# New estimations on the upper bounds for the nuclear norm of a tensor 

Xu Kong ${ }^{1 *}$ © , Jicheng $\mathrm{Li}^{2}$ and Xiaolong Wang ${ }^{3}$

"Correspondence:
xu.kong@hotmail.com
${ }^{1}$ School of Mathematical Sciences, Liaocheng University, Liaocheng, China
Full list of author information is available at the end of the article


#### Abstract

Using the orthogonal rank of the tensor, a new estimation method for the upper bounds on the nuclear norms is presented and some new tight upper bounds on the nuclear norms are established. Taking into account the structure information of the tensor, an important factor affecting the upper bounds is discussed and some corresponding properties related to the nuclear norms are given. Meanwhile, some new results of the upper bounds on the nuclear norms are obtained.

MSC: 15A18; 15A69 Keywords: Tensor; Nuclear norm; Orthogonal rank; Upper bound


## 1 Introduction

A tensor is a multidimensional array that can provide a natural and convenient way for representing the multidimensional data such as discrete forms of multivariate functions, images, video sequences, and so on $[1,2,10,14]$. With the successful and widespread use of the matrix nuclear norm (a sum of the singular values) in the information recovery, the research on the nuclear norm of the tensor (see Definition 2.3 in Sect. 2) has been a hot topic in both the theory and applications [5-8].

A natural problem is how to compute the nuclear norm of a tensor. Unfortunately, compared with the matrix nuclear norm, the nuclear norm of a tensor is closely related to the number field, and the computation of the tensor nuclear norm is NP-hard [5]. Thus, exploring some simple polynomial-time computable upper bounds on the nuclear norm is very important.

Relating to the nuclear norm of a tensor, Friedland and Lim established the following upper bound through the Frobenius norm of this tensor in [5].

Theorem 1.1 ([5]) Let $\mathcal{X} \in \mathbb{R}^{n_{1} \times \cdots \times n_{D}}$. Then

$$
\|\mathcal{X}\|_{*} \leq \sqrt{\prod_{i=1}^{D} n_{i}\|\mathcal{X}\|_{F} .}
$$

In [8], Hu established a tighter upper bound.

Theorem 1.2 (Lemma 5.1 in [8]) Let $\mathcal{X} \in \mathbb{R}^{n_{1} \times \cdots \times n_{D}}$. Then

$$
\begin{equation*}
\|\mathcal{X}\|_{*} \leq \sqrt{\frac{\prod_{i=1}^{D} n_{i}}{\max \left\{n_{1}, \ldots, n_{D}\right\}}}\|\mathcal{X}\|_{F} . \tag{1}
\end{equation*}
$$

Furthermore, Hu established another upper bound on the nuclear norm of a tensor through the nuclear norms of its unfolding matrices.

Theorem 1.3 (Theorem 5.2 in [8]) Let $\mathcal{X} \in \mathbb{R}^{n_{1} \times \cdots \times n_{D}}$. Then

$$
\begin{equation*}
\|\mathcal{X}\|_{*} \leq \sqrt{\frac{\prod_{i=2}^{D} n_{i}}{\max \left\{n_{2}, \ldots, n_{D}\right\}}}\left\|\mathbf{X}_{(1)}\right\|_{*} . \tag{2}
\end{equation*}
$$

In this paper, we present some new upper bounds on the nuclear norms through using the orthogonal rank of the tensor [9, 12]. Furthermore, taking into account the structure information of the tensor, some new results of the upper bounds on the nuclear norms are obtained.
Since the spectral norm and nuclear norm of a tensor are closely related to the number field [6], for the sake of simplicity, we always assume that the tensors discussed are nonzero real tensors, and unless mentioned otherwise, we just discuss the spectral and the nuclear norm over the real field. Some corresponding notations are as follows: a tensor is denoted by the calligraphic letter (e.g., $\mathcal{X}$ ), while the scalar is denoted by the plain letter, and matrices and vectors are denoted by bold letters (e.g., $\mathbf{X}$ and $\mathbf{x}$ ).
The rest of the paper is organized as follows. In Sect. 2, we recall some definitions and related results which are needed for the subsequent sections. In Sect. 3, we present the upper bounds on the nuclear norms of general tensors and discuss the factor affecting the upper bounds. Finally, some conclusions are made in Sect. 4.

## 2 Notations and preliminaries

This section is devoted to reviewing some conceptions and results related to tensors, which are needed for the following sections.

Firstly, we discuss the unfolding matrix or matrix representation of a tensor.
Let $\mathcal{X}=\left(x_{i_{1} \cdots i_{D}}\right) \in \mathbb{R}^{n_{1} \times \cdots \times n_{D}}$. Then, by organizing several indexes of $\mathcal{X}$ and the remaining indexes of $\mathcal{X}$ as the row index and the column index, respectively, the tensor $\mathcal{X}$ can be reshaped into a matrix form [13]. Especially, if the number of row index is equal to one, then we get the mode- $d$ matricization $\mathbf{X}_{(d)}$, the columns of which are the mode- $d$ fibers of the tensor (obtained by fixing every coordinate except one, $\left[x_{i_{1} \cdots i_{d-1} 1 i_{d+1} \cdots i_{D}}, \ldots, x_{i_{1} \cdots i_{d-1} n_{d} i_{d+1} \cdots i_{D}}\right]^{T}$ ), arranged in a cyclic ordering; see [3] for details. Contrary to the operation above, a matrix can also be reshaped into a tensor by using the opposite operation.

In what follows, we review some definitions of the tensor norms.

Definition 2.1 ([3]) Let $\mathcal{X}=\left(x_{i_{1} \cdots i_{D}}\right) \in \mathbb{R}^{n_{1} \times \cdots \times n_{D}}$, the Frobenius norm or HilbertSchmidt norm of the tensor $\mathcal{X}$ is defined as

$$
\|\mathcal{X}\|_{F}=\sqrt{\langle\mathcal{X}, \mathcal{X}\rangle}=\left(\sum_{i_{1}=1}^{n_{1}} \cdots \sum_{i_{D}=1}^{n_{D}} x_{i_{1} \cdots i_{D}}^{2}\right)^{1 / 2} .
$$

Definition 2.2 ([6]) Let " $\circ$ " denote the outer product operation. The spectral norm of $\mathcal{X} \in \mathbb{R}^{n_{1} \times \cdots \times n_{D}}$ is defined as

$$
\|\mathcal{X}\|_{2}=\max \left\{\left\langle\mathcal{X}, \mathbf{x}^{(1)} \circ \cdots \circ \mathbf{x}^{(D)}\right\rangle: \mathbf{x}^{(d)} \in \mathbb{R}^{n_{d}},\left\|\mathbf{x}^{(d)}\right\|_{2}=1,1 \leq d \leq D\right\} .
$$

Furthermore, $\|\mathcal{X}\|_{2}$ is equal to the Frobenius norm of the best rank-one approximation of the tensor $\mathcal{X}$.

Similar to the matrix case, the nuclear norm can be defined through the dual norm of the spectral norm.

Definition 2.3 ([6]) Let $\mathcal{X} \in \mathbb{R}^{n_{1} \times \cdots \times n_{D}}$, the nuclear norm of $\mathcal{X}$ is defined as the dual norm of the spectral norm. That is,

$$
\begin{equation*}
\|\mathcal{X}\|_{*}:=\max \left\{\langle\mathcal{X}, \mathcal{Y}\rangle: \mathcal{Y} \in \mathbb{R}^{n_{1} \times \cdots \times n_{D}},\|\mathcal{Y}\|_{2}=1\right\} \tag{3}
\end{equation*}
$$

For the nuclear norm defined by (3), it can be shown that

$$
\begin{aligned}
\|\mathcal{X}\|_{*}= & \min \left\{\sum_{p=1}^{P}\left|\lambda_{p}\right|: \mathcal{X}=\sum_{p=1}^{P} \lambda_{p} \mathbf{x}_{p}^{(1)} \circ \cdots \circ \mathbf{x}_{p}^{(D)},\right. \\
& \left.\left\|\mathbf{x}_{p}^{(d)}\right\|_{2}=1, \mathbf{x}_{p}^{(d)} \in \mathbb{R}^{n_{d}}, \lambda_{p} \in \mathbb{R}, P \in \mathbb{N}\right\} .
\end{aligned}
$$

Another important concept related to the matrix and the tensor is the mode- $d$ multiplication.

Definition 2.4 ([3]) Let $\mathcal{X}=\left(x_{i_{1} \cdots i_{D}}\right) \in \mathbb{R}^{n_{1 \times} \times \cdots \times n_{D}}$. Then the mode- $d$ multiplication of $\mathcal{X}$ by the matrix $\mathbf{U}=\left(u_{i j}\right) \in \mathbb{R}^{n_{d}^{\prime} \times n_{d}}$ is defined by

$$
\left(\mathcal{X} \times{ }_{d} \mathbf{U}\right)_{i_{1} \cdots i_{d-1} i_{d}^{\prime} i_{d+1} \cdots i_{D}}=\sum_{i_{d}=1}^{n_{d}} x_{i_{1} \cdots i_{d-1} i_{d} i_{d+1} \cdots i_{D}} u_{i_{d}^{\prime} i_{d}}, \quad 1 \leq d \leq D .
$$

It should be mentioned that the mode- $d$ multiplication is also available for $n_{d}^{\prime}=1$.
Furthermore, let

$$
\left(\mathbf{W}^{(1)}, \ldots, \mathbf{W}^{(D)}\right) \cdot \mathcal{X}=\mathcal{X} \times_{1} \mathbf{W}^{(1)} \times \cdots \times_{D} \mathbf{W}^{(D)}
$$

If for all $1 \leq d \leq D$ the matrices $\mathbf{W}^{(d)}$ are orthogonal matrices $\left(\mathbf{W}^{(d)} \mathbf{W}^{(d)^{T}}\right.$ is an identity matrix), then $\left(\mathbf{W}^{(1)}, \ldots, \mathbf{W}^{(D)}\right) \cdot \mathcal{X}$ is called a multi-linear orthogonal transformation of the tensor $\mathcal{X}$.
Finally in this section, we introduce the tool used in the paper for the estimation of the upper bounds.

Definition 2.5 ([9]) The orthogonal rank of $\mathcal{X} \in \mathbb{R}^{n_{1} \times \cdots n_{D}}$ is defined as the smallest number $R$ such that

$$
\begin{equation*}
\mathcal{X}=\sum_{r=1}^{R} \mathcal{U}_{r} \tag{4}
\end{equation*}
$$

where $\mathcal{U}_{r}(1 \leq r \leq R)$ are rank-one tensors such that $\left\langle\mathcal{U}_{r_{1}}, \mathcal{U}_{r_{2}}\right\rangle=0,\left(r_{1} \neq r_{2}\right)$ for $1 \leq r_{1} \leq R$ and $1 \leq r_{2} \leq R$.

The decomposition of $\mathcal{X}$ given by (4) is also called the orthogonal decomposition of the tensor $\mathcal{X}$.

For the orthogonal rank of a tensor, the following conclusion is true.
Theorem 2.1 ([11]) Let $n_{1} \leq \cdots \leq n_{D}$. Then, for any $\mathcal{X} \in \mathbb{R}^{n_{1} \times \cdots \times n_{D}}$, it holds

$$
r_{\perp}(\mathcal{X}) \leq \prod_{i=1}^{D-1} n_{i}
$$

where $r_{\perp}(\mathcal{X})$ denotes the orthogonal rank of $\mathcal{X}$.
Noting the fact that indices relabeling does not change the tensor orthogonal rank, then by Theorem 2.1, we have that, for any $\mathcal{X} \in \mathbb{R}^{n_{1} \times \cdots \times n_{D}}$, it holds

$$
\begin{equation*}
r_{\perp}(\mathcal{X}) \leq \frac{\prod_{i=1}^{D} n_{i}}{\max \left\{n_{1}, \ldots, n_{D}\right\}} \tag{5}
\end{equation*}
$$

Especially, for the third order tensor, the following result was established in [11].
Lemma 2.1 ([11]) Let $n \geq 2$. Then, for any $\mathcal{X} \in \mathbb{R}^{n \times n \times 2}$, the following holds:

$$
r_{\perp}(\mathcal{X}) \leq \begin{cases}2 n-1, & \text { if } n \text { is odd } ; \\ 2 n, & \text { if } n \text { is even } .\end{cases}
$$

## 3 Upper bounds of the nuclear norm

In this section, we discuss the upper bounds on the nuclear norm of a tensor. Meanwhile, some properties and polynomial-time computable bounds related to the nuclear norm will be given.

### 3.1 Upper bounds given by the Frobenius norm

In this subsection, we use the orthogonal rank of a general tensor to establish the upper bounds on the nuclear norm through the Frobenius norm of this tensor.

Theorem 3.1 Let $\mathcal{X} \in \mathbb{R}^{n_{1} \times \cdots \times n_{D}}$. Suppose that

$$
R=\max _{\mathcal{Y} \in \mathbb{R}^{n_{1} \times \cdots \times n_{D}}}\left\{\operatorname{rank}_{\perp}(\mathcal{Y})\right\} .
$$

Then

$$
\begin{equation*}
\|\mathcal{X}\|_{*} \leq \sqrt{R}\|\mathcal{X}\|_{F} . \tag{6}
\end{equation*}
$$

Proof Let $\mathcal{Y} \in \mathbb{R}^{n_{1} \times \cdots \times n_{D}}$ be an arbitrary nonzero tensor and the orthogonal rank of $\mathcal{Y}$ be $R_{y}$. Suppose that

$$
\mathcal{Y}=\sum_{r=1}^{R_{y}} \mathcal{U}_{r}
$$

is the orthogonal rank decomposition of the tensor $\mathcal{Y}$, where $R_{y} \leq R$.

Then, according to the properties of the orthogonal rank decomposition and the best rank-one approximation, we have

$$
\begin{equation*}
\|\mathcal{Y}\|_{2} \geq \max _{1 \leq r \leq R_{y}}\left\{\left\|\mathcal{U}_{r}\right\|_{F}\right\} \tag{7}
\end{equation*}
$$

and

$$
\begin{equation*}
\|\mathcal{Y}\|_{F}^{2}=\sum_{r=1}^{R_{y}}\left\|\mathcal{U}_{r}\right\|_{F}^{2} \tag{8}
\end{equation*}
$$

Without loss of generality, suppose that

$$
\left\|\mathcal{U}_{1}\right\|_{F}=\max _{1 \leq r \leq R_{y}}\left\{\left\|\mathcal{U}_{r}\right\|_{F}\right\} .
$$

Then it follows from (7) and (8) that

$$
\begin{align*}
\left\langle\mathcal{X}, \frac{\mathcal{Y}}{\|\mathcal{Y}\|_{2}}\right\rangle & \leq\|\mathcal{X}\|_{F} \frac{\|\mathcal{Y}\|_{F}}{\|\mathcal{Y}\|_{2}} \\
& \leq\|\mathcal{X}\|_{F} \frac{\sqrt{\sum_{r=1}^{R_{y}}\left\|\mathcal{U}_{r}\right\|_{F}^{2}}}{\left\|\mathcal{U}_{1}\right\|_{F}} \\
& \leq\|\mathcal{X}\|_{F} \frac{\sqrt{R_{y}\left\|\mathcal{U}_{1}\right\|_{F}^{2}}}{\left\|\mathcal{U}_{1}\right\|_{F}} \\
& =\sqrt{R_{y}}\|\mathcal{X}\|_{F} \\
& \leq \sqrt{R}\|\mathcal{X}\|_{F} . \tag{9}
\end{align*}
$$

Thus, according to the arbitrariness of $\mathcal{Y}$ and (9), we get

$$
\max _{\mathcal{Y} \in \mathbb{R}^{n_{1} \times \cdots \times n_{D}}}\left\{\left\langle\mathcal{X}, \frac{\mathcal{Y}}{\|\mathcal{Y}\|_{2}}\right\rangle\right\} \leq \sqrt{R}\|\mathcal{X}\|_{F}
$$

Noting the definition of the nuclear norm (Definition 2.3), the conclusion is established.

Remark 3.1 Comparing the upper bound given by (6) with the upper bound given by (1), which is obtained in [8], the new upper bound given by (6) is tighter.

Actually, it follows from (5) that the upper bound given by (6) improves the upper bound given by (1).
More specifically, we present a simple example to show that the upper bound given by Theorem 3.1 not only can be tighter than the upper bound given by (1) but also a sharp bound.

Example 3.1 Let

$$
\mathcal{A}=\left[\begin{array}{ccc|ccc}
0 & 1 & 0 & -1 & 0 & 0 \\
1 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 1
\end{array}\right] \in \mathbb{R}^{3 \times 3 \times 2}
$$

By Theorem 3.1 and Lemma 2.1, we get

$$
\|\mathcal{A}\|_{*} \leq \sqrt{5} \sqrt{1^{2}+1^{2}+(-1)^{2}+1^{2}+1^{2}}=5<\sqrt{\frac{3 \times 3 \times 2}{3}} \sqrt{5}=\sqrt{30} .
$$

This means that the upper bound given by (3.1) is tighter than the upper bound given by (1).

Furthermore, by a simple computation, we get $\|\mathcal{A}\|_{2}=1$. Then it follows from the definition of the nuclear norm that

$$
\|\mathcal{A}\|_{*} \geq\left\langle\mathcal{A}, \frac{\mathcal{A}}{\|\mathcal{A}\|_{2}}\right\rangle=\frac{\|\mathcal{A}\|_{F}^{2}}{\|\mathcal{A}\|_{2}}=\frac{5}{1}=5
$$

Thus, it holds

$$
\|\mathcal{A}\|_{*}=5
$$

Actually,

$$
\begin{align*}
\mathcal{A}= & \mathbf{e}_{1 ; 3} \circ \mathbf{e}_{2 ; 3} \circ \mathbf{e}_{1 ; 2}+\mathbf{e}_{2 ; 3} \circ \mathbf{e}_{1 ; 3} \circ \mathbf{e}_{1 ; 2}+\left(-\mathbf{e}_{1 ; 3}\right) \circ \mathbf{e}_{1 ; 3} \circ \mathbf{e}_{2 ; 2}+\mathbf{e}_{2 ; 3} \circ \mathbf{e}_{2 ; 3} \circ \mathbf{e}_{2 ; 2} \\
& +\mathbf{e}_{3 ; 3} \circ \mathbf{e}_{3 ; 3} \circ \mathbf{e}_{2 ; 2} \tag{10}
\end{align*}
$$

is a nuclear decomposition of $\mathcal{A}$, where $\mathbf{e}_{1 ; 3}=[1,0,0]^{T}, \mathbf{e}_{2 ; 3}=[0,1,0]^{T}, \mathbf{e}_{3 ; 3}=[0,0,1]^{T}$, $\mathbf{e}_{1 ; 2}=[1,0]^{T}$, and $\mathbf{e}_{2 ; 2}=[0,1]^{T}$. Since

$$
\begin{aligned}
\left\langle\mathcal{A}, \mathbf{e}_{1 ; 3} \circ \mathbf{e}_{2 ; 3} \circ \mathbf{e}_{1 ; 2}\right\rangle & =\left\langle\mathcal{A}, \mathbf{e}_{2 ; 3} \circ \mathbf{e}_{1 ; 3} \circ \mathbf{e}_{1 ; 2}\right\rangle=\left\langle\mathcal{A},\left(-\mathbf{e}_{1 ; 3}\right) \circ \mathbf{e}_{1 ; 3} \circ \mathbf{e}_{2 ; 2}\right\rangle \\
& =\left\langle\mathcal{A}, \mathbf{e}_{2 ; 3} \circ \mathbf{e}_{2 ; 3} \circ \mathbf{e}_{2 ; 2}\right\rangle=\left\langle\mathcal{A}, \mathbf{e}_{3 ; 3} \circ \mathbf{e}_{3 ; 3} \circ \mathbf{e}_{2 ; 2}\right\rangle,
\end{aligned}
$$

then, according to the sufficient and necessary conditions of the nuclear norm decomposition obtained in [6], we get that (10) is a nuclear decomposition of $\mathcal{A}$.
This also means that the upper bound given by Theorem 3.1 is a sharp upper bound of the nuclear norm.

### 3.2 Upper bounds given by nuclear norms of the unfolding matrices of a tensor

 In this subsection, we present a new way to establish the upper bounds on the nuclear norm of a tensor through the nuclear norms of the unfolding matrices of this tensor.Theorem 3.2 Let $\mathcal{X} \in \mathbb{R}^{n_{1} \times \cdots \times n_{D}}$. Suppose that

$$
\tilde{R}=\max _{\mathcal{Y} \in \mathbb{R}^{n_{2} \times \cdots \times n_{D}}}\left\{\operatorname{rank}_{\perp}(\mathcal{Y})\right\} .
$$

Then

$$
\begin{equation*}
\|\mathcal{X}\|_{*} \leq \sqrt{\tilde{R}}\left\|\mathbf{X}_{(1)}\right\|_{*} . \tag{11}
\end{equation*}
$$

Proof Let the singular value decomposition of the matrix $\mathbf{X}_{(1)}$ be

$$
\begin{equation*}
\mathbf{X}_{(1)}=\sigma_{1} \mathbf{u}_{1} \mathbf{v}_{1}^{T}+\cdots+\sigma_{S} \mathbf{u}_{S} \mathbf{v}_{S}^{T}, \tag{12}
\end{equation*}
$$

where $\mathbf{u}_{s} \in \mathbb{R}^{n_{1}}, \mathbf{v}_{s} \in \mathbb{R}^{n_{2} \times \cdots \times n_{D}},\left\|\mathbf{u}_{s}\right\|_{2}=1,\left\|\mathbf{v}_{s}\right\|_{2}=1, \sigma_{s}>0$, and $1 \leq s \leq S$.

Then equality (12) can be expressed as the following form:

$$
\begin{equation*}
\mathcal{X}=\sigma_{1} \mathbf{u}_{1} \circ \mathcal{V}_{1}+\cdots+\sigma_{S} \mathbf{u}_{S} \circ \mathcal{V}_{S} \tag{13}
\end{equation*}
$$

where $\mathcal{V}_{s} \in \mathbb{R}^{n_{2} \times \cdots \times n_{D}}$ are obtained by reordering the vector $\mathbf{v}_{s}$ into $(D-1)$ th order tensor with a certain order, $1 \leq s \leq S$. Suppose that the orthogonal rank decomposition of $\mathcal{V}_{s}$ is

$$
\mathcal{V}_{s}=\mathbf{v}_{1}^{(1, s)} \circ \cdots \circ \mathbf{v}_{1}^{(D-1, s)}+\cdots+\mathbf{v}_{P_{s}}^{(1, s)} \circ \cdots \circ \mathbf{v}_{P_{s}}^{(D-1, s)}, \quad 1 \leq s \leq S
$$

Then, by taking the expression of $\mathcal{V}_{s}$ into the right-hand side of (13), we get

$$
\begin{align*}
\mathcal{X}= & \sigma_{1} \sum_{i=1}^{P_{1}} \mathbf{u}_{1} \circ \mathbf{v}_{i}^{(1,1)} \circ \cdots \circ \mathbf{v}_{i}^{(D-1,1)}+\cdots+\sigma_{S} \sum_{i=1}^{P_{S}} \mathbf{u}_{1} \circ \mathbf{v}_{i}^{(1, S)} \circ \cdots \circ \mathbf{v}_{i}^{(D-1, S)} \\
= & \sigma_{1} \sum_{i=1}^{P_{1}}\left(\left\|\mathbf{v}_{i}^{(1,1)}\right\|_{2} \cdots\left\|\mathbf{v}_{i}^{(D-1,1)}\right\|_{2}\right) \mathbf{u}_{1} \circ \frac{\mathbf{v}_{i}^{(1,1)}}{\left\|\mathbf{v}_{i}^{(1,1)}\right\|_{2}} \circ \cdots \circ \frac{\mathbf{v}_{i}^{(D-1,1)}}{\left\|\mathbf{v}_{i}^{(D-1,1)}\right\|_{2}}+\cdots \\
& +\sigma_{S} \sum_{i=1}^{P_{S}}\left(\left\|\mathbf{v}_{i}^{(1, s)}\right\|_{2} \cdots\left\|\mathbf{v}_{i}^{(D-1, s)}\right\|_{2}\right) \mathbf{u}_{1} \circ \frac{\mathbf{v}_{i}^{(1, S)}}{\left\|\mathbf{v}_{i}^{(1, S)}\right\|_{2}} \circ \cdots \circ \frac{\mathbf{v}_{i}^{(D-1, S)}}{\left\|\mathbf{v}_{i}^{(D-1, S)}\right\|_{2}} \tag{14}
\end{align*}
$$

Noting that

$$
\begin{equation*}
\left\|\mathcal{V}_{s}\right\|_{F}^{2}=\sum_{i=1}^{P_{s}}\left\|\mathbf{v}_{i}^{(1, s)} \circ \cdots \circ \mathbf{v}_{i}^{(D-1, s)}\right\|_{F}^{2}=\sum_{i=1}^{P_{s}}\left\|\mathbf{v}_{i}^{(1, s)}\right\|_{F}^{2} \cdots\left\|\mathbf{v}_{i}^{(D-1, s)}\right\|_{F}^{2}=\left\|\mathbf{v}_{s}\right\|_{F}^{2}=1 \tag{15}
\end{equation*}
$$

where $P_{s} \leq \tilde{R}, 1 \leq s \leq S$, then it follows from the definition of the nuclear norm and (14) that

$$
\begin{aligned}
\|\mathcal{X}\|_{*} & \leq \sigma_{1} \sum_{i=1}^{P_{1}}\left\|\mathbf{v}_{i}^{(1,1)}\right\|_{2} \cdots\left\|\mathbf{v}_{i}^{(D-1,1)}\right\|_{2}+\cdots+\sigma_{S} \sum_{i=1}^{P_{S}}\left\|\mathbf{v}_{i}^{(1, s)}\right\|_{2} \cdots\left\|\mathbf{v}_{i}^{(D-1, s)}\right\|_{2} \\
& \leq \sigma_{1} \sqrt{P_{1}}+\cdots+\sigma_{S} \sqrt{P_{S}} \quad(\text { by }(15) \text { and Cauchy-Schwarz inequality }) \\
& \leq \sigma_{1} \sqrt{\tilde{R}}+\cdots+\sigma_{S} \sqrt{\tilde{R}} \\
& =\sqrt{\tilde{R}}\left\|\mathbf{X}_{(1)}\right\|_{*}
\end{aligned}
$$

Remark 3.2 Comparing the upper bound given by (11) with the upper bound given by (2), which is obtained in [8], the new upper bound given by (11) is smaller.

Actually, it follows from inequality (5) that the upper bound given by (11) improves the upper bound given by (2).
Similar to the discussion of Hu [8], the upper bounds can also be obtained by other unfolding ways and further improved by considering the multi-linear ranks of a tensor (ranks of the unfolding matrices).

Corollary 3.1 Let $\mathcal{X} \in \mathbb{R}^{n_{1} \times n_{2} \times n_{3}}$ and $r_{d}=\operatorname{rank}\left(\mathbf{X}_{(d)}\right), 1 \leq d \leq 3$. Then

$$
\|\mathcal{X}\|_{*} \leq \frac{\sqrt{\min \left\{r_{2}, r_{3}\right\}}\left\|\mathbf{X}_{(1)}\right\|_{*}+\sqrt{\min \left\{r_{3}, r_{1}\right\}}\left\|\mathbf{X}_{(2)}\right\|_{*}+\sqrt{\min \left\{r_{1}, r_{2}\right\}}\left\|\mathbf{X}_{(3)}\right\|_{*}}{3}
$$

Proof According to the conditions of the corollary and the higher order singular value decomposition of the tensor [3], the tensor $\mathcal{X}$ can be expressed as

$$
\mathcal{X}=\left(\mathbf{W}^{(1)}, \mathbf{W}^{(2)}, \mathbf{W}^{(3)}\right) \cdot \tilde{\mathcal{X}}
$$

where $\tilde{\mathcal{X}} \in \mathbb{R}^{r_{1} \times r_{2} \times r_{3}}$ and $\mathbf{W}^{(d)} \in \mathbb{R}^{n_{d} \times r_{d}}$ satisfying that $\mathbf{W}^{(d)^{T}} \mathbf{W}^{(d)}$ is an identity matrix for all $1 \leq d \leq 3$.
By the definition of the tensor nuclear norm (Definition 2.3), one can easily verify the following conclusions:

$$
\begin{equation*}
\|\tilde{\mathcal{X}}\|_{*}=\|\mathcal{X}\|_{*} \tag{16}
\end{equation*}
$$

and

$$
\begin{equation*}
\left\|\tilde{\mathbf{X}}_{(1)}\right\|_{*}=\left\|\mathbf{X}_{(1)}\right\|_{*} . \tag{17}
\end{equation*}
$$

It follows from Theorem 3.2 and (17) that

$$
\|\tilde{\mathcal{X}}\|_{*} \leq \sqrt{\max _{\mathcal{Y} \in \mathbb{R}^{r_{2} \times r_{3}}}\left\{\operatorname{rank}_{\perp}(\mathcal{Y})\right\}}\left\|\tilde{\mathbf{X}}_{(1)}\right\|_{*} \leq \sqrt{\min \left\{r_{2}, r_{3}\right\}}\left\|\mathbf{X}_{(1)}\right\|_{*} .
$$

Noting (16), we get

$$
\|\mathcal{X}\|_{*} \leq \sqrt{\min \left\{r_{2}, r_{3}\right\}}\left\|\mathbf{X}_{(1)}\right\|_{*} .
$$

Similarly, we have

$$
\|\mathcal{X}\|_{*} \leq \sqrt{\min \left\{r_{3}, r_{1}\right\}}\left\|\mathbf{X}_{(2)}\right\|_{*},
$$

and

$$
\|\mathcal{X}\|_{*} \leq \sqrt{\min \left\{r_{1}, r_{2}\right\}}\left\|\mathbf{X}_{(3)}\right\|_{*} .
$$

Thus, the conclusion is obtained.

### 3.3 Factors affecting the upper bounds on the nuclear norm and further results

 In this subsection, we discuss the factors affecting the nuclear norm of a tensor. Especially, we focus on the structure analysis of a tensor. Based on the discussion, some new upper bounds on the nuclear norms of tensors are presented.Firstly, we give a simple example to illustrate that the nuclear norm of a tensor is closely related to the structure of this tensor.

Example 3.2 Let

$$
\mathcal{A}=\left[\begin{array}{cc|cc}
0 & 1 & -1 & 0  \tag{18}\\
1 & 0 & 0 & 1
\end{array}\right]
$$

Similar to the discussion of Example 3.1, since

$$
\mathcal{A}=\mathbf{e}_{1 ; 2} \circ \mathbf{e}_{2 ; 2} \circ \mathbf{e}_{1 ; 2}+\mathbf{e}_{2 ; 2} \circ \mathbf{e}_{1 ; 2} \circ \mathbf{e}_{1 ; 2}+\left(-\mathbf{e}_{1 ; 2}\right) \circ \mathbf{e}_{1 ; 2} \circ \mathbf{e}_{2 ; 2}+\mathbf{e}_{2 ; 2} \circ \mathbf{e}_{2 ; 2} \circ \mathbf{e}_{2 ; 2}
$$

and

$$
\begin{aligned}
\left\langle\mathcal{A}, \mathbf{e}_{1 ; 2} \circ \mathbf{e}_{2 ; 2} \circ \mathbf{e}_{1 ; 2}\right\rangle & =\left\langle\mathcal{A}, \mathbf{e}_{2 ; 2} \circ \mathbf{e}_{1 ; 2} \circ \mathbf{e}_{1 ; 2}\right\rangle=\left\langle\mathcal{A},\left(-\mathbf{e}_{1 ; 2}\right) \circ \mathbf{e}_{1 ; 2} \circ \mathbf{e}_{2 ; 2}\right\rangle \\
& =\left\langle\mathcal{A}, \mathbf{e}_{2 ; 2} \circ \mathbf{e}_{2 ; 2} \circ \mathbf{e}_{2 ; 2}\right\rangle,
\end{aligned}
$$

then, according to the sufficient and necessary conditions of the nuclear norm decomposition obtained in [6], we get

$$
\|\mathcal{A}\|_{*}=4
$$

It is well known that the nuclear norm of a tensor is closely related to the number field [6]. Actually, the tensor $\mathcal{A}$ can be expressed as the following form:

$$
\mathcal{A}=\frac{1}{2}\left[\begin{array}{l}
-1 \\
-i
\end{array}\right] \circ\left[\begin{array}{l}
1 \\
i
\end{array}\right] \circ\left[\begin{array}{l}
i \\
1
\end{array}\right]+\frac{1}{2}\left[\begin{array}{c}
-1 \\
i
\end{array}\right] \circ\left[\begin{array}{c}
1 \\
-i
\end{array}\right] \circ\left[\begin{array}{c}
-i \\
1
\end{array}\right] .
$$

Let

$$
\mathbf{W}^{(1)}=\frac{1}{\sqrt{2}}\left[\begin{array}{cc}
-1 & i \\
-1 & -i
\end{array}\right], \quad \mathbf{W}^{(2)}=\frac{1}{\sqrt{2}}\left[\begin{array}{cc}
1 & -i \\
1 & i
\end{array}\right], \quad \mathbf{W}^{(3)}=\frac{1}{\sqrt{2}}\left[\begin{array}{cc}
-i & 1 \\
i & 1
\end{array}\right] .
$$

Then it holds

$$
\left(\mathbf{W}^{(1)}, \mathbf{W}^{(2)}, \mathbf{W}^{(3)}\right) \cdot \mathcal{A}=\left[\begin{array}{cc|cc}
\sqrt{2} & 0 & 0 & 0  \tag{19}\\
0 & 0 & 0 & \sqrt{2}
\end{array}\right] .
$$

For the sake of convenience, let $\hat{\mathcal{A}}=\left(\mathbf{W}^{(1)}, \mathbf{W}^{(2)}, \mathbf{W}^{(3)}\right) \cdot \mathcal{A}$. Then, using the same method as above, we have

$$
\left\langle\hat{\mathcal{A}}, \mathbf{e}_{1 ; 2} \circ \mathbf{e}_{1 ; 2} \circ \mathbf{e}_{1 ; 2}\right\rangle=\left\langle\hat{\mathcal{A}}, \mathbf{e}_{2 ; 2} \circ \mathbf{e}_{2 ; 2} \circ \mathbf{e}_{2 ; 2}\right\rangle
$$

Thus $\|\hat{\mathcal{A}}\|_{*}=2 \sqrt{2}$. Since all three matrices $\mathbf{W}^{(k)}(1 \leq k \leq 3)$ are unitary matrices, based on the invariance of the Frobenius norm of a tensor under the multi-linear orthogonal transformations, we get

$$
\|\mathcal{A}\|_{*}=\|\hat{\mathcal{A}}\|_{*}=2 \sqrt{2}
$$

Noting the structures of the tensors given by (18) and (19), the above derivation process implies that the nuclear norm of a tensor is closely related to the structure of this tensor. In what follows, we discuss the block diagonal tensor, which can be illustrated by Fig. 1. Furthermore, the block diagonal tensor can be expressed by using the direct sum operation " $\oplus$ " [4], which is defined as follows:

Figure 1 The block diagonal tensor with three diagonal blocks


Let $\mathcal{A}=\left(a_{i_{1} \cdots i_{D}}\right) \in \mathbb{R}^{n_{1} \times \cdots \times n_{D}}$ and $\mathcal{B}=\left(b_{j_{1} \cdots j_{D}}\right) \in \mathbb{R}^{n_{1}^{\prime} \times \cdots \times n_{D}^{\prime}}$, then the direct sum of $\mathcal{A}$ and $\mathcal{B}$ is an order- $D$ tensor $\mathcal{C}=\left(c_{i_{1} \cdots i_{D}}\right)=\mathcal{A} \oplus \mathcal{B} \in \mathbb{R}^{\left(n_{1}+n_{1}^{\prime}\right) \times \cdots \times\left(n_{D}+n_{D}^{\prime}\right)}$ defined by

$$
c_{i_{1} \cdots i_{D}}= \begin{cases}a_{i_{1} \cdots i_{D}}, & \text { if } 1 \leq i_{\alpha} \leq n_{\alpha}, \alpha=1,2, \ldots, D \\ b_{i_{1}-n_{1}, \ldots, i_{D}-n_{D}}, & \text { if } n_{\alpha}+1 \leq i_{\alpha} \leq n_{\alpha}+n_{\alpha}^{\prime}, \alpha=1,2, \ldots, D \\ 0, & \text { otherwise }\end{cases}
$$

Based on the discussion above, we present some properties of the spectral norm and nuclear norm of the tensor.

Lemma 3.1 Let $\mathcal{X}^{(l)} \in \mathbb{R}^{n_{1}^{(l)} \times \cdots \times n_{D}^{(l)}}, 1 \leq l \leq L$, and

$$
\mathcal{X}=\mathcal{X}^{(1)} \oplus \cdots \oplus \mathcal{X}^{(L)} \in \mathbb{R}^{\left(\sum_{l=1}^{L} n_{1}^{(l)}\right) \times \cdots \times\left(\sum_{l=1}^{L} n_{D}^{(l)}\right.} .
$$

Then

$$
\begin{equation*}
\|\mathcal{X}\|_{2}=\max _{1 \leq l \leq L}\left\{\left\|\mathcal{X}^{(l)}\right\|_{2}\right\} . \tag{20}
\end{equation*}
$$

Proof According to the definition of the spectral norm of a tensor (Definition 2.3), it is easy to get

$$
\|\mathcal{X}\|_{2} \geq \max _{1 \leq l \leq L}\left\{\left\|\mathcal{X}^{(l)}\right\|_{2}\right\} .
$$

Thus, the rest of the proof just needs to show

$$
\begin{equation*}
\|\mathcal{X}\|_{2} \leq \max _{1 \leq l \leq L}\left\{\left\|\mathcal{X}^{(l)}\right\|_{2}\right\} \tag{21}
\end{equation*}
$$

Firstly, we consider the case of the third order tensors.
Suppose that $\mathcal{X}^{(l)} \in \mathbb{R}^{n_{1}^{(l)} \times n_{2}^{(l)} \times n_{3}^{(l)}}(1 \leq l \leq L)$,

$$
\mathcal{X}=\mathcal{X}^{(1)} \oplus \cdots \oplus \mathcal{X}^{(L)} \in \mathbb{R}^{\left(\sum_{l=1}^{L} n_{1}^{(l)}\right) \times\left(\sum_{l=1}^{L} n_{2}^{(l)}\right) \times\left(\sum_{l=1}^{L} n_{3}^{(l)}\right)},
$$

and $\sigma \mathbf{u} \circ \mathbf{v} \circ \mathbf{w}$ is the best rank-one approximation of $\mathcal{X}$, where $\sigma=\|\mathcal{X}\|_{2}, \mathbf{u} \in \mathbb{R}^{\sum_{l=1}^{L} n_{1}^{(l)}}, \mathbf{v} \in$ $\mathbb{R}^{\sum_{l=1}^{L} n_{2}^{(l)}}, \mathbf{w}=\left[\mathbf{w}_{1}^{T}, \ldots, \mathbf{w}_{L}^{T}\right]^{T} \in \mathbb{R}^{\sum_{l=1}^{L} n_{3}^{(l)}}, \mathbf{w}_{l} \in \mathbb{R}_{3}^{n_{3}^{(l)}}(1 \leq l \leq L)$, and $\|\mathbf{u}\|_{2}=\|\mathbf{v}\|_{2}=\|\mathbf{w}\|_{2}=1$.

Then the following matrix

$$
\mathcal{X} \times{ }_{3} \mathbf{w}^{T}=\left[\begin{array}{lll}
\mathcal{X}^{(1)} \times_{3} \mathbf{w}_{1}^{T} & & \\
& \ddots & \\
& & \mathcal{X}^{(L)} \times_{3} \mathbf{w}_{L}^{T}
\end{array}\right]
$$

is a block diagonal matrix. It follows

$$
\|\mathcal{X}\|_{2}=\left\|\mathcal{X} \times_{3} \mathbf{w}^{T}\right\|_{2}=\max _{1 \leq l \leq L}\left\{\left\|\mathcal{X}^{(l)} \times_{3} \mathbf{w}_{l}^{T}\right\|_{2}\right\} \leq \max _{1 \leq l \leq L}\left\{\left\|\mathcal{X}^{(l)}\right\|_{2}\right\} .
$$

Hence, inequality (21) is proved. This also implies that equality (20) is true for the third order tensors.
Secondly, for the case of higher order tensors with order larger than or equal to four, the same result can be established by the recursive method.

In all, the conclusion is true.

Then, based on Lemma 3.1, the following two results related to the nuclear norms of tensors can be established.

Lemma 3.2 Let $\mathcal{O} \in \mathbb{R}^{n_{1}^{(1)} \times \cdots \times n_{D}^{(1)}}$ be a zero tensor, and $\mathcal{X} \in \mathbb{R}^{n_{1}^{(2)} \times \cdots \times n_{D}^{(2)}}$. Then

$$
\|\mathcal{O} \oplus \mathcal{X}\|_{*}=\|\mathcal{X} \oplus \mathcal{O}\|_{*}=\|\mathcal{X}\|_{*} .
$$

Proof Suppose that

$$
\|\mathcal{X}\|_{*}=\sum_{p=1}^{P}\left|\sigma_{p}\right|
$$

and

$$
\mathcal{X}=\sum_{p=1}^{P} \sigma_{p} \mathbf{x}_{p}^{(1)} \circ \cdots \circ \mathbf{x}_{p}^{(D)}
$$

where for all $1 \leq p \leq P, \mathbf{x}_{p}^{(1)} \circ \cdots \circ \mathbf{x}_{p}^{(D)}$ are rank-one tensors with $\left\|\mathbf{x}_{p}^{(1)}\right\|_{2}=\cdots=\left\|\mathbf{x}_{p}^{(D)}\right\|_{2}=1$. Then

$$
\mathcal{O} \oplus \mathcal{X}=\sum_{p=1}^{P} \sigma_{p}\binom{\mathbf{x}_{p}^{(1)}}{\mathbf{0}} \circ \cdots \circ\binom{\mathbf{x}_{p}^{(D)}}{\mathbf{0}},
$$

where all $0 s$ denote zero vectors with suitable dimensions. This implies

$$
\|\mathcal{O} \oplus \mathcal{X}\|_{*} \leq\|\mathcal{X}\|_{*} .
$$

Furthermore, assume that

$$
\|\mathcal{X}\|_{*}=\langle\mathcal{X}, \mathcal{Y}\rangle
$$

where $\mathcal{Y} \in \mathbb{R}^{n_{1}^{(2)} \times \cdots \times n_{D}^{(2)}}$ and $\|\mathcal{Y}\|_{2}=1$.

Then, by Lemma 3.1, we have

$$
\|\mathcal{O} \oplus \mathcal{Y}\|_{2}=\|\mathcal{Y}\|_{2}=1
$$

It follows from Definition 2.5 that

$$
\|\mathcal{O} \oplus \mathcal{X}\|_{*} \geq\langle\mathcal{O} \oplus \mathcal{X}, \mathcal{O} \oplus \mathcal{Y}\rangle=\langle\mathcal{X}, \mathcal{Y}\rangle=\|\mathcal{X}\|_{*} .
$$

Thus it holds $\|\mathcal{O} \oplus \mathcal{X}\|_{*}=\|\mathcal{X}\|_{*}$.
Using the same method, the equality $\|\mathcal{X} \oplus \mathcal{O}\|_{*}=\|\mathcal{X}\|_{*}$ can be proved.
Lemma 3.3 Let $\mathcal{X}^{(l)} \in \mathbb{R}^{n_{1}^{(l)} \times \cdots \times n_{D}^{(l)}}, 1 \leq l \leq L$, and

$$
\mathcal{X}=\mathcal{X}^{(1)} \oplus \cdots \oplus \mathcal{X}^{(L)} \in \mathbb{R}^{\left(\sum_{l=1}^{L} n_{1}^{(l)}\right) \times \cdots \times\left(\sum_{l=1}^{L} n_{D}^{(l)}\right)} .
$$

Then

$$
\|\mathcal{X}\|_{*}=\sum_{l=1}^{L}\left\|\mathcal{X}^{(l)}\right\|_{*}
$$

Proof We just need to prove the case of $L=2$. For the general case, the conclusion can be obtained in a recursive way.

Let

$$
\tilde{\mathcal{X}}_{1}=\mathcal{X}^{(1)} \oplus \mathcal{O}^{(1)}, \quad \tilde{\mathcal{X}}_{2}=\mathcal{O}^{(2)} \oplus \mathcal{X}^{(1)}
$$

where $\mathcal{O}^{(1)} \in \mathbb{R}^{n_{1}^{(2)} \times \cdots \times n_{D}^{(2)}}$ and $\mathcal{O}^{(2)} \in \mathbb{R}^{n_{1}^{(1)} \times \cdots \times n_{D}^{(1)}}$ are both zero tensors.
Then, by using Lemma 3.2, we get

$$
\begin{equation*}
\|\mathcal{X}\|_{*}=\left\|\tilde{\mathcal{X}}_{1}+\tilde{\mathcal{X}}_{2}\right\|_{*} \leq\left\|\tilde{\mathcal{X}}_{1}\right\|_{*}+\left\|\tilde{\mathcal{X}}_{2}\right\|_{*}=\left\|\mathcal{X}^{(1)}\right\|_{*}+\left\|\mathcal{X}^{(2)}\right\|_{*} \tag{22}
\end{equation*}
$$

Suppose that

$$
\left\|\mathcal{X}^{(l)}\right\|_{*}=\left\langle\mathcal{X}^{(l)}, \mathcal{Y}^{(l)}\right\rangle, \quad \text { and } \quad \mathcal{Y}^{(l)} \in \mathbb{R}^{n_{1}^{(l)} \times \cdots \times n_{D}^{(l)}},\left\|\mathcal{Y}^{(l)}\right\|_{2}=1 .
$$

Then, by Lemma 3.1, we get

$$
\left\|\mathcal{Y}^{(1)} \oplus \mathcal{Y}^{(2)}\right\|_{2}=1 .
$$

Thus, according to Definition 2.5, we have

$$
\begin{align*}
\|\mathcal{X}\|_{*} & =\max _{\substack{\mathcal{Y} \in \mathbb{R}^{\left(n_{1}^{(1)}+n_{1}^{(2)}\right) \times \cdots \times\left(n_{D}^{(1)}+n_{D}^{(2)}\right)} \\
\|\mathcal{Y}\|_{2}=1}}\{\langle\mathcal{X}, \mathcal{Y}\rangle\} \\
& \geq\left\langle\mathcal{X}^{(1)} \oplus \mathcal{X}^{(2)}, \mathcal{Y}^{(1)} \oplus \mathcal{Y}^{(2)}\right\rangle \\
& =\left\|\mathcal{X}^{(1)}\right\|_{*}+\left\|\mathcal{X}^{(2)}\right\|_{*} . \tag{23}
\end{align*}
$$

Combined (22) with (23), the results can be obtained.

Based on the fact that the nuclear norm of a tensor is also kept invariant under the multi-linear orthogonal transformation, we get the following result.

Corollary 3.2 Let $\mathcal{X} \in \mathbb{R}^{n_{1} \times \cdots \times n_{D}}$. If the tensor $\mathcal{X}$ admits a diagonal structure under the multi-linear orthogonal transformations, then

$$
\|\mathcal{X}\|_{*}=\sum_{p=1}^{P}\left|\sigma_{p}\right|,
$$

where $\sigma_{p}(1 \leq p \leq P)$ are the diagonal elements and $P \leq n_{1}$.

This case presented by Corollary 3.2 is consistent with the definition of the nuclear norm of the matrix case, and in this case, the nuclear norm of the tensor can be accurately calculated.

Taking into account the structure information of the tensor, some new results of the upper bounds on the nuclear norms can be obtained. For the convenience of comparison, we just present the upper bounds on the nuclear norms of tensors through the dimensions of the tensors, without considering the orthogonal rank of the tensors.

Theorem 3.3 Let $\mathcal{X} \in \mathbb{R}^{n_{1} \times \cdots \times n_{D}}$ and $L$ be the maximum number of diagonal blocks that the tensor $\mathcal{X}$ can attain under the multi-linear orthogonal transformations. Suppose that the size of each diagonal block is $n_{1}^{(l)} \times \cdots \times n_{D}^{(l)}$ and

$$
\tilde{n}_{l}=\frac{\prod_{i=1}^{D} n_{i}^{(l)}}{\max \left\{n_{1}^{(l)} \cdots n_{D}^{(l)}\right\}}, \quad 1 \leq l \leq L .
$$

Then it holds

$$
\begin{equation*}
\|\mathcal{X}\|_{*} \leq \sqrt{\sum_{l=1}^{L} \tilde{n}_{l}\|\mathcal{X}\|_{F} . . . . . . . .} \tag{24}
\end{equation*}
$$

Proof Assume that

$$
\mathcal{X}=\mathcal{D}(\mathcal{X}) \times_{1} \mathbf{W}^{(1)} \cdots \times_{D} \mathbf{W}^{(D)},
$$

where

$$
\mathcal{D}(\mathcal{X})=\mathcal{D}^{(1)} \oplus \cdots \oplus \mathcal{D}^{(L)}
$$

and $\mathbf{W}^{(d)} \in \mathbb{R}^{n_{d} \times n_{d}}(1 \leq d \leq D)$ are orthogonal matrices, $\mathcal{D}^{(l)} \in \mathbb{R}_{1}^{n_{1}^{(l)} \times \cdots \times n_{D}^{(l)}}, 1 \leq l \leq L$.
Then it follows from the invariance of the Frobenius norm of a tensor under the multilinear orthogonal transformation that

$$
\begin{equation*}
\|\mathcal{X}\|_{F}^{2}=\sum_{l=1}^{L}\left\|\mathcal{D}^{(l)}\right\|_{F}^{2} \tag{25}
\end{equation*}
$$

Furthermore, since the nuclear norm of a tensor is also kept invariant under the multilinear orthogonal transformation, we get

$$
\|\mathcal{X}\|_{*}=\|\mathcal{D}(\mathcal{X})\|_{*} .
$$

Hence, by Lemma 3.3 and (25), we get

$$
\begin{aligned}
\|\mathcal{X}\|_{*} & =\|\mathcal{D}(\mathcal{X})\|_{*} \\
& =\sum_{l=1}^{L}\left\|\mathcal{D}^{(l)}\right\|_{*} \\
& \leq \sum_{l=1}^{L} \sqrt{\tilde{n}_{l}}\left\|\mathcal{D}^{(l)}\right\|_{F} \\
& \leq \sqrt{\sum_{l=1}^{L} \tilde{n}_{l}} \sqrt{\sum_{l=1}^{L}\left\|\mathcal{D}^{(l)}\right\|_{F}^{2}} \quad \text { (by Cauchy-Schwarz inequality) } \\
& =\sqrt{\sum_{l=1}^{L} \tilde{n}_{l}\|\mathcal{X}\|_{F}}
\end{aligned}
$$

Without loss of generality, suppose that

$$
n_{D}=\max \left\{n_{1}, \ldots, n_{D}\right\} .
$$

Since

$$
\begin{aligned}
& n_{1}=\sum_{l=1}^{L} n_{1}^{(l)} \\
& \vdots \\
& n_{D-1}=\sum_{l=1}^{L} n_{D-1}^{(l)}
\end{aligned}
$$

it is easy to get

$$
\begin{aligned}
\frac{\prod_{i=1}^{D} n_{i}}{\max \left\{n_{1}, \ldots, n_{D}\right\}} & =\prod_{i=1}^{D-1} n_{i} \\
& =\left(\sum_{l=1}^{L} n_{1}^{(l)}\right) \cdots\left(\sum_{l=1}^{L} n_{D-1}^{(l)}\right) \\
& \geq \sum_{l=1}^{L} \tilde{n}_{l} .
\end{aligned}
$$

Thus, the upper bound given by (24) improves (1). Theorem 3.3 also shows that the upper bound on the nuclear norm can be improved by using the structural information.
Similarly, the following upper bound can also be obtained.

Theorem 3.4 Let $\mathcal{X} \in \mathbb{R}^{n_{1} \times \cdots \times n_{D}}$, and $L$ be the maximum number of diagonal blocks that the tensor $\mathcal{X}$ can attain under the multi-linear orthogonal transformations. Suppose that the size of each diagonal block is $n_{1}^{(l)} \times \cdots \times n_{D}^{(l)}$, and

$$
\tilde{n}_{l}=\frac{\prod_{i=2}^{D} n_{i}^{(l)}}{\max \left\{n_{2}^{(l)} \cdots n_{D}^{(l)}\right\}}, \quad 1 \leq l \leq L,
$$

and

$$
\tilde{n}=\max _{1 \leq l \leq L}\left\{\tilde{n}_{l}\right\} .
$$

Then it holds

$$
\begin{equation*}
\|\mathcal{X}\|_{*} \leq \sqrt{\tilde{n}}\left\|\mathbf{X}_{(1)}\right\|_{*} . \tag{26}
\end{equation*}
$$

Proof Similar to the proof of Theorem 3.3, assume

$$
\mathcal{X}=\mathcal{D}(\mathcal{X}) \times_{1} \mathbf{W}^{(1)} \cdots \times_{D} \mathbf{W}^{(D)}
$$

where

$$
\mathcal{D}(\mathcal{X})=\mathcal{D}^{(1)} \oplus \cdots \oplus \mathcal{D}^{(L)}
$$

and $\mathbf{W}^{(d)} \in \mathbb{R}^{n_{d} \times n_{d}}(1 \leq d \leq D)$ are orthogonal matrices.
Then it holds

$$
\begin{aligned}
\|\mathcal{X}\|_{*} & =\|\mathcal{D}(\mathcal{X})\|_{*} \\
& =\sum_{l=1}^{L}\left\|\mathcal{D}^{(l)}\right\|_{*} \\
& \leq \sum_{l=1}^{L} \sqrt{\tilde{n}_{l}}\left\|\mathbf{D}_{(1)}^{(l)}\right\|_{*} \\
& \leq \sqrt{\tilde{n}} \sum_{l=1}^{L}\left\|\mathbf{D}_{(1)}^{(l)}\right\|_{*} \\
& \leq \sqrt{\tilde{n}} \sum_{l=1}^{L}\left\|\mathbf{D}_{(1)}\right\|_{*} \\
& =\sqrt{\tilde{n}}\left\|\mathbf{X}_{(1)}\right\|_{*} .
\end{aligned}
$$

Theorem 3.4 implies that the upper bound given by nuclear norms of the unfolding matrices is more closely related to the structure. For the sake of clarity, we give a simple example to illustrate.

Example 3.3 Let $\mathcal{A}$ be defined in Example 3.2, and

$$
\begin{aligned}
\mathcal{B} & =\mathcal{A} \oplus \mathcal{A} \\
& =\left[\begin{array}{llll|cccc|cccc|cccc}
0 & 1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & -1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1
\end{array}\right] .
\end{aligned}
$$

Then, by Theorem 1.2, we get

$$
\|\mathcal{B}\|_{*} \leq \sqrt{4 * 4}\|\mathcal{B}\|_{F}=8 \sqrt{2} .
$$

It follows from Theorem 3.4 that

$$
\|\mathcal{B}\|_{*} \leq \sqrt{2 * 2}\|\mathcal{B}\|_{F}=4 \sqrt{2}
$$

There has been a marked improvement in the upper bounds on the nuclear norm.

## 4 Conclusions

In this paper, we provide a new estimation method for the upper bounds on the nuclear norms and obtain some new upper bounds related to the nuclear norms. Meanwhile, it is found that the upper bounds on the nuclear norms are not only related to the dimensions of the tensor but also to the structure of the tensor. Taking into consideration the structure information of the tensor, the upper bounds on the nuclear norms can be improved.

## Funding

This research work was supported by the Natural Science Foundation of China (NSFC) (Nos: 11401286, 11671318, 11401472), Natural Science Foundation of Shaanxi Province (No: 2014JM1029), and Scientific Research Foundation of Liaocheng University.

## Competing interests

All three authors declare that they have no competing interests.

## Authors' contributions

All three authors contributed equally to this work. All authors read and approved the final manuscript.

## Author details

'School of Mathematical Sciences, Liaocheng University, Liaocheng, China. ${ }^{2}$ School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an, China. ${ }^{3}$ School of Science, Northwestern Polytechnical University, Xi'an, China.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
Received: 11 July 2018 Accepted: 20 September 2018 Published online: 11 October 2018

## References

1. Che, M.L., Cichocki, A., Wei, Y.M.: Neural networks for computing best rank-one approximations of tensors and its applications. Neurocomputing 267, 114-133 (2017)
2. Cichocki, A., Mandicand, D., Pan, A.-H., Caiafa, C., Zhou, G., Zhao, Q., De Lathauwer, L.: Tensor decompositions for signal processing applications: from two-way to multiway component analysis. IEEE Signal Process. Mag. 32(2) 145-163 (2015)
3. De Lathauwer, L., De Moor, B., Vandewalle, J.: A multilinear singular value decomposition. SIAM J. Matrix Anal. Appl. 21(4), 1253-1278 (2000)
4. De Silva, V., Lim, L.H.: Tensor rank and the ill-posedness of the best low-rank approximation problem. SIAM J. Matrix Anal. Appl. 30(3), 1084-1127 (2008)
5. Friedland, S., Lim, L.-H.: The computational complexity of duality. SIAM J. Optim. 26(4), 2378-2393 (2016)
6. Friedland, S., Lim, L.H.: Nuclear norm of higher-order tensors. Math. Comput. 87(311), 1255-1281 (2018)
7. Gandy, S., Recht, B., Yamada, I.: Tensor completion and low-n-rank tensor recovery via convex optimization. Inverse Probl. 27(2), 025010 (2011)
8. Hu, S.L.: Relations of the nuclear norm of a tensor and its matrix flattenings. Linear Algebra Appl. 478, 188-199 (2015)
9. Kolda, T.G., Bader, B.W.: Orthogonal tensor decomposition. SIAM J. Matrix Anal. Appl. 23, 243-255 (2001)
10. Kolda, T.G., Bader, B.W.: Tensor decompositions and applications. SIAM Rev. 51(3), 455-500 (2009)
11. Kong, X., Meng, D.Y: The bounds for the best rank-1 approximation ratio of a finite dimensional tensor space. Pac. J. Optim. 11, 323-337 (2015)
12. Li, Z., Nakatsukasa, Y., Soma, T., Uschmajew, A.: On orthogonal tensors and best rank-one approximation ratio. SIAM J. Matrix Anal. Appl. 39(1), 400-425 (2018)
13. Oseledets, I.V., Tyrtyshnikov, E.E.: Breaking the curse of dimensionality, or how to use SVD in many dimensions. SIAM J. Sci. Comput. 31(5), 3744-3759 (2009)
14. Qi, L., Luo, Z.: Tensor Analysis: Spectral Theory and Special Tensor. SIAM, Philadelphia (2017)

## Submit your manuscript to a SpringerOpen ${ }^{\ominus}$ journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

