ON THE LARGEST MULTILINEAR SINGULAR VALUES OF HIGHER-ORDER TENSORS

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

Abstract. Let \( \sigma_n \) denote the largest mode-\( n \) multilinear singular value of an \( I_1 \times \cdots \times I_N \) tensor \( T \). We prove that \( \sigma_1^2 + \cdots + \sigma_{n-1}^2 + \sigma_n^2 + \cdots + \sigma_N^2 \leq (N-2)\|T\|^2 + \sigma_n^2 \), \( n = 1, \ldots, N \), 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 \( \sigma_1, \sigma_2, \) and \( \sigma_3 \) that satisfy \( \sigma_1^2 + \sigma_2^2 \leq \|T\|^2 + \sigma_3^2 \), \( \sigma_1^2 + \sigma_3^2 \leq \|T\|^2 + \sigma_2^2 \), \( \sigma_2^2 + \sigma_3^2 \leq \|T\|^2 + \sigma_1^2 \), and the trivial inequalities \( \sigma_1 \geq \frac{1}{\sqrt{N}}\|T\| \), \( \sigma_2 \geq \frac{1}{\sqrt{N}}\|T\| \), \( \sigma_3 \geq \frac{1}{\sqrt{N}}\|T\| \), there always exists an \( n \times n \times n \) tensor whose largest multilinear singular values are equal to \( \sigma_1, \sigma_2, \) and \( \sigma_3 \). We also show that if the equality \( \sigma_1^2 + \sigma_2^2 = \|T\|^2 + \sigma_3^2 \) holds, then \( T \) is necessarily equal to a sum of multilinear rank-\( (1,1,1) \) and multilinear rank-\( (1,2,2) \) 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 scare.

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 \( \sum_n^m (\cdot) = 0 \) and \( \prod_n^m (\cdot) = 1 \).

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

\[
\text{the} \left( \sum_{k=1 \atop k \neq n}^N (i_k - 1) \prod_{l=1 \atop l \neq n}^{k-1} I_l \right) \text{th entry of } T_{(n)} = \text{the} (i_1, \ldots, i_N) \text{th entry of } T.
\]
For instance, if \( N = 3 \), i.e., \( T \in \mathbb{C}^{I_1 \times I_2 \times I_3} \), then (1) implies that
\[
T_{(1)} = [T_1 \ldots T_{I_3}] \in \mathbb{C}^{I_2 \times I_1 I_3},
\]
\[
T_{(2)} = [T_1^T \ldots T_{I_3}^T] \in \mathbb{C}^{I_1 \times I_2 I_3},
\]
\[
T_{(3)} = [\text{vec}(T_1) \ldots \text{vec}(T_{I_3})]^T \in \mathbb{C}^{I_3 \times I_1 I_2},
\]
where \( T_1, \ldots, T_{I_3} \in \mathbb{C}^{I_1 \times I_2} \) denote the frontal slices of \( T \).

Tensor \( T \in \mathbb{C}^{I_1 \times \cdots \times I_N} \) is all-orthogonal if the matrices \( T_{(1)} T_{(1)}^H, \ldots, T_{(N)} T_{(N)}^H \) are diagonal. The multiLinear (ML) singular value decomposition (SVD) (aka higher-order SVD) is a factorization of \( T \) into the product of an all-orthogonal tensor \( S \in \mathbb{C}^{I_1 \times \cdots \times I_N} \) and \( N \) unitary matrices \( U_1 \in \mathbb{C}^{I_1 \times I_1}, \ldots, U_N \in \mathbb{C}^{I_N \times I_N}, \)
\[
T = S_{(1)} U_1 : U_2 : \cdots : U_N N,
\]
where “\(^n\)” denotes the \( n \)-mode product of \( S \) and \( U_n \). Rather than giving the formal definition of “\(^n\)”, for which we refer the reader to [2, 3, 12], we present \( N \) equivalent matricized versions of (2):
\[
T_{(n)} = U_n S_{(n)} (U_N \otimes \cdots \otimes U_{n+1} \otimes U_{n-1} \otimes \cdots \otimes U_1)^T, \quad n = 1, \ldots, N,
\]
where “\( \otimes \)” denotes the Kronecker product. For \( N = 2 \), i.e., for \( T = T_1 \in \mathbb{C}^{I_1 \times I_2} \), the MLSVD reduces up to trivial indeterminacies, to the classical SVD of a matrix, \( T_{(1)} = T_1 = USV^H \), where \( U = U_1, S = S_{(1)} \), and \( V = U_2^* \otimes 1 \). 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 \( T_{(n)} \), are called the mode-\( n \) singular values of \( T \). Since \( S_{(1)} S_{(1)}^H, \ldots, S_{(N)} S_{(N)}^H \) are diagonal, it follows from (3) that the ML singular values of \( T \) coincide with the ML singular values of \( S \), which are just the Frobenius norms of the rows of \( S_{(1)}, \ldots, S_{(N)} \). Throughout the paper,
\[
\sigma_n \quad \text{denotes the largest singular value of } T_{(n)}.
\]

In the matrix case, i.e., for \( N = 2 \), the description of MLSVD is trivial. Indeed, the singular values of \( T_{(1)} = T_1 \) and \( T_{(2)} = T_1^T \) coincide and \( T_{(3)} = \text{vec}(T_1)^T \) has a single singular value \( ||T|| \). Thus, the singular values of \( T_{(1)} \) completely define the singular values of \( T_{(2)} \) and \( T_{(3)} \). 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, \ldots, 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 \( \sigma_1, \ldots, \sigma_N \). Our results are valid for real and complex tensors. The following theorem presents simple necessary conditions for \( \sigma_1, \sigma_2, \) and \( \sigma_3 \) 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 $\sigma_1$, $\sigma_2$, and $\sigma_3$ denote the largest ML singular values of an $I_1 \times I_2 \times I_3$ tensor $T$. Then

\begin{align*}
(4) & \quad \sigma_1^2 + \sigma_2^2 \leq \|T\|^2 + \sigma_3^2, \quad \sigma_1^2 + \sigma_3^2 \leq \|T\|^2 + \sigma_2^2, \quad \sigma_2^2 + \sigma_3^2 \leq \|T\|^2 + \sigma_1^2,
(5) & \quad \sigma_1 \geq \frac{1}{\sqrt{I_1}} \|T\|, \quad \sigma_2 \geq \frac{1}{\sqrt{I_2}} \|T\|, \quad \sigma_3 \geq \frac{1}{\sqrt{I_3}} \|T\|.
\end{align*}

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

One can easily verify that if $\sigma_1$, $\sigma_2$ and $\sigma_3$ satisfy (4)-(5) for $I_1 = I_2 = I_3 = 2$ and $\|T\| = 1$, then $\sigma_1$, $\sigma_2$, and $\sigma_3$ are the largest ML singular values of the $2 \times 2 \times 2$ tensor $T$ with mode-1 matrix unfolding

$$T_{(1)} = [T_1 \ T_2] = \begin{bmatrix}
\sqrt{\sigma_1^2 + \sigma_2^2 + \sigma_3^2 - 1} / \sqrt{2} & 0 & \sqrt{1 + \sigma_1^2 - \sigma_2^2 - \sigma_3^2} / \sqrt{2} \\
0 & \sqrt{1 + \sigma_2^2 - \sigma_1^2 - \sigma_3^2} / \sqrt{2} & 0 \\
\sqrt{1 + \sigma_3^2 - \sigma_1^2 - \sigma_2^2} / \sqrt{2} & 0 & 0
\end{bmatrix}.$$ 

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

**Theorem 1.2.** Let $I_1 \leq I_2 \leq I_3$ and $\sigma_1$, $\sigma_2$, $\sigma_3$ satisfy (4) and the following three inequalities:

\begin{align*}
(6) & \quad \sigma_1 \geq \frac{1}{\sqrt{I_1}} \|T\|, \\
(7) & \quad (I_2 - I_1)\sigma_1^2 + (I_1I_2 - I_2)\sigma_3^2 + (1 - I_2)\|T\|^2 \geq 0, \\
(8) & \quad (I_2 - I_1)\sigma_1^2 + (I_1I_2 - I_2)\sigma_2^2 + (1 - I_2)\|T\|^2 \geq 0.
\end{align*}

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

1. all entries of $T$ are nonnegative;
2. $T$ is all-orthogonal;
3. the largest ML singular values of $T$ are equal to $\sigma_1$, $\sigma_2$, and $\sigma_3$.

Conditions (5) and (6)-(8) mean that the point $(\sigma_1^2, \sigma_2^2, \sigma_3^2)$ belongs to the trihedral angles $S_X Y_1 Z_1$ and $S_2 X_2 Y_2 Z_2$, respectively, where $S_2$ has coordinates $(\frac{1}{I_1}, \frac{1}{I_2}, \frac{1}{I_3})$. The gap between the necessary conditions in Theorem 1.1 and the sufficient conditions in Theorem 1.2, i.e., the set

\begin{align*}
(9) & \quad \{(\sigma_1^2, \sigma_2^2, \sigma_3^2) : \text{(4)-(5) hold and at least one of (6)-(8) does not hold}\},
\end{align*}

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 $\sigma_1$, $\sigma_2$, and $\sigma_3$ satisfy (4)-(5) for $I_1 = I_2 = I_3 = I \geq 2$.

Then there exists an $I \times I \times I$ tensor $T$ such that

1. all entries of $T$ are nonnegative;
2. $T$ is all-orthogonal;
3. the largest ML singular values of $T$ are equal to $\sigma_1$, $\sigma_2$, and $\sigma_3$.

Thus, the conditions in Theorem 1.1 are not only necessary but also sufficient for $\sigma_1$, $\sigma_2$, and $\sigma_3$ 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.
Fig. 1. The typical shapes of the set \( \{ (\sigma_1^2, \sigma_2^2, \sigma_3^2) : \sigma_1, \sigma_2, \sigma_3 \text{ satisfy (4)-(5)} \} \) for \( I_1 \leq I_2 \leq I_3 \) (drawn for \( I_1 = 2, I_2 = 3, I_3 = 5 \), and \( \| T \| = 1 \)). 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 \((\frac{1}{2}, \frac{1}{2}, \frac{1}{2}), (1 - \frac{1}{2I_1}, \frac{1}{2I_1}, \frac{1}{2J_1}), (\frac{1}{2}, 1 - \frac{1}{2I_1}, \frac{1}{2I_1}), (\frac{1}{2}, \frac{1}{2I_1}, 1 - \frac{1}{2J_1}), (\frac{1}{2}, \frac{1}{2J_1}, \frac{1}{2I_1}), \) and \((1, 1, 1)\)

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 \( S(\frac{1}{I_1}, \frac{1}{I_2}, \frac{1}{I_3}), X_1(1 - \frac{1}{I_2}, \frac{1}{I_1}, \frac{1}{I_3}), \) and \( Y_1(\frac{1}{I_1}, 1 - \frac{1}{I_2}, \frac{1}{I_3})\). 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 \( I_2 = I_1 I_3 \), 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 \( T \in \mathbb{C}^{I_1 \times \cdots \times I_N} \) is \textit{scaled all-orthonormal} [5, Definition 2] if at least \( N - 1 \) of the \( N \) matrices \( T_{(1)}^H T_{(1)}^H, \ldots, T_{(N)}^H T_{(N)}^H \) are multiples of the identity matrix. It is clear that if the largest mode-\( n \) singular value of a norm-1 tensor is \( \frac{1}{\sqrt{n}} \), then
The following results generalize Theorem 1.1 and Corollary 1.3 for $N$th-order tensors.

**Theorem 1.4.** Let $\sigma_1, \ldots, \sigma_N$ denote the largest ML singular values of an $I_1 \times \cdots \times I_N$ tensor $T$. Then

\[
\sigma_1^2 + \cdots + \sigma_{n-1}^2 + \sigma_{n+1}^2 + \cdots + \sigma_N^2 \leq (N - 2)\|T\|^2 + \sigma_n^2, \quad n = 1, \ldots, N,
\]

and

\[
\|T\| \geq \sigma_1 \geq \frac{1}{\sqrt{I_1}} \|T\|, \ldots, \|T\| \geq \sigma_N \geq \frac{1}{\sqrt{I_N}} \|T\|.
\]
THEOREM 1.5. Let \( \sigma_1, \ldots, \sigma_N \) satisfy (10)–(11) for \( I_1 = \cdots = I_N = I \geq 2 \). Then there exists an \( I \times \cdots \times I \) tensor \( T \) such that

1. all entries of \( T \) are nonnegative;
2. \( T \) is all-orthogonal;
3. the largest ML singular values of \( T \) are equal to \( \sigma_1, \ldots, \sigma_N \).

Thus, the conditions in Theorem 1.4 are not only necessary but also sufficient for \( \sigma_1, \ldots, \sigma_N \) 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 \( \sigma_1 = \sigma_2 = \sigma_3 = \|T\| \) implying that \( T_{(1)}, T_{(2)}, \) and \( T_{(3)} \) are rank-1 matrices. Hence all the remaining ML singular values of \( T \) are zero. Similarly, the two equalities \( \sigma_1^2 + \sigma_2^2 = \|T\|^2 + \sigma_3^2 \) and \( \sigma_1^2 + \sigma_2^2 = \|T\|^2 + \sigma_3^2 \) are equivalent to \( \sigma_1 = \|T\| \) and \( \sigma_2 = \sigma_3 \), implying that \( \text{rank}(T_{(1)}) = 1 \) and \( \text{rank}(T_{(2)}) = \text{rank}(T_{(3)}) = : L \), i.e., \( T \) 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 \( T \) also coincide and may take any positive values whose squares sum up to \( \|T\|^2 - \sigma_3^2 \). In section 4 we characterize the tensors \( T \) for which the single equality \( \sigma_1^2 + \sigma_2^2 = \|T\|^2 + \sigma_3^2 \) holds. We show that \( T \) 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 \( A \) and \( B \), what are all the possible eigenvalues of \( A + B \)? 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

\[ \lambda_i(\cdot) \]

denote the \( i \)th largest eigenvalue of a Hermitian matrix.

If

\[ \alpha_i = \lambda_i(A), \quad \beta_i = \lambda_i(B), \quad \gamma_i = \lambda_i(A + B), \]

then \( \alpha_i, \beta_i, \) and \( \gamma_i \) satisfy the trivial equality

\[ \gamma_1 + \cdots + \gamma_n = \alpha_1 + \cdots + \alpha_n + \beta_1 + \cdots + \beta_n \]

and the list of linear inequalities

\[ \sum_{k \in K} \gamma_k \leq \sum_{i \in I} \alpha_i + \sum_{j \in J} \beta_j, \quad (I, J, K) \in T^n, \quad 1 \leq r \leq n - 1, \]

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

We have the following results.

THEOREM 1.6. Let \( \sigma_1^2 + \sigma_2^2 = \|T\|^2 + \sigma_3^2 \). Then \( T \) 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 $\sigma_1^2 + \sigma_2^2 = \|T\|^2 + \sigma_3^2$. Then the values

$$
\begin{align*}
\sigma_1 &= \sigma_{11} \geq \sigma_{12} \geq \cdots \geq \sigma_{1i} \geq 0, \\
\sigma_2 &= \sigma_{21} \geq \sigma_{22} \geq \cdots \geq \sigma_{2i} \geq 0, \\
\sigma_3 &= \sigma_{31} \geq \sigma_{32} \geq \cdots \geq \sigma_{3i} \geq 0,
\end{align*}
$$

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

$$
\sigma_{11}^2 + \cdots + \sigma_{1i}^2 = \sigma_{21}^2 + \cdots + \sigma_{2i}^2 = \sigma_{31}^2 + \cdots + \sigma_{3i}^2 = \|T\|^2,
$$

and (13) and (14) hold for

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

and $n = I_3 - 1$.

Example 1.8. If $n = 2$, then $T^2 = \{(i, j, k) : k = i + j - 1, 1 \leq i, j, k \leq 2\} = \{(1, 1, 1), (1, 2, 2), (2, 1, 2)\}$ (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)

$$
\begin{align*}
\gamma_1 &\leq \alpha_1 + \beta_1, & \gamma_2 &\leq \alpha_1 + \beta_2, & \gamma_2 &\leq \alpha_2 + \beta_1
\end{align*}
$$

characterize the values $\alpha_1, \alpha_2, \beta_1, \beta_2, \gamma_1, \gamma_2$ that can be eigenvalues of $2 \times 2$ Hermitian matrices $A$, $B$, and $A + B$. Let $\sigma_1^2 + \sigma_2^2 = \|T\|^2 + \sigma_3^2$. 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 $T$ if and only if

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

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 $H = (H_{ij})_{i,j=1}^{I_3} \in C^{I_5 \times I_5 \times I_1}$ be a positive semidefinite matrix consisting of the blocks $H_{ij} \in C^{I_1 \times I_1}$. Then

$$
\lambda_{\max}(H_{11} + \cdots + H_{1k}) + \lambda_{\max}(H) \leq \text{tr}(H) + \lambda_{\max}(\Phi(H)),
$$

where $\Phi(H)$ denotes the $I_3 \times I_3$ matrix with the entries $(\Phi(H))_{ij} = \text{tr}(H_{ij})$ 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 $H = \sum_{r=1}^{R} w_r w_r^H$, where $w_r$ are
orthogonal and $w_r = [w_{1r}^T \ldots w_{Ir}^T]^T$ with $w_{kr} \in \mathbb{C}^{I_3}$. First, we rewrite (17) in terms of $w_{kr}$, $1 \leq k \leq I_3$, $1 \leq r \leq R$. Without loss of generality, we can assume that $\|w_1\| = \max_r \|w_r\|$. Hence,

$$\lambda_{\text{max}}(H) = \|w_1\|^2 = \sum_{k=1}^{I_3} \|w_{k1}\|^2. \tag{18}$$

It is clear that

$$H_{ij} = \sum_{r=1}^{R} w_{ir} w_{jr}^H, \quad 1 \leq i, j \leq I_3. \tag{19}$$

Hence

$$\lambda_{\text{max}}(H_{11} + \ldots + H_{I_3I_3}) = \max_{\|x\|=1} \sum_{k=1}^{I_3} (H_{kk} x, x) = \max_{\|x\|=1} \sum_{k=1}^{I_3} \sum_{r=1}^{R} |(w_{kr}, x)|^2. \tag{20}$$

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

$$\text{tr}(H) = \sum_{r=1}^{R} \|w_r\|^2. \tag{21}$$

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

$$\lambda_{\text{max}}(H) + \lambda_{\text{max}}(H_{11} + \ldots + H_{I_3I_3})$$

$$= \|w_1\|^2 + \max_{\|x\|=1} \left[ \sum_{k=1}^{I_3} |(w_{k1}, x)|^2 + \sum_{k=1}^{I_3} \sum_{r=2}^{R} |(w_{kr}, x)|^2 \right]$$

$$\leq \|w_1\|^2 + \max_{\|x\|=1} \left[ \sum_{k=1}^{I_3} |(w_{k1}, x)|^2 \right] + \sum_{r=2}^{R} \|w_r\|^2 = \text{tr}(H) + \max_{\|x\|=1} \left[ \sum_{k=1}^{I_3} |(w_{k1}, x)|^2 \right].$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.
To complete the proof of (17) we should show that

$$\max_{\|x\|=1} \sum_{k=1}^{I_3} \|w_{k1}^T x\|^2 \leq \lambda_{\text{max}}(\Phi(H)).$$

This can be done as follows:

$$\max_{\|x\|=1} \sum_{k=1}^{I_3} \|w_{k1}^T x\|^2 = \max_{\|x\|=1} \sum_{k=1}^{I_3} x^H w_{k1}^T w_{k1}^H x = \lambda_{\text{max}}(W_1 W_1^H)$$

$$= \lambda_{\text{max}}(W_1 W_1^H) \leq \lambda_{\text{max}}(\sum_{r=1}^{R} W_r^H W_r)$$

$$= \lambda_{\text{max}}(\Phi(H)^*) = \lambda_{\text{max}}(\Phi(H)).$$

Now we are ready to prove Theorem 1.1.

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

By definition of ML singular values,

$$\sigma_1^2 = \lambda_{\text{max}}(\mathbf{T}(1)^H \mathbf{T}(1)) = \lambda_{\text{max}}(\mathbf{T}_1^H \mathbf{T}_1^H + \cdots + \mathbf{T}_{I_3}^H \mathbf{T}_{I_3}^H),$$

$$\sigma_2^2 = \lambda_{\text{max}}(\mathbf{T}(2)^H \mathbf{T}(2)) = \lambda_{\text{max}}(\mathbf{T}_{I_2}^T \mathbf{T}_{I_2}^T) = \lambda_{\text{max}}(\Phi(H)),$$

where

$$H = \mathbf{T}(2)^T \mathbf{T}(2) = \begin{bmatrix} T_{I_1}^H & \cdots & T_{I_1}^H \\ \vdots & \ddots & \vdots \\ T_{I_3}^H & \cdots & T_{I_3}^H \end{bmatrix}.$$

Since vec($\mathbf{T}(1)^T$ vec($\mathbf{T}(2)^T$)) = tr($\mathbf{T}_1^H$), it follows that

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

where

$$\Phi(H) = \begin{bmatrix} \text{tr} (\mathbf{T}_1^H) & \cdots & \text{tr} (\mathbf{T}_1^H) \\ \vdots & \ddots & \vdots \\ \text{tr} (\mathbf{T}_{I_3}^H) & \cdots & \text{tr} (\mathbf{T}_{I_3}^H) \end{bmatrix}.$$

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

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

which holds by Lemma 2.1. \hfill \Box

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 $S_2$, $X_3$, $Y_2$, $Z_2$, and $\mathcal{N}$ whose squared largest ML singular values are the coordinates of $S_2(I_1^\perp, I_2^\perp, I_3^\perp)$, $X_3(I_1^\perp, I_2^\perp, I_3^\perp)$, $Y_2(I_1^\perp, I_2^\perp, I_3^\perp)$, $Z_2(I_1^\perp, I_2^\perp, I_3^\perp)$, and $\mathcal{N}(I_1^\perp, I_2^\perp, I_3^\perp)$, respectively.
Then we show that because of the zero patterns of $S_2$, $X_2$, $Y_2$, $Z_2$, and $N$, the tensor

$$\mathcal{T} = (t_{S_2} S_2^2 + t_{X_2} X_2^2 + t_{Y_2} Y_2^2 + t_{Z_2} Z_2^2 + t_N N^2)^{\frac{1}{2}}$$

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 “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 $\sigma_1^2$, $\sigma_2^2$, and $\sigma_3^2$.

**Step 1.** Let $\pi$ denote the cyclic permutation $\pi : 1 \to I_1 \to I_1 - 1 \to \cdots \to 2 \to 1$. The tensors $S_2$, $X_2$, $Y_2$, and $Z_2$ are defined by

$$S_{2,ijk} = \begin{cases} 
\frac{1}{I_1} & \text{if } j = \pi^{k-1}(i) \text{ and } 1 \leq i, k \leq I_1, \\
0 & \text{otherwise},
\end{cases}$$

$$X_{2,ijk} = \begin{cases} 
\frac{1}{\sqrt{I_2}} & \text{if } j = \pi^{k-1}(i), \ i = 1, \text{ and } 1 \leq k \leq I_1, \\
0 & \text{otherwise},
\end{cases}$$

$$Y_{2,ijk} = \begin{cases} 
\frac{1}{\sqrt{I_2}} & \text{if } i = 1 \text{ and } I_1 < j = k \leq I_2, \\
0 & \text{otherwise},
\end{cases}$$

$$Z_{2,ijk} = \begin{cases} 
\frac{1}{\sqrt{I_1}} & \text{if } j = \pi^{k-1}(i), \ j = 1, \text{ and } 1 \leq k \leq I_1, \\
0 & \text{otherwise},
\end{cases}$$

and the tensor $N$, by definition, has only one nonzero entry, $N_{111} = 1$. For instance, if $I_1 = I_2 = I_3 = 2$, then the first matrix unfoldings of $S_2$, $X_2$, $Y_2$, $Z_2$, and $N$ have the form

$$S_{2,(1)} = \frac{1}{2} \begin{bmatrix} 1 & 0 & 0 & 1 \\
0 & 1 & 1 & 0
\end{bmatrix}, \quad X_{2,(1)} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0
\end{bmatrix},$$

$$Y_{2,(1)} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 0 & 0 \\
0 & 0 & 1 & 0
\end{bmatrix}, \quad Z_{2,(1)} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 & 0 & 1 \\
0 & 0 & 0 & 0
\end{bmatrix}, \quad N_{(1)} = \begin{bmatrix} 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0
\end{bmatrix}.$$

**Step 2.** It is clear that the $(i, j, k)$th entry of a linear combination of $S_2^2$, $X_2^2$, $Y_2^2$, $Z_2^2$, and $N^2$ may be nonzero only if

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

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

**Step 3.** From the construction of the all-orthogonal tensors $S_2$, $X_2$, $Y_2$, $Z_2$, and $N$ 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 $t_{S_2}, t_{X_2}, t_{Y_2},$ and $t_{Z_2}$ explicitly. We set

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

If $(\sigma_1^2, \sigma_2^2, \sigma_3^2)$ belongs to the tetrahedron $X_2 Y_2 Z_2 N$, i.e., $f(\sigma_1^2, \sigma_2^2, \sigma_3^2) \geq 0$, then

$$t_{X_2} = \frac{I_2}{2(I_2 - 1)}(1 + \sigma_1^2 - \sigma_2^2 - \sigma_3^2), \quad t_{Y_2} = \frac{I_1}{2(I_1 - 1)}(1 + \sigma_2^2 - \sigma_1^2 - \sigma_3^2),$$

$$t_{Z_2} = \frac{I_1}{2(I_1 - 1)}(1 + \sigma_3^2 - \sigma_1^2 - \sigma_2^2),$$

$$t_N = 1 - t_{X_2} - t_{Y_2} - t_{Z_2} = \frac{f(\sigma_1^2, \sigma_2^2, \sigma_3^2)}{2(I_1 - 1)(I_2 - 1)}, \quad t_{S_2} = 0.$$

If $(\sigma_1^2, \sigma_2^2, \sigma_3^2)$ belongs to the tetrahedron $X_2 Y_2 Z_2 S_2$, i.e., $f(\sigma_1^2, \sigma_2^2, \sigma_3^2) \leq 0$, then

$$t_{X_2} = \frac{I_1}{I_1 - 1} \left( \sigma_1^2 - \frac{1}{I_1} \right),$$

$$t_{Y_2} = \frac{I_1}{I_1 - 1} \left( \sigma_2^2 - \frac{1}{I_1} \right) + \frac{(I_2 - I_1)I_1}{(I_1 - 1)I_2} \left( \sigma_1^2 - \frac{1}{I_1} \right),$$

$$t_{Z_2} = \frac{I_1}{I_1 - 1} \left( \sigma_3^2 - \frac{1}{I_1} \right) + \frac{(I_2 - I_1)I_1}{(I_1 - 1)I_2} \left( \sigma_1^2 - \frac{1}{I_1} \right),$$

$$t_{S_2} = 1 - t_{X_2} - t_{Y_2} - t_{Z_2} = \frac{-f(\sigma_1^2, \sigma_2^2, \sigma_3^2) I_1}{I_2(I_1 - 1)^2}, \quad t_N = 0.$$

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

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

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 $T$ into third-order tensors $T^{[1]}, \ldots, T^{[N-2]}$ 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 $k$

$$\text{the (i, j, (k - 1))th entry of } \mathcal{X}_{[1]} = \text{the (j, i + (k - 1)I)th entry of } \mathcal{X}_{[2]}$$

$$= \text{the (k, i + (j - 1)I)th entry of } \mathcal{X}_{[3]} = \text{the (i, j, k)th entry of } \mathcal{X}.$$ \tag{26}

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

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

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

is equal to the $(i_1, \ldots, i_N)$th entry of $T$.\[\Box\]
Now we apply (26) for $\mathcal{X} = T^{[n]}$ and

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

After simple algebraic manipulations, we obtain that

\[
\begin{align*}
\text{the} & \left( i_1 + \sum_{k=2}^{n} (i_k - 1) \prod_{l=1}^{k-1} I_l, i_{n+1} + \sum_{k=n+2}^{N} (i_k - 1) \prod_{l=n+1}^{k-1} I_l \right) \text{th entry of } T^{[n]}_{(1)} \\
= & \left( i_{n+1}, 1 + \sum_{k=2}^{n} (i_k - 1) \prod_{l=1}^{k-1} I_l \right) \text{th entry of } T^{[n]}_{(2)} \\
= & \left( i_{n+2} + \sum_{k=n+3}^{N} (i_k - 1) \prod_{l=n+2}^{k-1} I_l, i_1 + \sum_{k=2}^{n+1} (i_k - 1) \prod_{l=1}^{k-1} I_l \right) \text{th entry of } T^{[n]}_{(3)}
\end{align*}
\]

(27) = the $(i_1, \ldots, i_N)$th entry of $T$.

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

\[
\begin{align*}
T^{[1]}_{(1)} & = T_{(1)}, \\
T^{[n]}_{(2)} & = T_{(n+1)}, \quad 1 \leq n \leq N - 2, \\
T^{[N-2]}_{(3)} & = T_{(N)}.
\end{align*}
\]

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

\[
T^{[n]}_{(3)} = \left(T^{[n+1]}_{(1)}\right)^T, \quad 1 \leq n \leq N - 3.
\]

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

\[
\sigma_{\text{max}}^2\left(T^{[n]}_{(1)}\right) + \sigma_{\text{max}}^2\left(T^{[n]}_{(2)}\right) \leq \|T^{[n]}\|^2 + \sigma_{\text{max}}^2\left(T^{[n]}_{(3)}\right) = \|T\|^2 + \sigma_{\text{max}}^2\left(T^{[n]}_{(1)}\right),
\]

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

\[
\begin{align*}
\sigma_1^2 + \sigma_2^2 & \leq \|T\|^2 + \sigma_{\text{max}}^2\left(T^{[1]}_{(3)}\right) = \|T\|^2 + \sigma_{\text{max}}^2\left(T^{[2]}_{(1)}\right), \quad n = 1, \\
\sigma_{\text{max}}^2\left(T^{[2]}_{(1)}\right) + \sigma_3^2 & \leq \|T\|^2 + \sigma_{\text{max}}^2\left(T^{[2]}_{(3)}\right) = \|T\|^2 + \sigma_{\text{max}}^2\left(T^{[3]}_{(1)}\right), \quad n = 2, \\
& \vdots \\
\sigma_{\text{max}}^2\left(T^{[N-3]}_{(1)}\right) + \sigma_{N-2}^2 & \leq \|T\|^2 + \sigma_{\text{max}}^2\left(T^{[N-3]}_{(3)}\right) = \|T\|^2 + \sigma_{\text{max}}^2\left(T^{[N-2]}_{(1)}\right), \quad n = N - 3, \\
\sigma_{\text{max}}^2\left(T^{[N-2]}_{(1)}\right) + \sigma_{N-1}^2 & \leq \|T\|^2 + \sigma_{\text{max}}^2\left(T^{[N-2]}_{(3)}\right) = \|T\|^2 + \sigma_N^2, \quad n = N - 2.
\end{align*}
\]

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 points in $2N - N$ points,

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

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

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

Let $\pi$ denote the cyclic permutation $\pi : 1 \rightarrow I \rightarrow I - 1 \rightarrow \cdots \rightarrow 2 \rightarrow 1$. The tensor $P^{\pi_{1 \rightarrow \cdots \rightarrow 1}}$ is defined by

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

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

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

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

$$P^{\pi_{1 \rightarrow \cdots \rightarrow 1}}_{1} = \frac{1}{2\sqrt{2}} \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 1 \end{bmatrix}$$

and the first matrix unfoldings of the remaining tensors $P^{\alpha_1, \ldots, \alpha_N}$ can be obtained from $P^{\pi_{1 \rightarrow \cdots \rightarrow 1}}$ by rescaling and introducing more zeros.

It is clear that the $(i_1, \ldots, i_N)$th entry of a linear combination of $P^{\pi_{1 \rightarrow \cdots \rightarrow 1}}, \ldots, P^{\pi_{1 \rightarrow \cdots \rightarrow 1}}$ may be nonzero only if

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

The same is also true for $T$ defined by

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

where, as before, the superscripts "$\circ 2"$ and "$\circ 1$" denote the entrywise operations. One can easily check that each column of $T_{(1)}, \ldots, T_{(N)}$ contains at most one nonzero entry, implying that $T$ is all-orthogonal tensor. Finally, from the construction of the all-orthogonal tensors $P^{\alpha_1, \ldots, \alpha_N}$ 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 $T$ whenever the values $t_{\alpha_1, \ldots, \alpha_N}$ 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 $V$. $\square$

Note that in the proof of Theorem 1.5 the constructed tensor $T$ has squared singular values in the $n$th mode equal to $\sigma_n^2, \frac{1}{I-1}(1 - \sigma_n^2), \ldots, \frac{1}{I-1}(1 - \sigma_n^2)$, i.e., the $I - 1$ smallest singular values in the $n$th mode are equal.
3. Results on feasibility and nonfeasibility of the points $S$, $X_1$, and $Y_1$.

Throughout this subsection we assume that $T$ 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 $\sigma^2_3 = \frac{1}{I_3}$.

**Example 3.1.** Let $I_3 = I_1 I_2$ and $T \in \mathbb{C}^{I_1 \times I_2 \times I_3}$. Assume that $\sigma^2_3 = \frac{1}{I_3}$. Then $T^{H}_3 T_3 = \frac{1}{I_3} I_3$. Since $T_3$ is a square matrix, it follows that $T_3$ is a scalar multiple of a unitary matrix, $T_3 = \frac{1}{\sqrt{I_3}} U$. One can easily verify (see [5, p. 65]) that $T^{H}_{(1)} T_{(1)} = \frac{1}{I_1} I_1$ and $T^{H}_{(2)} T_{(2)} = \frac{1}{I_2} I_2$. Hence, $\sigma^2_1 = \frac{1}{I_2}$ 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 $I_3 = 3$. Indeed, let $T$ be an $2 \times 3 \times 4$ tensor with mode-3 matrix unfolding

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

Then one can also easily verify that $T_{(1)} T^{H}_{(1)} = \frac{1}{I_1} I_1$, $T_{(2)} T^{H}_{(2)} = \frac{1}{I_2} I_2$, and $T_{(3)} T^{H}_{(3)} = \frac{1}{I_3} I_3$. The following result implies that in the “intermediate” case $I_1 = 2$, $I_2 = 3$, and $I_3 = 3$ the point $S$ is not feasible.

**Theorem 3.2.** Let $I_3 = I_1 I_2 - 1$, $T \in \mathbb{C}^{I_1 \times I_2 \times I_3}$, and $T_{(3)} T^{H}_{(3)} = \frac{1}{I_3} I_3$. Then the following statements hold:

(i) if $T_{(1)} T^{H}_{(1)} = \frac{1}{I_1} I_1$, then $I_1 \leq I_2$;

(ii) if $T_{(2)} T^{H}_{(2)} = \frac{1}{I_2} I_2$, then $I_2 \leq I_3$;

(iii) if the point $S$ is feasible, then $I_1 = I_2$.

**Proof.**

(i) Let $T_{(3)} = [t_1 \ldots t_{I_1 I_2}]$. Then the identity $T_{(1)} T^{H}_{(1)} = \frac{1}{I_1} I_1$ is equivalent to the system

$$
\|t_{i_1}\|^2 + \|t_{i_1+i_2}\|^2 + \cdots + \|t_{i_1(I_2-1)+i_2}\|^2 = \frac{1}{I_1}, \quad 1 \leq i_1 < i_2 \leq I_1.
$$

(34)

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

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

or

$$
(35) \ t^H_{i_1} t_{j} + a_i a_j = 0 \quad \text{for } i \neq j \quad \text{and } \|t_{i}\|^2 + |a_i|^2 = \frac{1}{I_3}, \quad 1 \leq i < j \leq I_1 I_2.
$$
From (34)–(35) it follows that
\[ a_1 a_2 + a_1 a_2 + \cdots + a_1 a_2 = 0, \]
\[ |a_i| + |a_i + i| + \cdots + |a_i + i| = \frac{1}{I_1}, \quad 1 \leq i < i_2 \leq I_1. \]

Thus, the vectors
\[ [a_i a_i \cdots a_i + i_1 + i_2] \in \mathbb{C}, \quad 1 \leq i \leq I_1, \]
are nonzero and mutually orthogonal. Hence, \( I_1 \leq I_2 \).

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

(iii) Since \( S \) is feasible, it follows that \( T_{(1)} T_{(1)}^H = \frac{1}{T_1} I_{T_1} \) and \( T_{(2)} T_{(2)}^H = \frac{1}{T_2} I_{T_2} \).

Hence, by (i) and (ii), \( I_1 = I_2 \).

\[ \square \]

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 \( H \) and \( \Phi(H) \) be as in Lemma 2.1. Then the equality in (17) holds if and only if \( H \) can be factorized as
\[ H = [\text{vec}(W_1)^T \ G \ x] \text{vec}(W_1)^T, \]
where
(i) \( W_1 \in \mathbb{C}^{I_1 \times I_3} \) and \( x \) is a principal eigenvector of \( W_1 W_1^H \), i.e.,
\[ W_1 W_1^H x = \lambda_{max}(W_1 W_1^H) x, \quad \|x\| = 1; \]

(ii) the matrix \( G = [g_2 \cdots g_R] \in \mathbb{C}^{I_3 \times (R - 1)} \) has orthogonal columns;

(iii) \( G^T W_1^H = 0; \)

(iv) \( \lambda_{max}(W_1^H W_1) = \lambda_{max}(W_1^H W_1 + G^T G^T). \)

Moreover, if (36) and (i)–(iv) hold, then
\[ \sigma \left( \sum_{k=1}^{I_3} H_{kk} \right) = \sigma \left( W_1 W_1^H + \|G\|^2 xx^H \right), \]
\[ \sigma(H) = \{ \|W_1\|^2, \|g_2\|^2, \ldots, \|g_R\|^2, 0, \ldots, 0 \}, \]
\[ \sigma(\Phi(H)) = \sigma(W_1^H W_1 + G^T G^T), \]
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
\[ H = \sum_{r=1}^{R} \text{vec}(W_r) \text{vec}(W_r)^H, \quad \text{where } W_r = x g_r^T \text{ for } r = 2, \ldots, R. \]

Hence
\[ \sum_{k=1}^{I_3} H_{kk} = \sum_{r=1}^{R} W_r W_r^H = W_1 W_1^H + \sum_{r=2}^{R} x g_r^T g_r^* x^H = W_1 W_1^H + \|G\|^2 xx^H, \]
which implies (37). By (ii), (iii), and the convention \( \|x\| = 1 \) in (i), the vectors \( \text{vec}(W_r) \) are mutually orthogonal, which implies (38). Finally, by (21),

\[
\Phi(H) = \sum_{r=1}^{R} W_r^T W_r^* = W_1^T W_1^* + \sum_{r=1}^{R} g_r x^T x^r g_r^H = W_1^T W_1^* + G^H G^T,
\]

which implies (39).

**Sufficiency.** By (i) and (37),

\[
\lambda_{\max} \left( \sum_{k=1}^{I_3} H_{kk} \right) = \lambda_{\max} (W_1^H W_1^H) + \|g\|^2.
\]

By (iv) and (ii),

\[
\|W_1\|^2 = \lambda_{\max} (W_1^H W_1) \geq \lambda_{\max} (G^* G^T) = \max_{2 \leq r \leq R} \|g_r\|^2.
\]

Thus, by (38), \( \lambda_{\max}(H) = \|W_1\|^2 \) and \( \text{tr}(H) = \|W_1\|^2 + \|G\|^2 \). By (iv) and (39), \( \lambda_{\max}(\Phi(H)) = \lambda_{\max}(W_1^H W_1) \). Thus, the left- and right-hand sides of (17) are equal to \( \lambda_{\max}(W_1^H W_1^H) + \|W_1\|^2 + \|G\|^2 \).

**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 \( x \) satisfies (i). By the Cauchy inequality, the equality

\[
\sum_{k=1}^{I_3} \sum_{r=2}^{R} |(w_{kr}, x)|^2 = \sum_{r=2}^{R} \|w_r\|^2
\]

in (22) would imply that

\[
w_{kr} = c_{kr} x, \quad k = 1, \ldots, I_3, \quad r = 2, \ldots, R,
\]

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

\[
(40) \quad w_r = [w_1^T \ldots w_r^T]^T = [c_{1r} \ldots c_{I_3 r}]^T \otimes x = g_r \otimes x, \quad r = 2, \ldots, R. \quad \square
\]

Since \( H = \sum_{r=1}^{R} w_r w_r^H \), it follows that

\[
H = [w_1 \ldots w_R] [w_1 \ldots w_R]^H = [w_1 g_2 \otimes x \ldots g_R \otimes x] [w_1 g_2 \otimes x \ldots g_R \otimes x]^H,
\]

which coincides with (36). The mutual orthogonality of \( w_2, \ldots, w_R \) and the orthogonality of \( w_1 \) to \( w_2, \ldots, w_R \) imply (ii) and (iii), respectively. By (40), \( W_r = x g_r^T \) for \( r = 2, \ldots, R \). Hence, the equality

\[
\lambda_{\max}(W_1^H W_1) = \lambda_{\max} \left( \sum_{r=1}^{R} W_r^H W_r \right)
\]

in (23) would imply (iv):

\[
\lambda_{\max}(W_1^H W_1) = \lambda_{\max} (W_1^H W_1 + \sum_{r=1}^{R} g_r^* x^r g_r^T) = \lambda_{\max}(W_1^H W_1 + G^* G^T).
\]
Lemma 4.2.

(i) Let $W_1$, $G$, and $x$ satisfy conditions (i)–(iv) of Lemma 4.1, $H$ be defined as in (36), and
$$L := \lambda_{\text{max}}(W_1^H W_1).$$
Then there exist $(I_3 - 1) \times (I_3 - 1)$ positive semidefinite matrices $A$ and $B$ such that

\[
\text{rank}(A) \leq \min(I_1, I_3) - 1, \quad \text{rank}(B) = R - 1, \quad L \geq \lambda_{\text{max}}(A + B)
\]

and

\[
\sigma \left( \sum_{k=1}^{I_3} H_{kk} \right) = \left\{ L + \text{tr}(B), \lambda_1(A), \ldots, \lambda_{\min(I_1, I_3) - 1}(A), 0, \ldots, 0 \right\},
\]

\[
\sigma(H) = \left\{ L + \text{tr}(A), \lambda_1(B), \ldots, \lambda_{R-1}(B), 0, \ldots, 0 \right\},
\]

\[
\sigma(\Phi(H)) = \{ L \} \cup \sigma(A + B).
\]

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

Proof.

(i) Let $p$ be a principal eigenvector of $W_1 W_1^H$, i.e., $W_1^H W_1 p = L p$, $\|p\| = 1$. Then, by (iv), $G^* G^T p = 0$. Let $U_p$ be an $I_3 \times I_3$ unitary matrix whose first column is $p$. Then

\[
U_p^H W_1^H W_1 U_p = \begin{bmatrix} L & 0 \\ 0 & A \end{bmatrix}, \quad U_p^H G^* G^T U_p = \begin{bmatrix} 0 & 0 \\ 0 & B \end{bmatrix},
\]

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

\[
\lambda_k(A) = \lambda_{k+1}(W_1^H W_1), \quad k = 1, \ldots, I_3 - 1,
\]

\[
\lambda_k(B) = \lambda_k(G^* G^T) = \begin{cases} \|g_{k+1}\|^2, & k = 1, \ldots, R - 1, \\ 0, & k = R, \ldots, I_3 - 1. \end{cases}
\]

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

\[
\sigma \left( \sum_{k=1}^{I_3} H_{kk} \right) = \left\{ \lambda_{\text{max}}(W_1 W_1^H)
\right.
\]

\[
+ \|G\|^2, \lambda_2(W_1 W_1^H), \ldots, \lambda_{I_1}(W_1 W_1^H) \}.
\]

Since the nonzero eigenvalues of $W_1 W_1^H$ coincide with those of $W_1^H W_1$ and $\|G\|^2 = \text{tr}(B)$ it follows that (48) is equivalent to (42).

(ii) Let $H$ be defined as in (36), where

$W_1 = \begin{bmatrix} \sqrt{L} & 0 \\ 0 & \tilde{W}_1 \end{bmatrix}$, \quad $G = US^\dagger$, \quad $x = [1 \ 0 \ \ldots \ 0]^T$, the nonzero eigenvalues of $W_1$.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.
\( \mathbf{\bar{W}}_1 \) is an \((I_1 - 1) \times (I_3 - 1)\) matrix such that \( \mathbf{\bar{W}}_1^H \mathbf{\bar{W}}_1 = \mathbf{A} \), and \( \mathbf{USU}^H \) is the reduced SVD of \([\mathbf{U} \mathbf{0}] \). 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 \( \mathbf{H} = \mathbf{T}_{(2)}^* \mathbf{T}_{(2)}^* \). Since \( \sigma_1^2 + \sigma_2^2 = \| \mathbf{T} \|^2 + \sigma_3^2 \), it follows that equality (17) holds. Hence, by Horn’s conjecture, (13) and (14) hold. Therefore, there exists an \( I_2 \times R \) matrix \( \mathbf{V} \) whose columns are orthonormal and such that \( \mathbf{T}_{(2)}^* = [\vec(\mathbf{W}_1) \mathbf{G} \otimes \mathbf{x}] \mathbf{V}^H \), or equivalently,

\[
\mathbf{T}_k = [\mathbf{w}_{1k} \mathbf{g}_{k1} \cdots \mathbf{g}_{kR}] \mathbf{V}^H, \quad k = 1, \ldots, I_3.
\]

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

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

\[
\begin{align*}
\sigma_1^2 & \geq \sigma_2^2 \geq \cdots \geq \sigma_{I_1}^2 \geq 0 & \text{are the eigenvalues of } & \sum_{i=1}^{I_3} \mathbf{H}_i = \mathbf{T}_{(1)} \mathbf{T}_{(1)}^H, \\
\sigma_2^2 & \geq \sigma_3^2 \geq \cdots \geq \sigma_{I_2}^2 \geq 0 & \text{are the first } I_2 & \text{eigenvalues of } \mathbf{H}, \\
\sigma_3^2 & \geq \sigma_4^2 \geq \cdots \geq \sigma_{I_3}^2 \geq 0 & \text{are the eigenvalues of } & \Phi(\mathbf{H}) = \mathbf{T}_{(3)} \mathbf{T}_{(3)}^H.
\end{align*}
\]

**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 \( \alpha_i, \beta_i, \) and \( \gamma_i \) 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 \( \alpha_i, \beta_i, \) and \( \gamma_i \) 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 (49), \( \mathbf{H} \) is 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 \( \mathbf{T}_{(2)} = \mathbf{V}^*[\vec(\mathbf{W}_1) \mathbf{G} \otimes \mathbf{x}]^T \). Then, \( \mathbf{H} = \mathbf{T}_{(2)}^* \mathbf{T}_{(2)}^* \). The proof now follows from (49)–(51).

**5. Conclusion.** In the paper we studied geometrical properties of the set

\[
\Sigma_{I_1, I_2, I_3} := \left\{ (\sigma_{11}, \ldots, \sigma_{I_1}^2, \sigma_{I_2}^2, \sigma_{I_3}^2, \sigma_{I_3}^2, \ldots, \sigma_{I_3}^2) : \sigma_{nk} \text{ is the kth largest mode-n singular value of an } I_1 \times I_2 \times I_3 \text{ norm-1 tensor } \mathcal{T} \right\},
\]

where for each \( n = 1, 2, 3 \) the values \( \sigma_{nk} \) are sorted in descending order.

Let \( \pi \) denote a projection of \( \mathbb{R}^1 + I_2 + I_3 \) 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 $\Sigma_{I_1,I_2,I_3}$ 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 $\Sigma_{I_1,I_2,I_3}$ 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 $T^n_r$. In our presentation we follow [1, p. 302]. The set $T^n_r$ 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 $T^n_1$ if $k_1 = j_1 = 1$. For $r > 1$, $(I, J, K)$ is in $T^n_r$ if

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

and, for all $1 \leq p \leq r - 1$ and all $(U, V, W) \in T^n_p$,

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

Thus, $T^n_r$ is defined recursively in terms of $T^n_1, \ldots, T^n_{r-1}$.

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

REFERENCES


