Numerical Algorithms (2021) 87:1767–1797
https://doi.org/10.1007/s11075-021-01074-0

ORIGINAL PAPER

Numerical study on Moore-Penrose inverse of tensors via Einstein product

Baohua Huang[1,2]

Received: 31 December 2019 / Accepted: 19 January 2021 / Published online: 24 March 2021

Abstract

The notation of Moore-Penrose inverse of matrices has been extended from matrix space to even-order tensor space with Einstein product. In this paper, we give the numerical study on the Moore-Penrose inverse of tensors via the Einstein product. More precisely, we transform the calculation of Moore-Penrose inverse of tensors via the Einstein product into solving a class of tensor equations via the Einstein product. Then, by means of the conjugate gradient method, we obtain the approximate Moore-Penrose inverse of tensors via the Einstein product. Finally, we report some numerical examples to show the efficiency of the proposed methods and testify the conclusion suggested in this paper.

Keywords Moore-Penrose inverse · Einstein product · Tensor equations · Conjugate gradient method

1 Introduction

1.1 Background and motivation

Tensor is a multidimensional array, which is exactly the higher order generalization of matrix. Research on tensors has been very active recently [5, 6, 9, 16, 21] as tensors have a lot of applications in many different fields, such as graph analysis, computer vision, signal processing, data mining, human motion recognition, and chemometrics; one can see [4, 7, 8, 10, 11, 31, 33] and the references cited

✉ Baohua Huang
baohuahuang@126.com

1 School of Mathematical Sciences, South China Normal University, Guangzhou 510631, People's Republic of China

2 College of Mathematics and Informatics & FJKLMAA, Fujian Normal University, Fuzhou 350117, People's Republic of China

therein. Tensor models are employed in numerous disciplines addressing the problem of finding the multilinear structure in multiway data sets. Especially, multilinear systems model many phenomena in mechanics, physics, Markov process, control theory, partial differential equations, and engineering problems [13–15, 17–20, 32], such as the radiation transfer equation, high-dimensional Possion equation, Einstein gravitational field equation, and piezoelectric effect equation.

A lot of works have been done on tensors for the past four decades. However, there are few research works on the theory and applications of the generalized inverse of tensors. In 2016, Sun et al. [32] first introduced a generalized inverse called the Moore-Penrose inverse of an even-order tensor via the Einstein product, and a concrete representation for the Moore-Penrose inverse was given by using the singular value decomposition (SVD) of a tensor. Then, the authors obtained the minimum-norm least-squares solution of some multilinear systems by means of the Moore-Penrose inverse of tensors. Behera and Mishra [3] continued the same study and proposed some different types of generalized inverses of tensors. Different from the work in [3, 32], Jin et al. [24] defined a Moore-Penrose inverse of higher order tensors by using the t-product constructed in [27]. They proved that the Moore-Penrose inverse of an arbitrary tensor exists and is unique by using the technique of fast Fourier transform. They also presented some properties of the Moore-Penrose inverse of tensors and established some representations of {1}-inverses, {1, 3}-inverses and {1, 4}-inverses of tensors.

In [3], Behera and Mishra proposed two open problems. The first one is about the reverse-order law for the Moore-Penrose inverse of tensors and the second one is about full rank factorization of tensors. In 2018, Panigrahy et al. [28] answered the first problem for even-order tensors. Again, Panigrahy and Mishra [29] added more results to the same theory but for arbitrary order tensors. In the same year, Liang and Zheng [25] solved the other problem and proposed the full rank decomposition of tensors by introducing a new kind of tensor rank, which gives a novel representation of the Moore-Penrose inverse of tensors. They also obtained several formulae related to the Moore-Penrose inverse of tensors and their $\{i, j, k\}$-inverses.

In 2017, Ji and Wei [22] studied another extension of the Moore-Penrose inverse of an even-order tensor, called weighted Moore-Penrose inverse, and established the relation between the minimum-norm least-squares solution of a multilinear system and the weighted Moore-Penrose inverse of tensors. Behera et al. [2] further studied the weighted Moore-Penrose inverse of tensors and proposed a few identities involving the weighted Moore-Penrose inverse of tensors. They also obtained some necessary and sufficient conditions of the reverse-order law for the weighted Moore-Penrose inverse of arbitrary-order tensors. In 2018, Panigrahy and Mishra [30] improved the definition of Moore-Penrose inverse of an even-order tensor to a tensor of any order, which is called product Moore-Penrose inverse. A necessary and sufficient condition for the coincidence of the Moore-Penrose inverse and the product Moore-Penrose inverse is also proposed. At the same year, Ji and Wei [23] extended the notion of Drazin inverse of a square matrix to an even-order square tensor, and obtained an expression for the Drazin inverse through the core-nilpotent decomposition for a tensor of even-order. As an application, the Drazin inverse solution of the singular linear tensor equation $\mathcal{A} *_N \mathcal{X} = \mathcal{B}$ has been included. Ma et al. [26] derived some representations of tensor expressions involving the Moore-Penrose inverse and investigated the perturbation theory for the Moore-Penrose inverse of tensors via the Einstein product.

Since the Moore-Penrose inverse of tensors plays an important role in solving multilinear systems [3, 24, 32], it needs to further study this subject. The definitions of Moore-Penrose inverse are based on the singular value decomposition (SVD) of tensors, t-product of tensors, and full rank decomposition of tensors. All the existing works mainly analyze the properties of the Moore-Penrose inverse of tensors from an analytical perspective. The research work about the algorithm for computing the Moore-Penrose inverse of tensors is relatively rare. Although Jin et al. [24] established an algorithm for computing the Moore-Penrose inverse of a tensor by using the technique of fast Fourier transform and inverse fast Fourier transform of a matrix, the algorithm involves matrix computation. Naturally, a question arises: Does there exists an iterative algorithm, which involves only tensor computation, for computing the Moore-Penrose inverse of tensors?

In this paper, we will answer this question. This study can lead to the enhancement of the computation of Moore-Penrose inverse of tensors along with solutions of multilinear structure in multidimensional systems. In this regard, we reformulate the calculation of Moore-Penrose inverse of real tensors as the solution of linear tensor equation or equations. For the linear tensor equation or equations, we apply the conjugate gradient method to find the solution.

The main contribution of this paper is as follows:

- to give the new equivalent characterization of Moore-Penrose inverse of tensors via the Einstein product.
- to present two numerical methods to obtain the Moore-Penrose inverse of tensors via the Einstein product.

1.2 Outline

This paper is organized as follows. In the next subsection, we present some notations and definitions which will be useful for proving the main results in Sections 2 and 3. In Section 2, we first transform the computation of $\{1,3\}$-inverse and $\{1,4\}$-inverse of a tensor $\mathcal{A}$ into solving a class of tensor equations with one variable. Then, we develop the conjugate gradient method for the solution of tensor equation in order to obtain $\mathcal{A}^{(1,4)}$ and $\mathcal{A}^{(1,3)}$. Meanwhile, we give the convergence analysis and prove the finite termination property of the proposed method. By means of the numerical results of $\{1,4\}$-inverse $\mathcal{A}^{(1,4)}$ and $\{1,3\}$-inverse $\mathcal{A}^{(1,3)}$, and the relationship $\mathcal{A}^{\dagger} = \mathcal{A}^{(1,4)} *_N \mathcal{A} *_M \mathcal{A}^{(1,3)}$, we obtain the unique Moore-Penrose inverse $\mathcal{A}^{\dagger}$. In Section 3, we first derive an important property of the Moore-Penrose inverse of tensors. Based on the property of the Moore-Penrose inverse of tensors, we transform the computation of Moore-Penrose inverse of tensors into solving a class of tensor equations with two variables. Then, we present the conjugate gradient method to solve the tensor equation and give the convergence analysis. Numerical examples of the proposed algorithm are shown and analyzed in Section 4. The paper ends up with some concluding remarks in Section 5.

1.3 Notations and definitions

For a positive integer N, let $[N] = \{1, \cdots, N\}$. An order N tensor $\mathcal{A} = (a_{i_1\cdots i_N})_{1\le i_j\le I_j}(j = 1, \cdots, N)$ is a multidimensional array with $I_1 I_2 \cdots I_N$ entries. Let $\mathbb{C}^{I_1\times\cdots\times I_N}$ and $\mathbb{R}^{I_1\times\cdots\times I_N}$ be the sets of the order N dimension $I_1 \times \cdots \times I_N$ tensors over the complex number field $\mathbb{C}$ and the real number field $\mathbb{R}$, respectively. Here $I_1, I_2, \cdots, I_N$ are dimensions of the first, second, $\cdots$, Nth way, respectively. The order of a tensor is the number of dimensions. A first-order tensor is a vector while a second-order tensor is a matrix. Higher-order tensors are tensors of order three or higher. Each entry of $\mathcal{A}$ is denoted by $a_{i_1\cdots i_N}$. For $N = 3$, $\mathcal{A} \in \mathbb{C}^{I_1\times I_2\times I_3}$ is a third-order tensor, and $a_{i_1 i_2 i_3}$ denotes the entry of that tensor.

For a tensor $\mathcal{A} = (a_{i_1\cdots i_N j_1\cdots j_M}) \in \mathbb{C}^{I_1\times\cdots\times I_N\times J_1\times\cdots\times J_M}$, let $\mathcal{B} = (b_{i_1\cdots i_M j_1\cdots j_N}) \in \mathbb{C}^{J_1\times\cdots\times J_M\times I_1\times\cdots\times I_N}$ be the conjugate transpose of $\mathcal{A}$, where $b_{i_1\cdots i_M j_1\cdots j_N} = \overline{a}_{j_1\cdots j_N i_1\cdots i_M}$. The tensor $\mathcal{B}$ is denoted by $\mathcal{A}^H$. When $b_{i_1\cdots i_M j_1\cdots j_N} = a_{j_1\cdots j_N i_1\cdots i_M}$, the tensor $\mathcal{B}$ is called the transpose of $\mathcal{A}$, denoted by $\mathcal{A}^T$. We say a tensor $\mathcal{D} = (d_{i_1\cdots i_N j_1\cdots j_N}) \in \mathbb{C}^{I_1\times\cdots\times I_N\times I_1\times\cdots\times I_N}$ is a diagonal tensor if $d_{i_1\cdots i_N j_1\cdots j_N} = 0$ in the case that the indices $i_1\cdots i_N$ are different from $j_1\cdots j_N$. The unit tensor $\mathcal{I}$ is the one with all entries being zero except for the diagonal entries $d_{i_1\cdots i_N i_1\cdots i_N} = 1$. The zero tensor $\mathcal{O}$ is the one with all entries being zero. The trace of a tensor $\mathcal{A} \in \mathbb{R}^{I_1\times\cdots\times I_N\times I_1\times\cdots\times I_N}$ is $\mathrm{tr}(\mathcal{A}) = \sum_{i_1\cdots i_N} a_{i_1\cdots i_N i_1\cdots i_N}$ [4]. A few more notations and definitions are introduced below.

Let $\mathcal{A} \in \mathbb{C}^{I_1\times\cdots\times I_N\times K_1\times\cdots\times K_N}$ and $\mathcal{B} \in \mathbb{C}^{K_1\times\cdots\times K_N\times J_1\times\cdots\times J_M}$; the Einstein product [4] of tensors $\mathcal{A}$ and $\mathcal{B}$ is defined by the operation $*_N$ via

$$(\mathcal{A} *_N \mathcal{B})_{i_1\cdots i_N j_1\cdots j_M} = \sum_{k_1\cdots k_N} a_{i_1\cdots i_N k_1\cdots k_N} b_{k_1\cdots k_N j_1\cdots j_M}. \tag{1.1}$$

It is obvious that $\mathcal{A}*_N\mathcal{B} \in \mathbb{C}^{I_1\times\cdots\times I_N\times J_1\times\cdots\times J_M}$ and the associative law of this tensor product are true. By (1.1), if $\mathcal{B} \in \mathbb{C}^{K_1\times\cdots\times K_N}$, then

$$(\mathcal{A} *_N \mathcal{B})_{i_1\cdots i_N} = \sum_{k_1\cdots k_N} a_{i_1\cdots i_N k_1\cdots k_N} b_{k_1\cdots k_N},$$

where $\mathcal{A}*_N\mathcal{B} \in \mathbb{C}^{I_1\times\cdots\times I_N}$. This product is used in the study of the theory of relativity and in the area of continuum mechanics [17]. We comment here that the Einstein product $*_1$ reduces to the standard matrix multiplication as:

$$(A *_1 B)_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$$

for $A \in \mathbb{C}^{m\times n}$ and $B \in \mathbb{C}^{n\times l}$. We also know that when $\mathcal{A} \in \mathbb{C}^{I_1\times\cdots\times I_N}$ and $\mathcal{B}$ is a vector $b = (b_{i_N}) \in \mathbb{C}^{I_N}$, the Einstein product is defined by the operation $\times_1$ via

$$(\mathcal{A} \times_1 \mathcal{B})_{i_1\cdots i_{N-1}} = \sum_{i_N} a_{i_1\cdots i_N} b_{i_N},$$

where $\mathcal{A} \times_1 \mathcal{B} \in \mathbb{C}^{I_1\times\cdots\times I_{N-1}}$.

Lemma 1.1 [32] *Let* $\mathcal{A} \in \mathbb{C}^{I_1\times\cdots\times I_N\times K_1\times\cdots\times K_N}$ *and* $\mathcal{B} \in \mathbb{R}^{K_1\times\cdots\times K_N\times J_1\times\cdots\times J_M}$. *Then*
(1) $(\mathcal{A} *_N \mathcal{B})^H = \mathcal{B}^H *_N \mathcal{A}^H$;
(2) $(\mathcal{I}_N *_N \mathcal{B}) = \mathcal{B}$ *and* $(\mathcal{B} *_M \mathcal{I}_M) = \mathcal{B}$, *where the unit tensors* $\mathcal{I}_N \in \mathbb{C}^{K_1\times\cdots\times K_N\times K_1\times\cdots\times K_N}$ *and* $\mathcal{I}_M \in \mathbb{C}^{J_1\times\cdots\times J_M\times J_1\times\cdots\times J_M}$.

Definition 1.2 [4] Let $\mathcal{A}, \mathcal{B} \in \mathbb{R}^{I_1\times\cdots\times I_N\times J_1\times\cdots\times J_M}$. The inner product of two tensors $\mathcal{A}$ and $\mathcal{B}$ is defined by:

$$\langle \mathcal{A}, \mathcal{B}\rangle = \operatorname{tr}(\mathcal{B}^T *_N \mathcal{A}). \tag{1.2}$$

For a tensor $\mathcal{A} \in \mathbb{R}^{I_1\times\cdots\times I_N\times I_1\times\cdots\times I_N}$, the tensor norm induced by this inner product is Frobenius norm $\|\mathcal{A}\| = (\sum_{i_1\cdots i_N j_1\cdots j_N} |a_{i_1\cdots i_N j_1\cdots j_N}|^2)^{\frac{1}{2}}$. Wang and Xu [34] derived the symmetric property of this inner product as follows.

Lemma 1.3 [34] *Let* $\mathcal{A}, \mathcal{B} \in \mathbb{R}^{I_1\times\cdots\times I_N\times J_1\times\cdots\times J_M}$. *Then*

$$\langle \mathcal{A}, \mathcal{B}\rangle = \operatorname{tr}(\mathcal{B}^T *_N \mathcal{A}) = \operatorname{tr}(\mathcal{A} *_M \mathcal{B}^T) = \operatorname{tr}(\mathcal{B} *_M \mathcal{A}^T) = \operatorname{tr}(\mathcal{A}^T *_N \mathcal{B}) = \langle \mathcal{B}, \mathcal{A}\rangle.$$

In addition, it follows from the definition of the inner product and the properties of the tensor trace that for any $\mathcal{A}, \mathcal{B} \in \mathbb{C}^{I_1\times\cdots\times I_N\times J_1\times\cdots\times J_M}$ and any scalar $\alpha \in \mathbb{C}$,

(1) linearity in the first argument:

$$\langle \alpha\mathcal{A}, \mathcal{B}\rangle = \alpha\langle \mathcal{A}, \mathcal{B}\rangle,\ \ \langle \mathcal{A}+\mathcal{C}, \mathcal{B}\rangle = \langle \mathcal{A}, \mathcal{B}\rangle + \langle \mathcal{C}, \mathcal{B}\rangle,$$

(2) positive definiteness: $\langle \mathcal{A}, \mathcal{A}\rangle > 0$ for all nonzero tensor $\mathcal{A}$.

And, for two tensors $\mathcal{X}, \mathcal{Y} \in \mathbb{C}^{I_1\times\cdots\times I_N\times J_1\times\cdots\times J_M}$, by simple calculations, we obtain

$$\|\mathcal{X} - \mathcal{Y}\|^2 = \|\mathcal{X}\|^2 - 2\langle \mathcal{X}, \mathcal{Y}\rangle + \|\mathcal{Y}\|^2.$$

Lemma 1.4 [12] *Let* $\mathcal{A} \in \mathbb{R}^{I_1\times\cdots\times I_N\times J_1\times\cdots\times J_N}$ *and* $\mathcal{B} \in \mathbb{R}^{K_1\times\cdots\times K_M\times L_1\times\cdots\times L_M}$. *Then for any* $\mathcal{X} \in \mathbb{R}^{J_1\times\cdots\times J_N\times K_1\times\cdots\times K_M}$ *and* $\mathcal{Y} \in \mathbb{R}^{I_1\times\cdots\times I_N\times L_1\times\cdots\times L_M}$, *we have*

$$\langle \mathcal{A} *_N \mathcal{X} *_M \mathcal{B}, \mathcal{Y}\rangle = \langle \mathcal{X}, \mathcal{A}^T *_N \mathcal{Y} *_M \mathcal{B}^T\rangle.$$

Lemma 1.5 *Let* $\mathcal{D} \in \mathbb{R}^{J_1\times\cdots\times J_N\times I_1\times\cdots\times I_N}$ *and* $\mathcal{C} \in \mathbb{R}^{L_1\times\cdots\times L_M\times K_1\times\cdots\times K_M}$. *Then for any* $\mathcal{X} \in \mathbb{R}^{J_1\times\cdots\times J_N\ \times K_1\times\cdots\times K_M}$ *and* $\mathcal{Y} \in \mathbb{R}^{L_1\times\cdots\times L_M\times I_1\times\cdots\times I_N}$, *we have*

$$\langle \mathcal{C} *_M \mathcal{X}^T *_N \mathcal{D}, \mathcal{Y}\rangle = \langle \mathcal{X}, \mathcal{D} *_N \mathcal{Y}^T *_M \mathcal{C}\rangle.$$

Proof The result can be established by using an argument similar to the one used in Lemma 1.4. □

For a tensor $\mathcal{A} \in \mathbb{C}^{I_1\times\cdots\times I_N\times I_1\times\cdots\times I_N}$, if there exists a tensor $\mathcal{B} \in \mathbb{C}^{I_1\times\cdots\times I_N\times I_1\times\cdots\times I_N}$ such that $\mathcal{A} *_N \mathcal{B} = \mathcal{B} *_N \mathcal{A} = \mathcal{I}$, then $\mathcal{A}$ is said to be invertible and $\mathcal{B}$ is called the inverse of $\mathcal{A}$ and denoted by $\mathcal{A}^{-1}$ [4]. For a general tensor $\mathcal{A}$ in $\mathbb{C}^{I_1\times\cdots\times I_N\times J_1\times\cdots\times J_M}$, its inverse may not exist. But it is shown in [25] that there exists a unique Moore-Penrose inverse $\mathcal{X}$ in $\mathbb{C}^{J_1\times\cdots\times J_M\times I_1\times\cdots\times I_N}$. The definition of the

Moore-Penrose inverse of tensors in $\mathbb{C}^{I_1\times\cdots\times I_N\times J_1\times\cdots\times J_M}$ via the Einstein product is as follows.

Definition 1.6 [25] Let $\mathcal{A} \in \mathbb{C}^{I_1\times\cdots\times I_N\times J_1\times\cdots\times J_M}$. The tensor $\mathcal{X}$ in $\mathbb{C}^{J_1\times\cdots\times J_M\times I_1\times\cdots\times I_N}$ satisfied

$$\begin{array}{ll}(1)\ \mathcal{A} *_M \mathcal{X} *_N \mathcal{A} = \mathcal{A}, & (2)\ \mathcal{X} *_N \mathcal{A} *_M \mathcal{X} = \mathcal{X}, \\ (3)\ (\mathcal{A} *_M \mathcal{X})^H = \mathcal{A} *_M \mathcal{X}, & (4)\ (\mathcal{X} *_N \mathcal{A})^H = \mathcal{X} *_N \mathcal{A}\end{array} \tag{1.3}$$

is called the Moore-Penrose inverse of $\mathcal{A}$, denoted by $\mathcal{A}^\dagger$.

Particularly, if a tensor $\mathcal{X}$ satisfies tensor equations (i), (j), $\cdots$, (k) in (1.3), it is called an $\{i, j, \cdots, k\}$-inverse of $\mathcal{A}$, denoted by $\mathcal{A}^{(i,j,\cdots,k)}$. The set of $\{i, j, \cdots, k\}$-inverses of a tensor $\mathcal{A}$ is denoted by $\mathcal{A}\{i, j, \cdots, k\}$.

Clearly, we have $\mathcal{A}^\dagger = \mathcal{A}^{-1} = \mathcal{A}^{(i,j,\cdots,k)}$ if $\mathcal{A}$ is invertible. Especially, if $M = N$, it reduces to the Moore-Penrose inverse defined in [32].

2 Reformulation as tensor equations with one variable

In [3], Behera and Mishra studied the properties of the generalized inverse of tensors and obtained the following results.

Lemma 2.1 [3, Theorem 2.32.] *Let* $\mathcal{A} \in \mathbb{R}^{I_1\times\cdots\times I_N\times J_1\times\cdots\times J_N}$. *Then*

$$\mathcal{A}^\dagger = \mathcal{A}^{(1,4)} *_N \mathcal{A} *_N \mathcal{A}^{(1,3)}.$$

Similar to Lemma 2.1, we are easy to see the following fact. For completeness, we give a detailed proof.

Proposition 2.2 *Let* $\mathcal{A} \in \mathbb{C}^{I_1\times\cdots\times I_N\times J_1\times\cdots\times J_M}$. *Then*

$$\mathcal{A}^\dagger = \mathcal{A}^{(1,4)} *_N \mathcal{A} *_M \mathcal{A}^{(1,3)}.$$

Proof Let $\mathcal{X} = \mathcal{A}^{(1,4)} *_N \mathcal{A} *_M \mathcal{A}^{(1,3)}$. Then

$$\begin{aligned}\mathcal{A} *_M \mathcal{X} *_N \mathcal{A} &= \mathcal{A} *_M (\mathcal{A}^{(1,4)} *_N \mathcal{A} *_M \mathcal{A}^{(1,3)}) *_N \mathcal{A} \\ &= (\mathcal{A} *_M \mathcal{A}^{(1,4)} *_N \mathcal{A}) *_M \mathcal{A}^{(1,3)} *_N \mathcal{A} \\ &= \mathcal{A} *_M \mathcal{A}^{(1,3)} *_N \mathcal{A} = \mathcal{A}\end{aligned}$$

and

$$\begin{aligned}\mathcal{X} *_N \mathcal{A} *_M \mathcal{X} &= (\mathcal{A}^{(1,4)} *_N \mathcal{A} *_M \mathcal{A}^{(1,3)}) *_N \mathcal{A} *_M (\mathcal{A}^{(1,4)} *_N \mathcal{A} *_M \mathcal{A}^{(1,3)}) \\ &= \mathcal{A}^{(1,4)} *_N (\mathcal{A} *_M \mathcal{A}^{(1,3)} *_N \mathcal{A}) *_M \mathcal{A}^{(1,4)} *_N \mathcal{A} *_M \mathcal{A}^{(1,3)} \\ &= A^{(1,4)} *_N \mathcal{A} *_M \mathcal{A}^{(1,4)} *_N \mathcal{A} *_M \mathcal{A}^{(1,3)} \\ &= A^{(1,4)} *_N (\mathcal{A} *_M \mathcal{A}^{(1,4)} *_N \mathcal{A}) *_M \mathcal{A}^{(1,3)} \\ &= \mathcal{A}^{(1,4)} *_N \mathcal{A} *_M \mathcal{A}^{(1,3)} = \mathcal{X}.\end{aligned}$$

Since

$$\mathcal{A}*_M\mathcal{X} = \mathcal{A}*_M(\mathcal{A}^{(1,4)}*_N\mathcal{A}*_M\mathcal{A}^{(1,3)}) = (\mathcal{A}*_M\mathcal{A}^{(1,4)}*_N\mathcal{A})*_M\mathcal{A}^{(1,3)} = \mathcal{A}*_M\mathcal{A}^{(1,3)},$$

it follows that

$$(\mathcal{A} *_M \mathcal{X})^H = (\mathcal{A} *_M \mathcal{A}^{(1,3)})^H = \mathcal{A} *_M \mathcal{A}^{(1,3)} = \mathcal{A} *_M \mathcal{X}.$$

By some calculations, we know

$$\mathcal{X}*_N\mathcal{A} = (\mathcal{A}^{(1,4)}*_N\mathcal{A}*_M\mathcal{A}^{(1,3)})*_N\mathcal{A} = \mathcal{A}^{(1,4)}*_N(\mathcal{A}*_M\mathcal{A}^{(1,3)}*_N\mathcal{A}) = \mathcal{A}^{(1,4)}*_N\mathcal{A}.$$

Therefore, we conclude that

$$(\mathcal{X} *_N \mathcal{A})^H = (\mathcal{A}^{(1,4)} *_N \mathcal{A})^H = \mathcal{A}^{(1,4)} *_N \mathcal{A} = \mathcal{X} *_N \mathcal{A}.$$

Hence, $\mathcal{X} = \mathcal{A}^{\dagger}$, which completes the proof. □

Let

$$\begin{array}{l} \mathcal{A}_1 = \mathcal{B}_1 = \mathcal{F}_1 = \mathcal{A} \in \mathbb{R}^{I_1\times\cdots\times I_N\times J_1\times\cdots\times J_M}, \\ \mathcal{C}_1 = \mathcal{O} \in \mathbb{R}^{I_1\times\cdots\times I_N\times I_1\times\cdots\times I_N}, \mathcal{D}_1 = \mathcal{O} \in \mathbb{R}^{J_1\times\cdots\times J_M\times J_1\times\cdots\times J_M}, \\ \mathcal{A}_2 = \mathcal{A} \in \mathbb{R}^{I_1\times\cdots\times I_N\times J_1\times\cdots\times J_M}, \mathcal{B}_2 = \mathcal{C}_2 = \mathcal{I} \in \mathbb{R}^{I_1\times\cdots\times I_N\times I_1\times\cdots\times I_N}, \\ \mathcal{D}_2 = -\mathcal{A}^T \in \mathbb{R}^{J_1\times\cdots\times J_M\times I_1\times\cdots\times I_N}, \mathcal{F}_2 = \mathcal{O} \in \mathbb{R}^{I_1\times\cdots\times I_N\times I_1\times\cdots\times I_N}. \end{array} \tag{2.1}$$

Then the computation of $\{1,3\}$-inverse in $\mathbb{R}^{I_1\times\cdots\times I_N\times J_1\times\cdots\times J_M}$ becomes the following tensor equations

$$\begin{cases} \mathcal{A}_1 *_M \mathcal{X} *_N \mathcal{B}_1 + \mathcal{C}_1 *_N \mathcal{X}^T *_M \mathcal{D}_1 = \mathcal{F}_1, \\ \mathcal{A}_2 *_M \mathcal{X} *_N \mathcal{B}_2 + \mathcal{C}_2 *_N \mathcal{X}^T *_M \mathcal{D}_2 = \mathcal{F}_2. \end{cases} \tag{2.2}$$

If

$$\begin{array}{l} \mathcal{A}_1 = \mathcal{B}_1 = \mathcal{F}_1 = \mathcal{A} \in \mathbb{R}^{I_1\times\cdots\times I_N\times J_1\times\cdots\times J_M}, \\ \mathcal{C}_1 = \mathcal{O} \in \mathbb{R}^{I_1\times\cdots\times I_N\times I_1\times\cdots\times I_N}, \mathcal{D}_1 = \mathcal{O} \in \mathbb{R}^{J_1\times\cdots\times J_M\times J_1\times\cdots\times J_M}, \\ \mathcal{A}_2 = \mathcal{D}_2 = \mathcal{I} \in \mathbb{R}^{J_1\times\cdots\times J_M\times J_1\times\cdots\times J_M}, \mathcal{B}_2 = \mathcal{A} \in \mathbb{R}^{I_1\times\cdots\times I_N\times J_1\times\cdots\times J_M}, \\ \mathcal{C}_2 = -\mathcal{A}^T \in \mathbb{R}^{J_1\times\cdots\times J_M\times I_1\times\cdots\times I_N}, \mathcal{F}_2 = \mathcal{O} \in \mathbb{R}^{J_1\times\cdots\times J_M\times J_1\times\cdots\times J_M}, \end{array} \tag{2.3}$$

then the computation of $\{1,4\}$-inverse in $\mathbb{R}^{I_1\times\cdots\times I_N\times J_1\times\cdots\times J_M}$ becomes the tensor equations (2.2). In other words, the computation of $\{1,3\}$-inverse of a tensor can be transformed into finding the solution of tensor equations (2.2) with the coefficient tensors given in (2.1), and the computation of $\{1,4\}$-inverse of a tensor can be transformed into finding the solution of tensor equations (2.2) with the coefficient tensors given in (2.3). Once the $\{1,4\}$-inverse and $\{1,3\}$-inverse of a tensor $\mathcal{A}$ are obtained, it follows from Proposition 2.2 that the unique Moore-Penrose inverse of $\mathcal{A}$ equals to $\mathcal{A}^{\dagger} = \mathcal{A}^{(1,4)} *_N \mathcal{A} *_M \mathcal{A}^{(1,3)}$.

For the tensor equations (2.2), we apply the tensor form of the conjugate gradient method to find the solution, which is described as in Algorithm 2.3.

Algorithm 2.3 (Conjugate gradient method with tensor form for solving (2.2))

Step 0 *Input appropriate dimensionality tensors* $\mathcal{A}_1, \mathcal{B}_1, \mathcal{F}_1 \in \mathbb{R}^{I_1\times\cdots\times I_N\times J_1\times\cdots\times J_M}$, $\mathcal{C}_1 \in \mathbb{R}^{I_1\times\cdots\times I_N\times I_1\times\cdots\times I_N}$, $\mathcal{D}_1 \in \mathbb{R}^{J_1\times\cdots\times J_M\times J_1\times\cdots\times J_M}$, $\mathcal{A}_2 \in \mathbb{R}^{I_1\times\cdots\times I_N\times J_1\times\cdots\times J_M}$, $\mathcal{B}_2, \mathcal{C}_2, \mathcal{F}_2 \in \mathbb{R}^{I_1\times\cdots\times I_N\times I_1\times\cdots\times I_N}$ *and* $\mathcal{D}_2$

$\in \mathbb{R}^{J_1\times\cdots\times J_M\times I_1\times\cdots\times I_N}$ *in* (2.2). *Choose the initial tensor* $\mathcal{X}_0 \in \mathbb{R}^{J_1\times\cdots\times J_M\times I_1\times\cdots\times I_N}$.

Step 1 *Compute*

$$\mathcal{R}_0^{(1)} = \mathcal{F}_1 - \mathcal{A}_1 *_M \mathcal{X}_0 *_N \mathcal{B}_1 - \mathcal{C}_1 *_N \mathcal{X}_0^T *_M \mathcal{D}_1,\ \mathcal{R}_0^{(2)} = \mathcal{F}_2 - \mathcal{A}_2 *_M \mathcal{X}_0 *_N \mathcal{B}_2 - \mathcal{C}_2 *_N \mathcal{X}_0^T *_M \mathcal{D}_2,$$

and

$$\widetilde{\mathcal{R}}_0 = \mathcal{A}_1^T *_N \mathcal{R}_0^{(1)} *_M \mathcal{B}_1^T + \mathcal{A}_2^T *_M \mathcal{R}_0^{(2)} *_N \mathcal{B}_2^T + \mathcal{D}_1 *_M (\mathcal{R}_0^{(1)})^T *_N \mathcal{C}_1 + \mathcal{D}_2 *_M (\mathcal{R}_0^{(2)})^T *_M \mathcal{C}_2.$$

Set $\mathcal{Q}_0 = \widetilde{\mathcal{R}}_0$ *and* $k = 0$.

Step 2 *If* $\|\mathcal{R}_k^{(1)}\|^2 + \|\mathcal{R}_k^{(2)}\|^2 = 0$, *stop. Otherwise, go to step 3.*

Step 3 *Update the sequence*

$$\mathcal{X}_{k+1} = \mathcal{X}_k + \alpha_k \mathcal{Q}_k. \tag{2.4}$$

where

$$\alpha_k = \frac{\|\mathcal{R}_k^{(1)}\|^2 + \|\mathcal{R}_k^{(2)}\|^2}{\|\mathcal{Q}_k\|^2}. \tag{2.5}$$

Step 4 *Compute*

$$\begin{aligned} \mathcal{R}_{k+1}^{(1)} &= \mathcal{F}_1 - \mathcal{A}_1 *_M \mathcal{X}_{k+1} *_N \mathcal{B}_1 - \mathcal{C}_1 *_N \mathcal{X}_{k+1}^T *_M \mathcal{D}_1, \\ \mathcal{R}_{k+1}^{(2)} &= \mathcal{F}_2 - \mathcal{A}_2 *_M \mathcal{X}_{k+1} *_N \mathcal{B}_2 - \mathcal{C}_2 *_N \mathcal{X}_{k+1}^T *_M \mathcal{D}_2 \end{aligned} \tag{2.6}$$

and

$$\widetilde{\mathcal{R}}_{k+1} = \mathcal{A}_1^T *_N \mathcal{R}_{k+1}^{(1)} *_M \mathcal{B}_1^T + \mathcal{A}_2^T *_M \mathcal{R}_{k+1}^{(2)} *_M \mathcal{B}_2^T + \mathcal{D}_1 *_M (\mathcal{R}_{k+1}^{(1)})^T *_N \mathcal{C}_1 + \mathcal{D}_2 *_M (\mathcal{R}_{k+1}^{(2)})^T *_M \mathcal{C}_2. \tag{2.7}$$

Step 5 *Update the sequences*

$$\mathcal{Q}_{k+1} = \widetilde{\mathcal{R}}_{k+1} + \beta_k \mathcal{Q}_k, \tag{2.8}$$

where:

$$\beta_k = \frac{\|\mathcal{R}_{k+1}^{(1)}\|^2 + \|\mathcal{R}_{k+1}^{(2)}\|^2}{\|\mathcal{R}_k^{(1)}\|^2 + \|\mathcal{R}_k^{(2)}\|^2}. \tag{2.9}$$

Step 6 *Set* $k := k + 1$, *return to step 2.*

In what follows, we give the convergence analysis of Algorithm 2.3.

Lemma 2.4 *Let* $\{\mathcal{R}_k^{(1)}\}$, $\{\mathcal{R}_k^{(2)}\}$, $\{\widetilde{\mathcal{R}}_k\}$ *and* $\{\mathcal{Q}_k\}$ *be generated by Algorithm 2.3. Then*

$$\langle \mathcal{R}_{k+1}^{(1)}, \mathcal{R}_j^{(1)}\rangle + \langle \mathcal{R}_{k+1}^{(2)}, \mathcal{R}_j^{(2)}\rangle = \langle \mathcal{R}_k^{(1)}, \mathcal{R}_j^{(1)}\rangle + \langle \mathcal{R}_k^{(2)}, \mathcal{R}_j^{(2)}\rangle - \alpha_k \langle \mathcal{Q}_k, \widetilde{\mathcal{R}}_j\rangle. \tag{2.10}$$

Proof By the relations (2.4) and (2.6), we immediately have

$$\begin{aligned}\mathcal{R}_{k+1}^{(1)} &= \mathcal{F}_1 - \mathcal{A}_1 *_M \mathcal{X}_{k+1} *_N \mathcal{B}_1 - \mathcal{C}_1 *_N \mathcal{X}_{k+1}^T *_M \mathcal{D}_1 \\ &= \mathcal{F}_1 - \mathcal{A}_1 *_M (\mathcal{X}_k + \alpha_k \mathcal{Q}_k) *_N \mathcal{B}_1 - \mathcal{C}_1 *_N (\mathcal{X}_k + \alpha_k \mathcal{Q}_k)^T *_M \mathcal{D}_1 \\ &= \mathcal{R}_k^{(1)} - \alpha_k (\mathcal{A}_1 *_M \mathcal{Q}_k *_N \mathcal{B}_1 + \mathcal{C}_1 *_N \mathcal{Q}_k^T *_M \mathcal{D}_1). \end{aligned} \tag{2.11}$$

Similarly, we obtain

$$\mathcal{R}_{k+1}^{(2)} = \mathcal{R}_k^{(2)} - \alpha_k (\mathcal{A}_2 *_M \mathcal{Q}_k *_N \mathcal{B}_2 + \mathcal{C}_2 *_N \mathcal{Q}_k^T *_M \mathcal{D}_2). \tag{2.12}$$

It then follows from the relations (2.11), (2.12) and Lemmas 1.4 and 1.5 that

$$\begin{aligned} &\langle \mathcal{R}_{k+1}^{(1)}, \mathcal{R}_j^{(1)} \rangle + \langle \mathcal{R}_{k+1}^{(2)}, \mathcal{R}_j^{(2)} \rangle \\ = &\langle \mathcal{R}_k^{(1)}, \mathcal{R}_j^{(1)} \rangle + \langle \mathcal{R}_k^{(2)}, \mathcal{R}_j^{(2)} \rangle - \alpha_k [\langle \mathcal{A}_1 *_M \mathcal{Q}_k *_N \mathcal{B}_1 + \mathcal{C}_1 *_N \mathcal{Q}_k^T *_M \mathcal{D}_1, \mathcal{R}_j^{(1)} \rangle \\ &+ \langle \mathcal{A}_2 *_M \mathcal{Q}_k *_N \mathcal{B}_2 + \mathcal{C}_2 *_N \mathcal{Q}_k^T *_M \mathcal{D}_2, \mathcal{R}_j^{(2)} \rangle] \\ = &\langle \mathcal{R}_k^{(1)}, \mathcal{R}_j^{(1)} \rangle + \langle \mathcal{R}_k^{(2)}, \mathcal{R}_j^{(2)} \rangle - \alpha_k [\langle \mathcal{Q}_k, \mathcal{A}_1^T *_N \mathcal{R}_j^{(1)} *_M \mathcal{B}_1^T + \mathcal{A}_2^T *_M \mathcal{R}_j^{(2)} *_M \mathcal{B}_2^T \\ &+ \mathcal{D}_1 *_M (\mathcal{R}_j^{(1)})^T *_N \mathcal{C}_1 + \mathcal{D}_2 *_M (\mathcal{R}_j^{(2)})^T *_M \mathcal{C}_2 \rangle] \\ = &\langle \mathcal{R}_k^{(1)}, \mathcal{R}_j^{(1)} \rangle + \langle \mathcal{R}_k^{(2)}, \mathcal{R}_j^{(2)} \rangle - \alpha_k \langle \mathcal{Q}_k, \widetilde{\mathcal{R}}_j \rangle, \end{aligned}$$

which completes the proof. □

Lemma 2.5 *Let* $\{\mathcal{R}_k^{(1)}\}$, $\{\mathcal{R}_k^{(2)}\}$ *and* $\{\mathcal{Q}_k\}$ *be generated by Algorithm 2.3. Then*

$$\langle \mathcal{R}_i^{(1)}, \mathcal{R}_j^{(1)} \rangle + \langle \mathcal{R}_i^{(2)}, \mathcal{R}_j^{(2)} \rangle = 0, \ \langle \mathcal{Q}_i, \mathcal{Q}_j \rangle = 0, \ i, j = 0, 1, \cdots, k, \ i \neq j. \tag{2.13}$$

Proof First, we prove that

$$\langle \mathcal{R}_i^{(1)}, \mathcal{R}_j^{(1)} \rangle + \langle \mathcal{R}_i^{(2)}, \mathcal{R}_j^{(2)} \rangle = 0, \ \langle \mathcal{Q}_i, \mathcal{Q}_j \rangle = 0, \ 0 \leq j < i \leq k. \tag{2.14}$$

This will be shown by induction on k. For $k = 1$, by Lemma 2.4, the definition of α_0 and the fact $\mathcal{Q}_0 = \widetilde{\mathcal{R}}_0$, we obtain

$$\begin{aligned} \langle \mathcal{R}_1^{(1)}, \mathcal{R}_0^{(1)} \rangle + \langle \mathcal{R}_1^{(2)}, \mathcal{R}_0^{(2)} \rangle &= \langle \mathcal{R}_0^{(1)}, \mathcal{R}_0^{(1)} \rangle + \langle \mathcal{R}_0^{(2)}, \mathcal{R}_0^{(2)} \rangle - \alpha_0 \langle \mathcal{Q}_0, \widetilde{\mathcal{R}}_0 \rangle \\ &= \|\mathcal{R}_0^{(1)}\|^2 + \|\mathcal{R}_0^{(2)}\|^2 - \frac{\|\mathcal{R}_0^{(1)}\|^2 + \|\mathcal{R}_0^{(2)}\|^2}{\|\mathcal{Q}_0\|^2} \|\mathcal{Q}_0\|^2 \\ &= 0. \end{aligned}$$

Meanwhile, by Lemmas 1.3 and 2.4, the relation (2.8) and the definitions of α_0 and β_0, we get

$$\begin{aligned}
\langle \mathcal{Q}_1, \mathcal{Q}_0\rangle &= \langle \widetilde{\mathcal{R}}_1 + \beta_0 \mathcal{Q}_0, \mathcal{Q}_0\rangle \\
&= \langle \widetilde{\mathcal{R}}_1, \mathcal{Q}_0\rangle + \beta_0\langle \mathcal{Q}_0, \mathcal{Q}_0\rangle \\
&= \langle \mathcal{Q}_0, \widetilde{\mathcal{R}}_1\rangle + \beta_0\langle \mathcal{Q}_0, \mathcal{Q}_0\rangle \\
&= \frac{[\langle \mathcal{R}_1^{(1)}, \mathcal{R}_0^{(1)}\rangle + \langle \mathcal{R}_1^{(2)}, \mathcal{R}_0^{(2)}\rangle] - [\langle \mathcal{R}_1^{(1)}, \mathcal{R}_1^{(1)}\rangle + \langle \mathcal{R}_1^{(2)}, \mathcal{R}_1^{(2)}\rangle]}{\alpha_0} + \beta_0\|\mathcal{Q}_0\|^2 \\
&= -\frac{\|\mathcal{R}_1^{(1)}\|^2 + \|\mathcal{R}_1^{(2)}\|^2}{\alpha_0} + \beta_0\|\mathcal{Q}_0\|^2 \\
&= -(\|\mathcal{R}_1^{(1)}\|^2 + \|\mathcal{R}_1^{(2)}\|^2)\frac{\|\mathcal{Q}_0\|^2}{\|\mathcal{R}_0^{(1)}\|^2 + \|\mathcal{R}_0^{(2)}\|^2} + \frac{\|\mathcal{R}_1^{(1)}\|^2 + \|\mathcal{R}_1^{(2)}\|^2}{\|\mathcal{R}_0^{(1)}\|^2 + \|\mathcal{R}_0^{(2)}\|^2}\|\mathcal{Q}_0\|^2 \\
&= 0.
\end{aligned}$$

Therefore, the relation (2.14) is true for $k = 1$. Assume now that the relation (2.14) is true for some k. Then for $k + 1$, it follows from Lemma 2.4, the relation (2.8), the definition of α_k and the induction principle that

$$\begin{aligned}
\langle \mathcal{R}_{k+1}^{(1)}, \mathcal{R}_k^{(1)}\rangle + \langle \mathcal{R}_{k+1}^{(2)}, \mathcal{R}_k^{(2)}\rangle &= \langle \mathcal{R}_k^{(1)}, \mathcal{R}_k^{(1)}\rangle + \langle \mathcal{R}_k^{(2)}, \mathcal{R}_k^{(2)}\rangle - \alpha_k\langle \mathcal{Q}_k, \widetilde{\mathcal{R}}_k\rangle \\
&= \|\mathcal{R}_k^{(1)}\|^2 + \|\mathcal{R}_k^{(2)}\|^2 - \alpha_k\langle \mathcal{Q}_k, \mathcal{Q}_k - \beta_{k-1}\mathcal{Q}_{k-1}\rangle \\
&= \|\mathcal{R}_k^{(1)}\|^2 + \|\mathcal{R}_k^{(2)}\|^2 - \frac{\|\mathcal{R}_k^{(1)}\|^2 + \|\mathcal{R}_k^{(2)}\|^2}{\|\mathcal{Q}_k\|^2}\|\mathcal{Q}_k\|^2 \\
&= 0.
\end{aligned}$$

Meanwhile, according to Lemmas 1.3 and 2.4, the relation (2.8) and the definitions of α_k and β_k, we get

$$\begin{aligned}
&\langle \mathcal{Q}_{k+1}, \mathcal{Q}_k\rangle = \langle \widetilde{\mathcal{R}}_{k+1} + \beta_k \mathcal{Q}_k, \mathcal{Q}_k\rangle \\
&= \langle \widetilde{\mathcal{R}}_{k+1}, \mathcal{Q}_k\rangle + \beta_k\langle \mathcal{Q}_k, \mathcal{Q}_k\rangle \\
&= \langle \mathcal{Q}_k, \widetilde{\mathcal{R}}_{k+1}\rangle + \beta_k\langle \mathcal{Q}_k, \mathcal{Q}_k\rangle \\
&= \frac{[\langle \mathcal{R}_{k+1}^{(1)}, \mathcal{R}_k^{(1)}\rangle + \langle \mathcal{R}_{k+1}^{(2)}, \mathcal{R}_k^{(2)}\rangle] - [\langle \mathcal{R}_{k+1}^{(1)}, \mathcal{R}_{k+1}^{(1)}\rangle + \langle \mathcal{R}_{k+1}^{(2)}, \mathcal{R}_{k+1}^{(2)}\rangle]}{\alpha_k} + \beta_k\langle \mathcal{Q}_k, \mathcal{Q}_k\rangle \\
&= -\frac{\|\mathcal{R}_{k+1}^{(1)}\|^2 + \|\mathcal{R}_{k+1}^{(2)}\|^2}{\alpha_k} + \beta_k\|\mathcal{Q}_k\|^2 \\
&= -(\|\mathcal{R}_{k+1}^{(1)}\|^2 + \|\mathcal{R}_{k+1}^{(2)}\|^2)\frac{\|\mathcal{Q}_k\|^2}{\|\mathcal{R}_k^{(1)}\|^2 + \|\mathcal{R}_k^{(2)}\|^2} + \frac{\|\mathcal{R}_{k+1}^{(1)}\|^2 + \|\mathcal{R}_{k+1}^{(2)}\|^2}{\|\mathcal{R}_k^{(1)}\|^2 + \|\mathcal{R}_k^{(2)}\|^2}\|\mathcal{Q}_k\|^2 \\
&= 0.
\end{aligned}$$

In addition, by Lemma 2.4, the relation (2.8), the definition of α_k, and the induction principle, for $j = 0, 1, 2, \cdots, k-1$, we obtain

$$\begin{aligned}\langle \mathcal{R}_{k+1}^{(1)}, \mathcal{R}_j^{(1)}\rangle + \langle \mathcal{R}_{k+1}^{(2)}, \mathcal{R}_j^{(2)}\rangle &= \langle \mathcal{R}_k^{(1)}, \mathcal{R}_j^{(1)}\rangle + \langle \mathcal{R}_k^{(2)}, \mathcal{R}_j^{(2)}\rangle - \alpha_k\langle \mathcal{Q}_k, \widetilde{\mathcal{R}}_j\rangle \\ &= -\alpha_k\langle \mathcal{Q}_k, \mathcal{Q}_j - \beta_{j-1}\mathcal{Q}_{j-1}\rangle \\ &= 0.\end{aligned}$$

Meanwhile, according to Lemmas 1.3 and 2.4, the relation (2.8), and the induction principle, for $j = 0, 1, 2, \cdots, k-1$, we get

$$\begin{aligned}&\langle \mathcal{Q}_{k+1}, \mathcal{Q}_j\rangle = \langle \widetilde{\mathcal{R}}_{k+1} + \beta_k\mathcal{Q}_k, \mathcal{Q}_j\rangle \\ &= \langle \widetilde{\mathcal{R}}_{k+1}, \mathcal{Q}_j\rangle + \beta_k\langle \mathcal{Q}_k, \mathcal{Q}_j\rangle \\ &= \langle \mathcal{Q}_j, \widetilde{\mathcal{R}}_{k+1}\rangle + \beta_k\langle \mathcal{Q}_k, \mathcal{Q}_j\rangle \\ &= \frac{[\langle \mathcal{R}_j^{(1)}, \mathcal{R}_{k+1}^{(1)}\rangle + \langle \mathcal{R}_j^{(2)}, \mathcal{R}_{k+1}^{(2)}\rangle] - [\langle \mathcal{R}_{j+1}^{(1)}, \mathcal{R}_{k+1}^{(1)}\rangle + \langle \mathcal{R}_{j+1}^{(2)}, \mathcal{R}_{k+1}^{(2)}\rangle]}{\alpha_j} + \beta_k\langle \mathcal{Q}_k, \mathcal{Q}_j\rangle \\ &= 0.\end{aligned}$$

So the relation (2.14) is true for $k+1$. By the induction principle, the relation (2.14) is true for all $0 \le j < i \le k$. For $j > i$, by Lemma 1.3, one has

$$\langle \mathcal{R}_i^{(1)}, \mathcal{R}_j^{(1)}\rangle + \langle \mathcal{R}_i^{(2)}, \mathcal{R}_j^{(2)}\rangle = \langle \mathcal{R}_j^{(1)}, \mathcal{R}_i^{(1)}\rangle + \langle \mathcal{R}_j^{(2)}, \mathcal{R}_i^{(2)}\rangle = 0,\ \langle \mathcal{Q}_i, \mathcal{Q}_j\rangle = \langle \mathcal{Q}_j, \mathcal{Q}_i\rangle = 0,$$

which completes the proof. □

Lemma 2.6 *Let* $\{\mathcal{R}_k^{(1)}\}$, $\{\mathcal{R}_k^{(2)}\}$ *and* $\{\mathcal{Q}_k\}$ *be generated by Algorithm 2.3. If* $\overline{\mathcal{X}}$ *is a solution of the tensor equations* (2.2)*, then*

$$\langle \overline{\mathcal{X}} - \mathcal{X}_k, \mathcal{Q}_k\rangle = \|\mathcal{R}_k^{(1)}\|^2 + \|\mathcal{R}_k^{(2)}\|^2. \tag{2.15}$$

Proof This will be shown by induction on k. For $k = 0$, by Algorithm 2.3 and Lemma 1.5, we have

$$\begin{aligned}&\langle \overline{\mathcal{X}} - \mathcal{X}_0, \mathcal{Q}_0\rangle = \langle \overline{\mathcal{X}} - \mathcal{X}_0, \widetilde{\mathcal{R}}_0\rangle \\ &= \langle \overline{\mathcal{X}} - \mathcal{X}_0, \mathcal{A}_1^T *_N \mathcal{R}_0^{(1)} *_M \mathcal{B}_1^T + \mathcal{A}_2^T *_M \mathcal{R}_0^{(2)} *_M \mathcal{B}_2^T + \mathcal{D}_1 *_M (\mathcal{R}_0^{(1)})^T *_N \mathcal{C}_1 + \mathcal{D}_2 *_M (\mathcal{R}_0^{(2)})^T *_M \mathcal{C}_2\rangle \\ &= \langle \mathcal{A}_1 *_M (\overline{\mathcal{X}} - \mathcal{X}_0) *_N \mathcal{B}_1 + \mathcal{C}_1 *_N (\overline{\mathcal{X}} - \mathcal{X}_0)^T *_M \mathcal{D}_1, \mathcal{R}_0^{(1)}\rangle \\ &\quad + \langle \mathcal{A}_2 *_M (\overline{\mathcal{X}} - \mathcal{X}_0) *_N \mathcal{B}_2 + \mathcal{C}_2 *_N (\overline{\mathcal{X}} - \mathcal{X}_0)^T *_M \mathcal{D}_2, \mathcal{R}_0^{(2)}\rangle \\ &= \langle \mathcal{F}_1 - \mathcal{A}_1 *_M \mathcal{X}_0 *_N \mathcal{B}_1 - \mathcal{C}_1 *_N \mathcal{X}_0^T *_M \mathcal{D}_1, \mathcal{R}_0^{(1)}\rangle \\ &\quad + \langle \mathcal{F}_2 - \mathcal{A}_2 *_M \mathcal{X}_0 *_N \mathcal{B}_2 - \mathcal{C}_2 *_N \mathcal{X}_0^T *_M \mathcal{D}_2, \mathcal{R}_0^{(2)}\rangle \\ &= \|\mathcal{R}_0^{(1)}\|^2 + \|\mathcal{R}_0^{(2)}\|^2.\end{aligned}$$

Therefore, the relation (2.15) is true for $k=0$. Assume now that the relation (2.15) is true for some k. Then for $k+1$, it follows from the relation (2.4), the definition of α_k, and the induction principle that

$$\begin{aligned}\langle \overline{\mathcal{X}}-\mathcal{X}_{k+1}, \mathcal{Q}_k\rangle &= \langle \overline{\mathcal{X}}-\mathcal{X}_k, \mathcal{Q}_k\rangle - \alpha_k\|\mathcal{Q}_k\|^2 \\ &= \|\mathcal{R}_k^{(1)}\|^2+\|\mathcal{R}_k^{(2)}\|^2 - \frac{\|\mathcal{R}_k^{(1)}\|^2+\|\mathcal{R}_k^{(2)}\|^2}{\|\mathcal{Q}_k\|^2}\|\mathcal{Q}_k\|^2=0.\end{aligned}$$

Hence

$$\begin{aligned}&\langle \overline{\mathcal{X}}-\mathcal{X}_{k+1}, \mathcal{Q}_{k+1}\rangle = \langle \overline{\mathcal{X}}-\mathcal{X}_{k+1}, \widetilde{\mathcal{R}}_{k+1}\rangle + \beta_k\langle \overline{\mathcal{X}}-\mathcal{X}_{k+1}, \mathcal{Q}_k\rangle \\ &= \langle \overline{\mathcal{X}}-\mathcal{X}_{k+1}, \widetilde{\mathcal{R}}_{k+1}\rangle \\ &= \langle \overline{\mathcal{X}}-\mathcal{X}_{k+1}, \mathcal{A}_1^T *_N \mathcal{R}_{k+1}^{(1)} *_M \mathcal{B}_1^T + \mathcal{A}_2^T *_M \mathcal{R}_{k+1}^{(2)} *_M \mathcal{B}_2^T \\ &\quad +\mathcal{D}_1 *_M (\mathcal{R}_{k+1}^{(1)})^T *_N \mathcal{C}_1 + \mathcal{D}_2 *_M (\mathcal{R}_{k+1}^{(2)})^T *_M \mathcal{C}_2\rangle \\ &= \langle \mathcal{A}_1 *_M (\overline{\mathcal{X}}-\mathcal{X}_{k+1}) *_N \mathcal{B}_1 + \mathcal{C}_1 *_N (\overline{\mathcal{X}}-\mathcal{X}_{k+1})^T *_M \mathcal{D}_1, \mathcal{R}_{k+1}^{(1)}\rangle \\ &\quad +\langle \mathcal{A}_2 *_M (\overline{\mathcal{X}}-\mathcal{X}_{k+1}) *_N \mathcal{B}_2 + \mathcal{C}_2 *_N (\overline{\mathcal{X}}-\mathcal{X}_{k+1})^T *_M \mathcal{D}_2, \mathcal{R}_{k+1}^{(2)}\rangle \\ &= \langle \mathcal{F}_1 - \mathcal{A}_1 *_M \mathcal{X}_{k+1} *_N \mathcal{B}_1 - \mathcal{C}_1 *_N \mathcal{X}_{k+1}^T *_M \mathcal{D}_1, \mathcal{R}_{k+1}^{(1)}\rangle \\ &\quad +\langle \mathcal{F}_2 - \mathcal{A}_2 *_M \mathcal{X}_{k+1} *_N \mathcal{B}_2 - \mathcal{C}_2 *_N \mathcal{X}_{k+1}^T *_M \mathcal{D}_2, \mathcal{R}_{k+1}^{(2)}\rangle \\ &= \|\mathcal{R}_{k+1}^{(1)}\|^2 + \|\mathcal{R}_{k+1}^{(2)}\|^2,\end{aligned}$$

which completes the proof. □

Theorem 2.7 *For any initial tensor $\mathcal{X}_0$, the solution of the tensor equations (2.2) can be derived in at most $IJ+I^2$ iterative steps by Algorithm 2.3.*

Proof If there exists some $i_0 (0 \le i_0 \le IJ+I^2-1)$ such that $\|\mathcal{R}_{i_0}^{(1)}\|^2+\|\mathcal{R}_{i_0}^{(2)}\|^2=0$, then $\mathcal{X}_{i_0}$ is the solution of the tensor equations (2.2). If $\|\mathcal{R}_k^{(1)}\|^2+\|\mathcal{R}_k^{(2)}\|^2 \neq 0 (k = 0, 1, \cdots, IJ+I^2-1)$, by Lemma 2.6, we know $\mathcal{Q}_k \neq \mathcal{O} (k=0,1,\cdots,IJ+I^2-1)$. Then $\mathcal{X}_{IJ+I^2}$ and $\mathcal{Q}_{IJ+I^2}$ can be generated by Algorithm 2.3. By Lemma 2.5, we know that $\langle \mathcal{R}_i^{(1)}, \mathcal{R}_j^{(1)}\rangle + \langle \mathcal{R}_i^{(2)}, \mathcal{R}_j^{(2)}\rangle = 0$ and $\langle \mathcal{Q}_i, \mathcal{Q}_j\rangle = 0$ for all $i, j = 0, 1, 2, \cdots, IJ+I^2-1$, and $i \neq j$. So the tensor sequence of $\left(\mathcal{R}_0^{(1)}, \mathcal{R}_0^{(2)}\right), \cdots, \left(\mathcal{R}_{IJ+I^2-1}^{(1)}, \mathcal{R}_{IJ+I^2-1}^{(2)}\right)$ is an orthogonal basis of the linear space $\mathbb{R}^{I_1\times\cdots\times I_N\times J_1\times\cdots\times J_M} \times\mathbb{R}^{I_1\times\cdots\times I_N\times I_1\times\cdots\times I_N}$. Since $\left(\mathcal{R}_{IJ+I^2}^{(1)}, \mathcal{R}_{IJ+I^2}^{(2)}\right) \in \mathbb{R}^{I_1\times\cdots\times I_N\times J_1\times\cdots\times J_M} \times \mathbb{R}^{I_1\times\cdots\times I_N\times I_1\times\cdots\times I_N}$ and $\langle \mathcal{R}_{IJ+I^2}^{(1)}, \mathcal{R}_k^{(1)}\rangle + \langle \mathcal{R}_{IJ+I^2}^{(2)}, \mathcal{R}_k^{(2)}\rangle = 0$ for $k = 0, 1, \cdots, IJ+I^2-1$, we have $\mathcal{R}_{IJ+I^2}^{(1)} = \mathcal{O}$ and $\mathcal{R}_{IJ+I^2}^{(2)} = \mathcal{O}$, which completes the proof. □

3 Reformulation as tensor equations with two variables

In order to obtain the main result of this section, we first introduce the following lemma.

Lemma 3.1 *Let $\mathcal{A} \in \mathbb{C}^{I_1\times\cdots\times I_N\times J_1\times\cdots\times J_M}$. If $\mathcal{A}^H *_N \mathcal{A} = \mathcal{O}$, then $\mathcal{A} = \mathcal{O}$.*

Proof Let $\mathcal{A} = (a_{i_1\cdots i_N j_1\cdots j_M}) \in \mathbb{C}^{I_1\times\cdots\times I_N\times J_1\times\cdots\times J_M}$ and $\mathcal{A}^H = (b_{j_1\cdots j_M i_1\cdots i_N}) \in \mathbb{C}^{J_1\times\cdots\times J_M\times I_1\times\cdots\times I_N}$, where $b_{j_1\cdots j_M i_1\cdots i_N} = \bar{a}_{i_1\cdots i_N j_1\cdots j_M}$. Then, for any given $j_1 \in J_1, \cdots, j_M \in J_M$, we have

$$\begin{aligned}(\mathcal{A}^H *_N \mathcal{A})_{j_1\cdots j_M j_1\cdots j_M} &= \sum_{i_1\cdots i_N} b_{j_1\cdots j_M i_1\cdots i_N} a_{i_1\cdots i_N j_1\cdots j_M}\\ &= \sum_{i_1\cdots i_N} \bar{a}_{i_1\cdots i_N j_1\cdots j_M} a_{i_1\cdots i_N j_1\cdots j_M} = \sum_{i_1\cdots i_N} |a_{i_1\cdots i_N j_1\cdots j_M}|^2.\end{aligned}$$

Since $\mathcal{A}^H *_N \mathcal{A} = \mathcal{O}$, it follows that $a_{i_1\cdots i_N j_1\cdots j_M} = 0$ for all $i_1 \in I_1, \cdots, i_N \in I_N$, $j_1 \in J_1, \cdots, j_M \in J_M$. Hence $\mathcal{A} = \mathcal{O}$. The proof is completed. □

Theorem 3.2 *Let $\mathcal{A}\in\mathbb{C}^{I_1\times\cdots\times I_N\times J_1\times\cdots\times J_M}$. If there exist $\mathcal{X}\in\mathbb{C}^{J_1\times\cdots\times J_M\times I_1\times\cdots\times I_N}$, $\mathcal{U} \in \mathbb{C}^{I_1\times\cdots\times I_N\times I_1\times\cdots\times I_N}$ and $\mathcal{V} \in \mathbb{C}^{J_1\times\cdots\times J_M\times J_1\times\cdots\times J_M}$ satisfying*

$$\mathcal{A} *_M \mathcal{X} *_N \mathcal{A} = \mathcal{A},\ \mathcal{X} = \mathcal{A}^H *_N \mathcal{U}\ \text{ and }\ \mathcal{X} = \mathcal{V} *_M \mathcal{A}^H, \tag{3.1}$$

then $\mathcal{X}$ is unique and $\mathcal{X} = \mathcal{A}^\dagger$.

Proof According to the Moore-Penrose equation (1.3), we have

$$\begin{aligned}&\mathcal{A} *_M \mathcal{A}^\dagger *_N \mathcal{A} = \mathcal{A},\\ &\mathcal{A}^\dagger = \mathcal{A}^\dagger *_N \mathcal{A} *_M \mathcal{A}^\dagger = (\mathcal{A}^\dagger *_N \mathcal{A})^H *_M \mathcal{A}^\dagger = \mathcal{A}^H *_N (\mathcal{A}^\dagger)^H *_M \mathcal{A}^\dagger,\\ &\mathcal{A}^\dagger = \mathcal{A}^\dagger *_N \mathcal{A} *_M \mathcal{A}^\dagger = \mathcal{A}^\dagger *_N (\mathcal{A} *_M \mathcal{A}^\dagger)^H = \mathcal{A}^\dagger *_N (\mathcal{A}^\dagger)^H *_M \mathcal{A}^H.\end{aligned} \tag{3.2}$$

Denote $\mathcal{U} = (\mathcal{A}^\dagger)^H *_M \mathcal{A}^\dagger \in \mathbb{C}^{I_1\times\cdots\times I_N\times I_1\times\cdots\times I_N}$ and $\mathcal{V} = \mathcal{A}^\dagger *_N (\mathcal{A}^\dagger)^H \in \mathbb{C}^{J_1\times\cdots\times J_M\times J_1\times\cdots\times J_M}$. By the relation (3.2), it is easy to see that $\mathcal{A}^\dagger$ satisfies (3.1).

Now, we show that there exists a unique $\mathcal{X}$ satisfying (3.1). Suppose not, there exists $\mathcal{X}_1 \in \mathbb{C}^{J_1\times\cdots\times J_M\times I_1\times\cdots\times I_N}$ such that $\mathcal{X}_1 \neq \mathcal{X}$ and

$$\mathcal{A} *_M \mathcal{X}_1 *_N \mathcal{A} = \mathcal{A},\ \mathcal{X}_1 = \mathcal{A}^H *_N \mathcal{U}_1,\ \mathcal{X}_1 = \mathcal{V}_1 *_M \mathcal{A}^H, \tag{3.3}$$

where $\mathcal{U}_1 \in \mathbb{C}^{I_1\times\cdots\times I_N\times I_1\times\cdots\times I_N}$ and $\mathcal{V}_1 \in \mathbb{C}^{J_1\times\cdots\times J_M\times J_1\times\cdots\times J_M}$. Let

$$\mathcal{X}_2 = \mathcal{X} - \mathcal{X}_1,\ \mathcal{U}_2 = \mathcal{U} - \mathcal{U}_1,\ \mathcal{V}_2 = \mathcal{V} - \mathcal{V}_1. \tag{3.4}$$

It then follows from (3.1), (3.3), and (3.4) that

$$\mathcal{A} *_M \mathcal{X}_2 *_N \mathcal{A} = \mathcal{O},\ \mathcal{X}_2 = \mathcal{A}^H *_N \mathcal{U}_2,\ \mathcal{X}_2 = \mathcal{V}_2 *_M \mathcal{A}^H.$$

By some calculations, we immediately have

$$\begin{aligned}(\mathcal{X}_2 *_N \mathcal{A})^H *_M (\mathcal{X}_2 *_N \mathcal{A}) &= \mathcal{A}^H *_N (\mathcal{X}_2)^H *_M \mathcal{X}_2 *_N \mathcal{A}\\ &= \mathcal{A}^H *_N (\mathcal{A}^H *_N \mathcal{U}_2)^H *_M \mathcal{X}_2 *_N \mathcal{A}\\ &= (\mathcal{A}^H *_N \mathcal{U}_2^H) *_N (\mathcal{A} *_M \mathcal{X}_2 *_N \mathcal{A}) = \mathcal{O},\end{aligned}$$

which, together with Lemma 3.1, yields that $\mathcal{X}_2 *_N \mathcal{A} = \mathcal{O}$. Meanwhile, we find

$$\mathcal{X}_2 *_N \mathcal{X}_2^H = \mathcal{X}_2 *_N (\mathcal{V}_2 *_M \mathcal{A}^H)^H = (\mathcal{X}_2 *_N \mathcal{A}) *_M \mathcal{V}_2^H = \mathcal{O},$$

which, together with Lemma 3.1, yields that $\mathcal{X}_2 = \mathcal{O}$. According to the relation (3.4), it follows that $\mathcal{X} = \mathcal{X}_1$. This obtains a contradiction. Hence, we conclude that there exists a unique $\mathcal{X}$ satisfying (3.1) and $\mathcal{X} = \mathcal{A}^\dagger$. The proof is completed. □

Corollary 3.3 *Let $\mathcal{A} \in \mathbb{C}^{I_1\times\cdots\times I_N\times J_1\times\cdots\times J_M}$. If there exist $\mathcal{X} \in \mathbb{C}^{J_1\times\cdots\times J_M\times I_1\times\cdots\times I_N}$ and $\mathcal{Y} \in \mathbb{C}^{I_1\times\cdots\times I_N\times J_1\times\cdots\times J_M}$ satisfying*

$$\mathcal{A} *_M \mathcal{X} *_N \mathcal{A} = \mathcal{A},\ \mathcal{X} = \mathcal{A}^H *_N \mathcal{Y} *_M \mathcal{A}^H, \tag{3.5}$$

then $\mathcal{X} = \mathcal{A}^\dagger$.

Proof Let $\mathcal{U} = \mathcal{Y} *_M \mathcal{A}^H$ and $\mathcal{V} = \mathcal{A}^H *_N \mathcal{Y}$. Then, by (3.5), we have

$$\mathcal{A} *_M \mathcal{X} *_N \mathcal{A} = \mathcal{A},\ \mathcal{X} = \mathcal{A}^H *_N \mathcal{U},\ \mathcal{X} = \mathcal{V} *_M \mathcal{A}^H.$$

It then follows from Theorem 3.2 that $\mathcal{X} = \mathcal{A}^\dagger$. The proof is completed. □

Let

$$\begin{aligned}
&\mathcal{A}_1 = \mathcal{B}_1 = \mathcal{F}_1 = \mathcal{A} \in \mathbb{R}^{I_1\times\cdots\times I_N\times J_1\times\cdots\times J_M},\\
&\mathcal{C}_1 = \mathcal{O} \in \mathbb{R}^{I_1\times\cdots\times I_N\times I_1\times\cdots\times I_N},\ \mathcal{D}_1 = \mathcal{O} \in \mathbb{R}^{J_1\times\cdots\times J_M\times J_1\times\cdots\times J_M},\\
&\mathcal{A}_2 = \mathcal{I} \in \mathbb{R}^{J_1\times\cdots\times J_M\times J_1\times\cdots\times J_M},\ \mathcal{B}_2 = \mathcal{I} \in \mathbb{R}^{I_1\times\cdots\times I_N\times I_1\times\cdots\times I_N},\\
&\mathcal{C}_2 = -\mathcal{D}_2 = \mathcal{A}^T \in \mathbb{R}^{J_1\times\cdots\times J_M\times I_1\times\cdots\times I_N},\ \mathcal{F}_2 = \mathcal{O} \in \mathbb{R}^{J_1\times\cdots\times J_M\times I_1\times\cdots\times I_N}.
\end{aligned} \tag{3.6}$$

By Corollary 3.3, the computation of Moore-Penrose inverse of tensors in $\mathbb{R}^{I_1\times\cdots\times I_N\times J_1\times\cdots\times J_M}$ becomes the following tensor equations with two variables

$$\begin{cases} \mathcal{A}_1 *_M \mathcal{X} *_N \mathcal{B}_1 + \mathcal{C}_1 *_N \mathcal{Y} *_M \mathcal{D}_1 = \mathcal{F}_1,\\ \mathcal{A}_2 *_M \mathcal{X} *_N \mathcal{B}_2 + \mathcal{C}_2 *_N \mathcal{Y} *_M \mathcal{D}_2 = \mathcal{F}_2. \end{cases} \tag{3.7}$$

For the tensor equations (3.7), we apply the tensor form of the conjugate gradient method to find the solution, which is described as in Algorithm 3.4.

Algorithm 3.4 (Conjugate gradient method with tensor form for solving (3.7))

Step 0 *Input appropriate dimensionality tensors $\mathcal{A}_1$, $\mathcal{B}_1$, $\mathcal{C}_1$, $\mathcal{D}_1$, $\mathcal{F}_1$ and $\mathcal{A}_2$, $\mathcal{B}_2$, $\mathcal{C}_2$, $\mathcal{D}_2$, $\mathcal{F}_2$ in (2.2). Choose the initial tensors $\mathcal{X}_0 \in \mathbb{R}^{J_1\times\cdots\times J_M\times I_1\times\cdots\times I_N}$ and $\mathcal{Y}_0 \in \mathbb{R}^{I_1\times\cdots\times I_N\times J_1\times\cdots\times J_M}$.*

Step 1 *Compute*

$$\mathcal{R}_0^{(1)} = \mathcal{F}_1 - \mathcal{A}_1 *_M \mathcal{X}_0 *_N \mathcal{B}_1 - \mathcal{C}_1 *_N \mathcal{Y}_0 *_M \mathcal{D}_1,$$
$$\mathcal{R}_0^{(2)} = \mathcal{F}_2 - \mathcal{A}_2 *_M \mathcal{X}_0 *_N \mathcal{B}_2 - \mathcal{C}_2 *_N \mathcal{Y}_0 *_M \mathcal{D}_2,$$

and

$$\widetilde{\mathcal{R}}_0^{(1)} = \mathcal{A}_1^T *_N \mathcal{R}_0^{(1)} *_M \mathcal{B}_1^T + \mathcal{A}_2^T *_M \mathcal{R}_0^{(2)} *_N \mathcal{B}_2^T,$$
$$\widetilde{\mathcal{R}}_0^{(2)} = \mathcal{C}_1^T *_N \mathcal{R}_0^{(1)} *_M \mathcal{D}_1^T + \mathcal{C}_2^T *_M \mathcal{R}_0^{(2)} *_N \mathcal{D}_2^T.$$

Set $\mathcal{P}_0 = \widetilde{\mathcal{R}}_0^{(1)}$, $\mathcal{Q}_0 = \widetilde{\mathcal{R}}_0^{(2)}$ and $k = 0$.

Step 2 *If $\|\mathcal{R}_k^{(1)}\|^2 + \|\mathcal{R}_k^{(2)}\|^2 = 0$, stop. Otherwise, go to step 3.*

Step 3 *Update the sequences*

$$\mathcal{X}_{k+1} = \mathcal{X}_k + \alpha_k \mathcal{P}_k,\ \mathcal{Y}_{k+1} = \mathcal{Y}_k + \alpha_k \mathcal{Q}_k, \tag{3.8}$$

where

$$\alpha_k = \frac{\|\mathcal{R}_k^{(1)}\|^2 + \|\mathcal{R}_k^{(2)}\|^2}{\|\mathcal{P}_k\|^2 + \|\mathcal{Q}_k\|^2}. \tag{3.9}$$

Step 4 *Compute*

$$\mathcal{R}_{k+1}^{(1)} = \mathcal{F}_1 - \mathcal{A}_1 *_M \mathcal{X}_{k+1} *_N \mathcal{B}_1 - \mathcal{C}_1 *_N \mathcal{Y}_{k+1} *_M \mathcal{D}_1,$$
$$\mathcal{R}_{k+1}^{(2)} = \mathcal{F}_2 - \mathcal{A}_2 *_M \mathcal{X}_{k+1} *_N \mathcal{B}_2 - \mathcal{C}_2 *_N \mathcal{Y}_{k+1} *_M \mathcal{D}_2,$$
$$\widetilde{\mathcal{R}}_{k+1}^{(1)} = \mathcal{A}_1^T *_N \mathcal{R}_{k+1}^{(1)} *_M \mathcal{B}_1^T + \mathcal{A}_2^T *_M \mathcal{R}_{k+1}^{(2)} *_N \mathcal{B}_2^T,$$
$$\widetilde{\mathcal{R}}_{k+1}^{(2)} = \mathcal{C}_1^T *_N \mathcal{R}_{k+1}^{(1)} *_M \mathcal{D}_1^T + \mathcal{C}_2^T *_M \mathcal{R}_{k+1}^{(2)} *_N \mathcal{D}_2^T.$$

Step 5 *Update the sequences*

$$\mathcal{P}_{k+1} = \widetilde{\mathcal{R}}_{k+1}^{(1)} + \beta_k \mathcal{P}_k, \ \mathcal{Q}_{k+1} = \widetilde{\mathcal{R}}_{k+1}^{(2)} + \beta_k \mathcal{Q}_k, \tag{3.10}$$

where

$$\beta_k = \frac{\|\mathcal{R}_{k+1}^{(1)}\|^2 + \|\mathcal{R}_{k+1}^{(2)}\|^2}{\|\mathcal{R}_k^{(1)}\|^2 + \|\mathcal{R}_k^{(2)}\|^2}. \tag{3.11}$$

Step 6 *Set $k := k + 1$, return to step 2.*

Lemma 3.5 *Let $\{\mathcal{R}_k^{(1)}\}$, $\{\mathcal{R}_k^{(2)}\}$, $\{\widetilde{\mathcal{R}}_k^{(1)}\}$, $\{\widetilde{\mathcal{R}}_k^{(2)}\}$, $\{\mathcal{P}_k\}$ and $\{\mathcal{Q}_k\}$ be generated by Algorithm 3.4. Then*

$$\langle \mathcal{R}_{k+1}^{(1)}, \mathcal{R}_j^{(1)} \rangle + \langle \mathcal{R}_{k+1}^{(2)}, \mathcal{R}_j^{(2)} \rangle = \langle \mathcal{R}_k^{(1)}, \mathcal{R}_j^{(1)} \rangle + \langle \mathcal{R}_k^{(2)}, \mathcal{R}_j^{(2)} \rangle - \alpha_k [\langle \mathcal{P}_k, \widetilde{\mathcal{R}}_j^{(1)} \rangle + \langle \mathcal{Q}_k, \widetilde{\mathcal{R}}_j^{(2)} \rangle].$$

Proof By step 4 of Algorithm 3.4 and the relation (3.8), we immediately have

$$\begin{aligned} \mathcal{R}_{k+1}^{(1)} &= \mathcal{F}_1 - \mathcal{A}_1 *_M \mathcal{X}_{k+1} *_N \mathcal{B}_1 - \mathcal{C}_1 *_N \mathcal{Y}_{k+1} *_M \mathcal{D}_1 \\ &= \mathcal{F}_1 - \mathcal{A}_1 *_M (\mathcal{X}_k + \alpha_k \mathcal{P}_k) *_N \mathcal{B}_1 - \mathcal{C}_1 *_N (\mathcal{Y}_k + \alpha_k \mathcal{Q}_k) *_M \mathcal{D}_1 \\ &= \mathcal{R}_k^{(1)} - \alpha_k (\mathcal{A}_1 *_M \mathcal{P}_k *_N \mathcal{B}_1 + \mathcal{C}_1 *_N \mathcal{Q}_k *_M \mathcal{D}_1). \end{aligned} \tag{3.12}$$

Similarly, we obtain

$$\mathcal{R}_{k+1}^{(2)} = \mathcal{R}_k^{(2)} - \alpha_k (\mathcal{A}_2 *_M \mathcal{P}_k *_N \mathcal{B}_2 + \mathcal{C}_2 *_N \mathcal{Q}_k *_M \mathcal{D}_2). \tag{3.13}$$

It then follows from the relations (3.12), (3.13) and Lemma 1.4 that

$$\begin{aligned} &\langle \mathcal{R}_{k+1}^{(1)}, \mathcal{R}_j^{(1)} \rangle + \langle \mathcal{R}_{k+1}^{(2)}, \mathcal{R}_j^{(2)} \rangle \\ &= \langle \mathcal{R}_k^{(1)}, \mathcal{R}_j^{(1)} \rangle + \langle \mathcal{R}_k^{(2)}, \mathcal{R}_j^{(2)} \rangle - \alpha_k [\langle \mathcal{A}_1 *_M \mathcal{P}_k *_N \mathcal{B}_1 + \mathcal{C}_1 *_N \mathcal{Q}_k *_M \mathcal{D}_1, \mathcal{R}_j^{(1)} \rangle \\ &\quad + \langle \mathcal{A}_2 *_M \mathcal{P}_k *_N \mathcal{B}_2 + \mathcal{C}_2 *_N \mathcal{Q}_k *_M \mathcal{D}_2, \mathcal{R}_j^{(2)} \rangle] \\ &= \langle \mathcal{R}_k^{(1)}, \mathcal{R}_j^{(1)} \rangle + \langle \mathcal{R}_k^{(2)}, \mathcal{R}_j^{(2)} \rangle - \alpha_k [\langle \mathcal{P}_k, \mathcal{A}_1^T *_N \mathcal{R}_j^{(1)} *_M \mathcal{B}_1^T + \mathcal{A}_2^T *_M \mathcal{R}_j^{(2)} *_N \mathcal{B}_2^T \rangle \\ &\quad + \langle \mathcal{Q}_k, \mathcal{C}_1^T *_N \mathcal{R}_j^{(1)} *_M \mathcal{D}_1^T + \mathcal{C}_2^T *_M \mathcal{R}_j^{(2)} *_N \mathcal{C}_2^T \rangle] \\ &= \langle \mathcal{R}_k^{(1)}, \mathcal{R}_j^{(1)} \rangle + \langle \mathcal{R}_k^{(2)}, \mathcal{R}_j^{(2)} \rangle - \alpha_k [\langle \mathcal{P}_k, \widetilde{\mathcal{R}}_j^{(1)} \rangle + \langle \mathcal{Q}_k, \widetilde{\mathcal{R}}_j^{(2)} \rangle], \end{aligned}$$

which completes the proof. □

Lemma 3.6 *Let* $\{\mathcal{R}_k^{(1)}\}$, $\{\mathcal{R}_k^{(2)}\}$, $\{\mathcal{P}_k\}$ *and* $\{\mathcal{Q}_k\}$ *be generated by Algorithm 3.4. Then* $\langle\mathcal{R}_i^{(1)},\mathcal{R}_j^{(1)}\rangle+\langle\mathcal{R}_i^{(2)},\mathcal{R}_j^{(2)}\rangle=0,\ \langle\mathcal{P}_i,\mathcal{P}_j\rangle+\langle\mathcal{Q}_i,\mathcal{Q}_j\rangle=0,\ i,j=0,1,\cdots,k,\ i\neq j.$

Proof First, we prove that

$$\langle\mathcal{R}_i^{(1)},\mathcal{R}_j^{(1)}\rangle+\langle\mathcal{R}_i^{(2)},\mathcal{R}_j^{(2)}\rangle=0,\ \langle\mathcal{P}_i,\mathcal{P}_j\rangle+\langle\mathcal{Q}_i,\mathcal{Q}_j\rangle=0,\ 0\leq j<i\leq k. \tag{3.14}$$

This will be shown by induction on k. For $k=1$, by Lemma 3.5, the fact $\mathcal{P}_0=\widetilde{\mathcal{R}}_0^{(1)}$ and $\mathcal{Q}_0=\widetilde{\mathcal{R}}_0^{(2)}$ and the definition of α_0, we obtain

$$\begin{aligned}
&\langle\mathcal{R}_1^{(1)},\mathcal{R}_0^{(1)}\rangle+\langle\mathcal{R}_1^{(2)},\mathcal{R}_0^{(2)}\rangle\\
&=\langle\mathcal{R}_0^{(1)},\mathcal{R}_0^{(1)}\rangle+\langle\mathcal{R}_0^{(2)},\mathcal{R}_0^{(2)}\rangle-\alpha_0[\langle\mathcal{P}_0,\widetilde{\mathcal{R}}_0^{(1)}\rangle+\langle\mathcal{Q}_0,\widetilde{\mathcal{R}}_0^{(2)}\rangle]\\
&=\|\mathcal{R}_0^{(1)}\|^2+\|\mathcal{R}_0^{(2)}\|^2-\alpha_0[\langle\mathcal{P}_0,\mathcal{P}_0\rangle+\langle\mathcal{Q}_0,\mathcal{Q}_0\rangle]\\
&=\|\mathcal{R}_0^{(1)}\|^2+\|\mathcal{R}_0^{(2)}\|^2-\frac{\|\mathcal{R}_0^{(1)}\|^2+\|\mathcal{R}_0^{(2)}\|^2}{\|\mathcal{P}_0\|^2+\|\mathcal{Q}_0\|^2}[\|\mathcal{P}_0\|^2+\|\mathcal{Q}_0\|^2]\\
&=0.
\end{aligned}$$

Meanwhile, by Lemmas 1.3 and 3.5 and the relations (3.9)–(3.11), we get

$$\begin{aligned}
&\langle\mathcal{P}_1,\mathcal{P}_0\rangle+\langle\mathcal{Q}_1,\mathcal{Q}_0\rangle\\
&=\langle\widetilde{\mathcal{R}}_1^{(1)}+\beta_0\mathcal{P}_0,\mathcal{P}_0\rangle+\langle\widetilde{\mathcal{R}}_1^{(2)}+\beta_0\mathcal{Q}_0,\mathcal{Q}_0\rangle\\
&=\langle\mathcal{P}_0,\widetilde{\mathcal{R}}_1^{(1)}\rangle+\langle\mathcal{Q}_0,\widetilde{\mathcal{R}}_1^{(2)}\rangle+\beta_0[\|\mathcal{P}_0\|^2+\|\mathcal{Q}_0\|^2]\\
&=\frac{[\langle\mathcal{R}_1^{(1)},\mathcal{R}_0^{(1)}\rangle+\langle\mathcal{R}_1^{(2)},\mathcal{R}_0^{(2)}\rangle]-[\langle\mathcal{R}_1^{(1)},\mathcal{R}_1^{(1)}\rangle+\langle\mathcal{R}_1^{(2)},\mathcal{R}_1^{(2)}\rangle]}{\alpha_0}+\beta_0[\|\mathcal{P}_0\|^2+\|\mathcal{Q}_0\|^2]\\
&=-\frac{\|\mathcal{R}_1^{(1)}\|^2+\|\mathcal{R}_1^{(2)}\|^2}{\alpha_0}+\beta_0[\|\mathcal{P}_0\|^2+\|\mathcal{Q}_0\|^2]\\
&=-[\|\mathcal{R}_1^{(1)}\|^2+\|\mathcal{R}_1^{(2)}\|^2]\frac{\|\mathcal{P}_0\|^2+\|\mathcal{Q}_0\|^2}{\|\mathcal{R}_0^{(1)}\|^2+\|\mathcal{R}_0^{(2)}\|^2}+\frac{\|\mathcal{R}_1^{(1)}\|^2+\|\mathcal{R}_1^{(2)}\|^2}{\|\mathcal{R}_0^{(1)}\|^2+\|\mathcal{R}_0^{(2)}\|^2}[\|\mathcal{P}_0\|^2+\|\mathcal{Q}_0\|^2]\\
&=0.
\end{aligned}$$

Therefore, the relation (3.14) is true for $k=1$. Assume now that the relation (3.14) is true for some k. Then for $k+1$, it follows from Lemma 3.5, the relations (3.9) and (3.10), and the induction principle that

$$\begin{aligned}
&\langle\mathcal{R}_{k+1}^{(1)},\mathcal{R}_k^{(1)}\rangle+\langle\mathcal{R}_{k+1}^{(2)},\mathcal{R}_k^{(2)}\rangle\\
&=\langle\mathcal{R}_k^{(1)},\mathcal{R}_k^{(1)}\rangle+\langle\mathcal{R}_k^{(2)},\mathcal{R}_k^{(2)}\rangle-\alpha_k[\langle\mathcal{P}_k,\widetilde{\mathcal{R}}_k^{(1)}\rangle+\langle\mathcal{Q}_k,\widetilde{\mathcal{R}}_k^{(2)}\rangle]\\
&=\|\mathcal{R}_k^{(1)}\|^2+\|\mathcal{R}_k^{(2)}\|^2-\alpha_k[\langle\mathcal{P}_k,\mathcal{P}_k-\beta_{k-1}\mathcal{P}_{k-1}\rangle+\langle\mathcal{Q}_k,\mathcal{Q}_k-\beta_{k-1}\mathcal{Q}_{k-1}\rangle]\\
&=\|\mathcal{R}_k^{(1)}\|^2+\|\mathcal{R}_k^{(2)}\|^2-\alpha_k[\|\mathcal{P}_k\|^2+\|\mathcal{Q}_k\|^2]+\alpha_k\beta_{k-1}[\langle\mathcal{P}_k,\mathcal{P}_{k-1}\rangle+\langle\mathcal{Q}_k,\mathcal{Q}_{k-1}\rangle]\\
&=\|\mathcal{R}_k^{(1)}\|^2+\|\mathcal{R}_k^{(2)}\|^2-\frac{\|\mathcal{R}_k^{(1)}\|^2+\|\mathcal{R}_k^{(2)}\|^2}{\|\mathcal{P}_k\|^2+\|\mathcal{Q}_k\|^2}[\|\mathcal{P}_k\|^2+\|\mathcal{Q}_k\|^2]\\
&=0.
\end{aligned}$$

Meanwhile, according to Lemmas 1.3 and 3.5, the relation (3.10), and the definitions of α_k and β_k, it follows that

$$\begin{aligned}
&\langle \mathcal{P}_{k+1}, \mathcal{P}_k\rangle + \langle \mathcal{Q}_{k+1}, \mathcal{Q}_k\rangle \\
&= \langle \widetilde{\mathcal{R}}^{(1)}_{k+1} + \beta_k \mathcal{P}_k, \mathcal{P}_k\rangle + \langle \widetilde{\mathcal{R}}^{(2)}_{k+1} + \beta_k \mathcal{Q}_k, \mathcal{Q}_k\rangle \\
&= \langle \widetilde{\mathcal{R}}^{(1)}_{k+1}, \mathcal{P}_k\rangle + \langle \widetilde{\mathcal{R}}^{(2)}_{k+1}, \mathcal{Q}_k\rangle + \beta_k[\|\mathcal{P}_k\|^2 + \|\mathcal{Q}_k\|^2] \\
&= \langle \mathcal{P}_k, \widetilde{\mathcal{R}}^{(1)}_{k+1}\rangle + \langle \mathcal{Q}_k, \widetilde{\mathcal{R}}^{(2)}_{k+1}\rangle + \beta_k[\|\mathcal{P}_k\|^2 + \|\mathcal{Q}_k\|^2] \\
&= \frac{[\langle \mathcal{R}^{(1)}_k, \mathcal{R}^{(1)}_{k+1}\rangle + \langle \mathcal{R}^{(2)}_k, \mathcal{R}^{(2)}_{k+1}\rangle] - [\langle \mathcal{R}^{(1)}_{k+1}, \mathcal{R}^{(1)}_{k+1}\rangle + \langle \mathcal{R}^{(2)}_{k+1}, \mathcal{R}^{(2)}_{k+1}\rangle]}{\alpha_k} \\
&\quad + \beta_k[\|\mathcal{P}_k\|^2 + \|\mathcal{Q}_k\|^2] \\
&= -\frac{\|\mathcal{R}^{(1)}_{k+1}\|^2 + \|\mathcal{R}^{(2)}_{k+1}\|^2}{\alpha_k} + \beta_k[\|\mathcal{P}_k\|^2 + \|\mathcal{Q}_k\|^2] \\
&= -[\|\mathcal{R}^{(1)}_{k+1}\|^2 + \|\mathcal{R}^{(2)}_{k+1}\|^2]\frac{\|\mathcal{P}_k\|^2 + \|\mathcal{Q}_k\|^2}{\|\mathcal{R}^{(1)}_k\|^2 + \|\mathcal{R}^{(2)}_k\|^2} + \frac{\|\mathcal{R}^{(1)}_{k+1}\|^2 + \|\mathcal{R}^{(2)}_{k+1}\|^2}{\|\mathcal{R}^{(1)}_k\|^2 + \|\mathcal{R}^{(2)}_k\|^2}[\|\mathcal{P}_k\|^2 + \|\mathcal{Q}_k\|^2] \\
&= 0.
\end{aligned}$$

In addition, by Lemma 3.5, the relation (3.10), the definition of α_k, and the induction principle, for $j = 0, 1, 2, \cdots, k-1$, we obtain

$$\begin{aligned}
&\langle \mathcal{R}^{(1)}_{k+1}, \mathcal{R}^{(1)}_j\rangle + \langle \mathcal{R}^{(2)}_{k+1}, \mathcal{R}^{(2)}_j\rangle \\
&= \langle \mathcal{R}^{(1)}_k, \mathcal{R}^{(1)}_j\rangle + \langle \mathcal{R}^{(2)}_k, \mathcal{R}^{(2)}_j\rangle - \alpha_k[\langle \mathcal{P}_k, \widetilde{\mathcal{R}}^{(1)}_j\rangle + \langle \mathcal{Q}_k, \widetilde{\mathcal{R}}^{(2)}_j\rangle] \\
&= -\alpha_k[\langle \mathcal{P}_k, \mathcal{P}_j - \beta_{j-1}\mathcal{P}_{j-1}\rangle + \langle \mathcal{Q}_k, \mathcal{Q}_j - \beta_{j-1}\mathcal{Q}_{j-1}\rangle] \\
&= -\alpha_k[\langle \mathcal{P}_k, \mathcal{P}_j\rangle + \langle \mathcal{Q}_k, \mathcal{Q}_j\rangle] + \alpha_k\beta_{j-1}[\langle \mathcal{P}_k, \mathcal{P}_{j-1}\rangle + \langle \mathcal{Q}_k, \mathcal{Q}_{j-1}\rangle] \\
&= 0.
\end{aligned}$$

Meanwhile, according to Lemma 3.5, the relation (3.10) and the induction principle, for $j = 0, 1, 2, \cdots, k-1$, we get

$$\begin{aligned}
&\langle \mathcal{P}_{k+1}, \mathcal{P}_j\rangle + \langle \mathcal{Q}_{k+1}, \mathcal{Q}_j\rangle = \langle \widetilde{\mathcal{R}}^{(1)}_{k+1} + \beta_k\mathcal{P}_k, \mathcal{P}_j\rangle + \langle \widetilde{\mathcal{R}}^{(2)}_{k+1} + \beta_k\mathcal{Q}_k, \mathcal{Q}_j\rangle \\
&= \langle \widetilde{\mathcal{R}}^{(1)}_{k+1}, \mathcal{P}_j\rangle + \langle \widetilde{\mathcal{R}}^{(2)}_{k+1}, \mathcal{Q}_j\rangle + \beta_k[\langle \mathcal{P}_k, \mathcal{P}_j\rangle + \langle \mathcal{Q}_k, \mathcal{Q}_j\rangle] \\
&= \langle \mathcal{P}_j, \widetilde{\mathcal{R}}^{(1)}_{k+1}\rangle + \langle \mathcal{Q}_j, \widetilde{\mathcal{R}}^{(2)}_{k+1}\rangle \\
&= \frac{[\langle \mathcal{R}^{(1)}_j, \mathcal{R}^{(1)}_{k+1}\rangle + \langle \mathcal{R}^{(2)}_j, \mathcal{R}^{(2)}_{k+1}\rangle] - [\langle \mathcal{R}^{(1)}_{j+1}, \mathcal{R}^{(1)}_{k+1}\rangle + \langle \mathcal{R}^{(2)}_{j+1}, \mathcal{R}^{(2)}_{k+1}\rangle]}{\alpha_j} \\
&= 0.
\end{aligned}$$

So the relation (3.14) is true for $k+1$. By the induction principle, the relation (3.14) is true for all $0 \le j < i \le k$. For $j > i$, by Lemma 1.3, it follows that

$$\langle \mathcal{R}^{(1)}_i, \mathcal{R}^{(1)}_j\rangle + \langle \mathcal{R}^{(2)}_i, \mathcal{R}^{(2)}_j\rangle = \langle \mathcal{R}^{(1)}_j, \mathcal{R}^{(1)}_i\rangle + \langle \mathcal{R}^{(2)}_j, \mathcal{R}^{(2)}_i\rangle = 0$$

and

$$\langle \mathcal{P}_i, \mathcal{P}_j\rangle + \langle \mathcal{Q}_i, \mathcal{Q}_j\rangle = \langle \mathcal{P}_j, \mathcal{P}_i\rangle + \langle \mathcal{Q}_j, \mathcal{Q}_i\rangle = 0,$$

which completes the proof. □

Lemma 3.7 *Let* $(\mathcal{X}, \mathcal{Y}) \in \mathbb{R}^{J_1\times\cdots\times J_M\times I_1\times\cdots\times I_N} \times \mathbb{R}^{I_1\times\cdots\times I_N\times J_1\times\cdots\times J_M}$ *be a solution of the tensor equations (3.7). Then for any initial tensors* $\mathcal{X}_0 \in \mathbb{R}^{J_1\times\cdots\times J_M\times I_1\times\cdots\times I_N}$ *and* $\mathcal{Y}_0 \in \mathbb{R}^{I_1\times\cdots\times I_N\times J_1\times\cdots\times J_M}$, *we have*

$$\langle \overline{\mathcal{X}} - \mathcal{X}_k, \mathcal{P}_k\rangle + \langle \overline{\mathcal{Y}} - \mathcal{Y}_k, \mathcal{Q}_k\rangle = \|\mathcal{R}_k^{(1)}\|^2 + \|\mathcal{R}_k^{(2)}\|^2, \tag{3.15}$$

where the sequences $\{\mathcal{X}_k\}$, $\{\mathcal{Y}_k\}$, $\{\mathcal{P}_k\}$, $\{\mathcal{Q}_k\}$, $\{\mathcal{R}_k^{(1)}\}$ *and* $\{\mathcal{R}_k^{(2)}\}$ *are generated by Algorithm 3.4.*

Proof This will be shown by induction on k. For $k = 0$, by Lemmas 1.4 and 1.5, it follows that

$$\begin{aligned}
&\langle \overline{\mathcal{X}} - \mathcal{X}_0, \mathcal{P}_0\rangle + \langle \overline{\mathcal{Y}} - \mathcal{Y}_0, \mathcal{Q}_0\rangle = \langle \overline{\mathcal{X}} - \mathcal{X}_0, \widetilde{\mathcal{R}}_0^{(1)}\rangle + \langle \overline{\mathcal{Y}} - \mathcal{Y}_0, \widetilde{\mathcal{R}}_0^{(2)}\rangle\\
&= \langle \overline{\mathcal{X}} - \mathcal{X}_0, \mathcal{A}_1^T *_N \mathcal{R}_0^{(1)} *_M \mathcal{B}_1^T + \mathcal{A}_2^T *_M \mathcal{R}_0^{(2)} *_N \mathcal{B}_2^T\rangle\\
&\quad + \langle \overline{\mathcal{Y}} - \mathcal{Y}_0, \mathcal{C}_1^T *_N \mathcal{R}_0^{(1)} *_M \mathcal{D}_1^T + \mathcal{C}_2^T *_M \mathcal{R}_0^{(2)} *_N \mathcal{D}_2^T\rangle\\
&= \langle \mathcal{A}_1 *_M (\overline{\mathcal{X}} - \mathcal{X}_0) *_N \mathcal{B}_1 + \mathcal{C}_1 *_N (\overline{\mathcal{Y}} - \mathcal{Y}_0) *_M \mathcal{D}_1, \mathcal{R}_0^{(1)}\rangle\\
&\quad + \langle \mathcal{A}_2 *_M (\overline{\mathcal{X}} - \mathcal{X}_0) *_N \mathcal{B}_2 + \mathcal{C}_2 *_N (\overline{\mathcal{Y}} - \mathcal{Y}_0) *_M \mathcal{D}_2, \mathcal{R}_0^{(2)}\rangle\\
&= \langle \mathcal{F}_1 - \mathcal{A}_1 *_M \mathcal{X}_0 *_N \mathcal{B}_1 - \mathcal{C}_1 *_N \mathcal{Y}_0 *_M \mathcal{D}_1, \mathcal{R}_0^{(1)}\rangle\\
&\quad + \langle \mathcal{F}_2 - \mathcal{A}_2 *_M \mathcal{X}_0 *_N \mathcal{B}_2 - \mathcal{C}_2 *_N \mathcal{Y}_0 *_M \mathcal{D}_2, \mathcal{R}_0^{(2)}\rangle\\
&= \|\mathcal{R}_0^{(1)}\|^2 + \|\mathcal{R}_0^{(2)}\|^2.
\end{aligned}$$

Therefore, the relation (3.15) is true for $k = 0$. Assume now that the relation (3.15) is true for some k. Then for $k + 1$, by the relation (3.8), the definition of α_k, and the induction principle, we have

$$\begin{aligned}
&\langle \overline{\mathcal{X}} - \mathcal{X}_{k+1}, \mathcal{P}_k\rangle + \langle \overline{\mathcal{Y}} - \mathcal{Y}_{k+1}, \mathcal{Q}_k\rangle\\
&= \langle \overline{\mathcal{X}} - \mathcal{X}_k, \mathcal{P}_k\rangle + \langle \overline{\mathcal{Y}} - \mathcal{Y}_k, \mathcal{Q}_k\rangle - \alpha_k[\|\mathcal{P}_k\|^2 + \|\mathcal{Q}_k\|^2]\\
&= \|\mathcal{R}_k^{(1)}\|^2 + \|\mathcal{R}_k^{(2)}\|^2 - \frac{\|\mathcal{R}_k^{(1)}\|^2 + \|\mathcal{R}_k^{(2)}\|^2}{\|\mathcal{P}_k\|^2 + \|\mathcal{Q}_k\|^2}[\|\mathcal{P}_k\|^2 + \|\mathcal{Q}_k\|^2]\\
&= 0.
\end{aligned}$$

Hence

$$\begin{aligned}
&\langle \overline{\mathcal{X}} - \mathcal{X}_{k+1}, \mathcal{P}_{k+1}\rangle + \langle \overline{\mathcal{Y}} - \mathcal{Y}_{k+1}, \mathcal{Q}_{k+1}\rangle\\
&= [\langle \overline{\mathcal{X}} - \mathcal{X}_{k+1}, \widetilde{\mathcal{R}}_{k+1}^{(1)}\rangle + \beta_k\langle \overline{\mathcal{X}} - \mathcal{X}_{k+1}, \mathcal{P}_k\rangle] + [\langle \overline{\mathcal{Y}} - \mathcal{Y}_{k+1}, \widetilde{\mathcal{R}}_{k+1}^{(2)}\rangle + \beta_k\langle \overline{\mathcal{Y}} - \mathcal{Y}_{k+1}, \mathcal{Q}_k\rangle]\\
&= \langle \overline{\mathcal{X}} - \mathcal{X}_{k+1}, \widetilde{\mathcal{R}}_{k+1}^{(1)}\rangle + \langle \overline{\mathcal{Y}} - \mathcal{Y}_{k+1}, \widetilde{\mathcal{R}}_{k+1}^{(2)}\rangle
\end{aligned}$$

$$
\begin{aligned}
&= \langle \overline{\mathcal{X}} - \mathcal{X}_{k+1}, \mathcal{A}_1^T *_N \mathcal{R}_{k+1}^{(1)} *_M \mathcal{B}_1^T + \mathcal{A}_2^T *_M \mathcal{R}_{k+1}^{(2)} *_N \mathcal{B}_2^T \rangle \\
&\quad + \langle \overline{\mathcal{Y}} - \mathcal{Y}_{k+1}, \mathcal{C}_1^T *_N \mathcal{R}_{k+1}^{(1)} *_M \mathcal{D}_1^T + \mathcal{C}_2^T *_M \mathcal{R}_{k+1}^{(2)} *_N \mathcal{D}_2^T \rangle \\
&= \langle \mathcal{A}_1 *_M (\overline{\mathcal{X}} - \mathcal{X}_{k+1}) *_N \mathcal{B}_1 + \mathcal{C}_1 *_N (\overline{\mathcal{Y}} - \mathcal{Y}_{k+1}) *_M \mathcal{D}_1, \mathcal{R}_{k+1}^{(1)} \rangle \\
&\quad + \langle \mathcal{A}_2 *_M (\overline{\mathcal{X}} - \mathcal{X}_{k+1}) *_N \mathcal{B}_2 + \mathcal{C}_2 *_N (\overline{\mathcal{Y}} - \mathcal{Y}_{k+1}) *_M \mathcal{D}_2, \mathcal{R}_{k+1}^{(2)} \rangle \\
&= \langle \mathcal{F}_1 - \mathcal{A}_1 *_M \mathcal{X}_{k+1} *_N \mathcal{B}_1 - \mathcal{C}_1 *_N \mathcal{X}_{k+1}^T *_M \mathcal{D}_1, \mathcal{R}_{k+1}^{(1)} \rangle \\
&\quad + \langle \mathcal{F}_2 - \mathcal{A}_2 *_M \mathcal{X}_{k+1} *_N \mathcal{B}_2 - \mathcal{C}_2 *_N \mathcal{X}_{k+1}^T *_M \mathcal{D}_2, \mathcal{R}_{k+1}^{(2)} \rangle \\
&= \|\mathcal{R}_{k+1}^{(1)}\|^2 + \|\mathcal{R}_{k+1}^{(2)}\|^2,
\end{aligned}
$$

which completes the proof. □

Theorem 3.8 *For any initial tensors $\mathcal{X}_0 \in \mathbb{R}^{J_1 \times \cdots J_M \times I_1 \times \cdots \times I_N}$ and $\mathcal{Y}_0 \in \mathbb{R}^{I_1 \times \cdots \times I_N \times J_1 \times \cdots \times J_M}$, the solution of the tensor equations (3.7) can be derived in at most $2IJ$ iterative steps by Algorithm 3.4.*

Proof If there exists some $i_0 (0 \leq i_0 \leq 2IJ - 1)$ such that $\|\mathcal{R}_{i_0}^{(1)}\|^2 + \|\mathcal{R}_{i_0}^{(2)}\|^2 = 0$, then $\left(\mathcal{X}_{i_0}, \mathcal{Y}_{i_0}\right)$ is a solution of the tensor equations (3.7). If $\|\mathcal{R}_k^{(1)}\|^2 + \|\mathcal{R}_k^{(2)}\|^2 \neq 0 (k = 0, 1, \cdots, 2IJ - 1)$, by Lemma 3.7, we know $(\mathcal{P}_k, \mathcal{Q}_k) \neq (\mathcal{O}, \mathcal{O}) \in \mathbb{R}^{J_1 \times \cdots \times J_M \times I_1 \quad \times \cdots \times I_N} \times \mathbb{R}^{I_1 \times \cdots \times I_N \times J_1 \times \cdots \times J_M} (k = 0, 1, \cdots, 2IJ - 1)$. Then $(\mathcal{X}_{2IJ}, \mathcal{Y}_{2IJ})$ and $(\mathcal{P}_{2IJ}, \mathcal{Q}_{2IJ})$ can be generated by Algorithm 3.4. By Lemma 3.6, we know that $\langle \mathcal{R}_i^{(1)}, \mathcal{R}_j^{(1)} \rangle + \langle \mathcal{R}_i^{(2)}, \mathcal{R}_j^{(2)} \rangle = 0$ and $\langle \mathcal{P}_i, \mathcal{P}_j \rangle + \langle \mathcal{Q}_i, \mathcal{Q}_j \rangle = 0$ for all $i, j = 0, 1, 2, \cdots, 2IJ - 1$, and $i \neq j$. So the tensor sequence of $\left(\mathcal{R}_0^{(1)}, \mathcal{R}_0^{(2)}\right), \cdots, \left(\mathcal{R}_{2IJ-1}^{(1)}, \mathcal{R}_{2IJ-1}^{(2)}\right)$ is an orthogonal basis of the linear space $\mathbb{R}^{I_1 \times \cdots \times I_N \times J_1 \times \cdots \times J_M} \times \mathbb{R}^{J_1 \times \cdots \times J_M \times I_1 \times \cdots \times I_N}$. Since $\left(\mathcal{R}_{2IJ}^{(1)}, \mathcal{R}_{2IJ}^{(2)}\right) \in \mathbb{R}^{I_1 \times \cdots \times I_N \times J_1 \times \cdots \times J_M} \times \mathbb{R}^{J_1 \times \cdots \times J_M \times I_1 \times \cdots \times I_N}$ and $\langle \mathcal{R}_{2IJ}^{(1)}, \mathcal{R}_k^{(1)} \rangle + \langle \mathcal{R}_{2IJ}^{(2)}, \mathcal{R}_k^{(2)} \rangle = 0$ for $k = 0, 1, \cdots, 2IJ - 1$, we have $\mathcal{R}_{2IJ}^{(1)} = \mathcal{O} \in \mathbb{R}^{I_1 \times \cdots \times I_N \times J_1 \times \cdots \times J_M}$ and $\mathcal{R}_{2IJ}^{(2)} = \mathcal{O} \in \mathbb{R}^{J_1 \times \cdots \times J_M \times I_1 \times \cdots \times I_N}$. The proof is completed. □

4 Numerical experiments

In this section, we give some numerical examples to show the performance of our proposed iterative algorithms. All of the tests are implemented in MATLAB R2015a with the machine precision 10^{-16} on a personal computer (Intel (R) Core (TM)i7-5500U), where the CPU is 2.40 GHz and the memory is 8.0 GB. All the implementations are based on the functions from the MATLAB Tensor Toolbox developed by Bader and Kolda [1]. For example, we illustrate how to use the Einstein product in Matlab tensor toolbox. For tensors $\mathcal{A} \in \mathbb{R}^{I_1 \times \cdots \times I_N \times K_1 \times \cdots \times K_N}$ and $\mathcal{B} \in \mathbb{R}^{K_1 \times \cdots \times K_N \times J_1 \times \cdots \times J_M}$, the Einstein product $\mathcal{A} *_N \mathcal{B}$ can be implemented by using the tensor command "ttt". More concretely

$$\begin{aligned} A &= \text{tenrand}([I_1\ I_2\ \cdots\ I_N\ K_1\ K_2\ \cdots\ K_N]); && //\text{Input tensor } \mathcal{A} \\ B &= \text{tenrand}([K_1\ K_2\ \cdots\ K_N\ J_1\ J_2\ \cdots\ J_M]); && //\text{Input tensor } \mathcal{B} \\ \mathcal{A} *_N \mathcal{B} &:= \text{ttt}(A, B, [N+1\ N+2\ \cdots\ 2N], [1\ 2\ \cdots\ N]). && //\text{Comput Einstein product} \end{aligned}$$

In this section, we first verify the validity of the proposed numerical algorithm. Second, we demonstrate the accuracy of the proposed numerical algorithm.

4.1 Validity of the proposed method

Example 1 In this example, we use Maltab function tenrand($[m\ n\ p\ q]$) to generate the tensor $\mathcal{A}$. Such as, we take $m = 4$, $n = 6$ and $p = q = 5$, and we let the tensor $\mathcal{A}$ as follows.

$$\mathcal{A}(:,:,1,1) = \begin{bmatrix} 0.4578 & 0.5270 & 0.2788 & 0.2399 & 0.9347 & 0.8997 \\ 0.7222 & 0.8942 & 0.3794 & 0.5977 & 0.8179 & 0.0652 \\ 0.3390 & 0.7784 & 0.8647 & 0.4794 & 0.7089 & 0.3359 \\ 0.4012 & 0.0694 & 0.4200 & 0.8985 & 0.7432 & 0.0043 \end{bmatrix},$$

$$\mathcal{A}(:,:,2,1) = \begin{bmatrix} 0.8281 & 0.5348 & 0.0683 & 0.8989 & 0.8750 & 0.0766 \\ 0.5074 & 0.2895 & 0.4098 & 0.3536 & 0.3486 & 0.7405 \\ 0.3662 & 0.0684 & 0.1234 & 0.1202 & 0.0419 & 0.4565 \\ 0.2266 & 0.0850 & 0.4430 & 0.5691 & 0.1423 & 0.6682 \end{bmatrix},$$

$$\mathcal{A}(:,:,3,1) = \begin{bmatrix} 0.6992 & 0.6624 & 0.0681 & 0.6524 & 0.4880 & 0.1171 \\ 0.5714 & 0.8754 & 0.7142 & 0.5310 & 0.4978 & 0.2404 \\ 0.6287 & 0.4675 & 0.3080 & 0.7151 & 0.9360 & 0.6849 \\ 0.8778 & 0.1413 & 0.6712 & 0.5048 & 0.3893 & 0.8393 \end{bmatrix},$$

$$\mathcal{A}(:,:,4,1) = \begin{bmatrix} 0.9701 & 0.4030 & 0.7437 & 0.3896 & 0.9052 & 0.0369 \\ 0.2152 & 0.5100 & 0.3020 & 0.7753 & 0.8764 & 0.5447 \\ 0.7603 & 0.4956 & 0.0896 & 0.1794 & 0.9998 & 0.9976 \\ 0.5841 & 0.6514 & 0.8260 & 0.1094 & 0.8643 & 0.5110 \end{bmatrix},$$

$$\mathcal{A}(:,:,5,1) = \begin{bmatrix} 0.8735 & 0.5643 & 0.0137 & 0.4739 & 0.9820 & 0.4260 \\ 0.0702 & 0.4315 & 0.3741 & 0.4965 & 0.5136 & 0.2132 \\ 0.9875 & 0.3378 & 0.9227 & 0.3090 & 0.9926 & 0.1932 \\ 0.9227 & 0.7207 & 0.5465 & 0.9508 & 0.4558 & 0.8328 \end{bmatrix},$$

$$\mathcal{A}(:,:,1,2) = \begin{bmatrix} 0.7266 & 0.5520 & 0.0522 & 0.4063 & 0.1692 & 0.1599 \\ 0.5297 & 0.2133 & 0.6833 & 0.6299 & 0.0010 & 0.6668 \\ 0.8291 & 0.5878 & 0.6086 & 0.5553 & 0.4182 & 0.0179 \\ 0.5119 & 0.1428 & 0.2197 & 0.1276 & 0.4885 & 0.1197 \end{bmatrix},$$

$$\mathcal{A}(:,:,2,2)=\begin{bmatrix}0.9521 & 0.8627 & 0.6290 & 0.6804 & 0.5714 & 0.6825\\ 0.9759 & 0.2429 & 0.0022 & 0.3788 & 0.9817 & 0.3581\\ 0.0309 & 0.8343 & 0.3797 & 0.6320 & 0.8497 & 0.9869\\ 0.4939 & 0.8136 & 0.9044 & 0.2433 & 0.2834 & 0.0840\end{bmatrix},$$

$$\mathcal{A}(:,:,3,2)=\begin{bmatrix}0.2503 & 0.8006 & 0.3504 & 0.9053 & 0.9316 & 0.2603\\ 0.8113 & 0.7388 & 0.4785 & 0.6402 & 0.7831 & 0.5693\\ 0.0844 & 0.1417 & 0.5874 & 0.1629 & 0.6857 & 0.2488\\ 0.5313 & 0.4379 & 0.1458 & 0.5659 & 0.4662 & 0.3193\end{bmatrix},$$

$$\mathcal{A}(:,:,4,2)=\begin{bmatrix}0.9108 & 0.1788 & 0.6640 & 0.6003 & 0.5270 & 0.2441\\ 0.8852 & 0.5175 & 0.3892 & 0.0850 & 0.1189 & 0.8844\\ 0.7946 & 0.6270 & 0.7400 & 0.9224 & 0.3801 & 0.7126\\ 0.9258 & 0.9132 & 0.8176 & 0.0536 & 0.8128 & 0.3781\end{bmatrix},$$

$$\mathcal{A}(:,:,5,2)=\begin{bmatrix}0.2489 & 0.6853 & 0.3839 & 0.4611 & 0.5317 & 0.5158\\ 0.2529 & 0.6203 & 0.2602 & 0.0910 & 0.3550 & 0.7906\\ 0.7672 & 0.7467 & 0.8775 & 0.5643 & 0.3148 & 0.2045\\ 0.0499 & 0.9773 & 0.8061 & 0.1874 & 0.7267 & 0.6781\end{bmatrix},$$

$$\mathcal{A}(:,:,1,3)=\begin{bmatrix}0.0525 & 0.0916 & 0.3161 & 0.3705 & 0.9905 & 0.0646\\ 0.8012 & 0.9084 & 0.0775 & 0.6224 & 0.2265 & 0.7477\\ 0.6786 & 0.5100 & 0.8506 & 0.9976 & 0.3980 & 0.4204\\ 0.9460 & 0.6149 & 0.1445 & 0.5173 & 0.6966 & 0.8113\end{bmatrix},$$

$$\mathcal{A}(:,:,2,3)=\begin{bmatrix}0.3796 & 0.4132 & 0.0738 & 0.0224 & 0.4658 & 0.8976\\ 0.3191 & 0.0986 & 0.1205 & 0.0538 & 0.5609 & 0.2886\\ 0.9861 & 0.7346 & 0.9816 & 0.1409 & 0.4945 & 0.2690\\ 0.7182 & 0.6373 & 0.4968 & 0.8935 & 0.0678 & 0.5942\end{bmatrix},$$

$$\mathcal{A}(:,:,3,3)=\begin{bmatrix}0.4759 & 0.6204 & 0.0273 & 0.5199 & 0.5480 & 0.0782\\ 0.3683 & 0.2828 & 0.8762 & 0.0538 & 0.5669 & 0.4564\\ 0.6556 & 0.2052 & 0.6101 & 0.8622 & 0.6804 & 0.0478\\ 0.9382 & 0.4391 & 0.2036 & 0.4429 & 0.3714 & 0.7383\end{bmatrix},$$

$$\mathcal{A}(:,:,4,3)=\begin{bmatrix}0.0380 & 0.5134 & 0.9933 & 0.5970 & 0.0948 & 0.4528\\ 0.9542 & 0.2409 & 0.3567 & 0.4306 & 0.4510 & 0.6522\\ 0.7424 & 0.2600 & 0.7529 & 0.7307 & 0.6401 & 0.8270\\ 0.9374 & 0.7590 & 0.1100 & 0.2612 & 0.1320 & 0.3081\end{bmatrix},$$

$$\mathcal{A}(:,:,5,3)=\begin{bmatrix}0.4024 & 0.7598 & 0.3747 & 0.2425 & 0.4794 & 0.9897\\ 0.8842 & 0.2909 & 0.4369 & 0.9367 & 0.5650 & 0.1837\\ 0.7006 & 0.2774 & 0.3043 & 0.8602 & 0.4896 & 0.8617\\ 0.2419 & 0.0061 & 0.2909 & 0.3972 & 0.2698 & 0.0326\end{bmatrix},$$

$$\mathcal{A}(:,:,1,4)=\begin{bmatrix}0.3320 & 0.9522 & 0.9155 & 0.3118 & 0.3559 & 0.8441\\ 0.7487 & 0.5433 & 0.8956 & 0.0553 & 0.3959 & 0.2881\\ 0.6444 & 0.2514 & 0.4825 & 0.7538 & 0.8855 & 0.2503\\ 0.1692 & 0.5786 & 0.4427 & 0.1319 & 0.0212 & 0.4884\end{bmatrix},$$

$$\mathcal{A}(:,:,2,4) = \begin{bmatrix} 0.7290 & 0.5932 & 0.1900 & 0.1183 & 0.5164 & 0.8735 \\ 0.2026 & 0.3044 & 0.0018 & 0.0390 & 0.0075 & 0.1133 \\ 0.2163 & 0.9677 & 0.7118 & 0.5982 & 0.6889 & 0.3546 \\ 0.9763 & 0.8960 & 0.8677 & 0.6043 & 0.9460 & 0.2419 \end{bmatrix},$$

$$\mathcal{A}(:,:,3,4) = \begin{bmatrix} 0.5603 & 0.7956 & 0.7087 & 0.9129 & 0.1868 & 0.7772 \\ 0.6127 & 0.7811 & 0.9929 & 0.4817 & 0.2472 & 0.5111 \\ 0.3008 & 0.3511 & 0.1625 & 0.8518 & 0.0542 & 0.0278 \\ 0.7981 & 0.0543 & 0.1136 & 0.8099 & 0.6090 & 0.9904 \end{bmatrix},$$

$$\mathcal{A}(:,:,4,4) = \begin{bmatrix} 0.5009 & 0.5751 & 0.1440 & 0.0060 & 0.8709 & 0.5505 \\ 0.3320 & 0.7510 & 0.8506 & 0.8019 & 0.7228 & 0.9599 \\ 0.1739 & 0.1535 & 0.3379 & 0.4974 & 0.6681 & 0.5960 \\ 0.6256 & 0.3568 & 0.2752 & 0.5378 & 0.1788 & 0.8086 \end{bmatrix},$$

$$\mathcal{A}(:,:,5,4) = \begin{bmatrix} 0.9845 & 0.4511 & 0.2490 & 0.8246 & 0.7408 & 0.7919 \\ 0.8859 & 0.0138 & 0.3864 & 0.4530 & 0.7376 & 0.4522 \\ 0.2138 & 0.4737 & 0.4314 & 0.3806 & 0.9469 & 0.8492 \\ 0.0346 & 0.9512 & 0.8309 & 0.9259 & 0.5101 & 0.3904 \end{bmatrix},$$

$$\mathcal{A}(:,:,1,5) = \begin{bmatrix} 0.7384 & 0.2072 & 0.3299 & 0.5211 & 0.3466 & 0.1986 \\ 0.9764 & 0.3234 & 0.3421 & 0.7743 & 0.3346 & 0.6725 \\ 0.5233 & 0.1109 & 0.8171 & 0.1203 & 0.5746 & 0.9018 \\ 0.4299 & 0.3752 & 0.5317 & 0.6255 & 0.8639 & 0.1992 \end{bmatrix},$$

$$\mathcal{A}(:,:,2,5) = \begin{bmatrix} 0.2983 & 0.2770 & 0.0148 & 0.0649 & 0.8138 & 0.7750 \\ 0.4965 & 0.5340 & 0.7028 & 0.3586 & 0.3934 & 0.1653 \\ 0.8899 & 0.5742 & 0.5067 & 0.2343 & 0.0536 & 0.9122 \\ 0.5014 & 0.4128 & 0.3813 & 0.2035 & 0.3751 & 0.3192 \end{bmatrix},$$

$$\mathcal{A}(:,:,3,5) = \begin{bmatrix} 0.3298 & 0.9500 & 0.1411 & 0.7321 & 0.5209 & 0.8162 \\ 0.2042 & 0.1582 & 0.5121 & 0.7498 & 0.2191 & 0.7939 \\ 0.7672 & 0.2864 & 0.7213 & 0.4073 & 0.8424 & 0.4691 \\ 0.0700 & 0.6871 & 0.9288 & 0.2395 & 0.6629 & 0.3095 \end{bmatrix},$$

$$\mathcal{A}(:,:,4,5) = \begin{bmatrix} 0.6876 & 0.7061 & 0.8651 & 0.5470 & 0.6137 & 0.5768 \\ 0.9869 & 0.5953 & 0.0680 & 0.4030 & 0.7830 & 0.9440 \\ 0.7699 & 0.7529 & 0.9685 & 0.1070 & 0.5666 & 0.8715 \\ 0.8296 & 0.4967 & 0.0988 & 0.7242 & 0.8113 & 0.5076 \end{bmatrix},$$

$$\mathcal{A}(:,:,5,5) = \begin{bmatrix} 0.7888 & 0.9761 & 0.8312 & 0.5383 & 0.0763 & 0.2681 \\ 0.4730 & 0.2782 & 0.9223 & 0.4633 & 0.7087 & 0.8325 \\ 0.8288 & 0.0728 & 0.3270 & 0.8208 & 0.2349 & 0.9954 \\ 0.3225 & 0.7512 & 0.8041 & 0.9519 & 0.3989 & 0.6498 \end{bmatrix}.$$

In this example, we test the numerical performance of Algorithm 2.3. By using Algorithm 2.3, after 367 and 374 iterative steps, we obtain $\mathcal{A}^{\{1,4\}}$ and $\mathcal{A}^{\{1,3\}}$, respectively.

According to Lemma 2.1, we know $\mathcal{A}^{\dagger} = \mathcal{A}^{(1,4)} *_N \mathcal{A} *_N \mathcal{A}^{(1,3)}$. Hence, we obtain the Moore-Penrose inverse $\mathcal{A}^{\dagger}$ as follows.

$$\mathcal{A}^{\dagger}(:,:,1,1) = \begin{bmatrix} -0.1650 & 0.2027 & -0.4917 & -0.0635 & 0.0108 \\ -0.1870 & 0.3486 & -0.6409 & -0.2833 & -0.0595 \\ -0.1216 & -0.3074 & -0.2688 & 0.0834 & -0.4525 \\ -0.3919 & 0.5205 & -0.3217 & 0.3211 & 0.3720 \\ 1.1242 & 0.2336 & 0.2863 & 0.0535 & -0.0713 \end{bmatrix},$$

$$\mathcal{A}^{\dagger}(:,:,2,1) = \begin{bmatrix} -0.9010 & 0.1303 & 0.3721 & 0.5459 & 0.6833 \\ -0.1782 & -0.1101 & 1.0716 & 0.1785 & -0.1557 \\ 0.0873 & 1.0675 & 0.0246 & -0.0287 & -0.5160 \\ 0.1697 & 0.0981 & -0.8803 & -0.3417 & -0.2832 \\ -0.9055 & -0.1478 & 0.5920 & 0.0556 & -0.1256 \end{bmatrix},$$

$$\mathcal{A}^{\dagger}(:,:,3,1) = \begin{bmatrix} -0.4435 & 0.0511 & -0.1973 & -0.1310 & -0.7637 \\ -0.4798 & -0.8853 & 0.6652 & -0.3036 & -0.2062 \\ 0.4506 & 0.8518 & -0.3432 & -0.2416 & -0.0075 \\ 0.1724 & 0.9684 & -0.5355 & -0.3699 & 0.3296 \\ -0.0021 & -0.1024 & 0.8937 & 0.2602 & 0.2134 \end{bmatrix},$$

$$\mathcal{A}^{\dagger}(:,:,4,1) = \begin{bmatrix} -0.2761 & -0.1614 & -0.1691 & -0.2627 & 0.0039 \\ -0.0822 & -0.0396 & 0.6648 & 0.3001 & -0.0674 \\ 0.0392 & 0.6461 & 0.0926 & 0.1535 & 0.0478 \\ 0.2022 & 0.4509 & 0.0256 & 0.0426 & -0.2562 \\ -0.2608 & -0.4896 & 0.1374 & -0.4639 & -0.0575 \end{bmatrix},$$

$$\mathcal{A}^{\dagger}(:,:,1,2) = \begin{bmatrix} -0.6090 & 0.1793 & 0.1039 & 0.3610 & 0.0791 \\ 0.2166 & 0.2035 & -0.1433 & 0.6577 & 0.0585 \\ 0.0324 & 0.5073 & 0.1181 & -0.3131 & -0.0864 \\ -0.1662 & -0.6636 & -0.4281 & -0.2144 & 0.3138 \\ -0.1127 & 0.0799 & 0.3400 & -0.5156 & 0.3424 \end{bmatrix},$$

$$\mathcal{A}^{\dagger}(:,:,2,2) = \begin{bmatrix} -0.1597 & -0.2053 & -0.5800 & -0.3341 & -0.9316 \\ -0.9755 & -0.6186 & -0.1843 & -0.4290 & -0.1818 \\ 1.0385 & 1.0987 & -1.1674 & -0.1795 & -0.2265 \\ -0.6114 & 1.5006 & -0.5162 & 0.3749 & 0.5336 \\ 0.6901 & 0.2967 & 0.5718 & 0.2058 & 0.2494 \end{bmatrix},$$

$$\mathcal{A}^{\dagger}(:,:,3,2) = \begin{bmatrix} 0.5921 & 0.9141 & 0.4628 & 0.0341 & -0.3236 \\ 0.4898 & 0.3495 & -0.1110 & 0.2251 & 0.4050 \\ 0.0435 & -0.8240 & -0.0758 & 0.0223 & -0.0460 \\ 0.2822 & -1.1232 & 0.6683 & -0.1715 & -0.0617 \\ -0.5628 & 0.0300 & -1.0425 & 0.2956 & -0.0428 \end{bmatrix},$$

$$\mathcal{A}^{\dagger}(:,:,4,2) = \begin{bmatrix} -0.9575 & 0.2672 & -0.0563 & -0.1871 & -0.5010 \\ -1.0241 & -0.2146 & 0.0047 & 0.0287 & 0.2527 \\ 0.0063 & 1.2335 & -0.6322 & -0.0840 & -0.4929 \\ -0.2027 & 0.5238 & -0.3670 & 0.0756 & -0.0260 \\ 0.5156 & 0.2705 & 0.3611 & 0.6077 & 0.3587, \end{bmatrix},$$

$$\mathcal{A}^{\dagger}(:,:,1,3)=\begin{bmatrix} 0.7520 & 0.0339 & 0.3298 & 0.6124 & 0.4138 \\ 0.7793 & 0.3601 & -0.2483 & 0.1399 & -0.0164 \\ -0.7446 & -0.9138 & -0.0016 & 0.2688 & 0.0221 \\ 0.5424 & -0.7570 & 0.7873 & -0.1951 & -0.3468 \\ -0.0221 & -0.2060 & -0.8240 & -0.4502 & 0.0267 \end{bmatrix},$$

$$\mathcal{A}^{\dagger}(:,:,2,3)=\begin{bmatrix} 0.5161 & 0.4243 & -0.0161 & 0.3095 & 0.3319 \\ 0.2934 & 0.0380 & -0.2584 & 0.1065 & 0.6544 \\ -0.3006 & -0.4222 & 0.4013 & 0.1149 & 0.0199 \\ 0.1800 & -0.6548 & 0.5170 & 0.0554 & -0.4816 \\ -0.3570 & -0.2000 & -0.9477 & 0.0862 & 0.0915, \end{bmatrix},$$

$$\mathcal{A}^{\dagger}(:,:,3,3)=\begin{bmatrix} 0.9592 & -0.0468 & 0.0727 & 0.1849 & 1.3013 \\ 0.6942 & 1.2206 & -1.0399 & -0.2262 & 0.4381 \\ -0.6946 & -1.7124 & 0.5753 & 0.0836 & 0.1253 \\ -0.5336 & -1.1621 & 0.9747 & 0.1427 & -0.2455 \\ 0.8918 & 0.5425 & -1.0873 & -0.7959 & -0.3085 \end{bmatrix},$$

$$\mathcal{A}^{\dagger}(:,:,4,3)=\begin{bmatrix} 0.2335 & -0.3094 & 0.1437 & 0.2384 & 0.7913 \\ 0.5035 & 0.2563 & 1.1888 & 0.1659 & -0.3190 \\ -0.0065 & 0.1974 & 0.0062 & 0.0926 & 0.2870 \\ 0.5023 & 0.0728 & -0.1827 & -0.2035 & -1.1228 \\ -0.8722 & -0.1802 & -0.2485 & -0.7127 & 0.0638, \end{bmatrix},$$

$$\mathcal{A}^{\dagger}(:,:,1,4)=\begin{bmatrix} 0.4082 & -0.1481 & -0.3557 & -0.5691 & -0.5925 \\ 0.1236 & 0.0147 & -0.5449 & -0.4761 & 0.1051 \\ 0.2996 & -0.1667 & -0.1174 & 0.2659 & 0.4607 \\ -0.2770 & 0.3382 & 0.6606 & -0.2697 & 0.1527 \\ 0.4836 & 0.0788 & -0.2746 & 0.3598 & -0.2040 \end{bmatrix},$$

$$\mathcal{A}^{\dagger}(:,:,2,4)=\begin{bmatrix} -0.1305 & 0.5603 & 0.4128 & -0.1145 & 0.6413 \\ -0.0567 & 0.6542 & 0.0252 & -0.1156 & 0.1523 \\ -0.4477 & -0.1209 & -0.4187 & 0.4166 & -0.1104 \\ 0.2846 & -0.8090 & 0.2765 & 0.0913 & -0.7551 \\ 0.3291 & 0.1399 & -0.1396 & -0.3001 & -0.0667 \end{bmatrix},$$

$$\mathcal{A}^{\dagger}(:,:,3,4)=\begin{bmatrix} 0.3950 & -0.1220 & 0.2081 & -0.2312 & -0.3872 \\ -0.0153 & 0.1496 & -0.5861 & -0.1463 & -0.2429 \\ -0.0572 & -0.6463 & 0.3612 & -0.0706 & 0.0700 \\ -0.2974 & 0.1963 & 0.1814 & 0.0920 & 0.0028 \\ 0.2920 & 0.1397 & 0.3145 & 0.1219 & 0.1218 \end{bmatrix},$$

$$\mathcal{A}^{\dagger}(:,:,4,4)=\begin{bmatrix} 0.3369 & 0.1163 & 0.1994 & -0.0846 & -0.5629 \\ 0.0924 & -0.7382 & 0.3167 & 0.3064 & -0.3058 \\ 0.2745 & 0.3343 & -0.3969 & -0.2541 & -0.0313 \\ -0.0815 & 0.0782 & 0.0218 & -0.1213 & 0.1183 \\ -0.2658 & -0.4654 & -0.0823 & 0.5073 & 0.6370 \end{bmatrix},$$

$$\mathcal{A}^{\dagger}(:,:,1,5) = \begin{bmatrix} 0.4397 & -0.0176 & 0.5296 & 0.4599 & -0.0324 \\ 1.0507 & 0.0393 & -0.2436 & 0.3919 & 0.2789 \\ -0.6587 & -0.3424 & 0.3403 & -0.2642 & 0.1042 \\ 0.3359 & -0.5341 & 0.2227 & -0.1540 & -0.2531 \\ -0.1688 & -0.3595 & -0.3299 & -0.1675 & -0.1181 \end{bmatrix},$$

$$\mathcal{A}^{\dagger}(:,:,2,5) = \begin{bmatrix} 0.0224 & -0.4009 & -0.4317 & -0.5839 & 0.0262 \\ -0.6712 & 0.3281 & 0.5403 & -0.7178 & -0.1902 \\ -0.1185 & 0.4254 & 0.5749 & 0.2216 & -0.1631 \\ 0.2311 & 0.4513 & -0.3790 & 0.0939 & -0.0618 \\ -0.0636 & 0.3826 & 0.4528 & -0.1004 & -0.0248 \end{bmatrix},$$

$$\mathcal{A}^{\dagger}(:,:,3,5) = \begin{bmatrix} 0.1197 & 0.0072 & -0.0751 & 0.2398 & -0.4373 \\ -0.2151 & -0.5813 & -0.0754 & 0.0536 & -0.4559 \\ 0.5955 & -0.0759 & -0.0188 & -0.3637 & 0.3061 \\ 0.0830 & 0.1614 & 0.1200 & 0.1321 & 0.4228 \\ -0.1449 & -0.3196 & -0.0258 & 0.5674 & -0.2185 \end{bmatrix},$$

$$\mathcal{A}^{\dagger}(:,:,4,5) = \begin{bmatrix} -0.3146 & -0.3220 & 0.0133 & -0.1091 & 0.4053 \\ -0.4362 & -0.3763 & -0.0259 & 0.3029 & -0.0290 \\ -0.1539 & 0.2850 & 0.6627 & 0.1984 & 0.0662 \\ 0.3572 & -0.0053 & -0.6424 & -0.3082 & 0.3088 \\ -0.4501 & 0.1404 & 0.4499 & 0.0825 & -0.0761 \end{bmatrix},$$

$$\mathcal{A}^{\dagger}(:,:,1,6) = \begin{bmatrix} -0.1956 & -0.5102 & -0.5748 & -0.2705 & -0.3718 \\ -0.4826 & -0.2995 & 0.2475 & -0.2447 & -0.1052 \\ 0.0334 & 0.2164 & -0.2347 & 0.3760 & 0.2937 \\ -0.2309 & 0.9275 & -0.2238 & 0.3363 & 0.2757 \\ 0.2339 & 0.0176 & 0.6876 & 0.2462 & -0.4007 \end{bmatrix},$$

$$\mathcal{A}^{\dagger}(:,:,2,6) = \begin{bmatrix} 0.0015 & 0.0470 & -0.2327 & -0.3243 & -0.7554 \\ -0.0483 & -0.6943 & 0.4605 & -0.1089 & -0.6002 \\ 0.2432 & 0.3996 & -0.1962 & -0.2933 & 0.4090 \\ -0.0680 & 0.9630 & -0.1710 & 0.5575 & 0.4801 \\ -0.5964 & -0.2240 & 0.2118 & 0.2861 & 0.0966 \end{bmatrix},$$

$$\mathcal{A}^{\dagger}(:,:,3,6) = \begin{bmatrix} 0.3232 & -0.2699 & -0.0351 & -0.1559 & -0.0509 \\ 0.3543 & 0.1471 & -1.0734 & 0.1403 & 0.5888 \\ 0.3096 & -1.0174 & 0.0977 & -0.3176 & 0.2315 \\ -0.3849 & -0.5096 & 0.6915 & 0.1364 & 0.5564 \\ 0.2651 & 0.0030 & -0.3870 & 0.1268 & 0.1110 \end{bmatrix},$$

$$\mathcal{A}^{\dagger}(:,:,4,6) = \begin{bmatrix} -0.4576 & -0.2473 & 0.5401 & 0.4492 & 1.1727 \\ 0.4009 & 0.7367 & -0.0064 & -0.0077 & 0.1946 \\ -0.2390 & -1.0392 & 0.5533 & 0.5573 & -0.1850 \\ 0.2200 & -1.2994 & -0.0130 & -0.0102 & -0.1988 \\ -0.0158 & 0.5075 & -0.3594 & -0.2779 & -0.5157 \end{bmatrix}.$$

At this case, the residuals of tensor equations (1), (2), (3), and (4) in (1.3) (see Definition 1.6) are $8.3508e-07$, $4.5770e-07$, $1.1107e-06$, and $1.0290e-06$, respectively. This confirms that $\mathcal{A}^{\dagger}$ is indeed the numerical result of the Moore-Penrose inverse of

$\mathcal{A}$. Hence, the Moore-Penrose inverse of a tensor can be obtained by Algorithm 2.3 effectively.

Example 2 In this example, we use Maltab function tenrand([m n p q]) to generate the tensor $\mathcal{A}$. Such as, we take $m = 4$, $n = 3$, $p = 5$ and $q = 4$, and we let the tensor $\mathcal{A}$ as follows.

$$\mathcal{A}(:,:,1,1) = \begin{bmatrix} 0.0855 & 0.9289 & 0.2373 \\ 0.2625 & 0.7303 & 0.4588 \\ 0.8010 & 0.4886 & 0.9631 \\ 0.0292 & 0.5785 & 0.5468 \end{bmatrix}, \mathcal{A}(:,:,2,1) = \begin{bmatrix} 0.5211 & 0.6791 & 0.0377 \\ 0.2316 & 0.3955 & 0.8852 \\ 0.4889 & 0.3674 & 0.9133 \\ 0.6241 & 0.9880 & 0.7962 \end{bmatrix},$$

$$\mathcal{A}(:,:,3,1) = \begin{bmatrix} 0.0987 & 0.1366 & 0.4942 \\ 0.2619 & 0.7212 & 0.7791 \\ 0.3354 & 0.1068 & 0.7150 \\ 0.6797 & 0.6538 & 0.9037 \end{bmatrix}, \mathcal{A}(:,:,4,1) = \begin{bmatrix} 0.8909 & 0.0305 & 0.9047 \\ 0.3342 & 0.7441 & 0.6099 \\ 0.6987 & 0.5000 & 0.6177 \\ 0.1978 & 0.4799 & 0.8594 \end{bmatrix},$$

$$\mathcal{A}(:,:,5,1) = \begin{bmatrix} 0.8055 & 0.8865 & 0.9787 \\ 0.5767 & 0.0287 & 0.7127 \\ 0.1829 & 0.4899 & 0.5005 \\ 0.2399 & 0.1679 & 0.4711 \end{bmatrix}, \mathcal{A}(:,:,1,2) = \begin{bmatrix} 0.0596 & 0.5216 & 0.7224 \\ 0.6820 & 0.0967 & 0.1499 \\ 0.0424 & 0.8181 & 0.6596 \\ 0.0714 & 0.8175 & 0.5186 \end{bmatrix},$$

$$\mathcal{A}(:,:,2,2) = \begin{bmatrix} 0.9730 & 0.4324 & 0.1734 \\ 0.6490 & 0.8253 & 0.3909 \\ 0.8003 & 0.0835 & 0.8314 \\ 0.4538 & 0.1332 & 0.8034 \end{bmatrix}, \mathcal{A}(:,:,3,2) = \begin{bmatrix} 0.0605 & 0.6569 & 0.0155 \\ 0.3993 & 0.6280 & 0.9841 \\ 0.5269 & 0.2920 & 0.1672 \\ 0.4168 & 0.4317 & 0.1062 \end{bmatrix},$$

$$\mathcal{A}(:,:,4,2) = \begin{bmatrix} 0.3724 & 0.9516 & 0.2691 \\ 0.1981 & 0.9203 & 0.4228 \\ 0.4897 & 0.0527 & 0.5479 \\ 0.3395 & 0.7379 & 0.9427 \end{bmatrix}, \mathcal{A}(:,:,5,2) = \begin{bmatrix} 0.4177 & 0.6663 & 0.1781 \\ 0.9831 & 0.5391 & 0.1280 \\ 0.3015 & 0.6981 & 0.9991 \\ 0.7011 & 0.6665 & 0.1711 \end{bmatrix},$$

$$\mathcal{A}(:,:,1,3) = \begin{bmatrix} 0.0326 & 0.1904 & 0.1564 \\ 0.5612 & 0.3689 & 0.8555 \\ 0.8819 & 0.4607 & 0.6448 \\ 0.6692 & 0.9816 & 0.3763 \end{bmatrix}, \mathcal{A}(:,:,2,3) = \begin{bmatrix} 0.1909 & 0.5895 & 0.2518 \\ 0.4283 & 0.2262 & 0.2904 \\ 0.4820 & 0.3846 & 0.6171 \\ 0.1206 & 0.5830 & 0.2653 \end{bmatrix},$$

$$\mathcal{A}(:,:,3,3) = \begin{bmatrix} 0.8244 & 0.5841 & 0.8178 \\ 0.9827 & 0.1078 & 0.2607 \\ 0.7302 & 0.9063 & 0.5944 \\ 0.3439 & 0.8797 & 0.0225 \end{bmatrix}, \mathcal{A}(:,:,4,3) = \begin{bmatrix} 0.4253 & 0.4229 & 0.6959 \\ 0.3127 & 0.0942 & 0.6999 \\ 0.1615 & 0.5985 & 0.6385 \\ 0.1788 & 0.4709 & 0.0336 \end{bmatrix},$$

$$\mathcal{A}(:,:,5,3) = \begin{bmatrix} 0.0688 & 0.4076 & 0.5313 \\ 0.3196 & 0.8200 & 0.3251 \\ 0.5309 & 0.7184 & 0.1056 \\ 0.6544 & 0.9686 & 0.6110 \end{bmatrix}, \mathcal{A}(:,:,1,4) = \begin{bmatrix} 0.7788 & 0.1537 & 0.4574 \\ 0.4235 & 0.2810 & 0.8754 \\ 0.0908 & 0.4401 & 0.5181 \\ 0.2665 & 0.5271 & 0.9436 \end{bmatrix},$$

$$\mathcal{A}(:,:,2,4) = \begin{bmatrix} 0.6377 & 0.2891 & 0.2548 \\ 0.9577 & 0.6718 & 0.2240 \\ 0.2407 & 0.6951 & 0.6678 \\ 0.6761 & 0.0680 & 0.8444 \end{bmatrix}, \mathcal{A}(:,:,3,4) = \begin{bmatrix} 0.3445 & 0.6022 & 0.4624 \\ 0.7805 & 0.3868 & 0.4243 \\ 0.6753 & 0.9160 & 0.4609 \\ 0.0067 & 0.0012 & 0.7702 \end{bmatrix},$$

$$\mathcal{A}(:,:,4,4) = \begin{bmatrix} 0.3225 & 0.1759 & 0.3411 \\ 0.7847 & 0.7218 & 0.6074 \\ 0.4714 & 0.4735 & 0.1917 \\ 0.0358 & 0.1527 & 0.7384 \end{bmatrix}, \mathcal{A}(:,:,5,4) = \begin{bmatrix} 0.2428 & 0.1887 & 0.6834 \\ 0.9174 & 0.2875 & 0.5466 \\ 0.2691 & 0.0911 & 0.4257 \\ 0.7655 & 0.5762 & 0.6444 \end{bmatrix},$$

In this example, we test the numerical performance of Algorithm 3.4. The stopping criterion is considered as the kth iteration satisfies:

$$\sqrt{\|\mathcal{R}_k^{(1)}\|^2 + \|\mathcal{R}_k^{(2)}\|^2} < \varepsilon$$

where $\varepsilon = 10^{-6}$. After 89 iterative steps, we obtain the Moore-Penrose inverse of $\mathcal{A}$ as follows.

$$\mathcal{A}^{\dagger}(:,:,1,1)=\begin{bmatrix} -0.3064 & -0.2184 & -0.2186 & 0.3635 \\ 0.2181 & 0.2738 & -0.0579 & 0.0292 \\ -0.3670 & 0.0032 & 0.3552 & -0.2814 \\ 0.2007 & 0.1369 & -0.0070 & 0.0103 \\ 0.0301 & 0.0664 & -0.0032 & -0.2719 \end{bmatrix},$$

$$\mathcal{A}^{\dagger}(:,:,2,1)=\begin{bmatrix} -0.1498 & 0.2778 & 0.0160 & 0.1147 \\ -0.2800 & 0.0907 & 0.1534 & -0.0895 \\ -0.1939 & 0.1749 & 0.1847 & -0.1213 \\ -0.2793 & 0.0642 & -0.2207 & 0.6269 \\ -0.1975 & 0.1952 & -0.4326 & 0.4692 \end{bmatrix},$$

$$\mathcal{A}^{\dagger}(:,:,3,1)=\begin{bmatrix} 0.1761 & -0.3443 & 0.5137 & -0.4129 \\ 0.0326 & 0.2707 & 0.1229 & -0.1705 \\ -0.0626 & -0.1871 & 0.2665 & 0.4749 \\ 0.1354 & -0.1454 & -0.2961 & -0.2537 \\ 0.0429 & -0.4201 & 0.0864 & 0.1058 \end{bmatrix},$$

$$\mathcal{A}^{\dagger}(:,:,4,1)=\begin{bmatrix} -0.2261 & -0.4486 & 0.1906 & -0.3614 \\ 0.2059 & 0.0159 & -0.2356 & 0.4632 \\ 0.2621 & -0.0573 & -0.1056 & 0.1862 \\ -0.1512 & -0.2039 & -0.0209 & -0.7339 \\ 0.3780 & 0.0509 & 0.5058 & 0.1820 \end{bmatrix},$$

$$\mathcal{A}^{\dagger}(:,:,1,2)=\begin{bmatrix} 0.0989 & -0.0327 & -0.2332 & -0.1730 \\ 0.1174 & -0.0202 & 0.0701 & -0.0327 \\ -0.1870 & 0.1805 & -0.0107 & 0.1573 \\ -0.3851 & 0.3879 & -0.1063 & -0.1993 \\ 0.4603 & -0.0293 & 0.0999 & 0.0275 \end{bmatrix},$$

$$\mathcal{A}^{\dagger}(:,:,2,2)=\begin{bmatrix} 0.1396 & -0.0102 & -0.4104 & -0.0563 \\ -0.4600 & 0.0458 & -0.1248 & 0.0507 \\ 0.1403 & 0.4038 & -0.1434 & -0.4833 \\ 0.3141 & 0.2069 & 0.2874 & 0.4553 \\ -0.3335 & 0.4097 & 0.1619 & -0.2495 \end{bmatrix},$$

$$\mathcal{A}^{\dagger}(:,:,3,2)=\begin{bmatrix} -0.0150 & 0.0037 & 0.0435 & 0.1022 \\ 0.1849 & -0.3217 & -0.1596 & 0.4179 \\ -0.1012 & -0.0018 & -0.0627 & 0.4236 \\ 0.0547 & -0.3849 & 0.1964 & -0.1647 \\ -0.0477 & 0.0721 & 0.4580 & -0.5863 \end{bmatrix},$$

$$\mathcal{A}^{\dagger}(:,:,4,2)=\begin{bmatrix}-0.1557 & 0.3371 & 0.0299 & 0.3451\\ 0.1637 & -0.1067 & 0.1550 & -0.3683\\ -0.2171 & -0.0319 & 0.3401 & -0.4753\\ 0.0068 & 0.3142 & -0.1516 & 0.2416\\ -0.3882 & 0.0489 & 0.0697 & 0.0208\end{bmatrix},$$

$$\mathcal{A}^{\dagger}(:,:,1,3)=\begin{bmatrix}0.1128 & 0.0567 & -0.1304 & -0.3620\\ -0.4507 & -0.0682 & -0.0477 & -0.1561\\ 0.3513 & -0.1480 & 0.0188 & -0.0340\\ 0.3230 & -0.0155 & 0.2537 & -0.1709\\ 0.3642 & -0.1473 & 0.1386 & 0.3186\end{bmatrix},$$

$$\mathcal{A}^{\dagger}(:,:,2,3)=\begin{bmatrix}-0.0317 & -0.1102 & 0.1362 & 0.2732\\ 0.0699 & -0.1464 & -0.0503 & -0.1254\\ 0.0886 & 0.4903 & -0.1916 & -0.0460\\ -0.0399 & -0.1880 & 0.3071 & 0.2789\\ 0.0875 & -0.0896 & -0.2571 & -0.0299\end{bmatrix},$$

$$\mathcal{A}^{\dagger}(:,:,3,3)=\begin{bmatrix}0.4173 & 0.1402 & 0.0638 & -0.1280\\ -0.0094 & 0.1225 & 0.0982 & 0.0742\\ 0.3757 & -0.2753 & -0.2514 & -0.0900\\ 0.1441 & -0.2595 & 0.3440 & -0.2092\\ -0.1228 & 0.3156 & -0.4954 & -0.0760\end{bmatrix},$$

$$\mathcal{A}^{\dagger}(:,:,4,3)=\begin{bmatrix}-0.0614 & 0.1726 & 0.0716 & 0.2152\\ 0.3118 & -0.0202 & 0.0290 & 0.1280\\ 0.0109 & -0.4160 & -0.2031 & 0.4374\\ -0.1458 & 0.1627 & -0.4594 & -0.0615\\ 0.0553 & -0.3960 & 0.0780 & 0.1473\end{bmatrix},$$

At this case, the residuals of tensor equations (1), (2), (3), and (4) in (1.3) (see Definition 1.6) are $3.9308e-07$, $1.1978e-07$, $2.5707e-07$, and $4.9847e-07$, respectively. This indicates that $\mathcal{A}^{\dagger}$ is indeed the Moore-Penrose inverse of $\mathcal{A}$. Hence, Algorithm 3.4 can calculate the Moore-Penrose inverse effectively.

4.2 Accuracy of numerical solutions

If a good approximation to the Moore-Penrose inverse $\mathcal{A}^{\dagger}$ of a tensor $\mathcal{A}$ is known in advance by using another method, then we can perform Algorithm 3.4 in order to obtain solutions of high accuracy.

Let ϕ be a map which satisfying

$$\begin{aligned}\phi:\ &\mathbb{R}^{I_1\times\cdots\times I_N\times J_1\times\cdots\times J_M}\rightarrow\mathbb{R}^{(I_1\cdots I_N)\times(J_1\cdots J_M)}\\ &\mathcal{A}=(\mathcal{A}_{i_1\cdots i_N j_1\cdots j_M})\rightarrow\mathbf{A}=(\mathbf{A}_{\text{ivec}(\mathbf{i},\mathbb{I}),\text{ivec}(\mathbf{j},\mathbb{J})}),\end{aligned}$$

where $\text{ivec}(\mathbf{i}, \mathbb{I}) = i_1 + \sum_{r=2}^{N} (i_r - 1) \prod_{u=1}^{r-1} I_u$ and $\text{ivec}(\mathbf{j}, \mathbb{J}) = j_1 + \sum_{s=2}^{M} (j_s - 1) \prod_{v=1}^{s-1} J_v$. Let $\mathcal{A} \in \mathbb{R}^{I_1 \times \cdots \times I_N \times J_1 \times \cdots \times J_M}$. Assume that the SVD of $\phi(\mathcal{A})$ satisfies $\phi(\mathcal{A}) = \mathbf{U}\mathbf{S}\mathbf{V}^H$. According to [4, 25], the SVD of $\mathcal{A}$ is

$$\mathcal{A} = \mathcal{U} *_M \mathcal{S} *_N \mathcal{V}^H,$$

where $\mathcal{U} = \phi^{-1}(\mathbf{U})$, $\mathcal{S} = \phi^{-1}(\mathbf{S})$ and $\mathcal{V} = \phi^{-1}(\mathbf{V})$. Then, by Theorem 3.4 of [25], the Moore-Penrose inverse of $\mathcal{A}$ exists uniquely and satisfies

$$\mathcal{A}^\dagger = \mathcal{V} *_N \mathcal{S}^\dagger *_M \mathcal{U}^H, \tag{4.1}$$

in which $\mathcal{S}^\dagger = (\mathcal{S}^\dagger_{j_1 \cdots j_M i_1 \cdots i_N}) \in \mathbb{R}^{J_1 \times \cdots \times J_M \times I_1 \times \cdots \times I_N}$ is defined by

$$\mathcal{S}^\dagger_{j_1 \cdots j_M i_1 \cdots i_N} = \begin{cases} \mathcal{S}^{-1}_{i_1 \cdots i_N j_1 \cdots j_M}, & \text{if } \mathcal{S}_{i_1 \cdots i_N j_1 \cdots j_M} \neq 0, \\ 0, & \text{otherwise.} \end{cases}$$

Hence, we only need the "reshape" command and the singular value decomposition of $\phi(\mathcal{A})$ to obtain the Moore Penrose inverse of a tensor $\mathcal{A}$.

By applying MATLAB's svd function to the matrix $\phi(\mathcal{A})$ and using the "reshape" command, we obtain the Moore-Penrose inverse of $\mathcal{A}$, denoted by $\mathcal{A}^\dagger_{\text{svd}}$. We can perform Algorithm 3.4 with the initial tensor $\mathcal{X}_0 = \mathcal{A}^\dagger_{\text{svd}}$ and $\mathcal{Y}_0$ being arbitrary in order to obtain a more accurate Moore-Penrose inverse $\mathcal{A}^\dagger_{\text{new}}$. To see the degree of accuracy, we compare the mean of the Fronenius norms

$$\begin{aligned} \text{RES}_{\text{svd}} = \frac{1}{4} \Big(& \|\mathcal{A} *_M \mathcal{A}^\dagger_{\text{svd}} *_N \mathcal{A} - \mathcal{A}\| + \|\mathcal{A}^\dagger_{\text{svd}} *_N \mathcal{A} *_M \mathcal{A}^\dagger_{\text{svd}} - \mathcal{A}^\dagger_{\text{svd}}\| \\ & + \|(\mathcal{A} *_M \mathcal{A}^\dagger_{\text{svd}})^T - \mathcal{A} *_M \mathcal{A}^\dagger_{\text{svd}}\| + \|(\mathcal{A}^\dagger_{\text{svd}} *_N \mathcal{A})^T - \mathcal{A}^\dagger_{\text{svd}} *_N \mathcal{A}\| \Big) \end{aligned}$$

and

$$\begin{aligned} \text{RES}_{\text{new}} = \frac{1}{4} \Big(& \|\mathcal{A} *_M \mathcal{A}^\dagger_{\text{new}} *_N \mathcal{A} - \mathcal{A}\| + \|\mathcal{A}^\dagger_{\text{new}} *_N \mathcal{A} *_M \mathcal{A}^\dagger_{\text{new}} - \mathcal{A}^\dagger_{\text{new}}\| \\ & + \|(\mathcal{A} *_M \mathcal{A}^\dagger_{\text{new}})^T - \mathcal{A} *_M \mathcal{A}^\dagger_{\text{new}}\| + \|(\mathcal{A}^\dagger_{\text{new}} *_N \mathcal{A})^T - \mathcal{A}^\dagger_{\text{new}} *_N \mathcal{A}\| \Big) \end{aligned}$$

If Algorithm 3.4 gives a more accurate Moore-Penrose inverse of tensor $\mathcal{A}$, the following inequality is expected to hold

$$\text{RES}_{\text{new}} \leq \text{RES}_{\text{svd}}. \tag{4.2}$$

Our as many as 200 experiments with randomly chosen tensor $\mathcal{A} = 100 * \text{tenrand}([4\ 3\ 5\ 3]) - 50$ show that (4.2) holds for all 200 tensors of $\mathcal{A}$. This means that Algorithm 3.4 enhances the accuracy of the Moore-Penrose inverse obtained by MATLAB's svd function. In other words, Algorithm 3.4 always generates the Moore-Penrose inverse of high accuracy in the end.

5 Concluding remarks

The aim of this paper is to give the numerical study on the Moore-Penrose inverse of tensors via the Einstein product. For this purpose, we first transform the calculation of

$\{1,4\}$-inverse of a tensor $\mathcal{A}$ ($\mathcal{A}^{\{1,4\}}$) and the $\{1,3\}$-inverse of a tensor $\mathcal{A}$ ($\mathcal{A}^{\{1,3\}}$) into finding the solution of a class of tensor equations via the Einstein product with one variable. Then, by the known property $\mathcal{A}^{\dagger} = \mathcal{A}^{(1,4)} *_N \mathcal{A} *_N \mathcal{A}^{(1,3)}$ (see [3, Theorem 2.32]), we obtain the Moore-Penrose inverse $\mathcal{A}^{\dagger}$. Unlike the above method, we derive the equivalent characterization of the Moore-Penrose inverse of tensors. Then, we solve the tensor equations via the Einstein product with two variables by means of the conjugate gradient method. In this way, we implement the numerical calculation of the Moore-Penrose inverse of tensors via the Einstein product.

Funding This research is supported by China Postdoctoral Science Foundation (Grant No. 2019M660203) and National Natural Science Foundation of China (Grant Nos. 12001211, 12071159, 61976053)

References

1. Bader, B.W., Kolda, T.G., et al.: MATLAB Tensor Toolbox Version 2.5, 2012. http://www.sandia.gov/tgkolda/TensorToolbox/
2. Behera, R., Maji, S., Mohapatra, R.N.: Weighted Moore-Penrose inverses of arbitrary order tensors. Comput. Appl. Math. **39**, 284 (2020)
3. Behera, R., Mishra, D.: Further results on generalized inverses of tensors via the Einstein product. Linear Multilinear Algebra **65**, 1662–1682 (2017)
4. Brazell, M., Li, N., Navasca, C., Tamon, C.: Solving multilinear systems via tensor inversion. SIAM J. Matrix Anal. Appl. **34**, 542–570 (2013)
5. Bu, C., Zhang, X., Zhou, J., Wang, W., Wei, Y.: The inverse, rank and product of tensors. Linear Algebra Appl. **446**, 269–280 (2014)
6. Bu, C., Zhou, J., Wei, Y.: E-cospectral hypergraphs and some hypergraphs determined by their spectra. Linear Algebra Appl. **459**, 397–403 (2014)
7. Burdick, D.S., Tu, X.M., McGown, L.B., Millican, D.W.: Resolution of multicomponent fluorescent mixtures by analysis of the excitation-emission frequency array. J. Chemom. **4**, 15–28 (1990)
8. Cooper, J., Dutle, A.: Spectra of uniform hypergraphs. Linear Algebra Appl. **436**, 3268–3292 (2012)
9. Ding, W., Wei, Y.: Fast Hankel tensor-vector product and its application to exponential data fitting. Numer. Linear Algebra Appl. **22**, 814–832 (2015)
10. Eldén, L.: Matrix Methods in Data Mining and Pattern Recognition. SIAM, Philadelphia (2007)
11. Hu, S., Qi, L.: Algebraic connectivity of an even uniform hypergraph. J. Comb. Optim. **24**, 564–579 (2012)
12. Huang, B.H., Ma, C.F.: An iterative algorithm to solve the generalized Sylvester tensor equations. Linear Multilinear Algebra **68**, 1175–1200 (2020)
13. Huang, B.H., Xie, Y.J., Ma, C.F.: Krylov subspace methods to solve a class of tensor equations via the Einstein product. Numer. Linear Algebra Appl. **26**, e2254 (2019)
14. Huang, B.H., Ma, C.F.: Global least squares methods based on tensor form to solve a class of generalized Sylvester tensor equations. Appl. Math. Comput. **369**, 124892 (2020)
15. Huang, B.H., Li, W.: Numerical subspace algorithms for solving the tensor equations involving Einstein product. Numer. Linear Algebra Appl (2020). https://doi.org/10.1002/nla.2351
16. Kolda, T., Bader, B.: Tensor decompositions and applications. SIAM Rev. **51**, 455–500 (2009)
17. Lai, W.M., Rubin, D., Krempl, E.: Introduction to Continuum Mechanics. Butterworth Heinemann, Oxford (2009)
18. Li, B.W., Tian, S., Sun, Y.S., Hu, Z.M.: Schur-decomposition for 3D matrix equations and its applications in solving radiative discrete ordinates equations discretized by Chebyshev collocation spectral method. J. Comput. Phys. **229**, 1198–1212 (2010)
19. Li, B.W., Sun, Y.S., Zhang, D.W.: Chebyshev collocation spectral methods for coupled radiation and conduction in a concentric spherical participating medium. ASME J. Heat Transfer. **131**, 062701–062709 (2009)

20. Li, W., Ng, M.: On the limiting probability distribution of a transition probability tensor. Linear Multilinear Algebra **62**, 362–385 (2014)
21. Li, Z., Ling, C., Wang, Y., Yang, Q.: Some advances in tensor analysis and polynomial optimization. Oper. Res. Trans. **18**, 134–148 (2014)
22. Ji, J., Wei, Y.: Weighted Moore-Penrose inverses and fundamental theorem of even-order tensors with Einstein product. Front. Math. China **12**, 1319–1337 (2017)
23. Ji, J., Wei, Y.: The Drazin inverse of an even-order tensor and its application to singular tensor equations. Comput. Math. Appl. **75**, 3402–3413 (2018)
24. Jin, H., Bai, M., Bentez, J., Liu, X.: The generalized inverses of tensors and an application to linear models. Comput. Math. Appl. **74**, 385–397 (2017)
25. Liang, M., Zheng, B.: Further results on Moore-Penrose inverses of tensors with application to tensor nearness problems. Comput. Math. Appl. **77**, 1282–1293 (2019)
26. Ma, H.F., Li, N., Stanimirović, P.S., Katsikis, V.N.: Perturbation theory for Moore-Penrose inverse of tensor via Einstein product. Comput. Appl. Math. **38**, 111 (2019)
27. Martin, C.D., Shafer, R., LaRue, B.: An order-p tensor factorization with applications in imaging. SIAM J. Sci. Comput. **35**, 474–490 (2013)
28. Panigrahy, K., Behera, R., Mishra, D.: Reverse-order law for the Moore-Penrose inverses of tensors. Linear Multilinear Algebra **68**, 246–264 (2020)
29. Panigrahy, K., Mishra, D.: On reverse-order law of tensors and its application to additive results on Moore–Penrose inverse. RACSAM **184**, 114 (2020)
30. Panigrahy, K., Mishra, D.: Extension of Moore-Penrose inverse of tensor via Einstein product. Linear Multilinear Algebra (2020) https://doi.org/10.1080/03081087.2020.1748848
31. Smilde, A., Bro, R., Geladi, P.: Multi-Way Analysis: Applications in the Chemical Sciences. Wiley, West Sussex (2004)
32. Sun, L., Zheng, B., Bu, C., Wei, Y.: Moore-Penrose inverse of tensors via Einstein product. Linear Multilinear Algebra **64**, 686–698 (2016)
33. Vlasic, D., Brand, M., Pfister, H., Popovic, J.: Face transfer with multilinear models. ACM Trans. Graph. **24**, 426–433 (2005)
34. Wang, Q.W., Xu, X.: Iterative algorithms for solving some tensor equations. Linear Multilinear Algebra **67**, 1325–1349 (2019)