



Perturbation theory for Moore–Penrose inverse of tensor via Einstein product

Haifeng Ma¹ · Na Li¹ · Predrag S. Stanimirović² · Vasilios N. Katsikis³

Received: 12 January 2019 / Revised: 28 March 2019 / Accepted: 16 May 2019 / Published online: 22 May 2019
© SBMAC - Sociedade Brasileira de Matemática Aplicada e Computacional 2019

Abstract

The spectral norm of an even-order tensor is defined and investigated. An equivalence between the spectral norm of tensors and matrices is given. Using derived representations of some tensor expressions involving the Moore–Penrose inverse, we investigate the perturbation theory for the Moore–Penrose inverse of tensor via Einstein product. The classical results derived by Stewart (SIAM Rev 19:634–662, 1977) and Wedin (BIT 13:217–232, 1973) for the matrix case are extended to even-order tensors. An implementation in the *Matlab* programming language is developed and used in deriving appropriate numerical examples.

Keywords Moore–Penrose inverse · condition number · Perturbation bound · Tensor · Einstein product

Mathematics Subject Classification 15A09 · 65F20

Communicated by Jinyun Yuan.

Haifeng Ma is supported by the bilateral project between China and Poland (No. 37-18). Predrag S. Stanimirović gratefully acknowledges support from the Research Project 174013 of the Serbian Ministry of Science.

✉ Predrag S. Stanimirović
pecko@pmf.ni.ac.rs

Haifeng Ma
haifengma@aliyun.com

Na Li
1755203895@qq.com

Vasilios N. Katsikis
vaskatsikis@econ.uoa.gr

¹ School of Mathematical Science, Harbin Normal University, Harbin 150025, People's Republic of China

² Faculty of Sciences and Mathematics, University of Niš, Višegradska 33, Niš 18000, Serbia

³ National and Kapodistrian University of Athens, Sofokleous 1 Street, 10559 Athens, Greece

1 Introduction

For a positive integer N , let $\mathbf{I}_1, \dots, \mathbf{I}_N$ be positive integers. An order N tensor $\mathcal{A} = (\mathcal{A}_{i_1 i_2 \dots i_N})_{1 \leq i_j \leq \mathbf{I}_j, (j = 1, \dots, N)}$ is a multidimensional array with $\mathfrak{J} = \mathbf{I}_1 \mathbf{I}_2 \dots \mathbf{I}_N$ entries, where $\mathbf{I}_1, \dots, \mathbf{I}_N$ are positive integers. Let $\mathbb{C}^{\mathbf{I}_1 \times \dots \times \mathbf{I}_N}$ (resp. $\mathbb{R}^{\mathbf{I}_1 \times \dots \times \mathbf{I}_N}$) be the set of the order N tensors of dimension $\mathbf{I}_1 \times \dots \times \mathbf{I}_N$ over complex numbers \mathbb{C} (resp. real numbers \mathbb{R}).

The conjugate transpose of a tensor $\mathcal{A} = (\mathcal{A}_{i_1 \dots i_M j_1 \dots j_N}) \in \mathbb{C}^{\mathbf{I}_1 \times \dots \times \mathbf{I}_M \times \mathbf{J}_1 \times \dots \times \mathbf{J}_N}$ is denoted by \mathcal{A}^* and elementwise defined as $(\mathcal{A}^*)_{j_1 \dots j_N i_1 \dots i_M} = \overline{(\mathcal{A})_{i_1 \dots i_M j_1 \dots j_N}} \in \mathbb{C}^{\mathbf{J}_1 \times \dots \times \mathbf{J}_N \times \mathbf{I}_1 \times \dots \times \mathbf{I}_M}$, where the overline means the conjugate operator. When the tensors are defined over \mathbb{R} , the tensor \mathcal{A}^T satisfying $(\mathcal{A}^T)_{j_1 \dots j_N i_1 \dots i_M} = (\mathcal{A})_{i_1 \dots i_M j_1 \dots j_N} \in \mathbb{C}^{\mathbf{J}_1 \times \dots \times \mathbf{J}_N \times \mathbf{I}_1 \times \dots \times \mathbf{I}_M}$ is called the transpose of \mathcal{A} .

The Einstein product of tensors is defined in Einstein (2007) by the operation $*_N$ via

$$(\mathcal{A} *_N \mathcal{B})_{i_1 \dots i_N j_1 \dots j_M} = \sum_{k_1 \dots k_N} \mathcal{A}_{i_1 \dots i_N k_1 \dots k_N} \mathcal{B}_{k_1 \dots k_N j_1 \dots j_M}, \tag{1.1}$$

where $\mathcal{A} \in \mathbb{C}^{\mathbf{I}_1 \times \dots \times \mathbf{I}_N \times \mathbf{K}_1 \times \dots \times \mathbf{K}_N}$, $\mathcal{B} \in \mathbb{C}^{\mathbf{K}_1 \times \dots \times \mathbf{K}_N \times \mathbf{J}_1 \times \dots \times \mathbf{J}_M}$ and $\mathcal{A} *_N \mathcal{B} \in \mathbb{C}^{\mathbf{I}_1 \times \dots \times \mathbf{I}_N \times \mathbf{J}_1 \times \dots \times \mathbf{J}_M}$. The associative law of this tensor product holds. In the above formula, when $\mathcal{B} \in \mathbb{C}^{\mathbf{K}_1 \times \dots \times \mathbf{K}_N}$, then

$$(\mathcal{A} *_N \mathcal{B})_{i_1 i_2 \dots i_N} = \sum_{k_1, \dots, k_N} \mathcal{A}_{i_1 \dots i_N k_1 \dots k_N} \mathcal{B}_{k_1 \dots k_N},$$

where $\mathcal{A} *_N \mathcal{B} \in \mathbb{C}^{\mathbf{I}_1 \times \dots \times \mathbf{I}_N}$. When $\mathcal{A} \in \mathbb{C}^{\mathbf{I}_1 \times \dots \times \mathbf{I}_N}$ and \mathcal{B} is a vector $\mathbf{b} = (b_i) \in \mathbb{C}^{\mathbf{I}_N}$, the product is defined by operation \times_N via

$$(\mathcal{A} \times_N \mathcal{B})_{i_1 i_2 \dots i_{N-1}} = \sum_{i_N} \mathcal{A}_{i_1 \dots i_N} b_{i_N},$$

where $\mathcal{A} \times_N \mathcal{B} \in \mathbb{C}^{\mathbf{I}_1 \times \dots \times \mathbf{I}_{N-1}}$.

Definition 1.1 (Sun et al. 2016) Let $\mathcal{A} \in \mathbb{C}^{\mathbf{I}_1 \times \dots \times \mathbf{I}_N \times \mathbf{K}_1 \times \dots \times \mathbf{K}_N}$. The tensor $\mathcal{X} \in \mathbb{C}^{\mathbf{K}_1 \times \dots \times \mathbf{K}_N \times \mathbf{I}_1 \times \dots \times \mathbf{I}_N}$ which satisfies

$$\begin{aligned} (1^T) \quad & \mathcal{A} *_N \mathcal{X} *_N \mathcal{A} = \mathcal{A}; & (2^T) \quad & \mathcal{X} *_N \mathcal{A} *_N \mathcal{X} = \mathcal{X}; \\ (3^T) \quad & (\mathcal{A} *_N \mathcal{X})^* = \mathcal{A} *_N \mathcal{X}; & (4^T) \quad & (\mathcal{X} *_N \mathcal{A})^* = \mathcal{X} *_N \mathcal{A} \end{aligned}$$

is called the Moore–Penrose inverse of \mathcal{A} , abbreviated by M-P inverse, denoted by \mathcal{A}^\dagger . If the equation (i) of the above equations (1^T)–(4^T) holds, then \mathcal{X} is called an (i)-inverse of \mathcal{A} , denoted by $\mathcal{A}^{(i)}$.

For a tensor $\mathcal{A} \in \mathbb{C}^{\mathbf{I}_1 \times \dots \times \mathbf{I}_N \times \mathbf{I}_1 \times \dots \times \mathbf{I}_N}$, if there exists a tensor \mathcal{X} , such that $\mathcal{A} *_N \mathcal{X} = \mathcal{X} *_N \mathcal{A} = \mathcal{I}$, then \mathcal{X} is called the inverse of \mathcal{A} , denoted by \mathcal{A}^{-1} . Clearly, if \mathcal{A} is invertible, then $\mathcal{A}^\dagger = \mathcal{A}^{-1}$.

The Moore–Penrose inverse of matrices and linear operators plays an important role in theoretical study and numerical analysis in many areas, such as the singular matrix problems, ill-posed problems, optimization problems, total least-squares problem (Xie et al. 2019; Zheng et al. 2017), and statistical problems (Ben-Israel and Greville 2003; Cvetkovic-Ilicic and Wei 2017; Wang et al. 2018; Wei 2014). As a continuation of these results, operations with tensors have become increasingly prevalent in recent years (Che and Wei 2019; Che et al. 2019; Ding and Wei 2016; Harrison and Joseph 2016; Medellin et al. 2016; Qi and Luo 2017; Wei and Ding 2016). Brazell et al. (2013) introduced the notion of ordinary tensor inverse.

Sun et al. (2014, 2016) prove the existence and uniqueness of the Moore–Penrose inverse and $\{i, j, k\}$ -inverses of even-order tensors with the Einstein product. Panigrahy and Mishra (2018) and Sun et al. (2014, 2016) defined the Moore–Penrose inverse and $\{i\}$ -inverses ($i = 1, 2, 3, 4$) of even-order tensors with the Einstein product. In addition, the general solutions of some multilinear systems were given in terms of defined generalized inverses. A few further characterizations of different generalized inverses of tensors in conjunction with the new method to compute the Moore–Penrose inverse of tensors were considered in Behera and Mishra (2017). The weighted Moore–Penrose inverse in tensor spaces was introduced in Ji and Wei (2017). In addition, a characterization of the least-squares solutions to a multilinear system as well as the relationship between the weighted minimum-norm least-squares solution of a multilinear system and the weighted Moore–Penrose inverse of its coefficient tensor were considered in Ji and Wei (2017). Sun et al. (2018) defined $\{i\}$ -inverses for $i = 1, 2, 5$ and the group inverse of tensors, assuming a general tensor product. Panigrahy et al. (2018) proved some additional properties of the Moore–Penrose inverse of tensors via the Einstein product and also derived a few necessary and sufficient conditions for the reverse-order law for the Moore–Penrose inverse of tensors. Several new sufficient conditions which ensure the reverse-order law of the weighted Moore–Penrose inverse for even-order square tensors were presented in Panigrahy and Mishra (2019). Recently, Ji and Wei (2018) investigated the Drazin inverse of even-order tensors with Einstein product. Liang and Zheng (2018) defined an iterative algorithm for solving Sylvester tensor equation based on the Einstein product.

Using another definition of the tensor product, some basic properties for order 2 left (right) inverse and product of tensors were given in Bu et al. (2014). The generalized inverse of tensors was established in Jin et al. (2017) using tensor equations and the t -product of tensors. The definition of generalized tensor function via the tensor singular value decomposition based on the t -product was introduced in Miao et al. (2019). In addition, the least-squares solutions of tensor equations as well as an algorithm for generating the Moore–Penrose inverse of a tensor were proposed in Jin et al. (2017), Shi et al. (2013).

On the other hand, the additive and multiplicative perturbation models have been investigated frequently during the past decades. For more details, the reader is referred to the references (Cai et al. 2011; Liu et al. 2008; Meng and Zheng 2010; Stewart 1977; Wedin 1973; Wei and Ling 2010; Wei 1999). The classical results derived by Stewart (1977) and Wedin (1973) have been improved in Li et al. (2013), Xu et al. (2010b). The acute perturbation of the group inverse was investigated in Wei (2017). Some results related to the perturbation of the oblique projectors which include the weighted pseudoinverse were presented in Xu et al. (2008, 2010a). Some optimal perturbation bounds of the weighted Moore–Penrose inverse under the weighted unitary invariant norms, the weighted Q-norms and the weighted F-norms were obtained in Xu et al. (2011). A sharp estimation for the perturbation bounds of weighted Moore–Penrose inverse was considered in Ma (2018). Meyer (1980) presented a perturbation formula with application to Markov chains. The authors extended the formula to the Drazin inverse A^D . Two finite-time convergent Zhang neural network models for time-varying complex matrix Drazin inverse have been presented in Qiao et al. (2018). An explicit formula for perturbations of an outer inverse under certain conditions was given in Zhang and Wei (2008). The perturbation analysis for the nearest $\{1\}$, $\{1, i\}$, and $\{1, 2, i\}$ -inverses with respect to the multiplicative perturbation model was considered in Meng et al. (2017).

In addition, the perturbation theory for the tensor eigenvalue and singular value problems of tensors has been investigated recently. The perturbation bounds of the tensor eigenvalue and singular value problems of even-order tensors were considered in Che et al. (2016).

The explicit estimation of the backward error for the largest eigenvalue of an irreducible nonnegative tensor was given in Li and Ng (2014).

Our intention in the present paper is to extend the results concerning the perturbation of the Moore–Penrose inverse from the complex matrix space to more general results in the tensor space. According to this goal, our intention is to extend the classical results derived in Stewart (1977) and Wedin (1973) for the matrix case to even-order tensors.

The null spaces and the ranges of tensors were introduced in Ji and Wei (2017).

Definition 1.2 (Ji and Wei 2017) For $\mathcal{A} \in \mathbb{C}^{\mathbf{I}_1 \times \dots \times \mathbf{I}_N \times \mathbf{K}_1 \times \dots \times \mathbf{K}_N}$, the range $\mathcal{R}(\mathcal{A})$ and the null space $\mathcal{N}(\mathcal{A})$ of \mathcal{A} are defined by

$$\begin{aligned} \mathcal{R}(\mathcal{A}) &= \{\mathcal{Y} \in \mathbb{C}^{\mathbf{I}_1 \times \dots \times \mathbf{I}_N} : \mathcal{Y} = \mathcal{A} *_N \mathcal{X}, \mathcal{X} \in \mathbb{C}^{\mathbf{K}_1 \times \dots \times \mathbf{K}_N}\} \\ \mathcal{N}(\mathcal{A}) &= \{\mathcal{X} \in \mathbb{C}^{\mathbf{K}_1 \times \dots \times \mathbf{K}_N} : \mathcal{A} *_N \mathcal{X} = \mathcal{O}\}, \end{aligned}$$

where \mathcal{O} is an appropriate zero tensor.

Definition 1.3 (*Orthogonal Projection*) The orthogonal projection onto a subspace $\mathcal{R}(\mathcal{A})$ is denoted by $P_{\mathcal{A}}$ and defined as

$$P_{\mathcal{A}} = \mathcal{A} *_N \mathcal{A}^\dagger$$

Clearly, $P_{\mathcal{A}}$ is the Hermitian and idempotent, and $\mathcal{R}(P_{\mathcal{A}}) = \mathcal{R}(\mathcal{A})$. Similarly

$$R_{\mathcal{A}} = \mathcal{A}^\dagger *_N \mathcal{A}$$

is the projection onto $\mathcal{R}(\mathcal{A}^*)$.

Definition 1.4 (*Complement of projection*) The projection onto $\mathcal{R}(\mathcal{A})^\perp$ will be denoted by

$$P_{\mathcal{A}}^\perp \equiv \mathcal{I} - P_{\mathcal{A}}.$$

Likewise

$$R_{\mathcal{A}}^\perp \equiv \mathcal{I} - R_{\mathcal{A}}.$$

will denote the projection onto $\mathcal{R}(\mathcal{A}^*)^\perp$.

Main contributions of the manuscript can be summarized as follows.

- (1) The spectral norm of a tensor is defined and investigated.
- (2) Useful representations of $\mathcal{A} *_N \mathcal{A}^\dagger$ and $\mathcal{I} - \mathcal{A} *_N \mathcal{A}^\dagger$ are derived.
- (3) The perturbation theory for the Moore–Penrose inverse of even-order tensors via Einstein product is considered using derived representations of some tensor expressions involving the Moore–Penrose inverse. Therefore, derived results represent the first contribution to the perturbation of the Moore–Penrose inverse of tensors.

The rest of this paper is organized as follows. The spectral tensor norm is defined and investigated in Sect. 2. Useful representations of $\mathcal{A} *_N \mathcal{A}^\dagger$ and $\mathcal{I} - \mathcal{A} *_N \mathcal{A}^\dagger$ are derived in Sect. 3. Section 4 generalizes some results from the matrix theory to the perturbation theory for the Moore–Penrose inverse of even-order tensor via Einstein product. Numerical examples are presented in Sect. 5.

2 Spectral norm of tensors

To simplify presentation, we use the additional notation

$$\mathbf{I}(N) = \mathbf{I}_1 \times \cdots \times \mathbf{I}_N, \quad \mathbb{I} = \{I_1, \dots, I_N\},$$

where $\mathbf{I}_1, \dots, \mathbf{I}_N$ are positive integers. Then, the tensor $\mathcal{A} \in \mathbb{C}^{\mathbf{I}_1 \times \cdots \times \mathbf{I}_M \times \mathbf{K}_1 \times \cdots \times \mathbf{K}_N}$ is denoted shortly by $\mathcal{A} \in \mathbb{C}^{\mathbf{I}(M) \times \mathbf{K}(N)}$. The identity tensor \mathcal{I} of the order $\mathbf{I}(N) \times \mathbf{I}(N)$ tensor is defined as in Brazell et al. (2013) by

$$\mathcal{I}_{i_1 \dots i_N j_1 \dots j_N} = \prod_{k=1}^N \delta_{i_k j_k},$$

where

$$\delta_{ij} = \begin{cases} 1, & i = j, \\ 0, & i \neq j \end{cases}$$

denotes the Kronecker delta operator.

The Frobenius inner product of two tensors $\mathcal{A}, \mathcal{B} \in \mathbb{C}^{\mathbf{I}(N)}$ is defined as

$$(\mathcal{A}, \mathcal{B}) = \sum_{i_1=1}^{\mathbf{I}_1} \sum_{i_2=1}^{\mathbf{I}_2} \cdots \sum_{i_N=1}^{\mathbf{I}_N} \mathcal{A}_{i_1 i_2 \dots i_N} \mathcal{B}_{i_1 i_2 \dots i_N}.$$

If $(\mathcal{A}, \mathcal{B}) = 0$, then \mathcal{A} is orthogonal to \mathcal{B} . The Frobenius norm of \mathcal{A} is defined by $\|\mathcal{A}\|_F = \sqrt{(\mathcal{A}, \mathcal{A})}$.

A complex (real) tensor of order m dimension n is defined by $\mathcal{A} = (\mathcal{A}_{i_1 \dots i_m})$, $\mathcal{A}_{i_1 \dots i_m} \in \mathbb{C}(\mathbb{R})$, where $i_j = 1, \dots, n$ for each $j = 1, \dots, m$. If $\mathbf{x} = (x_1, \dots, x_n)^T$ is an n -dimensional vector, then $\mathbf{x}^m = x \otimes x \otimes \cdots \otimes x$ is considered as an m th order n -dimensional rank-one tensor with entries $(x^m)_{i_1, \dots, i_m} = x_{i_1} \dots x_{i_m}$, where “ \otimes ” is the Kronecker product of vectors. Then

$$\mathcal{A}\mathbf{x}^m = \sum_{i_1, i_2, \dots, i_m=1}^n \mathcal{A}_{i_1 i_2 \dots i_m} x_{i_1 i_2 \dots i_m}$$

is the tensor product of \mathcal{A} and \mathbf{x}^m . A tensor–vector multiplication of a tensor $\mathcal{A} = (a_{i_1, \dots, i_m}) \in \mathbb{C}^{n \times \cdots \times n}$ of order m dimension n and an n -dimensional vector $x = (x_1, x_2, \dots, x_n)^T$ is an n -dimensional vector $\mathcal{A}x^{m-1}$, whose i th component is equal to

$$(\mathcal{A}x^{m-1})_i = \sum_{i_2, \dots, i_m=1}^n a_{i i_2 \dots i_m} x_{i_2 \dots i_m}.$$

The eigenvalue of a tensor was introduced in Qi (2005). A complex number λ is called an *eigenvalue* of \mathcal{A} and \mathbf{x} is an *eigenvector* of \mathcal{A} associated with λ if the equation

$$(\mathcal{A}\mathbf{x}^{m-1})_i = \lambda x_i^{[m-1]}, \quad \mathbf{x}^{[m-1]} = (x_1^{m-1}, x_2^{m-1}, \dots, x_n^{m-1}),$$

is satisfied.

Recently, Liang et al. (2019) proposed a new definition of the eigenvalue of an even-order square tensor. Let $\mathcal{A} \in \mathbb{C}^{\mathbf{I}_1 \times \cdots \times \mathbf{I}_N \times \mathbf{I}_1 \times \cdots \times \mathbf{I}_N}$ be given. If a nonzero tensor $\mathcal{X} \in \mathbb{C}^{\mathbf{I}_1 \times \cdots \times \mathbf{I}_N}$ and a complex number λ satisfy

$$\mathcal{A} *_{\mathbf{N}} \mathcal{X} = \lambda \mathcal{X}, \tag{2.1}$$

then λ is an eigenvalue of the tensor \mathcal{A} and \mathcal{X} the eigentensor with respect to λ .

In addition, we found the following definition of the tensor spectral norm from Li (2016).

Definition 2.1 (Li 2016) For a given tensor $\mathcal{T} \in \mathbb{R}^{\mathbf{I}_1 \times \mathbf{I}_2 \times \dots \times \mathbf{I}_N}$, the spectral norm of \mathcal{T} , denoted by $\|\mathcal{T}\|_\sigma$, is defined as

$$\|\mathcal{T}\|_\sigma := \max \left\{ \langle \mathcal{T}, \mathbf{x}_1 \otimes \mathbf{x}_2 \otimes \dots \otimes \mathbf{x}_N \rangle : \mathbf{x}_k \in \mathbb{R}^{\mathbf{I}_k}, \|\mathbf{x}_k\|_F = 1, k = 1, \dots, N \right\},$$

where $\|\mathbf{x}\|_F$ denotes the Frobenius norm of the vector \mathbf{x} and $\mathbf{x}_1 \otimes \dots \otimes \mathbf{x}_N$ means the outer product of vectors: $(\mathbf{x}_1 \otimes \dots \otimes \mathbf{x}_N)_{i_1, \dots, i_N} = (\mathbf{x}_1)_{i_1} \dots (\mathbf{x}_N)_{i_N}$.

Essentially, $\|\mathcal{T}\|_\sigma$ is the maximal value of the Frobenius inner product between \mathcal{T} and the rank-one tensor $\mathbf{x}_1 \otimes \dots \otimes \mathbf{x}_N$ whose Frobenius norm is one.

Let the eigenvalues of a complex even-order square tensor are defined as in (2.1). By $\lambda_{\min}(\mathcal{K})$ and $\lambda_{\max}(\mathcal{K})$, we denote the smallest and largest eigenvalue of a tensor \mathcal{K} , respectively. Similarly, $\mu_1(\mathcal{K})$ stands for the largest singular value of a tensor \mathcal{K} .

Lemma 2.1 Let $\mathcal{A} \in \mathbb{C}^{\mathbf{I}^{(N)} \times \mathbf{K}^{(N)}}$. Then, the spectral norm of \mathcal{A} can be defined as

$$\|\mathcal{A}\|_2 = \sqrt{\lambda_{\max}(\mathcal{A}^* *_N \mathcal{A})} = \mu_1(\mathcal{A}), \tag{2.2}$$

where $\lambda_{\max}(\mathcal{A}^* *_N \mathcal{A})$ denotes the largest eigenvalue of $\mathcal{A}^* *_N \mathcal{A}$ and $\mu_1(\mathcal{A})$ is the largest singular value of \mathcal{A} .

Proof It is necessary to verify that the definition (2.2) satisfies properties of a norm function.

- (1) Clearly, $\|\mathcal{A}\|_2 \geq 0$, and $\|\mathcal{A}\|_2 = 0$ if and only if $\mathcal{A} = 0$,
- (2) The second property of $\|\mathcal{A}\|_2$ can be verified using

$$\begin{aligned} \|k\mathcal{A}\|_2 &= \sqrt{\lambda_{\max}((k\mathcal{A})^* *_N (k\mathcal{A}))} \\ &= \sqrt{k^2 \lambda_{\max}(\mathcal{A}^* *_N \mathcal{A})} \\ &= |k| \mu_1(\mathcal{A}) \\ &= |k| \|\mathcal{A}\|_2. \end{aligned}$$

- (3) Since

$$\mu_1(\mathcal{A} + \mathcal{B}) \leq \mu_1(\mathcal{A}) + \mu_1(\mathcal{B}),$$

immediately from the definition of the spectral norm it follows that

$$\|\mathcal{A} + \mathcal{B}\|_2 \leq \|\mathcal{A}\|_2 + \|\mathcal{B}\|_2.$$

Therefore, (2.2) is a valid definition of the matrix norm. □

Our intention is to determine the spectral norm of a tensor explicitly using the approach based on the *matricization* or *unfolding*. Matricization is the transformation that transforms a tensor into a matrix. Let $\mathbf{I}_1, \dots, \mathbf{I}_M, \mathbf{K}_1, \dots, \mathbf{K}_N$ be positive integers. Assume that $\mathfrak{J}, \mathfrak{K}$ are positive integers defined by

$$\mathfrak{J} = \mathbf{I}_1 \mathbf{I}_2 \dots \mathbf{I}_M, \quad \mathfrak{K} = \mathbf{K}_1 \mathbf{K}_2 \dots \mathbf{K}_N. \tag{2.3}$$

Denote by $\text{Mat}(\mathcal{A})$ the matrix obtained after the matricization

$$\text{Mat} : \mathbb{C}^{\mathbf{I}^{(M)} \times \mathbf{K}^{(N)}} \mapsto \mathbb{C}^{\mathfrak{J} \times \mathfrak{K}},$$

which transforms a tensor $\mathcal{A} \in \mathbb{C}^{\mathbf{I}(M) \times \mathbf{K}(N)}$ into the matrix $A \in \mathbb{C}^{\mathcal{J} \times \mathcal{K}}$. An arbitrary tensor \mathcal{A} can be unfolded into an appropriate matrix A in different ways.

It is known that the spectral norm of the tensor is bounded by the spectral norm of the matricized tensor, i.e., $\|\text{Mat}(\mathcal{A})\|_\sigma \geq \|\mathcal{A}\|_\sigma$ (see, for example, Li 2016).

One approach in the matricization, denoted by $\psi : \mathbb{C}^{\mathbf{I}(M) \times \mathbf{K}(N)} \mapsto \mathbb{C}^{\mathcal{J} \times \mathcal{K}}$ was proposed in Liang and Zheng (2018, Definition 2.4) (see also Brazell et al. 2013). The matricization ψ is defined by

$$\psi(\mathcal{A}_{i_1, \dots, i_M, k_1, \dots, k_N}) = A_{\text{ivec}(i, \mathbb{I}), \text{ivec}(k, \mathbb{K})},$$

where $i = (i_1, \dots, i_M)^T$ and

$$\text{ivec}(i, \mathbb{I}) = i_1 + \sum_{j=2}^M (i_j - 1) \prod_{s=1}^{j-1} I_s.$$

To define an effective procedure for the tensor matricization, we use the reshaping operation denoted as *rsh*, which was introduced in Panigrahy et al. (2018). Later, we define the spectral norm of a tensor by means of the spectral norm of the reshaped tensor. This operation can be implemented by means of the standard *Matlab* function `reshape`.

Definition 2.2 (Panigrahy et al. 2018) Let $\mathbf{I}_1, \dots, \mathbf{I}_M, \mathbf{K}_1, \dots, \mathbf{K}_N$ be given integers. Assume that \mathcal{J}, \mathcal{K} are the integers defined (2.3). The reshaping operation

$$\text{rsh} : \mathbb{C}^{\mathbf{I}(M) \times \mathbf{K}(N)} \mapsto \mathbb{C}^{\mathcal{J} \times \mathcal{K}},$$

transforms a tensor $\mathcal{A} \in \mathbb{C}^{\mathbf{I}(M) \times \mathbf{K}(N)}$ into the matrix $A \in \mathbb{C}^{\mathcal{J} \times \mathcal{K}}$ using the *Matlab* function `reshape` as follows:

$$\text{rsh}(\mathcal{A}) = A = \text{reshape}(\mathcal{A}, \mathcal{J}, \mathcal{K}), \quad \mathcal{A} \in \mathbb{C}^{\mathbf{I}(M) \times \mathbf{K}(N)}, \quad A \in \mathbb{C}^{\mathcal{J} \times \mathcal{K}}.$$

The inverse reshaping is the mapping defined by

$$\text{rsh}^{-1} : \mathbb{C}^{\mathcal{J} \times \mathcal{K}} \mapsto \mathbb{C}^{\mathbf{I}(M) \times \mathbf{K}(N)},$$

$$\text{rsh}^{-1}(A) = \mathcal{A} = \text{reshape}(A, \mathbf{I}_1, \dots, \mathbf{I}_M, \mathbf{K}_1, \dots, \mathbf{K}_N), \quad A \in \mathbb{C}^{\mathcal{J} \times \mathcal{K}}, \quad \mathcal{A} \in \mathbb{C}^{\mathbf{I}(M) \times \mathbf{K}(N)}.$$

The following result from Panigrahy et al. (2018) will be useful.

Lemma 2.2 (Panigrahy et al. 2018) Let $\mathcal{A} \in \mathbb{C}^{\mathbf{I}_1 \times \dots \times \mathbf{I}_N \times \mathbf{K}_1 \times \dots \times \mathbf{K}_N}$ and $\mathcal{B} \in \mathbb{C}^{\mathbf{K}_1 \times \dots \times \mathbf{K}_N \times \mathbf{L}_1 \times \dots \times \mathbf{L}_N}$ be given tensors, integers \mathcal{J}, \mathcal{K} are computed as in (2.3) and $\mathcal{L} = \mathbf{L}_1 \mathbf{L}_2 \dots \mathbf{L}_N$. Then

$$\text{rsh}(\mathcal{A} *_N \mathcal{B}) = \text{rsh}(\mathcal{A}) \text{rsh}(\mathcal{B}) = AB, \tag{2.4}$$

where $A = \text{rsh}(\mathcal{A}) \in \mathbb{C}^{\mathcal{J} \times \mathcal{K}}, B = \text{rsh}(\mathcal{B}) \in \mathbb{C}^{\mathcal{K} \times \mathcal{L}}$.

Applying the inverse reshaping operator $\text{rsh}^{-1}(\cdot)$ on both sides in (2.4), it can be concluded that $\text{rsh}^{-1}(AB) = \text{rsh}^{-1}(A) *_N \text{rsh}^{-1}(B) = \mathcal{A} *_N \mathcal{B}$.

Now, our intention is to approximate the tensor norm $\|\mathcal{A}\|_2$ by an effective computational procedure. For this purpose, we propose Algorithm 1 for computing $\text{rsh}^{-1}(A)$ in terms of the Singular Value Decomposition (SVD) of A . Since eigenvalues in Liang et al. (2019) are defined for even-order square tensors, our further investigation will be restricted to even-order tensors.

Algorithm 1 Computation of $\text{rsh}^{-1}(A)$

Require: Even integer N , integers $\mathbf{I}_1, \dots, \mathbf{I}_N, \mathbf{K}_1, \dots, \mathbf{K}_N$ positive integers \mathfrak{J} and \mathfrak{R} satisfying (2.3), and $A \in \mathbb{C}^{\mathfrak{J} \times \mathfrak{R}}$.

1: Compute Singular Value Decomposition (SVD) of A in the form:

$$[U, D, V] = \text{SVD}(A) = UDV^*,$$

where $U \in \mathbb{C}^{\mathfrak{J} \times \mathfrak{J}}$ and $V \in \mathbb{C}^{\mathfrak{R} \times \mathfrak{R}}$ are unitary matrices and $D \in \mathbb{C}^{\mathfrak{J} \times \mathfrak{R}}$ is diagonal with singular values of A on the main diagonal.

2: Perform the reshaping operations

$$\text{rsh}^{-1}(U) = U \in \mathbb{C}^{\mathbf{I}(N) \times \mathbf{I}(N)}, \quad \text{rsh}^{-1}(V^*) = V^* \in \mathbb{C}^{\mathbf{K}(N) \times \mathbf{K}(N)}, \quad \text{rsh}^{-1}(D) = \mathcal{D} \in \mathbb{C}^{\mathbf{I}(N) \times \mathbf{I}(N)}.$$

3: Compute the output

$$\mathcal{A} = U *_N \mathcal{D} *_N V^* \in \mathbb{C}^{\mathbf{I}(N) \times \mathbf{K}(N)}.$$

In Lemma 2.3, we show that the tensor \mathcal{A} in Algorithm 1 is defined well.

Lemma 2.3 *The tensor \mathcal{A} in Algorithm 1 is defined well.*

Proof Under the assumptions of Algorithm 1, an application of Lemma 2.2 gives

$$\begin{aligned} \text{rsh}(\mathcal{A}) &= \text{rsh}(U *_N \mathcal{D} *_N V^*) \\ &= \text{rsh}(U) \text{rsh}(\mathcal{D}) \text{rsh}(V^*) \\ &= UDV^* \\ &= A, \end{aligned}$$

which confirms $\mathcal{A} = \text{rsh}^{-1}(A)$. □

As a consequence of Algorithm 1, Lemma 2.4 shows that the spectral norm is invariant with respect to the function rsh .

Lemma 2.4 *Let $\mathcal{A} \in \mathbb{C}^{\mathbf{I}_1 \times \dots \times \mathbf{I}_N \times \mathbf{K}_1 \times \dots \times \mathbf{K}_N}$ be a given tensor and integers $\mathfrak{J}, \mathfrak{R}$ are computed as in (2.3). Then*

$$\|\mathcal{A}\|_2 = \|\text{rsh}^{-1}(\mathcal{A})\|_2 = \|A\|_2, \tag{2.5}$$

where $A \in \mathbb{C}^{\mathfrak{J} \times \mathfrak{R}}$.

Proof According to Algorithm 1, the tensor $\mathcal{A} \in \mathbb{C}^{\mathbf{I}_1 \times \dots \times \mathbf{I}_N \times \mathbf{K}_1 \times \dots \times \mathbf{K}_N}$ possesses the same singular values as the matrix $A \in \mathbb{C}^{\mathfrak{J} \times \mathfrak{R}}$. □

Example 2.1 Let $\mathcal{A} = \text{rand}(2, 2, 2, 2)$ is defined by

$$\begin{aligned} \mathcal{A}(:, :, 1, 1) &= \begin{bmatrix} 0.8147 & 0.1270 \\ 0.9058 & 0.9134 \end{bmatrix}, \quad \mathcal{A}(:, :, 2, 1) = \begin{bmatrix} 0.6324 & 0.2785 \\ 0.0975 & 0.5469 \end{bmatrix}, \\ \mathcal{A}(:, :, 1, 2) &= \begin{bmatrix} 0.9575 & 0.1576 \\ 0.9649 & 0.9706 \end{bmatrix}, \quad \mathcal{A}(:, :, 2, 2) = \begin{bmatrix} 0.9572 & 0.8003 \\ 0.4854 & 0.1419 \end{bmatrix}, \end{aligned}$$

then

$$A = \text{rsh}^{-1}(\mathcal{A}) = \text{reshape}(\mathcal{A}, 4, 4) = \begin{bmatrix} 0.8147 & 0.6324 & 0.9575 & 0.9572 \\ 0.9058 & 0.0975 & 0.9649 & 0.4854 \\ 0.1270 & 0.2785 & 0.1576 & 0.8003 \\ 0.9134 & 0.5469 & 0.9706 & 0.1419 \end{bmatrix}.$$

Simple verification shows that $\|\mathcal{A}\|_2 = \|A\|_2 = 2.6201$.

Various definitions of the tensor rank can be found in the relevant literature. For more details see Brazell et al. (2013), and Comon et al. (2009). An alternative definition of the tensor rank was introduced in Panigrahy et al. (2018).

Definition 2.3 (Panigrahy et al. 2018) Let $\mathcal{A} \in \mathbb{C}^{\mathbf{I}(N) \times \mathbf{K}(N)}$ and $A = \text{reshape}(\mathcal{A}, \mathfrak{J}, \mathfrak{K}) = \text{rsh}(\mathcal{A}) \in \mathbb{C}^{\mathfrak{J} \times \mathfrak{K}}$ are defined as in Algorithm 1. Then, the tensor rank of \mathcal{A} is denoted by $\text{rshrank}(\mathcal{A})$ and defined by $\text{rshrank}(\mathcal{A}) = \text{rank}(A)$.

3 Preliminary results

For $a \in \mathbb{C}$, let $a^\dagger = a^{-1}$, if $a \neq 0$ and $a^\dagger = 0$, if $a = 0$. Following this notation, the tensor $\mathcal{D} \in \mathbb{C}^{\mathbf{I}_1 \times \dots \times \mathbf{I}_N \times \mathbf{I}_1 \times \dots \times \mathbf{I}_N}$ is called diagonal if all its entries are zero except $\mathcal{D}_{i_1 \dots i_N i_1 \dots i_N}$, that is

$$\mathcal{D}_{i_1 \dots i_N j_1 \dots j_N} = \begin{cases} 0, & (i_1, \dots, i_N) \neq (j_1, \dots, j_N), \\ \mathcal{D}_{i_1 \dots i_N i_1 \dots i_N}, & (i_1, \dots, i_N) = (j_1, \dots, j_N), \end{cases} \tag{3.1}$$

where $\mathcal{D}_{i_1 \dots i_N i_1 \dots i_N}$ is a complex number. Particularly, a diagonal tensor becomes a unit tensor in this case $\mathcal{D}_{i_1 \dots i_N j_1 \dots j_N} = \delta_{i_1 j_1} \dots \delta_{i_N j_N}$, where

$$\delta_{lk} = \begin{cases} 1, & l = k, \\ 0, & l \neq k \end{cases}$$

is the Kronecker delta, then \mathcal{D} is a unit tensor, denoted by \mathcal{I} . It follows from Definition 1.1 that the Moore–Penrose inverse $\mathcal{D}^\dagger \in \mathbb{C}^{\mathbf{I}_1 \times \dots \times \mathbf{I}_N \times \mathbf{I}_1 \times \dots \times \mathbf{I}_N}$ of the diagonal tensor defined in (3.1) is equal to

$$(\mathcal{D}^\dagger)_{j_1 \dots j_N i_1 \dots i_N} = \begin{cases} \frac{1}{\mathcal{D}_{i_1 \dots i_N j_1 \dots j_N}}, & \mathcal{D}_{i_1 \dots i_N j_1 \dots j_N} \neq 0, \\ 0, & \mathcal{D}_{i_1 \dots i_N j_1 \dots j_N} = 0. \end{cases}$$

It is easy to see that if \mathcal{D} is a diagonal tensor, then $\mathcal{D} *_N \mathcal{D}^\dagger$ and $\mathcal{D}^\dagger *_N \mathcal{D}$ are diagonal tensors, whose diagonal entries are 1 or 0.

The tensor $\mathcal{A} \in \mathbb{C}^{\mathbf{I}_1 \times \dots \times \mathbf{I}_N \times \mathbf{I}_1 \times \dots \times \mathbf{I}_N}$ is orthogonal if $\mathcal{A} *_N \mathcal{A}^* = \mathcal{A}^* *_N \mathcal{A} = \mathcal{I}$.

The computation of the Moore–Penrose inverse of a tensor was proposed in Brazell et al. (2013), Sun et al. (2016). This method is restated in Lemma 3.1.

Lemma 3.1 (Brazell et al. 2013; Sun et al. 2016) For a tensor $\mathcal{A} \in \mathbb{C}^{\mathbf{I}(N) \times \mathbf{K}(N)}$, the singular value decomposition (SVD) of \mathcal{A} has the form:

$$\mathcal{A} = \mathcal{U} *_N \mathcal{D} *_N \mathcal{V}^*, \tag{3.2}$$

where $\mathcal{U} \in \mathbb{C}^{\mathbf{I}(N) \times \mathbf{I}(N)}$ and $\mathcal{V} \in \mathbb{C}^{\mathbf{K}(N) \times \mathbf{K}(N)}$ are orthogonal tensors, $\mathcal{D} \in \mathbb{C}^{\mathbf{I}(N) \times \mathbf{K}(N)}$ is a diagonal tensor satisfying

$$\mathcal{D}_{i_1 \dots i_N k_1 \dots k_N} = \begin{cases} 0, & (i_1, \dots, i_N) \neq (k_1, \dots, k_N), \\ \mu_{i_1 \dots i_N}, & (i_1, \dots, i_N) = (k_1, \dots, k_N), \end{cases}$$

wherein $\mu_{i_1 \dots i_N}$ are the singular values of \mathcal{A} . Then

$$\mathcal{A}^\dagger = \mathcal{V} *_N \mathcal{D}^\dagger *_N \mathcal{U}^*, \tag{3.3}$$

where

$$(\mathcal{D}^\dagger)_{k_1 \dots k_N i_1 \dots i_N} = \begin{cases} 0, & (i_1 \dots i_N) \neq (k_1 \dots k_N), \\ (\mu_{i_1 \dots i_N})^\dagger, & (i_1 \dots i_N) = (k_1 \dots k_N), \end{cases}$$

and

$$(\mu_{i_1 \dots i_N})^\dagger = \begin{cases} 0, & \mu_{i_1 \dots i_N} = 0, \\ (\mu_{i_1 \dots i_N})^{-1}, & \mu_{i_1 \dots i_N} \neq 0. \end{cases}$$

An effective algorithm for computing the Moore–Penrose inverse of a tensor in the form (3.3) was presented in Algorithm 1 from Huang et al. (2018). To compute the Moore–Penrose inverse by means of (3.3), it is necessary to compute the transpose of a tensor. For this purpose, we developed the following Algorithm 2.

Algorithm 2 Computation of the tensor transpose \mathcal{A}^T

Require: Tensor $\mathcal{A} \in \mathbb{C}^{\mathbf{I}_1 \times \dots \times \mathbf{I}_N \times \mathbf{K}_1 \times \dots \times \mathbf{K}_N}$, where N is a positive integer, integers $\mathbf{I}_1, \dots, \mathbf{I}_N, \mathbf{K}_1, \dots, \mathbf{K}_N$ integers \mathfrak{J} and \mathfrak{K} satisfying (2.3).

- 1: Compute $A = \text{rsh}(\mathcal{A}) \in \mathbb{C}^{\mathfrak{J} \times \mathfrak{K}}$.
- 2: Compute the matrix transpose A^T of A .
- 3: Perform the reshaping operation and compute the output

$$\mathcal{A}^T = \text{rsh}^{-1}(A^T) \in \mathbb{C}^{\mathbf{K}_1 \times \dots \times \mathbf{K}_N \times \mathbf{I}_1 \times \dots \times \mathbf{I}_N}.$$

Example 3.1 This example is aimed to verification of Algorithm 2. Consider $\mathcal{A} = \text{rand}(2, 2, 2, 2)$ equal to

$$\begin{aligned} \mathcal{A}(:, :, 1, 1) &= \begin{bmatrix} 0.8147 & 0.1270 \\ 0.9058 & 0.9134 \end{bmatrix}; & \mathcal{A}(:, :, 2, 1) &= \begin{bmatrix} 0.6324 & 0.2785 \\ 0.0975 & 0.5469 \end{bmatrix}; \\ \mathcal{A}(:, :, 1, 2) &= \begin{bmatrix} 0.9575 & 0.1576 \\ 0.9649 & 0.9706 \end{bmatrix}; & \mathcal{A}(:, :, 2, 2) &= \begin{bmatrix} 0.9572 & 0.8003 \\ 0.4854 & 0.1419 \end{bmatrix}. \end{aligned}$$

Then, $A = \text{rsh}(\mathcal{A}, 4, 4)$ is equal to

$$A = \begin{bmatrix} 0.8147 & 0.6324 & 0.9575 & 0.9572 \\ 0.9058 & 0.0975 & 0.9649 & 0.4854 \\ 0.1270 & 0.2785 & 0.1576 & 0.8003 \\ 0.9134 & 0.5469 & 0.9706 & 0.1419 \end{bmatrix}.$$

Furthermore, the results of Algorithm 2 is equal to $\mathcal{A} = \text{reshape}(A^T, 2, 2, 2, 2)$, which gives

$$\begin{aligned} \mathcal{A}^T(:, :, 1, 1) &= \begin{bmatrix} 0.8147 & 0.9575 \\ 0.6324 & 0.9572 \end{bmatrix}; & \mathcal{A}^T(:, :, 2, 1) &= \begin{bmatrix} 0.9058 & 0.9649 \\ 0.0975 & 0.4854 \end{bmatrix}; \\ \mathcal{A}^T(:, :, 1, 2) &= \begin{bmatrix} 0.1270 & 0.1576 \\ 0.2785 & 0.8003 \end{bmatrix}; & \mathcal{A}^T(:, :, 2, 2) &= \begin{bmatrix} 0.9134 & 0.9706 \\ 0.5469 & 0.1419 \end{bmatrix}. \end{aligned} \tag{3.4}$$

On the other hand, a direct calculation gives

$$\begin{aligned}
 a_{1111} &= 0.8147 = \mathcal{A}_{1111}^T; & a_{2111} &= 0.9058 = \mathcal{A}_{1121}^T; \\
 a_{1211} &= 0.1270 = \mathcal{A}_{1112}^T; & a_{2211} &= 0.9134 = \mathcal{A}_{1122}^T; \\
 a_{1121} &= 0.6324 = \mathcal{A}_{2111}^T; & a_{2121} &= 0.0975 = \mathcal{A}_{2121}^T; \\
 a_{1221} &= 0.2785 = \mathcal{A}_{2112}^T; & a_{2221} &= 0.5469 = \mathcal{A}_{2122}^T; \\
 a_{1112} &= 0.9575 = \mathcal{A}_{1211}^T; & a_{2112} &= 0.9649 = \mathcal{A}_{1221}^T; \\
 a_{1212} &= 0.1576 = \mathcal{A}_{1212}^T; & a_{2212} &= 0.9706 = \mathcal{A}_{1222}^T; \\
 a_{1122} &= 0.9572 = \mathcal{A}_{2211}^T; & a_{2122} &= 0.4854 = \mathcal{A}_{2221}^T; \\
 a_{1222} &= 0.8003 = \mathcal{A}_{2212}^T; & a_{2222} &= 0.1419 = \mathcal{A}_{2222}^T.
 \end{aligned}$$

Therefore, the result of direct calculation coincides with (3.4).

Lemma 3.2 *Let $\mathcal{A} \in \mathbb{C}^{\mathbf{I}(N) \times \mathbf{K}(N)}$ be a given tensor, let μ_{i_1, \dots, i_N} be the singular values of \mathcal{A} and v_{i_1, \dots, i_L} be the nonzero singular values of \mathcal{A} . Then*

$$\|\mathcal{A}\|_2 = \mu_1(\mathcal{A}); \quad \|\mathcal{A}^\dagger\|_2 = \frac{1}{v_{\min}(\mathcal{A})},$$

where $v_{\min}(\mathcal{A})$ denotes the smallest nonzero singular value of \mathcal{A} .

Proof The identity $\|\mathcal{A}\|_2 = \mu_1(\mathcal{A})$ follows from Lemma 2.1. Since v_{i_1, \dots, i_L} are nonzero singular values of \mathcal{A} defined in (3.2), it follows from (3.3) that $(v_{i_1, \dots, i_L})^{-1} > 0$ are the nonzero singular values of \mathcal{A}^\dagger . Accordingly, $\|\mathcal{A}^\dagger\|_2 = \mu_1(\mathcal{A}^\dagger) = \frac{1}{v_{\min}(\mathcal{A})}$. \square

A useful representation for $\mathcal{A} *_N \mathcal{A}^\dagger$ is derived in Lemma 3.3.

Lemma 3.3 *Let $\mathcal{A} \in \mathbb{C}^{\mathbf{I}(N) \times \mathbf{K}(N)}$ be an arbitrary tensor and the positive integers $\mathfrak{J}, \mathfrak{K}$ are defined in (2.3). Then*

$$\mathcal{A} *_N \mathcal{A}^\dagger = \mathcal{U}_A *_N \text{rsh}^{-1} \left(\begin{bmatrix} I_{\mathfrak{J}_R \times \mathfrak{J}_R} & O_{\mathfrak{J}_R \times (\mathfrak{J} - \mathfrak{J}_R)} \\ O_{(\mathfrak{J} - \mathfrak{J}_R) \times \mathfrak{J}_R} & O_{(\mathfrak{J} - \mathfrak{J}_R) \times (\mathfrak{K} - \mathfrak{J}_R)} \end{bmatrix} \right) *_N \mathcal{U}_A^*. \tag{3.5}$$

Proof We follow Algorithm 1 from Huang et al. (2018) to define \mathcal{A}^\dagger and \mathcal{B}^\dagger . According to Step 1, it is necessary to reshape $\mathcal{A} \in \mathbb{C}^{\mathbf{I}(N) \times \mathbf{K}(N)}$ into a matrix $A \in \mathbb{C}^{\mathfrak{J} \times \mathfrak{K}}$, where $\mathfrak{J}, \mathfrak{K}$ are defined in (2.3). This transformation is denoted by $\text{rsh}(\mathcal{A}) = A$. Step 2 assumes the Singular Value Decomposition (SVD) of A of the form $[U_A, D_A, V_A] = \text{SVD}(A)$, which implies $A = U_A D_A V_A^*$, where $U_A \in \mathbb{C}^{\mathfrak{J} \times \mathfrak{J}}$ and $V_A \in \mathbb{C}^{\mathfrak{K} \times \mathfrak{K}}$ are unitary and the matrix $D_A \in \mathbb{C}^{\mathfrak{J} \times \mathfrak{K}}$ is of the diagonal form:

$$D_A = \begin{bmatrix} \Sigma_A & O_{\mathfrak{J}_R \times (\mathfrak{K} - \mathfrak{J}_R)} \\ O_{(\mathfrak{J} - \mathfrak{J}_R) \times \mathfrak{J}_R} & O_{(\mathfrak{J} - \mathfrak{J}_R) \times (\mathfrak{K} - \mathfrak{J}_R)} \end{bmatrix},$$

where

$$\Sigma_A \in \mathbb{C}^{\mathfrak{J}_R \times \mathfrak{J}_R}, \quad \mathfrak{J}_R = \text{rshrank}(A)$$

is diagonal with singular values of A on the main diagonal and

$$O_{\mathfrak{J}_R \times (\mathfrak{K} - \mathfrak{J}_R)} \in \mathbb{C}^{\mathfrak{J}_R \times (\mathfrak{K} - \mathfrak{J}_R)}, \quad O_{(\mathfrak{J} - \mathfrak{J}_R) \times \mathfrak{J}_R} \in \mathbb{C}^{(\mathfrak{J} - \mathfrak{J}_R) \times \mathfrak{J}_R}, \quad O_{(\mathfrak{J} - \mathfrak{J}_R) \times (\mathfrak{K} - \mathfrak{J}_R)} \in \mathbb{C}^{(\mathfrak{J} - \mathfrak{J}_R) \times (\mathfrak{K} - \mathfrak{J}_R)}$$

are appropriate zero blocks. According to Step 3, we perform the reshaping operations:

$$\text{rsh}(U_A) = U_A \in \mathbb{C}^{\mathbf{I}^{(N)} \times \mathbf{I}^{(N)}}, \text{rsh}(V_A^*) = V_A^* \in \mathbb{C}^{\mathbf{K}^{(N)} \times \mathbf{K}^{(N)}}, \text{rsh}(D_A) = D_A \in \mathbb{C}^{\mathbf{I}^{(N)} \times \mathbf{K}^{(N)}}.$$

Then, compute

$$D_A^\dagger = \begin{bmatrix} \Sigma_A^{-1} & O_{\mathcal{J}_R \times (\mathcal{J} - \mathcal{J}_R)} \\ O_{(\mathcal{K} - \mathcal{J}_R) \times \mathcal{J}_R} & O_{(\mathcal{K} - \mathcal{J}_R) \times (\mathcal{J} - \mathcal{J}_R)} \end{bmatrix} \in \mathbb{C}^{\mathcal{K} \times \mathcal{J}}$$

and

$$\mathcal{D}_A^\dagger = \text{rsh}^{-1}(D_A^\dagger) \in \mathbb{C}^{\mathbf{K}^{(N)} \times \mathbf{I}^{(N)}}.$$

According to Step 4 of Algorithm 1 from Huang et al. (2018)

$$\mathcal{A}^\dagger = V_A *_{\mathcal{N}} \mathcal{D}_A^\dagger *_{\mathcal{N}} U_A^*.$$

Now, the tensor \mathcal{A} possesses the representation:

$$\mathcal{A} = U_A *_{\mathcal{N}} D_A *_{\mathcal{N}} V_A^*.$$

Later, one can verify

$$D_A *_{\mathcal{N}} \mathcal{D}_A^\dagger = \text{rsh}^{-1} \left(\begin{bmatrix} I_{\mathcal{J}_R \times \mathcal{J}_R} & O_{\mathcal{J}_R \times (\mathcal{J} - \mathcal{J}_R)} \\ O_{(\mathcal{J} - \mathcal{J}_R) \times \mathcal{J}_R} & O_{(\mathcal{J} - \mathcal{J}_R) \times (\mathcal{J} - \mathcal{J}_R)} \end{bmatrix} \right) \in \mathbb{C}^{\mathbf{I}^{(N)} \times \mathbf{I}^{(N)}},$$

where $I_{\mathcal{J}_R \times \mathcal{J}_R}$ is the identity $\mathcal{J}_R \times \mathcal{J}_R$ matrix. Consequently, $\mathcal{A} *_{\mathcal{N}} \mathcal{A}^\dagger$ possesses the representation (3.5). □

The result of Proposition 3.1 will be useful.

Proposition 3.1 (Meng and Zheng 2010) *Let $W \in \mathbb{C}^{n \times n}$ be a unitary matrix with the block form:*

$$W = \begin{bmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{bmatrix}, W_{11} \in \mathbb{C}^{r \times r}, W_{22} \in \mathbb{C}^{(n-r) \times (n-r)}, 1 \leq r < n.$$

Then, $\|W_{12}\| = \|W_{21}\|$ for any unitarily invariant norm.

4 Main results

For the sake of convenience, we assume that the following condition holds

$$\begin{aligned} \mathcal{A}, \mathcal{E} \in \mathbb{C}^{\mathbf{I}^{(N)} \times \mathbf{K}^{(N)}}, \mathcal{B} = \mathcal{A} + \mathcal{E}, \text{rshrank}(\mathcal{A}) = \text{rshrank}(\mathcal{B}) = r \\ \Delta = \|\mathcal{A}^\dagger\|_2 \|\mathcal{E}\|_2 < 1. \end{aligned} \tag{4.1}$$

Lemma 4.1 *If the Condition (4.1) is satisfied, then*

$$\|\mathcal{B}^\dagger\|_2 \leq \frac{\|\mathcal{A}^\dagger\|_2}{1 - \Delta} \tag{4.2}$$

Proof According to the Lemma 3.2, we get

$$\|\mathcal{B}^\dagger\|_2 = \frac{1}{\nu_{\min}(\mathcal{A} + \mathcal{E})},$$

so

$$\begin{aligned} \frac{1}{\|\mathcal{B}^\dagger\|_2} &= \nu_{\min}(\mathcal{A} + \mathcal{E}) \geq \nu_{\min}(\mathcal{A}) - \nu_{\min}(\mathcal{E}) \geq \nu_{\min}(\mathcal{A}) - \mu_1(\mathcal{E}) \\ &\geq \nu_{\min}(\mathcal{A}) - \|\mathcal{E}\|_2 \\ &= \frac{1}{\|\mathcal{A}^\dagger\|_2} - \|\mathcal{E}\|_2. \end{aligned}$$

Then

$$\|\mathcal{B}^\dagger\|_2 \leq \frac{\|\mathcal{A}^\dagger\|_2}{1 - \|\mathcal{A}^\dagger\|_2 \|\mathcal{E}\|_2} = \frac{\|\mathcal{A}^\dagger\|_2}{1 - \Delta},$$

which completes the proof. □

Next, we give the decomposition of $\mathcal{B}^\dagger - \mathcal{A}^\dagger$.

Theorem 4.1 *Let $\mathcal{A}, \mathcal{E} \in \mathbb{C}^{\mathbf{I}(N) \times \mathbf{K}(N)}$ and $\mathcal{B} = \mathcal{A} + \mathcal{E}$. Then*

$$\mathcal{B}^\dagger - \mathcal{A}^\dagger = -\mathcal{B}^\dagger *_N \mathcal{E} *_N \mathcal{A}^\dagger + \mathcal{B}^\dagger *_N (\mathcal{B}^\dagger)^* *_N \mathcal{E}^* *_N P_{\mathcal{A}}^\perp - R_{\mathcal{B}}^\perp *_N \mathcal{E}^* *_N (\mathcal{A}^\dagger)^* *_N \mathcal{A}^\dagger.$$

Proof After some verifications, one can obtain

$$\begin{aligned} \mathcal{B}^\dagger - \mathcal{A}^\dagger &= -\mathcal{B}^\dagger *_N \mathcal{E} *_N \mathcal{A}^\dagger + (\mathcal{B}^\dagger - \mathcal{A}^\dagger) + \mathcal{B}^\dagger *_N (\mathcal{B} - \mathcal{A}) *_N \mathcal{A}^\dagger \\ &= -\mathcal{B}^\dagger *_N \mathcal{E} *_N \mathcal{A}^\dagger + \mathcal{B}^\dagger *_N (\mathcal{I} - \mathcal{A} *_N \mathcal{A}^\dagger) - (\mathcal{I} - \mathcal{B}^\dagger *_N \mathcal{B}) *_N \mathcal{A}^\dagger. \end{aligned}$$

According to the properties (3^T) and (1^T) from Definition 1.1, it follows that

$$\mathcal{A}^* *_N (\mathcal{I} - \mathcal{A} *_N \mathcal{A}^\dagger) = \mathcal{A}^* - \mathcal{A}^* *_N (\mathcal{A} *_N \mathcal{A}^\dagger)^* = \mathcal{A}^* - (\mathcal{A} *_N \mathcal{A}^\dagger *_N \mathcal{A})^* = \mathcal{O},$$

where $\mathcal{O} \in \mathbb{C}^{\mathbf{K}(N) \times \mathbf{I}(N)}$ is an appropriate zero tensor. Consequently

$$\begin{aligned} \mathcal{B}^\dagger *_N (\mathcal{I} - \mathcal{A} *_N \mathcal{A}^\dagger) &= \mathcal{B}^\dagger *_N \mathcal{B} *_N \mathcal{B}^\dagger *_N (\mathcal{I} - \mathcal{A} *_N \mathcal{A}^\dagger) \\ &= \mathcal{B}^\dagger *_N (\mathcal{B} *_N \mathcal{B}^\dagger)^* *_N (\mathcal{I} - \mathcal{A} *_N \mathcal{A}^\dagger) \\ &= \mathcal{B}^\dagger *_N (\mathcal{B}^\dagger)^* *_N (\mathcal{A} + \mathcal{E})^* *_N (\mathcal{I} - \mathcal{A} *_N \mathcal{A}^\dagger) \\ &= \mathcal{B}^\dagger *_N (\mathcal{B}^\dagger)^* *_N \mathcal{E}^* *_N (\mathcal{I} - \mathcal{A} *_N \mathcal{A}^\dagger). \end{aligned}$$

Analogously, we arrive to

$$(\mathcal{I} - \mathcal{B}^\dagger *_N \mathcal{B}) *_N \mathcal{B}^* = \mathcal{O},$$

which further implies

$$(\mathcal{I} - \mathcal{B}^\dagger *_N \mathcal{B}) *_N \mathcal{A}^\dagger = -(\mathcal{I} - \mathcal{B}^\dagger *_N \mathcal{B}) *_N \mathcal{E}^* *_N (\mathcal{A}^\dagger)^* *_N \mathcal{A}^*.$$

The conclusion can be obtained. □

Lemma 4.2 *If $\mathcal{O} \neq \mathcal{P} \in \mathbb{C}^{\mathbf{K}(N) \times \mathbf{K}(N)}$, and $\mathcal{P}^2 = \mathcal{P} = \mathcal{P}^*$, then*

$$\|\mathcal{P}\|_2 = 1.$$

Proof Since

$$\|\mathcal{P}\|_2^2 = \|\mathcal{P}^* *_N \mathcal{P}\|_2 = \|\mathcal{P}^2\|_2 = \|\mathcal{P}\|_2,$$

it follows that

$$\|\mathcal{P}\|_2 (\|\mathcal{P}\|_2 - 1) = 0.$$

Therefore, $\|\mathcal{P}\|_2 = 1$ in the case $\mathcal{P} \neq \mathcal{O}$. □

Theorem 4.2 Let $\mathcal{A}, \mathcal{E} \in \mathbb{C}^{\mathbf{I}(N) \times \mathbf{K}(N)}, \mathcal{B} = \mathcal{A} + \mathcal{E}$. If the Condition (4.1) is satisfied, then

$$\frac{\|\mathcal{B}^\dagger - \mathcal{A}^\dagger\|_2}{\|\mathcal{A}^\dagger\|_2} \leq \left(1 + \frac{1}{1 - \Delta} + \frac{1}{(1 - \Delta)^2}\right) \Delta. \tag{4.3}$$

Proof Since $(\mathcal{I} - \mathcal{A} *_N \mathcal{A}^\dagger)^2 = (\mathcal{I} - \mathcal{A} *_N \mathcal{A}^\dagger) = (\mathcal{I} - \mathcal{A} *_N \mathcal{A}^\dagger)^*$ and $(\mathcal{I} - \mathcal{B}^\dagger *_N \mathcal{B})^2 = (\mathcal{I} - \mathcal{B}^\dagger *_N \mathcal{B}) = (\mathcal{I} - \mathcal{B}^\dagger *_N \mathcal{B})^*$, by Lemma 4.2

$$\|\mathcal{I} - \mathcal{A} *_N \mathcal{A}^\dagger\|_2 = 1, \quad \|\mathcal{I} - \mathcal{B}^\dagger *_N \mathcal{B}\|_2 = 1,$$

and from Theorem 4.1

$$\|\mathcal{B}^\dagger - \mathcal{A}^\dagger\|_2 \leq (\|\mathcal{A}^\dagger\|_2 \|\mathcal{B}^\dagger\|_2 + \|\mathcal{B}^\dagger\|_2^2 + \|\mathcal{A}^\dagger\|_2^2) \|\mathcal{E}\|_2.$$

An application of Lemma 4.1 initiates

$$\|\mathcal{B}^\dagger - \mathcal{A}^\dagger\|_2 \leq \left(\frac{\|\mathcal{A}^\dagger\|_2^2}{1 - \Delta} + \frac{\|\mathcal{A}^\dagger\|_2^2}{(1 - \Delta)^2} + \|\mathcal{A}^\dagger\|_2^2\right) \|\mathcal{E}\|_2.$$

Furthermore, the inequality (4.3) can be verified taking into account $\Delta = \|\mathcal{A}^\dagger\|_2 \|\mathcal{E}\|_2$. \square

Theorem 4.3 If $\mathcal{A}, \mathcal{E} \in \mathbb{C}^{\mathbf{I}(N) \times \mathbf{K}(N)}, \mathcal{B} = \mathcal{A} + \mathcal{E}$, and $\text{rshrank}(\mathcal{A}) = \text{rshrank}(\mathcal{B})$, then

$$\|\mathcal{B} *_N \mathcal{B}^\dagger *_N (\mathcal{I} - \mathcal{A} *_N \mathcal{A}^\dagger)\|_2 = \|\mathcal{A} *_N \mathcal{A}^\dagger *_N (\mathcal{I} - \mathcal{B} *_N \mathcal{B}^\dagger)\|_2, \tag{4.4}$$

where \mathcal{I} is the identity $\mathbf{I}(N) \times \mathbf{I}(N)$ tensor.

Proof According to Lemma 3.3, it follows that

$$\mathcal{A} *_N \mathcal{A}^\dagger = \mathcal{U}_A *_N \text{rsh}^{-1} \left(\begin{bmatrix} I_{\mathcal{J}_R \times \mathcal{J}_R} & O_{\mathcal{J}_R \times (\mathcal{J} - \mathcal{J}_R)} \\ O_{(\mathcal{J} - \mathcal{J}_R) \times \mathcal{J}_R} & O_{(\mathcal{J} - \mathcal{J}_R) \times (\mathcal{J} - \mathcal{J}_R)} \end{bmatrix} \right) *_N \mathcal{U}_A^*.$$

Furthermore, using Lemma 2.2 and

$$\mathcal{I} = \text{rsh}^{-1} (I_{\mathcal{J} \times \mathcal{J}}) \in \mathbb{C}^{\mathbf{I}(N) \times \mathbf{I}(N)},$$

it follows that

$$\begin{aligned} \mathcal{I} - \mathcal{A} *_N \mathcal{A}^\dagger &= \mathcal{U}_A *_N \left(\mathcal{I} - \text{rsh}^{-1} \left(\begin{bmatrix} I_{\mathcal{J}_R \times \mathcal{J}_R} & O_{\mathcal{J}_R \times (\mathcal{J} - \mathcal{J}_R)} \\ O_{(\mathcal{J} - \mathcal{J}_R) \times \mathcal{J}_R} & O_{(\mathcal{J} - \mathcal{J}_R) \times (\mathcal{J} - \mathcal{J}_R)} \end{bmatrix} \right) \right) *_N \mathcal{U}_A^* \\ &= \mathcal{U}_A *_N \text{rsh}^{-1} \left(\begin{bmatrix} O_{\mathcal{J}_R \times \mathcal{J}_R} & O_{\mathcal{J}_R \times (\mathcal{J} - \mathcal{J}_R)} \\ O_{(\mathcal{J} - \mathcal{J}_R) \times \mathcal{J}_R} & I_{(\mathcal{J} - \mathcal{J}_R) \times (\mathcal{J} - \mathcal{J}_R)} \end{bmatrix} \right) *_N \mathcal{U}_A^*. \end{aligned}$$

Similarly, in view of $\text{rshrank}(\mathcal{B}) = \text{rshrank}(\mathcal{A})$, it follows $\text{rank}(\mathcal{B}) = \text{rank}(\mathcal{A})$, where $\text{rsh}(\mathcal{A}) = \mathcal{A}$ and $\text{rsh}(\mathcal{B}) = \mathcal{B}$. The SVD of \mathcal{B} is given by $[\mathcal{U}_B, \mathcal{D}_B, \mathcal{V}_B] = \text{SVD}(\mathcal{B})$. Now, consider the reshaping operations

$$\begin{aligned} \text{rsh}(\mathcal{U}_B) &= \mathcal{U}_B \in \mathbb{C}^{\mathbf{I}(N) \times \mathbf{I}(N)}, \quad \text{rsh}(\mathcal{V}_B^*) = \mathcal{V}_B^* \in \mathbb{C}^{\mathbf{K}(N) \times \mathbf{K}(N)}, \\ \text{rsh}(\mathcal{D}_B) &= \text{rsh}^{-1} \left(\begin{bmatrix} \Sigma_B & O_{\mathcal{J}_R \times (\mathcal{R} - \mathcal{J}_R)} \\ O_{(\mathcal{J} - \mathcal{J}_R) \times \mathcal{J}_R} & O_{(\mathcal{J} - \mathcal{J}_R) \times (\mathcal{R} - \mathcal{J}_R)} \end{bmatrix} \right) = \mathcal{D}_B \in \mathbb{C}^{\mathbf{I}(N) \times \mathbf{I}(N)}, \end{aligned}$$

where

$$\Sigma_B \in \mathbb{C}^{\mathcal{J}_R \times \mathcal{J}_R}, \quad \mathcal{J}_R = \text{rank}(\mathcal{A})$$

is diagonal with singular values of B on the main diagonal. This causes

$$\begin{aligned} \mathcal{B} &= U_B *_{\mathcal{N}} \mathcal{D}_B *_{\mathcal{N}} \mathcal{V}_B^*, \quad \mathcal{B} *_{\mathcal{N}} \mathcal{B}^\dagger \\ &= U_B *_{\mathcal{N}} \text{rsh}^{-1} \left(\begin{bmatrix} I_{\mathcal{J}_R \times \mathcal{J}_R} & O_{\mathcal{J}_R \times (\mathcal{J} - \mathcal{J}_R)} \\ O_{(\mathcal{J} - \mathcal{J}_R) \times \mathcal{J}_R} & O_{(\mathcal{J} - \mathcal{J}_R) \times (\mathcal{J} - \mathcal{J}_R)} \end{bmatrix} \right) *_{\mathcal{N}} U_B^*, \end{aligned}$$

and further

$$\begin{aligned} \mathcal{I} - \mathcal{B} *_{\mathcal{N}} \mathcal{B}^\dagger &= U_B *_{\mathcal{N}} \left(\mathcal{I} - \text{rsh}^{-1} \left(\begin{bmatrix} I_{\mathcal{J}_R \times \mathcal{J}_R} & O_{\mathcal{J}_R \times (\mathcal{J} - \mathcal{J}_R)} \\ O_{(\mathcal{J} - \mathcal{J}_R) \times \mathcal{J}_R} & O_{(\mathcal{J} - \mathcal{J}_R) \times (\mathcal{J} - \mathcal{J}_R)} \end{bmatrix} \right) \right) *_{\mathcal{N}} U_B^* \\ &= U_B *_{\mathcal{N}} \text{rsh}^{-1} \left(\begin{bmatrix} O_{\mathcal{J}_R \times \mathcal{J}_R} & O_{\mathcal{J}_R \times (\mathcal{J} - \mathcal{J}_R)} \\ O_{(\mathcal{J} - \mathcal{J}_R) \times \mathcal{J}_R} & I_{(\mathcal{J} - \mathcal{J}_R) \times (\mathcal{J} - \mathcal{J}_R)} \end{bmatrix} \right) *_{\mathcal{N}} U_B^*. \end{aligned}$$

Now, observe the tensor products $U_A^* *_{\mathcal{N}} U_B$ and $U_B^* *_{\mathcal{N}} U_A$. They are also unitary and equal to

$$\begin{aligned} U_A^* *_{\mathcal{N}} U_B &= \text{rsh} (U_A^* U_B) = \text{rsh} \left(\begin{bmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{bmatrix} \right) \\ U_B^* *_{\mathcal{N}} U_A &= \text{rsh} (U_B^* U_A) = \text{rsh} \left(\begin{bmatrix} W_{11}^* & W_{21}^* \\ W_{12}^* & W_{22}^* \end{bmatrix} \right), \end{aligned}$$

where

$$W_{11} \in \mathbb{C}^{\mathcal{J}_R \times \mathcal{J}_R}, W_{12} \in \mathbb{C}^{\mathcal{J}_R \times (\mathcal{J} - \mathcal{J}_R)}, W_{21} \in \mathbb{C}^{(\mathcal{J} - \mathcal{J}_R) \times \mathcal{J}_R}, W_{22} \in \mathbb{C}^{(\mathcal{J} - \mathcal{J}_R) \times (\mathcal{J} - \mathcal{J}_R)},$$

In addition, it can be verified that $\|\cdot\|_2$ is a unitary invariant tensor norm (Govaerts and Pryce 1989), which implies in conjunction with Lemma 2.2:

$$\begin{aligned} &\|\mathcal{B} *_{\mathcal{N}} \mathcal{B}^\dagger *_{\mathcal{N}} (\mathcal{I} - \mathcal{A} *_{\mathcal{N}} \mathcal{A}^\dagger)\|_2 \\ &= \left\| \text{rsh}^{-1} \left(\begin{bmatrix} I_{\mathcal{J}_R \times \mathcal{J}_R} & O_{\mathcal{J}_R \times (\mathcal{J} - \mathcal{J}_R)} \\ O_{(\mathcal{J} - \mathcal{J}_R) \times \mathcal{J}_R} & O_{(\mathcal{J} - \mathcal{J}_R) \times (\mathcal{J} - \mathcal{J}_R)} \end{bmatrix} \begin{bmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{bmatrix} \begin{bmatrix} O_{\mathcal{J}_R \times \mathcal{J}_R} & O_{\mathcal{J}_R \times (\mathcal{J} - \mathcal{J}_R)} \\ O_{(\mathcal{J} - \mathcal{J}_R) \times \mathcal{J}_R} & I_{(\mathcal{J} - \mathcal{J}_R) \times (\mathcal{J} - \mathcal{J}_R)} \end{bmatrix} \right) \right\|_2. \end{aligned}$$

An application of Lemma 2.4 further implies

$$\begin{aligned} &\|\mathcal{B} *_{\mathcal{N}} \mathcal{B}^\dagger *_{\mathcal{N}} (\mathcal{I} - \mathcal{A} *_{\mathcal{N}} \mathcal{A}^\dagger)\|_2 \\ &= \left\| \begin{bmatrix} I_{\mathcal{J}_R \times \mathcal{J}_R} & O_{\mathcal{J}_R \times (\mathcal{J} - \mathcal{J}_R)} \\ O_{(\mathcal{J} - \mathcal{J}_R) \times \mathcal{J}_R} & O_{(\mathcal{J} - \mathcal{J}_R) \times (\mathcal{J} - \mathcal{J}_R)} \end{bmatrix} \begin{bmatrix} W_{11} & W_{12} \\ W_{21} & W_{22} \end{bmatrix} \begin{bmatrix} O_{\mathcal{J}_R \times \mathcal{J}_R} & O_{\mathcal{J}_R \times (\mathcal{J} - \mathcal{J}_R)} \\ O_{(\mathcal{J} - \mathcal{J}_R) \times \mathcal{J}_R} & I_{(\mathcal{J} - \mathcal{J}_R) \times (\mathcal{J} - \mathcal{J}_R)} \end{bmatrix} \right\|_2 \\ &= \left\| \begin{bmatrix} O & W_{12} \\ O & O \end{bmatrix} \right\|_2. \end{aligned}$$

Finally, using the result from Govaerts and Pryce (1989), it follows that

$$\|\mathcal{B} *_{\mathcal{N}} \mathcal{B}^\dagger *_{\mathcal{N}} (\mathcal{I} - \mathcal{A} *_{\mathcal{N}} \mathcal{A}^\dagger)\|_2 = \|W_{12}\|_2.$$

On the other hand, in dual case, it follows that

$$\begin{aligned} &\|\mathcal{A} *_{\mathcal{N}} \mathcal{A}^\dagger *_{\mathcal{N}} (\mathcal{I} - \mathcal{B} *_{\mathcal{N}} \mathcal{B}^\dagger)\|_2 \\ &= \left\| \begin{bmatrix} I_{\mathcal{J}_R \times \mathcal{J}_R} & O_{\mathcal{J}_R \times (\mathcal{J} - \mathcal{J}_R)} \\ O_{(\mathcal{J} - \mathcal{J}_R) \times \mathcal{J}_R} & O_{(\mathcal{J} - \mathcal{J}_R) \times (\mathcal{J} - \mathcal{J}_R)} \end{bmatrix} \begin{bmatrix} W_{11}^* & W_{21}^* \\ W_{12}^* & W_{22}^* \end{bmatrix} \begin{bmatrix} O_{\mathcal{J}_R \times \mathcal{J}_R} & O_{\mathcal{J}_R \times (\mathcal{J} - \mathcal{J}_R)} \\ O_{(\mathcal{J} - \mathcal{J}_R) \times \mathcal{J}_R} & I_{(\mathcal{J} - \mathcal{J}_R) \times (\mathcal{J} - \mathcal{J}_R)} \end{bmatrix} \right\|_2 \\ &= \left\| \begin{bmatrix} O & W_{21}^* \\ O & O \end{bmatrix} \right\|_2 \\ &= \|W_{21}^*\|_2. \end{aligned}$$

The proof can be completed by verifying $\mu_1(W_{12}) = \mu_1(W_{21}^*)$. Indeed, according to Proposition 3.1, it follows that $\|W_{12}\|_2 = \|W_{21}^*\|_2$. The proof can be completed using the result $\|W_{21}\|_2 = \|W_{21}^*\|_2$ from Govaerts and Pryce (1989). \square

Corollary 4.1 *If $\mathcal{A}, \mathcal{E} \in \mathbb{C}^{\mathbf{I}(N) \times \mathbf{K}(N)}$, $\mathcal{B} = \mathcal{A} + \mathcal{E}$, and $\text{rshrank}(\mathcal{A}) = \text{rshrank}(\mathcal{B})$, let*

$$\mathcal{G} = \mathcal{B}^\dagger *_N (\mathcal{I} - \mathcal{A} *_N \mathcal{A}^\dagger),$$

then

$$\|\mathcal{G}\|_2 \leq \|\mathcal{E}\|_2 \|\mathcal{A}^\dagger\|_2 \|\mathcal{B}^\dagger\|_2. \tag{4.5}$$

Proof Clearly, $\mathcal{G} = \mathcal{B}^\dagger *_N \mathcal{B} *_N \mathcal{B}^\dagger *_N (\mathcal{I} - \mathcal{A} *_N \mathcal{A}^\dagger)$. Then

$$\|\mathcal{G}\|_2 \leq \|\mathcal{B}^\dagger\|_2 \|\mathcal{B} *_N \mathcal{B}^\dagger *_N (\mathcal{I} - \mathcal{A} *_N \mathcal{A}^\dagger)\|_2,$$

since

$$\begin{aligned} \mathcal{B}^* *_N (\mathcal{I} - \mathcal{B} *_N \mathcal{B}^\dagger) &= \mathcal{O}, \\ (\mathcal{I} - \mathcal{B} *_N \mathcal{B}^\dagger)^2 &= (\mathcal{I} - \mathcal{B} *_N \mathcal{B}^\dagger) = (\mathcal{I} - \mathcal{B} *_N \mathcal{B}^\dagger)^*. \end{aligned}$$

Therefore, $\|\mathcal{I} - \mathcal{B} *_N \mathcal{B}^\dagger\|_2 = 1$, applying Theorem 4.3, one can obtain

$$\begin{aligned} \|\mathcal{B} *_N \mathcal{B}^\dagger *_N (\mathcal{I} - \mathcal{A} *_N \mathcal{A}^\dagger)\|_2 &= \|\mathcal{A} *_N \mathcal{A}^\dagger *_N (\mathcal{I} - \mathcal{B} *_N \mathcal{B}^\dagger)\|_2 \\ &= \|(\mathcal{A} *_N \mathcal{A}^\dagger)^* *_N (\mathcal{I} - \mathcal{B} *_N \mathcal{B}^\dagger)\|_2 \\ &= \|(\mathcal{A}^\dagger)^* *_N \mathcal{E}^* *_N (\mathcal{I} - \mathcal{B} *_N \mathcal{B}^\dagger)\|_2 \\ &\leq \|(\mathcal{A}^\dagger)^* *_N \mathcal{E}^*\|_2 = \|\mathcal{E} *_N \mathcal{A}^\dagger\|_2 \\ &\leq \|\mathcal{E}\|_2 \|\mathcal{A}^\dagger\|_2. \end{aligned}$$

Thus, the statement can be obtained. □

Theorem 4.4 *Let $\mathcal{A}, \mathcal{E} \in \mathbb{C}^{\mathbf{I}(N) \times \mathbf{K}(N)}$, $\mathcal{B} = \mathcal{A} + \mathcal{E}$ and $\mathfrak{J}, \mathfrak{K}$ are defined as in (2.3). If the Condition (4.1) is satisfied, then*

$$\frac{\|\mathcal{B}^\dagger - \mathcal{A}^\dagger\|_2}{\|\mathcal{A}^\dagger\|_2} \leq k \frac{\Delta}{1 - \Delta}, \tag{4.6}$$

where $\Delta = \|\mathcal{A}^\dagger\|_2 \|\mathcal{E}\|_2$ and the parameter k is defined as follows:

- (1) if $\text{rshrank}(\mathcal{A}) < \min(\mathfrak{J}, \mathfrak{K})$, then $k = \frac{1+\sqrt{5}}{2}$;
- (2) if $\text{rshrank}(\mathcal{A}) = \min(\mathfrak{J}, \mathfrak{K})$, then $k = \sqrt{2}$;
- (3) if $\text{rshrank}(\mathcal{A}) = \mathfrak{J} = \mathfrak{K}$, then $k = 1$.

Proof Let

$$\mathcal{F} = -\mathcal{B}^\dagger *_N \mathcal{E} *_N \mathcal{A}^\dagger, \mathcal{G} = \mathcal{B}^\dagger *_N (\mathcal{I} - \mathcal{A} *_N \mathcal{A}^\dagger), \mathcal{H} = -(\mathcal{I} - \mathcal{B}^\dagger *_N \mathcal{B}) *_N \mathcal{A}^\dagger.$$

By the Lemma 4.1, we can get

$$\begin{aligned} \|\mathcal{F}\|_2 &\leq \frac{\Delta}{1 - \Delta} \|\mathcal{A}^\dagger\|_2, \\ \|\mathcal{G}\|_2 &\leq \frac{\Delta}{1 - \Delta} \|\mathcal{A}^\dagger\|_2, \\ \|\mathcal{H}\|_2 &\leq \Delta \|\mathcal{A}^\dagger\|_2, \end{aligned}$$

where $\Delta = \|\mathcal{A}^\dagger\|_2 \|\mathcal{E}\|_2 = \|(\mathcal{A}^\dagger)^* *_N \mathcal{E}^*\|_2$. Let

$$\alpha = \frac{\Delta}{1 - \Delta},$$

then

$$\|\mathcal{F}\|_2, \|\mathcal{G}\|_2, \|\mathcal{H}\|_2 \leq \alpha \|\mathcal{A}^\dagger\|_2.$$

(1) Let $\mathcal{X} \in \mathbb{C}^{I_1 \times \dots \times I_N \times K_1 \times \dots \times K_N}$, $\|\mathcal{X}\|_2 = 1$, and $\mathcal{X} = \mathcal{X}_1 + \mathcal{X}_2$, where

$$\mathcal{X}_1 = \mathcal{A} *_N \mathcal{A}^\dagger *_N \mathcal{X}, \quad \mathcal{X}_2 = (\mathcal{I} - \mathcal{A} *_N \mathcal{A}^\dagger) *_N \mathcal{X}.$$

Clearly, \mathcal{X}_1 and \mathcal{X}_2 are orthogonal; hence,

$$1 = \|\mathcal{X}\|_2^2 = \|\mathcal{X}_1\|_2^2 + \|\mathcal{X}_2\|_2^2.$$

Therefore, there must be an angle φ which makes

$$\cos \varphi = \|\mathcal{X}_1\|_2, \quad \sin \varphi = \|\mathcal{X}_2\|_2.$$

Then

$$\begin{aligned} & (\mathcal{B}^\dagger - \mathcal{A}^\dagger) *_N \mathcal{X} \\ &= -\mathcal{B}^\dagger *_N \mathcal{E} *_N \mathcal{A}^\dagger *_N \mathcal{A} *_N \mathcal{A}^\dagger *_N \mathcal{X} + \mathcal{B}^\dagger *_N (\mathcal{I} - \mathcal{A} *_N \mathcal{A}^\dagger) *_N (\mathcal{I} - \mathcal{A} *_N \mathcal{A}^\dagger) *_N \mathcal{X} \\ &\quad - (\mathcal{I} - \mathcal{B}^\dagger *_N \mathcal{B}) *_N \mathcal{A}^\dagger *_N \mathcal{A} *_N \mathcal{A}^\dagger *_N \mathcal{X} = \\ & \mathcal{F} *_N \mathcal{X}_1 + \mathcal{G} *_N \mathcal{X}_2 + \mathcal{H} *_N \mathcal{X}_1 \equiv \mathcal{Y}_1 + \mathcal{Y}_2 + \mathcal{Y}_3, \end{aligned}$$

where $\mathcal{Y}_1 = \mathcal{F} *_N \mathcal{X}_1$, $\mathcal{Y}_2 = \mathcal{G} *_N \mathcal{X}_2$, $\mathcal{Y}_3 = \mathcal{H} *_N \mathcal{X}_1$.

Since

$$(\mathcal{I} - \mathcal{B}^\dagger *_N \mathcal{B}) *_N \mathcal{B}^* = \mathcal{O}.$$

It is easy to verify that \mathcal{Y}_3 is orthogonal to \mathcal{Y}_1 and \mathcal{Y}_2 ; therefore

$$\begin{aligned} \|\mathcal{B}^\dagger - \mathcal{A}^\dagger\|_2 \|\mathcal{X}\|_2 &\leq \|\mathcal{Y}_1 + \mathcal{Y}_2\|_2 + \|\mathcal{Y}_3\|_2 \\ &\leq \alpha^2 \|\mathcal{A}^\dagger\|_2^2 [(\|\mathcal{X}_1\|_2 + \|\mathcal{X}_2\|_2)^2 + \|\mathcal{X}_1\|_2^2] \\ &= \alpha^2 \|\mathcal{A}^\dagger\|_2^2 [(\cos \varphi + \sin \varphi)^2 + \cos^2 \varphi] \\ &= \alpha^2 \|\mathcal{A}^\dagger\|_2^2 (3 + 2 \sin 2\varphi + \cos 2\varphi) / 2 \\ &\leq \left(\frac{3 + \sqrt{5}}{2} \right) \alpha^2 \|\mathcal{A}^\dagger\|_2^2. \end{aligned}$$

Therefore

$$\|\mathcal{B}^\dagger - \mathcal{A}^\dagger\|_2 = \max_{\|\mathcal{X}\|_2=1} \|(\mathcal{B}^\dagger - \mathcal{A}^\dagger) *_N \mathcal{X}\|_2 \leq \frac{1 + \sqrt{5}}{2} \alpha \|\mathcal{A}^\dagger\|_2.$$

(2) If $\text{rshrank}(\mathcal{A}) = \text{rshrank}(\mathcal{B}) = \mathfrak{R} < \mathfrak{J}$, owing to

$$\mathcal{B}^\dagger = (\mathcal{B}^* *_N \mathcal{B})^{-1} \mathcal{B}^*,$$

then $(\mathcal{I} - \mathcal{B}^\dagger *_N \mathcal{B}) = \mathcal{O}$; therefore, $\mathcal{H} = \mathcal{O}$, $\mathcal{Y}_3 = \mathcal{O}$. If $\text{rank}(\mathcal{A}) = \text{rank}(\mathcal{B}) = \mathfrak{J} < \mathfrak{R}$, owing to

$$\mathcal{A}^\dagger = \mathcal{A}^* *_N (\mathcal{A} *_N \mathcal{A}^*)^{-1},$$

then $(\mathcal{I} - \mathcal{A} *_N \mathcal{A}^\dagger) = \mathcal{O}$, so $\mathcal{G} = \mathcal{O}$, $\mathcal{Y}_2 = \mathcal{O}$. When one of \mathcal{Y}_2 or \mathcal{Y}_3 is the zero tensor

$$\|(\mathcal{B}^\dagger - \mathcal{A}^\dagger) *_N \mathcal{X}\|_2^2 \leq 2\alpha^2 \|\mathcal{A}^\dagger\|_2^2.$$

Hence

$$\|\mathcal{B}^\dagger - \mathcal{A}^\dagger\|_2 = \max_{\|\mathcal{X}\|_2=1} \|(\mathcal{B}^\dagger - \mathcal{A}^\dagger) *_{\mathcal{N}} \mathcal{X}\|_2 \leq \sqrt{2}\alpha \|\mathcal{A}^\dagger\|_2.$$

(3) In the case $\text{rshrank}(\mathcal{A}) = \text{rshrank}(\mathcal{B}) = \mathcal{J} = \mathfrak{K}$, according to (2), we know $\mathcal{G} = \mathcal{H} = \mathcal{O}$, so the conclusion is established. □

Next, we introduce the condition number of the Moore–Penrose inverse for tensor \mathcal{A} :

$$\mathbb{K}_2(\mathcal{A}) = \|\mathcal{A}\|_2 \|\mathcal{A}^\dagger\|_2.$$

Theorem 4.5 *If the Condition (4.1) is satisfied, then*

$$\frac{\|\mathcal{B}^\dagger - \mathcal{A}^\dagger\|_2}{\|\mathcal{A}^\dagger\|_2} \leq k \mathbb{K}_2(\mathcal{A}) \frac{\frac{\|\mathcal{E}\|_2}{\|\mathcal{A}\|_2}}{1 - \mathbb{K}_2(\mathcal{A}) \frac{\|\mathcal{E}\|_2}{\|\mathcal{A}\|_2}}.$$

Proof The statement can be verified using Theorem 4.4 and the definition of $\mathbb{K}_2(\mathcal{A})$. □

Theorem 4.5 shows that the perturbation \mathcal{E} of \mathcal{A} has little influence on \mathcal{A}^\dagger when the condition number $\mathbb{K}_2(\mathcal{A})$ is small, and when the condition number $\mathbb{K}_2(\mathcal{A})$ is large, the influence of \mathcal{E} on the disturbance to \mathcal{A}^\dagger may be larger.

5 Examples

Example 5.1 This example is aimed to the verification of the inequality (4.2). Let the tensor $\mathcal{A} = 10^3 * \text{rand}(2, 2, 2, 2)$ be defined by

$$\begin{aligned} \mathcal{A}(:, :, 1, 1) &= \begin{bmatrix} 950.9152 & 400.0797 \\ 722.3485 & 831.8713 \end{bmatrix}, \mathcal{A}(:, :, 2, 1) = \begin{bmatrix} 134.3383 & 84.2471 \\ 60.4668 & 163.8983 \end{bmatrix}, \\ \mathcal{A}(:, :, 1, 2) &= \begin{bmatrix} 324.2199 & 11.6810 \\ 301.7268 & 539.9051 \end{bmatrix}, \mathcal{A}(:, :, 2, 2) = \begin{bmatrix} 95.3727 & 631.1412 \\ 146.5149 & 859.3204 \end{bmatrix}, \end{aligned}$$

and let $\mathcal{E} = 10^{-1} * \text{rand}(2, 2, 2, 2)$ be defined by

$$\begin{aligned} \mathcal{E}(:, :, 1, 1) &= \begin{bmatrix} 0.0974 & 0.0997 \\ 0.0571 & 0.0554 \end{bmatrix}, \mathcal{E}(:, :, 2, 1) = \begin{bmatrix} 0.0515 & 0.0430 \\ 0.0331 & 0.0492 \end{bmatrix}, \\ \mathcal{E}(:, :, 1, 2) &= \begin{bmatrix} 0.0071 & 0.0065 \\ 0.0888 & 0.0436 \end{bmatrix}, \mathcal{E}(:, :, 2, 2) = \begin{bmatrix} 0.0827 & 0.0613 \\ 0.0395 & 0.0819 \end{bmatrix}. \end{aligned}$$

Then, $\mathcal{B} = \mathcal{A} + \mathcal{E}$ is defined by

$$\begin{aligned} \mathcal{B}(:, :, 1, 1) &= \begin{bmatrix} 951.0126 & 400.1794 \\ 722.4056 & 831.9267 \end{bmatrix}, \mathcal{B}(:, :, 2, 1) = \begin{bmatrix} 134.3899 & 84.2901 \\ 60.4998 & 163.9475 \end{bmatrix}, \\ \mathcal{B}(:, :, 1, 2) &= \begin{bmatrix} 324.2270 & 11.6875 \\ 301.8156 & 539.9487 \end{bmatrix}, \mathcal{B}(:, :, 2, 2) = \begin{bmatrix} 95.4554 & 631.2026 \\ 146.5543 & 859.4023 \end{bmatrix}. \end{aligned}$$

It holds, $\text{rshrank}(\mathcal{A}) = \text{rshrank}(\mathcal{B}) = 4$. In addition, an application of Huang et al. (2018, Algorithm 1) gives the Moore–Penrose inverse of \mathcal{A} :

$$\begin{aligned} \mathcal{A}^\dagger(:, :, 1, 1) &= \begin{bmatrix} -0.000384784660649 & -0.001099286197235 \\ 0.013955577455799 & -0.001598581983579 \end{bmatrix}, \\ \mathcal{A}^\dagger(:, :, 2, 1) &= \begin{bmatrix} 0.002737547517780 & 0.000508206301570 \\ -0.021392949244820 & 0.001110875409922 \end{bmatrix}, \\ \mathcal{A}^\dagger(:, :, 1, 2) &= \begin{bmatrix} 0.001227173049100 & -0.003076391592687 \\ -0.002071095392062 & 0.001139922248389 \end{bmatrix}, \\ \mathcal{A}^\dagger(:, :, 2, 2) &= \begin{bmatrix} -0.001325364668035 & 0.002294859505387 \\ 0.003619787791101 & 0.000314492021375 \end{bmatrix} \end{aligned}$$

and the following Moore–Penrose inverse of \mathcal{B} :

$$\begin{aligned} \mathcal{B}^\dagger(:, :, 1, 1) &= \begin{bmatrix} -0.000385255308227 & -0.001098543826394 \\ 0.013953167583606 & -0.001598698848909 \end{bmatrix}, \\ \mathcal{B}^\dagger(:, :, 2, 1) &= \begin{bmatrix} 0.002738190383336 & 0.000507663916870 \\ -0.021390844989888 & 0.001111108748350 \end{bmatrix}, \\ \mathcal{B}^\dagger(:, :, 1, 2) &= \begin{bmatrix} 0.001227624787307 & -0.003075755111089 \\ -0.002076686203945 & 0.001140238650466 \end{bmatrix}, \\ \mathcal{B}^\dagger(:, :, 2, 2) &= \begin{bmatrix} -0.001325803821792 & 0.002294485476614 \\ 0.003623245688721 & 0.000314224257971 \end{bmatrix}. \end{aligned}$$

In view of (2.5), it is easy to check that the tensor norms are equal to $\|\mathcal{A}^\dagger\|_2 = 0.026095036211067$, $\|\mathcal{E}\|_2 = 0.235145716909881$ and $\Delta = \|\mathcal{A}^\dagger\|_2 \|\mathcal{E}\|_2 = 0.006136135997641 < 1$. Therefore, Condition (4.1) is satisfied. Then, $\|\mathcal{B}^\dagger\|_2 = 0.026093083833995$ and $\frac{\|\mathcal{A}^\dagger\|_2}{1-\Delta} = 0.026256147502919$. Hence, the inequality (4.2) in Lemma 4.1 is verified.

Example 5.2 The tensors in Example 5.1 are invertible. Example 5.2 is aimed to the verification of the inequality (4.2) in singular tensor case. To this end, let $\mathcal{A}, \mathcal{E} \in \mathbb{R}^{(2 \times 2) \times (2 \times 2)}$ with

$$\begin{aligned} \mathcal{A}(:, :, 1, 1) &= 10^2 \cdot \begin{bmatrix} 0.985940927109977 & 1.682512984915278 \\ 1.420272484319284 & 1.962489222569553 \end{bmatrix}, \quad \mathcal{A}(:, :, 2, 1) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \\ \mathcal{A}(:, :, 1, 2) &= 10^2 \cdot \begin{bmatrix} 8.929224052859770 & 5.557379427193866 \\ 7.0322322245562910 & 1.844336677576532 \end{bmatrix}, \quad \mathcal{A}(:, :, 2, 2) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \end{aligned}$$

and

$$\begin{aligned} \mathcal{E}(:, :, 1, 1) &= \begin{bmatrix} 0.055778896675488 & 0.016620356290215 \\ 0.031342898993659 & 0.062249725927990 \end{bmatrix}, \quad \mathcal{E}(:, :, 2, 1) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \\ \mathcal{E}(:, :, 1, 2) &= \begin{bmatrix} 0.007399476957694 & 0.040238833269616 \\ 0.068409606696201 & 0.098283520139395 \end{bmatrix}, \quad \mathcal{E}(:, :, 2, 2) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}. \end{aligned}$$

Then, the tensor $\mathcal{B} = \mathcal{A} + \mathcal{E}$ is defined by

$$\begin{aligned} \mathcal{B}(:, :, 1, 1) &= 10^2 \cdot \begin{bmatrix} 0.986498716076732 & 1.682679188478180 \\ 1.420585913309220 & 1.963111719828833 \end{bmatrix}, \quad \mathcal{B}(:, :, 2, 1) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \\ \mathcal{B}(:, :, 1, 2) &= 10^2 \cdot \begin{bmatrix} 8.929298047629347 & 5.557781815526563 \\ 7.032916341629872 & 1.845319512777926 \end{bmatrix}, \quad \mathcal{B}(:, :, 2, 2) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}. \end{aligned}$$

It holds, $\text{rshrank}(\mathcal{A}) = \text{rshrank}(\mathcal{B}) = 2$. In addition, an application of Huan et al. (2018, Algorithm 1) gives the following Moore–Penrose inverse of \mathcal{A} :

$$\begin{aligned} \mathcal{A}^\dagger(:, :, 1, 1) &= \begin{bmatrix} -0.002139621590719 & 0.000961949275511 \\ 0 & 0 \end{bmatrix}, \\ \mathcal{A}^\dagger(:, :, 2, 1) &= 10^{-3} \cdot \begin{bmatrix} 0.154116174683625 & 0.400242571625946 \\ 0 & 0 \end{bmatrix}, \\ \mathcal{A}^\dagger(:, :, 1, 2) &= \begin{bmatrix} 0.001721913469812 & 0.000005405961529 \\ 0 & 0 \end{bmatrix}, \\ \mathcal{A}^\dagger(:, :, 2, 2) &= \begin{bmatrix} 0.004582706318709 & -0.000777570778470 \\ 0 & 0 \end{bmatrix}. \end{aligned}$$

Similarly

$$\begin{aligned} \mathcal{B}^\dagger(:, :, 1, 1) &= \begin{bmatrix} -0.002139080520305 & 0.000961905529252 \\ 0 & 0 \end{bmatrix}, \\ \mathcal{B}^\dagger(:, :, 2, 1) &= 10^{-3} \cdot \begin{bmatrix} 0.153469746601847 & 0.400351263240931 \\ 0 & 0 \end{bmatrix}, \\ \mathcal{B}^\dagger(:, :, 1, 2) &= \begin{bmatrix} 0.001720928456951 & 0.000005485433595 \\ 0 & 0 \end{bmatrix}, \\ \mathcal{B}^\dagger(:, :, 2, 2) &= \begin{bmatrix} 0.004582730894265 & -0.000777786686340 \\ 0 & 0 \end{bmatrix}. \end{aligned}$$

Following (2.5), it is easy to check $\|\mathcal{A}^\dagger\|_2 = 0.005446932213520$, $\|\mathcal{E}\|_2 = 0.149158220173799$ and $\Delta = \|\mathcal{A}^\dagger\|_2 \|\mathcal{E}\|_2 = 8.124547143759799 \cdot 10^{-4} < 1$. Therefore, Condition (4.1) is satisfied. Then, $\|\mathcal{B}^\dagger\|_2 = 0.005446449437497$ and $\frac{\|\mathcal{A}^\dagger\|_2}{1-\Delta} = 0.005451361197625$, which confirms the inequality (4.2) in Lemma 4.1.

Example 5.3 This example is a continuation of Example 5.1 to verify the inequality (4.3) proved in Theorem 4.2. Therefore, for the tensors \mathcal{A} and \mathcal{E} defined in Example 5.1, we have

$$\frac{\|\mathcal{B}^\dagger - \mathcal{A}^\dagger\|_2}{\|\mathcal{A}^\dagger\|_2} = 2.834195478378557 \cdot 10^{-4}$$

and

$$\left(1 + \frac{1}{1-\Delta} + \frac{1}{(1-\Delta)^2}\right) \Delta = 0.018295682536782.$$

Hence, inequality (4.3) of Theorem 4.2 is valid.

Example 5.4 This example is a continuation of Example 5.2 to verify the inequality (4.3) proved in Theorem 4.2. Therefore, for the tensors \mathcal{A} and \mathcal{E} defined in Example 5.2, we have

$$\frac{\|\mathcal{B}^\dagger - \mathcal{A}^\dagger\|_2}{\|\mathcal{A}^\dagger\|_2} = 2.394253885112045 \cdot 10^{-4}$$

and

$$\left(1 + \frac{1}{1-\Delta} + \frac{1}{(1-\Delta)^2}\right) \Delta = 0.002435384431426.$$

Hence, inequality (4.3) of Theorem 4.2 is confirmed.

Example 5.5 We shall use again the settings of Example 5.2 to verify the validity of equality (4.4).

After appropriate calculations, one can verify

$$\begin{aligned}
 & (\mathcal{B} *_N \mathcal{B}^\dagger *_N (\mathcal{I} - \mathcal{A} *_N \mathcal{A}^\dagger)) (:, :, 1, 1) = 10^{-4} \\
 & \cdot \begin{bmatrix} -0.489468003874380 & 0.346424530042466 \\ 0.006594102139740 & 0.970661633334091 \end{bmatrix}, \\
 & (\mathcal{B} *_N \mathcal{B}^\dagger *_N (\mathcal{I} - \mathcal{A} *_N \mathcal{A}^\dagger)) (:, :, 2, 1) = 10^{-4} \\
 & \cdot \begin{bmatrix} 0.441733836613403 & -0.163043151723344 \\ 0.084143275517548 & -0.629731697791430 \end{bmatrix}, \\
 & (\mathcal{B} *_N \mathcal{B}^\dagger *_N (\mathcal{I} - \mathcal{A} *_N \mathcal{A}^\dagger)) (:, :, 1, 2) = 10^{-3} \\
 & \cdot \begin{bmatrix} 0.035216688972592 & -0.046360736953049 \\ -0.013384105875716 & -0.105126013940471 \end{bmatrix}, \\
 & (\mathcal{B} *_N \mathcal{B}^\dagger *_N (\mathcal{I} - \mathcal{A} *_N \mathcal{A}^\dagger)) (:, :, 2, 2) = 10^{-4} \\
 & \cdot \begin{bmatrix} -0.375707154131807 & 0.341422001841479 \\ 0.050538644491456 & 0.869372625158238 \end{bmatrix}.
 \end{aligned}$$

In addition

$$\begin{aligned}
 & (\mathcal{A} *_N \mathcal{A}^\dagger *_N (\mathcal{I} - \mathcal{B} *_N \mathcal{B}^\dagger)) (:, :, 1, 1) = 10^{-4} \\
 & \cdot \begin{bmatrix} 0.489668212727348 & -0.346613307126986 \\ -0.006685674852000 & -0.970692034327758 \end{bmatrix}, \\
 & (\mathcal{A} *_N \mathcal{A}^\dagger *_N (\mathcal{I} - \mathcal{B} *_N \mathcal{B}^\dagger)) (:, :, 2, 1) = 10^{-4} \\
 & \cdot \begin{bmatrix} 0.441825409325317 & 0.163138141635252 \\ -0.084077978807235 & 0.629669045010689 \end{bmatrix}, \\
 & (\mathcal{A} *_N \mathcal{A}^\dagger *_N (\mathcal{I} - \mathcal{B} *_N \mathcal{B}^\dagger)) (:, :, 1, 2) = 10^{-3} \\
 & \cdot \begin{bmatrix} -0.035235566680988 & 0.046380141446686 \\ 0.013393604866949 & 0.105126135280298 \end{bmatrix}, \\
 & (\mathcal{A} *_N \mathcal{A}^\dagger *_N (\mathcal{I} - \mathcal{B} *_N \mathcal{B}^\dagger)) (:, :, 2, 2) = 10^{-4} \\
 & \cdot \begin{bmatrix} 0.375676753138765 & -0.341420788443209 \\ -0.050601297272405 & -0.868951121121841 \end{bmatrix}.
 \end{aligned}$$

Hence

$$\begin{aligned}
 \|\mathcal{B} *_N \mathcal{B}^\dagger *_N (\mathcal{I} - \mathcal{A} *_N \mathcal{A}^\dagger)\|_2 &= \|(\mathcal{A} *_N \mathcal{A}^\dagger *_N (\mathcal{I} - \mathcal{B} *_N \mathcal{B}^\dagger))\|_2 \\
 &= 2.0813844590544 \cdot 10^{-4}.
 \end{aligned}$$

Example 5.6 This example is a continuation of Example 5.2 with the aim to verify the validity of inequality (4.5). It is possible to compute

$$\|\mathcal{G}\|_2 = \|\mathcal{B}^\dagger *_N (\mathcal{I} - \mathcal{A} *_N \mathcal{A}^\dagger)\|_2 = 1.132716947645736 \cdot 10^{-6}.$$

In addition, $\|\mathcal{E}\|_2 \|\mathcal{A}^\dagger\|_2 \|\mathcal{B}^\dagger\|_2 = 4.424993522104842 \cdot 10^{-6}$. Therefore, the inequality (4.5) is verified.

Example 5.7 In this example, we verify cases (1)–(3) of Theorem 4.4.

Case (1) The tensors \mathcal{A} , \mathcal{E} and \mathcal{B} are reused from Example 5.2. It can be verified that $\text{rshrank}(\mathcal{A}) = 2 < 4 = \min(\mathcal{J}, \mathcal{K})$. From example 5.4, we have

$$\frac{\|\mathcal{B}^\dagger - \mathcal{A}^\dagger\|_2}{\|\mathcal{A}^\dagger\|_2} = 2.394253885112045 \cdot 10^{-4}$$

and since $\Delta = \|\mathcal{A}^\dagger\|_2 \|\mathcal{E}\|_2 = 8.124547143759799 \cdot 10^{-4}$ (see Example 5.2), it follows

$$\frac{1 + \sqrt{5}}{2} \cdot \frac{\Delta}{1 - \Delta} = 0.001315648246801.$$

Hence, inequality (4.6) of Theorem 4.4 is valid.

Case (2) We consider the tensors \mathcal{A} , \mathcal{E} and \mathcal{B} from Example 5.1. It is clear that $\text{rshrank}(\mathcal{A}) = 4 = \min(\mathcal{J}, \mathcal{K})$. From Example 5.3, we have

$$\frac{\|\mathcal{B}^\dagger - \mathcal{A}^\dagger\|_2}{\|\mathcal{A}^\dagger\|_2} = 2.834195478378557 \cdot 10^{-4},$$

and since $\Delta = \|\mathcal{A}^\dagger\|_2 \|\mathcal{E}\|_2 = 0.006136135997641$ (see Example 5.1), we have

$$\sqrt{2} \cdot \frac{\Delta}{1 - \Delta} = 0.008731383706299.$$

Hence, inequality (4.6) of Theorem 4.4 is valid.

Case (3) We consider the tensors \mathcal{A} , \mathcal{E} and \mathcal{B} from Example 5.1. Then, $\text{rshrank}(\mathcal{A}) = 4 = \mathcal{J} = \mathcal{K}$. As in the previous case [Case (2)], we have

$$\frac{\|\mathcal{B}^\dagger - \mathcal{A}^\dagger\|_2}{\|\mathcal{A}^\dagger\|_2} = 2.834195478378557 \cdot 10^{-4},$$

and

$$1 \cdot \frac{\Delta}{1 - \Delta} = 0.006174020627866.$$

Hence, inequality (4.6) of Theorem 4.4 is valid.

6 Concluding remarks

The aim of this paper is to generalize some results about the perturbation theory of the matrix pseudoinverse to tensors. For this purpose, we derive several useful representations and introduce some notions. The spectral norm of even-order tensors is defined by a computationally effective definition and investigated. In addition, useful representations of $\mathcal{A} *_{\mathcal{N}} \mathcal{A}^\dagger$ and $\mathcal{I} - \mathcal{A} *_{\mathcal{N}} \mathcal{A}^\dagger$ are derived. As a result, we explore the perturbation bounds for Moore–Penrose inverse of tensor via Einstein product. Unlike to so far exploited approaches which were developed only in the tensor or in the matrix case, our approach assumes an exact transition from one to another space. In this way, it is possible to extend many of known results from the matrix case into the multiarray case. The results derived in current research extend the classical results in the matrix case, derived by Stewart (1977) and Wedin (1973). It is shown that the influence of the perturbation in the tensors depends on exactly defined condition number. Illustrative numerical examples also confirm derived theoretical results.

Recently, Ji and Wei (2017, 2018) investigated the weighted Moore–Penrose inverses and the Drazin inverse of even-order tensors with Einstein product. It is natural to investigate

possible extensions of derived results to the perturbation bounds for the weighted Moore–Penrose inverses and the Drazin inverse of tensors via Einstein product.

Acknowledgements The authors would like to thank the editor and two referees for their detailed comments. This research is supported by the bilateral project between China and Serbia “The theory of tensors, operator matrices and applications (No. 4-5)”. H. Ma would like to thank Prof. Dragana S. Cvetković Ilić for her kind invitation and great hospitality; thank Prof. D.S. Djordjević and Prof. V. Rakočević for their nice monograph (Djordjević and Rakočević 2008). Partial work is completed during her visiting at University of Niš in 2017.

References

- Behera R, Mishra D (2017) Further results on generalized inverses of tensors via the Einstein product. *Linear Multilinear Algebra* 65:1662–1682
- Ben-Israel A, Greville TNE (2003) *Generalized inverses: theory and applications*, 2nd edn. Springer, New York
- Brazell M, Li N, Navasca C, Tamon C (2013) Solving multilinear systems via tensor inversion. *SIAM J Matrix Anal Appl* 34:542–570
- Bu C, Zhang X, Zhou J, Wang W, Wei Y (2014) The inverse, rank and product of tensors. *Linear Algebra Appl* 446:269–280
- Cai LX, Xu WW, Li W (2011) Additive and multiplicative perturbation bounds for the Moore–Penrose inverse. *Linear Algebra Appl* 434:480–489
- Che M, Wei Y (2019) Randomized algorithms for the approximations of Tucker and the tensor train decompositions. *Adv Comput Math* 45:395–428
- Che M, Qi L, Wei Y (2016) Perturbation bounds of tensor eigenvalue and singular value problems with even order. *Linear Multilinear Algebra* 64:622–652
- Che M, Qi L, Wei Y (2019) Stochastic R_0 tensors to stochastic tensor complementarity problems. *Optim Lett* 13:261–279
- Comon P, ten Berge JMF, De Lathauwer L, Castaing J (2009) Generic and typical ranks of multi-way arrays. *Linear Algebra Appl* 430:2997–3007
- Cvetkovic-Ilic DS, Wei Y (2017) *Algebraic properties of generalized inverses*. Springer, Singapore
- Ding W, Wei Y (2016) Solving multi-linear systems with \mathcal{M} -tensors. *J Sci Comput* 68:689–715
- Djordjević DS, Rakočević V (2008) *Lectures on generalized inverses*. Faculty of Sciences and Mathematics, University of Niš, Niš
- Einstein A (2007) The foundation of the general theory of relativity. In: Kox A, Klein M, Schulmann R (eds) *The collected papers of Albert Einstein*, vol 6. Princeton University Press, Princeton, pp 146–200
- Govaerts W, Pryce JD (1989) A Singular value inequality for block matrices. *Linear Algebra Appl* 125:141–148
- Harrison AP, Joseph D (2016) Numeric tensor framework: exploiting and extending Einstein notation. *J Comput Sci* 16:128–139
- Huang S, Zhao G, Chen M (2018) Tensor extreme learning design via generalized Moore–Penrose inverse and triangular type-2 fuzzy sets. *Neural Comput Appl*. <https://doi.org/10.1007/s00521-018-3385-5>
- Ji J, Wei Y (2017) Weighted Moore–Penrose inverses and the fundamental theorem of even-order tensors with Einstein product. *Front Math China* 12:1317–1337
- Ji J, Wei Y (2018) The Drazin inverse of an even-order tensor and its application to singular tensor equations. *Comput Math Appl* 75:3402–3413
- Jin H, Bai M, Benítez J, Liu X (2017) The generalized inverses of tensors and an application to linear models. *Comput Math Appl* 74:385–397
- Li Z (2016) Bounds on the spectral norm and the nuclear norm of a tensor based on tensor partitions. *SIAM J Matrix Anal Appl* 37:1440–1452
- Li W, Ng MK (2014) The perturbation bound for the spectral radius of a nonnegative tensor. *Adv Numer Anal* 2014, Article ID 109525. <https://doi.org/10.1155/2014/109525>
- Li Z, Xu Q, Wei Y (2013) A note on stable perturbations of Moore–Penrose inverses. *Numer Linear Algebra Appl* 20:18–26
- Liang M-L, Zheng B (2018) Gradient-based iterative algorithms for solving Sylvester tensor equations and the associated tensor nearness problems. [arXiv:1811.10378v1](https://arxiv.org/abs/1811.10378v1)
- Liang M-L, Zheng B, Zhao R-J (2019) Tensor inversion and its application to the tensor equations with Einstein product. *Linear Multilinear Algebra* 67:843–870
- Liu X, Wang W, Wei Y (2008) Continuity properties of the $\{1\}$ -inverse and perturbation bounds for the Drazin inverse. *Linear Algebra Appl* 429:1026–1037

- Ma H (2018) Acute perturbation bounds of weighted Moore–Penrose inverse. *Int J Comput Math* 95:710–720
- Medellin D, Ravi VR, Torres-Verdin C (2016) Multidimensional NMR inversion without Kronecker products: multilinear inversion. *J Magn Reson* 269:24–35
- Meng L, Zheng B (2010) The optimal perturbation bounds of the Moore–Penrose inverse under the Frobenius norm. *Linear Algebra Appl* 432:956–963
- Meng L, Zheng B, Ma P (2017) Perturbation bounds of generalized inverses. *Appl Math Comput* 298:88–100
- Meyer C (1980) The condition number of a finite Markov chain and perturbation bounds for the limiting probabilities. *SIAM J Algebr Discrete Methods* 1:273–283
- Miao Y, Qi L, Wei Y (2019) Generalized tensor function via the tensor singular value decomposition based on the T-product. [arXiv:1901.0425v2](https://arxiv.org/abs/1901.0425v2)
- Panigrahy K, Mishra D (2018) An extension of the Moore–Penrose inverse of a tensor via the Einstein product. [arXiv:1806.03655v1](https://arxiv.org/abs/1806.03655v1)
- Panigrahy K, Mishra D (2019) Reverse order law for weighted Moore–Penrose inverses of tensors. [arXiv:1901.01527v1](https://arxiv.org/abs/1901.01527v1)
- Panigrahy K, Behera R, Mishra D (2018) Reverse order law for the Moore–Penrose inverses of tensors. *Linear Multilinear Algebra*. <https://doi.org/10.1080/03081087.2018.1502252>
- Qi L (2005) Eigenvalues of a real supersymmetric tensor. *J Symb Comput* 40:1302–1324
- Qi L, Luo Z (2017) Tensor analysis: spectral theory and special tensors. SIAM, Philadelphia
- Qiao S, Wang X, Wei Y (2018) Two finite-time convergent Zhang neural network models for time-varying complex matrix Drazin inverse. *Linear Algebra Appl* 542:101–117
- Shi X, Wei Y, Ling S (2013) Backward error and perturbation bounds for high order Sylvester tensor equation. *Linear Multilinear Algebra* 61:1436–1446
- Stanimirović PS, Ćirić M, Katsikis VN, Li C, Ma H (2018) Outer and (b, c) inverses of tensors. *Linear Multilinear Algebra*. <https://doi.org/10.1080/03081087.2018.1521783>
- Stewart G (1977) On the perturbation of pseudoinverses, projections and linear least squares problems. *SIAM Rev* 19:634–662
- Sun L, Zheng B, Bu C, Wei Y (2014) Some results on the generalized inverse of tensors and idempotent tensors. [arXiv:1412.7430v1](https://arxiv.org/abs/1412.7430v1)
- Sun L, Zheng B, Bu C, Wei Y (2016) Moore–Penrose inverse of tensors via Einstein product. *Linear Multilinear Algebra* 64:686–698
- Sun L, Zheng B, Bu C, Wei Y (2018) Generalized inverses of tensors via a general product of tensors. *Front Math China* 13:893–911
- Wang G, Wei Y, Qiao S (2018) Generalized inverses: theory and computations, 2nd edn, *Developments in Mathematics*, vol 53. Springer, Singapore; Science Press, Beijing
- Wedin P (1973) Perturbation theory for pseudo-inverses. *BIT* 13:217–232
- Wei Y (1999) On the perturbation of the group inverse and oblique projection. *Appl Math Comput* 98:29–42
- Wei Y (2014) Generalized inverses of matrices, chapter 27. In: Hogben L (ed) *Handbook of linear algebra*, 2nd edn. Chapman and Hall/CRC, Boca Raton
- Wei Y (2017) Acute perturbation of the group inverse. *Linear Algebra Appl* 534:135–157
- Wei Y, Ding W (2016) Theory and computation of tensors. Academic Press, London
- Wei M, Ling S (2010) On the perturbation bounds of g-inverse and oblique projections. *Linear Algebra Appl* 433:1778–1792
- Xie P, Xiang H, Wei Y (2019) Randomized algorithms for total least squares problems. *Numer Linear Algebra Appl* 26(1):e2219 16 pp
- Xu Z, Sun J, Gu C (2008) Perturbation for a pair of oblique projectors $AA_{M,N}^{\dagger}$ and $AA_{M,N}^{\ddagger}$. *Appl Math Comput* 203:432–446
- Xu Z, Gu C, Feng B (2010a) Weighted acute perturbation for two matrices. *Arab J Sci Eng* 35:129–143
- Xu Q, Wei Y, Gu Y (2010b) Sharp norm-estimations for Moore–Penrose inverses of stable perturbations of Hilbert C^* -module operators. *SIAM J Numer Anal* 47:4735–4758
- Xu W, Cai L, Li W (2011) The optimal perturbation bounds for the weighted Moore–Penrose inverse. *Electron J Linear Algebra* 22:521–538
- Zhang N, Wei Y (2008) A note on the perturbation of an outer inverse. *Calcolo* 45:263–273
- Zheng B, Meng L, Wei Y (2017) Condition numbers of the multidimensional total least squares problem. *SIAM J Matrix Anal Appl* 38:924–948