SIAM J. MATRIX ANAL. APPL.
Vol. 38, No. 4, pp. 1434–1453

ON THE LARGEST MULTILINEAR SINGULAR VALUES OF HIGHER-ORDER TENSORS*

IGNAT DOMANOV†, ALWIN STEGEMAN†, AND LIEVEN DE LATHAUWER†

Abstract. Let σ_n denote the largest mode-n multilinear singular value of an $I_1 \times \cdots \times I_N$ tensor $\mathcal{T}$. We prove that $\sigma_1^2 + \cdots + \sigma_{n-1}^2 + \sigma_{n+1}^2 + \cdots + \sigma_N^2 \leq (N-2)\|\mathcal{T}\|^2 + \sigma_n^2,\ n = 1, \ldots, N$, where $\|\cdot\|$ denotes the Frobenius norm. We also show that at least for third-order cubic tensors the inverse problem always has a solution. Namely, for each σ_1, σ_2, and σ_3 that satisfy $\sigma_1^2 + \sigma_2^2 \leq \|\mathcal{T}\|^2 + \sigma_3^2$, $\sigma_1^2 + \sigma_3^2 \leq \|\mathcal{T}\|^2 + \sigma_2^2$, $\sigma_2^2 + \sigma_3^2 \leq \|\mathcal{T}\|^2 + \sigma_1^2$, and the trivial inequalities $\sigma_1 \geq \frac{1}{\sqrt{n}}\|\mathcal{T}\|$, $\sigma_2 \geq \frac{1}{\sqrt{n}}\|\mathcal{T}\|$, $\sigma_3 \geq \frac{1}{\sqrt{n}}\|\mathcal{T}\|$, there always exists an $n \times n \times n$ tensor whose largest multilinear singular values are equal to σ_1, σ_2, and σ_3. We also show that if the equality $\sigma_1^2 + \sigma_2^2 = \|\mathcal{T}\|^2 + \sigma_3^2$ holds, then $\mathcal{T}$ is necessarily equal to a sum of multilinear rank-$(L_1, 1, L_1)$ and multilinear rank-$(1, L_2, L_2)$ tensors and we give a complete description of all its multilinear singular values. We establish a connection with honeycombs and eigenvalues of the sum of two Hermitian matrices. This seems to give at least a partial explanation of why results on the joint distribution of multilinear singular values are scarce.

Key words. multilinear singular value decomposition, multilinear rank, singular value decomposition, tensor

AMS subject classifications. 15A69, 15A23

DOI. 10.1137/16M110770X

1. Introduction.

Throughout the paper $\|\cdot\|$ denotes the Frobenius norm of a vector, matrix, or tensor and the superscripts ^{T}, ^{H}, and * denote transpose, hermitian transpose, and conjugation, respectively. We also use the "empty sum/product" convention, i.e., if $m > n$, then $\sum_m^n (\cdot) = 0$ and $\prod_m^n (\cdot) = 1$.

Let $\mathcal{T} \in \mathbb{C}^{I_1 \times \cdots \times I_N}$. A *mode-$n$ fiber* of $\mathcal{T}$ is a column vector obtained by fixing indices $i_1, \ldots, i_{n-1}, i_{n+1}, \ldots, i_N$. A matrix $\mathbf{T}_{(n)} \in \mathbb{C}^{I_n \times I_1 \cdots I_{n-1} I_{n+1} \cdots I_N}$ formed by all mode-n fibers is called a *mode-n matrix unfolding* (aka flattening or matricization) of $\mathcal{T}$. For notational convenience we assume that the columns of $\mathbf{T}_{(n)}$ are ordered such that

$$\text{(1)} \quad \text{the } \left(i_n, 1 + \sum_{\substack{k=1\\k\neq n}}^{N} (i_k - 1) \prod_{\substack{l=1\\l\neq n}}^{k-1} I_l \right) \text{th entry of } \mathbf{T}_{(n)} = \text{ the } (i_1, \ldots, i_N)\text{th entry of } \mathcal{T}.$$

*Received by the editors December 13, 2016; accepted for publication (in revised form) by D. B. Szyld June 20, 2017; published electronically November 28, 2017. This paper reflects only the authors' views and the European Union is not liable for any use that may be made of the contained information.

http://www.siam.org/journals/simax/38-4/M110770.html

Funding: The work of the authors was funded by The Research Council KU Leuven, C1 project c16/15/059-nD; F.W.O. project G.0830.14N, G.0881.14N; the Belgian Federal Science Policy Office, IUAP P7 (DYSCO II, Dynamical systems, control and optimization, 2012–2017); and the EU. The research leading to these results received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013)/ERC Advanced Grant, BIOTENSORS (339804).

†Group Science, Engineering and Technology, KU Leuven–Kulak, 8500 Kortrijk, Belgium, and Department of Electrical Engineering ESAT/STADIUS KU Leuven, B-3001 Leuven-Heverlee, Belgium (ignat.domanov@kuleuven.be, alwin.stegeman@kuleuven.be, lieven.delathauwer@kuleuven.be).

For instance, if $N = 3$, i.e., $\mathcal{T} \in \mathbb{C}^{I_1 \times I_2 \times I_3}$, then (1) implies that

$$\begin{aligned}\mathbf{T}_{(1)} &= [\mathbf{T}_1 \ \dots \ \mathbf{T}_{I_3}] \in \mathbb{C}^{I_1 \times I_2 I_3},\\ \mathbf{T}_{(2)} &= [\mathbf{T}_1^T \ \dots \ \mathbf{T}_{I_3}^T] \in \mathbb{C}^{I_2 \times I_1 I_3},\\ \mathbf{T}_{(3)} &= [\operatorname{vec}(\mathbf{T}_1) \ \dots \operatorname{vec}(\mathbf{T}_{I_3})]^T \in \mathbb{C}^{I_3 \times I_1 I_2},\end{aligned}$$

where $\mathbf{T}_1, \dots, \mathbf{T}_{I_3} \in \mathbb{C}^{I_1 \times I_2}$ denote the frontal slices of $\mathcal{T}$.

Tensor $\mathcal{T} \in \mathbb{C}^{I_1 \times \cdots \times I_N}$ is *all-orthogonal* if the matrices $\mathbf{T}_{(1)}\mathbf{T}_{(1)}^H, \dots, \mathbf{T}_{(N)}\mathbf{T}_{(N)}^H$ are diagonal. *The multiLinear (ML) singular value decomposition (SVD)* (aka higher-order SVD) is a factorization of $\mathcal{T}$ into the product of an all-orthogonal tensor $\mathcal{S} \in \mathbb{C}^{I_1 \times \cdots \times I_N}$ and N unitary matrices $\mathbf{U}_1 \in \mathbb{C}^{I_1 \times I_1}, \dots, \mathbf{U}_N \in \mathbb{C}^{I_N \times I_N}$,

$$\mathcal{T} = \mathcal{S} \cdot_1 \mathbf{U}_1 \cdot_2 \mathbf{U}_2 \dots \cdot_N \mathbf{U}_N, \tag{2}$$

where "$\cdot_n$" denotes the n-mode product of $\mathcal{S}$ and $\mathbf{U}_n$. Rather than giving the formal definition of "$\cdot_n$", for which we refer the reader to [2, 3, 12], we present N equivalent matricized versions of (2):

$$\mathbf{T}_{(n)} = \mathbf{U}_n \mathbf{S}_{(n)} (\mathbf{U}_N \otimes \cdots \otimes \mathbf{U}_{n+1} \otimes \mathbf{U}_{n-1} \otimes \cdots \otimes \mathbf{U}_1)^T, \qquad n = 1, \dots, N, \tag{3}$$

where "$\otimes$" denotes the Kronecker product. For $N = 2$, i.e., for $\mathcal{T} = \mathbf{T}_1 \in \mathbb{C}^{I_1 \times I_2}$, the MLSVD reduces, up to trivial indeterminacies, to the classical SVD of a matrix, $\mathbf{T}_{(1)} = \mathbf{T}_1 = \mathbf{U}\mathbf{S}\mathbf{V}^H$, where $\mathbf{U} = \mathbf{U}_1$, $\mathbf{S} = \mathbf{S}_{(1)}$, and $\mathbf{V} = \mathbf{U}_2^* \otimes 1$. It is known [3] that MLSVD always exists and that its uniqueness properties are similar to those of the matrix SVD.

The MLSVD has many applications in signal processing, data analysis, and machine learning (see, for instance, the overview papers [12, subsection 4.4], [16]). Here we just mention that as principal component analysis (PCA) can be done by SVD of a data matrix, MLPCA can be done by MLSVD of a data tensor [4, 14, 17].

The singular values of $\mathbf{T}_{(n)}$, are called *the mode-n singular values* of $\mathcal{T}$. Since $\mathbf{S}_{(1)}\mathbf{S}_{(1)}^H, \dots, \mathbf{S}_{(N)}\mathbf{S}_{(N)}^H$ are diagonal, it follows from (3) that the ML singular values of $\mathcal{T}$ coincide with the ML singular values of $\mathcal{S}$, which are just the Frobenius norms of the rows of $\mathbf{S}_{(1)}, \dots, \mathbf{S}_{(N)}$. Throughout the paper,

$$\sigma_n \ \text{ denotes the largest singular value of } \mathbf{T}_{(n)}.$$

In the matrix case, i.e., for $N = 2$, the description of MLSVD is trivial. Indeed, the singular values of $\mathbf{T}_{(1)} = \mathbf{T}_1$ and $\mathbf{T}_{(2)} = \mathbf{T}_1^T$ coincide and $\mathbf{T}_{(3)} = \operatorname{vec}(\mathbf{T}_1)^T$ has a single singular value $\|\mathcal{T}\|$. Thus, the singular values of $\mathbf{T}_{(1)}$ completely define the singular values of $\mathbf{T}_{(2)}$ and $\mathbf{T}_{(3)}$. In particular, the set of triplets $(\sigma_1, \sigma_2, \sigma_3)$ coincides with the set $\{(x, x, y) : \ y \geq x \geq 0\} \subset \mathbb{R}^3$ whose Lebesgue measure is zero. The situation for tensors is much more complicated. It is clear that in the general case $N \geq 2$, the sets of the mode-$1, \dots,$ mode-N singular values are not independent either. The study of topological properties of the set of ML singular values of real tensors has been initiated only recently in [7] and [6]. In particular, it has been shown in [7] and [6] that, as in the matrix case, some configurations of ML singular values are not possible but, nevertheless, at least for $n \times \cdots \times n$ tensors the set of ML singular values has a positive Lebesgue measure.

In this paper we study possible configurations for the largest ML singular values, i.e., for $\sigma_1, \dots, \sigma_N$. Our results are valid for real and complex tensors. The following theorem presents simple necessary conditions for σ_1, σ_2, and σ_3 to be the largest ML singular values of a third-order tensor. For instance, it implies that a norm-1 tensor whose largest ML singular values are equal to 0.9, 0.9, and 0.7 does not exist.

THEOREM 1.1. *Let σ_1, σ_2, and σ_3 denote the largest ML singular values of an $I_1 \times I_2 \times I_3$ tensor $\mathcal{T}$. Then*

$$\sigma_1^2 + \sigma_2^2 \leq \|\mathcal{T}\|^2 + \sigma_3^2, \quad \sigma_1^2 + \sigma_3^2 \leq \|\mathcal{T}\|^2 + \sigma_2^2, \quad \sigma_2^2 + \sigma_3^2 \leq \|\mathcal{T}\|^2 + \sigma_1^2, \tag{4}$$

$$\sigma_1 \geq \frac{1}{\sqrt{I_1}}\|\mathcal{T}\|, \ \sigma_2 \geq \frac{1}{\sqrt{I_2}}\|\mathcal{T}\|, \ \sigma_3 \geq \frac{1}{\sqrt{I_3}}\|\mathcal{T}\|. \tag{5}$$

Figure 1 shows four typical shapes of the set $\{(\sigma_1^2, \sigma_2^2, \sigma_3^2) : \ \sigma_1, \ \sigma_2, \ \sigma_3$ satisfy (4)–(5)$\}$ (without loss of generality, we assumed that $I_1 \leq I_2 \leq I_3$).

One can easily verify that if σ_1, σ_2 and σ_3 satisfy (4)–(5) for $I_1 = I_2 = I_3 = 2$ and $\|\mathcal{T}\| = 1$, then σ_1, σ_2, and σ_3 are the largest ML singular values of the $2 \times 2 \times 2$ tensor $\mathcal{T}$ with mode-1 matrix unfolding

$$\mathbf{T}_{(1)} = [\mathbf{T}_1 \ \mathbf{T}_2] = \begin{bmatrix} \frac{\sqrt{\sigma_1^2+\sigma_2^2+\sigma_3^2-1}}{\sqrt{2}} & 0 & 0 & \frac{\sqrt{1+\sigma_1^2-\sigma_2^2-\sigma_3^2}}{\sqrt{2}} \\ 0 & \frac{\sqrt{1+\sigma_3^2-\sigma_1^2-\sigma_2^2}}{\sqrt{2}} & \frac{\sqrt{1+\sigma_2^2-\sigma_1^2-\sigma_3^2}}{\sqrt{2}} & 0 \end{bmatrix}.$$

The proof of the following result relies on a similar explicit construction of an $I_1 \times I_2 \times I_3$ tensor $\mathcal{T}$.

THEOREM 1.2. *Let $I_1 \leq I_2 \leq I_3$ and σ_1, σ_2, σ_3 satisfy* (4) *and the following three inequalities:*

$$\sigma_1 \geq \frac{1}{\sqrt{I_1}}\|\mathcal{T}\|, \tag{6}$$

$$(I_2 - I_1)\sigma_1^2 + (I_1 I_2 - I_2)\sigma_3^2 + (1 - I_2)\|\mathcal{T}\|^2 \geq 0, \tag{7}$$

$$(I_2 - I_1)\sigma_1^2 + (I_1 I_2 - I_2)\sigma_2^2 + (1 - I_2)\|\mathcal{T}\|^2 \geq 0. \tag{8}$$

Then there exists an $I_1 \times I_2 \times I_3$ tensor $\mathcal{T}$ such that

1. *all entries of $\mathcal{T}$ are nonnegative;*
2. *$\mathcal{T}$ is all-orthogonal;*
3. *the largest ML singular values of $\mathcal{T}$ are equal to σ_1, σ_2, and σ_3.*

Conditions (5) and (6)–(8) mean that the point $(\sigma_1^2, \sigma_2^2, \sigma_3^2)$ belongs to the trihedral angles $SX_1Y_1Z_1$ and $S_2X_2Y_2Z_2$, respectively, where S_2 has coordinates $(\frac{1}{I_1}, \frac{1}{I_1}, \frac{1}{I_1})$. The gap between the necessary conditions in Theorem 1.1 and the sufficient conditions in Theorem 1.2, i.e., the set

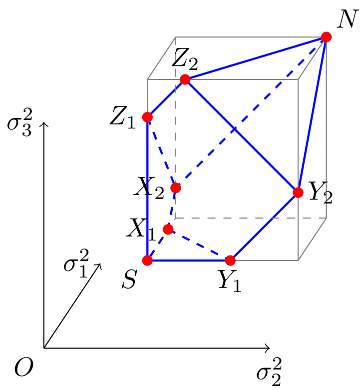
$$\{(\sigma_1^2, \sigma_2^2, \sigma_3^2) : \text{(4)–(5) hold and at least one of (6)–(8) does not hold}\}, \tag{9}$$

is shown in Figure 2(c). One can easily verify that the gap is empty only for $I_1 = I_2 = I_3$.

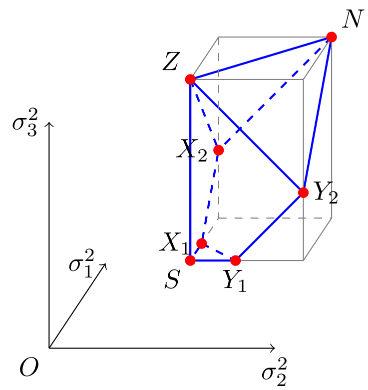
COROLLARY 1.3. *Let σ_1, σ_2, and σ_3 satisfy* (4)–(5) *for $I_1 = I_2 = I_3 = I \geq 2$. Then there exists an $I \times I \times I$ tensor $\mathcal{T}$ such that*

1. *all entries of $\mathcal{T}$ are nonnegative;*
2. *$\mathcal{T}$ is all-orthogonal;*
3. *the largest ML singular values of $\mathcal{T}$ are equal to σ_1, σ_2, and σ_3.*

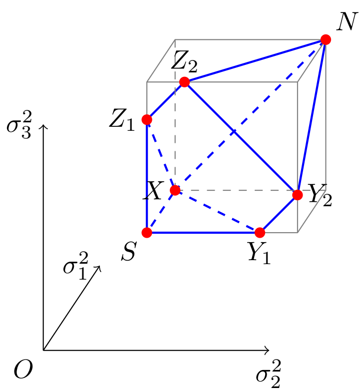
Thus, the conditions in Theorem 1.1 are not only necessary but also sufficient for σ_1, σ_2, and σ_3 to be feasible largest ML singular values of a cubic third-order tensor. Figure 1(d) shows the set of feasible triplets $(\sigma_1^2, \sigma_2^2, \sigma_3^2)$ of an $I \times I \times I$ tensor.



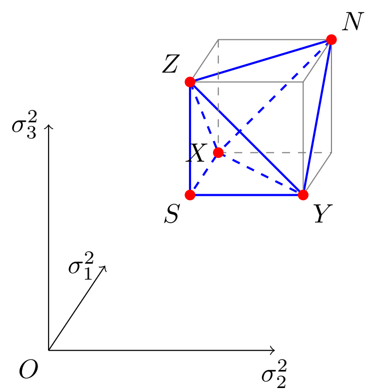
(a) $I_1 < I_2 < I_3$



(b) $I_1 = I_2 < I_3$



(c) $I_1 < I_2 = I_3$

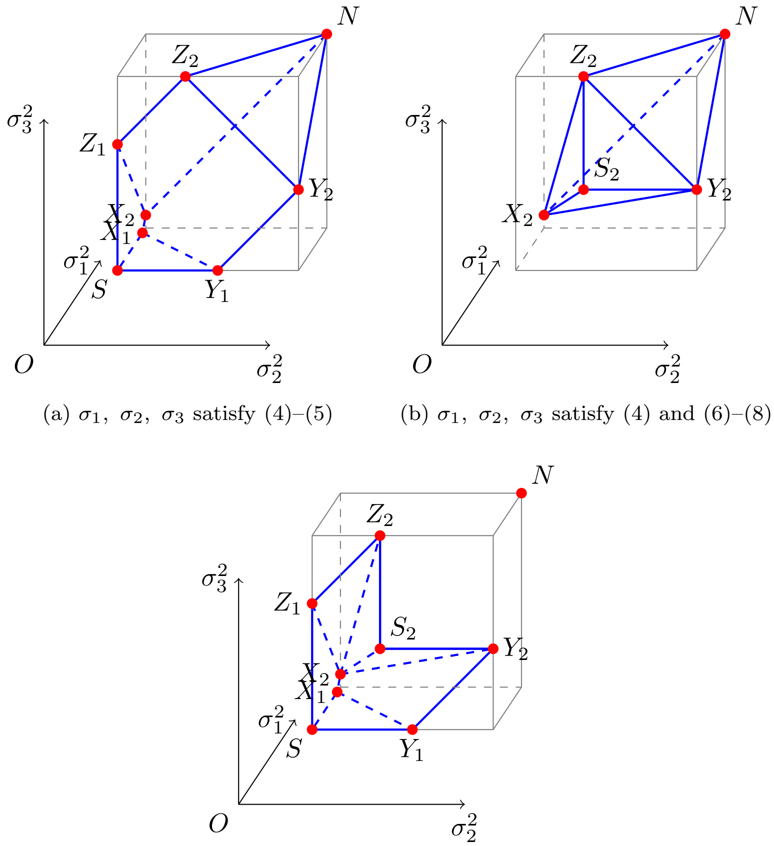


(d) $I_1 = I_2 = I_3 = I$

FIG. 1. *The typical shapes of the set* $\{(\sigma_1^2, \sigma_2^2, \sigma_3^2): \ \sigma_1, \ \sigma_2, \ \sigma_3$ *satisfy* (4)–(5)$\}$ *for* $I_1 \le I_2 \le I_3$ *(drawn for* $I_1 = 2$*,* $I_2 = 3$*,* $I_3 = 5$*, and* $\|\mathcal{T}\| = 1$*). Plot* (a) *is the case where all dimensions of a tensor are distinct. The points* S*,* X_1*,* X_2*,* Y_1*,* Y_2*,* Z_1*,* Z_2*, and* N *have coordinates* $(\frac{1}{I_1}, \frac{1}{I_2}, \frac{1}{I_3})$*,* $(1-\frac{1}{I_2}+\frac{1}{I_3}, \frac{1}{I_2}, \frac{1}{I_3})$*,* $(1, \frac{1}{I_2}, \frac{1}{I_2})$*,* $(\frac{1}{I_1}, 1-\frac{1}{I_1}+\frac{1}{I_3}, \frac{1}{I_3})$*,* $(\frac{1}{I_1}, 1, \frac{1}{I_1})$*,* $(\frac{1}{I_1}, \frac{1}{I_2}, 1-\frac{1}{I_1}+\frac{1}{I_2})$*,* $(\frac{1}{I_1}, \frac{1}{I_1}, 1)$*, and* $(1, 1, 1)$*, respectively. Plots* (b)–(c) *are the cases where a tensor has exactly two equal dimensions, the points* Z_1 *and* Z_2 *were merged into one point* Z*, and the points* X_1 *and* X_2 *were merged into one point* X*. Plot* (d) *is the case where all three dimensions of a tensor are equal to each other,* $I_1 = I_2 = I_3 = I$*. In this case, the points* Y_1 *and* Y_2 *were merged into one point* Y*, so* S*,* X*,* Y*, and* Z *have the coordinates* $(\frac{1}{I}, \frac{1}{I}, \frac{1}{I})$*,* $(1, \frac{1}{I}, \frac{1}{I})$*,* $(\frac{1}{I}, 1, \frac{1}{I})$*, and* $(\frac{1}{I}, \frac{1}{I}, 1)$*, respectively. By Corollary* 1.3*, any point* $(\sigma_1^2, \sigma_2^2, \sigma_3^2)$ *of the polyhedron* $SXYZN$ *in plot* (d) *is feasible, i.e., there exists a norm-*1 *tensor* $\mathcal{T} \in \mathbb{C}^{I\times I\times I}$ *whose squared largest ML singular values are* σ_1^2*,* σ_2^2*, and* σ_3^2*. The volume of* $SXYZN$ *equals half of the volume of the cube, i.e.,* $\frac{1}{2}(1-\frac{1}{I})^3$.

We do not have a complete view on the feasibility of points in (9). In section 3 we obtain particular results on the (non)feasibility of the points $S(\frac{1}{I_1}, \frac{1}{I_2}, \frac{1}{I_3})$, $X_1(1-\frac{1}{I_2}+\frac{1}{I_3}, \frac{1}{I_2}, \frac{1}{I_3})$, and $Y_1(\frac{1}{I_1}, 1-\frac{1}{I_1}+\frac{1}{I_3}, \frac{1}{I_3})$. Namely, we show that if $I_1 < I_2$ and $I_3 = I_1 I_2 - 1$, then the point S is not feasible, and if $I_3 = I_1 I_2$, then the point S is feasible but the points X_1 and Y_1 not.

It worth mentioning a link with scaled all-orthonormal tensors introduced recently in [5]. Tensor $\mathcal{T} \in \mathbb{C}^{I_1\times\cdots\times I_N}$ is *scaled all-orthonormal* [5, Definition 2] if at least $N-1$ of the N matrices $\mathbf{T}_{(1)}\mathbf{T}_{(1)}^H, \dots, \mathbf{T}_{(N)}\mathbf{T}_{(N)}^H$ are multiples of the identity matrix. It is clear that if the largest mode-n singular value of a norm-1 tensor is $\frac{1}{\sqrt{I_n}}$, then



(a) $\sigma_1,\ \sigma_2,\ \sigma_3$ satisfy (4)–(5) (b) $\sigma_1,\ \sigma_2,\ \sigma_3$ satisfy (4) and (6)–(8)

(c) the set in (9)

FIG. 2. *Gap between the necessary conditions in Theorem* 1.1 *and the sufficient conditions in Theorem* 1.2 *for $I_1 < I_2 < I_3$ (drawn for $I_1 = 2$, $I_2 = 5$, $I_3 = 7$, and $\|\mathcal{T}\| = 1$). The point S_2 has coordinates $(\frac{1}{I_1}, \frac{1}{I_1}, \frac{1}{I_1})$. The set in plot* (c) *is the difference of the set in plot* (a) *and the set in plot* (b).

all mode-n singular values are also $\frac{1}{\sqrt{I_n}}$. Thus, feasibility of a point belonging to the segment SX_1 (resp., SY_1 or SZ_1) is equivalent to the existence of a norm-1 $I_1 \times I_2 \times I_3$ tensor $\mathcal{T}$ such that

$$\mathbf{T}_{(2)}\mathbf{T}_{(2)}^H = \frac{1}{I_2}\mathbf{I}_{I_2}, \mathbf{T}_{(3)}\mathbf{T}_{(3)}^H = \frac{1}{I_3}\mathbf{I}_{I_3}$$

$$\left(\text{resp., } \mathbf{T}_{(1)}\mathbf{T}_{(1)}^H = \frac{1}{I_1}\mathbf{I}_{I_1}, \mathbf{T}_{(3)}\mathbf{T}_{(3)}^H = \frac{1}{I_3}\mathbf{I}_{I_3} \text{ or } \mathbf{T}_{(1)}\mathbf{T}_{(1)}^H = \frac{1}{I_1}\mathbf{I}_{I_1}, \mathbf{T}_{(2)}\mathbf{T}_{(2)}^H = \frac{1}{I_2}\mathbf{I}_{I_2}\right),$$

i.e., to the existence of a scaled all-orthonormal tensor $\mathcal{T}$.

The following results generalize Theorem 1.1 and Corollary 1.3 for Nth-order tensors.

THEOREM 1.4. *Let $\sigma_1, \dots, \sigma_N$ denote the largest ML singular values of an $I_1 \times \cdots \times I_N$ tensor $\mathcal{T}$. Then*

$$\sigma_1^2 + \cdots + \sigma_{n-1}^2 + \sigma_{n+1}^2 + \cdots + \sigma_N^2 \le (N-2)\|\mathcal{T}\|^2 + \sigma_n^2, \quad n = 1, \dots, N, \tag{10}$$

$$\|\mathcal{T}\| \ge \sigma_1 \ge \frac{1}{\sqrt{I_1}}\|\mathcal{T}\|, \dots, \|\mathcal{T}\| \ge \sigma_N \ge \frac{1}{\sqrt{I_N}}\|\mathcal{T}\|. \tag{11}$$

THEOREM 1.5. *Let $\sigma_1, \dots, \sigma_N$ satisfy* (10)–(11) *for $I_1 = \cdots = I_N = I \geq 2$. Then there exists an $I \times \cdots \times I$ tensor $\mathcal{T}$ such that*

1. *all entries of $\mathcal{T}$ are nonnegative;*
2. *$\mathcal{T}$ is all-orthogonal;*
3. *the largest ML singular values of $\mathcal{T}$ are equal to $\sigma_1, \dots, \sigma_N$.*

Thus, the conditions in Theorem 1.4 are not only necessary but also sufficient for $\sigma_1, \dots, \sigma_N$ to be feasible largest ML singular values of an $I \times \cdots \times I$ tensor. This result was independently proved for real $2 \times \cdots \times 2$ tensors in [15].

Theorems 1.1, 1.2, 1.4, and 1.5 are proved in section 2.

It is natural to ask what happens if some inequalities in (4) are replaced by equalities. Obviously, the three equalities in (4) hold if and only if $\sigma_1 = \sigma_2 = \sigma_3 = \|\mathcal{T}\|$, implying that $\mathbf{T}_{(1)}$, $\mathbf{T}_{(2)}$, and $\mathbf{T}_{(3)}$ are rank-1 matrices. Hence all the remaining ML singular values of $\mathcal{T}$ are zero. Similarly, the two equalities $\sigma_1^2 + \sigma_2^2 = \|\mathcal{T}\|^2 + \sigma_3^2$ and $\sigma_1^2 + \sigma_3^2 = \|\mathcal{T}\|^2 + \sigma_2^2$ are equivalent to $\sigma_1 = \|\mathcal{T}\|$ and $\sigma_2 = \sigma_3$, implying that $\operatorname{rank}(\mathbf{T}_{(1)}) = 1$ and $\operatorname{rank}(\mathbf{T}_{(2)}) = \operatorname{rank}(\mathbf{T}_{(3)}) =: L$, i.e., $\mathcal{T}$ is an *ML rank*-$(1, L, L)$ *tensor*, where $L \leq \min(I_2, I_3)$. It is clear that in this case the remaining nonzero mode-2 and mode-3 singular values of $\mathcal{T}$ also coincide and may take any positive values whose squares sum up to $\|\mathcal{T}\|^2 - \sigma_2^2$. In section 4 we characterize the tensors $\mathcal{T}$ for which the single equality $\sigma_1^2 + \sigma_2^2 = \|\mathcal{T}\|^2 + \sigma_3^2$ holds. We show that $\mathcal{T}$ is necessarily equal to a sum of ML rank-$(L_1, 1, L_1)$ and ML rank-$(1, L_2, L_2)$ tensors and we give a complete description of all its ML singular values. The description relies on a problem posed by Weyl in 1912: given the eigenvalues of two $n \times n$ Hermitian matrices $\mathbf{A}$ and $\mathbf{B}$, what are all the possible eigenvalues of $\mathbf{A} + \mathbf{B}$? The following answer was conjectured by Horn in 1962 [8] and has been proved through the development of the theory of honeycombs in [9, 10] (see also [1, 11]). Let

$$\lambda_i(\cdot) \text{ denote the } i\text{th largest eigenvalue of a Hermitian matrix.}$$

If

$$\alpha_i = \lambda_i(\mathbf{A}), \qquad \beta_i = \lambda_i(\mathbf{B}), \qquad \gamma_i = \lambda_i(\mathbf{A} + \mathbf{B}), \tag{12}$$

then α_i, β_i, and γ_i satisfy the trivial equality

$$\gamma_1 + \cdots + \gamma_n = \alpha_1 + \cdots + \alpha_n + \beta_1 + \cdots + \beta_n \tag{13}$$

and the list of linear inequalities

$$\sum_{k \in K} \gamma_k \leq \sum_{i \in I} \alpha_i + \sum_{j \in J} \beta_j, \qquad (I, J, K) \in T_r^n, \qquad 1 \leq r \leq n - 1, \tag{14}$$

where $I = \{i_1, \dots, i_r\}$, $J = \{j_1, \dots, j_r\}$, $K = \{k_1, \dots, k_r\}$ are subsets of $\{1, \dots, n\}$ and T_r^n denotes a particular finite set of triplets (I, J, K). (The construction of T_r^n is given in Appendix A.) The inverse statement also holds: if α_i, β_i, and γ_i satisfy (13) and (14), then there exist $n \times n$ Hermitian matrices $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ such that (12) holds.

We have the following results.

THEOREM 1.6. *Let $\sigma_1^2 + \sigma_2^2 = \|\mathcal{T}\|^2 + \sigma_3^2$. Then $\mathcal{T}$ is a sum of ML rank-$(L_1, 1, L_1)$ and ML rank-$(1, L_2, L_2)$ tensors, where $L_1 \leq \min(I_1, I_3)$ and $L_2 \leq \min(I_2 - 1, I_3)$.*

THEOREM 1.7. *Let $\sigma_1^2+\sigma_2^2=\|\mathcal{T}\|^2+\sigma_3^2$. Then the values*

$$\begin{aligned}\sigma_1=\sigma_{11}\geq\sigma_{12}\geq\cdots\geq\sigma_{1I_1}\geq 0,\\ \sigma_2=\sigma_{21}\geq\sigma_{22}\geq\cdots\geq\sigma_{2I_2}\geq 0,\\ \sigma_3=\sigma_{31}\geq\sigma_{32}\geq\cdots\geq\sigma_{3I_3}\geq 0,\end{aligned}$$

*are the mode-*1*, mode-*2*, and mode-*3 *singular values of an $I_1\times I_2\times I_3$ tensor $\mathcal{T}$, respectively, if and only if*

$$\begin{aligned}\sigma_{11}^2+\cdots+\sigma_{1I_1}^2=\sigma_{21}^2+\cdots+\sigma_{2I_2}^2=\sigma_{31}^2+\cdots+\sigma_{3I_3}^2=\|\mathcal{T}\|^2,\\ \sigma_{1i}=0 \quad \textit{for} \quad i>\min(I_1,I_3),\\ \sigma_{2i}=0 \quad \textit{for} \quad i>\min(I_2,I_3),\end{aligned}$$

and (13) *and* (14) *hold for*

$$\alpha_i=\begin{cases}\sigma_{1i+1}^2, & i\leq\min(I_1,I_3),\\ 0 & \textit{otherwise},\end{cases}\qquad \beta_i=\begin{cases}\sigma_{2i+1}^2, & i\leq\min(I_2,I_3),\\ 0 & \textit{otherwise},\end{cases}\gamma_i=\sigma_{3i+1}^2, \tag{15}$$

and $n=I_3-1$.

Example 1.8. If $n=2$, then $T_1^2=\{(i,j,k): k=i+j-1, 1\leq i,j,k\leq 2\}=\{(1,1,1),(1,2,2),(2,1,2)\}$ (see Appendix A). By Horn's conjecture, the equality $\gamma_1+\gamma_2=\alpha_1+\alpha_2+\beta_1+\beta_2$ together with the inequalities (also known as the Weyl inequalities)

$$\gamma_1\leq\alpha_1+\beta_1,\qquad \gamma_2\leq\alpha_1+\beta_2,\qquad \gamma_2\leq\alpha_2+\beta_1 \tag{16}$$

characterize the values $\alpha_1,\alpha_2,\beta_1,\beta_2,\gamma_1,\gamma_2$ that can be eigenvalues of 2×2 Hermitian matrices $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{A}+\mathbf{B}$. Let $\sigma_{11}^2+\sigma_{21}^2=\|\mathcal{T}\|^2+\sigma_{31}^2$. From Theorem 1.7 and (16) it follows that the values $\sigma_{11}\geq\sigma_{12}\geq\sigma_{13}\geq 0$, $\sigma_{21}\geq\sigma_{22}\geq\sigma_{23}\geq 0$, and $\sigma_{31}\geq\sigma_{32}\geq\sigma_{33}\geq 0$ are the mode-1, mode-2, and mode-3 singular values, respectively, of a $3\times 3\times 3$ tensor $\mathcal{T}$ if and only if

$$\begin{aligned}\sigma_{11}^2+\sigma_{12}^2+\sigma_{13}^2=\sigma_{21}^2+\sigma_{22}^2+\sigma_{23}^2=\sigma_{31}^2+\sigma_{32}^2+\sigma_{33}^2=\|\mathcal{T}\|^2,\\ \sigma_{32}^2\leq\sigma_{12}^2+\sigma_{22}^2,\qquad \sigma_{33}^2\leq\sigma_{12}^2+\sigma_{23}^2,\qquad \sigma_{33}^2\leq\sigma_{13}^2+\sigma_{22}^2.\end{aligned}$$

Horn's conjecture has recently also been linked to singular values of matrix unfoldings in the tensor train format [13].

2. Proofs of Theorems 1.1, 1.2, 1.4, and 1.5.

The following lemma will be used in the proof of Theorem 1.1.

LEMMA 2.1. *Let $\mathbf{H}=(\mathbf{H}_{ij})_{i,j=1}^{I_3}\in\mathbb{C}^{I_3I_1\times I_3I_1}$ be a positive semidefinite matrix consisting of the blocks $\mathbf{H}_{ij}\in\mathbb{C}^{I_1\times I_1}$. Then*

$$\lambda_{max}(\mathbf{H}_{11}+\cdots+\mathbf{H}_{I_3I_3})+\lambda_{max}(\mathbf{H})\leq\operatorname{tr}(\mathbf{H})+\lambda_{max}(\mathbf{\Phi}(\mathbf{H})), \tag{17}$$

where $\mathbf{\Phi}(\mathbf{H})$ denotes the $I_3\times I_3$ matrix with the entries $(\mathbf{\Phi}(\mathbf{H}))_{ij}=\operatorname{tr}(\mathbf{H}_{ij})$ and $\lambda_{max}(\cdot)$ denotes the largest eigenvalue of a matrix.

Proof. To get an idea of the proof we refer the reader to the mathoverflow page (http://mathoverflow.net/questions/248975/) where the case $I_3=2$ was discussed. Here we present a formal proof for $I_3\geq 2$. Let $\mathbf{H}=\sum_{r=1}^{R}\mathbf{w}_r\mathbf{w}_r^H$, where $\mathbf{w}_r$ are orthogonal and $\mathbf{w}_r = [\mathbf{w}_{1r}^T \ \dots \ \mathbf{w}_{I_3 r}^T]^T$ with $\mathbf{w}_{kr} \in \mathbb{C}^{I_3}$. First, we rewrite (17) in terms of $\mathbf{w}_{kr}$, $1 \le k \le I_3$, $1 \le r \le R$. Without loss of generality, we can assume that $\|\mathbf{w}_1\| = \max_r \|\mathbf{w}_r\|$. Hence,

$$\lambda_{max}(\mathbf{H}) = \|\mathbf{w}_1\|^2 = \sum_{k=1}^{I_3} \|\mathbf{w}_{k1}\|^2. \tag{18}$$

It is clear that

$$\mathbf{H}_{ij} = \sum_{r=1}^{R} \mathbf{w}_{ir}\mathbf{w}_{jr}^H, \qquad 1 \le i, j \le I_3.$$

Hence

$$\lambda_{max}(\mathbf{H}_{11} + \dots + \mathbf{H}_{I_3 I_3}) = \max_{\|\mathbf{x}\|=1} \sum_{k=1}^{I_3} (\mathbf{H}_{kk}\mathbf{x}, \mathbf{x}) = \max_{\|\mathbf{x}\|=1} \sum_{k=1}^{I_3}\sum_{r=1}^{R} |(\mathbf{w}_{kr}, \mathbf{x})|^2. \tag{19}$$

Since $\mathbf{H} = \sum_{r=1}^R \mathbf{w}_r \mathbf{w}_r^H$, it follows that

$$\operatorname{tr}(\mathbf{H}) = \sum_{r=1}^{R} \|\mathbf{w}_r\|^2. \tag{20}$$

Since

$$\mathbf{\Phi}(\mathbf{H})_{ij} = \operatorname{tr}(\mathbf{H}_{ij}) = \operatorname{tr}\left(\sum_{r=1}^{R} \mathbf{w}_{ir}\mathbf{w}_{jr}^H\right) = \sum_{r=1}^{R} \mathbf{w}_{jr}^H \mathbf{w}_{ir} = \sum_{r=1}^{R} \mathbf{w}_{ir}^T \mathbf{w}_{jr}^*,$$

it follows that

$$\begin{aligned}\mathbf{\Phi}(\mathbf{H}) &= \sum_{r=1}^{R} \begin{bmatrix} \mathbf{w}_{1r}^T\mathbf{w}_{1r}^* & \dots & \mathbf{w}_{1r}^T\mathbf{w}_{I_3 r}^* \\ \vdots & \dots & \vdots \\ \mathbf{w}_{I_3 r}^T\mathbf{w}_{1r}^* & \dots & \mathbf{w}_{I_3 r}^T\mathbf{w}_{I_3 r}^* \end{bmatrix} \\ &= \sum_{r=1}^{R} \begin{bmatrix} \mathbf{w}_{1r}^T \\ \vdots \\ \mathbf{w}_{I_3 r}^T \end{bmatrix} \begin{bmatrix} \mathbf{w}_{1r}^* & \dots & \mathbf{w}_{I_3 r}^* \end{bmatrix} = \sum_{r=1}^{R} \mathbf{W}_r^T \mathbf{W}_r^*,\end{aligned} \tag{21}$$

where

$$\mathbf{W}_r := [\mathbf{w}_{1r} \ \dots \ \mathbf{w}_{I_3 r}] \in \mathbb{C}^{I_1 \times I_3}.$$

Now we prove (17). By (18), (19), the Cauchy inequality, and (20),

$$\begin{aligned}&\lambda_{max}(\mathbf{H}) + \lambda_{max}(\mathbf{H}_{11} + \dots \mathbf{H}_{I_3 I_3}) \\ &\quad = \|\mathbf{w}_1\|^2 + \max_{\|\mathbf{x}\|=1}\left[\sum_{k=1}^{I_3} |(\mathbf{w}_{k1}, \mathbf{x})|^2 + \sum_{k=1}^{I_3}\sum_{r=2}^{R} |(\mathbf{w}_{kr}, \mathbf{x})|^2\right] \\ &\quad \le \|\mathbf{w}_1\|^2 + \max_{\|\mathbf{x}\|=1}\left[\sum_{k=1}^{I_3} |(\mathbf{w}_{k1}, \mathbf{x})|^2\right] + \sum_{r=2}^{R} \|\mathbf{w}_r\|^2 = \operatorname{tr}(\mathbf{H}) + \max_{\|\mathbf{x}\|=1}\left[\sum_{k=1}^{I_3} |(\mathbf{w}_{k1}, \mathbf{x})|^2\right].\end{aligned} \tag{22}$$

To complete the proof of (17) we should show that

$$\max_{\|\mathbf{x}\|=1}\left[\sum_{k=1}^{I_3}|(\mathbf{w}_{k1},\mathbf{x})|^2\right]\leq\lambda_{\max}(\mathbf{\Phi}(\mathbf{H})).$$

This can be done as follows:

$$\begin{aligned}\max_{\|\mathbf{x}\|=1}\left[\sum_{k=1}^{I_3}|(\mathbf{w}_{k1},\mathbf{x})|^2\right]&=\max_{\|\mathbf{x}\|=1}\left[\sum_{k=1}^{I_3}\mathbf{x}^H\mathbf{w}_{k1}\mathbf{w}_{k1}^H\mathbf{x}\right]=\lambda_{max}\left(\mathbf{W}_1\mathbf{W}_1^H\right)\\&=\lambda_{max}\left(\mathbf{W}_1^H\mathbf{W}_1\right)\leq\lambda_{max}\left(\sum_{r=1}^{R}\mathbf{W}_r^H\mathbf{W}_r\right)\\&=\lambda_{max}(\mathbf{\Phi}(\mathbf{H})^*)=\lambda_{max}(\mathbf{\Phi}(\mathbf{H})).\end{aligned}\tag{23}$$

∎

Now we are ready to prove Theorem 1.1.

Proof of Theorem 1.1. The three inequalities in (5) are obvious. We prove that $\sigma_1^2+\sigma_2^2\leq\|\mathcal{T}\|^2+\sigma_3^2$. The proofs of the inequalities $\sigma_1^2+\sigma_3^2\leq\|\mathcal{T}\|^2+\sigma_2^2$ and $\sigma_2^2+\sigma_3^2\leq\|\mathcal{T}\|^2+\sigma_1^2$ can be obtained in a similar way.

By definition of ML singular values,

$$\begin{aligned}\sigma_1^2&=\lambda_{max}(\mathbf{T}_{(1)}\mathbf{T}_{(1)}^H)=\lambda_{max}(\mathbf{T}_1\mathbf{T}_1^H+\cdots+\mathbf{T}_{I_3}\mathbf{T}_{I_3}^H),\\\sigma_2^2&=\lambda_{max}(\mathbf{T}_{(2)}^H\mathbf{T}_{(2)})=\lambda_{max}(\mathbf{T}_{(2)}^T\mathbf{T}_{(2)}^*)=\lambda_{max}(\mathbf{H}),\end{aligned}$$

where

$$\mathbf{H}=\mathbf{T}_{(2)}^T\mathbf{T}_{(2)}^*=\begin{bmatrix}\mathbf{T}_1\mathbf{T}_1^H&\dots&\mathbf{T}_1\mathbf{T}_{I_3}^H\\\vdots&\dots&\vdots\\\mathbf{T}_{I_3}\mathbf{T}_1^H&\dots&\mathbf{T}_{I_3}\mathbf{T}_{I_3}^H\end{bmatrix}.$$

Since $\operatorname{vec}(\mathbf{T}_i)^T(\operatorname{vec}(\mathbf{T}_j)^T)^H=\operatorname{tr}(\mathbf{T}_i\mathbf{T}_j^H)$, it follows that

$$\sigma_3^2=\lambda_{max}(\mathbf{T}_{(3)}\mathbf{T}_{(3)}^H)=\lambda_{max}(\mathbf{\Phi}(\mathbf{H})),$$

where

$$\mathbf{\Phi}(\mathbf{H})=\begin{bmatrix}\operatorname{tr}\left(\mathbf{T}_1\mathbf{T}_1^H\right)&\dots&\operatorname{tr}\left(\mathbf{T}_1\mathbf{T}_{I_3}^H\right)\\\vdots&\dots&\vdots\\\operatorname{tr}\left(\mathbf{T}_{I_3}\mathbf{T}_1^H\right)&\dots&\operatorname{tr}\left(\mathbf{T}_{I_3}\mathbf{T}_{I_3}^H\right)\end{bmatrix}.$$

Since $\|\mathcal{T}\|^2=\operatorname{tr}(\mathbf{H})$, the inequality $\sigma_1^2+\sigma_2^2\leq\|\mathcal{T}\|^2+\sigma_3^2$ is equivalent to

$$\lambda_{max}\left(\mathbf{T}_1\mathbf{T}_1^H+\cdots+\mathbf{T}_{I_3}\mathbf{T}_{I_3}^H\right)+\lambda_{max}(\mathbf{H})\leq\operatorname{tr}(\mathbf{H})+\lambda_{max}(\mathbf{\Phi}(\mathbf{H})),$$

which holds by Lemma 2.1. ∎

Proof of Theorem 1.2. The proof consists of three steps. In the first step we construct all-orthogonal and nonnegative $I_1\times I_2\times I_3$ tensors $\mathcal{S}_2$, $\mathcal{X}_2$, $\mathcal{Y}_2$, $\mathcal{Z}_2$, and $\mathcal{N}$ whose squared largest ML singular values are the coordinates of $S_2(\frac{1}{I_1},\frac{1}{I_1},\frac{1}{I_1})$,

$X_2(1, \frac{1}{I_2}, \frac{1}{I_2})$, $Y_2(\frac{1}{I_1}, 1, \frac{1}{I_1})$, $Z_2(\frac{1}{I_1}, \frac{1}{I_1}, 1)$, and $N(1, 1, 1)$, respectively (see Figure 2(b)). Then we show that because of the zero patterns of $\mathcal{S}_2$, $\mathcal{X}_2$, $\mathcal{Y}_2$, $\mathcal{Z}_2$, and $\mathcal{N}$, the tensor

$$\mathcal{T} = \left(t_{S_2}\mathcal{S}_2^2 + t_{X_2}\mathcal{X}_2^2 + t_{Y_2}\mathcal{Y}_2^2 + t_{Z_2}\mathcal{Z}_2^2 + t_N\mathcal{N}^2\right)^{\frac{1}{2}} \tag{24}$$

is all-orthogonal for any nonnegative values t_{S_2}, t_{X_2}, t_{Y_2}, t_{Z_2}, t_N. The superscripts "2" and "$\frac{1}{2}$" in (24) denote the entrywise operations. Finally, in the third step, we find nonnegative values t_{S_2}, t_{X_2}, t_{Y_2}, t_{Z_2}, t_N such that $\mathcal{T}$ is a norm-1 tensor whose squared largest ML singular values are equal to σ_1^2, σ_2^2, and σ_3^2.

Step 1. Let π denote the cyclic permutation $\pi : 1 \to I_1 \to I_1 - 1 \to \cdots \to 2 \to 1$. The tensors $\mathcal{S}_2$, $\mathcal{X}_2$, $\mathcal{Y}_2$, and $\mathcal{Z}_2$ are defined by

$$\mathcal{S}_{2,ijk} = \begin{cases} \dfrac{1}{I_1} & \text{if } j = \pi^{k-1}(i) \text{ and } 1 \le i, k \le I_1, \\ 0 & \text{otherwise}, \end{cases}$$

$$\mathcal{X}_{2,ijk} = \begin{cases} \dfrac{1}{\sqrt{I_2}} & \text{if } j = \pi^{k-1}(i),\ i = 1,\ \text{and } 1 \le k \le I_1, \\ \dfrac{1}{\sqrt{I_2}} & \text{if } i = 1 \text{ and } I_1 < j = k \le I_2, \\ 0 & \text{otherwise}, \end{cases}$$

$$\mathcal{Y}_{2,ijk} = \begin{cases} \dfrac{1}{\sqrt{I_1}} & \text{if } j = \pi^{k-1}(i),\ j = 1,\ \text{and } 1 \le k \le I_1, \\ 0 & \text{otherwise}, \end{cases}$$

$$\mathcal{Z}_{2,ijk} = \begin{cases} \dfrac{1}{\sqrt{I_1}} & \text{if } j = \pi^{k-1}(i),\ k = 1,\ \text{and } 1 \le i \le I_1, \\ 0 & \text{otherwise}, \end{cases}$$

and the tensor $\mathcal{N}$, by definition, has only one nonzero entry, $\mathcal{N}_{111} = 1$. For instance, if $I_1 = I_2 = I_3 = 2$, then the first matrix unfoldings of $\mathcal{S}_2$, $\mathcal{X}_2$, $\mathcal{Y}_2$, $\mathcal{Z}_2$, and $\mathcal{N}$ have the form

$$\mathbf{S}_{2,(1)} = \frac{1}{2}\begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}, \ \mathbf{X}_{2,(1)} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix},$$

$$\mathbf{Y}_{2,(1)} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \ \mathbf{Z}_{2,(1)} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \ \mathbf{N}_{(1)} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Step 2. It is clear that the (i, j, k)th entry of a linear combination of $\mathcal{S}_2^2$, $\mathcal{X}_2^2$, $\mathcal{Y}_2^2$, $\mathcal{Z}_2^2$, and $\mathcal{N}^2$ may be nonzero only if

$$j = \pi^{k-1}(i) \text{ and } 1 \le i, k \le I_1 \qquad \text{or} \qquad i = 1 \text{ and } I_1 < j = k \le I_2.$$

The same is also true for $\mathcal{T}$ defined in (24). One can easily check that each column of $\mathbf{T}_{(1)}$, $\mathbf{T}_{(2)}$, and $\mathbf{T}_{(3)}$ contains at most one nonzero entry, implying that $\mathcal{T}$ is all-orthogonal tensor.

Step 3. From the construction of the all-orthogonal tensors $\mathcal{S}_2$, $\mathcal{X}_2$, $\mathcal{Y}_2$, $\mathcal{Z}_2$, and $\mathcal{N}$ it follows that their largest ML singular values are equal to the Frobenius norms of the first rows of their matrix unfoldings. Thus, the same property should also hold for $\mathcal{T}$ whenever the values t_{S_2}, t_{X_2}, t_{Y_2}, t_{Z_2}, and t_N are nonnegative. Now the result follows from the fact that the polyhedron in Figure 2(b) is the convex hull of the points S_2, X_2, Y_2, Z_2, and N. We can also write the values of t_{S_2}, t_{X_2}, t_{Y_2}, t_{Z_2}, and t_N explicitly. We set

$$f(\sigma_1^2,\sigma_2^2,\sigma_3^2) := (I_1I_2 + I_2 - 2I_1)\sigma_1^2 + (I_1-1)I_2\sigma_2^2 + (I_1-1)I_2\sigma_3^2 + (2 - I_1I_2 - I_2).$$

If $(\sigma_1^2,\sigma_2^2,\sigma_3^2)$ belongs to the tetrahedron $X_2Y_2Z_2N$, i.e., $f(\sigma_1^2,\sigma_2^2,\sigma_3^2)\geq 0$, then

$$\begin{aligned}
t_{X_2} &= \frac{I_2}{2(I_2-1)}(1+\sigma_1^2-\sigma_2^2-\sigma_3^2),\ t_{Y_2} = \frac{I_1}{2(I_1-1)}(1+\sigma_2^2-\sigma_1^2-\sigma_3^2),\\
t_{Z_2} &= \frac{I_1}{2(I_1-1)}(1+\sigma_3^2-\sigma_1^2-\sigma_2^2),\\
t_N &= 1 - t_{X_2} - t_{Y_2} - t_{Z_2} = \frac{f(\sigma_1^2,\sigma_2^2,\sigma_3^2)}{2(I_1-1)(I_2-1)}, \qquad t_{S_2} = 0.
\end{aligned}$$

If $(\sigma_1^2,\sigma_2^2,\sigma_3^2)$ belongs to the tetrahedron $X_2Y_2Z_2S_2$, i.e., $f(\sigma_1^2,\sigma_2^2,\sigma_3^2)\leq 0$, then

$$\begin{aligned}
t_{X_2} &= \frac{I_1}{I_1-1}\left(\sigma_1^2 - \frac{1}{I_1}\right),\\
t_{Y_2} &= \frac{I_1}{I_1-1}\left(\sigma_2^2 - \frac{1}{I_1}\right) + \frac{(I_2-I_1)I_1}{(I_1^2-1)I_2}\left(\sigma_1^2 - \frac{1}{I_1}\right),\\
t_{Z_2} &= \frac{I_1}{I_1-1}\left(\sigma_3^2 - \frac{1}{I_1}\right) + \frac{(I_2-I_1)I_1}{(I_1^2-1)I_2}\left(\sigma_1^2 - \frac{1}{I_1}\right),\\
t_{S_2} &= 1 - t_{X_2} - t_{Y_2} - t_{Z_2} = \frac{-f\left(\sigma_1^2,\sigma_2^2,\sigma_3^2\right)I_1}{I_2(I_1-1)^2}, \qquad t_N = 0.
\end{aligned}$$

□

Proof of Theorem 1.4. The inequalities in (11) are obvious. We prove that

$$\sigma_1^2 + \cdots + \sigma_{N-1}^2 \leq (N-2)\|\mathcal{T}\|^2 + \sigma_N^2. \tag{25}$$

The proofs of the remaining $N-1$ inequalities in (10) can be obtained in a similar way.

The proof of (25) consists of two steps. In the first step we reshape $\mathcal{T}$ into third-order tensors $\mathcal{T}^{[1]},\dots,\mathcal{T}^{[N-2]}$ and compute their matrix unfoldings. In this step we will make use of (1) for $N=3$. For the reader's convenience and for future reference here we write a third-order version of (1) explicitly: if $\mathcal{X}\in\mathbb{C}^{I\times J\times K}$, then for all values of indices i, j, and k

$$\begin{aligned}
&\text{the } (i, j+(k-1)J)\text{th entry of } \mathbf{X}_{(1)} = \text{the } (j, i+(k-1)I)\text{th entry of } \mathbf{X}_{(2)}\\
&\quad = \text{the } (k, i+(j-1)I)\text{th entry of } \mathbf{X}_{(3)} = \text{the } (i,j,k)\text{th entry of } \mathcal{X}.
\end{aligned} \tag{26}$$

In the second step, we apply the first inequality in (4) to each tensor $\mathcal{T}^{[n]}$, then we sum up the obtained inequalities and show that the result coincides with inequality (25).

Step 1. Let $n\in\{1,\dots,N-2\}$. A third-order tensor $\mathcal{T}^{[n]}\in\mathbb{C}^{I_1\cdots I_n\times I_{n+1}\times I_{n+2}\cdots I_N}$ is constructed as follows:

$$\text{the } \left(i_1 + \sum_{k=2}^{n}(i_k-1)\prod_{l=1}^{k-1} I_l,\, i_{n+1},\, i_{n+2} + \sum_{k=n+3}^{N}(i_k-1)\prod_{l=n+2}^{k-1} I_l\right)\text{th entry of } \mathcal{T}^{[n]}$$
$$\text{is equal to the } (i_1,\dots,i_N)\text{th entry of } \mathcal{T}.$$

Now we apply (26) for $\mathcal{X} = \mathcal{T}^{[n]}$ and

$$i = i_1 + \sum_{k=2}^{n}(i_k - 1)\prod_{l=1}^{k-1} I_l, \quad j = i_{n+1}, \quad k = i_{n+2} + \sum_{k=n+3}^{N}(i_k - 1)\prod_{l=n+2}^{k-1} I_l.$$

After simple algebraic manipulations, we obtain that

$$\begin{aligned}
&\text{the } \left(i_1 + \sum_{k=2}^{n}(i_k - 1)\prod_{l=1}^{k-1} I_l, i_{n+1} + \sum_{k=n+2}^{N}(i_k - 1)\prod_{l=n+1}^{k-1} I_l\right)\text{th entry of } \mathbf{T}_{(1)}^{[n]}\\
&\quad = \text{the } \left(i_{n+1}, 1 + \sum_{\substack{k=2\\ k\neq n+1}}^{N}(i_k - 1)\prod_{l=1 l\neq n+1}^{k-1} I_l\right)\text{th entry of } \mathbf{T}_{(2)}^{[n]}\\
&\quad = \text{the } \left(i_{n+2} + \sum_{k=n+3}^{N}(i_k - 1)\prod_{l=n+2}^{k-1} I_l, i_1 + \sum_{k=2}^{n+1}(i_k - 1)\prod_{l=1}^{k-1} I_l\right)\text{th entry of } \mathbf{T}_{(3)}^{[n]}\\
&\quad = \text{the } (i_1, \dots, i_N)\text{th entry of } \mathcal{T}.
\end{aligned} \tag{27}$$

Step 2. From (27) and (1) it follows that

$$\mathbf{T}_{(1)}^{[1]} = \mathbf{T}_{(1)}, \tag{28}$$

$$\mathbf{T}_{(2)}^{[n]} = \mathbf{T}_{(n+1)}, \qquad 1 \le n \le N-2, \tag{29}$$

$$\mathbf{T}_{(3)}^{[N-2]} = \mathbf{T}_{(N)}. \tag{30}$$

Comparing the expressions of $\mathbf{T}_{(1)}^{[n]}$ and $\mathbf{T}_{(3)}^{[n]}$ in (27), we obtain that

$$\mathbf{T}_{(3)}^{[n]} = \left(\mathbf{T}_{(1)}^{[n+1]}\right)^T, \qquad 1 \le n \le N-3. \tag{31}$$

By Theorem 1.1, for every $n \in \{1, \dots, N-2\}$

$$\sigma_{max}^2\left(\mathbf{T}_{(1)}^{[n]}\right) + \sigma_{max}^2\left(\mathbf{T}_{(2)}^{[n]}\right) \le \|\mathcal{T}^{[n]}\|^2 + \sigma_{max}^2\left(\mathbf{T}_{(3)}^{[n]}\right) = \|\mathcal{T}\|^2 + \sigma_{max}^2\left(\mathbf{T}_{(3)}^{[n]}\right), \tag{32}$$

where $\sigma_{max}(\cdot)$ denotes the largest singular value of a matrix. Substituting (28)–(31) into (32) we obtain

$$\begin{gathered}
\sigma_1^2 + \sigma_2^2 \le \|\mathcal{T}\|^2 + \sigma_{max}^2(\mathbf{T}_{(3)}^{[1]}) = \|\mathcal{T}\|^2 + \sigma_{max}^2\left(\mathbf{T}_{(1)}^{[2]}\right),\ n = 1,\\
\sigma_{max}^2\left(\mathbf{T}_{(1)}^{[2]}\right) + \sigma_3^2 \le \|\mathcal{T}\|^2 + \sigma_{max}^2\left(\mathbf{T}_{(3)}^{[2]}\right) = \|\mathcal{T}\|^2 + \sigma_{max}^2\left(\mathbf{T}_{(1)}^{[3]}\right),\ n = 2,\\
\vdots\\
\sigma_{max}^2\left(\mathbf{T}_{(1)}^{[N-3]}\right) + \sigma_{N-2}^2 \le \|\mathcal{T}\|^2 + \sigma_{max}^2\left(\mathbf{T}_{(3)}^{[N-3]}\right) = \|\mathcal{T}\|^2 + \sigma_{max}^2\left(\mathbf{T}_{(1)}^{[N-2]}\right),\ n = N-3,\\
\sigma_{max}^2\left(\mathbf{T}_{(1)}^{[N-2]}\right) + \sigma_{N-1}^2 \le \|\mathcal{T}\|^2 + \sigma_{max}^2\left(\mathbf{T}_{(3)}^{[N-2]}\right) = \|\mathcal{T}\|^2 + \sigma_N^2,\ n = N-2.
\end{gathered}$$

Summing up the above inequalities and canceling identical terms on the left- and right-hand sides we obtain (25). ∎

Proof of Theorem 1.5. It can be checked that a polyhedron described by the inequalities in (10)–(11) is a convex hull of $2^N - N$ points,

$$V = \left\{(\alpha_1, \dots, \alpha_N), \quad \alpha_n \in \left\{\frac{1}{I}, 1\right\} \text{ and at least two of } \alpha\text{-s are equal to } \frac{1}{I}\right\}. \tag{33}$$

To show that each point of the polyhedron is feasible we proceed as in the proof of Theorem 1.2.

First, for each $(\alpha_1, \dots, \alpha_N) \in V$ we construct an all-orthogonal and nonnegative $I \times \cdots \times I$ tensor $P^{\alpha_1,\dots,\alpha_N}$ whose squared largest ML singular values are $\alpha_1, \dots, \alpha_N$.

Let π denote the cyclic permutation $\pi : 1 \to I \to I-1 \to \cdots \to 2 \to 1$. The tensor $\mathcal{P}^{\frac{1}{I},\dots,\frac{1}{I}}$ is defined by

$$\mathcal{P}^{\frac{1}{I},\dots,\frac{1}{I}}_{i_1,\dots,i_N} = \begin{cases} I^{-\frac{N-1}{2}} & \text{if } i_2 = \pi^{i_3+\cdots+i_N-N+2}(i_1), \\ 0 & \text{otherwise,} \end{cases}$$

and the tensor $\mathcal{P}^{1,\dots,1}$, by definition, has only one nonzero entry, $\mathcal{P}^{1\dots 1}_{1,\dots,1} = 1$. Let $(\alpha_1, \dots, \alpha_N) \in V \setminus \{(\frac{1}{I}, \dots, \frac{1}{I}), (1, \dots, 1)\}$ and $j_1, \dots, j_k$ denote all indices such that $\alpha_{j_1} = \cdots = \alpha_{j_k} = 1$. Then the tensor $P^{\alpha_1,\dots,\alpha_N}$ is defined by

$$\mathcal{P}^{\alpha_1,\dots,\alpha_N}_{i_1,\dots,i_N} = \begin{cases} I^{-\frac{N-1-k}{2}} & \text{if } i_2 = \pi^{i_3+\cdots+i_N-N+2}(i_1) \text{ and } i_{j_1} = \cdots = i_{j_k} = 1, \\ 0 & \text{otherwise.} \end{cases}$$

For instance, if $N = 4$ and $I = 2$, then the first matrix unfolding of $\mathcal{P}^{\frac{1}{I},\dots,\frac{1}{I}}$ is given by

$$\mathcal{P}^{\frac{1}{I},\dots,\frac{1}{I}}_{(1)} = \frac{1}{2\sqrt{2}} \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}$$

and the first matrix unfoldings of the remaining tensors $\mathcal{P}^{\alpha_1,\dots,\alpha_N}$ can be obtained from $\mathcal{P}^{\frac{1}{I},\dots,\frac{1}{I}}_{(1)}$ by rescaling and introducing more zeros.

It is clear that the $(i_1, \dots, i_N)$th entry of a linear combination of $\mathcal{P}^{\frac{1}{I},\dots,\frac{1}{I}}, \dots, \mathcal{P}^{1,\dots,1}$ may be nonzero only if

$$i_2 = \pi^{i_3+\cdots+i_N-N+2}(i_1).$$

The same is also true for $\mathcal{T}$ defined by

$$\mathcal{T} = \left(\sum_{(\alpha_1,\dots,\alpha_N)\in V} t_{\alpha_1,\dots,\alpha_N} {P^{\alpha_1,\dots,\alpha_N}}^2 \right)^{\frac{1}{2}},$$

where, as before, the superscripts "2" and "$\frac{1}{2}$" denote the entrywise operations. One can easily check that each column of $\mathbf{T}_{(1)}, \dots, \mathbf{T}_{(N)}$ contains at most one nonzero entry, implying that $\mathcal{T}$ is all-orthogonal tensor. Finally, from the construction of the all-orthogonal tensors $P^{\alpha_1,\dots,\alpha_N}$ it follows that their largest ML singular values are equal to the Frobenius norms of the first rows of their matrix unfoldings. Thus, the same property should also hold for $\mathcal{T}$ whenever the values $t_{\alpha_1,\dots,\alpha_N}$ are nonnegative. Now the result follows from the fact that the polyhedron described by the inequalities in (10)–(11) is a convex hull of points in V. ◻

Note that in the proof of Theorem 1.5 the constructed tensor $\mathcal{T}$ has squared singular values in the nth mode equal to $\sigma_n^2, \frac{1}{I-1}(1-\sigma_n^2), \dots, \frac{1}{I-1}(1-\sigma_n^2)$, i.e., the $I-1$ smallest singular values in the nth mode are equal.

3. Results on feasibility and nonfeasibility of the points S, X_1, and Y_1. Throughout this subsection we assume that $\mathcal{T}$ is a norm-1 tensor.

In the following example we show that it may happen that S is the only feasible point in the plane through the points S, X_1, and Y_1, i.e., the plane $\sigma_3^2 = \frac{1}{I_3}$.

Example 3.1. Let $I_3 = I_1 I_2$ and $\mathcal{T} \in \mathbb{C}^{I_1 \times I_2 \times I_3}$. Assume that $\sigma_3^2 = \frac{1}{I_3}$. Then $\mathbf{T}_{(3)}^H \mathbf{T}_{(3)} = \frac{1}{I_3}\mathbf{I}_{I_3}$. Since $\mathbf{T}_{(3)}$ is a square matrix, it follows that $\mathbf{T}_{(3)}$ is a scalar multiple of a unitary matrix, $\mathbf{T}_{(3)} = \frac{1}{\sqrt{I_3}}\mathbf{U}$. One can easily verify (see [5, p. 65]) that $\mathbf{T}_{(1)}^H \mathbf{T}_{(1)} = \frac{1}{I_1}\mathbf{I}_{I_1}$ and $\mathbf{T}_{(2)}^H \mathbf{T}_{(2)} = \frac{1}{I_2}\mathbf{I}_{I_2}$. Hence, $\sigma_1^2 = \frac{1}{I_1}$ and $\sigma_2^2 = \frac{1}{I_2}$. Thus, the points X_1 and Y_1 are not feasible.

From Example 3.1 it follows that the point S is feasible if $I_1 = 2$, $I_2 = 3$, and $I_3 = 6$. The point S is also feasible if $I_1 = 2$, $I_2 = 3$, and $I_3 = 4$. Indeed, let $\mathcal{T}$ be an $2 \times 3 \times 4$ tensor with mode-3 matrix unfolding

$$\mathbf{T}_{(3)} = \frac{1}{2\sqrt{3}} \begin{bmatrix} 1+\sqrt{3} & 0 & 0 & 1-\sqrt{3} & -2 & 0 \\ 0 & 1+\sqrt{3} & 1-\sqrt{3} & 0 & 0 & 2 \\ 0 & 1-\sqrt{3} & 1+\sqrt{3} & 0 & 0 & 2 \\ 1-\sqrt{3} & 0 & 0 & 1+\sqrt{3} & -2 & 0 \end{bmatrix}.$$

Then one can also easily verify that $\mathbf{T}_{(1)}\mathbf{T}_{(1)}^H = \frac{1}{2}\mathbf{I}_2$, $\mathbf{T}_{(2)}\mathbf{T}_{(2)}^H = \frac{1}{3}\mathbf{I}_3$, and $\mathbf{T}_{(3)}\mathbf{T}_{(3)}^H = \frac{1}{4}\mathbf{I}_4$. The following result implies that in the "intermediate" case $I_1 = 2$, $I_2 = 3$, and $I_3 = 5$ the point S is not feasible.

THEOREM 3.2. *Let $I_3 = I_1 I_2 - 1$, $\mathcal{T} \in \mathbb{C}^{I_1 \times I_2 \times I_3}$, and $\mathbf{T}_{(3)}\mathbf{T}_{(3)}^H = \frac{1}{I_3}\mathbf{I}_{I_3}$. Then the following statements hold:*
- (i) *if $\mathbf{T}_{(1)}\mathbf{T}_{(1)}^H = \frac{1}{I_1}\mathbf{I}_{I_1}$, then $I_1 \le I_2$;*
- (ii) *if $\mathbf{T}_{(2)}\mathbf{T}_{(2)}^H = \frac{1}{I_2}\mathbf{I}_{I_2}$, then $I_2 \le I_1$;*
- (iii) *if the point S is feasible, then $I_1 = I_2$.*

Proof.
- (i) Let $\mathbf{T}_{(3)} = [\mathbf{t}_1 \ \dots \ \mathbf{t}_{I_1 I_2}]$. Then the identity $\mathbf{T}_{(1)}\mathbf{T}_{(1)}^H = \frac{1}{I_1}\mathbf{I}_{I_1}$ is equivalent to the system

$$\begin{aligned} &\mathbf{t}_{i_1}^H \mathbf{t}_{i_2} + \mathbf{t}_{I_1+i_1}^H \mathbf{t}_{I_1+i_2} + \cdots + \mathbf{t}_{I_1(I_2-1)+i_1}^H \mathbf{t}_{I_1(I_2-1)+i_2} = 0, \\ &\|\mathbf{t}_{i_1}\|^2 + \|\mathbf{t}_{I_1+i_1}\|^2 + \cdots + \|\mathbf{t}_{I_1(I_2-1)+i_1}\|^2 = \frac{1}{I_1}, \qquad 1 \le i_1 < i_2 \le I_1. \end{aligned} \tag{34}$$

Since $\mathbf{T}_{(3)}\mathbf{T}_{(3)}^H = \frac{1}{I_3}\mathbf{I}_{I_3}$, the matrix $\sqrt{I_3}\mathbf{T}_{(3)} \in \mathbb{C}^{I_3 \times I_1 I_2}$ can be extended to a unitary matrix $\sqrt{I_3}\left[\begin{smallmatrix}\mathbf{T}_{(3)} \\ \mathbf{a}^T\end{smallmatrix}\right] \in \mathbb{C}^{I_1 I_2 \times I_1 I_2}$, where $\mathbf{a} \in \mathbb{C}^{I_1 I_2}$ is a vector such that $\mathbf{T}_{(3)}\mathbf{a}^* = \mathbf{0}$ and $\|\mathbf{a}\|^2 = \frac{1}{I_3}$. Hence,

$$\begin{bmatrix}\mathbf{T}_{(3)}^H & \mathbf{a}^*\end{bmatrix}\begin{bmatrix}\mathbf{T}_{(3)} \\ \mathbf{a}^T\end{bmatrix} = \frac{1}{I_3}\mathbf{I}_{I_2 I_3}$$

or

$$\mathbf{t}_i^H \mathbf{t}_j + \bar{a}_i a_j = 0 \ \text{ for } i \ne j \ \text{ and } \ \|\mathbf{t}_i\|^2 + |a_i|^2 = \frac{1}{I_3}, \quad 1 \le i < j \le I_1 I_2. \tag{35}$$

From (34)–(35) it follows that

$$\bar{a}_{i_1}a_{i_2} + \bar{a}_{I_1+i_1}a_{I_1+i_2} + \cdots + \bar{a}_{I_1(I_2-1)+i_1}a_{I_1(I_2-1)+i_2} = 0,$$
$$|a_{i_1}|^2 + |a_{I_1+i_1}|^2 + \cdots + |a_{I_1(I_2-1)+i_1}|^2 = \frac{1}{I_1}, \qquad 1 \le i_1 < i_2 \le I_1.$$

Thus, the vectors

$$[a_i \ a_{I_1+i} \ \dots \ a_{I_1(I_2-1)+i}]^T \in \mathbb{C}^{I_2}, \qquad 1 \le i \le I_1,$$

are nonzero and mutually orthogonal. Hence, $I_1 \le I_2$.

(ii) The proof is similar to the proof of (i).

(iii) Since S is feasible, it follows that $\mathbf{T}_{(1)}\mathbf{T}_{(1)}^H = \frac{1}{I_1}\mathbf{I}_{I_1}$ and $\mathbf{T}_{(2)}\mathbf{T}_{(2)}^H = \frac{1}{I_2}\mathbf{I}_{I_2}$. Hence, by (i) and (ii), $I_1 = I_2$. ∎

4. The case of at least one equality in (4). The following two lemmas will be used in the proof of Theorem 1.7.

LEMMA 4.1. *Let* $\mathbf{H}$ *and* $\mathbf{\Phi}(\mathbf{H})$ *be as in Lemma* 2.1*. Then the equality in* (17) *holds if and only if* $\mathbf{H}$ *can be factorized as*

$$\mathbf{H} = [\operatorname{vec}(\mathbf{W}_1)\ \mathbf{G}\otimes\mathbf{x}][\operatorname{vec}(\mathbf{W}_1)\ \mathbf{G}\otimes\mathbf{x}]^H, \tag{36}$$

where

(i) $\mathbf{W}_1 \in \mathbb{C}^{I_1\times I_3}$ *and* $\mathbf{x}$ *is a principal eigenvector of* $\mathbf{W}_1\mathbf{W}_1^H$*, i.e.,*

$$\mathbf{W}_1\mathbf{W}_1^H\mathbf{x} = \lambda_{max}(\mathbf{W}_1\mathbf{W}_1^H)\mathbf{x}, \quad \|\mathbf{x}\| = 1;$$

(ii) *the matrix* $\mathbf{G} = [\mathbf{g}_2 \dots \mathbf{g}_R] \in \mathbb{C}^{I_3\times(R-1)}$ *has orthogonal columns;*

(iii) $\mathbf{G}^T\mathbf{W}_1^H\mathbf{x} = \mathbf{0}$*;*

(iv) $\lambda_{max}(\mathbf{W}_1^H\mathbf{W}_1) = \lambda_{max}(\mathbf{W}_1^H\mathbf{W}_1 + \mathbf{G}^*\mathbf{G}^T)$*.*

Moreover, if (36) *and* (i)–(iv) *hold, then*

$$\sigma\left(\sum_{k=1}^{I_3}\mathbf{H}_{kk}\right) = \sigma\left(\mathbf{W}_1\mathbf{W}_1^H + \|\mathbf{G}\|^2\mathbf{x}\mathbf{x}^H\right), \tag{37}$$

$$\sigma(\mathbf{H}) = \left\{\|\mathbf{W}_1\|^2, \|\mathbf{g}_2\|^2, \dots, \|\mathbf{g}_R\|^2, 0, \dots, 0\right\}, \tag{38}$$

$$\sigma(\mathbf{\Phi}(\mathbf{H})) = \sigma\left(\mathbf{W}_1^H\mathbf{W}_1 + \mathbf{G}^*\mathbf{G}^T\right), \tag{39}$$

where $\sigma(\cdot)$ *denotes the spectrum of a matrix.*

Proof. The proof essentially relies on the proof of Lemma 2.1 so we use the same notation and conventions as in the proof of Lemma 2.1.

Derivation of (37)–(39). Assume that 36 and (i)–(iv) hold. Then

$$\mathbf{H} = \sum_{r=1}^{R}\operatorname{vec}(\mathbf{W}_r)\operatorname{vec}(\mathbf{W}_r)^H, \qquad \text{where } \mathbf{W}_r = \mathbf{x}\mathbf{g}_r^T \text{ for } r = 2,\dots,R.$$

Hence

$$\sum_{k=1}^{I_3}\mathbf{H}_{kk} = \sum_{r=1}^{R}\mathbf{W}_r\mathbf{W}_r^H = \mathbf{W}_1\mathbf{W}_1^H + \sum_{r=2}^{R}\mathbf{x}\mathbf{g}_r^T\mathbf{g}_r^*\mathbf{x}^H = \mathbf{W}_1\mathbf{W}_1^H + \|\mathbf{G}\|^2\mathbf{x}\mathbf{x}^H,$$

which implies (37). By (ii), (iii), and the convention $\|x\| = 1$ in (i), the vectors $\mathrm{vec}(\mathbf{W}_r)$ are mutually orthogonal, which implies (38). Finally, by (21),

$$\mathbf{\Phi}(\mathbf{H}) = \sum_{r=1}^{R} \mathbf{W}_r^T\mathbf{W}_r^* = \mathbf{W}_1^T\mathbf{W}_1^* + \sum_{r=1}^{R} \mathbf{g}_r\mathbf{x}^T\mathbf{x}^*\mathbf{g}_r^H = \mathbf{W}_1^T\mathbf{W}_1^* + \mathbf{G}\mathbf{G}^H,$$

which implies (39).

Sufficiency. By (i) and (37),

$$\lambda_{max}\left(\sum_{k=1}^{I_3} \mathbf{H}_{kk}\right) = \lambda_{max}\left(\mathbf{W}_1\mathbf{W}_1^H\right) + \|\mathbf{G}\|^2.$$

By (iv) and (ii),

$$\|\mathbf{W}_1\|^2 \geq \lambda_{max}\left(\mathbf{W}_1^H\mathbf{W}_1\right) \geq \lambda_{max}\left(\mathbf{G}^*\mathbf{G}^T\right) = \max_{2\leq r\leq R} \|\mathbf{g}_r\|^2.$$

Thus, by (38), $\lambda_{max}(\mathbf{H}) = \|\mathbf{W}_1\|^2$ and $\mathrm{tr}(\mathbf{H}) = \|\mathbf{W}_1\|^2 + \|\mathbf{G}\|^2$. By (iv) and (39), $\lambda_{max}(\mathbf{\Phi}(\mathbf{H})) = \lambda_{max}(\mathbf{W}_1^H\mathbf{W}_1)$. Thus, the left- and right-hand sides of (17) are equal to $\lambda_{max}(\mathbf{W}_1\mathbf{W}_1^H) + \|\mathbf{W}_1\|^2 + \|\mathbf{G}\|^2$.

Necessity. It is clear that the equality in (17) holds if and only if it holds in (22) and (23). So we replace the inequality signs in (22) and (23) with an equality sign.

From the first line of (23) it follows that $\mathbf{x}$ satisfies (i). By the Cauchy inequality, the equality

$$\sum_{k=1}^{I_3}\sum_{r=2}^{R} |(\mathbf{w}_{kr}, \mathbf{x})|^2 = \sum_{r=2}^{R} \|\mathbf{w}_r\|^2$$

in (22) would imply that

$$\mathbf{w}_{kr} = c_{kr}\mathbf{x}, \qquad k = 1, \dots, I_3, \quad r = 2, \dots, R,$$

for some $c_{kr} \in \mathbb{C}$. Hence,

$$\mathbf{w}_r = \left[\mathbf{w}_{1r}^T \ \dots \ \mathbf{w}_{I_3r}^T\right]^T = [c_{1r} \ \dots \ c_{I_3r}]^T \otimes \mathbf{x} = \mathbf{g}_r \otimes \mathbf{x}, \qquad r = 2, \dots, R. \tag{40}$$ □

Since $\mathbf{H} = \sum_{r=1}^{R} \mathbf{w}_r\mathbf{w}_r^H$, it follows that

$$\mathbf{H} = [\mathbf{w}_1 \ \dots \ \mathbf{w}_R][\mathbf{w}_1 \ \dots \ \mathbf{w}_R]^H = [\mathbf{w}_1 \ \mathbf{g}_2 \otimes \mathbf{x} \ \dots \ \mathbf{g}_R \otimes \mathbf{x}][\mathbf{w}_1 \ \mathbf{g}_2 \otimes \mathbf{x} \ \dots \ \mathbf{g}_R \otimes \mathbf{x}]^H,$$

which coincides with (36). The mutual orthogonality of $\mathbf{w}_2, \dots, \mathbf{w}_R$ and the orthogonality of $\mathbf{w}_1$ to $\mathbf{w}_2, \dots, \mathbf{w}_R$ imply (ii) and (iii), respectively. By (40), $\mathbf{W}_r = \mathbf{x}\mathbf{g}_r^T$ for $r = 2, \dots, R$. Hence, the equality

$$\lambda_{max}\left(\mathbf{W}_1^H\mathbf{W}_1\right) = \lambda_{max}\left(\sum_{r=1}^{R} \mathbf{W}_r^H\mathbf{W}_r\right)$$

in (23) would imply (iv):

$$\lambda_{max}\left(\mathbf{W}_1^H\mathbf{W}_1\right) = \lambda_{max}\left(\mathbf{W}_1^H\mathbf{W}_1 + \sum_{r=1}^{R} \mathbf{g}_r^*\mathbf{x}^H\mathbf{x}\mathbf{g}_r^T\right) = \lambda_{max}(\mathbf{W}_1^H\mathbf{W}_1 + \mathbf{G}^*\mathbf{G}^T).$$

LEMMA 4.2.

(i) *Let* $\mathbf{W}_1$*,* $\mathbf{G}$*, and* $\mathbf{x}$ *satisfy conditions* (i)–(iv) *of Lemma* 4.1*,* $\mathbf{H}$ *be defined as in* (36)*, and* $L := \lambda_{max}(\mathbf{W}_1^H\mathbf{W}_1)$*. Then there exist* $(I_3-1)\times(I_3-1)$ *positive semidefinite matrices* $\mathbf{A}$ *and* $\mathbf{B}$ *such that*

$$\text{rank}(\mathbf{A}) \le \min(I_1, I_3) - 1, \qquad \text{rank}(\mathbf{B}) = R-1, \qquad L \ge \lambda_{max}(\mathbf{A}+\mathbf{B}) \tag{41}$$

and

$$\sigma\left(\sum_{k=1}^{I_3}\mathbf{H}_{kk}\right) = \left\{L + \text{tr}(\mathbf{B}), \lambda_1(\mathbf{A}), \dots, \lambda_{\min(I_1,I_3)-1}(\mathbf{A}), \underbrace{0,\dots,0}_{I_1-I_3}\right\}, \tag{42}$$

$$\sigma(\mathbf{H}) = \left\{L + \text{tr}(\mathbf{A}), \lambda_1(\mathbf{B}), \dots, \lambda_{R-1}(\mathbf{B}), \underbrace{0,\dots,0}_{I_1I_3-R}\right\}, \tag{43}$$

$$\sigma\left(\mathbf{\Phi}(\mathbf{H})\right) = \{L\} \cup \sigma\left(\mathbf{A}+\mathbf{B}\right). \tag{44}$$

(ii) *Let a positive value* L *and* $(I_3-1)\times(I_3-1)$ *positive semidefinite matrices* $\mathbf{A}$ *and* $\mathbf{B}$ *satisfy* (41)*. Then there exists a matrix* $\mathbf{H}$ *of form* (36) *such that* (42)–(44) *hold.*

Proof.

(i) Let $\mathbf{p}$ be a principal eigenvector of $\mathbf{W}_1\mathbf{W}_1^H$, i.e., $\mathbf{W}_1^H\mathbf{W}_1\mathbf{p} = L\mathbf{p}$, $\|\mathbf{p}\| = 1$. Then, by (iv), $\mathbf{G}^*\mathbf{G}^T\mathbf{p} = 0$. Let $\mathbf{U}_\mathbf{p}$ be an $I_3 \times I_3$ unitary matrix whose first column is $\mathbf{p}$. Then

$$\mathbf{U}_\mathbf{p}^H\mathbf{W}_1^H\mathbf{W}_1\mathbf{U}_\mathbf{p} = \begin{bmatrix} L & \mathbf{0} \\ \mathbf{0} & \mathbf{A}\end{bmatrix}, \qquad \mathbf{U}_\mathbf{p}^H\mathbf{G}^*\mathbf{G}^T\mathbf{U}_\mathbf{p} = \begin{bmatrix} 0 & \mathbf{0} \\ \mathbf{0} & \mathbf{B}\end{bmatrix}, \tag{45}$$

where $\mathbf{A}$ and $\mathbf{B}$ are $(I_3-1)\times(I_3-1)$ positive semidefinite matrices. It is clear that

$$\lambda_k(\mathbf{A}) = \lambda_{k+1}(\mathbf{W}_1^H\mathbf{W}_1), \qquad k = 1, \dots, I_3-1, \tag{46}$$

$$\lambda_k(\mathbf{B}) = \lambda_k(\mathbf{G}^*\mathbf{G}^T) = \begin{cases} \|\mathbf{g}_{k+1}\|^2, & k = 1, \dots, R-1, \\ 0, & k = R, \dots, I_3-1. \end{cases} \tag{47}$$

Now, (43) follows from (38) and (45) and (44) follows from (39) and (45). To prove (42) we rewrite (37) as

$$\sigma\left(\sum_{k=1}^{I_3}\mathbf{H}_{kk}\right) = \left\{\lambda_{max}\left(\mathbf{W}_1\mathbf{W}_1^H\right) + \|\mathbf{G}\|^2, \lambda_2\left(\mathbf{W}_1\mathbf{W}_1^H\right), \dots, \lambda_{I_1}\left(\mathbf{W}_1\mathbf{W}_1^H\right)\right\}. \tag{48}$$

Since the nonzero eigenvalues of $\mathbf{W}_1\mathbf{W}_1^H$ coincide with those of $\mathbf{W}_1^H\mathbf{W}_1$ and $\|\mathbf{G}\|^2 = \text{tr}(\mathbf{B})$ it follows that (48) is equivalent to (42).

(ii) Let $\mathbf{H}$ be defined as in (36), where

$$\mathbf{W}_1 = \begin{bmatrix}\sqrt{L} & \mathbf{0} \\ \mathbf{0} & \widetilde{\mathbf{W}}_1\end{bmatrix}, \qquad \mathbf{G} = \mathbf{U}\mathbf{S}^{\frac{1}{2}}, \qquad \mathbf{x} = [1\ 0\ \dots\ 0]^T,$$

$\widetilde{\mathbf{W}}_1$ is an $(I_1-1)\times(I_3-1)$ matrix such that $\widetilde{\mathbf{W}}_1^H\widetilde{\mathbf{W}}_1=\mathbf{A}$, and $\mathbf{USU}^H$ is the reduced SVD of $\left[\begin{smallmatrix}0 & \mathbf{0}\\ \mathbf{0} & \mathbf{B}\end{smallmatrix}\right]$. One can easily verify that conditions (i)–(iv) in Lemma 4.1 hold. Hence, by Lemma 4.1, (37)–(39) also hold. Substituting $\mathbf{W}_1$, $\mathbf{G}$, and $\mathbf{x}$ in (37)–(39) we obtain (42)–(44). ∎

Proof of Theorem 1.6. Let $\mathbf{H}=\mathbf{T}_{(2)}^T\mathbf{T}_{(2)}^*$. Since $\sigma_1^2+\sigma_2^2=\|\mathcal{T}\|^2+\sigma_3^2$, it follows that equality (17) holds. Hence, by Lemma 4.1, $\mathbf{H}$ can be factorized as in (36). Therefore, there exists an $I_2\times R$ matrix $\mathbf{V}$ whose columns are orthonormal and such that $\mathbf{T}_{(2)}^T=[\mathrm{vec}(\mathbf{W}_1)\ \mathbf{G}\otimes\mathbf{x}]\mathbf{V}^H$, or equivalently,

$$\mathbf{T}_k=[\mathbf{w}_{1k}\ \mathbf{x}[g_{k1}\ \ldots\ g_{kR}]]\mathbf{V}^H,\qquad k=1,\ldots,I_3.$$

Let $\mathcal{W}$ and $\mathcal{G}$ denote the $I_1\times I_2\times I_3$ tensors whose kth frontal slice is $[\mathbf{w}_{1k}\ \mathbf{0}\ \ldots\ \mathbf{0}]\mathbf{V}^H$ and $[\mathbf{0}\ \mathbf{x}[g_{k1}\ \ldots\ g_{kR}]]\mathbf{V}^H$, respectively. It is clear that $\mathcal{T}=\mathcal{W}+\mathcal{G}$, $\mathcal{W}$ is an ML rank-$(L_1,1,L_1)$ tensor, and $\mathcal{G}$ is ML rank-$(1,L_2,L_2)$ tensors, where $L_1\le\min(I_1,I_3)$ and $L_2\le\min(I_2-1,I_3)$. ∎

Proof of Theorem 1.7. Let $\mathbf{H}=\mathbf{T}_{(2)}^T\mathbf{T}_{(2)}^*$. Then

$$\sigma_{11}^2\ge\sigma_{12}^2\ge\cdots\ge\sigma_{1I_1}^2\ge0\qquad\text{are the eigenvalues of }\sum_{i=1}^{I_3}\mathbf{H}_{ii}=\mathbf{T}_{(1)}\mathbf{T}_{(1)}^H,\tag{49}$$

$$\sigma_{21}^2\ge\sigma_{22}^2\ge\cdots\ge\sigma_{2I_2}^2\ge0\qquad\text{are the first }I_2\text{ eigenvalues of }\mathbf{H},\tag{50}$$

$$\sigma_{31}^2\ge\sigma_{32}^2\ge\cdots\ge\sigma_{3I_3}^2\ge0\qquad\text{are the eigenvalues of }\mathbf{\Phi}(\mathbf{H})=\mathbf{T}_{(3)}\mathbf{T}_{(3)}^H.\tag{51}$$

Necessity. By Lemmas 4.1 and 4.2(i), there exist $(I_3-1)\times(I_3-1)$ positive semidefinite matrices $\mathbf{A}$ and $\mathbf{B}$ such that (41)–(44) hold. Thus, by (15) and (49)–(51), the values α_i, β_i, and γ_i are eigenvalues of $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{A}+\mathbf{B}$, respectively. Hence, by Horn's conjecture, (13) and (14) hold.

Sufficiency. Since (13) and (14) hold, from Horn's conjecture it follows that there exist $(I_3-1)\times(I_3-1)$ positive semidefinite matrices $\mathbf{A}$ and $\mathbf{B}$ such that α_i, β_i, and γ_i are eigenvalues of $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{A}+\mathbf{B}$, respectively. Hence, by Lemma 4.2(ii), there exists a matrix $\mathbf{H}$ of form (36) such that (42)–(44) hold. By (15) and (43), $\mathrm{rank}\,\mathbf{H}\le1+R\le I_2$. Let $\mathbf{V}$ be an $I_2\times R$ matrix whose columns are orthonormal and let $\mathcal{T}$ denote an $I_1\times I_2\times I_3$ tensor with mode-2 matrix unfolding $\mathbf{T}_{(2)}=\mathbf{V}^*[\mathrm{vec}(\mathbf{W}_1)\ \mathbf{G}\otimes\mathbf{x}]^T$. Then, $\mathbf{H}=\mathbf{T}_{(2)}^T\mathbf{T}_{(2)}^*$. The proof now follows from (49)–(51). ∎

5. Conclusion.

In the paper we studied geometrical properties of the set

$$\Sigma_{I_1,I_2,I_3}:=\big\{\big(\sigma_{11}^2,\ldots,\sigma_{1I_1}^2,\sigma_{21}^2,\ldots,\sigma_{2I_2}^2,\sigma_{31}^2,\ldots,\sigma_{3I_3}^2\big):$$
$$\sigma_{nk}\text{ is the }k\text{th largest mode-}n\text{ singular value of an }I_1\times I_2\times I_3\text{ norm-1 tensor }\mathcal{T}\big\},$$

where for each $n=1,2,3$ the values σ_{nk} are sorted in descending order.

Let π denote a projection of $\mathbb{R}^{I_1+I_2+I_3}$ onto the first, (I_1+1)th, and (I_1+I_2+1)th coordinates. We have shown that there exist two convex polyhedrons of positive volume such that the set $\pi(\Sigma_{I_1,I_2,I_3})\subset\mathbb{R}^3$ contains one polyhedron (Theorem 1.2) and is contained in another (Theorem 1.1). We have also shown that both polyhedrons coincide for cubic tensors, i.e., for $I_1=I_2=I_3$ (Corollary 1.3), and can be different in the noncubic case (Example 3.1 and Theorem 3.2).

In Theorem 1.7, we considered the case where the largest ML singular values of $\mathcal{T}$ satisfy the equality

$$\sigma_{11}^2+\sigma_{21}^2=1+\sigma_{31}^2\text{ or }\sigma_{11}^2+\sigma_{31}^2=1+\sigma_{21}^2\text{ or }\sigma_{21}^2+\sigma_{31}^2=1+\sigma_{11}^2$$

and described the preimage $\pi^{-1}(\Sigma_{I_1,I_2,I_3})$. The description implies that $\pi^{-1}(\Sigma_{I_1,I_2,I_3})$ is a convex polyhedron. This seems to indicate that the whole set Σ_{I_1,I_2,I_3} is also a convex polyhedron. As the description of $\pi^{-1}(\Sigma_{I_1,I_2,I_3})$ relies on a problem concerning the eigenvalues of the sum of two Hermitian matrices that has long been standing, the complete description of Σ_{I_1,I_2,I_3} could be an even harder problem.

We have also proved a higher-order generalizations of Theorem 1.1 (Theorem 1.4) and Corollary 1.3 (Theorem 1.5).

Appendix A. Definition of T_r^n. In our presentation we follow [1, p. 302].

The set T_r^n of triplets (I, J, K) of cardinality r can be described by induction on r as follows.

Let us write $I = \{i_1 < i_2 < \cdots < i_r\}$ and likewise for J and K. Then for $r = 1$, (I, J, K) is in T_1^n if $k_1 = i_1 + j_1 - 1$. For $r > 1$, (I, J, K) is in T_r^n if

$$\sum_{i\in I} i + \sum_{j\in J} j = \sum_{k\in K} k + \frac{r(r+1)}{2},$$

and, for all $1 \le p \le r-1$ and all $(U, V, W) \in T_p^r$,

$$\sum_{u\in U} i_u + \sum_{v\in V} j_v = \sum_{w\in W} k_w + \frac{p(p+1)}{2}.$$

Thus, T_r^n is defined recursively in terms of $T_1^r, \dots, T_{r-1}^r$.

Acknowledgment. The authors express their gratitude to the mathoverflow.net user with nickname @fedja for help in proving Lemma 2.1.

REFERENCES

[1] R. Bhatia, *Linear Algebra to Quantum Cohomology: The Story of Alfred Horn's Inequalities*, Amer. Math. Monthly, 108 (2001), pp. 289–318.

[2] A. Cichocki, D. Mandic, C. Caiafa, A.-H. Phan, G. Zhou, Q. Zhao, and L. de Lathauwer, *Tensor decompositions for signal processing applications. From two-way to multiway component analysis*, IEEE Signal Process. Mag., 32 (2015), pp. 145–163.

[3] L. de Lathauwer, B. de Moor, and J. Vandewalle, *A multilinear singular value decomposition*, SIAM J. Matrix Anal. Appl., 21 (2000), pp. 1253–1278.

[4] L. de Lathauwer and J. Vandewalle, *Dimensionality reduction in higher-order signal processing and rank-$(R_1, R_2, \dots, R_n)$ reduction in multilinear algebra*, Linear Algebra Appl., 391 (2004), pp. 31–55.

[5] D. Gerard and P. Hoff, *A higher-order LQ decomposition for separable covariance models*, Linear Algebra Appl., 505 (2016), pp. 57–84.

[6] W. Hackbusch, D. Kressner, and A. Uschmajew, *Perturbation of Higher-Order Singular Values*, INS Preprint 1616, 2016.

[7] W. Hackbusch and A. Uschmajew, *On the interconnection between the higher-order singular values of real tensors*, Numer. Math., 135 (2017), pp. 875–894.

[8] A. Horn, *Eigenvalues of sums of Hermitian matrices*, Pacific J. Math., 12 (1962), pp. 225–241.

[9] A. A. Klyachko, *Stable bundles, representation theory and Hermitian operators*, Selecta Math., 4 (1998), pp. 419–445.

[10] A. Knutson and T. Tao, *The honeycomb model of $GL_n(\mathbb{C})$ tensor products* I: *Proof of the saturation conjecture*, J. Amer. Math. Soc., 12 (1999), pp. 1055–1090.

[11] A. Knutson and T. Tao, *Honeycombs and sums of Hermitian matrices*, Notices Amer. Math. Soc., 48 (2001), pp. 175–186.

[12] T. G. Kolda and B. W. Bader, *Tensor decompositions and applications*, SIAM Rev., 51 (2009), pp. 455–500.

[13] S. Krämer, *The Geometrical Description of Feasible Singular Values in the Tensor Train Format*, arXiv:1701.08437, 2017.

[14] P. M. Kroonenberg, *Applied Multiway Data Analysis*, Wiley, Hoboken, NJ, 2008.
[15] A. Seigal, *Gram Determinants of Real Binary Tensors*, arXiv:1612.04420, 2016.
[16] N. D. Sidiropoulos, L. De. Lathauwer, X. Fu, K. Huang, E. E. Papalexakis, and C. Faloutsos, *Tensor decomposition for signal processing and machine learning*, IEEE Trans. Signal Process., 65 (2017), pp. 3551–3582.
[17] L. R. Tucker, *Some mathematical notes on three-mode factor analysis*, Psychometrika, 31 (1966), pp. 279–311.