Hot-SVD: higher order t-singular value decomposition for tensors based on tensor–tensor product

Ying Wang · Yuning Yang

Received: 13 September 2022 / Revised: 29 October 2022 / Accepted: 31 October 2022
© The Author(s) under exclusive licence to Sociedade Brasileira de Matemática Aplicada e Computacional 2022

Abstract
This paper considers a way of generalizing the t-SVD of third-order tensors (regarded as tubal matrices) to tensors of arbitrary order $N$ [which can be similarly regarded as tubal tensors of order $(N - 1)$]. Such a generalization is different from the t-SVD for tensors of order greater than three (Martin et al. in SIAM J Sci Comput 35(1):A474–A490, 2013). The decomposition is called Hot-SVD, since it can be recognized as a tensor–tensor product version of the celebrated higher order SVD (HOSVD). The existence of Hot-SVD is proved. To this end, the “small-t” transpose for third-order tensors is introduced. This transpose is crucial in the verification of Hot-SVD, since it serves as a bridge between tubal tensors and their unfolding tubal matrices. We establish some properties of Hot-SVD, analogous to those of HOSVD. The truncated Hot-SVD and sequentially truncated Hot-SVD are then introduced, with $\sqrt{N}$-error bounds established for an $(N + 1)$-th-order tensor. We provide numerical examples to validate Hot-SVD, truncated Hot-SVD, and sequentially truncated Hot-SVD.

Keywords
T-product · Tensor–tensor product · Tensor decomposition · T-SVD · HOSVD

Mathematics Subject Classification 90C26 · 15A18 · 15A69 · 41A50

1 Introduction

Research in tensor decomposition and approximation has seen increasing popularity with the exponential increase and availability of data in our world. Encoding these data in a tensor-based format allows us to exploit fully their inherent multi-dimensional features. Tensor decompositions, such as canonical polyadic decomposition, Tucker decomposition,
tensor-train decomposition, and t-SVD, have found various applications in signal processing, machine learning, computer vision, etc; see, e.g., (Kolda and Bader 2009; Comon 2014; Cichocki et al. 2015; Sidiropoulos et al. 2017; Oseledets 2011; Kilmer and Martin 2011) and the references therein. One of the powerful tensor-algebraic methods for processing tensor-type data is provided by the t-product-based tensor theory and computation (Braman 2010; Kilmer and Martin 2011; Kilmer et al. 2013; Kernfeld et al. 2015), for which an Eckart–Young-type result (t-SVD) exists. The t-product framework has recently generated a surge of research activities both in theory and applications, rendering available a large stock of the matrix algebra arsenal to the tensor-algebra community [see, e.g., (Braman 2010; Kilmer and Martin 2011; Kilmer et al. 2013; Kernfeld et al. 2015; Miao et al. 2021; Qi and Luo 2021; Ling et al. 2021; Qi and Yu 2021; Zhu and Wei 2022; Newman et al. 2018; Zhang and Aeron 2017; Zhang et al. 2014; Kong et al. 2018; Yin et al. 2019)].

In Kernfeld et al. (2015) and Qi and Luo (2021), a natural perspective on the t-product (and also the more general tensor–tensor product defined by an arbitrarily invertible linear transformation) is provided by tubal matrices. That is, one regards a third-order tensor $A$ of size $I_1 \times I_2 \times I_3$ as a tubal matrix of size $I_1 \times I_2$ whose entries are horizontal vectors (tubal scalars) of length $I_3$. A variation of the Hadamard product, which is obtained by composing the Hadamard product with an invertible linear transformation, gives the general tensor–tensor product of two tubal scalars (Kernfeld et al. 2015). When the linear transformation in action is the (non-normalized) discrete Fourier transform (DFT), this gives the most widely used t-product of tubal scalars, which can then be extended to tubal matrix multiplications to give the t-product of tubal matrices.

Based on the t-product, the t-SVD of a third-order tensor was proposed in Kilmer and Martin (2011). The idea is to decompose the frontal slices of a tubal matrix in the Fourier domain using the matrix SVD, in conjunction with DFT and IDFT (inverse discrete Fourier transform) operations on the tubal scalars before and after the matrix decompositions. t-SVD was then generalized to tensors of order higher than three via recursion (Martin et al. 2013). t-SVD was introduced in Kernfeld et al. (2015) and Kilmer et al. (2021) in the sense of the more general tensor–tensor product.

In this paper, we investigate an alternative way of generalizing t-SVD to tensors of order higher than three. The celebrated higher order singular value decomposition (HOSVD) is quite a successful extension of SVD to higher order tensors that decomposes the data tensor into a core tensor and a collection of factor matrices. HOSVD together with the tensor–tensor product naturally suggests us to consider a tensor–tensor product version of HOSVD. The new decomposition, called the higher order t-singular value decomposition (Hot-SVD), treats an $(N + 1)$-th-order tensor as an $N$-th-order tubal tensor, and factorizes it into a core tubal tensor and a collection of tubal matrices. This is essentially based on the unfolding in the tensor–tensor product sense and t-SVD for third-order tensors (tubal matrices). The connection of Hot-SVD with t-SVD and HOSVD is illustrated in Fig. 1. Inherited from HOSVD, a feature of Hot-SVD is that the truncation can be made on each mode (except the tubal mode).

We then prove the existence of Hot-SVD and establish several properties such as all-orthogonality and ordering that are similar to those of HOSVD. A crucial point to make the analysis go through is the use of the “small-t” transpose for third-order tensors, which is the transpose without inversion that is different from Kilmer and Martin (2011) and Kernfeld et al. (2015). We show that some necessary properties hold for this transpose. In particular, based on this transpose, the Kronecker product and corresponding properties can be generalized in the sense of tensor–tensor product. These enable us to build a bridge between tubal tensors
and their unfolding tubal matrices, such that the proof of the validity of Hot-SVD can go through.

We then introduce the truncated and sequentially truncated Hot-SVD, and establish their $\sqrt{N}$-error bounds for an $(N + 1)$-th-order tensor. The sequentially truncated Hot-SVD generalizes the sequentially truncated HOSVD to tubal tensors (Vannieuwenhoven et al. 2012).

Our motivation for introducing and studying Hot-SVD is to contribute another kind of generalization of the t-SVD of third-order tensors to higher order tensors in the setup of tensor–tensor product. As a new model of tensor decomposition and approximation, it adds to the variety of the available tensor processing methods and enriches the tensor–tensor product-based tensor theory and computation. We also expect potential applications of our truncated algorithms for Hot-SVD in fields such as image and video processing.

The rest is organized as follows. In Sect. 2, we collect some preliminary materials related to the tensor–tensor product of third-order tensors. In Sect. 3, we introduce the small-t transpose and prove its basic properties.

We also extend some familiar notions and results from matrices (such as the Kronecker product) to tubal matrices. The Hot-SVD model for tubal tensors is established and analyzed in Sect. 4. Section 5 is focused on the truncated Hot-SVD, the sequentially truncated Hot-SVD, and their error bounds. We analyze the computational complexity of Hot-SVD, truncated Hot-SVD, and sequentially truncated Hot-SVD in Sect. 6. Numerical examples are provided in Sect. 7.

Finally, we draw some conclusions in Sect. 8.

2 Preliminaries

A tensor of order $N$ is a multi-way array $A = (a_{i_1i_2...i_N})(i_1, i_2,..., i_N) \in I_1 \times I_2 \times \cdots \times I_N \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$. Complex tensors are defined similarly. The Matlab indexing notation is adopted in this article. Denote by $A(i_1, \ldots, i_N)$ the $(i_1, \ldots, i_N)$-th entry of $A$ and by $A(i_1, \ldots, i_{n-1}, \ldots, i_{n+1}, \ldots, i_N)$ the mode-$n$ fiber of $A$ obtained by varying the $n$-th index while fixing the other $N - 1$ indices to be $i_1, \ldots, i_{n-1}, i_{n+1}, \ldots, i_N$, respectively.

The Frobenius norm of a tensor $A \in \mathbb{C}^{I_1 \times I_2 \times \cdots \times I_N}$ is defined to be $\|A\|_F := \sqrt{\sum_{i_1=1}^{I_1} \cdots \sum_{i_N=1}^{I_N} |A(i_1, \ldots, i_N)||A(i_1, \ldots, i_N)|}$, where the overline indicates the conjugate of a complex number and $\| \cdot \|$ is the absolute value of a complex number. In this paper, we drop the subscript “$F$” in the notation as we do not
use any other norm. The Frobenius norm of a tubal matrix or a tubal tensor (to be introduced later) is defined to be the Frobenius norm of the underlying tensor.

The $n$-mode product of a tensor $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \times \cdots \times I_N}$ with a matrix $U \in \mathbb{C}^{J \times I_n}$ is the tensor $\mathcal{A} \times_n U \in \mathbb{C}^{I_1 \times \cdots \times I_{n-1} \times J \times I_{n+1} \times \cdots \times I_N}$ obtained by multiplying each mode-$n$ fiber of $\mathcal{A}$ by $U$.

**Definition 2.1** (Kilmer et al. 2013, Definition 2.1) An element $c \in \mathbb{R}^{1 \times 1 \times p}$ is called a tubal scalar of length $p$. That is, a tubal scalar of length $p$ is just a vector of dimension $p$ viewed as a third-order tensor with a single tubal fiber. Denote by $c^{(j)}$ the $j$-th frontal slice of $c$ (as a third-order tensor), i.e., the $j$-th component of $c$ as a vector. The set of tubal scalars of length $p$ is denoted by $\mathbb{R}_p$. Complex tubal scalars are defined similarly and the set of complex tubal scalars of length $p$ is denoted by $\mathbb{C}_p$.

For $a = (a^{(1)}, \ldots, a^{(p)})$ and $b = (b^{(1)}, \ldots, b^{(p)}) \in \mathbb{C}_p$, their Hadamard product is the tubal scalar $a \odot b := (a^{(1)} b^{(1)}, \ldots, a^{(p)} b^{(p)}) \in \mathbb{C}_p$, where $\odot$ refers to the Hadamard product. It can be easily verified that the Hadamard product is commutative, associative, unital (with the identity being $(1, 1, \ldots, 1) \in \mathbb{C}_p$), and distributive over addition.

The following definition comes from Kernfeld et al. (Kernfeld et al. 2015, Definition 4.2), specialized to the tubal scalar case.

**Definition 2.2** (Kernfeld et al. 2015) Let $L : \mathbb{C}_p \rightarrow \mathbb{C}_p$ be an invertible linear transformation. For tubal scalars $a, b \in \mathbb{C}_p$, their tensor–tubal product with respect to $L$ is the tubal scalar $a \ast_L b = L^{-1}(L(a) \odot L(b)) \in \mathbb{C}_p$.

When the invertible linear transformation $L$ is the (non-normalized) discrete Fourier transform (DFT), the resulting product is the popular t-product defined in Kilmer and Martin (2011). There are other choices of $L$, such as discrete cosine transform and discrete wavelet transform (Kernfeld et al. 2015). For ease of notation, in the sequel, we will drop the subscript $L$ and denote by $a \ast b$ the tensor–tubal product of $a$ and $b$ with respect to a fixed invertible linear transformation $L$.

The tensor–tubal product of tubal scalars is commutative, associative, unital (with the identity being $L^{-1}((1, 1, \ldots, 1)) \in \mathbb{C}_p$), and distributive over addition (Kernfeld et al. 2015). Therefore, $\mathbb{C}_p$ endowed with the tensor–tubal product is a commutative ring (Kernfeld et al. 2015, Proposition 4.2), but not a field in general.

**Definition 2.3** (Kilmer et al. 2013, Definition 2.4) A tubal matrix with entries in $\mathbb{C}_p$ is a two-dimensional array $A = (a_{ij})_{(i,j) \in I \times J} \in \mathbb{C}_p^{I \times J}$ where the entries $a_{ij} \in \mathbb{C}_p$ are tubal scalars of length $p$. Thus, a tubal matrix $A \in \mathbb{C}_p^{I \times J}$ is essentially a third-order tensor $A \in \mathbb{C}^{I \times J \times p}$.

**Definition 2.4** (Kernfeld et al. 2015, Definition 2.1) Let $A \in \mathbb{C}_p^{I \times J}$ and $B \in \mathbb{C}_p^{J \times K}$ be two tubal matrices. The face-wise product of $A$ and $B$ is the tubal matrix $A \Delta B \in \mathbb{C}_p^{I \times K}$ defined according to

$$(A \Delta B)^{(i)} = A^{(i)} B^{(i)}, \quad i = 1, \ldots, p,$$

where $A^{(i)}$ is the $i$-th frontal slice of $A$.

**Definition 2.5** (Kernfeld et al. 2015, Definition 4.2) Let $A \in \mathbb{C}_p^{I \times J}$ and $B \in \mathbb{C}_p^{J \times K}$ be two tubal matrices. The tensor–tubal product of $A$ and $B$ is the tubal matrix $A \ast B \in \mathbb{C}_p^{I \times K}$ given by

$$A \ast B = L^{-1}(L(A) \Delta L(B)),$$

where $L(A)$ means applying $L$ to each entry of $A$, and $\Delta$ is the face-wise product.
The tensor–tensor product of tubal matrices can also be defined using the standard matrix multiplication rule. Kernfeld et al. (2015, Lemma 4.1) states that for any $i \in I, k \in K$

$$(A \ast B) (i, k) = \sum_{j=1}^{J} A(i, j) \ast B(j, k). \tag{2.1}$$

When $L$ is the DFT, Definition 2.5 is equivalent to the t-product defined in Kilmer and Martin (2011).

The tensor–tensor product of tubal matrices, like the usual matrix multiplication, is associative and distributive over addition (see the proof of Kernfeld et al. (2015, Proposition 4.2)), but not commutative in general.

**Definition 2.6** (Kernfeld et al. 2015, Proposition 4.1) Let $\hat{L} \in \mathbb{C}_{p}^{I \times I}$ be the tubal matrix, such that $\hat{L}^{(i)}$ is the identity matrix for $i = 1, \ldots, p$. The identity tubal matrix with respect to the tensor–tensor product is defined to be $\mathcal{I} = L^{-1}(\hat{L})$. It holds that $A \ast \mathcal{I} = A \ast \hat{L}^{(i)}$ for any $A \in \mathbb{C}_{p}^{H \times I}$ and $\mathcal{I} \ast B = B \ast \hat{L}^{(i)}$ for any $B \in \mathbb{C}_{p}^{I \times J}$.

When $L$ is the DFT, the identity tubal matrix $\mathcal{I}$ is the third-order tensor whose first frontal slice is the identity matrix and whose other frontal slices are zero matrices.

**Definition 2.7** (Kernfeld et al. 2015, Definition 4.4) The Hermitian transpose $A^{H} \in \mathbb{C}_{p}^{I \times J}$ of a tubal matrix $A \in \mathbb{C}_{p}^{I \times J}$ is defined according to

$$(L(A^{H}))^{(i)} = (L(A)^{(i)})^{H}, \quad i = 1, \ldots, p,$$

where $(L(A)^{(i)})^{H}$ is the Hermitian transpose of the matrix $L(A)^{(i)}$. In particular, the Hermitian transpose of a tubal scalar $a \in \mathbb{C}_{p} = \mathbb{C}_{p}^{1 \times 1}$ is the tubal scalar $a^{H}$, such that $(L(a^{H}))^{(i)} = (L(a)^{(i)})^{H} = L(a^{(i)})$ for $i = 1, \ldots, p$.

When $A$ is real valued, we write $A^{T}$ for $A^{H}$ and $A^{T}$ is called the transpose of $A$. When $L$ is the DFT, the transpose $A^{T}$ can be obtained by transposing each frontal slice of $A$ and reversing the order of the frontal slices except for the first one (see Kernfeld et al. 2015, p. 560).

**Proposition 2.1** (Kernfeld et al. 2015, Proposition 4.3) Let $A \in \mathbb{C}_{p}^{I \times J}, B \in \mathbb{C}_{p}^{J \times K}$ be tubal matrices. The Hermitian transpose enjoys the rule $(A \ast B)^{H} = B^{H} \ast A^{H}$, where $A^{H} \in \mathbb{C}_{p}^{I \times J}, B^{H} \in \mathbb{C}_{p}^{K \times J}, (A \ast B)^{H} \in \mathbb{C}_{p}^{K \times I}$.

**Definition 2.8** (Kilmer et al. 2021, Definition 2.3) A tubal matrix $A \in \mathbb{C}_{p}^{I \times J}$ is said to be unitary with respect to the tensor–tensor product if $A \ast A^{H} = A^{H} \ast A = \mathcal{I}$, where $\mathcal{I}$ is the identity tubal matrix with respect to the tensor–tensor product. Orthogonality with respect to the tensor–tensor product is defined similarly using $A^{T}$ when $A$ is real valued.

**Definition 2.9** A tubal matrix $A \in \mathbb{C}_{p}^{I \times J}$ ($I \geq J$) is said to be partially unitary with respect to the tensor–tensor product if $A^{H} \ast A = \mathcal{I}$. Similarly, a real tubal matrix $A \in \mathbb{R}_{p}^{I \times J}$ ($I \geq J$) is said to be partially orthogonal with respect to the tensor–tensor product if $A^{T} \ast A = \mathcal{I}$.

Two column tubal vectors $\mathcal{X}, \mathcal{Y} \in \mathbb{C}_{p}^{I \times 1}$ are said to be orthogonal with respect to the tensor–tensor product if $\mathcal{X}^{H} \ast \mathcal{Y} = 0$, where $0 = (0, \ldots, 0) \in \mathbb{C}_{p}$ is the zero tubal scalar. Two row tubal vectors $\mathcal{X}, \mathcal{Y} \in \mathbb{C}_{p}^{1 \times I}$ are said to be orthogonal with respect to the tensor–tensor product if $\mathcal{X} \ast \mathcal{Y}^{H} = 0$. 
We quote here the singular value decomposition theorem for tubal matrices based on the tensor–tensor product.

**Theorem 2.1** (t-SVD of tubal matrices, Kernfeld et al. 2015, Theorem 5.1) Let \( A \in \mathbb{C}^{l \times J} \) be a tubal matrix. Then, there exist unitary tubal matrices \( U \in \mathbb{C}^{l \times I} \) and \( V \in \mathbb{C}^{J \times J} \) and an f-diagonal tubal matrix \( S \in \mathbb{C}^{I \times J} \), such that

\[
A = U \ast S \ast V^H.
\]

Here, an f-diagonal tubal matrix means a tubal matrix whose frontal slices are diagonal matrices.

**Definition 2.10** (Kilmer et al. 2021, Definition 3.4) Let \( A \in \mathbb{C}^{l \times J} \) be a tubal matrix and \( A = U \ast S \ast V^H \) a t-SVD of \( A \). The number of non-zero tubal scalars on the diagonal of \( S \), which does not depend on the particular decomposition, is called the t-rank of \( A \).

**Definition 2.11** (Kilmer et al. 2021, Definition 3.5) Let \( A \in \mathbb{C}^{l \times J} \) be a tubal matrix. The vector \( \rho = (\rho_1, \ldots, \rho_p) \), where \( \rho_i \) is the rank of the \( i \)-th frontal slice \( L(A)^{(i)} \), is called the multi-rank of \( A \).

An especially satisfying feature of the t-SVD is the following Eckart–Young-type result. The tubal matrix notation is used in the statement below. Throughout this paper, if not specified, the norm \( \| \cdot \| \) always means the Frobenius norm (i.e., the Frobenius norm of the underlying third-order tensor).

**Theorem 2.2** (Kilmer et al. 2021, Theorem 3.7) Let \( L \) be of the form \( L = cW \), where \( c \in \mathbb{C} \) is a non-zero scalar and \( W : \mathbb{C}_p \to \mathbb{C}_p \) is a unitary transformation (i.e., a unitary matrix). Let the t-SVD of \( A \in \mathbb{C}^{n_1 \times n_2} \) be given by \( A = U \ast S \ast V^H \), and for \( k < \min(n_1, n_2) \), define

\[
A_k = \sum_{i=1}^{k} U(:, i) \ast S(i, i) \ast V(:, i)^H.
\]

Then, \( A_k = \arg \min_{\tilde{A} \in M} \| A - \tilde{A} \| \), where \( M = \{ X \ast Y : X \in \mathbb{C}_p^{n_1 \times k}, Y \in \mathbb{C}_p^{k \times n_2} \} \).

The tubal matrix perspective naturally leads to the interpretation of a third-order tensor (i.e., a tubal matrix) as a linear operator on the space of (lateral) matrices [i.e., tubal (column) vectors]; see Kilmer et al. (2013) for a detailed exposition.

The following result is taken from Kilmer et al. (2013), where it was stated for the t-product (i.e., for \( L \) being the DFT), while the proof works for the tensor–tensor product defined by any invertible linear transformation \( L \).

**Theorem 2.3** (Kilmer et al. 2013, Theorem 4.6) Let \( A \in \mathbb{C}^{l \times m} \) be a tubal matrix and \( A = U \ast S \ast V^H \) its t-SVD. Let \( s_i = S(i, i) \), \( i = 1, \ldots, \min(l, m) \). Suppose that \( s_1, \ldots, s_j \) are invertible, \( s_{j+1}, \ldots, s_{j+k} \) are non-zero but not invertible, and the remaining \( s_{j+k+1}, \ldots, s_{\min(l, m)} \) are zero. Then, the range \( R(A) \) of \( A \) and the kernel \( N(A) \) of \( A \) are given by

\[
R(A) = \{ U(:, 1) \ast c_1 + \cdots + U(:, j + k) \ast c_{j+k} : c_i = s_i \ast d_i, d_i \in \mathbb{C}_p, j + 1 \leq i \leq j + k \},
\]

\[
N(A) = \{ V(:, j + 1) \ast c_{j+1} + \cdots + V(:, m) \ast c_m : s_i \ast c_i = 0, j + 1 \leq i \leq j + k \}.
\]

\(^1\) The DFT is of this form, with \( c = \sqrt{p} \) and \( W \) being the normalized DFT matrix.
The following theorem states that the Frobenius norm of a tubal matrix (i.e., the Frobenius norm of the underlying third-order tensor) is invariant when multiplied (in the sense of tensor–tensor product) by a partially unitary tubal matrix. The statement below is slightly more general than (Kilmer et al. 2021, Theorem 3.1) in that it allows partial unitarity, but can be proved in the same way.

**Theorem 2.4** (Kilmer et al. 2021, Theorem 3.1) Let \( L = cW \) where \( c \in \mathbb{C} \) is a non-zero scalar and \( W \) is a unitary transformation. Let \( A \in \mathbb{C}^{I \times J}_p \) be a tubal matrix. Let \( Q \in \mathbb{C}^{I' \times I}_p \), \( P \in \mathbb{C}^{J' \times J}_p \) be partially unitary tubal matrices. Then

\[
\| Q * A * P^H \| = \| A \| .
\]

### 3 The small-t transpose for tubal matrices

We first introduce the following transpose, and then, we extend some familiar notions and results in matrix algebra (such as the Kronecker product) to tubal matrices based on this transpose.

**Definition 3.1** Let \( A \in \mathbb{C}^{I \times J}_p \) be a tubal matrix. We denote by \( A^t \) the tubal matrix obtained by simply transposing the frontal slices of \( A \). Thus, \( A^t \) is obtained by transposing the tubal matrix \( A \), i.e., \( A^t(j, i) = A(i, j) \) for all \( i \in I, j \in J \).

The transpose defined above will be referred to as the small-t transpose (or the face-wise transpose) in contrast to the capital-T transpose introduced in the previous section. A series of basic properties concerning the small-t transpose will be established in the sequel.

**Proposition 3.1** The small-t transpose enjoys the rule

\[
(A * B)^t = B^t * A^t \quad \text{where} \quad A, B \in \mathbb{C}^{I \times J}_p.
\]

**Proof**

\[
(A * B)^t(k, i) = (A * B)(i, k)
\]

\[
\overset{(2.1)}{=} \sum_{j=1}^{J} A(i, j) * B(j, k)
\]

\[
= \sum_{j=1}^{J} A^t(j, i) * B^t(k, j)
\]

\[
= \sum_{j=1}^{J} B^t(k, j) * A^t(j, i)
\]

\[
\overset{(2.1)}{=} (B^t * A^t)(k, i),
\]

where the fourth equality holds, since the tensor–tensor product of tubal scalars is commutative.

We will make use of the following proposition in establishing our main theorem.

**Proposition 3.2** \( A \in \mathbb{C}^{I \times J}_p \) is unitary if and only if \( A^t \) is.
Proof By definition, $\mathcal{A}$ is unitary if $\mathcal{A} \ast \mathcal{A}^H = \mathcal{A}^H \ast \mathcal{A} = \mathcal{I}$, which is equivalent to

$$L(\mathcal{A} \ast \mathcal{A}^H) = L(\mathcal{A}^H \ast \mathcal{A}) = L(\mathcal{I}).$$

Note that

$$\left( L(\mathcal{A} \ast \mathcal{A}^H) \right)^{(i)} = L(\mathcal{A})^{(i)} L(\mathcal{A}^H)^{(i)} \overset{?}{=} L(\mathcal{A})^{(i)} \left( L(\mathcal{A})^{(i)} \right)^H,$$

$$\left( L(\mathcal{A}^H \ast \mathcal{A}) \right)^{(i)} = L(\mathcal{A})^{(i)} L(\mathcal{A})^{(i)} \overset{?}{=} \left( L(\mathcal{A})^{(i)} \right)^H L(\mathcal{A})^{(i)}.$$

Thus, $\mathcal{A}$ is unitary if and only if for $i = 1, \ldots, p$,

$$L(\mathcal{A})^{(i)} \left( L(\mathcal{A})^{(i)} \right)^H = \left( L(\mathcal{A})^{(i)} \right)^H L(\mathcal{A})^{(i)} = (L(\mathcal{I}))^{(i)} = I,$$

where $I \in \mathbb{C}^{I \times I}$ is the identity matrix. That is, $\mathcal{A}$ is unitary if and only if the frontal slices of $L(\mathcal{A})$ are unitary matrices.

Observe that

$$L(\mathcal{A}^i) = L(\mathcal{A})^i, \text{ i.e., } L(\mathcal{A}^i)^{(k)} = (L(\mathcal{A})^{(k)})^T, \ k = 1, \ldots, p,$$

(3.1)

since for any $j \in J, i \in I$,

$$L(\mathcal{A}^i)(j, i) = L(\mathcal{A}^i)(j, i)$$

$$= L(\mathcal{A}(i, j))$$

$$= L(\mathcal{A})(i, j)$$

$$= (L(\mathcal{A}))^i(j, i),$$

where the first and third equalities hold by definition of the operator $L$. Therefore, we have for $i = 1, \ldots, p$,

$$L(\mathcal{A}^i)^{(i)} \left( L(\mathcal{A}^i)^{(i)} \right)^H \overset{(3.1)}{=} \left( L(\mathcal{A})^{(i)} \right)^T \left( \left( L(\mathcal{A})^{(i)} \right)^T \right)^H$$

$$= \frac{\left( L(\mathcal{A})^{(i)} \right)^H L(\mathcal{A})^{(i)}}{L(\mathcal{A})^{(i)}}$$

where $\overline{L(\mathcal{A})^{(i)}}$ refers to the matrix obtained by taking the conjugate of $L(\mathcal{A})^{(i)}$ and the last equality comes from the fact that $\overline{A \ast B} = \overline{A} \ast \overline{B}$ for matrices $A, B$.

Similarly, we have for $i = 1, \ldots, p$,

$$\left( L(\mathcal{A}^i)^{(i)} \right)^H L(\mathcal{A}^i)^{(i)} = \overline{L(\mathcal{A})^{(i)}} \left( L(\mathcal{A})^{(i)} \right)^H.$$

Now, we are ready to prove the proposition.

If $\mathcal{A}$ is unitary, then for $i = 1, \ldots, p$,

$$L(\mathcal{A}^i)^{(i)} \left( L(\mathcal{A}^i)^{(i)} \right)^H = \overline{L(\mathcal{A})^{(i)}} \overline{L(\mathcal{A})^{(i)}} = \overline{T} = I,$$

and

$$\overline{L(\mathcal{A})^{(i)}} \overline{L(\mathcal{A})^{(i)}} = \overline{L(\mathcal{A})^{(i)}} \overline{L(\mathcal{A})^{(i)}} = T = I,$$

so $\mathcal{A}^i$ is unitary.
Conversely, if $A'$ is unitary, then for $i = 1, \ldots, p$,

\[
\left( L(A)^{(i)} \right)^H L(A)^{(i)} = L(A')^{(i)} \left( L(A')^{(i)} \right)^H = I = 1,
\]

and

\[
L(A)^{(i)} \left( L(A)^{(i)} \right)^H = (L(A')^{(i)})^H L(A')^{(i)} = I = 1,
\]

so $A$ is unitary. \hfill \Box

Next, we extend some familiar notions from matrices to tubal matrices.

**Definition 3.2** Let $A = (a_{ij}) \in \mathbb{C}^{I \times J}$ and $B = (b_{kl}) \in \mathbb{C}^{K \times L}$ be tubal matrices. The (tensor–tensor product-based) Kronecker product of $A$ and $B$ is defined to be the tubal matrix $A \otimes B \in \mathbb{C}^{IK \times JL}$ whose $(i, k)$-th entry is the tubal scalar $a_{ij} \ast b_{kl}$. That is

\[
A \otimes B = (a_{ij} \ast b_{kl})_{(i, j) \in I \times J},
\]

where $a_{ij} \ast B = (a_{ij} \ast b_{kl})_{(k, \ell) \in K \times L}$ is the tubal matrix obtained by taking the tensor–tensor product of $a_{ij}$ and each entry of $B$. Equivalently, the $(k + (i - 1)I, l + (j - 1)J)$-th entry of $A \otimes B$ is $a_{ij} \ast b_{kl}$.

In general, for tubal matrices $A_N \in \mathbb{C}^{I_N \times J_N}$, $\ldots$, $A_2 \in \mathbb{C}^{I_2 \times J_2}$, $A_1 \in \mathbb{C}^{I_1 \times J_1}$, we define their (tensor–tensor product-based) Kronecker product

\[
A_N \otimes \cdots \otimes A_2 \otimes A_1 \in \mathbb{C}^{I_N \cdots I_2 I_1 \times J_N \cdots J_2 J_1}
\]

to be the tubal matrix whose $(i_N, \ldots, i_2, i_1, j_N, \ldots, j_2, j_1)$-th entry (i.e., $(i_1 + (i_2 - 1)I_1 + \cdots + (i_N - 1)I_N\cdots I_{N-1}, j_1 + (j_2 - 1)J_1 + \cdots + (j_N - 1)J_N\cdots J_{N-1})$-th entry) is

\[
A_N(i_N, j_N) \ast \cdots \ast A_2(i_2, j_2) \ast A_1(i_1, j_1).
\] (3.2)

We need the following properties of the Kronecker product for tubal matrices.

**Proposition 3.3** Let $A, B, C, D$ be tubal matrices over $\mathbb{C}_p$ of appropriate size.

(i) $(A \otimes B)' = A' \otimes B'$.

(ii) $(A \otimes B)^H = A^H \otimes B^H$.

(iii) $(A \otimes B) \ast (C \otimes D) = (A \ast C) \otimes (B \ast D)$.

(iv) $(A \otimes B) \otimes C = A \otimes (B \otimes C)$.

(v) If $U_1 \in \mathbb{C}_p^{I_1 \times J_1}$, $U_2 \in \mathbb{C}_p^{J_2 \times L}$ are unitary, then $U_1 \otimes U_2$ is also unitary.

**Proof** (i) Let $A \in \mathbb{C}_p^{I \times J}$, $B \in \mathbb{C}_p^{K \times L}$. Then, $A \otimes B \in \mathbb{C}_p^{IK \times JL}$, $(A \otimes B)' \in \mathbb{C}_p^{JL \times IK}$. For $i \in I$, $j \in J$, $k \in K$, $l \in L$

\[
(A \otimes B)^{((j, l), (i, k)}) = (A \otimes B) ((i, k), (j, l)) = A(i, j) \ast B(k, l) = A'(j, i) \ast B'(l, k) = (A' \otimes B') ((j, l), (i, k)).
\]
(ii) We introduce an auxiliary notation used only in this proof.

For tubal matrices \( A \in \mathbb{C}^{I \times J} \) and \( B \in \mathbb{C}^{K \times L} \) we define \( A \otimes_{H} B \in \mathbb{C}^{IK \times JL} \) (with “H” referring to “Hadamard”) to be the tubal matrix whose \( ((i, k), (j, l)) \)-th entry is \( A(i, j) \odot B(k, l) \), where \( \odot \) is the Hadamard product. In other words, for \( i = 1, \ldots, p \), we have

\[
\left( A \otimes_{H} B \right)_{(i)}^{(j)} = A^{(i)} \otimes B^{(i)}, \tag{3.3}
\]

where \( A^{(i)} \otimes B^{(i)} \) refers to the Kronecker product in the usual sense between matrices \( A^{(i)} \) and \( B^{(i)} \).

Note that

\[
L(A \otimes B) = L(A) \otimes_{H} L(B), \tag{3.4}
\]

since the \( ((i, k), (j, l)) \)-th entry of \( L(A \otimes B) \) is

\[
L(A(i, j) \ast B(k, l)) = L(A(i, j)) \odot L(B(k, l)) = L(A)(i, j) \odot L(B)(k, l),
\]

which is also the \( ((i, k), (j, l)) \)-th entry of \( L(A) \otimes_{H} L(B) \).

The formula for Hermitian transpose then follows, since for \( i = 1, \ldots, p \),

\[
\left( L\left( (A \otimes B)^{H} \right) \right)_{(i)}^{(j)} = \frac{\left( L(A) \otimes_{H} L(B) \right)_{(i)}}{\left( L(A) \otimes_{H} L(B) \right)^{T}}
\]

\[
\frac{\left( L(A) \otimes_{H} L(B) \right)_{(i)}}{\left( L(A) \otimes_{H} L(B) \right)^{T}} = \frac{L(A)^{(i)} \otimes (L(B)^{(i)})^{T}}{(L(A)^{(i)})^{T} \otimes (L(B)^{(i)})^{T}}
\]

\[
= \left( L(A)^{(i)} \right)^{H} \otimes \left( L(B)^{(i)} \right)^{H}
\]

\[
= L(A)^{(i)} \ast L(B)^{(i)}
\]

\[
= \left( L(A)^{(i)} \otimes (L(B)^{(i)})^{T} \right)
\]

\[
= \left( L(A)^{(i)} \otimes L(B)^{(i)} \right)
\]

\[
= \frac{L(A)^{(i)} \otimes L(B)^{(i)}}{(L(A)^{(i)})^{T} \otimes (L(B)^{(i)})^{T}}
\]

\[
= \left( L(A)^{(i)} \right)^{H} \otimes \left( L(B)^{(i)} \right)^{H}
\]

\[
= L(A)^{(i)} \ast L(B)^{(i)}
\]

\[
= \left( L(A)^{(i)} \otimes (L(B)^{(i)})^{T} \right)
\]

\[
= \left( L(A)^{(i)} \otimes L(B)^{(i)} \right)
\]

which means \( L\left( (A \otimes B)^{H} \right) = L\left( A^{H} \otimes B^{H} \right) \), implying that \( (A \otimes B)^{H} = A^{H} \otimes_{t} B^{H} \). In the derivation, we used properties of the Kronecker product for matrices.
(iii) Suppose $A \in \mathbb{C}_p^{I \times J}$, $B \in \mathbb{C}_p^{L \times M}$, $C \in \mathbb{C}_p^{J \times K}$, $D \in \mathbb{C}_p^{M \times N}$. As block tubal matrices, we have
\[
\left( (A \otimes B) \ast (C \otimes D) \right) (i,k) = \sum_{j=1}^{J} (A(i,j) \ast B) \ast (C(j,k) \ast D)
\]
\[
= \sum_{j=1}^{J} (A(i,j) \ast C(j,k)) \ast (B \ast D)
\]
\[
= (A \ast C)(i,k) \ast (B \ast D)
\]
\[
= \left( (A \ast C) \otimes (B \ast D) \right) (i,k),
\]
where the second equality comes from the associativity and commutativity of the tensor–tensork product for tubal scalars.

(iv) Suppose $A \in \mathbb{C}_p^{I \times J}$, $B \in \mathbb{C}_p^{K \times L}$, $C \in \mathbb{C}_p^{M \times N}$. Both $(A \otimes B) \otimes C$ and $A \otimes (B \otimes C)$ are tubal matrices of size $IKM \times JLN$ whose $(i,k,m), (j,l,n))$-th entry is
\[
A(i,j) \ast B(k,l) \ast C(m,n).
\]

(v) Since $U_1 \in \mathbb{C}_p^{I \times I}$, $U_2 \in \mathbb{C}_p^{J \times J}$ are unitary, we have
\[
(U_1 \otimes U_2)^H \ast (U_1 \otimes U_2) = (U_1^H \otimes U_2^H) \ast (U_1 \otimes U_2)
\]
\[
= I_I \otimes I_J
\]
\[
= I_{IJ}.
\]
Similarly, $(U_1 \otimes U_2) \ast (U_1 \otimes U_2)^H = I_{IJ}$, so $U_1 \otimes U_2$ is unitary.

\[\square\]

**Definition 3.3** A tubal tensor of order $N$ with entries in $\mathbb{C}_p$ is an element $A \in \mathbb{C}_p^{I_1 \times I_2 \times \cdots \times I_N}$, that is, a multi-way array $(a_{i_1i_2...i_N})_{(i_1,i_2,...,i_N) \in I_1 \times I_2 \times \cdots \times I_N}$, where $a_{i_1i_2...i_N} \in \mathbb{C}_p$ are tubal scalars of length $p$. In particular, a tubal tensor of order 2 is a tubal matrix.

Since $\mathbb{C}_p$ is isomorphic to $\mathbb{C}^p$, $\mathbb{C}_p^{I_1 \times I_2 \times \cdots \times I_N}$ is isomorphic to $\mathbb{C}^{I_1 \times I_2 \times \cdots \times I_N \times p}$. Thus, a tubal tensor of order $N$ in $\mathbb{C}_p^{I_1 \times I_2 \times \cdots \times I_N}$ is essentially a tensor of order $N + 1$ in $\mathbb{C}^{I_1 \times I_2 \times \cdots \times I_N \times p}$. Conversely, every tensor of order $N + 1$ in $\mathbb{C}^{I_1 \times I_2 \times \cdots \times I_N \times p}$ can be regarded as a tubal tensor of order $N$ in $\mathbb{C}_p^{I_1 \times I_2 \times \cdots \times I_N}$. In particular, third-order tensors can be identified as tubal matrices.

The Frobenius norm of a tubal tensor of order $N$ is defined to be the Frobenius norm of its underlying tensor of order $N + 1$.

We can unfold a tubal tensor into a tubal matrix. For the mode-$n$ unfolding, we follow the convention of Kolda and Bader (2009).

**Definition 3.4** Let $A \in \mathbb{C}_p^{I_1 \times I_2 \times \cdots \times I_N}$ be a tubal tensor of order $N$ with entries in $\mathbb{C}_p$. We define its mode-$n$ unfolding (where $1 \leq n \leq N$) to be the tubal matrix $A_{(n)} \in \mathbb{C}_p^{I_1 \times I_2 \times \cdots \times I_{n-1} \times I_{n+1} \times \cdots \times I_N}$.
\[ \mathbb{C}_p^{l_n \times (l_1 l_2 \ldots l_{n-1} l_n \ldots l_N)} , \text{ such that the } (i_1, \ldots, i_N)\text{-th tubal tensor entry maps to the } (i_n, j)\text{-th tubal matrix entry where} \]

\[ j = 1 + \sum_{k=1,k \neq n}^{N} (i_k - 1) \prod_{m=1,m \neq n}^{k-1} I_m. \]

In particular, for the mode-1 unfolding, the \((i_1, \ldots, i_N)\)-th entry of \(\mathcal{A}\) is mapped to the \((i_1, j)\)-th entry of \(\mathcal{A}_{(1)}\), where

\[ j = i_2 + (i_3 - 1)i_2 + (i_4 - 1)i_2i_3 + \cdots + (i_N - 1)i_2i_3 \ldots i_{N-1}. \] (3.5)

Next, we introduce the notion of mode-\(n\) \(t\)-rank of a tubal tensor.

**Definition 3.5** Let \(\mathcal{A} \in \mathbb{C}_p^{l_1 \times l_2 \times \cdots \times l_N}\) be a tubal tensor of order \(N\) and let \(\mathcal{A}_{(n)} \in \mathbb{C}_p^{l_n \times (l_1 l_2 \ldots l_{n-1} l_n \ldots l_N)}\) be its mode-\(n\) unfolding. Then, the \(t\)-rank of the tubal matrix \(\mathcal{A}_{(n)}\) is called the mode-\(n\) \(t\)-rank of \(\mathcal{A}\).

Finally, we define the \(n\)-mode product based on the tensor–tensor product and prove a proposition that connects the \(n\)-mode product with the tensor–tensor product of unfolding tubal matrices.

**Definition 3.6** The \(n\)-mode product of a tubal tensor \(\mathcal{A} \in \mathbb{C}_p^{l_1 \times \cdots \times l_N}\) by a tubal matrix \(\mathcal{U} \in \mathbb{C}_p^{l_n \times l_n}\) is the tubal tensor \(\mathcal{A} \ast_n \mathcal{U} \in \mathbb{C}_p^{l_1 \times \cdots \times l_{n-1} \times l_n \times l_{n+1} \times \cdots \times l_N}\) obtained by taking the tensor–tensor product of \(\mathcal{U}\) and the mode-\(n\) tubal vectors of \(\mathcal{A}\). Entrywisely, we have

\[ (\mathcal{A} \ast_n \mathcal{U})(i_1, \ldots, i_{n-1}, j, i_{n+1}, \ldots, i_N) = \sum_{i_n=1}^{l_n} \mathcal{A}(i_1, \ldots, i_n, \ldots, i_N) \ast \mathcal{U}(j, i_n). \]

From the above definition, \(\mathcal{A} \ast_n \mathcal{U} = \mathcal{B}\) is the same as \(\mathcal{U} \ast \mathcal{A}_{(n)} = \mathcal{B}_{(n)}\).

**Proposition 3.4** Let \(\mathcal{A} \in \mathbb{C}_p^{l_1 \times \cdots \times l_N}\). The following properties of the \(n\)-mode product based on tensor–tensor product hold:

(i) \((\mathcal{A} \ast_n \mathcal{F}) \ast_m \mathcal{G} = (\mathcal{A} \ast_m \mathcal{G}) \ast_n \mathcal{F}\) if \(m \neq n\).

(ii) \((\mathcal{A} \ast_m \mathcal{F}) \ast_n \mathcal{G} = \mathcal{A} \ast_n (\mathcal{G} \ast \mathcal{F})\).

**Proof** (i) Let \(m > n\), \(\mathcal{F} \in \mathbb{C}_p^{l_m \times l_m}, \mathcal{G} \in \mathbb{C}_p^{l_k \times l_m}\). For the entry whose index is \((i_1, \ldots, i_{n-1}, j, i_{n+1}, \ldots, i_{m-1}, k, i_{m+1}, \ldots, i_N)\), we have

\[
\sum_{i_m=1}^{l_m} \left( \sum_{i_n=1}^{l_n} \mathcal{A}(i_1, \ldots, i_N) \ast \mathcal{F}(j, i_n) \right) \ast \mathcal{G}(k, i_m) = \sum_{i_n=1}^{l_n} \left( \sum_{i_m=1}^{l_m} \mathcal{A}(i_1, \ldots, i_N) \ast \mathcal{G}(k, i_m) \right) \ast \mathcal{F}(j, i_n),
\]

where the equality comes from the associativity and commutativity of the tensor–tensor product of tubal scalars.
(ii) Let $F \in \mathbb{C}^{I \times I}$, $G \in \mathbb{C}^{K \times J}$. For the $(i_1, \ldots, i_{n-1}, k, i_{n+1}, \ldots, i_N)$-th entry, we have

$$
\sum_{j=1}^{J} \left\{ \left( \sum_{i_n=1}^{I_n} A(i_1, \ldots, i_N) * F(j, i_n) \right) * G(k, j) \right\} 
= \sum_{i_n=1}^{I_n} \left\{ A(i_1, \ldots, i_N) * \left( \sum_{j=1}^{J} F(j, i_n) * G(k, j) \right) \right\} 
= \sum_{i_n=1}^{I_n} \left\{ A(i_1, \ldots, i_N) * \left( \sum_{j=1}^{J} G(k, j) * F(j, i_n) \right) \right\}.
$$

where the equalities come from the associativity and commutativity of the tensor–tensor product of tubal scalars, respectively.

\[\square\]

The following proposition is the most crucial one in deriving the existence of Hot-SVD for tubal tensors.

**Proposition 3.5** (Tubal matrix representation of the $n$-mode product based on tensor–tensor product) Let $A, S \in \mathbb{C}^{I_1 \times \cdots \times I_N}$, $U_1 \in \mathbb{C}^{I_1 \times I_1}$, ..., $U_N \in \mathbb{C}^{I_N \times I_N}$. Then, for $n = 1, \ldots, N$,

$$
A = S *_1 U_1 *_2 U_2 *_3 \cdots *_N U_N
$$

is equivalent to

$$
A_{(n)} = U_n * S_{(n)} * \left( U_N \otimes \cdots \otimes U_{n+1} \otimes U_{n-1} \otimes \cdots \otimes U_1 \right)^{t}.
$$

**Proof** First of all, observe that we can reduce the claim to the case $n = 1$.

To see this, for an arbitrary $n$, define $\widetilde{A} \in \mathbb{C}^{I_1 \times I_1 \times I_2 \times \cdots \times I_{n-1} \times I_{n+1} \times \cdots \times I_N}$ by

$$
\widetilde{A}(i_n, i_1, i_2, \ldots, i_{n-1}, i_{n+1}, \ldots, i_N) = A(i_1, i_2, \ldots, i_n, \ldots, i_N).
$$

Then

$$
A_{(n)} = \widetilde{A}(1).
$$

Moreover

$$
A = S *_1 U_1 *_2 \cdots *_N U_N
$$

is the same as

$$
\widetilde{A} = \widetilde{S} *_1 U_n *_2 \cdots *_n U_{n-1} *_{n+1} U_{n+1} \cdots *_N U_N,
$$

and

$$
A_{(n)} = U_n * S_{(n)} * \left( U_N \otimes \cdots \otimes U_{n+1} \otimes U_{n-1} \otimes \cdots \otimes U_1 \right)^{t}
$$

is the same as

$$
\widetilde{A}_{(1)} = U_n * \widetilde{S}_{(1)} * \left( U_N \otimes \cdots \otimes U_{n+1} \otimes U_{n-1} \otimes \cdots \otimes U_1 \right)^{t}.
$$

Thus, we have reduced the general statement to the case $n = 1$. 

\[\square\]
Now, we prove the equivalence for \( n = 1 \). That is
\[
A = S \ast_1 U_1 \ast_2 U_2 \ast_3 \cdots \ast_N U_N
\]
if and only if
\[
A_{(1)} = U_1 \ast S_{(1)} \ast \left( U_N \otimes U_{N-1} \otimes \cdots \otimes U_2 \right)^t.
\]

By our convention for unfolding Eq. (3.5), we have
\[
A_{(1)}(i_1, j) = A(i_1, i_2, \ldots, i_N), \tag{3.6}
\]
where
\[
j = i_2 + (i_3 - 1)I_2 + (i_4 - 1)I_2I_3 + \cdots + (i_N - 1)I_2I_3 \cdots I_{N-1}. \tag{3.7}
\]

Similarly, we have
\[
S_{(1)}(i'_1, j') = S(i'_1, i'_2, \ldots, i'_N), \tag{3.8}
\]
where
\[
j' = i'_2 + (i'_3 - 1)I_2 + (i'_4 - 1)I_2I_3 + \cdots + (i'_N - 1)I_2I_3 \cdots I_{N-1}. \tag{3.9}
\]

Define
\[
V = \left( U_N \otimes \cdots \otimes U_2 \right)^t \in C_{I_2 \times I_2}^{I_N}, \quad W = U_1 \ast S_{(1)} \ast V.
\]

According to the definition of Kronecker product Eq. (3.2), \( V = U_N^t \otimes \cdots \otimes U_2^t \in C_{I_2 \times I_2}^{(I_N \times I_{N-1} \times \cdots \times I_2)} \), and the tubal scalar \( U_N^t(i'_N, i_N) \ast \cdots \ast U_2^t(i'_2, i_2) \) is the entry of \( V \) whose index is
\[
\left( i'_2 + (i'_3 - 1)I_2 + \cdots + (i'_N - 1)I_2I_3 \cdots I_{N-1} \right).
\]

It is exactly the \((j', j)\)-th entry of \( V \) according to Eqs. (3.7) and (3.9). Therefore, we have
\[
V(j', j) = \left( U_N^t \otimes \cdots \otimes U_2^t \right)(j', j) = U_N^t(i'_N, i_N) \ast \cdots \ast U_2^t(i'_2, i_2). \tag{3.10}
\]
Then, we have
\[
\mathcal{W}(i_1, j) = (\mathcal{U}_1 * S_{(1)} * \mathcal{V})(i_1, j)
\]
\[
= \sum_{i_1', j'} \mathcal{U}_1(i_1, i_1') * S(i_1', i_2', \ldots, i_N') * \mathcal{U}_N(i_N') * \cdots * \mathcal{U}_2(i_2', i_2)
\]
\[
= \sum_{i_1', i_2', \ldots, i_N'} S(i_1', i_2', \ldots, i_N') * \mathcal{U}_1(i_1, i_1') * \cdots * \mathcal{U}_N(i_N, i_N')
\]
where the third equality holds due to Eqs. (3.8) and (3.10) and the last equality comes from the commutativity of the tensor–tensor product of tubal scalars.

Now, we are ready to prove that \( \mathcal{A} = S * \mathcal{U}_1 * \cdots * \mathcal{U}_N \) if and only if \( \mathcal{A}_{(1)} = \mathcal{W} = \mathcal{U}_1 * S_{(1)} * \mathcal{V} \).

If \( \mathcal{A} = S * \mathcal{U}_1 * \cdots * \mathcal{U}_N \), this means that entrywisely
\[
\mathcal{A}(i_1, \ldots, i_N) = \sum_{i_1', \ldots, i_N'} S(i_1', \ldots, i_N') * \mathcal{U}_1(i_1, i_1') * \cdots * \mathcal{U}_N(i_N, i_N'),
\]
then, according to Eq. (3.11)
\[
\mathcal{W}(i_1, j) = \mathcal{A}(i_1, \ldots, i_N) = \mathcal{A}_{(1)}(i_1, j),
\]
where \( j \) is as defined in Eq. (3.7). Therefore, we have \( \mathcal{A}_{(1)} = \mathcal{W} \).

Conversely, if \( \mathcal{W} = \mathcal{A}_{(1)} \), then
\[
\mathcal{A}(i_1, \ldots, i_N) = \mathcal{A}_{(1)}(i_1, j) = \mathcal{W}(i_1, j) = \sum_{i_1', \ldots, i_N'} S(i_1', \ldots, i_N') * \mathcal{U}_1(i_1, i_1') * \cdots * \mathcal{U}_N(i_N, i_N'),
\]
where \( j \) is as defined in Eq. (3.7) and the last equality is due to Eq. (3.11). Therefore, we have \( \mathcal{A} = S * \mathcal{U}_1 * \cdots * \mathcal{U}_N \).

\[\square\]

4 Higher order t-SVD of tubal tensors

With these preparations, we are ready to establish the Hot-SVD of tubal tensors. To facilitate the comparison with HOSVD of (usual) tensors, we quote the following theorem.

Theorem 4.1 (HOSVD of tensors, De Lathauwer et al. 2000, Theorem 2) Every tensor \( \mathcal{A} \in \mathbb{C}^{I_1 \times \cdots \times I_N} \) can be written as the product
\[
\mathcal{A} = S \times_1 \mathcal{U}_1 \cdots \times_N \mathcal{U}_N.
\]
where

1. $U_n$ is a unitary matrix for $n = 1, \ldots, N$,
2. $S \in \mathbb{C}^{I_1 \times \cdots \times I_N}$ has the following properties:
   (i) all-orthogonality:
       \[ \langle S_{in=\alpha}, S_{in=\beta} \rangle = 0 \]
       for $1 \leq \alpha \neq \beta \leq I_n$, where $S_{in=\alpha}$ is the $(N-1)$-th order tensor obtained by fixing the $n$-th index of $S$ to be $\alpha$.
   (ii) ordering:
       \[ \|S_{in=1}\| \geq \|S_{in=2}\| \geq \cdots \geq \|S_{in=I_n}\| \]
       for all possible values of $n$, where $\|S_{in=\alpha}\|$ is the same as the Frobenius norm of the $\alpha$-th row tubal vector of the mode-$n$ unfolding tubal matrix $S_{(n)}$.

Now, we state and prove the tubal version of the above theorem, based on tensor–tensor product.

**Theorem 4.2** (Hot-SVD of tubal tensors) Any tubal tensor $A \in \mathbb{C}^{I_1 \times \cdots \times I_N}$ can be written as the product

\[ A = S \ast_1 U_1 \cdots \ast_N U_N, \]

where

1. $U_n \in \mathbb{C}^{I_n \times I_n}$ is a unitary tubal matrix for $n = 1, \ldots, N$,
2. $S \in \mathbb{C}^{I_1 \times \cdots \times I_N}$ has the following properties:
   (i) all-orthogonality: for all $1 \leq n \leq N$ and $1 \leq \alpha \neq \beta \leq I_n$,
       \[ \sum S(i_1, \ldots, i_{n-1}, \alpha, i_{n+1}, \ldots, i_N) \ast S(i_1, \ldots, i_{n-1}, \beta, i_{n+1}, \ldots, i_N)^H = 0, \]
       where the sum is taken over $i_1, \ldots, i_{n-1}, i_{n+1}, \ldots, i_N$ and $0 = (0, \ldots, 0) \in \mathbb{C}_p$.
   (ii) ordering: for $L = cW$ where $W$ is a unitary transformation and $c \in \mathbb{C}$ is a non-zero scalar, we have further
       \[ \|S_{in=1}\| \geq \|S_{in=2}\| \geq \cdots \geq \|S_{in=I_n}\| \]
       for all possible values of $n$, where $\|S_{in=\alpha}\|$ is the same as the Frobenius norm of the $\alpha$-th row tubal vector of the mode-$n$ unfolding tubal matrix $S_{(n)}$.

Decomposition of the form in Theorem 4.2 is called higher order t-SVD (Hot-SVD) in this work.

**Proof** According to Proposition 3.5, we only need to prove that the equivalent tubal matrix representation

\[ A_{(n)} = U_n \ast S_{(n)} \ast \left( U_N \otimes \cdots \otimes U_{n+1} \otimes U_{n-1} \otimes \cdots \otimes U_1 \right)^t \]

holds for all $n$.

For $n = 1, \ldots, N$, let $A_{(n)} = U_n \ast \Sigma_n \ast V_n^H$ be a t-SVD of the tubal matrix $A_{(n)}$. Define

\[ S = A \ast_1 U_1^H \cdots \ast_N U_N^H. \]
Then, by Proposition 3.5

\[ S_{(n)} = U_n^H \star A_{(n)} \star \left( U_N^H \otimes \cdots \otimes U_{n+1}^H \otimes U_{n-1}^H \cdots \otimes U_1^H \right)^t. \]

Multiplying both sides by \( U_n \) and \( (U_N \otimes \cdots \otimes U_{n+1} \otimes U_{n-1} \otimes \cdots \otimes U_1)^t \), we have

\[
U_n \star S_{(n)} \star \left( U_N \otimes \cdots \otimes U_{n+1} \otimes U_{n-1} \otimes \cdots \otimes U_1 \right)^t = U_n \star U_n^H \star A_{(n)} \star \left( U_N^H \otimes \cdots \otimes U_{n+1}^H \otimes U_{n-1}^H \cdots \otimes U_1^H \right)^t \times \left( U_N \otimes \cdots \otimes U_{n+1} \otimes U_{n-1} \otimes \cdots \otimes U_1 \right) = I_{I_n} \star A_{(n)} \star \left( U_N \otimes \cdots \otimes U_{n+1} \otimes U_{n-1} \otimes \cdots \otimes U_1 \right).
\]

\[ (3.3, i) \]

\[
= I_{I_n} \star A_{(n)} \star \left( U_N^H \otimes \cdots \otimes U_{n+1}^H \otimes U_{n-1}^H \cdots \otimes U_1^H \otimes (U_{n+1}^H \star U_{n+1}^H) \right) \times (U_N \otimes \cdots \otimes U_{n+1} \otimes U_{n-1} \otimes \cdots \otimes U_1) = I_{I_n} \star A_{(n)} \star \left( I_{I_N} \otimes \cdots \otimes I_{I_{n+1}} \otimes I_{I_{n-1}} \otimes \cdots \otimes I_{I_1} \right) = A_{(n)}.
\]

where the tubal matrices \( U_n \)'s are unitary, since they come from t-SVD.

From

\[ A_{(n)} = U_n \star \Sigma_n \star \gamma_n^H. \]

and

\[ A_{(n)} = U_n \star S_{(n)} \star \left( U_N \otimes \cdots \otimes U_{n+1} \otimes U_{n-1} \otimes \cdots \otimes U_1 \right)^t, \]

we obtain

\[
U_n \star S_{(n)} \star \left( U_N \otimes \cdots \otimes U_{n+1} \otimes U_{n-1} \otimes \cdots \otimes U_1 \right)^t = U_n \star \Sigma_n \star \gamma_n^H.
\]
Multiplying both sides by $U_n^H$ and $(U_N \otimes \cdots \otimes U_{n+1} \otimes U_{n-1} \otimes \cdots \otimes U_1)^t H$ and noticing that the latter tubal matrix is unitary by Proposition 3.2 and Proposition 3.3, (v), we have

$$S_{(n)} = U_n^H \ast U_n \ast \Sigma_n \ast V_n^H \ast \left((U_N \otimes \cdots \otimes U_{n+1} \otimes U_{n-1} \otimes \cdots \otimes U_1)^t H\right)$$

$$= \Sigma_n \ast V_n^H \ast (U_N \otimes \cdots \otimes U_{n+1} \otimes U_{n-1} \otimes \cdots \otimes U_1)^t H$$

$$\equiv \Sigma_n \ast V_n^H \ast \left(U'_N \otimes \cdots \otimes U'_{n+1} \otimes U'_{n-1} \otimes \cdots \otimes U'_1\right)^H.$$

Since the tubal matrices $U'_i (i = 1, \ldots, N)$ are unitary (by Proposition 3.2), the tubal matrix

$$(U'_N \otimes \cdots \otimes U'_{n+1} \otimes U'_{n-1} \otimes \cdots \otimes U'_1)^H$$

is also unitary by Proposition 3.3, (v). Then, the tubal matrix

$$V = V_n^H \ast \left(U'_N \otimes \cdots \otimes U'_{n+1} \otimes U'_{n-1} \otimes \cdots \otimes U'_1\right)^H$$

is unitary. As $\Sigma_n$ is an $f$-diagonal tubal matrix, we conclude that the row tubal vectors of $S_{(n)}$, being tubal scalar multiples of the row tubal vectors of the unitary tubal matrix $V$, are orthogonal to each other with respect to the tensor–tensor product, whence the all-orthogonality of $S$.

For the ordering property, observe first that $\|S_{n=\alpha}\| = \|S_{(n)}(\alpha, :)\| = \|\Sigma_n(\alpha, \alpha)\ast W(\alpha, :)\| = \|\Sigma_n(\alpha, \alpha)\|$, where the last equality follows from the unitary invariance of the Frobenius norm when $L = c W$ (Theorem 2.4). From the construction of t-SVD, each frontal slice of $L(\Sigma_n)$ has non-increasing singular values, implying that the Frobenius norms $\|L(\Sigma_n)(\alpha, \alpha)\|$ are non-increasing, and consequently, the Frobenius norms

$$\|\Sigma_n(\alpha, \alpha)\| = \|L^{-1}(L(\Sigma_n)(\alpha, \alpha))\|$$

are non-increasing

$$\|\Sigma_n(1, 1)\| \geq \|\Sigma_n(2, 2)\| \geq \cdots \geq \|\Sigma_n(I_N, I_N)\|,$$

since $\|L^{-1}(\alpha)\| = \|c^{-1}W^H(\alpha)\| = \|c^{-1}\| \cdot \|W^H(\alpha)\| = \|c^{-1}\| \cdot \|\alpha\|$ for any tubal scalar $\alpha \in \mathbb{C}_p$. Then

$$\|S_{n=1}\| \geq \|S_{n=2}\| \geq \cdots \geq \|S_{n=I_n}\|,$$

since $\|S_{n=\alpha}\| = \|\Sigma_n(\alpha, \alpha)\|$ as noted above. \hfill \Box

Remark 4.1 The above proof tells us that Proposition 3.5 plays a key role to show the validity of Hot-SVD, and we emphasize that the link in Proposition 3.5 between the Hot-SVD of a tubal tensor $A$ and the t-SVD of the unfolding $A_{(n)}$ of $A$ does not hold if we replace the small-t transpose with the usual capital-T transpose.

The above proof actually indicates how Hot-SVD of a given tubal tensor can be computed: the tubal matrix $U_0$ can be directly found through the t-SVD of the unfolding tubal matrix $A_{(n)}$, and the core tubal tensor $S$ can be computed by the $n$-mode product. When $L$ is the DFT, the computational complexity of this procedure is of the same order as that of t-SVD for higher order tensors defined in Martin et al. (2013).
Many properties of HOSVD have clear counterparts in our model based on the tensor–
tensor product.

**Property 4.1** (Generalization) The Hot-SVD of a tubal matrix boils down to the t-SVD.

This is obvious from Proposition 3.5.

**Property 4.2** [n-rank] The mode-n t-rank of $\mathcal{A}$ is the same as the highest index $r_n$ for which $\|S_{i_{n}=r_n}\| > 0$.

**Proof** We use the notations from the proof of Theorem 4.2. By definition, the mode-n t-
rank of $\mathcal{A}$ is the number of non-zero tubal scalars on the diagonal of $\Sigma_n$. From the proof of Theorem 4.2, we have

$$\|S_{i_{n}=r_n}\| = \|\Sigma_n(r_n, r_n)\|.$$ Thus, the highest index $r_n$ for which $\|S_{i_{n}=r_n}\| > 0$ is the same as the number of non-zero tubal scalars on the diagonal of $\Sigma_n$, i.e., the mode-n t-rank of $\mathcal{A}$.

**Property 4.3** (Link between Hot-SVD and t-SVD) The Hot-SVD gives a thin t-SVD of $\mathcal{A}(n)$ by normalizing $S_{i_{n}}$ to extract a diagonal tubal matrix.

**Proof** If $\mathcal{A} = S \ast U_1 \cdots \ast U_N$, then

$$\mathcal{A}(n) = U_n \ast S_{i_{n}} \ast \left(U_N \otimes \cdots \otimes U_{n+1} \otimes U_{n-1} \otimes \cdots \otimes U_1\right)^t,$$

where $S_{i_{n}}$ has mutually orthogonal tubal rows, whose Frobenius norms are $\sigma_1(n), \ldots, \sigma_{i_{n}}(n)$.

Define

$$\Sigma_n = \text{diag}(\sigma_1(n) \mathbf{1}, \sigma_2(n) \mathbf{1}, \ldots, \sigma_{i_{n}}(n) \mathbf{1}),$$

where $\mathbf{1} = L^{-1}((1, 1, \ldots, 1)) \in \mathbb{C}_p$ is the identity tubal scalar in $\mathbb{C}_p$.

Let $\tilde{S}_n$ be the normalized version of $S_{i_{n}}$, i.e., $\tilde{S}_n = \Sigma_n \ast \tilde{S}_n$. Define

$$\mathcal{V}_n^H = \tilde{S}_n \ast \left(U_N \otimes \cdots \otimes U_{n+1} \otimes U_{n-1} \otimes \cdots \otimes U_1\right)^t,$$

which is a tubal matrix with orthonormal tubal rows. Then, $\mathcal{A}(n) = U_n \ast \Sigma_n \ast \mathcal{V}_n^H$ is a thin t-SVD of $\mathcal{A}(n) \in \mathbb{C}^{I_p \times (I_1 I_2 \cdots I_{n-1} I_{n+1} \cdots I_N)}_p$ with $U_n \in \mathbb{C}^{I_p \times I_n}$, $\Sigma_n \in \mathbb{C}^{I_n \times I_n}$ and $\mathcal{V}_n^H \in \mathbb{C}^{I_n \times (I_1 I_2 \cdots I_{n-1} I_{n+1} \cdots I_N)}$. We can modify appropriately the number of tubal columns of $\Sigma_n$ and the number of orthonormal tubal rows of $\mathcal{V}_n^H$ to get a full t-SVD $\mathcal{A} = U_h \ast \Sigma_n \ast \mathcal{V}_n^H$ with $U_h \in \mathbb{C}^{I_h \times I_n}$, $\Sigma_n \in \mathbb{C}^{I_n \times (I_1 I_2 \cdots I_{n-1} I_{n+1} \cdots I_N)}$ and $\mathcal{V}_n^H \in \mathbb{C}^{(I_1 I_2 \cdots I_{n-1} I_{n+1} \cdots I_N) \times (I_1 I_2 \cdots I_{n-1} I_{n+1} \cdots I_N)}$. □

**Property 4.4** (Structure) Let $\mathcal{A}(n) = U_n \ast \Sigma_n \ast \mathcal{V}_n^H$ be a full t-SVD of $\mathcal{A}(n) \in \mathbb{C}^{I_p \times (I_1 I_2 \cdots I_{n-1} I_{n+1} \cdots I_N)}_p$ obtained from the Hot-SVD of $\mathcal{A} \in \mathbb{C}^{I_h \times (I_1 I_2 \cdots I_{n-1} I_{n+1} \cdots I_N)}$. Denote the tubal scalars on the diagonal of $\Sigma_n$ by $s_i$. Suppose that $s_1, \ldots, s_j$ are invertible, $s_{j+1}, \ldots, s_{j+k}$ are non-zero but not invertible, and the remaining $s_{j+k+1}, \ldots, s_{\min(l,m)}$ are zero (where $l = I_h, m = I_1 I_2 \cdots I_{n-1} I_{n+1} \cdots I_N$). Then

i) The range of $\mathcal{A}(n)$ is

$$\{U_n(\cdot, 1) \ast c_1 + \cdots + U_n(\cdot, j + k) \ast c_{j+k} | c_i = s_i \ast d_i, d_i \in \mathbb{C}_p, j + 1 \leq i \leq j + k\}.$$
Algorithm 1 tr-Hot-SVD($\mathcal{A}, I'_1, I'_2, \ldots, I'_N$)

while $1 \leq n \leq N$ do
    $\hat{U}_n \leftarrow I'_n$ leading left singular tubal vectors of $\mathcal{A}(n)$
end while
$S \leftarrow \mathcal{A} \ast_1 \hat{U}_1^H \ast_2 \hat{U}_2^H \cdots \ast_N \hat{U}_N^H$
return $S, \hat{U}_1, \hat{U}_2, \ldots, \hat{U}_N$

(ii) The kernel of $\mathcal{A}(n)$ is
\[
\left\{ V_n(:, j+1) \ast c_{j+1} + \cdots + V_n(:, m) \ast c_m \mid s_i \ast c_i = 0, c_i \in \mathbb{C}_p, j+1 \leq i \leq j+k \right\}.
\]
This is exactly Theorem 2.3 as applied to the tubal matrix $\mathcal{A}(n)$.

Property 4.5 (Norm) When $L = cW$, where $W$ is a unitary transformation and $c \in \mathbb{C}$ is non-zero scalar, we have
\[
\| \mathcal{A} \| = \| S \|.
\]
This follows from the unitary invariance of the Frobenius norm of tubal matrices (Theorem 2.4).

5 Truncated Hot-SVD and sequentially truncated Hot-SVD

This section derives the truncated Hot-SVD and sequentially truncated Hot-SVD, which generalize those of De Lathauwer et al. (2000) and Vannieuwenhoven et al. (2012) to the tubal setting. The algorithms are depicted in Algorithms 1 and 2.

Algorithm 2 seq-tr-Hot-SVD($\mathcal{A}, I'_1, I'_2, \ldots, I'_N$)

$\hat{S} \leftarrow \mathcal{A}$
while $1 \leq n \leq N$ do
    $\hat{U}_n \leftarrow I'_n$ leading left singular tubal vectors of $\hat{S}(n)$
end while
$\tilde{S} \leftarrow \mathcal{A} \ast_1 \hat{U}_1^H \ast_2 \hat{U}_2^H \cdots \ast_N \hat{U}_N^H$
return $\tilde{S}, \hat{U}_1, \hat{U}_2, \ldots, \hat{U}_N$

To prove an error bound for truncated Hot-SVD and sequentially truncated Hot-SVD, we first need some technical preparations.

Proposition 5.1 Suppose $L = cW$, where $c \in \mathbb{C}$ is a non-zero scalar and $W$ is a unitary transformation. Let $\mathcal{A}_1, \ldots, \mathcal{A}_N \in \mathbb{C}_{I_1 \times I_2}^p$ be tubal matrices that are orthogonal to each other with respect to the tensor–tensor product: for $1 \leq m \neq n \leq N$
\[
\mathcal{A}_m^H \ast \mathcal{A}_n = 0.
\]
Then $\mathcal{A}_1, \ldots, \mathcal{A}_N$ are mutually orthogonal: for $1 \leq m \neq n \leq N$
\[
\langle \mathcal{A}_m, \mathcal{A}_n \rangle = 0.
\]
Consequently, we have
\[
\| \mathcal{A}_1 \|^2 + \| \mathcal{A}_2 \|^2 + \cdots + \| \mathcal{A}_N \|^2 = \| \mathcal{A}_1 + \mathcal{A}_2 + \cdots + \mathcal{A}_N \|^2.
\]
**Proof** Since $A_m^H \ast A_n = O$, we have for $k = 1, \ldots, p$,

$$L(A_m)^{(k)H}L(A_n)^{(k)} = 0,$$

where 0 above denotes the zero matrix of the proper size. This implies that

$$\langle L(A_m)^{(k)}, L(A_n)^{(k)} \rangle = \text{tr} \left( L(A_m)^{(k)H}L(A_n)^{(k)} \right) = 0. \quad (5.1)$$

Therefore, the matrices $L(A_1)^{(k)}, \ldots, L(A_N)^{(k)}$ are orthogonal to each other in the Frobenius norm. Then, it follows that:

$$\langle A_m, A_n \rangle = \sum_{i_1, i_2} A_m(i_1, i_2)^{(k)} A_n(i_1, i_2)^{(k)}$$

$$= \sum_{i_1, i_2} \left( \sum_{k=1}^{p} A_m(i_1, i_2)^{(k)} A_n(i_1, i_2)^{(k)} \right)$$

$$= \sum_{i_1, i_2} \langle A_m(i_1, i_2), A_n(i_1, i_2) \rangle$$

$$= \sum_{i_1, i_2} \frac{1}{|c|^2} \langle L(A_m(i_1, i_2)), L(A_n(i_1, i_2)) \rangle$$

$$= \sum_{i_1, i_2} \frac{1}{|c|^2} \langle L(A_m)(i_1, i_2), L(A_n)(i_1, i_2) \rangle$$

$$= \frac{1}{|c|^2} \sum_{i_1, i_2, k} (L(A_m)(i_1, i_2))^{(k)} (L(A_n)(i_1, i_2))^{(k)}$$

$$= \frac{1}{|c|^2} \sum_{k=1}^{p} \left( \sum_{i_1, i_2} (L(A_m)(i_1, i_2))^{(k)} (L(A_n)(i_1, i_2))^{(k)} \right)$$

$$= \frac{1}{|c|^2} \sum_{k=1}^{p} \left( L(A_m)^{(k)}, L(A_n)^{(k)} \right)$$

$$\quad (5.2)$$

Finally, the equality concerning the Frobenius norms follows from the bi-linearity of the inner product.

**Proposition 5.2** Suppose $L = cW$, where $c \in \mathbb{C}$ is a non-zero scalar and $W$ is a unitary transformation. Let $A \in \mathbb{C}_{I_p}^{I_1 \times \cdots \times I_N}$ be a tubal tensor and $\tilde{U}_n \in \mathbb{C}_{I_p}^{I_p \times R_n}$ ($I_p \geq R_n$) be a partially unitary tubal matrix (i.e., $\tilde{U}_n^H \ast \tilde{U}_n = I$). Then

$$\| A \ast_n \left( \tilde{U}_n \ast \tilde{U}_n^H \right) \|^2 + \| A \ast_n \left( I - \tilde{U}_n \ast \tilde{U}_n^H \right) \|^2 = \| A \|^2.$$  

In particular, $\| A \ast_n \left( \tilde{U}_n^{(n)} \ast \tilde{U}_n^H \right) \| \leq \| A \|$. 


Proof Note that \((\tilde{U}_n \ast \tilde{U}_n^H)^H \ast (I - \tilde{U}_n \ast \tilde{U}_n^H) = \tilde{U}_n \ast \tilde{U}_n^H - \tilde{U}_n \ast \tilde{U}_n^H \ast \tilde{U}_n \ast \tilde{U}_n^H = O\), so Proposition 5.1 implies that

\[
\|A \ast_n (\tilde{U}_n \ast \tilde{U}_n^H)\|_2^2 + \|A \ast_n (I - \tilde{U}_n \ast \tilde{U}_n^H)\|_2^2 = \| (\tilde{U}_n \ast \tilde{U}_n^H) \ast A(n)\|_2^2 + \| (I - \tilde{U}_n \ast \tilde{U}_n^H) \ast A(n)\|_2^2 = \|A(n)\|_2^2 = \|A\|_2^2.
\]

\[\square\]

Proposition 5.3 Suppose \(L = cW\), where \(c \in \mathbb{C}\) is a non-zero scalar and \(W\) is a unitary transformation. Let \(A \in \mathbb{C}^{l_1 \times l_2 \times \cdots \times l_N}\) be a tubal tensor of order \(N\). Let \(A\) be approximated by

\[
\tilde{A} = A \ast_1 (\tilde{U}_1 \ast \tilde{U}_1^H) \ast_2 (\tilde{U}_2 \ast \tilde{U}_2^H) \ast_3 (\tilde{U}_3 \ast \tilde{U}_3^H) \ast \cdots \ast_0 (\tilde{U}_0 \ast \tilde{U}_0^H),
\]

where \(\tilde{U}_n \in \mathbb{C}^{l_n \times r_n}\) (\(I_n \geq R_n, n = 1, \ldots, N\)) are partially unitary tubal matrices (i.e., \(\tilde{U}_n^H \ast \tilde{U}_n = I\)). Then, the squared approximation error is

\[
\|A - \tilde{A}\|_2^2 = \|A \ast_1 (I - \tilde{U}_1 \ast \tilde{U}_1^H)\|_2^2 + \|A \ast_1 (\tilde{U}_1 \ast \tilde{U}_1^H) \ast_2 (I - \tilde{U}_2 \ast \tilde{U}_2^H)\|_2^2 + \|A \ast_1 (\tilde{U}_1 \ast \tilde{U}_1^H) \ast_2 (\tilde{U}_2 \ast \tilde{U}_2^H) \ast_3 (I - \tilde{U}_3 \ast \tilde{U}_3^H)\|_2^2 + \cdots + \|A \ast_1 (\tilde{U}_1 \ast \tilde{U}_1^H) \ast_2 (\tilde{U}_2 \ast \tilde{U}_2^H) \ast_3 (\tilde{U}_3 \ast \tilde{U}_3^H) \ast_4 (I - \tilde{U}_4 \ast \tilde{U}_4^H)\|_2^2.
\]

Proof Note that

\[
A - \tilde{A} = A - A \ast_1 (\tilde{U}_1 \ast \tilde{U}_1^H) + A \ast_1 (\tilde{U}_1 \ast \tilde{U}_1^H)
\]

\[
- A \ast_1 (\tilde{U}_1 \ast \tilde{U}_1^H) \ast_2 (\tilde{U}_2 \ast \tilde{U}_2^H)
\]

\[
+ A \ast_1 (\tilde{U}_1 \ast \tilde{U}_1^H) \ast_2 (\tilde{U}_2 \ast \tilde{U}_2^H)
\]

\[
\ldots
\]

\[
- A \ast_1 (\tilde{U}_1 \ast \tilde{U}_1^H) \ast_2 (\tilde{U}_2 \ast \tilde{U}_2^H) \ast \cdots \ast_0 (\tilde{U}_0 \ast \tilde{U}_0^H)
\]

\[
= A \ast_1 (I - \tilde{U}_1 \ast \tilde{U}_1^H)
\]

\[
+ A \ast_1 (\tilde{U}_1 \ast \tilde{U}_1^H) \ast_2 (I - \tilde{U}_2 \ast \tilde{U}_2^H)
\]

\[
+ \cdots
\]

\[
+ A \ast_1 (\tilde{U}_1 \ast \tilde{U}_1^H) \ast_2 (\tilde{U}_2 \ast \tilde{U}_2^H) \ast \cdots \ast_0 (\tilde{U}_0 \ast \tilde{U}_0^H).
\]

\[\square\]
We now show that any two distinct terms in the above expression are orthogonal to each other in the Frobenius norm. For $i < j$, let $\tilde{B}, \tilde{C}$ be the $i$-th and $j$-th terms in the above expression, respectively. Consider

$$\tilde{B}_{(i)} := (I - \tilde{u}_i * \tilde{u}_i^H) * A_{(i)} * \left( I \otimes \cdots \otimes I \otimes \left( \tilde{u}_{i-1} * \tilde{u}_{i-1}^H \right) \otimes \cdots \otimes (\tilde{u}_1 * \tilde{u}_1^H) \right)^t$$

and

$$\tilde{C}_{(i)} := (\tilde{u}_i * \tilde{u}_i^H) * A_{(i)} * \left( I \otimes \cdots \otimes I \otimes \left( I - \tilde{u}_j * \tilde{u}_j^H \right) \otimes (\tilde{u}_{j-1} * \tilde{u}_{j-1}^H) \otimes \cdots \otimes \left( \tilde{u}_1 * \tilde{u}_1^H \right) \right)^t.$$

Then $\tilde{C}_{(i)}^H * \tilde{B}_{(i)} = 0$, since $(\tilde{u}_i * \tilde{u}_i^H) * (I - \tilde{u}_i * \tilde{u}_i^H) = 0$. Then, Proposition 5.1 implies that the $i$-th term and the $j$-th term are orthogonal in the Frobenius norm.

Now, we consider the truncated Hot-SVD.

**Theorem 5.1** (Error bound for tr-Hot-SVD) Suppose $L = cW$, where $c \in \mathbb{C}$ is a non-zero scalar and $W : \mathbb{C}_p \rightarrow \mathbb{C}_p$ is a unitary transformation. Let $A \in \mathbb{C}_p^{I_1 \times \cdots \times I_N}$ be a tubal tensor and $A = S *_1 U_1 \cdots *_N U_N$ its Hot-SVD. Define $\hat{A}$ to be

$$\hat{A} = \hat{S} *_1 U_1 \cdots *_N U_N,$$

where $\hat{S} \in \mathbb{C}_p^{I_1 \times \cdots \times I_N}$ is obtained by truncation of the first $I_1' \times \cdots \times I_N'$ tubal scalars of $S$ (with other tubal scalars being zero). Then

$$\| A - \hat{A} \| \leq \sqrt{N} \sum_{i_1 = I_1' + 1}^{R_1} (\sigma_{i_1}^{(1)})^2 + \cdots + \sqrt{N} \sum_{i_N = I_N' + 1}^{R_N} (\sigma_{i_N}^{(N)})^2,$$

where $R_n$ is the t-rank of $A_{(n)}$, $\sigma_{i_n}^{(n)}$ is the Frobenius norm of the $i_n$-th row of $S_{(n)}$ (which is equal to the Frobenius norm of the $i_n$-th tubal scalar on the diagonal of $\Sigma_n$, the $f$-diagonal tubal matrix appearing in the t-SVD of $A_{(n)}$), $A^* = S^* *_1 U_1^* \cdots *_N U_N^*$, and $(S^*, U_1^*, \ldots, U_N^*)$ is the optimal solution to the following (tensor–tensor product-based) low-rank approximation problem:

$$\min_{S, U_1, \ldots, U_N} \| A - S *_1 U_1 \cdots *_N U_N \|,$$

s.t. $U_1^H * U_1 = I_{I_1'}, \ldots, U_N^H * U_N = I_{I_N'},$

$U_n \in \mathbb{C}_p^{I_n' \times I_n'}$ (for $n = 1, \ldots, N$), $S \in \mathbb{C}_p^{I_1' \times \cdots \times I_N'}$.

**Proof** Let $A_{(n)} = U_n * \Sigma_n * V_n^H$ be the t-SVD of the mode-$n$ unfolding. Define $\overline{U}_n$ to be the tubal matrix obtained from the first $I_n'$ tubal column vectors of $U_n$. By definition

$$\hat{A} = \hat{S} *_1 U_1 \cdots *_N U_N = \overline{S} *_1 \overline{U}_1 \cdots *_N \overline{U}_N,$$
where \( \overline{S} \in C_{p}^{I_{1} \times \cdots \times I_{N}} \) is obtained by truncation of \( S \). Note also that

\[
\overline{S} = A \ast_{1} \overline{U}_{1}^{H} \cdots \ast_{N} \overline{U}_{N}^{H},
\]

since \( S = A \ast_{1} U_{1}^{H} \cdots \ast_{N} U_{N}^{H} \). Then, we have

\[
\hat{A} = (A \ast_{1} \overline{U}_{1}^{H} \cdots \ast_{N} \overline{U}_{N}^{H}) \ast_{1} U_{1} \cdots \ast_{N} U_{N}
\]

(5.4)

Then, by Proposition 5.3, we have

\[
\|A - \hat{A}\|^{2} = \|A - A \ast_{1} (\overline{U}_{1} \ast \overline{U}_{1}^{H}) \cdots \ast_{N} (\overline{U}_{N} \ast \overline{U}_{N}^{H})\|^{2}
\]

\[
= \|A \ast_{1} \left( I - \overline{U}_{1} \ast \overline{U}_{1}^{H} \right)\|^{2}
\]

\[
+ \|A \ast_{1} (\overline{U}_{1} \ast \overline{U}_{1}^{H}) \ast_{2} (I - \overline{U}_{2} \ast \overline{U}_{2}^{H})\|^{2}
\]

\[
+ \|A \ast_{1} (\overline{U}_{1} \ast \overline{U}_{1}^{H}) \ast_{2} (\overline{U}_{2} \ast \overline{U}_{2}^{H}) \ast_{3} (I - \overline{U}_{3} \ast \overline{U}_{3}^{H})\|^{2}
\]

\[
+ \cdots
\]

\[
+ \|A \ast_{1} (\overline{U}_{1} \ast \overline{U}_{1}^{H}) \cdots \ast_{N-1} (\overline{U}_{N-1} \ast \overline{U}_{N-1}^{H}) \ast_{N} (I - \overline{U}_{N} \ast \overline{U}_{N}^{H})\|^{2}
\]

(5.5)

For the \( n \)-th term in the above expression, by Proposition 5.2, we have

\[
\|A \ast_{1} (\overline{U}_{1} \ast \overline{U}_{1}^{H}) \cdots \ast_{n-1} (\overline{U}_{n-1} \ast \overline{U}_{n-1}^{H}) \ast_{n} (I - \overline{U}_{n} \ast \overline{U}_{n}^{H})\| \leq \|A \ast_{n} (I - \overline{U}_{n} \ast \overline{U}_{n}^{H})\|.
\]

(5.6)

Note that

\[
\|A \ast_{n} (I - \overline{U}_{n} \ast \overline{U}_{n}^{H})\|^{2}
\]

\[
= \| (I - \overline{U}_{n} \ast \overline{U}_{n}^{H}) \ast A_{(n)}\|^{2}
\]

\[
= \| (I - \overline{U}_{n} \ast \overline{U}_{n}^{H}) \ast \overline{U}_{n} \ast \Sigma_{n} \ast \mathcal{V}_{n}^{H}\|^{2}
\]

\[
= \| (\overline{U}_{n} - \overline{U}_{n} \ast \overline{U}_{n}^{H} \ast \overline{U}_{n}) \ast \Sigma_{n} \ast \mathcal{V}_{n}^{H}\|^{2}
\]

\[
= \| \left( U_{n} - (U_{n}(\cdot, 1) \cdots U_{n}(\cdot, I_{n}^{p})) \ast \left( U_{n}(\cdot, 1) \cdots U_{n}(\cdot, I_{n}^{p})^{H} \right) \right) \ast \Sigma_{n} \ast \mathcal{V}_{n}^{H}\|^{2}
\]

\[
= \| \sum_{i_{n}=I_{n}^{p}+1}^{R_{n}} U_{n}(\cdot, i_{n}) \ast \Sigma_{n}(\cdot, i_{n}) \ast \mathcal{V}_{n}(\cdot, i_{n})^{H}\|^{2}
\]

\[
\leq \|A \ast_{n} (I - \overline{U}_{n} \ast \overline{U}_{n}^{H})\|.
\]
Proposition 5.1\[\begin{align*}
&= \sum_{i_n = I_n^t + 1}^{R_n} \| \mathcal{U}_n(\cdot, i_n) \ast \Sigma_n(\cdot, i_n) \ast \mathcal{V}_n(\cdot, i_n)^H \|^2 \\
&= \sum_{i_n = I_n^t + 1}^{R_n} \| \Sigma_n(\cdot, i_n) \|^2 \\
&= \sum_{i_n = I_n^t + 1}^{R_n} (\sigma_{i_n}^{(n)})^2, \tag{5.7}
\end{align*}\]
where the second-to-last equality comes from the unitary invariance of the Frobenius norm (Theorem 2.4).

Set \( \tilde{A}_{(1)} = \mathcal{U}^* \ast \mathcal{S}^* \ast \mathcal{V}^* \), where \((\mathcal{S}^*, \mathcal{U}^*, \mathcal{V}^*)\) is the optimal solution to the following low-rank approximation problem:

\[
\min_{\mathcal{S}, \mathcal{U}, \mathcal{V}} \left\| \mathcal{A}_{(1)} - \mathcal{U} \ast \mathcal{S} \ast \mathcal{V}^T \right\|.
\]

s.t. \( \mathcal{U} \in \mathbb{C}^{I_1 \times I_1'}, \mathcal{S} \in \mathbb{C}^{I_p \times I_p'}, \mathcal{V} \in \mathbb{C}^{(I_2 \times \cdots \times I_N) \times I_1'}, \mathcal{U}^H \ast \mathcal{U} = I_1', \mathcal{V}^H \ast \mathcal{V} = I_1'. \)

Then, the equality \( \sum_{i_1 = I_1^t + 1}^{I_1} (\sigma_{i_1}^{(1)})^2 = \| \mathcal{A}_{(1)} - \tilde{A}_{(1)} \|^2 \) follows from the Eckart–Young type result for tubal matrices Theorem 2.2, since

\[
\tilde{A}_{(1)} = \sum_{i_1 = 1}^{I_1} \mathcal{U}_1(\cdot, i) \ast \Sigma_1(i, i) \ast \mathcal{V}_1(\cdot, i)^H
\]
is the optimal solution to the above problem.

Similarly, we have equalities

\[
\sum_{i_2 = I_2^t + 1}^{I_2} (\sigma_{i_2}^{(2)})^2 = \| \mathcal{A}_{(2)} - \tilde{A}_{(2)} \|^2, \ldots, \sum_{i_N = I_N^t + 1}^{I_N} (\sigma_{i_N}^{(N)})^2 = \| \mathcal{A}_{(N)} - \tilde{A}_{(N)} \|^2.
\]

Therefore, we have

\[
\| \mathcal{A} - \tilde{\mathcal{A}} \|^2 \\
\leq \sum_{i_1 = I_1^t + 1}^{I_1} (\sigma_{i_1}^{(1)})^2 + \cdots + \sum_{i_N = I_N^t + 1}^{I_N} (\sigma_{i_N}^{(N)})^2 \\
= \| \mathcal{A}_{(1)} - \tilde{\mathcal{A}}_{(1)} \|^2 + \cdots + \| \mathcal{A}_{(N)} - \tilde{\mathcal{A}}_{(N)} \|^2 \\
\leq \| \mathcal{A}_{(1)} - \mathcal{A}^*_{(1)} \|^2 + \cdots + \| \mathcal{A}_{(N)} - \mathcal{A}^*_{(N)} \|^2 \\
= N \| \mathcal{A} - \mathcal{A}^* \|^2,
\]
where the first inequality comes from Eqs. (5.5) to (5.7) and the second inequality is due to the Eckart–Young theorem for tubal matrices (Theorem 2.2). \(\square\)

Note that the truncated HOSVD of an \((N + 1)\)-th order tensor \(\mathcal{A}\) leads to an error bound \(\| \mathcal{A} - \tilde{\mathcal{A}} \| \leq \sqrt{N + 1} \| \mathcal{A} - \mathcal{A}^* \|\) (where \(\tilde{\mathcal{A}}\) is obtained by the truncated HOSVD and \(\mathcal{A}^*\)
is the solution to the usual (i.e., not based on tensor–tensor product) orthogonal low-rank approximation problem.

Now, we prove the error bound for sequentially truncated Hot-SVD. We remark that the sequentially truncated Hot-SVD can be run in the order \( p_1, \ldots, p_N \) defined by a permutation of \( 1, \ldots, N \). In practice, the processing order is important. For theoretical analysis of the error bound, we only need to consider the case \( 1, \ldots, N \).

**Theorem 5.2** (Error bound for sequentially truncated Hot-SVD) Suppose that \( A \in \mathbb{C}_p^{I_1 \times \cdots \times I_N} \) is a tubal tensor of order \( N \). Let \( A^{(n)} = U_n \ast \Sigma_n \ast V_n^H \) be a t-SVD of the mode-\( n \) unfolding of \( A \). Let \( R_n \) be the t-rank of \( A^{(n)} \), i.e., \( R_n \) is the number of non-zero tubal scalars on the diagonal of \( \Sigma_n \). Let \( \sigma_1^{(n)}, \ldots, \sigma_R^{(n)} \) be the Frobenius norms of the non-zero tubal scalars on the diagonal of \( \Sigma_n \). Let \( \hat{A} \) be a rank-\((I'_1, \ldots, I'_N)\) sequentially truncated Hot-SVD of \( A \) where \( I'_1 \leq R_1, \ldots, I'_N \leq R_N \). That is

\[
\hat{A} = \hat{S} \ast_1 \hat{U}_1 \ast_2 \hat{U}_2 \ast_3 \hat{U}_3 \ast \cdots \ast_N \hat{U}_N,
\]

where \( \hat{S}, \hat{U}_1, \ldots, \hat{U}_N \) are as constructed in Algorithm 2. In other words, \( \hat{U}_n \) is the tubal matrix formed by the first \( I'_n \) column tubal vectors of the left unitary tubal matrix of the t-SVD of \( A^{(n)} \)

\[
(A^\ast_1 \hat{U}_1^H \cdots \ast_{n-1} \hat{U}_{n-1}^H)^{(n)}.
\]

\[
\hat{S} = A^\ast_1 \hat{U}_1^H \cdots \ast_N \hat{U}_N^H.
\]

and

\[
\hat{A} = A^\ast_1 (\hat{U}_1^\ast \hat{U}_1^H) \cdots \ast_N \left( \hat{U}_N^\ast \hat{U}_N^H \right).
\]

Then, we have the following error bound:

\[
\| A - \hat{A} \| \leq \sqrt{N \| A - A^\ast \|},
\]

where \( A^\ast \) was defined in Theorem 5.1.

**Proof** By Proposition 5.3, we have

\[
\| A - \hat{A} \|^2 = \| A^\ast_1 (I - \hat{U}_1^\ast \hat{U}_1^H) \|^2 + \| A^\ast_1 \left( \hat{U}_1^\ast \hat{U}_1^H \right) \ast_2 (I - \hat{U}_2^\ast \hat{U}_2^H) \|^2 + \| A^\ast_1 \left( \hat{U}_1^\ast \hat{U}_1^H \right) \ast_2 \left( \hat{U}_2^\ast \hat{U}_2^H \right) \ast_3 (I - \hat{U}_3^\ast \hat{U}_3^H) \|^2 + \cdots
\]

\[
+ \| A^\ast_1 \left( \hat{U}_1^\ast \hat{U}_1^H \right) \cdots \ast_{N-1} \left( \hat{U}_{N-1}^\ast \hat{U}_{N-1}^H \right) \ast_N (I - \hat{U}_N^\ast \hat{U}_N^H) \|^2.
\]

(5.8)

Let \( \hat{U}_n \) be the tubal matrix formed by the first \( I'_n \) column tubal vectors of the left unitary tubal matrix of the t-SVD of \( A^{(n)} \). That is, \( \hat{U}_1, \ldots, \hat{U}_N \) are the unitary tubal matrices obtained...
from the truncated Hot-SVD of \( \mathcal{A} \). For the \( n \)-th term in the above expression, we have

\[
\| \mathcal{A} \ast_1 \left( \hat{U}_1 \ast \hat{U}_1^H \right) \cdots \ast_{n-1} \left( \hat{U}_{n-1} \ast \hat{U}_{n-1}^H \right) \ast_n \left( \mathcal{I} - \hat{U}_n \ast \hat{U}_n^H \right) \| ^2 \\
= \| \mathcal{A} \ast_1 \left( \hat{U}_1 \ast \hat{U}_1^H \right) \cdots \ast_{n-1} \left( \hat{U}_{n-1} \ast \hat{U}_{n-1}^H \right) \ast_n \left( \mathcal{I} - \hat{U}_n \ast \hat{U}_n^H \right) \| ^2 \\
\leq \| \mathcal{A} \ast_1 \left( \hat{U}_1 \ast \hat{U}_1^H \right) \cdots \ast_{n-1} \left( \hat{U}_{n-1} \ast \hat{U}_{n-1}^H \right) \ast_n \left( \mathcal{I} - \hat{U}_n \ast \hat{U}_n^H \right) \| ^2 \\
= \| \mathcal{A} \ast_1 \left( \hat{U}_1 \ast \hat{U}_1^H \right) \cdots \ast_{n-1} \left( \hat{U}_{n-1} \ast \hat{U}_{n-1}^H \right) \ast_n \left( \mathcal{I} - \hat{U}_n \ast \hat{U}_n^H \right) \| ^2 ,
\]

(5.9)

where the inequality holds due to the Eckart–Young theorem for t-SVD (since \( \hat{U}_n \), being obtained from the truncation of t-SVD gives the minimum) and the equalities hold due to the unitary invariance of the Frobenius norm under the tensor–tensor product (Theorem 2.4). By Proposition 5.2, we have

\[
\| \mathcal{A} \ast_1 \left( \hat{U}_1 \ast \hat{U}_1^H \right) \cdots \ast_{n-1} \left( \hat{U}_{n-1} \ast \hat{U}_{n-1}^H \right) \ast_n \left( \mathcal{I} - \hat{U}_n \ast \hat{U}_n^H \right) \| ^2 \\
\leq \| \mathcal{A} \ast_n \left( \mathcal{I} - \hat{U}_n \ast \hat{U}_n^H \right) \| ^2 .
\]

(5.10)

Note that by Eq. (5.7), we have

\[
\| \mathcal{A} \ast_n \left( \mathcal{I} - \hat{U}_n \ast \hat{U}_n^H \right) \| ^2 = \sum_{i_n = I_n'+1}^{K_n} (\sigma_{i_n}^{(n)})^2 .
\]

(5.11)

Combining Eqs. (5.8) to (5.11), we obtain the desired error bound. \( \square \)

Remark 5.1 The notion of the generalized HOSVD was previously introduced in Liao and Maybank (2020, Sect. 4.2) and Liao et al. (2022, Sect. V-C), in the context of generalized tensors (the notion of generalized tensors is a generalization of that of tubal tensors, and the generalized HOSVD was called THOSVD there). However, their focus is different from ours. Liao and Maybank (2020) and Liao et al. (2022) validated THOSVD via extensive numerical experiments, while no proof of the existence of THOSVD was given nor did the authors explore properties such as all-orthogonality and ordering of the core tubal tensor \( \mathcal{S} \). In fact, one of our main contributions is the rigorous proof of the validity of Hot-SVD, where the basic definitions and properties introduced and developed in Sect. 3 play important roles in the analysis. Moreover, we formally introduce the truncated Hot-SVD, sequentially truncated Hot-SVD, and establish their error bounds. In addition, several properties of HOSVD are also generalized to Hot-SVD. These were not studied in Liao et al. (2022) and Liao and Maybank (2020).

6 Computational complexity

In this section, assuming that \( L \) is the DFT, we analyze the computational complexity of Hot-SVD, tr-Hot-SVD, and seq-tr-Hot-SVD.

To obtain the Hot-SVD of \( \mathcal{A} \in \mathbb{C}^{I_1 \times \cdots \times I_N} \), we perform the following operations:

1. Apply FFT to \( \mathcal{A}_{(1)}, \ldots, \mathcal{A}_{(N)} \) to get \( L(\mathcal{A}_{(1)}), \ldots, L(\mathcal{A}_{(N)}) \). The complexity is \( O(I_1 \ldots I_N \log(p)) \).
2. Apply matrix SVD to the frontal slices of \( L(\mathcal{A}_{(1)}) \in \mathbb{C}^{I_1 \times (I_2 \ldots I_N)} \), \ldots, \( L(\mathcal{A}_{(N)}) \in \mathbb{C}^{I_N \times (I_1 \ldots I_{N-1})} \),
to get $L(U_1), \ldots, L(U_N)$. Then, apply IFFT to get $U_1, \ldots, U_N$.

3. Apply $n$-mode product to get the core tensor $S = A \ast_1 U_1^H \ast_n U_N^H$.

If we assume that $I_1 \leq I_2 \ldots I_N$, then step 2 requires

$$O(pI_1^2I_2 \ldots I_N + \cdots + pI_N^2N_1 \ldots I_{N-1}) \quad (6.1)$$

flops, as for each $U_n$ it computes $p$ SVD of size $I_n \times \prod_{j \neq n}^N I_j$ in the Fourier domain (if $A$ is real then only half of the SVDs need to be performed) and the operations concerning IFFT are dominated. Step 3 requires

$$O(pI_1^2I_2 \ldots I_N + \cdots + pI_N^2I_1 \ldots I_{N-1})$$

flops (Kilmer and Martin 2011, Fact 3).

For tr-Hot-SVD with $I_1' \leq I_1, \ldots, I_N' \leq I_N$, the only difference is that at step 3, the complexity is

$$O \left( pI_1'(I_1I_3 \ldots I_N) + pI_2'(I_1'I_3 \ldots I_N) + \cdots + pI_N'(I_1'I_3 \ldots I_{N-1}) \right). \quad (6.2)$$

Next, we consider seq-tr-Hot-SVD.

For seq-tr-Hot-SVD of $A$, we perform in step $n$ (for $n = 1, \ldots, N$) the following two operations.

First, we compute the t-SVD of

$$\left(A \ast_1 \hat{U}_1^H \ast_n \cdots \ast_{n-1} \hat{U}_{n-1}^H \right) \in \mathbb{C}_{p}^{I_1 \times (I_1'I_2 \ldots I_{n-1}'I_{n+1}I_{n+2} \ldots I_N)},$$

that get $U_n \in \mathbb{C}_{p}^{I_n \times I_n}$. Assuming that $I_n \leq I_1'I_2 \ldots I_{n-1}'I_{n+1}I_{n+2} \ldots I_N$, this step requires

$$O \left( pI_n^2(I_1'I_2 \ldots I_{n-1}'I_{n+1}I_{n+2} \ldots I_N) \right) \quad (6.3)$$

flops. The FFT and IFFT operations are dominated.

Next, we compute $\left(A \ast_1 \hat{U}_1^H \ast_n \cdots \ast_{n-1} \hat{U}_{n-1}^H \right) *U_n^H \ast_n U_n^H$, which can be obtained by

$$\hat{U}_n^H * \left( A \ast_1 \hat{U}_1^H \ast_n \cdots \ast_{n-1} \hat{U}_{n-1}^H \right) = \hat{U}_n^H * U_n^H * \Sigma_n * V_n^H = \hat{U}_n^H * \Sigma_n * V_n^H,$$

where $\Sigma_n \in \mathbb{C}_{p}^{I_n'I_1 \times I_n}$ and $V_n^H \in \mathbb{C}_{p}^{I_n'I_1 \times (I_1'I_2 \ldots I_{n-1}'I_{n+1}I_{n+2} \ldots I_N)}$ are obtained through truncation of the t-SVD of $\left( A \ast_1 \hat{U}_1^H \ast_n \cdots \ast_{n-1} \hat{U}_{n-1}^H \right)_{(p)}$. This step requires the computation of $I_n'(I_1'I_2 \ldots I_{n-1}'I_{n+1}I_{n+2} \ldots I_N)$ t-product of tubal scalars and needs

$$O \left( p \log(p)I_n'(I_1'I_2 \ldots I_{n-1}'I_{n+1}I_{n+2} \ldots I_N) \right)$$

flops.

In summary, a rank-$\left(I_1', \ldots, I_N'\right)$ seq-tr-Hot-SVD of $A \in \mathbb{C}^{I_1 \times I_2 \times \cdots \times I_N}$ requires

$$O \left( pI_1^2(I_1I_3 \ldots I_N) + pI_2^2(I_1'I_3 \ldots I_N) + \cdots + pI_N^2(I_1'I_3 \ldots I_{N-1}) \right.$$\hspace{2cm}$$\left. + p \log(p)I_1'(I_1I_3 \ldots I_N) + p \log(p)I_2'(I_1'\ldots I_N) + \cdots \right) \quad (6.4)$$

flops.

Comparing this with tr-Hot-SVD ((6.1) + (6.2)), we see that seq-tr-Hot-SVD requires fewer computations than tr-Hot-SVD, especially when $I_1', \ldots, I_N'$ are small.
7 Numerical examples

We only conduct preliminary experiments on Hot-SVD, tr-Hot-SVD, and seq-tr-Hot-SVD in this section, as the main focus of this paper is on the derivation and providing theoretical analysis of these models. All the examples are conducted on an Intel i7 CPU desktop computer with 32 GB of RAM. The supporting software is Matlab 2019b. We remark that our codes are modified from those in Tensorlab (Vervliet et al. 2016), namely, the codes are tubal versions of the corresponding ones in Vervliet et al. (2016). We set $L$ as DFT in this section.

We first illustrate the properties of Theorem 5.1 via a small example. Consider the tensor $A \in \mathbb{R}^{2 \times 2 \times 2 \times 2}$ given by $A(i, j, k, l) = (i + j + k + l - 3)^{-1}$ for each $(i, j, k, l)$. This is known as the Hilbert tensor (Song and Qi 2014). Applying Hot-SVD to $A$, we obtain $U_1 = U_2 = U_3 = U \in \mathbb{R}_2^{2 \times 2}$, with $U(:, 1) = \begin{bmatrix} -0.8924 & 0.4395 \\ -0.4395 & -0.8924 \end{bmatrix}$, $U(:, 2) = \begin{bmatrix} 0.0453 & 0.0920 \\ -0.0920 & 0.0453 \end{bmatrix}$, and $S \in \mathbb{R}_2^{2 \times 2 \times 2}$ with

$$S(1, 1, 1) = [-1.4734 - 0.8780]^T, \quad S(2, 2, 2) = [0.0102 0.0107]^T,$$
$$S(2, 1, 1) = S(1, 2, 1) = S(1, 1, 2) = [-0.0004 - 0.0004]^T,$$
$$S(2, 2, 1) = S(2, 1, 2) = S(1, 2, 2) = [-0.0612 - 0.0343]^T.$$

Then, noticing the symmetry of the tubal scalars and the fact that $a^T = a$ for a real tubal scalar $a \in \mathbb{R}_2$ of length 2, we have

$$\sum_{i_2, i_3=1}^2 S(1, i_2, i_3) \ast S(2, i_2, i_3)$$
$$= S(1, 1, 2) \ast (S(1, 1, 1) + 2S(1, 2, 2)) + S(1, 2, 2) \ast S(2, 2, 2) = [0 0]^T,$$

$$\sum_{i_1, i_3=1}^2 S(i_1, 1, i_3) \ast S(i_1, 2, i_3)$$
$$= \sum_{i_1, i_2=1}^2 S(i_1, i_2, 1) \ast S(i_1, i_2, 2) = \sum_{i_2, i_3=1}^2 S(1, i_2, i_3) \ast S(2, i_2, i_3) = [0 0]^T,$$

illustrating the all-orthogonality of $S$.

We then show the ordering of $S$. Note that

$$S(1) = \begin{bmatrix} S(1,1,1) & S(1,1,2) & S(1,2,1) & S(1,2,2) \\ S(2,1,1) & S(2,1,2) & S(2,2,1) & S(2,2,2) \end{bmatrix} \in \mathbb{R}_2^{2 \times 4},$$
$$S(2) = \begin{bmatrix} S(1,1,1) & S(1,1,2) & S(1,2,1) & S(1,2,2) \\ S(2,1,1) & S(2,1,2) & S(2,2,1) & S(2,2,2) \end{bmatrix} \in \mathbb{R}_2^{2 \times 4},$$
$$S(3) = \begin{bmatrix} S(1,1,1) & S(1,1,2) & S(1,2,1) & S(1,2,2) \\ S(2,1,1) & S(2,1,2) & S(2,2,1) & S(2,2,2) \end{bmatrix} \in \mathbb{R}_2^{2 \times 4}.$$

Then

$$(\|S_{1,2,1}\|, \|S_{2,1,2}\|) = (\|S_{1,2,1}\|, \|S_{3,2,1}\|) = (\|S_{3,2,1}\|, \|S_{1,2,2}\|) = (1.7166, 0.1002).$$

confirming the ordering property.

Next, we generate $A$ as

$$A = A^\circ + \beta \varepsilon / \|\varepsilon\|,$$

where $A^\circ = \sum_{i=1}^R a_{1,i} \cdots a_{N,i} \in \mathbb{R}_L^{I_1 \times R}, \ldots, [a_{1,1} \ldots a_{N,1}] \in \mathbb{R}_L^{I_N \times R}$ are randomly generated matrices obeying standard Gaussian distribution, and $\varepsilon$ is also a randomly
Table 1 Comparisons of tr-Hot-SVD and seq-tr-Hot-SVD on recovering randomly generated tensors

| $[I_1 \cdots I_N]$ | tr-Hot-SVD | seq-tr-Hot-SVD |
|-------------------|------------|---------------|
| Err               | Time       | Err           | Time       |
| [10 10 10 10]     | 0.04616    | 0.010         | 0.04595    | 0.005       |
| [15 15 15 10]     | 0.02913    | 0.015         | 0.02911    | 0.009       |
| [20 20 20 10]     | 0.02131    | 0.022         | 0.02121    | 0.013       |
| [25 25 25 10]     | 0.01671    | 0.036         | 0.01668    | 0.020       |
| [30 30 30 10]     | 0.01371    | 0.041         | 0.01370    | 0.030       |
| [35 35 35 10]     | 0.01165    | 0.048         | 0.01161    | 0.036       |
| [40 40 40 10]     | 0.01023    | 0.054         | 0.01013    | 0.039       |
| [10 10 10 10 10]  | 0.01291    | 0.042         | 0.01284    | 0.040       |
| [15 15 15 15 10]  | 0.00773    | 0.096         | 0.00768    | 0.066       |
| [20 20 20 20 10]  | 0.00518    | 0.212         | 0.00516    | 0.112       |
| [25 25 25 25 10]  | 0.00374    | 0.441         | 0.00373    | 0.205       |
| [30 30 30 30 10]  | 0.00289    | 0.771         | 0.00287    | 0.334       |
| [35 35 35 35 10]  | 0.00230    | 1.328         | 0.00229    | 0.548       |
| [40 40 40 40 10]  | 0.01812    | 0.084         | 0.01809    | 0.066       |
| [10 10 10 10 10 10] | 0.01182 | 0.542         | 0.01166    | 0.277       |
| [20 20 20 20 20 10] | 0.00332 | 2.203         | 0.00331    | 0.880       |
| [25 25 25 25 25 10] | 0.00194 | 7.134         | 0.00193    | 2.585       |
| [30 30 30 30 30 10] | 0.00126 | 15.889        | 0.00125    | 5.629       |

$I_i^1 = \cdots = I_{N-1}^N = R$

Boldface means that the error is the lowest among all the compared methods.

generated tensor obeying standard Gaussian distribution. We set $R = 5$ and $\beta = 0.1$ in our experiment. We apply tr-Hot-SVD and seq-tr-Hot-SVD on recovering $\mathcal{A}^2$. The size of the core tubal tensor $S$ is $I_1^1 = \cdots = I_{N-1}^N = R$, and the tubal length $p = I_N$. We evaluate $err = \|\mathcal{A}^2 - \tilde{\mathcal{A}}\|/\|\mathcal{A}^2\|$ and the CPU time, where $\tilde{\mathcal{A}} = S \ast_1 U_1 \cdots \ast_{N-1} U_{N-1}$ is generated by tr-Hot-SVD or seq-tr-Hot-SVD. The results are presented in Table 1, averaged over 50 instances for each case.

From the table, we observe that both algorithms can recover the true tensors well, and seq-tr-Hot-SVD is slightly better in terms of the recovery error. Considering the CPU time, we see that seq-tr-Hot-SVD usually performs 2–3 times faster than tr-Hot-SVD, which confirms the computational complexity analysis in Sect. 6. In particular, when the size of the tensor becomes larger and larger (relative to the truncation size $I_1^1, \ldots, I_{N-1}^N$), the advantage of seq-tr-Hot-SVD turns out to be more evident.

We then test tr-Hot-SVD and seq-tr-Hot-SVD on color video compression. The tested video “airport” was downloaded from [http://perception.i2r.a-star.edu.sg/bk_model/bk_index.html](http://perception.i2r.a-star.edu.sg/bk_model/bk_index.html). The original video consists of 4583 frames, each of size $144 \times 176$. We use 500 frames, resulting into a tensor of size $500 \times 144 \times 176 \times 3$. We treat it as a tubal tensor in $\mathbb{R}^{500\times144\times176}$.

We set different truncation sizes for the two algorithms to compress the data. We use tr-HOSVD (De Lathauwer et al. 2000) and seq-tr-HOSVD (Vannieuwenhoven et al. 2012) as baselines. The truncation sizes of tr-HOSVD and seq-tr-HOSVD are the same as the tubal counterparts, except that we do not truncate the fourth mode of both tr-HOSVD and seq-tr-HOSVD. We evaluate the reconstruction error $err = \|\mathcal{A}^2 - \tilde{\mathcal{A}}\|/\|\mathcal{A}^2\|$ and the CPU time.
Table 2  Comparisons of tr-Hot-SVD, seq-tr-Hot-SVD, tr-HOSVD, and seq-tr-HOSVD on color video compression

| $[t_1', t_2', t_3']$ | tr-Hot-SVD | seq-tr-Hot-SVD | tr-HOSVD | seq-tr-HOSVD |
|----------------------|-------------|----------------|----------|--------------|
|                      | Err  | Time | Err  | Time |Err  | Time | Err  | Time |
| [200 50 50]          | 0.0769 | 14.50 | **0.0762** | 7.37 | 0.0782 | 5.33 | 0.0776 | 3.20 |
| [100 50 50]          | 0.0841 | 13.85 | **0.0835** | 6.35 | 0.0854 | 5.59 | 0.0848 | 2.66 |
| [50 50 50]           | 0.0943 | 13.19 | **0.0937** | 9.98 | 0.0955 | 6.07 | 0.0950 | 2.95 |
| [30 30 30]           | 0.1171 | 17.75 | **0.1157** | 7.94 | 0.1187 | 5.68 | 0.1174 | 2.19 |
| [20 10 10]           | 0.1595 | 12.48 | **0.1574** | 3.39 | 0.1621 | 6.01 | 0.1601 | 1.76 |
| [10 5 5]             | 0.1841 | 13.65 | **0.1824** | 3.27 | 0.1881 | 6.03 | 0.1863 | 1.46 |

Results marked with underline mean that the errors are the second lowest among all the compared methods. Boldface means that the error is the lowest among all the compared methods.

Fig. 2  The first row: the original frames. The second-to-last rows: frames reconstructed by seq-tr-Hot-SVD with different truncation sizes.

time, where $\mathcal{A}^2$ is the data tensor and $\hat{\mathcal{A}}$ is reconstructed by the algorithms. The results are presented in Table 2. Some reconstructed frames by seq-tr-Hot-SVD are illustrated in Fig. 2.

From the table, we observe that concerning the reconstruction error, seq-tr-Hot-SVD is the best among the four algorithms, followed by tr-Hot-SVD. This shows the advantage of the tubal versions of HOSVD. Concerning the CPU time, seq-tr-HOSVD is the fastest one, while seq-tr-Hot-SVD and tr-Hot-SVD are slower than tr-HOSVD and seq-tr-HOSVD. This
is because the tubal versions need to perform additional Fourier transforms. We also see that seq-tr-Hot-SVD is about 2–4 times faster than tr-Hot-SVD.

8 Conclusions

In this paper, we studied the analogue of HOSVD in the setup of tubal tensors, which we call Hot-SVD. This is an alternative to the t-SVD for tensors of order higher than three (Martin et al. 2013). To prove the validity of Hot-SVD, we used the small-t transpose for third-order tensors and established some basic properties. Hot-SVD enjoys most of the properties of the HOSVD and manifests the efficacy of the language of tubal matrices. We also established an error bound $\sqrt{N}$ for the truncated Hot-SVD and sequentially truncated Hot-SVD of tensors of order $N + 1$. We remark that our purpose of studying Hot-SVD is not to compare it with t-SVD for tensors of order higher than three (Martin et al. 2013), but to contribute another kind of decomposition of higher order tensors in the tensor–tensors product setting.

In the future, randomized algorithms can be investigated for Hot-SVD.

Acknowledgements This work was supported by National Natural Science Foundation of China under Grant No. 12171105, Fok Ying Tong Education Foundation under Grant No. 171094, Guangxi Science and Technology base and Talent Project under Grant No. AD22080047, and the special foundation for Guangxi Ba Gui Scholars.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

Braman K (2010) Third-order tensors as linear operators on a space of matrices. Linear Algebra Appl 433(7):1241–1253
Cichocki A, Mandic D, De Lathauwer L, Zhou G, Zhao Q, Caiafa C, Phan HA (2015) Tensor decompositions for signal processing applications: from two-way to multiway component analysis. IEEE Signal Process Mag 32(2):145–163
Comon P (2014) Tensors: a brief introduction. IEEE Signal Process Mag 31(3):44–53
De Lathauwer L, De Moor B, Vandewalle J (2000) A multilinear singular value decomposition. SIAM J Matrix Anal Appl 21:1253–1278
Kernfeld E, Kilmer M, Aeron S (2015) Tensor-tensor products with invertible linear transforms. Linear Algebra Appl 485:545–570
Kilmer ME, Braman K, Hao N, Hoover RC (2013) Third-order tensors as operators on matrices: a theoretical and computational framework with applications in imaging. SIAM J Matrix Anal Appl 34(1):148–172
Kilmer ME, Horesh L, Avron H, Newman E (2021) Tensor-tensor algebra for optimal representation and compression of multiway data. PNAS USA 118(28):e2015851118
Kilmer ME, Martin CD (2011) Factorization strategies for third-order tensors. Linear Algebra Appl 435(3):641–658
Kolda TG, Bader BW (2009) Tensor decompositions and applications. SIAM Rev 51:455–500
Kong H, Xie X, Lin Z (2018) t-schatten-$p$ norm for low-rank tensor recovery. IEEE J Sel Top Signal Process 12(6):1405–1419
Liao L, Lin S, Li L, Zhang X, Zhao S, Wang Y, Wang X, Gao Q, Wang J (2022) Approximation of images via generalized higher order singular value decomposition over finite-dimensional commutative semisimple algebra. arXiv:2202.00040 (arXiv preprint)
Liao L, Maybank SJ (2020) Generalized visual information analysis via tensorial algebra. J Math Imaging Vis 62(4):560–584
Ling C, Liu J, Ouyang C, Qi L (2021) ST-SVD factorization and s-diagonal tensors. arXiv:2104.05329 (arXiv preprint)
Lund K (2020) The tensor t-function: a definition for functions of third-order tensors. Numer Linear Algebra Appl 27(3):e2288
Martin CD, Shafer R, Larue B (2013) An order-p tensor factorization with applications in imaging. SIAM J Sci Comput 35(1):A474–A490
Miao Y, Qi L, Wei Y (2020) Generalized tensor function via the tensor singular value decomposition based on the T-product. Linear Algebra Appl 590:258–303
Miao Y, Qi L, Wei Y (2021) T-Jordan canonical form and t-Drazin inverse based on the t-product. Commun Appl Math Comput 3(2):201–220
Newman E, Horesh L, Avron H, Kilmer M (2018) Stable tensor neural networks for rapid deep learning. arXiv:1811.06569 (arXiv preprint)
Oseledets IV (2011) Tensor-train decomposition. SIAM J Sci Comput 33(5):2295–2317
Qi L, Luo Z (2021) Tubal matrix. arXiv:2105.00793 (arXiv preprint)
Qi L, Yu G (2021) T-singular values and T-sketching for third order tensors. arXiv:2103.00976 (arXiv preprint)
Sidiropoulos ND, De Lathauwer L, Fu X, Huang K, Papalexakis EE, Faloutsos C (2017) Tensor decomposition for signal processing and machine learning. IEEE Trans Signal Process 65(13):3551–3582
Song Y, Qi L (2014) Infinite and finite dimensional hilbert tensors. Linear Algebra Appl 451:1–14
Vannieuwenhoven N, Vandebriel R, Meerbergen K (2012) A new truncation strategy for the higher-order singular value decomposition. SIAM J Sci Comput 34(2):A1027–A1052
Vervliet N, Debals O, De Lathauwer L (2016) Tensorlab 3.0-numerical optimization strategies for large-scale constrained and coupled matrix/tensor factorization. In: 2016 50th Asilomar conference on signals, systems and computers, pp 1733–1738. IEEE
Yin M, Gao J, Xie S, Guo Y (2019) Multiview subspace clustering via tensorial t-product representation. IEEE Trans Neural Netw Learn Syst 30(3):851–864
Zhang Z, Aeron S (2017) Exact tensor completion using t-svd. IEEE Trans Signal Process 65(6):1511–1526
Zhang Z, Ely G, Aeron S, Hao N, Kilmer M (2014) Novel methods for multilinear data completion and de-noising based on tensor-svd. In: 2014 IEEE conference on computer vision and pattern recognition, pp 3842–3849
Zheng MM, Huang ZH, Wang Y (2021) T-positive semidefiniteness of third-order symmetric tensors and T-semidefinite programming. Comput Optim Appl 78(1):239–272
Zhu Y, Wei Y (2022) Tensor LU and QR decompositions and their randomized algorithms. Comput Math Comput Model Appl (CMCMA) 1(1):1–16. https://doi.org/10.52547/CMCMA.1.1.1

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.