



On the tensor spectral p -norm and its dual norm via partitions

Bilian Chen^{1,2} · Zhening Li³

Received: 18 December 2018 / Published online: 20 February 2020
© The Author(s) 2020

Abstract

This paper presents a generalization of the spectral norm and the nuclear norm of a tensor via arbitrary tensor partitions, a much richer concept than block tensors. We show that the spectral p -norm and the nuclear p -norm of a tensor can be lower and upper bounded by manipulating the spectral p -norms and the nuclear p -norms of subtensors in an arbitrary partition of the tensor for $1 \leq p \leq \infty$. Hence, it generalizes and answers affirmatively the conjecture proposed by Li (SIAM J Matrix Anal Appl 37:1440–1452, 2016) for a tensor partition and $p = 2$. We study the relations of the norms of a tensor, the norms of matrix unfoldings of the tensor, and the bounds via the norms of matrix slices of the tensor. Various bounds of the tensor spectral and nuclear norms in the literature are implied by our results.

Keywords Tensor norm bound · Spectral norm · Nuclear norm · Tensor partition · Block tensor

Mathematics Subject Classification 15A60 · 15A69 · 90C59 · 68Q17

1 Introduction

The spectral p -norm of a tensor generalizes the spectral p -norm of a matrix. It can be defined by the L_p -sphere constrained multilinear form optimization problem:

Research of the first author was supported in part by National Natural Science Foundation of China (Grants 61772442 and 11671335)

✉ Zhening Li
zheningli@gmail.com

Bilian Chen
blchen@xmu.edu.cn

¹ Department of Automation, Xiamen University, Xiamen 361005, China

² Xiamen Key Laboratory of Big Data Intelligent Analysis and Decision, Xiamen 361005, China

³ School of Mathematics and Physics, University of Portsmouth, Portsmouth PO1 3HF, UK

$$\|\mathcal{T}\|_{p_\sigma} = \max \{ \mathcal{T}(\mathbf{x}^1, \mathbf{x}^2, \dots, \mathbf{x}^d) : \|\mathbf{x}^k\|_p = 1, \mathbf{x}^k \in \mathbb{R}^{n_k}, k = 1, 2, \dots, d \},$$

where $\|\mathcal{T}\|_{p_\sigma}$ denotes the spectral p -norm of a given tensor $\mathcal{T} = (t_{i_1 i_2 \dots i_d}) \in \mathbb{R}^{n_1 \times n_2 \times \dots \times n_d}$,

$$\mathcal{T}(\mathbf{x}^1, \mathbf{x}^2, \dots, \mathbf{x}^d) = \sum_{i_1=1}^{n_1} \sum_{i_2=1}^{n_2} \dots \sum_{i_d=1}^{n_d} t_{i_1 i_2 \dots i_d} x_{i_1}^1 x_{i_2}^2 \dots x_{i_d}^d \quad (1)$$

is a multilinear form of $(\mathbf{x}^1, \mathbf{x}^2, \dots, \mathbf{x}^d)$, and $\|\cdot\|_p$ denotes the L_p -norm of a vector for $1 \leq p \leq \infty$. When the order of the tensor \mathcal{T} is two, the problem is reduced to the spectral p -norm of a matrix, and in particular when $p = 2$, to the spectral norm or the largest singular value of a matrix. The spectral p -norm of a tensor was proposed by Lim [18] in terms of singular values of a tensor, and is closely related to the largest Z -eigenvalue (for the case $p = 2$) of a tensor proposed by Qi [24].

The matrix spectral p -norm is evidently important in many branches of mathematics as well as in various practical applications; see e.g., [6, 11]. The complexity and approximation methods of the matrix spectral p -norm were studied extensively [1, 21, 27], and they have particular applications in robust optimization [27]. When $p = 1, 2$, the matrix spectral p -norm can be computed easily, and when $2 < p \leq \infty$, computing the matrix spectral p -norm is NP-hard, while it remains unknown for the rest of p . The tensor spectral p -norm was studied mainly in approximation algorithms of polynomial optimization [15]. When the order of a tensor is larger than two, computing the tensor spectral norm ($p = 2$) is already NP-hard proved by He et al. [8] (see also [10]), a sharp contrast to the case of matrices. NP-hardness to compute the tensor spectral p -norm was also established when $2 < p \leq \infty$ by Hou and So [12]. Various approximation bounds of the tensor spectral p -norm were established in the literature [7–9, 12, 26]. Nikiforov [23] studied the tensor spectral p -norm using combinatorial methods and proposed several bounds. Li and Zhao [17] recently studied a more general tensor spectral p -norm and provided upper bounds via norm compression tensors.

The dual norm to the spectral p -norm of a tensor \mathcal{T} , called the nuclear p -norm, is defined as $\|\mathcal{T}\|_{p_*} = \max_{\|\mathcal{X}\|_{p_\sigma} \leq 1} \langle \mathcal{T}, \mathcal{X} \rangle$. In the case of matrices and $p = 2$, it is reduced to the nuclear norm of a matrix, which is equal to the sum of all the singular values of a matrix. The matrix nuclear norm was used widely as a convex envelope of the matrix rank for many rank minimization problems, such as matrix completion [2]. Friedland and Lim [4] studied the tensor nuclear p -norm systematically, and showed that computing the tensor nuclear norm ($p = 2$) is NP-hard when the order of the tensor is larger than two. They also proposed simple lower and upper bounds of the tensor spectral norm and nuclear norm. The study on the tensor nuclear p -norm has been mainly focused on the case $p = 2$, such as tensor completion [5, 20, 30]. Derksen [3] discussed the nuclear norm of various tensors based on orthogonality. Nie [22] studied symmetric tensor nuclear norms. Extremal properties of the tensor spectral norm and nuclear norm were studied in [16].

Most of the methods to tackle the tensor spectral p -norm and nuclear p -norm in the literature have been heavily relying on matrix unfoldings, no matter in theory

such as approximation methods [15] and in practice such as tensor completion [5]. Hu [13] established the relation of the tensor nuclear norm to the nuclear norms of its matrix unfoldings. Wang et al. [29] systematically studied the tensor spectral p -norm via various matrix unfoldings and tensor unfoldings. Li [14] proposed a novel approach to study the tensor spectral norm and nuclear norm via tensor partitions, a concept generalizing block tensors by Ragnarsson and Van Loan [25]. Some neat bounds of the tensor spectral norm (respectively, nuclear norm) via the spectral norms (respectively, nuclear norms) of subtensors in any regular partition were proposed, and a conjecture [14, Conjecture 3.5] on the bounds in any tensor partition was proposed.

In this paper, we systematically study the tensor spectral p -norm and nuclear p -norm via the partition approach in [14]. We prove that for the most general partition called arbitrary partition, the bounds of the tensor spectral p -norm and nuclear p -norm via subtensors can be established for any $1 \leq p \leq \infty$. It generalizes and answers affirmatively the Li's conjecture, which is the case $p = 2$ for a tensor partition. The novelty of the proof lies in establishing an index system to describe subtensors in an arbitrary partition. Based on these, we study the relations of the spectral p -norm of a tensor, the spectral p -norms of matrix unfoldings of the tensor, and the bounds via the spectral p -norms of matrix slices of the tensor. The same relation is studied for the tensor nuclear p -norm. Various bounds of these tensor norms in the literature can be derived from our results.

This paper is organized as follows. We start with the preparation of various notations, definitions and properties of tensor norms and tensor partitions in Sect. 2. In Sect. 3, we present our main result on bounding the tensor spectral p -norm and nuclear p -norm via partitioned subtensors. Section 4 is devoted to the discussion and theoretical applications, particularly on the relations among the tensor norms, the norms of matrix unfoldings, and the norms via matrix slices.

2 Preparation

Throughout this paper, we uniformly use the lower case letters (e.g., x), the boldface lower case letters (e.g., $\mathbf{x} = (x_i)$), the capital letters (e.g., $X = (x_{ij})$), and the calligraphic letters (e.g., $\mathcal{X} = (x_{i_1 i_2 \dots i_d})$) to denote scalars, vectors, matrices, and higher order (order three or more) tensors, respectively. Denote $\mathbb{R}^{n_1 \times n_2 \times \dots \times n_d}$ to be the space of d th order real tensors of dimension $n_1 \times n_2 \times \dots \times n_d$. The same notations apply for a vector space and a matrix space when $d = 1$ and $d = 2$, respectively. Denote \mathbb{N} to be the set of positive integers.

Given a d th order tensor space $\mathbb{R}^{n_1 \times n_2 \times \dots \times n_d}$, we denote $\mathbb{I}^k := \{1, 2, \dots, n_k\}$ to be the index set of mode- k for $k = 1, 2, \dots, d$. Trivially, $\mathbb{I}^1 \times \mathbb{I}^2 \times \dots \times \mathbb{I}^d$ becomes the index set of the entries of a tensor in the tensor space. The Frobenius inner product of two tensors $\mathcal{U}, \mathcal{V} \in \mathbb{R}^{n_1 \times n_2 \times \dots \times n_d}$ is defined as:

$$\langle \mathcal{U}, \mathcal{V} \rangle := \sum_{i_1=1}^{n_1} \sum_{i_2=1}^{n_2} \dots \sum_{i_d=1}^{n_d} u_{i_1 i_2 \dots i_d} v_{i_1 i_2 \dots i_d}.$$

Its induced Frobenius norm is naturally defined as $\|\mathcal{T}\|_2 := \sqrt{\langle \mathcal{T}, \mathcal{T} \rangle}$. When $d = 1$, the Frobenius norm is reduced to the Euclidean norm of a vector. In a similar vein, we may define the L_p -norm of a tensor (also known as the Hölder p -norm) for $1 \leq p \leq \infty$ to looking at a tensor as a vector, as follows:

$$\|\mathcal{T}\|_p = \left(\sum_{i_1=1}^{n_1} \sum_{i_2=1}^{n_2} \cdots \sum_{i_d=1}^{n_d} |t_{i_1 i_2 \dots i_d}|^p \right)^{\frac{1}{p}}.$$

A rank-one tensor, also called a simple tensor, is a tensor that can be written as outer products of vectors, i.e., $\mathcal{T} = \mathbf{x}^1 \otimes \mathbf{x}^2 \otimes \cdots \otimes \mathbf{x}^d$ where $\mathbf{x}^k \in \mathbb{R}^{n_k}$ for $k = 1, 2, \dots, d$. It can be equivalently represented by the entries as:

$$t_{i_1 i_2 \dots i_d} = \prod_{k=1}^d x_{i_k}^k \quad \forall (i_1, i_2, \dots, i_d) \in \mathbb{1}^1 \times \mathbb{1}^2 \times \cdots \times \mathbb{1}^d. \tag{2}$$

Here is a property of the L_p -norm of a rank-one tensor.

Proposition 2.1 *If a tensor $\mathcal{T} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$ is rank-one, say $\mathcal{T} = \mathbf{x}^1 \otimes \mathbf{x}^2 \otimes \cdots \otimes \mathbf{x}^d$, then $\|\mathcal{T}\|_p = \prod_{k=1}^d \|\mathbf{x}^k\|_p$ for any $1 \leq p \leq \infty$.*

Proof According to (2), we have

$$\|\mathcal{T}\|_p = \left(\sum_{i_1=1}^{n_1} \sum_{i_2=1}^{n_2} \cdots \sum_{i_d=1}^{n_d} \left| \prod_{k=1}^d x_{i_k}^k \right|^p \right)^{\frac{1}{p}} = \left(\prod_{k=1}^d \left(\sum_{i_k=1}^{n_k} |x_{i_k}^k|^p \right) \right)^{\frac{1}{p}} = \prod_{k=1}^d \|\mathbf{x}^k\|_p. \quad \square$$

2.1 The spectral p -norm and nuclear p -norm

Let us formally define the tensor spectral p -norm and its dual norm.

Definition 2.2 For a given tensor $\mathcal{T} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_d}$ and $1 \leq p \leq \infty$, the spectral p -norm of \mathcal{T} , denoted by $\|\mathcal{T}\|_{p_\sigma}$, is defined as

$$\|\mathcal{T}\|_{p_\sigma} := \max \{ \langle \mathcal{T}, \mathbf{x}^1 \otimes \mathbf{x}^2 \otimes \cdots \otimes \mathbf{x}^d \rangle : \|\mathbf{x}^k\|_p = 1, k = 1, 2, \dots, d \}. \tag{3}$$

Essentially, $\|\mathcal{T}\|_{p_\sigma}$ is the maximal value of the Frobenius inner product between \mathcal{T} and a rank-one tensor whose L_p -norm is one, according to Proposition 2.1. We remark that $\langle \mathcal{T}, \mathbf{x}^1 \otimes \mathbf{x}^2 \otimes \cdots \otimes \mathbf{x}^d \rangle$ in (3) is exactly the multilinear form $\mathcal{T}(\mathbf{x}^1, \mathbf{x}^2, \dots, \mathbf{x}^d)$ defined in (1). Hence, as mentioned in Sect. 1, the tensor spectral p -norm is more commonly known as the L_p -sphere constrained multilinear form optimization problem in the optimization community. When $p = 2$, the tensor spectral p -norm is often called the tensor spectral norm, and is also known to be the largest singular value of the tensor [18].

Definition 2.3 For a given tensor $\mathcal{T} \in \mathbb{R}^{n_1 \times n_2 \times \dots \times n_d}$ and $1 \leq p \leq \infty$, the nuclear p -norm of \mathcal{T} , denoted by $\|\mathcal{T}\|_{p_*}$, is defined as

$$\|\mathcal{T}\|_{p_*} := \min \left\{ \sum_{i=1}^r |\lambda_i| : \mathcal{T} = \sum_{i=1}^r \lambda_i \mathbf{x}_i^1 \otimes \mathbf{x}_i^2 \otimes \dots \otimes \mathbf{x}_i^d, \|\mathbf{x}_i^k\|_p = 1 \text{ for all } k \text{ and } i, r \in \mathbb{N} \right\}. \tag{4}$$

The decomposition of \mathcal{T} into a sum of rank-one tensors, such as that in (4), is called a rank-one decomposition of \mathcal{T} . Therefore, the tensor nuclear p -norm is the minimum of the sum of the L_p -norms of rank-one tensors in any rank-one decomposition. A rank-one decomposition of \mathcal{T} that attains $\|\mathcal{T}\|_{p_*}$ is called a nuclear p -decomposition of \mathcal{T} , similar to the nuclear decomposition of a tensor for $p = 2$ discussed in [4]. When $p = 2$, the tensor nuclear p -norm is commonly known as the tensor nuclear norm. The tensor nuclear norm is the convex envelope of the tensor rank and is widely used in tensor completion [30].

We provide some basic facts of the tensor spectral p -norm and nuclear p -norm. The proof is essentially based on the Hölder’s inequality.

Proposition 2.4 For any $1 \leq p, q \leq \infty$ with $\frac{1}{p} + \frac{1}{q} = 1$, we have the followings:

- For a scalar $t \in \mathbb{R}$, $\|t\|_{p_\sigma} = \|\mathcal{T}\|_{p_*} = |t|$;
- For a vector $\mathbf{t} \in \mathbb{R}^n$, $\|\mathbf{t}\|_{p_\sigma} = \|\mathbf{t}\|_q$ and $\|\mathbf{t}\|_{p_*} = \|\mathbf{t}\|_p$;
- For a rank-one tensor \mathcal{T} , $\|\mathcal{T}\|_{p_\sigma} = \|\mathcal{T}\|_q$ and $\|\mathcal{T}\|_{p_*} = \|\mathcal{T}\|_p$.

The tensor nuclear p -norm is the dual norm to the tensor spectral p -norm, and vice versa, for any $1 \leq p \leq \infty$.

Lemma 2.5 For given tensors \mathcal{T} and \mathcal{Z} in a same tensor space and $1 \leq p \leq \infty$, it follows that

$$\langle \mathcal{T}, \mathcal{Z} \rangle \leq \|\mathcal{T}\|_{p_\sigma} \|\mathcal{Z}\|_{p_*},$$

and further

$$\begin{aligned} \|\mathcal{T}\|_{p_\sigma} &= \max_{\|\mathcal{Z}\|_{p_*} \leq 1} \langle \mathcal{T}, \mathcal{Z} \rangle, \\ \|\mathcal{T}\|_{p_*} &= \max_{\|\mathcal{Z}\|_{p_\sigma} \leq 1} \langle \mathcal{T}, \mathcal{Z} \rangle. \end{aligned} \tag{5}$$

Proof Let $\mathcal{Z} = \sum_{i=1}^r \lambda_i \mathbf{x}_i^1 \otimes \mathbf{x}_i^2 \otimes \dots \otimes \mathbf{x}_i^d$ with $\|\mathbf{x}_i^k\|_p = 1$ for all k and i with $\|\mathcal{Z}\|_{p_*} = \sum_{i=1}^r |\lambda_i|$, i.e., a nuclear p -decomposition of \mathcal{Z} . By Definition 2.2,

$$\langle \mathcal{T}, \mathbf{x}_i^1 \otimes \mathbf{x}_i^2 \otimes \dots \otimes \mathbf{x}_i^d \rangle \leq \|\mathcal{T}\|_{p_\sigma} \quad \forall i = 1, 2, \dots, r,$$

which leads to

$$\langle \mathcal{T}, \mathcal{Z} \rangle = \sum_{i=1}^r \lambda_i \langle \mathcal{T}, \mathbf{x}_i^1 \otimes \mathbf{x}_i^2 \otimes \dots \otimes \mathbf{x}_i^d \rangle \leq \sum_{i=1}^r |\lambda_i| \cdot \|\mathcal{T}\|_{p_\sigma} = \|\mathcal{T}\|_{p_\sigma} \|\mathcal{Z}\|_{p_*}.$$

By choosing $\|\mathcal{Z}\|_{p_*} \leq 1$, we have

$$\max_{\|\mathcal{Z}\|_{p_*} \leq 1} \langle \mathcal{T}, \mathcal{Z} \rangle \leq \max_{\|\mathcal{Z}\|_{p_*} \leq 1} \|\mathcal{T}\|_{p_\sigma} \|\mathcal{Z}\|_{p_*} = \|\mathcal{T}\|_{p_\sigma}.$$

On the other hand, let $\|\mathcal{T}\|_{p_\sigma} = \langle \mathcal{T}, \mathbf{y}^1 \otimes \mathbf{y}^2 \otimes \dots \otimes \mathbf{y}^d \rangle$ with $\|\mathbf{y}^k\|_p = 1$ for all k . By Proposition 2.4, we have

$$\|\mathbf{y}^1 \otimes \mathbf{y}^2 \otimes \dots \otimes \mathbf{y}^d\|_{p_*} = \|\mathbf{y}^1 \otimes \mathbf{y}^2 \otimes \dots \otimes \mathbf{y}^d\|_p = \prod_{k=1}^d \|\mathbf{y}^k\|_p = 1,$$

which leads to

$$\max_{\|\mathcal{Z}\|_{p_*} \leq 1} \langle \mathcal{T}, \mathcal{Z} \rangle \geq \langle \mathcal{T}, \mathbf{y}^1 \otimes \mathbf{y}^2 \otimes \dots \otimes \mathbf{y}^d \rangle = \|\mathcal{T}\|_{p_\sigma}.$$

Therefore, $\max_{\|\mathcal{Z}\|_{p_*} \leq 1} \langle \mathcal{T}, \mathcal{Z} \rangle = \|\mathcal{T}\|_{p_\sigma}$ and so as to the other dual norm equality. \square

We remark that the proof of Lemma 2.5 for $p = 2$ can be found in [3, 19]. When $d = 2$, the tensor spectral p -norm and nuclear p -norm are reduced to the matrix spectral p -norm and nuclear p -norm, respectively. When $d = 1$, a vector, its spectral p -norm is the L_q -norm where $\frac{1}{p} + \frac{1}{q} = 1$ and its nuclear p -norm is the L_p -norm, as mentioned in Proposition 2.4. Two extreme cases of these norms worth mentioning, and they are the only known easy cases to compute.

Proposition 2.6 *For any tensor \mathcal{T} , it follows that $\|\mathcal{T}\|_{1_\sigma} = \|\mathcal{T}\|_\infty$ and $\|\mathcal{T}\|_{1_*} = \|\mathcal{T}\|_1$.*

Proof Let $|t_{s_1 s_2 \dots s_d}| = \max_{i_k \in \mathbb{I}^k, k=1,2,\dots,d} |t_{i_1 i_2 \dots i_d}| = \|\mathcal{T}\|_\infty$. For any $\mathbf{x}^k \in \mathbb{R}^{n_k}$ with $\|\mathbf{x}^k\|_1 = 1$ for $k = 1, 2, \dots, d$,

$$\begin{aligned} \langle \mathcal{T}, \mathbf{x}^1 \otimes \mathbf{x}^2 \otimes \dots \otimes \mathbf{x}^d \rangle &= \sum_{i_1=1}^{n_1} \sum_{i_2=1}^{n_2} \dots \sum_{i_d=1}^{n_d} t_{i_1 i_2 \dots i_d} x_{i_1}^1 x_{i_2}^2 \dots x_{i_d}^d \\ &\leq \sum_{i_1=1}^{n_1} \sum_{i_2=1}^{n_2} \dots \sum_{i_d=1}^{n_d} |t_{s_1 s_2 \dots s_d}| \cdot |x_{i_1}^1 x_{i_2}^2 \dots x_{i_d}^d| \\ &= |t_{s_1 s_2 \dots s_d}| \prod_{k=1}^n \|\mathbf{x}^k\|_1 \\ &= |t_{s_1 s_2 \dots s_d}|, \end{aligned}$$

implying that $\|\mathcal{T}\|_{1_\sigma} \leq \|\mathcal{T}\|_\infty$. On the other hand, denote \mathbf{e}^i to be the vector whose i th entry is one and others are zeros. Clearly $\|\mathbf{e}^i\|_1 = 1$, and we have

$$\langle \mathcal{T}, \mathbf{e}^{s_1} \otimes \mathbf{e}^{s_2} \otimes \dots \otimes \mathbf{e}^{s_d} \rangle = t_{s_1 s_2 \dots s_d},$$

implying that $\|\mathcal{T}\|_{1_\sigma} \geq |t_{s_1 s_2 \dots s_d}| = \|\mathcal{T}\|_\infty$. Therefore, $\|\mathcal{T}\|_{1_\sigma} = \|\mathcal{T}\|_\infty$, and the other identity follows since the dual norm of the tensor L_1 -norm is the tensor L_∞ -norm. \square

2.2 Tensor partitions

A matrix can be partitioned into submatrices, the same can be applied to a tensor. One important class of tensor partitions, block tensors, was proposed and studied in [25, 28]. It is a straightforward generalization of block matrices. Li [14] proposed three types of partitions for tensors, namely, modal partitions (an alternative name for block tensors), regular partitions, and tensor partitions, with the latter generalizing the former. Some neat bounds on the tensor spectral norm and nuclear norm based on *regular partitions* were proposed in [14]. The proofs heavily relied on the recursive structure in defining regular partitions. Since we are extending the results to a more general class of partitions than *tensor partitions*, we only discuss the definition of tensor partitions and refer modal partitions and regular partitions to [14].

Before presenting the partition concepts, we first discuss notations to describe subtensors of a tensor. It is also an essential step to prove our main bounds to be established in Sect. 3. Suppose that \mathcal{T}_j is a subtensor of a tensor $\mathcal{T} \in \mathbb{R}^{n_1 \times n_2 \times \dots \times n_d}$. We denote the set of its mode- k indices in the original tensor \mathcal{T} to be \mathbb{I}_j^k for $k = 1, 2, \dots, d$. We then let

$$\mathcal{T}_j = \mathcal{T}(\mathbb{I}_j^1, \mathbb{I}_j^2, \dots, \mathbb{I}_j^d) \text{ where } \mathbb{I}_j^k \subset \mathbb{I}^k \text{ for } k = 1, 2, \dots, d.$$

Specifically, \mathcal{T}_j is a subtensor of \mathcal{T} by keeping only the indices in \mathbb{I}_j^k of mode- k for $k = 1, 2, \dots, d$. Alternatively, \mathcal{T}_j is a subtensor by deleting all the indices in $\mathbb{I}^k \setminus \mathbb{I}_j^k$ of mode- k for $k = 1, 2, \dots, d$ from the original tensor \mathcal{T} . The dimension of the subtensor \mathcal{T}_j is $|\mathbb{I}_j^1| \times |\mathbb{I}_j^2| \times \dots \times |\mathbb{I}_j^d|$. In our analysis, we do not relabel the indices of some mode of \mathcal{T}_j , say \mathbb{I}_j^k , to $\{1, 2, \dots, |\mathbb{I}_j^k|\}$, but keep its original indices in \mathcal{T} .

Definition 2.7 [14, Definition 2.4] A partition $\{\mathcal{T}_1, \mathcal{T}_2, \dots, \mathcal{T}_m\}$ is called a *tensor partition* of a tensor \mathcal{T} , if

- every \mathcal{T}_j ($j = 1, 2, \dots, m$) can be written as $\mathcal{T}(\mathbb{I}_j^1, \mathbb{I}_j^2, \dots, \mathbb{I}_j^d)$ where the indices of every $\mathbb{I}_j^k \subset \mathbb{I}^k$ ($k = 1, 2, \dots, d$) are consecutive,
- every pair $\{\mathcal{T}_i, \mathcal{T}_j\}$ with $i \neq j$ has no common entry of \mathcal{T} , and
- every entry of \mathcal{T} belongs to one of $\{\mathcal{T}_1, \mathcal{T}_2, \dots, \mathcal{T}_m\}$.

We remark that as a tensor partition, every subtensor \mathcal{T}_j must be a whole block (not disconnected) from the original tensor \mathcal{T} . The following observation is straightforward from Definition 2.7.

Proposition 2.8 If $\{\mathcal{T}_1, \mathcal{T}_2, \dots, \mathcal{T}_m\}$ is a tensor partition of a tensor \mathcal{T} where

$$\mathcal{T}_j = \mathcal{T}(\mathbb{I}_j^1, \mathbb{I}_j^2, \dots, \mathbb{I}_j^d) \quad \forall j = 1, 2, \dots, m,$$

then $\{\mathbb{I}_j^1 \times \mathbb{I}_j^2 \times \dots \times \mathbb{I}_j^d : j = 1, 2, \dots, m\}$ is a partition of $\mathbb{I}^1 \times \mathbb{I}^2 \times \dots \times \mathbb{I}^d$, the index set of \mathcal{T} .

In a similar way, we denote $\mathbf{x}(\mathbb{I}_j^k) \in \mathbb{R}^{|\mathbb{I}_j^k|}$ to be the vector by keeping only the entries of \mathbf{x} with indices in \mathbb{I}_j^k , or the vector by deleting the entries of \mathbf{x} whose indices are not in \mathbb{I}_j^k . Again, in our analysis, we do not relabel these indices to $\{1, 2, \dots, |\mathbb{I}_j^k|\}$.

We remark that Proposition 2.8 indeed implies a more general partition concept than the tensor partition in Definition 2.7. We may further drop the requirement of the indices of \mathbb{I}_j^k to be consecutive for \mathcal{T}_j . In this case, \mathcal{T}_j may consist several disconnected pieces by viewing from the original tensor \mathcal{T} but can be put together to form a tensor by deleting empty entries from \mathcal{T} (see Example 2.10). Although one can relabel some mode- k indices (similar operations to swapping rows or columns in a matrix) to make one of \mathcal{T}_j 's to be a tensor with consecutive indices in every mode, it may break other \mathcal{T}_j 's into disconnected pieces. Hence, one can define a more general partition concept that allows disconnections.

Definition 2.9 A partition $\{\mathcal{T}_1, \mathcal{T}_2, \dots, \mathcal{T}_m\}$ where

$$\mathcal{T}_j = \mathcal{T}(\mathbb{I}_j^1, \mathbb{I}_j^2, \dots, \mathbb{I}_j^d) \quad \forall j = 1, 2, \dots, m$$

and $\mathbb{I}_j^k \subset \mathbb{I}^k$ for $k = 1, 2, \dots, d$ and $j = 1, 2, \dots, m$ is called an *arbitrary partition* of a tensor $\mathcal{T} \in \mathbb{R}^{n_1 \times n_2 \times \dots \times n_d}$ if $\{\mathbb{I}_j^1 \times \mathbb{I}_j^2 \times \dots \times \mathbb{I}_j^d : j = 1, 2, \dots, m\}$ is a partition of $\mathbb{I}^1 \times \mathbb{I}^2 \times \dots \times \mathbb{I}^d$.

Arbitrary partitions is the most general case of partitioning a tensor. The following example indicates the key difference between a tensor partition and an arbitrary partition for a matrix. Obviously, arbitrary partitions can be far more complicated than tensor partitions for higher order tensors.

Example 2.10 Let $M \in \mathbb{R}^{4 \times 6}$ be a matrix shown as 4×6 blocks in Fig. 1.

- For (a), $\{A, B, C, D, E, F\}$ is a tensor partition (a special arbitrary partition) of M with $A, B, C, D \in \mathbb{R}^{2 \times 2}$ and $E, F \in \mathbb{R}^{1 \times 4}$.
- For (b), $\{U, V, W, X, Y, Z\}$ is an arbitrary partition (but not a tensor partition) of M with $U, V, W \in \mathbb{R}^{2 \times 2}$ and $X, Y, Z \in \mathbb{R}^{1 \times 4}$. Here $V = \begin{pmatrix} V_1 \\ V_2 \end{pmatrix}$, $W = \begin{pmatrix} W_1 \\ W_2 \end{pmatrix}$, and $Y = (Y_1, Y_2)$ are disconnected in M .

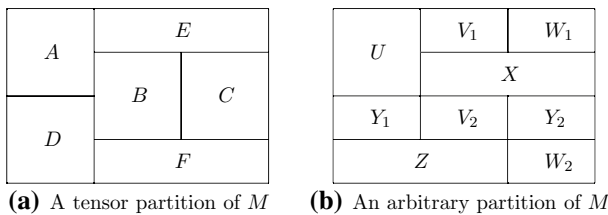


Fig. 1 Tensor partition and arbitrary partition of a second order tensor (matrix)

In particular, there is no way for a tensor partition of a 4×6 matrix consisting of exactly three 2×2 matrices and three 1×4 matrices. However, an arbitrary partition can make it, such as the partition in the right subfigure of Fig. 1.

Finally in this section, we remark that some \mathcal{T}_j (either connected or disconnected) in an arbitrary partition of a tensor may not have the same order of the original tensor \mathcal{T} . If some \mathbb{I}_j^k contains only one index, this causes the disappearance of mode- k and reduces the order of \mathcal{T}_j by one. However, we still treat this \mathcal{T}_j as a d th order tensor by keeping the dimension of mode- k to be one. For instance, we can always treat a scalar as a one-dimensional vector, or a one-by-one matrix.

3 Bounds of the tensor norms

With the establishment of the index system to describe subtensors in an arbitrary partition, we are now in a better position to present and prove the main results in this paper, bounding the spectral p -norm and the nuclear p -norm of a tensor via the spectral p -norms and the nuclear p -norms of subtensors in an arbitrary partition.

Theorem 3.1 *If $\{\mathcal{T}_1, \mathcal{T}_2, \dots, \mathcal{T}_m\}$ is an arbitrary partition of a tensor \mathcal{T} and $1 \leq p, q \leq \infty$ with $\frac{1}{p} + \frac{1}{q} = 1$, then*

$$\left\| (\|\mathcal{T}_1\|_{p_\sigma}, \|\mathcal{T}_2\|_{p_\sigma}, \dots, \|\mathcal{T}_m\|_{p_\sigma}) \right\|_\infty \leq \|\mathcal{T}\|_{p_\sigma} \leq \left\| (\|\mathcal{T}_1\|_{p_\sigma}, \|\mathcal{T}_2\|_{p_\sigma}, \dots, \|\mathcal{T}_m\|_{p_\sigma}) \right\|_q, \tag{6}$$

$$\left\| (\|\mathcal{T}_1\|_{p_*}, \|\mathcal{T}_2\|_{p_*}, \dots, \|\mathcal{T}_m\|_{p_*}) \right\|_p \leq \|\mathcal{T}\|_{p_*} \leq \left\| (\|\mathcal{T}_1\|_{p_*}, \|\mathcal{T}_2\|_{p_*}, \dots, \|\mathcal{T}_m\|_{p_*}) \right\|_1. \tag{7}$$

Proof For an arbitrary partition $\{\mathcal{T}_1, \mathcal{T}_2, \dots, \mathcal{T}_m\}$ of \mathcal{T} , let $\mathcal{T}_j = \mathcal{T}(\mathbb{I}_j^1, \mathbb{I}_j^2, \dots, \mathbb{I}_j^d)$, where $\mathbb{I}_j^k \subset \mathbb{I}^k$ for $k = 1, 2, \dots, d$ and $j = 1, 2, \dots, m$. The whole proof is divided into four steps, each one showing one bound in (6) and (7).

- (1) The lower bound of $\|\mathcal{T}\|_{p_\sigma}$ in (6).
 For any given \mathcal{T}_j , we let $\mathbf{y}^k \in \mathbb{R}^{|\mathbb{I}_j^k|}$ with $\|\mathbf{y}^k\|_p = 1$ for $k = 1, 2, \dots, d$ be an optimal solution of $\max \{ \langle \mathcal{T}_j, \mathbf{x}^1 \otimes \mathbf{x}^2 \otimes \dots \otimes \mathbf{x}^d \rangle : \|\mathbf{x}^k\|_p = 1, k = 1, 2, \dots, d \}$, i.e.,

$$\|\mathcal{T}_j\|_{p_\sigma} = \langle \mathcal{T}_j, \mathbf{y}^1 \otimes \mathbf{y}^2 \otimes \dots \otimes \mathbf{y}^d \rangle.$$

Instead of being $\{1, 2, \dots, |\mathbb{I}_j^k|\}$, the indices of \mathbf{y}^k are kept as that of \mathbb{I}_j^k for $k = 1, 2, \dots, d$. For every k , we define $\mathbf{x}^k \in \mathbb{R}^{n_k}$ where

$$x_i^k = \begin{cases} y_i^k & i \in \mathbb{I}_j^k, \\ 0 & i \in \mathbb{I}^k / \mathbb{I}_j^k. \end{cases}$$

Clearly we have $\|\mathbf{x}^k\|_p = \|\mathbf{y}^k\|_p = 1$. Therefore,

$$\|T_j\|_{p_\sigma} = \langle T_j, \mathbf{y}^1 \otimes \mathbf{y}^2 \otimes \dots \otimes \mathbf{y}^d \rangle = \langle T, \mathbf{x}^1 \otimes \mathbf{x}^2 \otimes \dots \otimes \mathbf{x}^d \rangle \leq \|T\|_{p_\sigma},$$

proving that $\max_{1 \leq j \leq m} \|T_j\|_{p_\sigma} \leq \|T\|_{p_\sigma}$.

- (2) The upper bound of $\|T\|_{p_\sigma}$ in (6).

Let $\mathbf{x}^k \in \mathbb{R}^{n_k}$ with $\|\mathbf{x}^k\|_p = 1$ for $k = 1, 2, \dots, d$ be an optimal solution of (3), i.e.,

$$\|T\|_{p_\sigma} = \langle T, \mathbf{x}^1 \otimes \mathbf{x}^2 \otimes \dots \otimes \mathbf{x}^d \rangle.$$

First, we observe that

$$\langle T_j, \mathbf{x}^1(\mathbb{1}_j^1) \otimes \mathbf{x}^2(\mathbb{1}_j^2) \otimes \dots \otimes \mathbf{x}^d(\mathbb{1}_j^d) \rangle \leq \|T_j\|_{p_\sigma} \prod_{k=1}^d \|\mathbf{x}^k(\mathbb{1}_j^k)\|_p. \tag{8}$$

It is obvious that (8) holds trivially if one of $\mathbf{x}^1(\mathbb{1}_j^1), \mathbf{x}^2(\mathbb{1}_j^2), \dots, \mathbf{x}^d(\mathbb{1}_j^d)$ is a zero vector. Otherwise, we get

$$\begin{aligned} \|T_j\|_{p_\sigma} &\geq \left\langle T_j, \frac{\mathbf{x}^1(\mathbb{1}_j^1)}{\|\mathbf{x}^1(\mathbb{1}_j^1)\|_p} \otimes \frac{\mathbf{x}^2(\mathbb{1}_j^2)}{\|\mathbf{x}^2(\mathbb{1}_j^2)\|_p} \otimes \dots \otimes \frac{\mathbf{x}^d(\mathbb{1}_j^d)}{\|\mathbf{x}^d(\mathbb{1}_j^d)\|_p} \right\rangle \\ &= \frac{1}{\prod_{k=1}^d \|\mathbf{x}^k(\mathbb{1}_j^k)\|_p} \langle T_j, \mathbf{x}^1(\mathbb{1}_j^1) \otimes \mathbf{x}^2(\mathbb{1}_j^2) \otimes \dots \otimes \mathbf{x}^d(\mathbb{1}_j^d) \rangle, \end{aligned}$$

proving that (8) holds in general. Since $\{T_1, T_2, \dots, T_m\}$ is an arbitrary partition of T , $\{\mathbb{1}_j^1 \times \mathbb{1}_j^2 \times \dots \times \mathbb{1}_j^d : j = 1, 2, \dots, m\}$ is a partition of $\{\mathbb{1}^1 \times \mathbb{1}^2 \times \dots \times \mathbb{1}^d\}$. Therefore,

$$\begin{aligned} \|T\|_{p_\sigma} &= \langle T, \mathbf{x}^1 \otimes \mathbf{x}^2 \otimes \dots \otimes \mathbf{x}^d \rangle \\ &= \langle T(\mathbb{1}^1, \mathbb{1}^2, \dots, \mathbb{1}^d), (\mathbf{x}^1 \otimes \mathbf{x}^2 \otimes \dots \otimes \mathbf{x}^d)(\mathbb{1}^1, \mathbb{1}^2, \dots, \mathbb{1}^d) \rangle \\ &= \sum_{j=1}^m \langle T(\mathbb{1}_j^1, \mathbb{1}_j^2, \dots, \mathbb{1}_j^d), (\mathbf{x}^1 \otimes \mathbf{x}^2 \otimes \dots \otimes \mathbf{x}^d)(\mathbb{1}_j^1, \mathbb{1}_j^2, \dots, \mathbb{1}_j^d) \rangle \\ &= \sum_{j=1}^m \langle T_j, \mathbf{x}^1(\mathbb{1}_j^1) \otimes \mathbf{x}^2(\mathbb{1}_j^2) \otimes \dots \otimes \mathbf{x}^d(\mathbb{1}_j^d) \rangle \\ &\leq \sum_{j=1}^m \left(\|T_j\|_{p_\sigma} \prod_{k=1}^d \|\mathbf{x}^k(\mathbb{1}_j^k)\|_p \right) \\ &\leq \left(\sum_{j=1}^m \|T_j\|_{p_\sigma}^q \right)^{\frac{1}{q}} \left(\sum_{j=1}^m \left(\prod_{k=1}^d \|\mathbf{x}^k(\mathbb{1}_j^k)\|_p \right)^p \right)^{\frac{1}{p}} \\ &= \left\| (\|T_1\|_{p_\sigma}, \|T_2\|_{p_\sigma}, \dots, \|T_m\|_{p_\sigma}) \right\|_q, \end{aligned}$$

where the first inequality is due to (8), the second inequality follows from the Hölder’s inequality, and the last equality holds due to Proposition 2.1 and

$$\begin{aligned} \sum_{j=1}^m \left(\prod_{k=1}^d \|\mathbf{x}^k(\mathbb{I}_j^k)\|_p \right)^p &= \sum_{j=1}^m \left\| \mathbf{x}^1(\mathbb{I}_j^1) \otimes \mathbf{x}^2(\mathbb{I}_j^2) \otimes \dots \otimes \mathbf{x}^d(\mathbb{I}_j^d) \right\|_p^p \\ &= \sum_{j=1}^m \left\| (\mathbf{x}^1 \otimes \mathbf{x}^2 \otimes \dots \otimes \mathbf{x}^d) (\mathbb{I}_j^1, \mathbb{I}_j^2, \dots, \mathbb{I}_j^d) \right\|_p^p \\ &= \left\| (\mathbf{x}^1 \otimes \mathbf{x}^2 \otimes \dots \otimes \mathbf{x}^d) (\mathbb{I}^1, \mathbb{I}^2, \dots, \mathbb{I}^d) \right\|_p^p \\ &= \left\| \mathbf{x}^1 \otimes \mathbf{x}^2 \otimes \dots \otimes \mathbf{x}^d \right\|_p^p \\ &= \left(\prod_{k=1}^d \|\mathbf{x}^k\|_p \right)^p \\ &= 1. \end{aligned}$$

(3) The lower bound of $\|\mathcal{T}\|_{p^*}$ in (7).

For any $\mathcal{X} \in \mathbb{R}^{n_1 \times n_2 \times \dots \times n_d}$, let $\mathcal{X}_j = \mathcal{X}(\mathbb{I}_j^1, \mathbb{I}_j^2, \dots, \mathbb{I}_j^d)$ for $j = 1, 2, \dots, m$, i.e., $\{\mathcal{X}_1, \mathcal{X}_2, \dots, \mathcal{X}_m\}$ is an arbitrary partition of \mathcal{X} . By the upper bound of (6) proved in (2), we have

$$\sum_{j=1}^m \|\mathcal{X}_j\|_{p_\sigma}^q \leq 1 \implies \|\mathcal{X}\|_{p_\sigma} \leq 1.$$

Therefore, according to the dual property in Lemma 2.5, we have

$$\|\mathcal{T}\|_{p^*} = \max_{\|\mathcal{X}\|_{p_\sigma} \leq 1} \langle \mathcal{T}, \mathcal{X} \rangle = \max_{\|\mathcal{X}\|_{p_\sigma} \leq 1} \sum_{j=1}^m \langle \mathcal{T}_j, \mathcal{X}_j \rangle \geq \max_{\sum_{j=1}^m \|\mathcal{X}_j\|_{p_\sigma}^q \leq 1} \sum_{j=1}^m \langle \mathcal{T}_j, \mathcal{X}_j \rangle. \tag{9}$$

For $j = 1, 2, \dots, m$, let $y_j = \|\mathcal{X}_j\|_{p_\sigma} \geq 0$ and further let $\mathcal{Z}_j = \frac{\mathcal{X}_j}{y_j}$ if $y_j > 0$ or $\mathcal{Z}_j = \mathcal{O}$ if $y_j = 0$. Clearly $\|\mathcal{Z}_j\|_{p_\sigma} \leq 1$ and we have

$$\sum_{j=1}^m \|\mathcal{X}_j\|_{p_\sigma}^q \leq 1 \iff \sum_{j=1}^m y_j^q \leq 1, y_j \geq 0, \|\mathcal{Z}_j\|_{p_\sigma} \leq 1, j = 1, 2, \dots, m.$$

Therefore, (9) further leads to

$$\begin{aligned}
 \|\mathcal{T}\|_{p_*} &\geq \max_{\sum_{j=1}^m y_j^q \leq 1, y_j \geq 0, \|\mathcal{Z}_j\|_{p_\sigma} \leq 1, j=1,2,\dots,m} \sum_{j=1}^m \langle \mathcal{T}_j, y_j \mathcal{Z}_j \rangle \\
 &= \sum_{\sum_{j=1}^m y_j^q \leq 1, y_j \geq 0, j=1,2,\dots,m} \max_{\|\mathcal{Z}_j\|_{p_\sigma} \leq 1, j=1,2,\dots,m} \left(\sum_{j=1}^m y_j \langle \mathcal{T}_j, \mathcal{Z}_j \rangle \right) \\
 &= \sum_{\sum_{j=1}^m y_j^q \leq 1, y_j \geq 0, j=1,2,\dots,m} \max_{\|\mathcal{Z}_j\|_{p_\sigma} \leq 1} \left(\sum_{j=1}^m y_j \max_{\|\mathcal{Z}_j\|_{p_\sigma} \leq 1} \langle \mathcal{T}_j, \mathcal{Z}_j \rangle \right) \\
 &= \sum_{\sum_{j=1}^m y_j^q \leq 1, y_j \geq 0, j=1,2,\dots,m} \sum_{j=1}^m y_j \|\mathcal{T}_j\|_{p_*} \\
 &= \left\| (\|\mathcal{T}_1\|_{p_*}, \|\mathcal{T}_2\|_{p_*}, \dots, \|\mathcal{T}_m\|_{p_*}) \right\|_p,
 \end{aligned}$$

where the second equality is due to the nonnegativity of y_j and $\max_{\|\mathcal{Z}_j\|_{p_\sigma} \leq 1} \langle \mathcal{T}_j, \mathcal{Z}_j \rangle$ for any $1 \leq j \leq m$, the third equality is due to the dual norm property, and the last equality is due to the tightness of the Hölder’s inequality.

(4) The upper bound of $\|\mathcal{T}\|_{p_*}$ in (7).

For every $j = 1, 2, \dots, m$, let $\mathcal{T}'_j \in \mathbb{R}^{n_1 \times n_2 \times \dots \times n_d}$ where

$$\left(\mathcal{T}'_j \right)_{i_1 i_2 \dots i_d} = \begin{cases} t_{i_1 i_2 \dots i_d} & (i_1, i_2, \dots, i_d) \in \mathbb{I}_j^1 \times \mathbb{I}_j^2 \times \dots \times \mathbb{I}_j^d, \\ 0 & (i_1, i_2, \dots, i_d) \notin \mathbb{I}_j^1 \times \mathbb{I}_j^2 \times \dots \times \mathbb{I}_j^d. \end{cases}$$

By applying a similar approach as we prove (1), it is not difficult to get $\|\mathcal{T}'_j\|_{p_*} = \|\mathcal{T}_j\|_{p_*}$ for any $1 \leq j \leq m$. Since $\left\{ \mathbb{I}_j^1 \times \mathbb{I}_j^2 \times \dots \times \mathbb{I}_j^d : j = 1, 2, \dots, m \right\}$ is a partition of $\left\{ \mathbb{I}^1 \times \mathbb{I}^2 \times \dots \times \mathbb{I}^d \right\}$, we have $\mathcal{T} = \sum_{j=1}^m \mathcal{T}'_j$. Therefore, by the triangle inequality, we have

$$\|\mathcal{T}\|_{p_*} = \left\| \sum_{j=1}^m \mathcal{T}'_j \right\|_{p_*} \leq \sum_{j=1}^m \|\mathcal{T}'_j\|_{p_*} = \sum_{j=1}^m \|\mathcal{T}_j\|_{p_*},$$

proving the last bound. □

Theorem 3.1 generalizes and answers affirmatively the conjecture in [14], which is for $p = 2$ and a tensor partition (a special case of arbitrary partition):

Conjecture 3.2 [14, Conjecture 3.5] *If $\{\mathcal{T}_1, \mathcal{T}_2, \dots, \mathcal{T}_m\}$ is a tensor partition of a tensor \mathcal{T} , then*

$$\begin{aligned}
 \left\| (\|\mathcal{T}_1\|_{2_\sigma}, \|\mathcal{T}_2\|_{2_\sigma}, \dots, \|\mathcal{T}_m\|_{2_\sigma}) \right\|_\infty &\leq \|\mathcal{T}\|_{2_\sigma} \leq \left\| (\|\mathcal{T}_1\|_{2_\sigma}, \|\mathcal{T}_2\|_{2_\sigma}, \dots, \|\mathcal{T}_m\|_{2_\sigma}) \right\|_2, \\
 \left\| (\|\mathcal{T}_1\|_{2_*}, \|\mathcal{T}_2\|_{2_*}, \dots, \|\mathcal{T}_m\|_{2_*}) \right\|_2 &\leq \|\mathcal{T}\|_{2_*} \leq \left\| (\|\mathcal{T}_1\|_{2_*}, \|\mathcal{T}_2\|_{2_*}, \dots, \|\mathcal{T}_m\|_{2_*}) \right\|_1.
 \end{aligned}$$

Theorem 3.1 also provides an alternative proof of a more special case which is for $p = 2$ and a regular partition (a special case of tensor partition) in [14,

Theorem 3.1], whose proof is based on mathematical induction and heavily relies on the recursive structure in the definition of a regular partition. The novelty of the proof of Theorem 3.1 lies in establishing an index system to describe arbitrary partitions. It also provides a clearer picture relating the subtensors to the original tensor.

4 Discussions and theoretical applications

The general bounds on the tensor spectral p -norm and nuclear p -norm in Theorem 3.1 provide more insights on dealing with particular tensor instances in practice. Unlike the traditional matrix unfolding technique in which one needs to unfold a tensor in a fixed way, the flexibility on arbitrary partitions of a tensor provides more tools to estimate tensor norms of given tensor data in applications. In particular, it is useful for some tensors comprised of pieces with known spectral or nuclear p -norms. Let us look into its theoretical applications and see how these bounds connect to other tensor norm bounds in the literature.

We first check the tightness of the bounds in Theorem 3.1. Given the flexibility of arbitrary partitions, it is impossible to provide a general necessary and sufficient condition for these bounds to be tight. A trial sufficient condition for all the bounds in Theorem 3.1 to be tight is that all but one of \mathcal{T}_j 's are zero tensors. The other obvious case is for $p = 1$ and $q = \infty$, under which Theorem 3.1 is reduced to

$$\begin{aligned} \|\mathcal{T}\|_{1_\sigma} &= \left\| \left(\|\mathcal{T}_1\|_{1_\sigma}, \|\mathcal{T}_2\|_{1_\sigma}, \dots, \|\mathcal{T}_m\|_{1_\sigma} \right) \right\|_\infty, \\ \|\mathcal{T}\|_{1_*} &= \left\| \left(\|\mathcal{T}_1\|_{1_*}, \|\mathcal{T}_2\|_{1_*}, \dots, \|\mathcal{T}_m\|_{1_*} \right) \right\|_1. \end{aligned}$$

These identities can also be verified by Proposition 2.6 where $\|\mathcal{T}\|_{1_\sigma} = \|\mathcal{T}\|_\infty$ and $\|\mathcal{T}\|_{1_*} = \|\mathcal{T}\|_1$.

One interesting case is for rank-one tensors, which was already observed in [14] for $p = 2$ and a regular partition.

Proposition 4.1 *If $\{\mathcal{T}_1, \mathcal{T}_2, \dots, \mathcal{T}_m\}$ is an arbitrary partition of a rank-one tensor \mathcal{T} , then*

$$\left\| \left(\|\mathcal{T}_1\|_{p_\sigma}, \|\mathcal{T}_2\|_{p_\sigma}, \dots, \|\mathcal{T}_m\|_{p_\sigma} \right) \right\|_q = \|\mathcal{T}\|_{p_\sigma} = \|\mathcal{T}\|_{q_*} = \left\| \left(\|\mathcal{T}_1\|_{q_*}, \|\mathcal{T}_2\|_{q_*}, \dots, \|\mathcal{T}_m\|_{q_*} \right) \right\|_q. \tag{10}$$

Proof Let $\mathcal{T} = (t_{i_1 i_2 \dots i_d}) \in \mathbb{R}^{n_1 \times n_2 \times \dots \times n_d}$ and $\mathcal{T}_j = \mathcal{T}(\mathbb{1}_j^1, \mathbb{1}_j^2, \dots, \mathbb{1}_j^d)$ where $\mathbb{1}_j^k \in \mathbb{1}^k$ for all k and all j . Observe that $\left\{ t_{i_1 i_2 \dots i_d} \in \mathbb{R}^{1 \times 1 \times \dots \times 1} : (i_1, i_2, \dots, i_d) \in \mathbb{1}_j^1 \times \mathbb{1}_j^2 \times \dots \times \mathbb{1}_j^d \right\}$ is an arbitrary partition of \mathcal{T}_j for every j . Noticing that any scalar $x \in \mathbb{R}$ has $\|x\|_{p_\sigma} = \|x\|_{p_*} = |x|$, by applying the upper bound of (6) for \mathcal{T} and every \mathcal{T}_j ($1 \leq j \leq m$), one has

$$\|\mathcal{T}\|_{p_\sigma} \leq \left\| \left(\|\mathcal{T}_1\|_{p_\sigma}, \|\mathcal{T}_2\|_{p_\sigma}, \dots, \|\mathcal{T}_m\|_{p_\sigma} \right) \right\|_q \leq \left(\sum_{i_1=1}^{n_1} \sum_{i_2=1}^{n_2} \dots \sum_{i_d=1}^{n_d} \|t_{i_1 i_2 \dots i_d}\|_{p_\sigma}^q \right)^{\frac{1}{q}} = \|\mathcal{T}\|_q, \tag{11}$$

and by applying the lower bound of (7) one also has

$$\|\mathcal{T}\|_q = \left(\sum_{i_1=1}^{n_1} \sum_{i_2=1}^{n_2} \dots \sum_{i_d=1}^{n_d} \|t_{i_1 i_2 \dots i_d}\|_{q_*}^q \right)^{\frac{1}{q}} \leq \left\| \left(\|\mathcal{T}_1\|_{q_*}, \|\mathcal{T}_2\|_{q_*}, \dots, \|\mathcal{T}_m\|_{q_*} \right) \right\|_q \leq \|\mathcal{T}\|_{q_*}. \tag{12}$$

On the other hand, as \mathcal{T} is rank-one, one has $\|\mathcal{T}\|_{p_\sigma} = \|\mathcal{T}\|_q = \|\mathcal{T}\|_{q_*}$ according to Proposition 2.4. By combining it with (11) and (12), we are lead to the final identity (10). □

As we see from the above discussion, both the upper and lower bounds in Theorem 3.1 can be obtained for various cases. In general, the more subtensors in an arbitrary partition, the larger gap between the lower and upper bounds for a generic tensor. In particular, if a partition has m subtensors, the largest possible gap between the lower and upper bounds can be m^q when all subtensors have the same spectral p -norm or nuclear p -norm. In an extreme though trivial case where there is only one subtensor in the partition (the original tensor itself), all the bounds become naturally tight. However, due to the curse of dimensionality and the NP-hardness to compute these norms, the larger the subtensors, the more difficulty and inaccuracy in estimating these norms.

We now discuss the main bounds in some special cases to relate existing bounds in the literature. By applying the finest partition $\mathcal{T} = \{t_{i_1 i_2 \dots i_d} \in \mathbb{R}^{1 \times 1 \times \dots \times 1} : (i_1, i_2, \dots, i_d) \in \mathbb{1} \times \mathbb{1}^2 \times \dots \times \mathbb{1}^d\}$ to Theorem 3.1, we obtain the following bounds among tensor norms.

Proposition 4.2 *For any tensor \mathcal{T} and $1 \leq p, q \leq \infty$ with $\frac{1}{p} + \frac{1}{q} = 1$,*

$$\|\mathcal{T}\|_\infty \leq \|\mathcal{T}\|_{p_\sigma} \leq \|\mathcal{T}\|_q \leq \|\mathcal{T}\|_{q_*} \leq \|\mathcal{T}\|_1. \tag{13}$$

The second inequality of (13), $\|\mathcal{T}\|_{p_\sigma} \leq \|\mathcal{T}\|_q$, is exactly the one in [23, Theorem 20], and hence it provides an alternatively proof of the upper bound of the tensor spectral p -norm. When $p = 2$, (13) also implies the bounds proposed in [4, Lemma 9.1]:

$$\frac{1}{\sqrt{\prod_{k=1}^d n_k}} \|\mathcal{T}\|_2 \leq \|\mathcal{T}\|_{2_\sigma} \leq \|\mathcal{T}\|_2 \leq \|\mathcal{T}\|_{2_*} \leq \sqrt{\prod_{k=1}^d n_k} \|\mathcal{T}\|_2.$$

Next, we apply partitions to vector fibers of \mathcal{T} to Theorem 3.1, say mode- d fibers, i.e.,

$$\mathcal{T} = \{t_{i_1 i_2 \dots i_{d-1}} \in \mathbb{R}^{n_d} : (i_1, i_2, \dots, i_{d-1}) \in \mathbb{I}^1 \times \mathbb{I}^2 \times \dots \times \mathbb{I}^{d-1}\}.$$

The bounds tighten that of (13) to the followings:

Proposition 4.3 *For any tensor \mathcal{T} and $1 \leq p, q \leq \infty$ with $\frac{1}{p} + \frac{1}{q} = 1$,*

$$\max_{i_k \in \mathbb{I}^k, k=1,2,\dots,d-1} \|t_{i_1 i_2 \dots i_{d-1}}\|_q \leq \|\mathcal{T}\|_{p_\sigma} \leq \|\mathcal{T}\|_q \leq \|\mathcal{T}\|_{q_\sigma} \leq \sum_{i_k \in \mathbb{I}^k, k=1,2,\dots,d-1} \|t_{i_1 i_2 \dots i_{d-1}}\|_q. \tag{14}$$

The first inequality of (14) is exactly the one in [23, Proposition 22]. When $p = 2$ and suppose that $n_d = \max_{1 \leq k \leq d} n_k$, the first inequality of (14) also implies the bound in [29, Corollary 4.9]:

$$\|\mathcal{T}\|_2 \leq \sqrt{\prod_{k=1}^{d-1} n_k} \|\mathcal{T}\|_{2_\sigma}.$$

This is because the largest gap between the lowest and highest bounds in (14) is $\sqrt{\prod_{k=1}^{d-1} n_k}$.

Let us now apply partitions to matrix slices and discuss their connections to matrix unfoldings. Matrix unfoldings of a tensor have been one of the main tools to study tensor computation and optimization problems, mainly due to the fact that most tensor problems are NP-hard [10] while the corresponding matrix problems are much easier. One important example is that for the tensor spectral norm and nuclear norm, both are NP-hard when the order of the tensor $d \geq 3$, while they can be computed in polynomial time for a matrix ($d = 2$). In practice, the tensor nuclear norm is widely used in tensor completion [5, 20] as a convex envelope of the tensor rank. In some literature, even the tensor nuclear norm is defined by the average nuclear norms of its matrix unfoldings, as this definition, albeit is different to the original definition, can be computed in polynomial time.

When $p = 2$, for the tensor spectral norm, the relations of a tensor and its matrix unfoldings have been studied widely, while that for the tensor nuclear norm was only addressed by Hu [13] and soon again by Friedland and Lim [4]. Wang et al. [29] studied comprehensively on the spectral p -norm based on various matrix unfoldings as well as tensor unfoldings. One obvious way to apply Theorem 3.1 is to partition a tensor into matrix slices. To make a clearer presentation, we mainly discuss third order tensors, which can be easily generalized to higher orders. Let $\mathcal{T} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$. Denote $\text{Mat}_1(\mathcal{T}) \in \mathbb{R}^{n_1 \times n_2 n_3}$, $\text{Mat}_2(\mathcal{T}) \in \mathbb{R}^{n_2 \times n_1 n_3}$, and $\text{Mat}_3(\mathcal{T}) \in \mathbb{R}^{n_3 \times n_1 n_2}$ to be the mode-1, mode-2, and mode-3 unfolding matrix of \mathcal{T} , respectively. For $k = 1, 2, 3$, denote T_i^k to be the i th mode- k matrix slice for $i = 1, 2, \dots, n_k$; see the following example.

Example 4.4 Let $\mathcal{T} = (t_{ij\ell}) \in \mathbb{R}^{2 \times 3 \times 4}$ where $i \in \{1, 2\}$, $j \in \{1, 2, 3\}$ and $\ell \in \{1, 2, 3, 4\}$, and we have

$$\begin{aligned} \text{Mat}_1(\mathcal{T}) &= \begin{pmatrix} t_{111} & t_{112} & t_{113} & t_{114} & t_{121} & t_{122} & t_{123} & t_{124} & t_{131} & t_{132} & t_{133} & t_{134} \\ t_{211} & t_{212} & t_{213} & t_{214} & t_{221} & t_{222} & t_{223} & t_{224} & t_{231} & t_{232} & t_{233} & t_{234} \end{pmatrix}, \\ \text{Mat}_2(\mathcal{T}) &= \begin{pmatrix} t_{111} & t_{112} & t_{113} & t_{114} & t_{211} & t_{212} & t_{213} & t_{214} \\ t_{121} & t_{122} & t_{123} & t_{124} & t_{221} & t_{222} & t_{223} & t_{224} \\ t_{131} & t_{132} & t_{133} & t_{134} & t_{231} & t_{232} & t_{233} & t_{234} \end{pmatrix}, \\ \text{Mat}_3(\mathcal{T}) &= \begin{pmatrix} t_{111} & t_{121} & t_{131} & t_{211} & t_{221} & t_{231} \\ t_{112} & t_{122} & t_{132} & t_{212} & t_{222} & t_{232} \\ t_{113} & t_{123} & t_{133} & t_{213} & t_{223} & t_{233} \\ t_{114} & t_{124} & t_{134} & t_{214} & t_{224} & t_{234} \end{pmatrix}, \\ T_1^1 &= \begin{pmatrix} t_{111} & t_{112} & t_{113} & t_{114} \\ t_{121} & t_{122} & t_{123} & t_{124} \\ t_{131} & t_{132} & t_{133} & t_{134} \end{pmatrix}, \\ T_2^1 &= \begin{pmatrix} t_{211} & t_{212} & t_{213} & t_{214} \\ t_{221} & t_{222} & t_{223} & t_{224} \\ t_{231} & t_{232} & t_{233} & t_{234} \end{pmatrix}. \end{aligned}$$

Let us first generalize the relations of the norms of a tensor and the norms of its matrix unfoldings, from the tensor spectral norm to the tensor spectral p -norm, and from the tensor nuclear norm [13] to the tensor nuclear p -norm.

Lemma 4.5 *If $\mathcal{T} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ and $1 \leq p \leq \infty$, then for any $\ell = 1, 2, 3$,*

$$\begin{aligned} \|\mathcal{T}\|_{p_\sigma} &\leq \|\text{Mat}_\ell(\mathcal{T})\|_{p_\sigma}, \\ \|\mathcal{T}\|_{p_*} &\geq \|\text{Mat}_\ell(\mathcal{T})\|_{p_*}. \end{aligned}$$

Proof We prove the case for $\ell = 1$ as the other two cases are similar. Let $\mathbf{x}^k \in \mathbb{R}^{n_k}$ with $\|\mathbf{x}^k\|_p = 1$ for $k = 1, 2, 3$, such that $\|\mathcal{T}\|_{p_\sigma} = \langle \mathcal{T}, \mathbf{x}^1 \otimes \mathbf{x}^2 \otimes \mathbf{x}^3 \rangle$. By Proposition 2.1, $\|\mathbf{x}^2 \otimes \mathbf{x}^3\|_p = 1$, and so $\|\text{vec}(\mathbf{x}^2 \otimes \mathbf{x}^3)\|_p = 1$, where $\text{vec}(\cdot)$ turns a tensor or a matrix to a vector. Therefore,

$$\|\mathcal{T}\|_{p_\sigma} = \langle \mathcal{T}, \mathbf{x}^1 \otimes \mathbf{x}^2 \otimes \mathbf{x}^3 \rangle = \langle \text{Mat}_1(\mathcal{T}), \mathbf{x}^1 \otimes \text{vec}(\mathbf{x}^2 \otimes \mathbf{x}^3) \rangle \leq \|\text{Mat}_1(\mathcal{T})\|_{p_\sigma}.$$

For the nuclear p -norm, let $\mathcal{T} = \sum_{i=1}^r \lambda_i \mathbf{y}_i^1 \otimes \mathbf{y}_i^2 \otimes \mathbf{y}_i^3$ with $\|\mathbf{y}_i^k\|_p = 1$ for all k and all i , such that $\|\mathcal{T}\|_{p_*} = \sum_{i=1}^r |\lambda_i|$. It is not difficulty to see that

$$\text{Mat}_1(\mathcal{T}) = \sum_{i=1}^r \lambda_i \mathbf{y}_i^1 \otimes \text{vec}(\mathbf{y}_i^2 \otimes \mathbf{y}_i^3),$$

and $\text{vec}(\mathbf{y}_i^2 \otimes \mathbf{y}_i^3) \in \mathbb{R}^{n_2 n_3}$ with $\|\text{vec}(\mathbf{y}_i^2 \otimes \mathbf{y}_i^3)\|_p = 1$ for all i . Therefore,

$$\|\text{Mat}_1(\mathcal{T})\|_{p_*} \leq \sum_{i=1}^r |\lambda_i| = \|\mathcal{T}\|_{p_*}.$$

□

Our main result in this section discusses the relations of the norms of a tensor, the norms of matrix unfoldings of the tensor, and the norms obtained by partitions to matrix slices of the tensor, as follows.

Theorem 4.6 *Let $\mathcal{T} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ and $1 \leq p, q \leq \infty$ with $\frac{1}{p} + \frac{1}{q} = 1$. For $k = 1, 2, 3$, denote*

$$\begin{aligned} \mathbf{t}_{p_\sigma}^k &= \left(\|T_1^k\|_{p_\sigma}, \|T_2^k\|_{p_\sigma}, \dots, \|T_{n_k}^k\|_{p_\sigma} \right) \in \mathbb{R}^{n_k}. \\ \mathbf{t}_{p_*}^k &= \left(\|T_1^k\|_{p_*}, \|T_2^k\|_{p_*}, \dots, \|T_{n_k}^k\|_{p_*} \right) \in \mathbb{R}^{n_k}. \end{aligned}$$

It follows that for any $k = 1, 2, 3$ and any $\ell \neq k$,

$$n_k^{-\frac{1}{q}} \|\text{Mat}_\ell(\mathcal{T})\|_{p_\sigma} \leq n_k^{-\frac{1}{q}} \|\mathbf{t}_{p_\sigma}^k\|_q \leq \|\mathbf{t}_{p_\sigma}^k\|_\infty \leq \|\mathcal{T}\|_{p_\sigma} \leq \|\text{Mat}_\ell(\mathcal{T})\|_{p_\sigma} \leq \|\mathbf{t}_{p_\sigma}^k\|_q, \tag{15}$$

$$n_k^{\frac{1}{q}} \|\text{Mat}_\ell(\mathcal{T})\|_{p_*} \geq n_k^{\frac{1}{q}} \|\mathbf{t}_{p_*}^k\|_p \geq \|\mathbf{t}_{p_*}^k\|_1 \geq \|\mathcal{T}\|_{p_*} \geq \|\text{Mat}_\ell(\mathcal{T})\|_{p_*} \geq \|\mathbf{t}_{p_*}^k\|_p. \tag{16}$$

Proof A key observation is that for $k \neq \ell$, $\{T_1^k, T_2^k, \dots, T_{n_k}^k\}$ or $\{(T_1^k)^T, (T_2^k)^T, \dots, (T_{n_k}^k)^T\}$ must be an arbitrary partition of the matrix $\text{Mat}_\ell(\mathcal{T})$ (see Example 4.4). By applying Theorem 3.1, the last inequality of (15) and the last inequality of (16) hold, and so as to the first inequality of (15) and the first inequality of (16). The fourth inequality of (15) and the fourth inequality of (16) hold by Lemma 4.5. The third inequality of (15) and the third inequality of (16) hold by Theorem 3.1. Finally, the second inequality of (15) holds by the largest gap between the L_q -norm and the L_∞ -norm of an n_k -dimensional vector, and the second inequality of (16) holds by the largest gap between the L_p -norm and the L_1 -norm of an n_k -dimensional vector. □

When $p = 2$, (16) provides tighter lower or upper bounds than that in [13, Theorem 4.4] and [4, Theorem 9.4]:

$$\begin{aligned} n_k^{-\frac{1}{2}} \|\text{Mat}_\ell(\mathcal{T})\|_{2_\sigma} &\leq \|\mathcal{T}\|_{2_\sigma} \leq \|\text{Mat}_\ell(\mathcal{T})\|_{2_\sigma}, \\ n_k^{\frac{1}{2}} \|\text{Mat}_\ell(\mathcal{T})\|_{2_*} &\geq \|\mathcal{T}\|_{2_*} \geq \|\text{Mat}_\ell(\mathcal{T})\|_{2_*}. \end{aligned}$$

In general, by Theorem 4.6, both $\|\mathbf{t}_{p_\sigma}^k\|_q$ obtained from partitions to matrix slices and $\|\text{Mat}_\ell(\mathcal{T})\|_{p_\sigma}$ obtained from matrix unfoldings, provide a bound with a factor $n_k^{\frac{1}{q}}$ for $\|\mathcal{T}\|_{p_\sigma}$. The same factor $n_k^{\frac{1}{q}}$ for $\|\mathcal{T}\|_{p_*}$ by both $\|\mathbf{t}_{p_*}^k\|_p$ from partitions to matrix slices

and $\|\text{Mat}_\ell(\mathcal{T})\|_{p_*}$ from matrix unfoldings. For the flexibility of n_k 's in Theorem 4.6, one may choose the tightest bound to be $\min_{1 \leq k \leq 3} n_k^q$. Finally, by choosing one bound from the best matrix unfolding and the other from the best partition to matrix slices would give the tightest bound of both the tensor spectral p -norm and the tensor nuclear p -norm.

It is not difficult to extend Theorem 4.6 to fourth or higher order tensors. Again, the bounds in terms of n_k 's, the dimensions of a tensor, are similarly obtained from matrix unfoldings and from partitions to matrix slices, and can be tighter by combining the two. We only list the following result to extend Theorem 4.6 to a general order, whose proof is left to interested readers.

Theorem 4.7 *Let $\mathcal{T} \in \mathbb{R}^{n_1 \times n_2 \times \dots \times n_d}$ and $1 \leq p, q \leq \infty$ with $\frac{1}{p} + \frac{1}{q} = 1$. Let $\{\mathbb{I}_1, \mathbb{I}_2\}$ be a partition of the set $\{1, 2, \dots, d\}$, and pick any $i \in \mathbb{I}_1$ and $j \in \mathbb{I}_2$. Denote $\text{Mat}(\mathcal{T})$ to be the matrix unfolding of \mathcal{T} by combining modes of \mathbb{I}_1 into the row index and modes of \mathbb{I}_2 into the column index, i.e., a $\left(\prod_{k \in \mathbb{I}_1} n_k\right) \times \left(\prod_{k \in \mathbb{I}_2} n_k\right)$ matrix. Consider the set of matrix slices of \mathcal{T} obtained by fixing all the mode- k indices except modes i and j , i.e., a set of $\prod_{1 \leq k \leq d, k \neq i, j} n_k$ number of $n_i \times n_j$ matrices. Further, denote $\mathbf{t}_{p_\sigma} \in \mathbb{R}^{\prod_{1 \leq k \leq d, k \neq i, j} n_k}$ to be the vector whose entries are the spectral p -norms of this set of matrix slices and $\mathbf{t}_{p_*} \in \mathbb{R}^{\prod_{1 \leq k \leq d, k \neq i, j} n_k}$ to be the vector whose entries are the nuclear p -norm of this set of matrix slices. It follows that*

$$\begin{aligned} \|\text{Mat}(\mathcal{T})\|_{p_\sigma} \prod_{1 \leq k \leq d, k \neq i, j} n_k^{-\frac{1}{q}} &\leq \|\mathbf{t}_{p_\sigma}\|_q \prod_{1 \leq k \leq d, k \neq i, j} n_k^{-\frac{1}{q}} \leq \|\mathbf{t}_{p_\sigma}\|_\infty \leq \|\mathcal{T}\|_{p_\sigma} \leq \|\text{Mat}(\mathcal{T})\|_{p_\sigma} \leq \|\mathbf{t}_{p_\sigma}\|_q, \\ \|\text{Mat}(\mathcal{T})\|_{p_*} \prod_{1 \leq k \leq d, k \neq i, j} n_k^{\frac{1}{q}} &\geq \|\mathbf{t}_{p_*}\|_p \prod_{1 \leq k \leq d, k \neq i, j} n_k^{\frac{1}{q}} \geq \|\mathbf{t}_{p_*}\|_1 \geq \|\mathcal{T}\|_{p_*} \geq \|\text{Mat}(\mathcal{T})\|_{p_*} \geq \|\mathbf{t}_{p_*}\|_p. \end{aligned}$$

We remark that Theorem 4.7 indicates any matrix unfolding, not necessarily having n_i rows and $\prod_{1 \leq k \leq d, k \neq i} n_k$ columns such as third order tensors. In this sense for $p = 2$, it extends the result in [13, Theorem 5.2]. Finally, we remark that one can even use the tensor unfolding technique [29] to derive more sophisticated bounds, but we do not pursue here as it involves heavy notations on the partition lattice of modes. The key point leading to all of these is the following fact: For any tensor unfolding of a tensor, there exists a partition of the original tensor, such that it is also a partition of the tensor unfolding.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Alon, N., Naor, A.: Approximating the cut-norm via Grothendieck's inequality. *SIAM J. Comput.* **35**, 787–803 (2006)
2. Candès, E.J., Recht, B.: Exact matrix completion via convex optimization. *Found. Comput. Math.* **9**, 717–772 (2009)
3. Derksen, H.: On the nuclear norm and the singular value decomposition of tensors. *Found. Comput. Math.* **16**, 779–811 (2016)
4. Friedland, S., Lim, L.-H.: Nuclear norm of higher-order tensors. *Math. Comput.* **87**, 1255–1281 (2018)
5. Gandy, S., Recht, B., Yamada, I.: Tensor completion and low- n -rank tensor recovery via convex optimization. *Inverse Probl.* **27**, 025010 (2011)
6. Golub, G.H., Van Loan, C.F.: *Matrix Computations*. Johns Hopkins University Press, Baltimore (1996)
7. He, S., Jiang, B., Li, Z., Zhang, S.: Probability bounds for polynomial functions in random variables. *Math. Oper. Res.* **39**, 889–907 (2014)
8. He, S., Li, Z., Zhang, S.: Approximation algorithms for homogeneous polynomial optimization with quadratic constraints. *Math. Program.* **125**, 353–383 (2010)
9. He, S., Li, Z., Zhang, S.: Approximation algorithms for discrete polynomial optimization. *J. Oper. Res. Soc. China* **1**, 3–36 (2013)
10. Hillar, C. J., Lim, L.-H.: Most tensor problems are NP-hard. *J. ACM* **60**, Artical 45 (2013)
11. Horn, R.A., Johnson, C.R.: *Matrix Analysis*. Cambridge University Press, New York (1985)
12. Hou, K., So, A.M.-C.: Hardness and approximation results for L_p -ball constrained homogeneous polynomial optimization problems. *Math. Oper. Res.* **39**, 1084–1108 (2014)
13. Hu, S.: Relations of the nuclear norm of a tensor and its matrix flattenings. *Linear Algebra Appl.* **478**, 188–199 (2015)
14. Li, Z.: Bounds on the spectral norm and the nuclear norm of a tensor based on tensor partitions. *SIAM J. Matrix Anal. Appl.* **37**, 1440–1452 (2016)
15. Li, Z., He, S., Zhang, S.: *Approximation Methods for Polynomial Optimization: Models, Algorithms, and Applications*. Springer, New York (2012)
16. Li, Z., Nakatsukasa, Y., Soma, T., Uschmajew, A.: On orthogonal tensors and best rank-one approximation ratio. *SIAM J. Matrix Anal. Appl.* **39**, 400–425 (2018)
17. Li, Z., Zhao, Y.-B.: On norm compression inequalities for partitioned block tensors. *Calcolo* **57**, 11 (2020)
18. Lim, L.-H.: Singular values and eigenvalues of tensors: a variational approach. In: *Proceedings of the IEEE International Workshop on Computational Advances in Multi-sensor Adaptive Processing*, vol. 1, pp. 129–132 (2005)
19. Lim, L.-H., Comon, P.: Blind multilinear identification. *IEEE Trans. Inf. Theory* **60**, 1260–1280 (2014)
20. Liu, J., Musialski, P., Wonka, P., Ye, J.: Tensor completion for estimating missing values in visual data. *IEEE Trans. Pattern Anal. Mach. Intell.* **35**, 208–220 (2013)
21. Nesterov, Y.: Global quadratic optimization via conic relaxation. In: *Wolkowicz, H., Saigal, R., Vandenberghe, L. (eds.) Handbook of Semidefinite Programming: Theory, Algorithms, and Applications*, pp. 363–387. Kluwer Academic Publishers, Boston (2000)
22. Nie, J.: Symmetric tensor nuclear norms. *SIAM J. Appl. Algebra Geom.* **1**, 599–625 (2017)
23. Nikiforov, V.: Combinatorial methods for the spectral p -norm of hypermatrices. *Linear Algebra Appl.* **529**, 324–354 (2017)
24. Qi, L.: Eigenvalues of a real supersymmetric tensor. *J. Symb. Comput.* **40**, 1302–1324 (2005)
25. Ragnarsson, S., Van Loan, C.F.: Block tensor unfoldings. *SIAM J. Matrix Anal. Appl.* **33**, 149–169 (2012)
26. So, A.M.-C.: Deterministic approximation algorithms for sphere constrained homogeneous polynomial optimization problems. *Math. Program.* **192**, 357–382 (2011)
27. Steinberg, D.: *Computation of Matrix Norms with Applications to Robust Optimization*. Master's Thesis, Technion—Israel Institute of Technology, Haifa (2005)
28. Vannieuwenhoven, N., Meerbergen, K., Vandebril, R.: Computing the gradient in optimization algorithms for the CP decomposition in constant memory through tensor blocking. *SIAM J. Sci. Comput.* **37**, C415–C438 (2015)
29. Wang, M., Dao Duc, K., Fischer, J., Song, Y.S.: Operator norm inequalities between tensor unfoldings on the partition lattice. *Linear Algebra Appl.* **520**, 44–66 (2017)

30. Yuan, M., Zhang, C.-H.: On tensor completion via nuclear norm minimization. *Found. Comput. Math.* **16**, 1031–1068 (2016)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.