2 + \|\mathcal{V}\|_F^2) + \frac{\lambda}{2} \|\mathcal{X} - \mathcal{U} *_{*L} \mathcal{V}\|_F^2 + \gamma \|\mathcal{X}\|_F^2 \\ & + \frac{\alpha_1}{2} \|H \cdot \text{Mat}_1(\mathcal{X})\|_F^2 + \frac{\alpha_2}{2} \|K \cdot \text{Mat}_2(\mathcal{X})\|_F^2 \\ \text{s.t.} \quad & P_{\Omega}(\mathcal{X}) = P_{\Omega}(\mathcal{M}). \end{aligned}$$

We will give a detail explanation about these two models in Section 3.

The rest of this study is organized as follows. Some notions and preliminaries are listed in Section 2. In Section 3, we establish a low-rank Internet traffic data recovery model without d -th order tensor SVD and solve it using the alternating direction method of multipliers (ADMM), proposing a computational framework for solving the model. In Section 4, we conduct a convergence analysis of Algorithm 1, ensuring that it converges to a stationary point. Numerical experiments are presented in Section 5, where we apply the proposed method to two real-world Internet traffic datasets and simulate potential structural missingness scenarios in practice. The experimental results demonstrate that our method exhibits significant advantages, both in terms of recovery accuracy and computational efficiency. Finally, we give the conclusion of this study in Section 6.

2 Preliminary

In this study, the real number field and the complex number field are denoted by $\mathbb{R}$ and $\mathbb{C}$, respectively. We use lowercase letters $a, b, \dots$ to denote scalars, bold lowercase letters $\mathbf{a}, \mathbf{b}, \dots$ to represent vectors, capital letters $A, B, \dots$ to signify matrices, and calligraphic letters $\mathcal{A}, \mathcal{B}, \dots$ to denote tensors. For any positive integer n , we define the set $[n] := \{1, 2, \dots, n\}$. For $A \in \mathbb{C}^{m \times n}$, A^H denotes the conjugate transpose of A (when A is in the real number field, its transpose is A^T), and its nuclear norm is denoted as $\|A\|_* = \sum_{i=1}^{\min\{m,n\}} \sigma_i(A)$, where σ_i is the i -th singular value of A . For a d -th order real tensor $\mathcal{A} = (a_{i_1 i_2 \dots i_d}) \in \mathbb{R}^{n_1 \times n_2 \times \dots \times n_d}$, its corresponding Frobenius norm is $\|\mathcal{A}\|_F = \left(\sum_{i_1=1}^{n_1} \dots \sum_{i_d=1}^{n_d} |a_{i_1 \dots i_d}|^2 \right)^{1/2}$. The inner product between

$\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times \dots \times n_d}$ and $\mathcal{B} \in \mathbb{R}^{n_1 \times n_2 \times \dots \times n_d}$ is denoted as $\langle \mathcal{A}, \mathcal{B} \rangle = \sum_{i_1=1}^{n_1} \dots \sum_{i_d=1}^{n_d} a_{i_1 \dots i_d} b_{i_1 \dots i_d}$. Denote $\mathcal{A}(i_1, \dots, i_{l-1}, :, i_{l+1}, \dots, i_d) \in \mathbb{R}^{n_l}$ as a column vector formed by fixing all modes except for mode- l , and $\mathcal{A}(i_1, \dots, i_{p-1}, :, i_{p+1}, \dots, i_{q-1}, :, i_{q+1}, \dots, i_d) \in \mathbb{R}^{n_p \times n_q}$ ($p < q$) as the matrix slice along mode- p and mode- q . In particular, the matrix slice encompassing the first two modes is designated as $\mathcal{A}^j = \mathcal{A}(:, :, i_3, \dots, i_d) \in \mathbb{R}^{n_1 \times n_2}$, where $j = (i_d - 1)n_3 \dots n_{d-1} + \dots + (i_4 - 1)n_3 + i_3, i_d \in [n_d]$. Denote

$$\text{bdiag}(\mathcal{A}) = \begin{bmatrix} \mathcal{A}^1 & & & \\ & \mathcal{A}^2 & & \\ & & \ddots & \\ & & & \mathcal{A}^J \end{bmatrix} \in \mathbb{R}^{(n_1 n_3 n_4 \dots n_d) \times (n_2 n_3 n_4 \dots n_d)},$$

where $J = n_3 \dots n_d$. Herein, elucidate the face-wise product for d -th order tensors.

Definition 1 (Face-wise product [14]). The face-wise product $\mathcal{A} \Delta \mathcal{B}$ of two d -th order tensors $\mathcal{A} \in \mathbb{C}^{n_1 \times l \times n_3 \times \dots \times n_d}$ and $\mathcal{B} \in \mathbb{C}^{l \times n_2 \times n_3 \times \dots \times n_d}$ is the element of $\mathbb{C}^{n_1 \times n_2 \times n_3 \times \dots \times n_d}$ defined according to

$$\mathcal{C} = \mathcal{A} \Delta \mathcal{B} \iff \text{bdiag}(\mathcal{C}) = \text{bdiag}(\mathcal{A}) \cdot \text{bdiag}(\mathcal{B}). \quad (1)$$

For a given set of invertible transformation matrices denoted as $\{L_i \in \mathbb{C}^{n_i \times n_i}\}_{i=3}^d$, the linear transformation of a d -th order tensor $\mathcal{A}$ under L is defined as $\mathcal{A}_L \triangleq L(\mathcal{A}) = \mathcal{A} \times_3 L_3 \times_4 \dots \times_d L_d$, where the symbol $\times_k$ represents the k -mode product of a tensor with a matrix defined in [15], and the inverse transformation of $\mathcal{A}$ under L is denoted as $L^{-1}(\mathcal{A}) = \mathcal{A} \times_3 L_3^{-1} \times_4 \dots \times_d L_d^{-1}$. Thus, when the corresponding matrices $\{L_i\}_{i=3}^d$ of invertible linear transforms L are unitary matrices, we can obtain

$$\|\mathcal{A}\|_F = \|\mathcal{A}_L\|_F. \quad (2)$$

Next, the clear definitions for the multiplication operation and related concepts of a d -th order tensor $\mathcal{A}$ within the domain of general invertible transformations.

Definition 2 (d -th order T-product [14]). Let $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times \dots \times n_d}$ and $\mathcal{B} \in \mathbb{R}^{n_2 \times l \times n_3 \times \dots \times n_d}$. Then, the invertible linear transforms L based T-product is defined as

$$\mathcal{C} = \mathcal{A} *_L \mathcal{B} = L^{-1}(\mathcal{A}_L \Delta \mathcal{B}_L).$$

Definition 3 (d -th order tensor conjugate transpose [14]). The invertible linear transforms L based conjugate transpose of a tensor $\mathcal{A} \in \mathbb{C}^{n_1 \times n_2 \times n_3 \times \dots \times n_d}$ is the tensor $\mathcal{B} \in \mathbb{C}^{n_2 \times n_1 \times n_3 \times \dots \times n_d}$, whose matrix slice encompassing the first two modes $\mathcal{B}(:, :, i_3, \dots, i_d) = (\mathcal{A}_L(:, :, i_3, \dots, i_d))^H$. We denote by $\mathcal{A}^H$ the conjugate transpose of a tensor $\mathcal{A}$.

Definition 4 (d -th order TNN [14]). Let $\mathcal{A} \in \mathbb{R}^{n_1 \times \dots \times n_d}$, the tensor nuclear norm of $\mathcal{A}$ is defined as

$$\|\mathcal{A}\|_{*,L} = \frac{1}{\rho} \|\text{bdiag}(\mathcal{A}_L)\|_* = \frac{1}{\rho} \sum_{i_3=1}^{n_3} \dots \sum_{i_d=1}^{n_d} \|\mathcal{A}_L(:, :, i_3, \dots, i_d)\|_*,$$

where $\rho > 0$ is a positive constant determined by the invertible linear transforms L .

Remark 1. The constant ρ in the key definition and theorem arises when the corresponding matrices $\{L_i\}_{i=3}^d$ of invertible linear transforms L fulfill the given equation:

$$\begin{aligned} & (L_d \otimes L_{d-1} \otimes \dots \otimes L_3) \cdot (L_d^H \otimes L_{d-1}^H \otimes \dots \otimes L_3^H) \\ &= (L_d^H \otimes L_{d-1}^H \otimes \dots \otimes L_3^H) \cdot (L_d \otimes L_{d-1} \otimes \dots \otimes L_3) \\ &= \rho I_{n_3 n_4 \dots n_d}, \end{aligned}$$

where $\otimes$ is the Kronecker product, I is the identity matrix, and $n_3 n_4 \dots n_d$ is its dimensions.

Based on Definition 4, we can draw the conclusion of the equivalence between d -th order tensor nuclear norm and the minimum of the sum of two factor tensors' squared Frobenius norms:

Theorem 1. If the corresponding matrices $\{L_i\}_{i=3}^d$ of invertible linear transforms L are unitary matrices, for $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times \dots \times n_d}$, we have

$$\|\mathcal{A}\|_{*,L} = \min_{\mathcal{X}, \mathcal{Y}} \left\{ \frac{1}{2} (\|\mathcal{X}\|_F^2 + \|\mathcal{Y}\|_F^2) : \mathcal{A} = \mathcal{X} *_L \mathcal{Y}^H \right\}.$$

Proof: By Definition 4, we have

$$\begin{aligned} \|\mathcal{A}\|_{*,L} &= \frac{1}{\rho} \sum_{i_3=1}^{n_3} \dots \sum_{i_d=1}^{n_d} \|\mathcal{A}_L(:, :, i_3, i_4, \dots, i_d)\|_* \\ &= \min_{\mathcal{X}_L, \mathcal{Y}_L} \left\{ \frac{1}{2\rho} \sum_{i_3=1}^{n_3} \dots \sum_{i_d=1}^{n_d} \left(\|\mathcal{X}_L(:, :, i_3, i_4, \dots, i_d)\|_F^2 \right. \right. \\ &\quad \left. \left. + \|\mathcal{Y}_L(:, :, i_3, \dots, i_d)\|_F^2 \right) : \right. \\ &\quad \left. \mathcal{X}_L(:, :, i_3, i_4, \dots, i_d) \cdot (\mathcal{Y}_L(:, :, i_3, \dots, i_d))^H \right. \\ &\quad \left. = \mathcal{A}_L(:, :, i_3, \dots, i_d) \right\} \\ &= \min_{\mathcal{X}, \mathcal{Y}} \left\{ \frac{1}{2} (\|\mathcal{X}\|_F^2 + \|\mathcal{Y}\|_F^2) : \mathcal{A} = \mathcal{X} *_L \mathcal{Y}^H \right\}, \end{aligned}$$

where the proof for the second equality can be deduced from the proofs of Lemma 5.1 and Proposition 2.1 of [16], and the reason for the third equality holding is that $\{L_i\}_{i=3}^d$ are unitary matrices.

3 Model and algorithm

As mentioned in the introduction, Internet traffic data consist of network traffic records measured T times daily from N origin nodes to N destination nodes. Consequently, data collected over D consecutive days can be organized as a fourth-order tensor in $\mathbb{R}^{D \times T \times N \times N}$ (i.e., $n_1 = D, n_2 = T, n_3 = n_4 = N$), comprehensively capturing the spatio-temporal dynamics of the network traffic.

3.1 Design of model and transformation

Based on low-rank property of observed incomplete tensor $\mathcal{M} \in \mathbb{R}^{n_1 \times n_2 \times n_3 \times n_4}$, we construct a data recovery model for

missing traffic data within the domain of unitary transformation (i.e., the corresponding matrices $\{L_i\}_{i=3}^4$ in L are unitary):

$$\begin{aligned} \min_{\mathcal{X}} \quad & \|\mathcal{X}\|_{*,L} + \frac{\alpha_1}{2} \|H \cdot \text{Mat}_1(\mathcal{X})\|_F^2 + \frac{\alpha_2}{2} \|K \cdot \text{Mat}_2(\mathcal{X})\|_F^2 + \gamma \|\mathcal{X}\|_F^2 \\ \text{s.t.} \quad & P_\Omega(\mathcal{X}) = P_\Omega(\mathcal{M}), \end{aligned} \quad (3)$$

where α_1, α_2 , and γ are positive parameters, $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times n_3 \times n_4}$ is the Internet traffic tensor needed to be estimated, Ω is the index set corresponding to the observed entries of $\mathcal{M}$, and $P_\Omega(\cdot)$ is the linear operator that keeps known elements in Ω while setting the others to be zeros. Here, $\text{Mat}_1(\mathcal{X}) = [\mathcal{X}(:, 1, 1, 1), \dots, \mathcal{X}(:, n_2, 1, 1), \dots, \mathcal{X}(:, n_2, n_3, 1), \dots, \mathcal{X}(:, n_2, n_3, n_4)] \in \mathbb{R}^{n_1 \times (n_2 n_3 n_4)}$, where adjacent rows represent flow data measurements from consecutive days. And $\text{Mat}_2(\mathcal{X}) = [\mathcal{X}(1, :, 1, 1), \dots, \mathcal{X}(n_1, :, 1, 1), \dots, \mathcal{X}(n_1, :, n_3, 1), \dots, \mathcal{X}(n_1, :, n_3, n_4)] \in \mathbb{R}^{n_2 \times (n_1 n_3 n_4)}$, where adjacent rows represent flow data measurements from consecutive time points. The regularization term $\gamma \|\mathcal{X}\|_F^2$ is employed to ensure the boundedness of the sequence generated by the proposed algorithm. To better capture the periodicity of the data measured on each observation day and the similarity of the data obtained between adjacent observation time points within each day, we choose the temporal constraint matrix $H = \text{Toeplitz}(0, 1, -1)$ of the size $(n_1 - 1) \times n_1$ and $K = \text{Toeplitz}(0, 1, -1)$ of size $(n_2 - 1) \times n_2$, respectively, i.e.,

$$H = \begin{bmatrix} 1 & -1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & -1 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -1 \end{bmatrix}_{(n_1-1) \times n_1},$$

$$K = \begin{bmatrix} 1 & -1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & -1 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -1 \end{bmatrix}_{(n_2-1) \times n_2}.$$

It is well-known that the Equation (Model) 3 requires the so-called SVD to compute the $\|\cdot\|_{*,L}$ minimization subproblem, which often takes much computing time for large-scale tensors. Therefore, with the employment of Theorem 1, we gainfully reformulate Equation (Model) 3 as the following tensor factorization version:

$$\begin{aligned} \min_{\mathcal{X}, \mathcal{U}, \mathcal{V}} \quad & \frac{1}{2} (\|\mathcal{U}\|_F^2 + \|\mathcal{V}\|_F^2) + \frac{\lambda}{2} \|\mathcal{X} - \mathcal{U} *_{L} \mathcal{V}\|_F^2 + \gamma \|\mathcal{X}\|_F^2 \\ & + \frac{\alpha_1}{2} \|H \cdot \text{Mat}_1(\mathcal{X})\|_F^2 \\ & + \frac{\alpha_2}{2} \|K \cdot \text{Mat}_2(\mathcal{X})\|_F^2 \\ \text{s.t.} \quad & P_\Omega(\mathcal{X}) = P_\Omega(\mathcal{M}), \end{aligned} \quad (4)$$

where $\mathcal{U} \in \mathbb{R}^{n_1 \times s \times n_3 \times n_4}$, $\mathcal{V} \in \mathbb{R}^{s \times n_2 \times n_3 \times n_4}$. Clearly, such an optimization model no longer requires SVD to find its solution. Therefore, we call Equation (Model) 4 a SVD-free model. Moreover, to facilitate the solution of Equation (Model) 4, we

introduce auxiliary variables $\mathcal{Y}, \mathcal{Z} \in \mathbb{R}^{n_1 \times n_2 \times n_3 \times n_4}$. Thus, the Equation (Model) 4 is equivalently written as

$$\begin{aligned} \min_{\mathcal{X}, \mathcal{U}, \mathcal{V}, \mathcal{Y}, \mathcal{Z}} \quad & f(\mathcal{X}, \mathcal{U}, \mathcal{V}, \mathcal{Y}, \mathcal{Z}) = \frac{1}{2} (\|\mathcal{U}\|_F^2 + \|\mathcal{V}\|_F^2) + \frac{\lambda}{2} \|\mathcal{X} - \mathcal{U} *_{L} \mathcal{V}\|_F^2 \\ & + \frac{\gamma}{2} \|\mathcal{Y}\|_F^2 + \frac{\gamma}{2} \|\mathcal{Z}\|_F^2 + \frac{\alpha_1}{2} \|H \cdot \text{Mat}_1(\mathcal{Y})\|_F^2 \\ & + \frac{\alpha_2}{2} \|K \cdot \text{Mat}_2(\mathcal{Z})\|_F^2 \\ \text{s.t.} \quad & P_\Omega(\mathcal{X}) = P_\Omega(\mathcal{M}), \mathcal{Y} = \mathcal{X}, \mathcal{Z} = \mathcal{X}. \end{aligned} \quad (5)$$

3.2 Description of algorithm

We apply the alternating direction method of multipliers (ADMM) algorithm to solve Equation (Model) 5. The augmented Lagrangian function of Equation (Model) 5 is written as follows:

$$\begin{aligned} \mathcal{L}(\mathcal{X}, \mathcal{U}, \mathcal{V}, \mathcal{Y}, \mathcal{Z}, \mathcal{W}_1, \mathcal{W}_2) = & f(\mathcal{X}, \mathcal{U}, \mathcal{V}, \mathcal{Y}, \mathcal{Z}) \\ & + \langle \mathcal{W}_1, \mathcal{X} - \mathcal{Y} \rangle + \frac{\beta_1}{2} \|\mathcal{X} - \mathcal{Y}\|_F^2 + \langle \mathcal{W}_2, \mathcal{X} - \mathcal{Z} \rangle + \frac{\beta_2}{2} \|\mathcal{X} - \mathcal{Z}\|_F^2, \end{aligned} \quad (6)$$

where $\mathcal{W}_1$ and $\mathcal{W}_2 \in \mathbb{R}^{n_1 \times n_2 \times n_3 \times n_4}$ are the Lagrangian multipliers, and β_1 and β_2 are positive penalty parameters. Below, given the q -th iterate $\Theta^q := (\mathcal{X}^q, \mathcal{U}^q, \mathcal{V}^q, \mathcal{Y}^q, \mathcal{Z}^q, \mathcal{W}_1^q, \mathcal{W}_2^q)$, we provide specific solutions for the updates of each variable.

► The $\mathcal{X}$ -subproblem

We update $\mathcal{X}$ by solving the following subproblem

$$\begin{aligned} \min_{\mathcal{X}} \quad & \frac{\lambda}{2} \|\mathcal{X} - \mathcal{U}^q *_{L} \mathcal{V}^q\|_F^2 + \langle \mathcal{W}_1^q, \mathcal{X} - \mathcal{Y}^q \rangle + \frac{\beta_1}{2} \|\mathcal{X} - \mathcal{Y}^q\|_F^2 \\ & + \langle \mathcal{W}_2^q, \mathcal{X} - \mathcal{Z}^q \rangle + \frac{\beta_2}{2} \|\mathcal{X} - \mathcal{Z}^q\|_F^2 \\ \text{s.t.} \quad & P_\Omega(\mathcal{X}) = P_\Omega(\mathcal{M}). \end{aligned} \quad (7)$$

The solution of Equation (Model) 7 is given by

$$\mathcal{X}_{ijkl}^{q+1} = \begin{cases} \frac{1}{\lambda + \beta_1 + \beta_2} (\lambda \mathcal{U}^q *_{L} \mathcal{V}^q + \beta_1 \mathcal{Y}^q + \beta_2 \mathcal{Z}^q - \mathcal{W}_1^q - \mathcal{W}_2^q)_{ijkl}, & (i, j, k, l) \notin \Omega, \\ \mathcal{M}_{ijkl}, & \text{otherwise.} \end{cases} \quad (8)$$

► The $\mathcal{U}$ -subproblem

We update $\mathcal{U}$ by

$$\mathcal{U}^{q+1} = \arg \min_{\mathcal{U}} \left\{ \frac{1}{2} \|\mathcal{U}\|_F^2 + \frac{\lambda}{2} \|\mathcal{X}^{q+1} - \mathcal{U} *_{L} \mathcal{V}^q\|_F^2 \right\}. \quad (9)$$

Because the corresponding matrices $\{L_i\}_{i=3}^4$ in L are unitary, it holds that for $j \in [n_3 n_4]$,

$$\begin{aligned} & \frac{1}{2} \|\mathcal{U}\|_F^2 + \frac{\lambda}{2} \|\mathcal{X}^{q+1} - \mathcal{U} *_{L} \mathcal{V}^q\|_F^2 \\ & = \frac{1}{2} \|\mathcal{U}_L\|_F^2 + \frac{\lambda}{2} \|\mathcal{X}_L^{q+1} - \mathcal{U}_L \Delta \mathcal{V}_L^q\|_F^2 \\ & = \frac{1}{2} \sum_{j=1}^{n_3 n_4} \left(\|\mathcal{U}_L^j\|_F^2 + \lambda \|(\mathcal{X}_L^j)^{q+1} - \mathcal{U}_L^j (\mathcal{V}_L^j)^q\|_F^2 \right). \end{aligned} \quad (10)$$

Due to the separable structure of Equation 10 with respect to decision variables U_L^j for $j \in [n_3 n_4]$, solving the Equation 9 is equivalent to solving

$$\min_{U_L^j} \frac{1}{2} \|U_L^j\|_F^2 + \frac{\lambda}{2} \|(X_L^j)^{q+1} - U_L^j (V_L^j)^q\|_F^2, \quad j \in [n_3 n_4], \quad (11)$$

whose solution is given by

$$(U_L^j)^{q+1} = \left(\lambda (X_L^j)^{q+1} ((V_L^j)^{q+1})^\top \right. \\ \left. \left(\lambda (V_L^j)^{q+1} ((V_L^j)^{q+1})^\top + I \right)^{-1} \right).$$

With the help of the obtained $(U_L^j)^{q+1}$ for $j \in [n_3 n_4]$, we can get $\mathcal{U}^{q+1}$ by

$$\mathcal{U}^{q+1} = L^{-1}(\mathcal{U}_L^{q+1}). \quad (12)$$

► The $\mathcal{V}$ -subproblem

We update $\mathcal{V}$ by

$$\mathcal{V}^{q+1} = \arg \min_{\mathcal{V}} \left\{ \frac{1}{2} \|\mathcal{V}\|_F^2 + \frac{\lambda}{2} \|\mathcal{X}^{q+1} - \mathcal{U}^{q+1} *_L \mathcal{V}\|_F^2 \right\}. \quad (13)$$

Similarly to $\mathcal{U}$ -subproblem, solving Equation 13 is equivalent to solving the following problem:

$$\min_{V_L^j} \frac{1}{2} \|V_L^j\|_F^2 + \frac{\lambda}{2} \|(X_L^j)^{q+1} - (U_L^j)^{q+1} \cdot V_L^j\|_F^2, \quad (14)$$

for every $j \in [n_3 n_4]$. Clearly, the solution of Equation (Model) 14 is given by

$$(V_L^j)^{q+1} = \left(\lambda ((U_L^j)^{q+1})^\top (U_L^j)^{q+1} + I \right)^{-1} \\ \left(\lambda ((U_L^j)^{q+1})^\top (X_L^j)^{q+1} \right),$$

for every $j \in [n_3 n_4]$. Thus, $\mathcal{V}^{q+1}$ can be obtained via

$$\mathcal{V}^{q+1} = L^{-1}(\mathcal{V}_L^{q+1}). \quad (15)$$

► The $\mathcal{Y}$ -subproblem

We update $\mathcal{Y}$ by

$$\mathcal{Y}^{q+1} = \arg \min_{\mathcal{Y}} \left\{ \frac{\alpha_1}{2} \|H \cdot \text{Mat}_1(\mathcal{Y})\|_F^2 + \langle \mathcal{W}_1^q, \mathcal{X}^{q+1} - \mathcal{Y} \rangle \right. \\ \left. + \frac{\beta_1}{2} \|\mathcal{X}^{q+1} - \mathcal{Y}\|_F^2 + \frac{\gamma}{2} \|\mathcal{Y}\|_F^2 \right\}, \quad (16)$$

which is equivalent to solving

$$\min_{\text{Mat}_1(\mathcal{Y})} \frac{\alpha_1}{2} \|H \cdot \text{Mat}_1(\mathcal{Y})\|_F^2 + \frac{\beta_1}{2} \|\text{Mat}_1(\mathcal{Y}) - (\text{Mat}_1(\mathcal{X}^{q+1}) \\ + \frac{1}{\beta_1} \text{Mat}_1(\mathcal{W}_1^q))\|_F^2 + \frac{\gamma}{2} \|\text{Mat}_1(\mathcal{Y})\|_F^2. \quad (17)$$

By the optimality condition of Equation (Model) 17, we have

$$\alpha_1 H^\top H \text{Mat}_1(\mathcal{Y}^{q+1}) + \beta_1 (\text{Mat}_1(\mathcal{Y}^{q+1}) - (\text{Mat}_1(\mathcal{X}^{q+1}) \\ + \frac{1}{\beta_1} \text{Mat}_1(\mathcal{W}_1^q))) + \gamma \text{Mat}_1(\mathcal{Y}^{q+1}) = 0, \quad (18)$$

```

1: Select parameters  $\lambda, \alpha_1, \alpha_2, \beta_1, \beta_2, \gamma, \varepsilon > 0$  and
   starting point  $\Theta^0 = (\mathcal{X}^0, \mathcal{U}^0, \mathcal{V}^0, \mathcal{Y}^0, \mathcal{Z}^0, \mathcal{W}_1^0, \mathcal{W}_2^0)$ .
2: for  $q = 0, 1, 2, \dots$  do
3:   Update  $\mathcal{X}^{q+1}$  via (Equation 8).
4:   Update  $\mathcal{U}^{q+1}$  via (Equation 12).
5:   Update  $\mathcal{V}^{q+1}$  via (Equation 15).
6:   Update  $\mathcal{Y}^{q+1}$  via (Equation 20).
7:   Update  $\mathcal{Z}^{q+1}$  via (Equation 22).
8:   Update  $\mathcal{W}_1^{q+1}$  and  $\mathcal{W}_2^{q+1}$  via (Equation 23).
9:   If  $\|\mathcal{X}^{q+1} - \mathcal{X}^q\|_F^2 / \|\mathcal{X}^q\|_F^2 < \varepsilon$ , then stop.
10: end for

```

Algorithm 1. Tensor completion algorithm with d -th order tensor product decomposition and spatio-temporal regularization (SRdTPD).

which implies that the optimal solution of Equation (Model) 17 is

$$\text{Mat}_1(\mathcal{Y}^{q+1}) = \left(\alpha_1 H^\top H + (\beta_1 + \gamma)I \right)^{-1} \\ \left(\beta_1 \text{Mat}_1(\mathcal{X}^{q+1}) + \text{Mat}_1(\mathcal{W}_1^q) \right). \quad (19)$$

As a result, the optimal solution of Equation 16 is

$$\mathcal{Y}^{q+1} = \text{IMat}_1 \left(\left(\alpha_1 H^\top H + (\beta_1 + \gamma)I \right)^{-1} \right. \\ \left. \left(\beta_1 \text{Mat}_1(\mathcal{X}^{q+1}) + \text{Mat}_1(\mathcal{W}_1^q) \right) \right), \quad (20)$$

where $\text{IMat}_1(\cdot)$ is the inverse of $\text{Mat}_1(\cdot)$.

► The $\mathcal{Z}$ -subproblem

We update $\mathcal{Z}$ by

$$\mathcal{Z} = \arg \min_{\mathcal{Z}} \left\{ \frac{\alpha_2}{2} \|K \cdot \text{Mat}_2(\mathcal{Z})\|_F^2 + \langle \mathcal{W}_2^q, \mathcal{X}^{q+1} - \mathcal{Z} \rangle \right. \\ \left. + \frac{\beta_2}{2} \|\mathcal{X}^{q+1} - \mathcal{Z}\|_F^2 + \frac{\gamma}{2} \|\mathcal{Z}\|_F^2 \right\}. \quad (21)$$

Similarly to $\mathcal{Y}$ -subproblem, the optimal solution of Equation 21 is

$$\mathcal{Z}^{q+1} = \text{IMat}_2 \left(\left(\alpha_2 K^\top K + (\beta_2 + \gamma)I \right)^{-1} \right. \\ \left. \left(\beta_2 \text{Mat}_2(\mathcal{X}^{q+1}) + \text{Mat}_2(\mathcal{W}_2^q) \right) \right), \quad (22)$$

where $\text{IMat}_2(\cdot)$ is the inverse of $\text{Mat}_2(\cdot)$.

► The $\mathcal{W}_1$ and $\mathcal{W}_2$ -subproblems

We update $\mathcal{W}_1$ and $\mathcal{W}_2$ by

$$\mathcal{W}_1^{q+1} = \mathcal{W}_1^q + \beta_1 (\mathcal{X}^{q+1} - \mathcal{Y}^{q+1}), \\ \mathcal{W}_2^{q+1} = \mathcal{W}_2^q + \beta_2 (\mathcal{X}^{q+1} - \mathcal{Z}^{q+1}). \quad (23)$$

Following the above analysis, the algorithmic framework for solving Equation (Model) 5 is presented at Algorithm 1.

Remark 2. In the $\mathcal{U}$ and $\mathcal{V}$ -subproblems, the inverses of two related matrices are computed directly. In the $\mathcal{Y}$ and $\mathcal{Z}$ -subproblems, we first utilize the special structure of H and K to perform Cholesky

decomposition on the two matrices $H^T H + (\beta_1 + \gamma)I$ and $K^T K + (\beta_2 + \gamma)I$, decomposing them into the product of two triangular matrices respectively, and then obtain their inverse matrices. This will significantly reduce the computational cost. In addition, considering that the inverses of the two related matrices remain constant in each round of solving the $\mathcal{Y}$ and $\mathcal{Z}$ -subproblems, we only need to compute the inverses once, and in the subsequent solving of the $\mathcal{Y}$ and $\mathcal{Z}$ -subproblems, we can simply call the inverse matrices directly.

4 Convergence analysis

In this section, we investigate the convergence properties of Algorithm 1. Although the convergence of ADMM algorithm for solving the general multi-block optimization cannot be guaranteed [17], we can still use the special structure of the Equation (Model) 5 to obtain the convergence of Algorithm 1, whose proofs are similar to the proofs in Section 3.2 of [18]. We first present the Karush-Kuhn-Tucker (KKT) conditions for Equation (Model) 5:

$$\nabla_{\mathcal{X}^c} f(\mathcal{X}, \mathcal{U}, \mathcal{V}, \mathcal{Y}, \mathcal{Z}) + \mathcal{W}_{1\Omega^c} + \mathcal{W}_{2\Omega^c} = \mathcal{O}_{\Omega^c}, \quad (24)$$

$$\nabla_{\mathcal{U}} f(\mathcal{X}, \mathcal{U}, \mathcal{V}, \mathcal{Y}, \mathcal{Z}) = \mathcal{O}, \quad (25)$$

$$\nabla_{\mathcal{V}} f(\mathcal{X}, \mathcal{U}, \mathcal{V}, \mathcal{Y}, \mathcal{Z}) = \mathcal{O}, \quad (26)$$

$$\nabla_{\mathcal{Y}} f(\mathcal{X}, \mathcal{U}, \mathcal{V}, \mathcal{Y}, \mathcal{Z}) - \mathcal{W}_1 = \mathcal{O}, \quad (27)$$

$$\nabla_{\mathcal{Z}} f(\mathcal{X}, \mathcal{U}, \mathcal{V}, \mathcal{Y}, \mathcal{Z}) - \mathcal{W}_2 = \mathcal{O}, \quad (28)$$

$$P_{\Omega}(\mathcal{X}) = P_{\Omega}(\mathcal{M}), \quad (29)$$

$$\mathcal{Y} = \mathcal{X}, \quad (30)$$

$$\mathcal{Z} = \mathcal{X}, \quad (31)$$

where Ω^c represents the complement of the set Ω . Moreover, we call $(\mathcal{X}^o, \mathcal{U}^o, \mathcal{V}^o, \mathcal{Z}^o, \mathcal{W}_1^o, \mathcal{W}_2^o)$ satisfying (Equations 24–31) a stationary point of Equation (Model) 5, where $\mathcal{W}_1^o$ and $\mathcal{W}_2^o$ are Lagrange multipliers associated with the constraints $\mathcal{Y}^o = \mathcal{X}^o$ and $\mathcal{Z}^o = \mathcal{X}^o$, respectively. For any integer $q \geq 0$, denote $\Theta^q = (\mathcal{X}^q, \mathcal{U}^q, \mathcal{V}^q, \mathcal{Y}^q, \mathcal{Z}^q, \mathcal{W}_1^q, \mathcal{W}_2^q)$.

Proposition 1. Let $\{\Theta^q = (\mathcal{X}^q, \mathcal{U}^q, \mathcal{V}^q, \mathcal{Y}^q, \mathcal{Z}^q, \mathcal{W}_1^q, \mathcal{W}_2^q)\}_{q=0}^{\infty}$ be the sequence generated by Algorithm 1. Then, we have

$$\begin{aligned} \mathcal{L}(\mathcal{X}^{q+1}, \mathcal{U}^q, \mathcal{V}^q, \mathcal{Y}^q, \mathcal{Z}^q, \mathcal{W}_1^q, \mathcal{W}_2^q) + \frac{\lambda + \beta_1 + \beta_2}{2} \|\mathcal{X}^{q+1} - \mathcal{X}^q\|_F^2 \\ \leq \mathcal{L}(\mathcal{X}^q, \mathcal{U}^q, \mathcal{V}^q, \mathcal{Y}^q, \mathcal{Z}^q, \mathcal{W}_1^q, \mathcal{W}_2^q). \end{aligned} \quad (32)$$

Proof: We consider three cases: (a) $\Omega = \emptyset$, (b) $\emptyset \subsetneq \Omega \subsetneq [n_1] \times [n_2] \times [n_3] \times [n_4]$, and (c) $\Omega = [n_1] \times [n_2] \times [n_3] \times [n_4]$. In case (a), the $\mathcal{X}$ -subproblem Equation (Model) 7 becomes an unconstrained optimization problem, which is equivalent to $\min_{\mathcal{X}} f_1(\mathcal{X})$, where

$$f_1(\mathcal{X}) = \frac{\lambda}{2} \|\mathcal{X} - \mathcal{C}^q\|_F^2 + \frac{\beta_1}{2} \|\mathcal{X} - \mathcal{D}^q\|_F^2 + \frac{\beta_2}{2} \|\mathcal{X} - \mathcal{E}^q\|_F^2 \quad (33)$$

with $\mathcal{C}^q = \mathcal{U}^q *_L \mathcal{V}^q$, $\mathcal{D}^q = \mathcal{Y}^q - (1/\beta_1)\mathcal{W}_1^q$ and $\mathcal{E}^q = \mathcal{Z}^q - (1/\beta_2)\mathcal{W}_2^q$. It is obvious that f_1 is strongly convex with modulus at least $\lambda + \beta_1 + \beta_2$. Consequently, by Theorem 5.24 in [19], it holds that

$$f_1(\mathcal{X}') \geq f_1(\mathcal{X}) + \langle \nabla f_1(\mathcal{X}), \mathcal{X}' - \mathcal{X} \rangle + \frac{\lambda + \beta_1 + \beta_2}{2} \|\mathcal{X}' - \mathcal{X}\|_F^2 \quad (34)$$

for any $\mathcal{X}, \mathcal{X}' \in \mathbb{R}^{n_1 \times n_2 \times n_3 \times n_4}$. Since $\mathcal{X}^{q+1}$ is the optimal solution of $\min_{\mathcal{X}} f_1(\mathcal{X})$, which means $\nabla f_1(\mathcal{X}^{q+1}) = \mathcal{O}$, from Equation (Inequality) 34, it holds that

$$f_1(\mathcal{X}^q) \geq f_1(\mathcal{X}^{q+1}) + \frac{\lambda + \beta_1 + \beta_2}{2} \|\mathcal{X}^q - \mathcal{X}^{q+1}\|_F^2, \quad (35)$$

which implies, together with the definition of $\mathcal{L}$, that Equation (Inequality) 32 holds.

In the case $\emptyset \subsetneq \Omega \subsetneq [n_1] \times [n_2] \times [n_3] \times [n_4]$, the simple equality constraints in Equation (Model) 7 can be eliminated by substituting them into its objective function, and the corresponding $\mathcal{X}$ -subproblem is converted into an unconstrained optimization problem, whose objective function is similar to the structure of f_1 , but $\mathcal{X}$, $\mathcal{C}^q$, $\mathcal{D}^q$, and $\mathcal{E}^q$ are replaced by $\mathcal{X}_{\Omega^c}$, $\mathcal{C}_{\Omega^c}^q$, $\mathcal{D}_{\Omega^c}^q$, and $\mathcal{E}_{\Omega^c}^q$, respectively. Noticing $\mathcal{X}_{\Omega^c}^{q+1} = \mathcal{M}_{\Omega^c}$, similar to the proof for the case (a), we know that Equation (Inequality) 32 still holds. In the case (c), since $\mathcal{X}^{q+1} = \mathcal{X}^q = \mathcal{M}$, the inequality [Equation (Inequality) 32] is obvious. We complete the proof.

Proposition 2. Let $\{\Theta^q\}_{q=0}^{\infty}$ be the sequence generated by Algorithm 1. Then we have

$$\begin{aligned} \mathcal{L}(\mathcal{X}^{q+1}, \mathcal{U}^{q+1}, \mathcal{V}^q, \mathcal{Y}^q, \mathcal{Z}^q, \mathcal{W}_1^q, \mathcal{W}_2^q) + \frac{1}{2} \|\mathcal{U}^{q+1} - \mathcal{U}^q\|_F^2 \\ \leq \mathcal{L}(\mathcal{X}^{q+1}, \mathcal{U}^q, \mathcal{V}^q, \mathcal{Y}^q, \mathcal{Z}^q, \mathcal{W}_1^q, \mathcal{W}_2^q) \end{aligned} \quad (36)$$

and

$$\begin{aligned} \mathcal{L}(\mathcal{X}^{q+1}, \mathcal{U}^{q+1}, \mathcal{V}^{q+1}, \mathcal{Y}^q, \mathcal{Z}^q, \mathcal{W}_1^q, \mathcal{W}_2^q) + \frac{1}{2} \|\mathcal{V}^{q+1} - \mathcal{V}^q\|_F^2 \\ \leq \mathcal{L}(\mathcal{X}^{q+1}, \mathcal{U}^{q+1}, \mathcal{V}^q, \mathcal{Y}^q, \mathcal{Z}^q, \mathcal{W}_1^q, \mathcal{W}_2^q). \end{aligned} \quad (37)$$

Proof: For every $j \in [n_3 n_4]$ and integer $q \geq 0$, we define the function $f_{jq} : \mathbb{C}^{n_1 \times s} \rightarrow \mathbb{R}$ by

$$f_{jq}(\mathcal{U}_L^j) = \frac{1}{2} \|\mathcal{U}_L^j\|_F^2 + \frac{\lambda}{2} \|(\mathcal{X}_L^j)^{q+1} - \mathcal{U}_L^j (\mathcal{V}_L^j)^q\|_F^2.$$

It is obvious that f_{jq} is strongly convex with modulus at least 1. Similar to the proof for (a) of Proposition 1, we have

$$f_{jq}((\mathcal{U}_L^j)^{q+1}) + \frac{1}{2} \|(\mathcal{U}_L^j)^{q+1} - (\mathcal{U}_L^j)^q\|_F^2 \leq f_{jq}((\mathcal{U}_L^j)^q),$$

which implies, together with the fact (Equation 2), that

$$\begin{aligned} \frac{1}{2} \|\mathcal{U}^{q+1}\|_F^2 + \frac{\lambda}{2} \|\mathcal{X}^{q+1} - \mathcal{U}^{q+1} *_L \mathcal{V}^q\|_F^2 + \frac{1}{2} \|\mathcal{U}^{q+1} - \mathcal{U}^q\|_F^2 \\ \leq \frac{1}{2} \|\mathcal{U}^q\|_F^2 + \frac{\lambda}{2} \|\mathcal{X}^{q+1} - \mathcal{U}^q *_L \mathcal{V}^q\|_F^2. \end{aligned}$$

Moreover, by the definition of $\mathcal{L}$, we know that the first inequality in Equation (Inequality) 36 holds. The inequality [Equation (Inequality) 37] can be proved similarly.

Employing a proof analogous to that of Proposition 1, we can derive the following proposition.

Proposition 3. Let $\{\Theta^q\}_{q=0}^{\infty}$ be the sequence generated by Algorithm 1. Then, we have

$$\begin{aligned} \mathcal{L}(\mathcal{X}^{q+1}, \mathcal{U}^{q+1}, \mathcal{V}^{q+1}, \mathcal{Y}^{q+1}, \mathcal{Z}^q, \mathcal{W}_1^q, \mathcal{W}_2^q) + \frac{\beta_1 + \gamma}{2} \|\mathcal{Y}^{q+1} - \mathcal{Y}^q\|_F^2 \\ \leq \mathcal{L}(\mathcal{X}^{q+1}, \mathcal{U}^{q+1}, \mathcal{V}^{q+1}, \mathcal{Y}^q, \mathcal{Z}^q, \mathcal{W}_1^q, \mathcal{W}_2^q), \end{aligned} \quad (38)$$

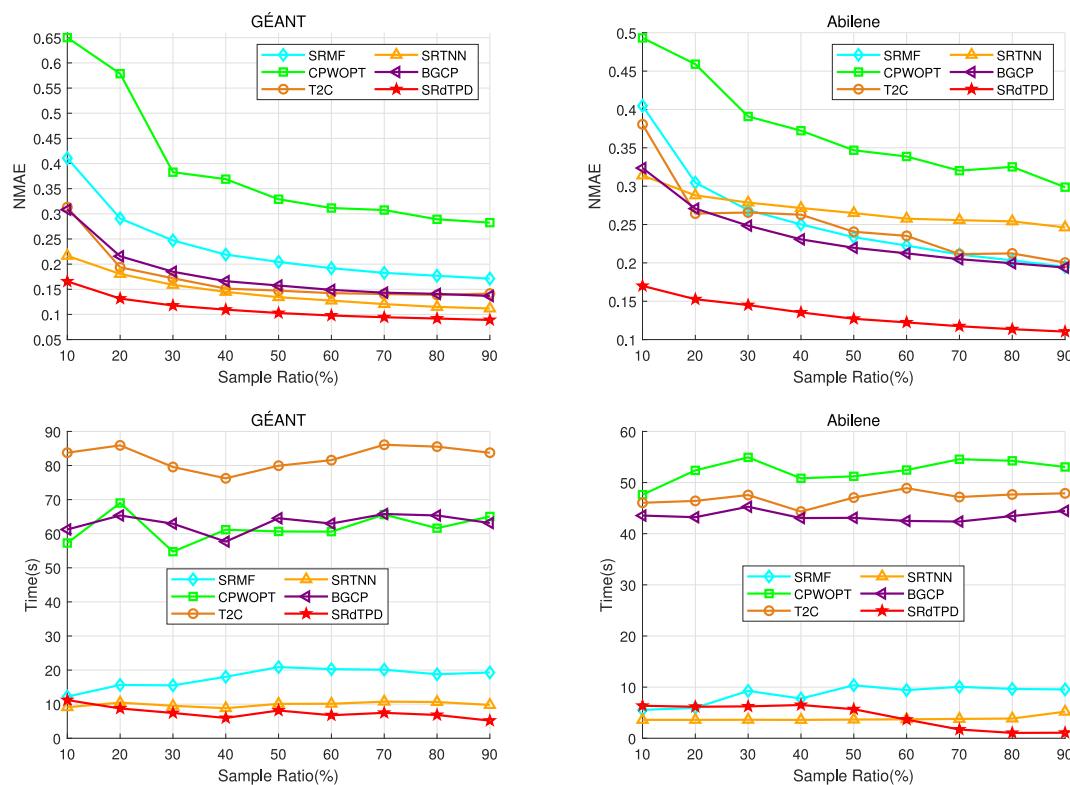


FIGURE 1
Numerical comparison on NMAE and computing time (seconds) with respect to different random sample ratios.

and

$$\begin{aligned} & \mathcal{L}(\mathcal{X}^{q+1}, \mathcal{U}^{q+1}, \mathcal{V}^{q+1}, \mathcal{Y}^{q+1}, \mathcal{Z}^{q+1}, \mathcal{W}_1^q, \mathcal{W}_2^q) \\ & + \frac{\beta_2 + \gamma}{2} \|\mathcal{Z}^{q+1} - \mathcal{Z}^q\|_F^2 \\ & \leq \mathcal{L}(\mathcal{X}^{q+1}, \mathcal{U}^{q+1}, \mathcal{V}^{q+1}, \mathcal{Y}^{q+1}, \mathcal{Z}^q, \mathcal{W}_1^q, \mathcal{W}_2^q). \end{aligned} \quad (39)$$

Denote

$$\begin{aligned} \xi_1 &= \beta_1(\beta_1 + \gamma) - 2(\alpha_1\sigma_1(H^\top H) + \gamma)^2 \text{ and} \\ \xi_2 &= \beta_2(\beta_2 + \gamma) - 2(\alpha_2\sigma_1(K^\top K) + \gamma)^2. \end{aligned}$$

From Propositions 1, 2, and 3, we have the following theorem, which characterizes the sufficient decrease property and boundedness of the sequence $\{\Theta^q\}_{q=0}^\infty$.

Theorem 2. Let $\{\Theta^q\}_{q=0}^\infty$ be the sequence generated by Algorithm 1. If $\xi_1 > 0$ and $\xi_2 > 0$, then the sequence $\{\Theta^q\}_{q=0}^\infty$ satisfies

$$\mathcal{L}(\Theta^{q+1}) + \mu \|\Lambda^{q+1} - \Lambda^q\|_F^2 \leq \mathcal{L}(\Theta^q), \quad (40)$$

where $\Lambda^q = (\mathcal{X}^q, \mathcal{U}^q, \mathcal{V}^q, \mathcal{Y}^q, \mathcal{Z}^q)$ and $\mu = (1/2)\min\{1, \lambda + \beta_1 + \beta_2, \xi_1/\beta_1, \xi_2/\beta_2\}$. Furthermore, if $\eta_1 := \beta_1\gamma - (\alpha_1\sigma_1(H^\top H) + \gamma)^2 > 0$ and $\eta_2 := \beta_2\gamma - (\alpha_2\sigma_1(K^\top K) + \gamma)^2 > 0$, then the sequence $\{\Theta^q\}_{q=0}^\infty$ is bounded.

Proof: From Equations 6, 23, we have

$$\begin{aligned} \mathcal{L}(\Theta^{q+1}) &= \mathcal{L}(\mathcal{X}^{q+1}, \mathcal{U}^{q+1}, \mathcal{V}^{q+1}, \mathcal{Y}^{q+1}, \mathcal{Z}^{q+1}, \mathcal{W}_1^q, \mathcal{W}_2^q) \\ & + \frac{1}{\beta_1} \|\mathcal{W}_1^{q+1} - \mathcal{W}_1^q\|_F^2 + \frac{1}{\beta_2} \|\mathcal{W}_2^{q+1} - \mathcal{W}_2^q\|_F^2. \end{aligned} \quad (41)$$

Moreover, from Equations (Inequalities) 32, 36, 38, we have

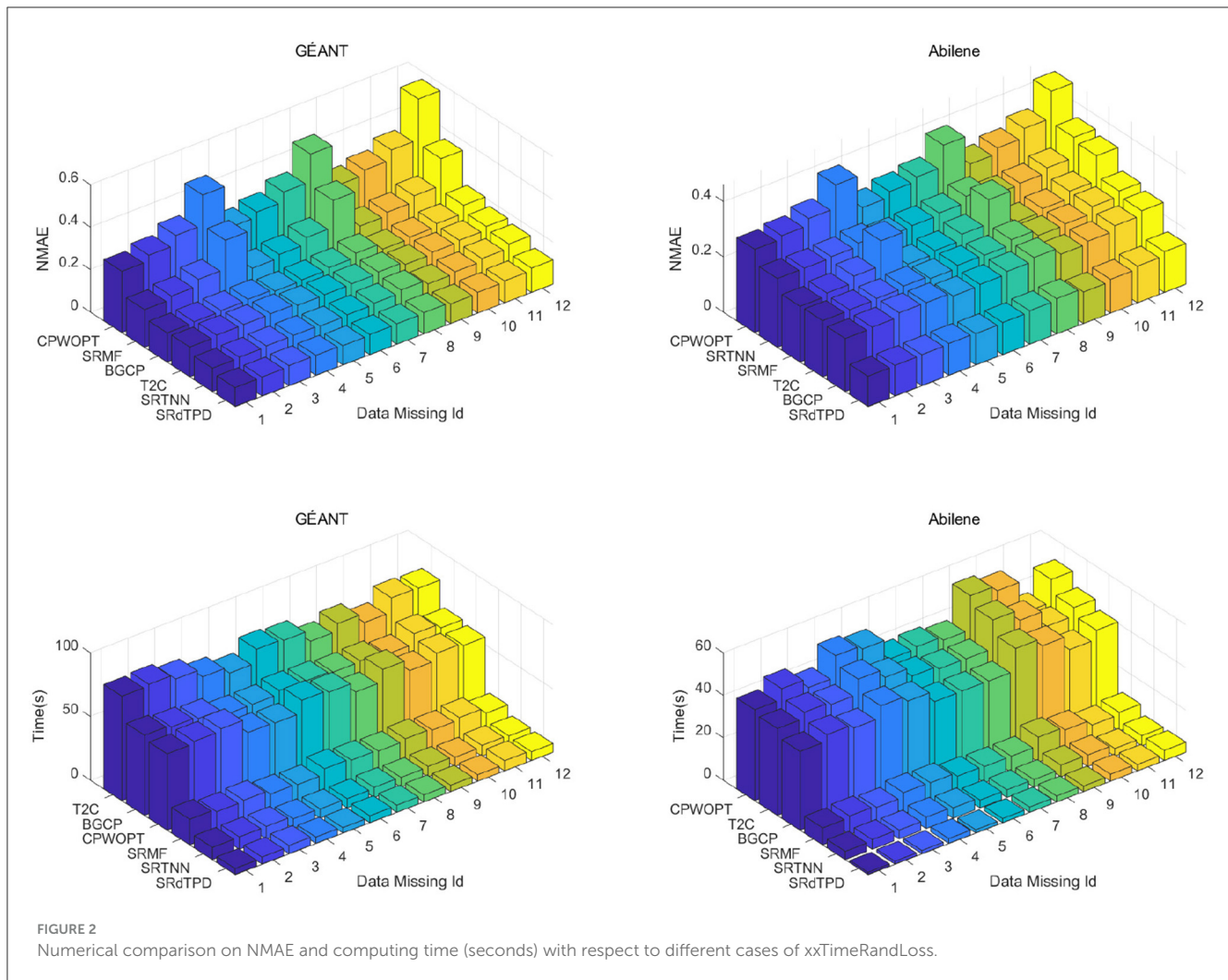
$$\begin{aligned} & \mathcal{L}(\mathcal{X}^{q+1}, \mathcal{U}^{q+1}, \mathcal{V}^{q+1}, \mathcal{Y}^{q+1}, \mathcal{Z}^{q+1}, \mathcal{W}_1^q, \mathcal{W}_2^q) \\ & + \frac{\lambda + \beta_1 + \beta_2}{2} \|\mathcal{X}^{q+1} - \mathcal{X}^q\|_F^2 \\ & + \frac{1}{2} \|\mathcal{U}^{q+1} - \mathcal{U}^q\|_F^2 + \frac{1}{2} \|\mathcal{V}^{q+1} - \mathcal{V}^q\|_F^2 + \frac{\beta_1 + \gamma}{2} \|\mathcal{Y}^{q+1} - \mathcal{Y}^q\|_F^2 \\ & + \frac{\beta_2 + \gamma}{2} \|\mathcal{Z}^{q+1} - \mathcal{Z}^q\|_F^2 \\ & \leq \mathcal{L}(\Theta^q). \end{aligned} \quad (42)$$

Consequently, by Equation 41 and Equation (Inequality) 42, it holds that

$$\begin{aligned} & \mathcal{L}(\Theta^{q+1}) + \frac{\lambda + \beta_1 + \beta_2}{2} \|\mathcal{X}^{q+1} - \mathcal{X}^q\|_F^2 + \frac{1}{2} \|\mathcal{U}^{q+1} - \mathcal{U}^q\|_F^2 \\ & + \frac{1}{2} \|\mathcal{V}^{q+1} - \mathcal{V}^q\|_F^2 + \frac{\beta_1 + \gamma}{2} \|\mathcal{Y}^{q+1} - \mathcal{Y}^q\|_F^2 \\ & + \frac{\beta_2 + \gamma}{2} \|\mathcal{Z}^{q+1} - \mathcal{Z}^q\|_F^2 \\ & \leq \mathcal{L}(\Theta^q) + \frac{1}{\beta_1} \|\mathcal{W}_1^{q+1} - \mathcal{W}_1^q\|_F^2 + \frac{1}{\beta_2} \|\mathcal{W}_2^{q+1} - \mathcal{W}_2^q\|_F^2. \end{aligned} \quad (43)$$

On the other hand, it holds that

$$\text{Mat}_1(\mathcal{W}_1^{q+1}) = (\alpha_1 H^\top H + \gamma I) \text{Mat}_1(\mathcal{Y}^{q+1}) \quad (44)$$



for every $q \geq 0$. Consequently, by Equation 44, we have

$$\begin{aligned} \|\mathcal{W}_1^{q+1} - \mathcal{W}_1^q\|_F^2 &= \|\text{Mat}_1(\mathcal{W}_1^{q+1}) - \text{Mat}_1(\mathcal{W}_1^q)\|_F^2 \\ &\leq (\alpha_1 \sigma_1(H^\top H) + \gamma)^2 \|\text{Mat}_1(\mathcal{Y}^{q+1}) \\ &\quad - \text{Mat}_1(\mathcal{Y}^q)\|_F^2 \\ &= (\alpha_1 \sigma_1(H^\top H) + \gamma)^2 \|\mathcal{Y}^{q+1} - \mathcal{Y}^q\|_F^2. \end{aligned} \quad (45)$$

Similarly, we have

$$\text{Mat}_1(\mathcal{W}_2^{q+1}) = (\alpha_2 K^\top K + \gamma I) \text{Mat}_2(\mathcal{Z}^{q+1}) \quad (46)$$

for every $q \geq 0$, which implies

$$\|\mathcal{W}_2^{q+1} - \mathcal{W}_2^q\|_F^2 \leq (\alpha_2 \sigma_1(K^\top K) + \gamma)^2 \|\mathcal{Z}^{q+1} - \mathcal{Z}^q\|_F^2. \quad (47)$$

Combining Equations (Inequalities) 43 with 45 and 47, we obtain

$$\begin{aligned} \mathcal{L}(\Theta^{q+1}) &+ \frac{\lambda + \beta_1 + \beta_2}{2} \|\mathcal{X}^{q+1} - \mathcal{X}^q\|_F^2 + \frac{1}{2} \|\mathcal{U}^{q+1} - \mathcal{U}^q\|_F^2 \\ &+ \frac{1}{2} \|\mathcal{V}^{q+1} - \mathcal{V}^q\|_F^2 + \frac{\xi_1}{2\beta_1} \|\mathcal{Y}^{q+1} - \mathcal{Y}^q\|_F^2 + \frac{\xi_2}{2\beta_2} \|\mathcal{Z}^{q+1} - \mathcal{Z}^q\|_F^2 \\ &\leq \mathcal{L}(\Theta^q). \end{aligned} \quad (48)$$

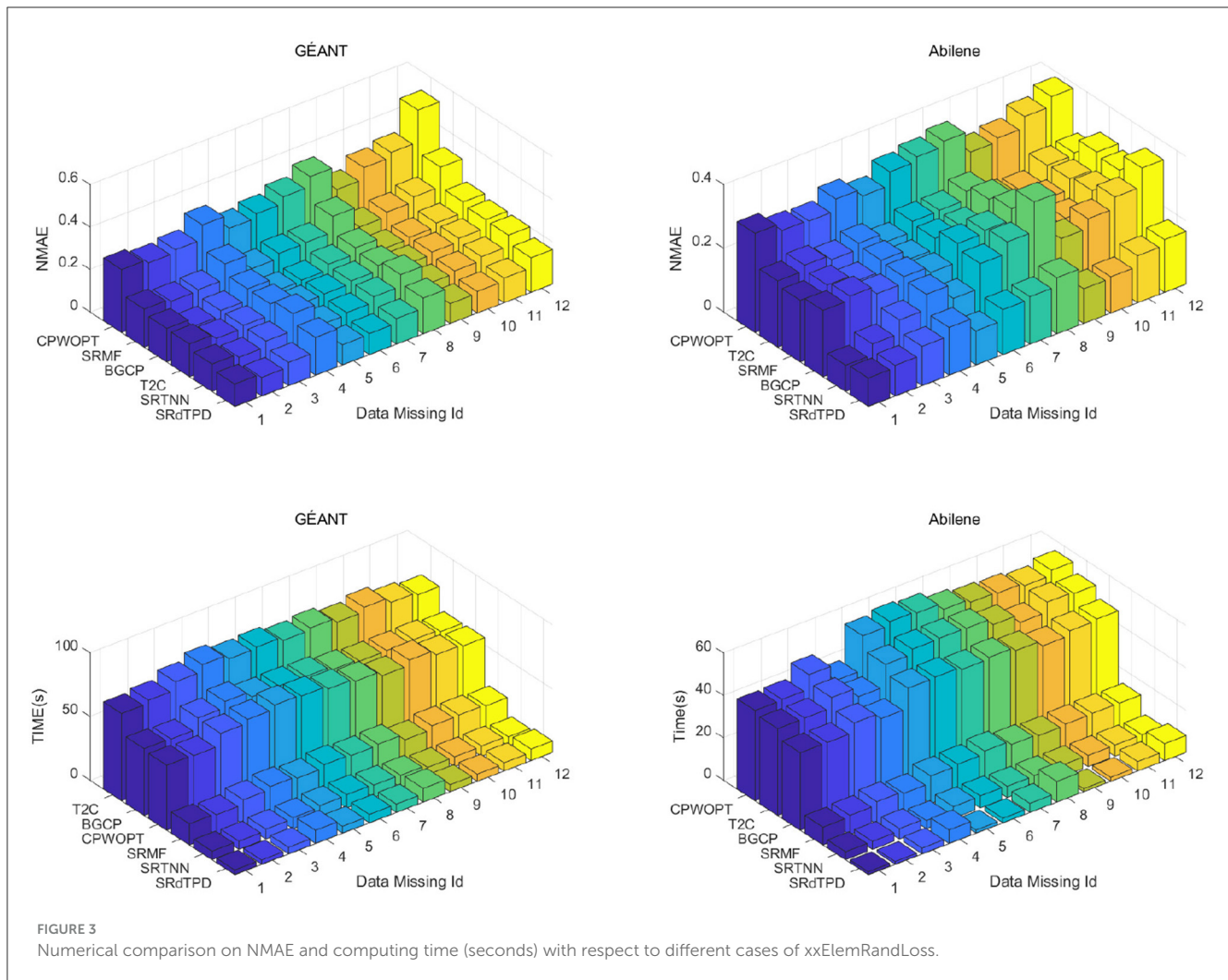
Finally, by the definition of μ , we know that Equation (Inequality) 40 holds.

Now, we prove the boundedness of $\{\Theta^q\}_{q=0}^\infty$. Because $\|\mathcal{W}_1^q\|_F^2 \leq (\alpha_1 \sigma_1(H^\top H) + \gamma)^2 \|\mathcal{Y}^q\|_F^2$ and $\|\mathcal{W}_2^q\|_F^2 \leq (\alpha_2 \sigma_1(K^\top K) + \gamma)^2 \|\mathcal{Z}^q\|_F^2$ for any $q \geq 0$, by Equation 6, we can obtain

$$\begin{aligned} \mathcal{L}(\Theta^q) &\geq f(\mathcal{X}^q, \mathcal{U}^q, \mathcal{V}^q, \mathcal{Y}^q, \mathcal{Z}^q) + \frac{\beta_1}{2} \|\mathcal{X}^q - \mathcal{Y}^q\|_F^2 + \frac{1}{\beta_1} \|\mathcal{W}_1^q\|_F^2 \\ &\quad - \frac{(\alpha_1 \sigma_1(H^\top H) + \gamma)^2}{2\beta_1} \|\mathcal{Y}^q\|_F^2 \\ &\quad + \frac{\beta_2}{2} \|\mathcal{X}^q - \mathcal{Z}^q\|_F^2 + \frac{1}{\beta_2} \|\mathcal{W}_2^q\|_F^2 - \frac{(\alpha_2 \sigma_1(K^\top K) + \gamma)^2}{2\beta_2} \|\mathcal{Z}^q\|_F^2 \end{aligned} \quad (49)$$

for every $q \geq 0$. According to Equation (Inequality) 40, for any $q \geq 0$, it is clear that

$$\mathcal{L}(\Theta^q) + \mu \sum_{s=0}^{q-1} \|\Lambda^{s+1} - \Lambda^s\|_F^2 \leq \mathcal{L}(\Theta^0), \quad (50)$$



Therefore, by the definition of f , as well as Equation 6 and Equation (Inequalities) 49, 50, it holds that for any $q \geq 0$,

$$\|\mathcal{U}^q\|_F^2 \leq 2\mathcal{L}(\Theta^0), \|\mathcal{V}^q\|_F^2 \leq 2\mathcal{L}(\Theta^0), \eta_1 \|\mathcal{Y}^q\|_F^2 \leq 2\beta_1 \mathcal{L}(\Theta^0) \text{ and } \eta_2 \|\mathcal{Z}^q\|_F^2 \leq 2\beta_2 \mathcal{L}(\Theta^0). \quad (51)$$

Since $\eta_1 > 0, \eta_2 > 0$, we know that $\{(\mathcal{U}^q, \mathcal{V}^q, \mathcal{Y}^q, \mathcal{Z}^q)\}_{q=0}^\infty$ is bounded, which implies the boundedness of $\{(\mathcal{W}_1^q, \mathcal{W}_2^q)\}_{q=0}^\infty$ from Equations 44, 46. In addition, since $\beta_1 \|\mathcal{X}^q - \mathcal{Y}^q + (1/\beta_1)\mathcal{W}_1^q\|_F^2 \leq 2\mathcal{L}(\Theta^0)$ for any $q \geq 0$, which from Equation (Inequality) 49, it follows that $\{\mathcal{X}^q\}_{q=0}^\infty$ is also bounded. Based on the comprehensive analysis presented above, we have completed the proof.

Remark 3. Given the predefined dimensions of matrices H and K , the singular values of $H^T H$ and $K^T K$ can be computed. Consequently, we only need to select appropriate parameters that satisfy the hypotheses stated in Theorem 2.

Theorem 3. For the sequence $\{\Theta^q\}_{q=0}^\infty$ generated by Algorithm 1. If the parameters satisfy the conditions in Theorem 2, then

any cluster point Θ^∞ of $\{\Theta^q\}_{q=0}^\infty$ is a stationary point of Equation (Model) 5.

Proof: According to Theorem 2, the sequence $\{\Theta^q\}_{q=0}^\infty$ is bounded. Let $\{\Theta^{q_i}\}_{i=0}^\infty$ be a convergent subsequence such that $\lim_{i \rightarrow \infty} \Theta^{q_i} = \Theta^\infty := (\mathcal{X}^\infty, \mathcal{U}^\infty, \mathcal{V}^\infty, \mathcal{Y}^\infty, \mathcal{Z}^\infty, \mathcal{W}_1^\infty, \mathcal{W}_2^\infty)$. First, from Equation (Inequality) 50, we have

$$\mu \sum_{q=0}^{q_i} \|\Lambda^{q+1} - \Lambda^q\|_F^2 \leq \mathcal{L}(\Theta^0) - \mathcal{L}(\Theta^{q_i}). \quad (52)$$

Taking the limit of both sides of the Equation 52 as i approaches ∞ , it holds that

$$\mu \sum_{q=0}^{\infty} \|\Lambda^{q+1} - \Lambda^q\|_F^2 \leq \mathcal{L}(\Theta^0) - \mathcal{L}(\Theta^\infty) \leq \infty, \quad (53)$$

which means $\lim_{q \rightarrow \infty} \|\Lambda^{q+1} - \Lambda^q\|_F = 0$. Moreover, from the optimality conditions of $\mathcal{X}$ -subproblem to $\mathcal{Z}$ -subproblem, as well

TABLE 1 Results on NMAE and computing time (seconds) with respect to diverse cases of xxElemSyncLoss.

Dataset	xx	q	SRMF		CPWOPT		T2C		SRTNN		BGCP		SRdTPD	
			NMAE	Time	NMAE	Time	NMAE	Time	NMAE	Time	NMAE	Time	NMAE	Time
GÉANT	25	30	0.232	21.56	0.355	69.25	0.162	93.25	0.126	10.64	0.152	66.99	0.097	2.64
		50	0.210	19.70	0.348	64.98	0.178	85.53	0.152	9.31	0.187	63.25	0.121	2.84
		70	0.238	19.02	0.347	62.27	0.173	85.26	0.163	8.84	0.177	62.59	0.124	5.35
		90	0.480	18.77	0.373	62.31	0.186	82.87	0.202	8.81	0.198	61.21	0.173	16.70
	50	30	0.223	22.72	0.329	70.23	0.155	94.54	0.126	11.41	0.156	69.24	0.100	2.82
		50	0.224	22.78	0.329	72.33	0.158	93.57	0.139	12.28	0.171	70.73	0.110	4.23
		70	0.448	23.20	0.345	72.25	0.148	91.18	0.150	12.10	0.172	70.06	0.112	7.89
		90	0.688	18.67	0.469	59.63	0.198	80.37	0.218	9.89	0.249	61.64	0.177	16.74
	75	30	0.364	20.37	0.359	68.11	0.146	84.66	0.121	10.53	0.148	62.71	0.092	2.94
		50	0.503	20.12	0.421	68.11	0.159	84.66	0.136	10.53	0.178	62.71	0.105	2.94
		70	0.688	20.29	0.462	65.46	0.155	80.70	0.147	10.53	0.226	62.34	0.116	8.56
		90	0.807	15.82	0.683	58.78	0.199	74.38	0.204	9.06	0.288	58.59	0.177	14.24
Abilene	25	30	0.219	11.22	0.306	59.76	0.225	53.93	0.231	5.99	0.234	47.10	0.132	3.35
		50	0.199	12.57	0.302	61.45	0.205	54.05	0.267	5.89	0.217	48.17	0.119	3.26
		70	0.241	13.32	0.325	62.76	0.228	55.02	0.291	5.94	0.237	48.76	0.136	3.60
		90	0.416	8.82	0.347	52.04	0.233	44.54	0.428	5.43	0.261	41.99	0.168	2.71
	50	30	0.236	11.52	0.340	56.86	0.211	49.41	0.262	5.51	0.211	44.93	0.115	3.80
		50	0.228	12.42	0.330	57.76	0.215	52.12	0.282	5.69	0.215	47.15	0.117	3.56
		70	0.426	11.61	0.379	55.83	0.245	50.32	0.326	5.44	0.271	56.14	0.150	3.87
		90	0.559	11.20	0.419	56.35	0.278	50.60	0.502	5.37	0.327	46.06	0.190	3.25
	75	30	0.344	11.51	0.349	65.25	0.211	56.38	0.277	6.00	0.210	49.40	0.115	4.56
		50	0.384	10.81	0.349	54.24	0.207	46.30	0.304	5.29	0.221	43.12	0.115	3.56
		70	0.635	10.24	0.398	54.70	0.243	45.12	0.353	5.34	0.277	53.18	0.147	3.39
		90	0.794	11.01	0.423	59.14	0.272	53.40	0.544	5.29	0.390	46.50	0.175	3.56

Bold values are the optimal values.

as the update of $\mathcal{W}_1$ and $\mathcal{W}_2$, it is easy to see that

$$\begin{cases}
 \nabla_{\mathcal{X}_{\Omega^c}} f(\mathcal{X}_{i+1}^{q_i+1}, \mathcal{U}^{q_i}, \mathcal{V}^{q_i}, \mathcal{Z}^{q_i}) + (\mathcal{W}_1^{q_i})_{\Omega^c} + \beta_1(\mathcal{X}_{i+1}^{q_i+1} - \mathcal{Y}^{q_i})_{\Omega^c} \\
 + (\mathcal{W}_2^{q_i})_{\Omega^c} + \beta_1(\mathcal{X}_{i+1}^{q_i+1} - \mathcal{Z}^{q_i})_{\Omega^c} = \mathcal{O}_{\Omega^c}, \\
 (\mathcal{X}_{i+1}^{q_i+1})_{\Omega} = (\mathcal{M})_{\Omega}, \\
 \nabla_{\mathcal{U}} f(\mathcal{X}_{i+1}^{q_i+1}, \mathcal{U}^{q_i+1}, \mathcal{V}^{q_i}, \mathcal{Z}^{q_i}) = \mathcal{O}, \\
 \nabla_{\mathcal{V}} f(\mathcal{X}_{i+1}^{q_i+1}, \mathcal{U}^{q_i+1}, \mathcal{V}^{q_i+1}, \mathcal{Z}^{q_i}) = \mathcal{O}, \\
 \nabla_{\mathcal{Z}} f(\mathcal{X}_{i+1}^{q_i+1}, \mathcal{U}^{q_i+1}, \mathcal{V}^{q_i+1}, \mathcal{Z}^{q_i+1}) - \mathcal{W}_1^{q_i} = \mathcal{O}, \\
 \nabla_{\mathcal{Z}} f(\mathcal{X}_{i+1}^{q_i+1}, \mathcal{U}^{q_i+1}, \mathcal{V}^{q_i+1}, \mathcal{Z}^{q_i+1}) - \mathcal{W}_2^{q_i} = \mathcal{O}, \\
 \mathcal{W}_1^{q_i+1} = \mathcal{W}_1^{q_i} + \beta_1(\mathcal{X}_{i+1}^{q_i+1} - \mathcal{Y}^{q_i+1}), \\
 \mathcal{W}_2^{q_i+1} = \mathcal{W}_2^{q_i} + \beta_1(\mathcal{X}_{i+1}^{q_i+1} - \mathcal{Z}^{q_i+1}).
 \end{cases} \quad (54)$$

Since $\lim_{i \rightarrow \infty} \|\Lambda^{q_i+1} - \Lambda^{q_i}\|_F = 0$ and $\lim_{i \rightarrow \infty} \Theta^{q_i} = \Theta^\infty$, it holds that $\lim_{i \rightarrow \infty} \|\mathcal{W}_1^{q_i+1} - \mathcal{W}_1^{q_i}\|_F = 0$ from Equation (Inequality) 45. Similarly, we have $\lim_{i \rightarrow \infty} \|\mathcal{W}_2^{q_i+1} - \mathcal{W}_2^{q_i}\|_F = 0$ from Equation (Inequality) 47. Consequently, by letting $i \rightarrow \infty$ in Equation 54, we know that Θ^∞ is a stationary point of Equation (Model) 5. We complete the proof.

5 Numerical experiments

In this section, we apply our approach (SRdTPD) to two authentic Internet traffic datasets: the GÉANT dataset [20], which logs Internet traffic data from 23 original nodes to 23 destination nodes every 15 min, and the Abilene¹ dataset, which records Internet traffic data from 11 original nodes to 11 destination nodes every 5 min. For both datasets, we select data spanning a period of 7 days. The sizes of these two datasets, arranged in fourth-order tensor format, are $7 \times 96 \times 23 \times 23$ and $7 \times 288 \times 11 \times 11$, respectively. For each experiment situation, we reconstruct the Internet traffic data 5 times, and all experimental results presented are the average of five repetitions. As demonstrated by the recovery test results for synthetic experiments and many real-world applications [14], algorithms with the discrete Fourier transform not only exhibit similar recovery accuracy but also incur lower computational costs, compared to methods utilizing the discrete cosine transform or random orthogonal transform. It is well-known that the most basic

1 <http://abilene.internet2.edu/observatory/data-collections.html>

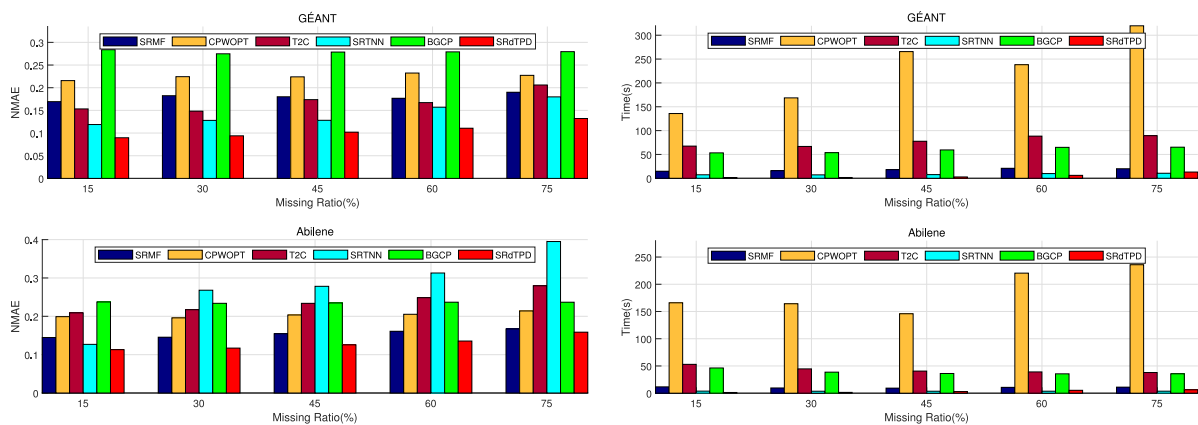


FIGURE 4
Numerical comparison on NMAE and computing time (seconds) with respect to diverse missing ratio of RowRandLoss.

Fast Fourier Transform (FFT) has a computational complexity of $O(N \log N)$ [21]. Using this result, it is not difficult to see that the computational complexity of SRdTPD using FFT is $O((n_1 + n_2)(n_1 n_2 + \text{slog}(n_3 n_4)) n_3 n_4)$, which is significantly lower than that of methods using other transforms. Considering the potential computational cost advantage of SRdTPD equipped with FFT when processing large-scale data, we choose normalized discrete Fourier transform as L_3 and L_4 in L , i.e.,

$$L_3 = L_4 = \frac{1}{\sqrt{N}} \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & w & w^2 & \dots & w^{N-1} \\ 1 & w^2 & w^4 & \dots & w^{2(N-1)} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & w^{N-1} & w^{2(N-1)} & \dots & w^{(N-1)^2} \end{bmatrix}$$

with the $w = e^{-i\frac{2\pi}{N}}$. All the experiments are performed on a notebook computer (13th Gen Intel(R) Core(TM) i9-13900HX CPU @2.20 GHz with 24G memory) running a 64-bit Windows Operating System. The code of SRdTPD is released at <https://github.com/Duan-Yuxuan/Traffic-data-recovery/tree/master>.

To verify the performance of SRdTPD, we compare it with five existing approaches designed for data recovery, the first of which is a matrix completion method, and the remaining four are tensor completion methods:

- **SRMF**, which is a low-rank matrix completion with a spatio-temporal regularization [4];
- **CPWOPT**, which is a tensor CP factorization completion approach [22];
- **T2C**, which is a tensor filling approach that decomposes third-order tensors into three third-order low-rank tensors in a balanced manner [23];
- **SRTNN**, which is a tensor-based approach with spatio-temporal regularization terms [11];
- **BGCP**, which is a Bayesian Gaussian CP tensor decomposition approach [24].

We use the Normalized Mean Absolute Error (NMAE) to measure the quality of the recovered data by models and algorithms, where the NMAE is defined as follows

$$\text{NMAE} = \frac{\sum_{(i,j,k,l) \notin \Omega} |(\mathcal{X}_{\text{true}})_{ijkl} - (\mathcal{X}_{\text{rec}})_{ijkl}|}{\sum_{(i,j,k,l) \notin \Omega} |(\mathcal{X}_{\text{true}})_{ijkl}|},$$

where $\mathcal{X}_{\text{true}}$ and $\mathcal{X}_{\text{rec}}$ are original data and recovered data, respectively. Clearly, lower NMAE value means better quality of the recovered data. We first consider the recovery effectiveness of various approaches for missing data with data sample rates ranging from 10% to 90%. Under random missingness in the GÉANT dataset, we set $\alpha_1 = 0.1, \alpha_2 = 200, \beta_1 = 0.01, \alpha_2 = 10, \lambda = 0.001, \gamma = 0.01$, while for the Abilene dataset, we choose $\alpha_1 = \beta_2 = 0.01, \alpha_2 = 10, \beta_1 = \lambda = \gamma = 0.001$. Figure 1 presents the NMAE values for missing data recovery using various methods at different sampling rates. It can be visually observed that SRdTPD outperforms other methods in terms of data recovery effectiveness, particularly as the missing data rate increases, where the superiority of our method's recovery performance becomes increasingly apparent.

In addition to the recovery of data missing at random, we next consider the following four types of structural missingness scenarios, which are given in [4], thereby demonstrating that our method fully exploits the structural information within the data.

- **xxTimeRandLoss** This scenario simulates the phenomenon of structured data absence at specific time points influenced by certain factors, such as the utilization of unreliable data transmission equipment. Specifically, these data are missing at these specific time points in a certain random proportion. In our simulation, we randomly select $xx\%$ of the matrix slices along the last two modes of the Internet traffic tensor $\mathcal{X}$ of size $n_1 \times n_2 \times n_3 \times n_4$, and subsequently, we further randomly delete $q\%$ of its elements in each selected matrix.
- **xxElemRandLoss** This scenario simulates the structured absence of data for specific OD (Origination-Destination) nodes under the influence of certain factors, such as the use

of unreliable data transmission equipment. In this context, OD node data are missing in a certain random proportion. To simulate this phenomenon, we randomly select $xx\%$ of the matrix slices along the first two modes of the Internet traffic tensor $\mathcal{X}$ of size $n_1 \times n_2 \times n_3 \times n_4$ and further randomly delete $q\%$ of the elements in each selected matrix.

- **xxElemSyncLoss** This scenario simulates the structured data missingness for specific OD nodes due to a uniform underlying cause, resulting in temporally synchronized missingness among these OD nodes. To emulate this condition, we randomly select $xx\%$ of the slices from the Internet traffic tensor $\mathcal{X}$ of size $n_1 \times n_2 \times n_3 \times n_4$, which are matrix representations expanded along its first two modes. Subsequently, for each selected matrix, we randomly choose a common set of time indices ($q\%$) and delete the corresponding elements at these indices.
- **RowRandLoss** With flow level measurements, data are collected by a router. If a router is unable to collect data for a period of time, all data collected during those specific time points will be missing. To simulate this scenario, we randomly select $p\%$ of the time points within a day and delete all data recorded at those time points across a 7-day period.

Under these scenarios of structural missingness, with $\beta_1 = \gamma = 0.01$, the parameters for the first two structural missingness cases are identical within the same dataset (GÉANT: $\alpha_1 = 0.1, \alpha_2 = 200, \beta_2 = 10, \lambda = 0.001$; Abilene: $\alpha_1 = 0.1, \alpha_2 = 2, \beta_2 = 1, \lambda = 0.001$), while the parameters for the last two structural missingness cases are consistent within the same dataset (GÉANT: $\alpha_1 = 0.01, \alpha_2 = 2, \beta_2 = 1, \lambda = 0.01$; Abilene: $\alpha_1 = 0.01, \alpha_2 = 1, \beta_2 = 0.1, \lambda = 0.01$). Among these four types of structured missingness scenarios, we simulate the first three by considering the following 12 specific missingness cases: Missing Id 1-4: Set $xx = 25$ and, correspondingly, set $q = 30, 50, 70$, and 90 , respectively; Missing Id 5-8: Set $xx = 50$ and, correspondingly, set $q = 30, 50, 70$, and 90 , respectively; Missing Id 9-12: Set $xx = 75$ and, correspondingly, set $q = 30, 50, 70$, and 90 , respectively. For the fourth type of structured missingness, we specifically select $p = 15, 30, 45, 50$, and 75 .

The recovery performance and computational time of different methods under distinct structural missingness scenarios are comprehensively presented in [Figures 2–4](#), [Table 1](#). Experimental results demonstrate that the proposed SRdTPD approach achieves substantial improvements in handling structurally missing data. Regarding recovery accuracy, SRdTPD demonstrates superior performance compared to listed methods, attaining competitive results across all test cases. In terms of computational efficiency, SRdTPD maintains favorable time — while a moderate increase in computational time is observed under extreme missingness conditions (attributable to iterative process requirements), the runtime remains within practical thresholds for real-world applications. These findings collectively suggest that SRdTPD effectively balances recovery accuracy and computational demands, providing a robust solution for structural missing data recovery tasks.

6 Conclusion

This study establishes an equivalence relationship between the d -th order tensor nuclear norm (TNN) in the unitary transformation domain and the squared sum of the Frobenius norms of its two factorization factors. Based on this relationship, we constructed a novel low-rank recovery method for Internet traffic data that effectively incorporates the spatio-temporal characteristics of traffic data while significantly reducing computational time. Auxiliary variables are introduced into the model solution process, which is then solved using the ADMM. Numerical experiments demonstrate that the proposed method exhibits significant advantages in the recovery of Internet traffic data, particularly in cases of structural missingness.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

YD: Investigation, Methodology, Writing – original draft. CL: Validation, Writing – review & editing. JL: Validation, Writing – review & editing. XY: Validation, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. CL's work was supported in part by National Natural Science Foundation of China (No. 11971138). JL's work was supported by the Science and Technology Research Program of Chongqing Municipal Education Commission (Grant No. KJZD-K202200506), by Chongqing Talent Program Lump-sum Project (No. cstc2022ycjh-bgzxm0040), and by the Foundation of Chongqing Normal University (22XLB005, ncamc2022-msxm02). XY's work was supported by the Major Program of National Natural Science Foundation of China (Nos. 11991020, 11991024).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated

organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Roughan M, Thorup M, Zhang Y. Traffic engineering with estimated traffic matrices. In: *Proceedings of the 3rd ACM SIGCOMM Conference on Internet Measurement*. Miami Beach, FL: Association for Computing Machinery (2003). p. 248–258.
2. Cahn RS. Wide area network design: concepts and tools for optimization. *IEEE Commun Mag.* (2000) 38:28–30. doi: 10.1109/MCOM.2000.867843
3. Zhang Y, Ge Z, Greenberg A, Roughan M. Network anomography. In: *Proceedings of the 5th ACM SIGCOMM conference on Internet Measurement*. Berkeley, CA: USENIX Association (2005). p. 30.
4. Zhang Y, Roughan M, Willinger W, Qiu L. Spatio-temporal compressive sensing and internet traffic matrices. In: *Proceedings of the ACM SIGCOMM 2009 Conference on Data Communication*. Barcelona: Association for Computing Machinery (2009). p. 267–278.
5. Carroll JD, Chang JJ. Analysis of individual differences in multidimensional scaling via an N-way generalization of “Eckart-Young” decomposition. *Psychometrika.* (1970) 35:283–319. doi: 10.1007/BF02310791
6. Harshman RA, et al. Foundations of the PARAFAC procedure: Models and conditions for an “explanatory” multi-modal factor analysis. In: *UCLA Work Papers Phonetics*. Ann Arbor; Los Angeles, CA: University Microfilms (1970). p. 84.
7. Oseledets IV. Tensor-train decomposition. *SIAM J Sci Comput.* (2011) 33:2295–317. doi: 10.1137/090752286
8. Kilmer ME, Martin CD. Factorization strategies for third-order tensors. *Linear Algebra Appl.* (2011) 435:641–58. doi: 10.1016/j.laa.2010.09.020
9. Candès E, Recht B. Exact matrix completion via convex optimization. *Found Comput Math.* (2009) 9:717–72. doi: 10.1007/s10208-009-9045-5
10. Zhang Z, Ely G, Aeron S, Hao N, Kilmer M. Novel methods for multilinear data completion and de-noising based on tensor-SVD. In: *2014 IEEE Conference on Computer Vision and Pattern Recognition*. Columbus, OH: IEEE (2014). p. 3842–3849.
11. Li C, Chen Y, Li D. Internet traffic tensor completion with tensor nuclear norm. *Comput Optim Appl.* (2024) 87:1033–57. doi: 10.1007/s10589-023-00545-5
12. He H, Ling C, Xie W. Tensor completion via a generalized transformed tensor T-product decomposition without t-SVD. *J Sci Comput.* (2022) 93:47. doi: 10.1007/s10915-022-02006-3
13. Martin CD, Shafer R, LaRue B. An order- p tensor factorization with applications in imaging. *SIAM J Sci Comput.* (2013) 35:A474–90. doi: 10.1137/110841229
14. Qin W, Wang H, Zhang F, Wang J, Luo X, Huang T. Low-rank high-order tensor completion with applications in visual data. *IEEE Trans Image Process.* (2022) 31:2433–48. doi: 10.1109/TIP.2022.3155949
15. Kolda TG, Bader BW. Tensor decompositions and applications. *SIAM Rev.* (2009) 51:455–500. doi: 10.1137/07070111X
16. Recht B, Fazel M, Parrilo PA. Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization. *SIAM Rev.* (2010) 52:471–501. doi: 10.1137/070697835
17. Chen C, He B, Ye Y, Yuan X. The direct extension of ADMM for multi-block convex minimization problems is not necessarily convergent. *Math Program.* (2016) 155:57–79. doi: 10.1007/s10107-014-0826-5
18. Popa J, Lou Y, Minkoff SE. Low-rank tensor data reconstruction and denoising via ADMM: algorithm and convergence analysis. *J Sci Comput.* (2023) 97:49. doi: 10.1007/s10915-023-02364-6
19. Beck A. *First-Order Methods in Optimization*. Philadelphia: SIAM. (2017). doi: 10.1137/1.9781611974997
20. Uhlig S, Quoitin B, Lepropre J, Balon S. Providing public intradomain traffic matrices to the research community. *ACM SIGCOMM Comput Commun Rev.* (2006) 36:83–6. doi: 10.1145/1111322.1111341
21. Cooley JW, Tukey JW. An algorithm for the machine calculation of complex Fourier series. *Math Comp.* (1965) 19:297–301. doi: 10.1090/S0025-5718-1965-0178586-1
22. Acar E, Dunlavy DM, Kolda TG, Mørup M. Scalable tensor factorizations for incomplete data. *Chemom Intell Lab Syst.* (2011) 106:41–56. doi: 10.1016/j.chemolab.2010.08.004
23. Chen X, He Z, Chen Y, Lu Y, Wang J. Missing traffic data imputation and pattern discovery with a Bayesian augmented tensor factorization model. *Transp Res Part C: Emerg Technol.* (2019) 104:66–77. doi: 10.1016/j.trc.2019.03.003
24. Chen X, He Z, Sun L, A. Bayesian tensor decomposition approach for spatiotemporal traffic data imputation. *Transp Res Part C: Emerg Technol.* (2019) 98:73–84. doi: 10.1016/j.trc.2018.11.003