Received: 16 May 2016 | Revised: 29 January 2017 | Accepted: 3 February 2017

DOI: 10.1002/nla.2097

WILEY

RESEARCH ARTICLE

A constructive arbitrary-degree Kronecker product decomposition of tensors

Kim Batselier | **Ngai Wong**

Department of Electrical and Electronic Engineering, The University of Hong Kong, Pokfulam Road Hong Kong, Hong Kong

Correspondence
Kim Batselier, Department of Electrical and Electronic Engineering, The University of Hong Kong, Pokfulam Road, Hong Kong, Hong Kong.
Email: kim.batselier@gmail.com

Funding information
Hong Kong Research Grants Council, Grant/Award Number: 17212315

Summary

We generalize the matrix Kronecker product to tensors and propose the tensor Kronecker product singular value decomposition that decomposes a real k-way tensor $\mathcal{A}$ into a linear combination of tensor Kronecker products with an arbitrary number of d factors. We show how to construct $\mathcal{A}=\sum_{j=1}^{R}\sigma_j\,\mathcal{A}_j^{(d)}\otimes\cdots\otimes\mathcal{A}_j^{(1)}$, where each factor $\mathcal{A}_j^{(i)}$ is also a k-way tensor, thus including matrices ($k=2$) as a special case. This problem is readily solved by reshaping and permuting $\mathcal{A}$ into a d-way tensor, followed by a orthogonal polyadic decomposition. Moreover, we introduce the new notion of general symmetric tensors (encompassing symmetric, persymmetric, centrosymmetric, Toeplitz and Hankel tensors, etc.) and prove that when $\mathcal{A}$ is structured then its factors $\mathcal{A}_j^{(1)},\ldots,\mathcal{A}_j^{(d)}$ will also inherit this structure.

KEYWORDS

generalized symmetric tensors, Hankel tensor, Kronecker product, structured tensors, tensor decomposition, Toeplitz tensor

1 | INTRODUCTION

Consider the following 16×16 centrosymmetric matrix[1]

$$A=\begin{pmatrix}
1&17&33&49&65&81&97&113&128&112&96&80&64&48&32&16\\
2&18&34&50&66&82&98&114&127&111&95&79&63&47&31&15\\
3&19&35&51&67&83&99&115&126&110&94&78&62&46&30&14\\
4&20&36&52&68&84&100&116&125&109&93&77&61&45&29&13\\
5&21&37&53&69&85&101&117&124&108&92&76&60&44&28&12\\
6&22&38&54&70&86&102&118&123&107&91&75&59&43&27&11\\
7&23&39&55&71&87&103&119&122&106&90&74&58&42&26&10\\
8&24&40&56&72&88&104&120&121&105&89&73&57&41&25&9\\
9&25&41&57&73&89&105&121&120&104&88&72&56&40&24&8\\
10&26&42&58&74&90&106&122&119&103&87&71&55&39&23&7\\
11&27&43&59&75&91&107&123&118&102&86&70&54&38&22&6\\
12&28&44&60&76&92&108&124&117&101&85&69&53&37&21&5\\
13&29&45&61&77&93&109&125&116&100&84&68&52&36&20&4\\
14&30&46&62&78&94&110&126&115&99&83&67&51&35&19&3\\
15&31&47&63&79&95&111&127&114&98&82&66&50&34&18&2\\
16&32&48&64&80&96&112&128&113&97&81&65&49&33&17&1
\end{pmatrix}.$$

The Kronecker product singular value decomposition (KPSVD)[2] decomposes A into the following sum of Kronecker products (KPs)

$$1154.2\begin{pmatrix}0.09&0.31&0.35&0.13\\0.10&0.33&0.34&0.12\\0.12&0.34&0.33&0.10\\0.13&0.35&0.31&0.09\end{pmatrix}\otimes\begin{pmatrix}0.25&0.25&0.25&0.25\\0.25&0.25&0.25&0.25\\0.25&0.25&0.25&0.25\\0.25&0.25&0.25&0.25\end{pmatrix}$$

$$+117.98\begin{pmatrix} 0.25 & 0.25 & -0.25 & -0.25 \\ 0.25 & 0.25 & -0.25 & -0.25 \\ 0.25 & 0.25 & -0.25 & -0.25 \\ 0.25 & 0.25 & -0.25 & -0.25 \end{pmatrix} \otimes \begin{pmatrix} -0.35 & -0.13 & 0.09 & 0.31 \\ -0.34 & -0.12 & 0.10 & 0.33 \\ -0.33 & -0.10 & 0.12 & 0.34 \\ -0.31 & -0.09 & 0.13 & 0.35 \end{pmatrix}.$$

The first and second terms consist of a KP of centrosymmetric and skew-centrosymmetric matrices, respectively. The extension of this kind of decomposition beyond three factors has been considered a difficult computational and theoretical problem. Indeed, in Hackbusch et al.,[3] the authors state that "… but for $m > 2$ the construction of the Kronecker tensor-product approximations becomes a much more difficult problem that requires quite intricate algorithms … and still needs an adequate theory … we avoid the above-mentioned theoretical and algorithmical difficulties of $m > 2$ by focusing on the case $m = 2$," where m denotes the number of KP factors.

In this article, we derive a constructive algorithm that computes an exact decomposition of any k-way tensor into a linear combination of d KPs. When applying our algorithm to the A matrix above, we decompose it into the following linear combination

$$\begin{aligned} &1033.98\begin{pmatrix} 0.47 & 0.53 \\ 0.53 & 0.47 \end{pmatrix} \otimes \begin{pmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{pmatrix} \otimes \begin{pmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{pmatrix} \otimes \begin{pmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{pmatrix} \\ &+513.00\begin{pmatrix} 0.5 & -0.5 \\ 0.5 & -0.5 \end{pmatrix} \otimes \begin{pmatrix} -0.53 & 0.47 \\ -0.47 & 0.53 \end{pmatrix} \otimes \begin{pmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{pmatrix} \otimes \begin{pmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{pmatrix} \\ &+256.50\begin{pmatrix} 0.5 & -0.5 \\ 0.5 & -0.5 \end{pmatrix} \otimes \begin{pmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{pmatrix} \otimes \begin{pmatrix} -0.53 & 0.47 \\ -0.47 & 0.53 \end{pmatrix} \otimes \begin{pmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{pmatrix} \\ &+128.25\begin{pmatrix} 0.5 & -0.5 \\ 0.5 & -0.5 \end{pmatrix} \otimes \begin{pmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{pmatrix} \otimes \begin{pmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{pmatrix} \otimes \begin{pmatrix} -0.53 & 0.47 \\ -0.47 & 0.53 \end{pmatrix}, \end{aligned}$$

where each term consists of four Kronecker factors. Again, the KP factors are either centrosymmetric or skew-centrosymmetric matrices. We show in this article that such a decomposition into an arbitrary number of d Kronecker factors can even be done for tensors of any order k by introducing the tensor-based KP singular value decomposition (TKPSVD). The TKPSVD decomposes an arbitrary real k-way tensor $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_k}$ into a finite linear combination

$$\mathcal{A} = \sum_{j=1}^{R} \sigma_j \, \mathcal{A}_j^{(d)} \otimes \cdots \otimes \mathcal{A}_j^{(1)}, \tag{1}$$

where $\otimes$ denotes the tensor KP, defined in Section 3, and the tensors $\mathcal{A}_j^{(i)} \in \mathbb{R}^{n_1^{(i)} \times \cdots \times n_k^{(i)}}$ satisfy

$$||\mathcal{A}_j^{(i)}||_F = 1 \quad \text{with} \quad \prod_{i=1}^{d} n_r^{(i)} = n_r \quad (r \in \{1, \dots, k\}). \tag{2}$$

The Kronecker rank,[3] which is generally hard to compute, is defined as the minimal R required in Equation 1 in order for the equality to hold. If for any $r \in \{1, \dots, k\}$ we have that $n \triangleq n_r^{(1)} = n_r^{(2)} = \cdots = n_r^{(d)}$, then $n_r = n^d$. For this reason, we call the number of factors d in Equation 1 the degree of the decomposition. The user of the TKPSVD algorithm is completely free to choose the degree d and the dimensions $n_r^{(i)}$ of each of the KP factors, as long as they satisfy Equation 2. In addition to the development of the TKPSVD algorithm, another major contribution of this article, shown for the first time in the literature, is the proof that by our proposed algorithm, we will have the following favorable structure-preserving properties when all $\mathcal{A}_j^{(i)}$ and $\mathcal{A}$ are cubical:

$$\text{if } \mathcal{A} \text{ is} \left\{ \begin{array}{l} \text{symmetric} \\ \text{persymmetric} \\ \text{centrosymmetric} \\ \text{Toeplitz} \\ \text{Hankel} \\ \text{general symmetric} \end{array} \right\} \text{then each } \mathcal{A}_j^{(i)} \text{ is} \left\{ \begin{array}{l} \text{(skew-)symmetric} \\ \text{(skew-)persymmetric} \\ \text{(skew-)centrosymmetric} \\ \text{Toeplitz} \\ \text{Hankel} \\ \text{general (skew-)symmetric} \end{array} \right\}.$$

The fact that each of the factors $\mathcal{A}_j^{(i)}$ inherits the structure of $\mathcal{A}$ is not trivial. In providing this proof, a very natural generalization of symmetric tensors is introduced, which we name *general symmetric tensors*. In addition, Toeplitz and Hankel tensors are also generalized into what we call *shifted-index structures*, which are special cases of general symmetries. In fact, when $\mathcal{A}$ is general symmetric, then all its cubical factors $\mathcal{A}_j^{(i)}$ will also be general symmetric. Another interesting feature of the TKPSVD algorithm is that if the summation in Equation 1 is limited to the first r terms, then the relative approximation error in the Frobenius norm is given by

$$\frac{||A - \sum_{j=1}^{r} \sigma_j\, A_j^{(d)} \otimes \cdots \otimes A_j^{(1)}||_F}{||A||_F} = \frac{\sqrt{\sigma_{r+1}^2 + \cdots + \sigma_R^2}}{\sqrt{\sigma_1^2 + \cdots + \sigma_R^2}}. \tag{3}$$

Equation 3 has the computational advantage that the relative approximation error can be easily obtained from the σ_j's without having to explicitly construct the approximant.

1.1 | Background

The TKPSVD is directly inspired by the work of Van Loan and Pitsianis.[4] In their paper, they solve the problem of finding matrices B, C such that $||A - B \otimes C||_F$ is minimized. The globally minimizing matrices B, C turn out to be the singular vectors corresponding with the largest singular value of a particular permutation of A. In Van Loan,[2] the full singular value decomposition (SVD) of the permuted A is considered and the corresponding decomposition of A into a linear combination of KPs is called the KPSVD. Applications of the $d = 2, k = 2$ KPSVD approximation in image restoration are described in Kamm and Nagy[5] and Nagy et al.,[6] whereas extensions to the $d = 3, k = 2$ case using the higher order singular value decomposition (HOSVD), also for imaging, are described in Nagy and Kilmer[7] and Rezghi et al.[8] In Hackbusch et al.,[3] the decomposition (Equation 1) for $k = d = 2$ is studied for the approximation of certain classes of functions and nonlocal operators.

One major contribution of this article is to show how a decomposition(Equation 1) that consists of an arbitrary number of d factors can be computed for any k-way tensor. The decomposition in this paper is hence a direct generalization of the KPSVD to an arbitrary number of KP factors d and to arbitrary k-way tensors. In the TKPSVD case, the SVD is replaced by either a canonical polyadic decomposition (CPD),[9,10] with orthogonal factor matrices, the HOSVD[11] or the tensor-train rank-1 SVD (TTr1SVD).[12] The TKPSVD reduces to the KPSVD for the cases $d = 2$ and $k = 2$. Contrary to previous work in the literature, we are not interested in minimizing $||\mathcal{A} - \sum_{j=1}^{r} \mathcal{A}_j^{(d)} \otimes \cdots \otimes \mathcal{A}_j^{(1)}||_F$. Instead, we derive an exact decomposition such that any structure of $\mathcal{A}$ is also present in the KP factors. Explicit decomposition algorithms for when $d \geq 4$ and $k \geq 3$ are not found in the literature. To this end, our proposed TKPSVD algorithm readily works for any degree d and any tensor order k, does not require any a priori knowledge of the number of KP terms, and preserves general symmetry in the KP factors $\mathcal{A}_j^{(i)}$. Furthermore, a Matlab/Octave implementation that uses the TTr1SVD and works for any arbitrary degree d and order k can be freely downloaded from https://github.com/kbatseli/TKPSVD. In brief, the contributions of this article are the following:

- an explicit formulation of the generalization of the KPSVD algorithm to $d > 2$ and $k > 2$ is presented for the first time in the literature,
- a new notion of general symmetric tensors, which describes many tensor structures, is developed, and
- we prove that for a general symmetric tensor $\mathcal{A}$, all KP factors $\mathcal{A}_j^{(i)}$ in Equation 1 are guaranteed to have the same general symmetry under a mild condition.

The outline of this article is as follows. In Section 2, we introduce some basic tensor concepts and notations. In Section 3, we generalize the matrix KP to the tensor KP and present some of its properties. In Section 4, we first derive the theorem that underlies the TKPSVD and then present the TKPSVD algorithm for both general and diagonal tensors. In Section 5, we introduce the framework of *general symmetry* that describes many different structured tensors such as symmetry, centrosymmetry, Hankel, Toeplitz, etc. Preservation of general symmetry in the KP factors when the original tensor is general symmetric is proven in Section 6. Numerical experiments that demonstrate different aspects of the TKPSVD are presented in Section 7, after which, conclusions follow.

2 | TENSOR BASICS AND NOTATION

We only consider real matrices and tensors in this paper. Scalars are denoted by greek letters ($\alpha, \beta, \ldots$), vectors by lowercase letters ($a, b, \ldots$), matrices by uppercase letters ($A, B, \ldots$), and higher-order tensors by uppercase caligraphic letters ($\mathcal{A}, \mathcal{B}, \ldots$). The notation $(\cdot)^T$ denotes the transpose of either a vector or a matrix. A kth-order or k-way tensor is a multiway array $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_k}$. A tensor is cubical if all dimensions are equal, for example, $n_1 = n_2 = \cdots = n_k$.

Entries of tensors are always denoted with square brackets around the indices. This enables an easy way of representing the grouping of indices. Suppose for example that $\mathcal{A}$ is a four-way tensor with entries $\mathcal{A}_{[i_1][i_2][i_3][i_4]}$. To improve readability, we do not write the square brackets when all indices are considered separate, therefore $\mathcal{A}_{i_1 i_2 i_3 i_4} \triangleq \mathcal{A}_{[i_1][i_2][i_3][i_4]}$. A three-way tensor can now be formed by grouping, for example, the first two indices together. The entries of this three-way tensor are then denoted by $\mathcal{A}_{[i_1 i_2][i_3][i_4]}$, where the grouped index $[i_1 i_2]$ is easily converted into the linear index $i_1 + n_1(i_2 - 1)$. Grouping the indices into the $[i_1]$ and $[i_2 i_3 i_4]$ results in a $n_1 \times n_2 n_3 n_4$ matrix with entries $\mathcal{A}_{[i_1][i_2 i_3 i_4]}$. The column index $[i_2 i_3 i_4]$ is equivalent to the linear index $i_2 + n_2(i_3 - 1) + n_2n_3(i_4 - 1)$. A very special case of grouping indices is obtained when all indices are grouped together. The resulting vector is then called the vectorization of $\mathcal{A}$, denoted vec($\mathcal{A}$), with entries $\mathcal{A}_{[i_1i_2i_3i_4]}$.

The r-mode product of a tensor $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_k}$ with a matrix $U \in \mathbb{R}^{p \times n_r}$ is defined by

$$(\mathcal{A} \times_r U)_{i_1 \cdots i_{k-1} j i_{k+1} \cdots i_d} = \sum_{i_r=1}^{n_r} U_{j i_r} \mathcal{A}_{i_1 \cdots i_r \cdots i_d}$$

so that $\mathcal{A} \times_r U \in \mathbb{R}^{n_1 \times \cdots \times n_{r-1} \times p \times n_{r+1} \times \cdots \times n_d}$. The multiplication of a k-way tensor $\mathcal{A}$ along all its modes with matrices $P_1, \ldots, P_k$

$$\mathcal{B} = \mathcal{A} \times_1 P_1 \times_2 P_2 \times_3 \cdots \times_k P_k$$

can be rewritten as the following linear system:

$$\text{vec}(\mathcal{B}) = (P_d \otimes \cdots \otimes P_2 \otimes P_1) \ \text{vec}(\mathcal{A}), \tag{4}$$

where $\otimes$ is the conventional matrix KP,[13] which is defined and generalized to the tensor case in Section 3. The inner product between two tensors $\mathcal{A}, \mathcal{B} \in \mathbb{R}^{n_1 \times \cdots \times n_d}$ is defined as

$$\langle \mathcal{A}, \mathcal{B} \rangle = \sum_{i_1, i_2, \cdots, i_d} \mathcal{A}_{i_1 i_2 \cdots i_d} \ \mathcal{B}_{i_1 i_2 \cdots i_d} = \text{vec}(\mathcal{A})^T \ \text{vec}(\mathcal{B}).$$

Two tensors $\mathcal{A}, \mathcal{B}$ are orthogonal with respect to one another when $\langle \mathcal{A}, \mathcal{B} \rangle = 0$. The norm of a tensor is taken to be the Frobenius norm $||\mathcal{A}||_F = \langle \mathcal{A}, \mathcal{A} \rangle^{1/2}$. A k-way rank-1 ten sor $\mathcal{A} \in \mathbb{R}^{n_1 \times \cdots \times n_k}$ is per definition the outer product,[14] denoted by $\circ$, of k vectors $a^{(i)} \in \mathbb{R}^{n_i}$ $(i \in \{1, \ldots, k\})$ such that

$$\mathcal{A} = a^{(1)} \circ a^{(2)} \circ \cdots \circ a^{(k)} \quad \text{with} \quad \mathcal{A}_{i_1 i_2 \cdots i_k} = a^{(1)}_{i_1} \ a^{(2)}_{i_2} \ \cdots \ a^{(k)}_{i_k}. \tag{5}$$

Any real k-way tensor $\mathcal{A}$ can always be written as a linear combination of rank-1 terms

$$\mathcal{A} = \sum_{j=1}^{R} \sigma_j \ a^{(1)}_j \circ a^{(2)}_j \circ \cdots \circ a^{(k)}_j, \tag{6}$$

where $\sigma_j \in \mathbb{R}$ and all the $a^{(i)}_j$ vectors satisfy $||a^{(i)}_j||_2 = 1$. Such a decomposition is called a polyadic decomposition (PD) of the tensor $\mathcal{A}$. When the equality in Equation 6 holds with a minimal number of terms R, then the PD is called canonical (CPD).[9,10] The number R in the CPD is called the *tensor rank*. Unlike the SVD, each of the rank-1 terms in the (C)PD is not necessarily orthogonal.

The Tucker decomposition[15,16] writes a k-way tensor $\mathcal{A}$ as the following multilinear transformation of a core tensor $\mathcal{S} \in \mathbb{R}^{r_1 \times r_2 \times \cdots \times r_k}$ by factor matrices $U^{(i)} \in \mathbb{R}^{n_i \times r_i}$

$$\mathcal{A} = \mathcal{S} \times_1 U^{(1)} \times_2 U^{(2)} \times_3 \cdots \times_k U^{(k)}, \tag{7}$$

which can also be written as Equation 6 where each σ_j is now an entry of the core tensor $\mathcal{S}$. Each rank-1 term of the Tucker decomposition is then given by

$$\mathcal{S}_{i_1 i_2 \cdots i_k} \ U^{(1)}(:, i_1) \circ U^{(2)}(:, i_2) \circ \cdots \circ U^{(k)}(:, i_k),$$

where we use Matlab colon notation to denote columns of the $U^{(i)}$ factor matrices. The minimal size of the core tensor $\mathcal{S}$ such that the equality in Equation 7 holds is called the *multilinear rank*. The higher-order SVD (HOSVD)[11] is a Tucker decomposition where the core tensor $\mathcal{S}$ has the same dimensions as the original tensor $\mathcal{A}$ and with the additional property that the factor matrices $U^{(i)}$ and the slices of $\mathcal{S}$ in the same mode are orthogonal. This implies that each rank-1 term is orthogonal to all other rank-1 terms in the HOSVD, which has the immediate advantage that the approximation error can be determined as in Equation 3.

The PARATREE/TTr1SVD decomposition[12,17] is another decomposition of a k-way tensor into orthogonal rank-1 terms as described by Equation 6. The total number of terms in the TTr1SVD is upperbounded by $R = \prod_{r=0}^{k-2} \min(n_{r+1}, \prod_{i=r+2}^{k} n_i)$ and therefore depends on the ordering of the indices. This decomposition is computed from repeated reshapings and SVD computations and is unique for a fixed order of indices. Note that although each of the rank-1 terms is orthogonal with respect to all others, unlike the HOSVD, the factor matrices $U^{(i)}$ are not orthogonal. In addition, the scalar σ_j coefficients obtained in the PARATREE/TTR1SVD decomposition are guaranteed to be nonnegative. This has the advantage that one can plot these σ_j coefficients in descending order and inspect the relative weight of each of the rank-1 terms in a very straightforward manner, just like one can do with the singular values of a matrix.

3 | TENSOR KP

3.1 | Definition

The definition of the KP for two matrices is well known. If $B \in \mathbb{R}^{m_1 \times m_2}$ and $C \in \mathbb{R}^{n_1 \times n_2}$, then their KP $B \otimes C$ is an $m_1 \times m_2$ block matrix whose (i_3, i_4)th block is the $n_1 \times n_2$ matrix $B_{i_3 i_4} C$

$$\begin{pmatrix} B_{11} & \cdots & B_{1n_1} \\ \vdots & \ddots & \vdots \\ B_{m_1 1} & \cdots & B_{m_1 n_1} \end{pmatrix} \otimes C = \begin{pmatrix} B_{11}C & \cdots & B_{1n_1}C \\ \vdots & \ddots & \vdots \\ B_{m_1 1}C & \cdots & B_{m_1 n_1}C \end{pmatrix}. \tag{8}$$

Generalizing this definition to the KP of k-way tensors is quite straightforward, although not in the form as given by Equation 8. Before giving the definition of the tensor KP, we first investigate how the entries of the matrix KP are described by the indices of the original matrices. The following lemma is easily verified.

Lemma 3.1. If the entries of the matrices $B \in \mathbb{R}^{m_1 \times m_2}, C \in \mathbb{R}^{n_1 \times n_2}$ are denoted $B_{i_3 i_4}$ and $C_{i_1 i_2}$, respectively, then the entries of their Kronecker product $A = B \otimes C$ are described by $A_{[i_1 i_3][i_2 i_4]} = B_{i_3 i_4}\ C_{i_1 i_2}$ for all possible values of i_1, i_2, i_3, i_4.

Remember that the grouped indices $[i_1 i_3], [i_2 i_4]$ are easily converted into the linear row index $i_1 + n_1(i_3 - 1)$ and the linear column index $i_2 + n_2(i_4 - 1)$, respectively. The definition of the tensor KP follows from the generalization of Lemma 3.1 to multiple indices.

Definition 1. Let $\mathcal{B} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_k}, \mathcal{C} \in \mathbb{R}^{m_1 \times m_2 \times \cdots \times m_k}$ be two k-way tensors with entries denoted by $\mathcal{B}_{i_{k+1} \cdots i_{2k}}, \mathcal{C}_{i_1 \cdots i_k}$, respectively. The tensor KP $\mathcal{A} = \mathcal{B} \otimes \mathcal{C} \in \mathbb{R}^{n_1 m_1 \times n_2 m_2 \times \cdots \times n_k m_k}$ is then defined from

$$\mathcal{A}_{[i_1 i_{k+1}][i_2 i_{k+2}] \cdots [i_k i_{2k}]} = \mathcal{B}_{i_{k+1} \cdots i_{2k}}\ \mathcal{C}_{i_1 \cdots i_k},$$

which needs to hold for all possible values of $i_1, \ldots, i_{2k}$.

An equivalent definition of the tensor KP can be found in Phan et al,[18] where a decomposition similar to the TKPSVD but limited to only two factors, is discussed. One possible implementation of the tensor KP would be to use Definition 1 over all possible values of the indices $i_1, \ldots, i_{2k}$, but this would not be very efficient. Instead, one can use an existing implementation of the vector or matrix KP ("kron.m" in Matlab) on the vectorized tensors vec($\mathcal{B}$), vec($\mathcal{C}$). Indeed, the entries of $c \triangleq \text{vec}(\mathcal{B}) \otimes \text{vec}(\mathcal{C})$ are indexed by the single grouped index $[i_1 \cdots i_k i_{k+1} \cdots i_{2k}]$. One can then reshape and permute the entries of c such that the desired $[i_1 i_{k+1}][i_2 i_{k+2}] \cdots [i_k i_{2k}]$ index structure is obtained. This is how the tensor KP is implemented in our Matlab/Octave TKPSVD package.

3.2 | Properties of the tensor KP

We briefly list some properties of the tensor KP without going into details. The following properties are easily verified

$$\begin{aligned} \mathcal{A} \otimes (\mathcal{B} + \mathcal{C}) &= \mathcal{A} \otimes \mathcal{B} + \mathcal{A} \otimes \mathcal{C}, \\ (\mathcal{A} + B) \otimes \mathcal{C} &= \mathcal{A} \otimes \mathcal{C} + \mathcal{B} \otimes \mathcal{C}, \\ (\alpha\mathcal{A}) \otimes \mathcal{B} &= \mathcal{A} \otimes (\alpha\mathcal{B}) = \alpha(\mathcal{A} \otimes \mathcal{B}), \\ (\mathcal{A} \otimes \mathcal{B}) \otimes \mathcal{C} &= \mathcal{A} \otimes (\mathcal{B} \otimes \mathcal{C}), \end{aligned}$$

where $\mathcal{A}, \mathcal{B}, \mathcal{C}$ are k-way tensors and α is a scalar. Just as in the matrix case, the tensor KP is not commutative but permutation equivalent. That is, there exists permutation matrices $P_1, \ldots, P_k$ such that

$$\mathcal{A} \otimes \mathcal{B} = (\mathcal{B} \otimes \mathcal{A}) \times_1 P_1 \times_2 \cdots \times_k P_k. \tag{9}$$

This is easily seen from the definition. Indeed, suppose we have that $\mathcal{C} = \mathcal{A} \otimes \mathcal{B}$ and $\tilde{\mathcal{C}} = \mathcal{B} \otimes \mathcal{A}$, then

$$\begin{aligned} \mathcal{C}_{[i_1 i_{k+1}][i_2 i_{k+2}] \cdots [i_k i_{2k}]} &= \mathcal{A}_{i_{k+1} \cdots i_{2k}}\ \mathcal{B}_{i_1 \cdots i_k}, \\ \tilde{\mathcal{C}}_{[i_{k+1} i_1][i_{k+2} i_2] \cdots [i_{2k} i_k]} &= \mathcal{A}_{i_1 \cdots i_k}\ \mathcal{B}_{i_{k+1} \cdots i_{2k}}. \end{aligned}$$

Now let $P_1, \ldots, P_k$ be the permutation matrices that swap $[i_{k+1} i_1]$ into $[i_1 i_{k+1}]$, $[i_{k+2} i_2]$ into $[i_2 i_{k+2}]$, $\ldots$, $[i_{2k} i_k]$ into $[i_k i_{2k}]$, respectively, then Equation 9 follows. Furthermore, if $\mathcal{A}, \mathcal{B}$ are cubical and of the same dimension then $\mathcal{C}$ and $\tilde{\mathcal{C}}$ are permutation similar, which means that $P_1 = P_2 = \cdots = P_k$.

The mixed-product property of the KP states that if A, B, C, D are matrices such that one can form the matrix products AC, BD, then $(A \otimes B)(C \otimes D) = (AC) \otimes (BD)$. This is called the mixed-product property, because it mixes the ordinary matrix product and the KP. We can also write the mixed-product property using the one-mode product as $(C \otimes D) \times_1 (A \otimes B) = (C \times_1 A) \otimes (D \times_1 B)$. Its generalization involves the mixing of the r-mode product with the tensor KP. Let A, B be matrices and $\mathcal{C}, \mathcal{D}$ be k-way tensors of appropriate dimensions then for any $r \in \{1, \ldots, k\}$

$$(\mathcal{C} \otimes \mathcal{D}) \times_r (A \otimes B) = (\mathcal{C} \times_r A) \otimes (\mathcal{D} \times_r B). \tag{10}$$

What the mixed-product property tells us is that we can obtain the r-mode product of the tensor $\mathcal{C} \otimes \mathcal{D}$ with the matrix $A \otimes B$ from the KP of $(\mathcal{C} \times_r A)$ with $(\mathcal{D} \times_r B)$. Choosing $r = 1$ and replacing $\mathcal{C}, \mathcal{D}$ with matrices in Equation 10 results in the matrix mixed-product property.

4 | TKPSVD THEOREM AND ALGORITHM

Using Definition 1, we can easily extend the tensor KP to multiple factors. Suppose we have three 3-way tensors $\mathcal{A}^{(1)}, \mathcal{A}^{(2)}, \mathcal{A}^{(3)}$, then their KP $\mathcal{A} = \mathcal{A}^{(3)} \otimes \mathcal{A}^{(2)} \otimes \mathcal{A}^{(1)}$ is completely characterized by the following relationship

$$\mathcal{A}_{[i_1 i_4 i_7][i_2 i_5 i_8][i_3 i_6 i_9]} = \mathcal{A}^{(1)}_{i_1 i_2 i_3}\ \mathcal{A}^{(2)}_{i_4 i_5 i_6}\ \mathcal{A}^{(3)}_{i_7 i_8 i_9}. \tag{11}$$

Now, suppose we permute all entries such that $\tilde{\mathcal{A}}_{[i_1 i_2 i_3][i_4 i_5 i_6][i_7 i_8 i_9]} \triangleq \mathcal{A}_{[i_1 i_4 i_7][i_2 i_5 i_8][i_3 i_6 i_9]}$ and that this is a rank-1 tensor. According to Equation 5, $\tilde{\mathcal{A}}$ can then be written as the following outer product of vectors $\tilde{\mathcal{A}} = a^{(1)} \circ a^{(2)} \circ a^{(3)}$ with

$$\tilde{\mathcal{A}}_{[i_1 i_2 i_3][i_4 i_5 i_6][i_7 i_8 i_9]} = a^{(1)}_{[i_1 i_2 i_3]}\ a^{(2)}_{[i_4 i_5 i_6]}\ a^{(3)}_{[i_7 i_8 i_9]}. \tag{12}$$

Comparison of Equation 11 with Equation 12 allows us to conclude that $a^{(1)} = \mathrm{vec}(\mathcal{A}^{(1)})$, $a^{(2)} = \mathrm{vec}(\mathcal{A}^{(2)})$ and $a^{(3)} = \mathrm{vec}(\mathcal{A}^{(3)})$. We formalize this observation in the following theorem.

Theorem 1. For a given k-way tensor $\mathcal{A} \in \mathbb{R}^{n_1 \times n_2 \times \cdots \times n_k}$, if

$$\mathcal{A} = \sum_{j=1}^{R} \sigma_j\ \mathcal{A}_j^{(d)} \otimes \cdots \otimes \mathcal{A}_j^{(2)} \otimes \mathcal{A}_j^{(1)} \text{ and } \tilde{\mathcal{A}} = \sum_{j=1}^{R} \sigma_j\ a_j^{(1)} \circ a_j^{(2)} \circ \cdots \circ a_j^{(d)}, \tag{13}$$

where $\tilde{\mathcal{A}}$ is the permutation of $\mathcal{A}$ such that the indices of the $a_j^{(i)}$ vectors are identical to those of the k-way $\mathcal{A}_j^{(i)}$ tensors, then $a_j^{(i)} = \mathrm{vec}(\mathcal{A}_j^{(i)})$ for all $i \in \{1, \ldots, d\}, j \in \{1, \ldots, R\}$.

Observe that the order of the KPs in Equation 13 is reversed with respect to the order of the outer products. Theorem 1 is crucial for the TKPSVD algorithm, because it tells us that the desired decomposition (Equation 1) can be computed from a PD of $\tilde{\mathcal{A}}$. We now derive the TKPSVD algorithm by means of a simple example. Suppose we have a three-way tensor $\mathcal{A}$ for which we want to find a degree-3 decomposition. This implies, as shown in Equation 11, that each entry of $\mathcal{A}$ is labeled as $\mathcal{A}_{[i_1 i_4 i_7][i_2 i_5 i_8][i_3 i_6 i_9]}$. Figure 1(a) illustrates how the grouped indices of the tensor $\mathcal{A}$ relate to those of the KP factors $\mathcal{A}_j^{(i)}$. Theorem 1 tells us that the desired TKSPVD can be obtained from a PD of the permuted tensor $\tilde{\mathcal{A}}$. The first step in the TKPSVD algorithm is then to permute the indices of $\mathcal{A}$ such that their order corresponds with $i_1, i_2, i_3, i_4, i_5, i_6, i_7, i_8, i_9$. In order to do this, we first reshape the three-way tensor $\mathcal{A}$ into the nine-way tensor $\mathcal{A}$ with entries $\mathcal{A}_{i_1 i_4 i_7 i_2 i_5 i_8 i_3 i_6 i_9}$. The indices of $\mathcal{A}$ are then permuted into the desired order $\tilde{\mathcal{A}}_{i_1 i_2 i_3 i_4 i_5 i_6 i_7 i_8 i_9}$. The next step of the TKPSVD algorithm is to compute the KP factors $\mathcal{A}_j^{(i)}$, each of which is computed as a vector in a PD. We therefore group the indices such that we obtain the three-way tensor $\tilde{\mathcal{A}}$ with entries $\tilde{\mathcal{A}}_{[i_1 i_2 i_3][i_4 i_5 i_6][i_7 i_8 i_9]}$. The steps prior to the computation of the PD are hence summarized as

$$\mathcal{A}_{[i_1 i_4 i_7][i_2 i_5 i_8][i_3 i_6 i_9]} \xrightarrow{\text{reshape}} \mathcal{A}_{i_1 i_4 i_7 i_2 i_5 i_8 i_3 i_6 i_9} \xrightarrow{\text{permute}} \tilde{\mathcal{A}}_{i_1 i_2 i_3 i_4 i_5 i_6 i_7 i_8 i_9}$$
$$\xrightarrow{\text{reshape}} \tilde{\mathcal{A}}_{[i_1 i_2 i_3][i_4 i_5 i_6][i_7 i_8 i_9]}.$$

Each of the KP factors $\mathcal{A}_j^{(i)}$ is obtained from reshaping the $a_j^{(i)}$ vectors of the PD (Figure 2) into a three-way tensor of the correct dimensions. In order to make sure this procedure of reshaping and applying the permutation is clear, we also demonstrate it for a simple matrix example. Suppose we have a 12×12 matrix A, which we want to decompose into a sum of KPs of a 4×4 matrix with a 3×3 matrix. If the entries of the KP factors $A_j^{(1)}, A_j^{(2)}$ are labeled by $i_1 i_2, i_3 i_4$, respectively, then the row index of A is $[i_1 i_3]$ and the column index is $[i_2 i_4]$, shown in Figure 1(b). The steps prior to the computation of the PD are now

$$\mathcal{A}_{[i_1 i_3][i_2 i_4]} \xrightarrow{\text{reshape}} \mathcal{A}_{i_1 i_3 i_2 i_4} \xrightarrow{\text{permute}} \tilde{\mathcal{A}}_{i_1 i_2 i_3 i_4} \xrightarrow{\text{reshape}} \tilde{\mathcal{A}}_{[i_1 i_2][i_3 i_4]}.$$

The dimensions of the tensors in each of these steps are $12 \times 12, 4 \times 3 \times 4 \times 3, 4 \times 4 \times 3 \times 3$, and 16×9, respectively. The final step is to compute a PD with orthogonal rank-1 terms, which for the matrix case is the SVD. The previous two examples

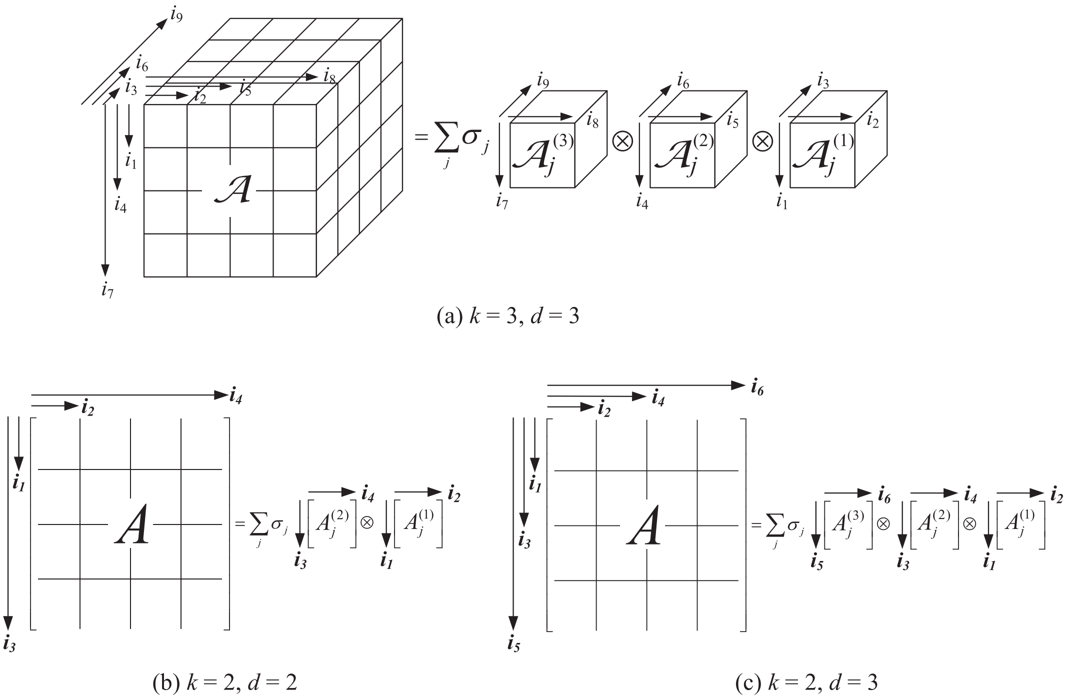


FIGURE 1 How the grouped indices of $\mathcal{A}$ relate to the indices of the Kronecker product factors $\mathcal{A}_j^i$

might give the impression that d and k need to be equal in the TKPSVD. This is not the case. In Figure 1(c), we show a degree-3 TKPSVD of a matrix. Here, the steps prior to the computation of the PD are

$$\mathcal{A}_{[i_1 i_3 i_5][i_2 i_4 i_6]} \xrightarrow{\text{reshape}} \mathcal{A}_{i_1 i_3 i_5 i_2 i_4 i_6} \xrightarrow{\text{permute}} \tilde{\mathcal{A}}_{i_1 i_2 i_3 i_4 i_5 i_6} \xrightarrow{\text{reshape}} \tilde{\mathcal{A}}_{[i_1 i_2][i_3 i_4][i_5 i_6]}.$$

The pseudocode for our general TKPSVD algorithm is presented in Algorithm 4.1. As we will show in Section 6, the structure-preserving property of the TKPSVD critically depends on the fact that the rank-1 terms of the computed PD are orthogonal with respect to one another. Another consequence of this orthogonality is that

$$||\mathcal{A}||_F = \sqrt{\sigma_1^2 + \cdots + \sigma_R^2},$$

where R is the total number of terms in the decomposition. The relative approximation error obtained from truncating the number of terms to $r < R$ is then easily computed as

$$\frac{||A - \sum_{j=1}^{r} \sigma_j \, A_j^{(d)} \otimes \cdots \otimes A_j^{(1)}||_F}{||A||_F} = \frac{\sqrt{\sigma_{r+1}^2 + \cdots + \sigma_R^2}}{\sqrt{\sigma_1^2 + \cdots + \sigma_R^2}}.$$

This is especially convenient when using the TTr1SVD[12] to compute the PD, since then all σ_j's are positive and can be sorted in descending order just like singular values in the matrix case.

Algorithm 4.1
TKPSVD Algorithm
Input: tensor $\mathcal{A}$, dimensions $n_1^{(1)}, \ldots, n_1^{(d)}, n_2^{(1)}, \ldots, n_2^{(d)}, \ldots, n_k^{(1)}, \ldots, n_k^{(d)}$
Output: $\sigma_1, \ldots, \sigma_R$, tensors $\mathcal{A}_j^{(i)}$

$\mathcal{A} \leftarrow$ reshape $\mathcal{A}$ into a (kd)-way tensor according to $n_1^{(1)}, \ldots, n_k^{(d)}$
$\tilde{\mathcal{A}} \leftarrow$ permute $\mathcal{A}$ to indices $i_1 i_2 i_3 i_4 \cdots i_{kd-1} i_{kd}$
$\tilde{\mathcal{A}} \leftarrow$ reshape $\tilde{\mathcal{A}}$ into a d-way tensor by grouping every k indices together
$a_1^{(1)}, \ldots, a_R^{(d)}, \sigma_1, \ldots, \sigma_R \leftarrow$ compute a PD of $\tilde{\mathcal{A}}$ with orthogonal rank-1 terms
for all nonzero σ_j **do**
  $\mathcal{A}_j^{(i)} \leftarrow$ reshape $a_j^{(i)}$ into a $n_1^{(i)} \times \cdots \times n_k^{(i)}$ tensor
end for

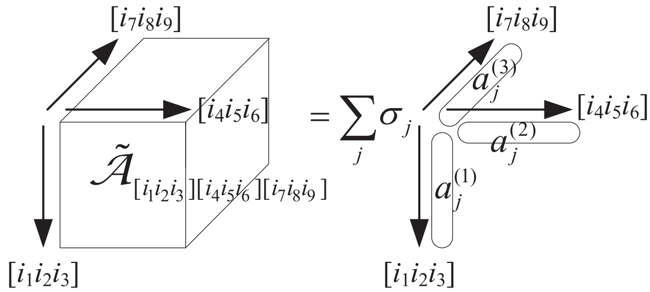


FIGURE 2 Decomposition of the three-way tensor $\tilde{\mathcal{A}}$ into a linear combination of rank-1 terms

The PD with orthogonal rank-1 terms is easily computed for the case $d = 2$ as the SVD. When $d \geq 3$, several options are available. A first option is to compute the CPD with the additional constraint of orthogonal factor matrices. This orthogonality constraint limits the size of the factor matrices and consequently also the total number of rank-1 terms that are possible to find. As a result, the CPD with orthogonality constraints does not lend itself very well to applications. We demonstrate this with a worked out example in Section 7. Alternatively, one could compute the HOSVD or the TTr1SVD of $\tilde{\mathcal{A}}$, as these decompositions have orthogonal rank-1 terms. The CPD with orthogonal factor matrices and the HOSVD can be computed in Matlab using Tensorlab,[19] freely available from http://www.tensorlab.net/. A Matlab/Octave implementation of Algorithm 4.1 that uses the TTr1SVD and works for any arbitrary degree d and tensor order k can be freely downloaded from https://github.com/kbatseli/TKPSVD.

In Section 5, we introduce a new framework in which many different structured tensors (symmetric, persymmetric, centrosymmetric, Toeplitz, Hankel, …) can be described and then prove that the TKPSVD algorithm guarantees to preserve these structures in the KP factors. But first, we present a small modification of Algorithm 4.1 for the case of diagonal tensors.

4.1 | Diagonal tensors

A diagonal tensor $\mathcal{D}$ is an extremely simple symmetric tensor (see Section 5 for the definition). If we define the main diagonal of a cubical tensor as the entries $A_{i_1 i_2 \cdots i_k}$ with $i_1 = i_2 = \cdots = i_k$, then the entries not on the main diagonal of a diagonal tensor are per definition zero. It is easy to see that the KP of two diagonal tensors is also diagonal. This motivates us to adjust Algorithm 4.1 such that only the main diagonal entries $\mathcal{D}_{i_1 i_1 \cdots i_1}$ are considered. This reduces the number of entries to store in memory from n^d to n. As a result, the diagonal tensor $\mathcal{D}$ is decomposed into a KP of diagonal factors $\mathcal{D}_j^{(i)}$. Suppose a degree d TKPSVD of a diagonal tensor $\mathcal{D}$ is required. We then consider the vector a that contains all main diagonal entries with entries $a_{i_1 i_2 \cdots i_d}$ and reshape it into a d-way tensor $\mathcal{A}$. Note that because the indices are already in the right order, no permutation of indices is required and the PD decomposition can be directly computed from $\mathcal{A}$. Each KP factor $\mathcal{D}_j^{(i)}$ of Equation 1 is then an $n_i \times \cdots \times n_i$ diagonal tensor with main diagonal entries $a_j^{(i)}$. The pseudocode for the diagonal TKPSVD algorithm is summarized in Algorithm 4.2.

Algorithm 4.2
TKPSVD Algorithm for a diagonal tensor
Input: diagonal tensor $\mathcal{D}$, dimensions $n_1, n_2, \ldots, n_d$
Output: $\sigma_1, \ldots, \sigma_R$, diagonal tensors $D_j^{(i)}$

```
a ← main diagonal entries of D
A ← reshape a into an n_1 × n_2 × … × n_d tensor
a_1^(1), … , a_R^(d), σ_1, … , σ_R ← compute a PD of A with orthogonal rank-1 terms
for all nonzero σ_j do
    D_j^(i) ← a n_i × ⋯ × n_i diagonal tensor with main diagonal entries a_j^(i)
end for
```

It is interesting to investigate whether it is possible to adjust Algorithm 4.1 to exploit other specific structures, such as the general symmetries, which we define in Section 5. At first sight, this does not seem to be straightforward to exploit, because $\tilde{\mathcal{A}}$ will not retain the original structure. We keep this problem for future research.

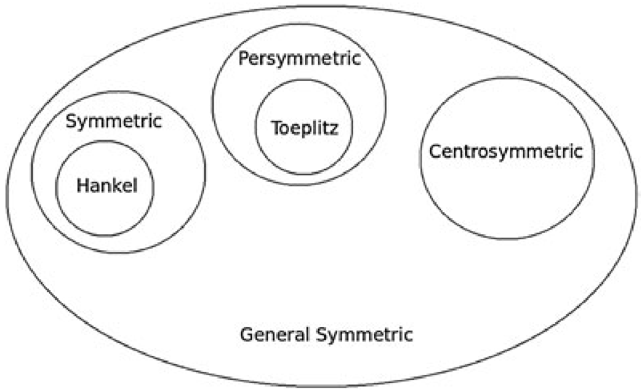


FIGURE 3 Overview general symmetric tensors

5 | GENERAL SYMMETRIC TENSORS

It turns out that the d-way tensor $\tilde{\mathcal{A}}$ in Algorithm 4.1 allows us to generalize the notion of symmetric tensors in a very natural way. The motivation of introducing general symmetry lies in the fact that then only one proof suffices to show the preservation of symmetry, persymmetry, centrosymmetry, and many other symmetries in the KP factors of the TKPSVD. This new framework also provides a different perspective of describing and investigating these different tensor structures. Figure 3 shows an overview of how the notion of general symmetry encapsulates symmetric, persymmetric, centrosymmetric, Toeplitz, and Hankel tensors. The key idea of the general symmetric structure is that it involves particular permutations P of the entries of vec($\mathcal{A}$) that can be decomposed into KP of smaller permutations along each mode of $\tilde{\mathcal{A}}$. We discuss and demonstrate this decomposition of the permutation P for three particular cases. In the remainder of this section, we always assume that $\mathcal{A}$ is a k-way cubical tensor of dimensions n.

5.1 | Symmetry

The symmetric structure of a k-way cubical tensor $\mathcal{A}$ can be defined as a particular permutation of the entries of vec($\mathcal{A}$). This permutation is described by the perfect shuffle matrix.

Definition 2. The perfect shuffle matrix S is the $n^k \times n^k$ permutation matrix

$$S = \begin{pmatrix} I(1 : n^{k-1} : n^k, :) \\ I(2 : n^{k-1} : n^k, :) \\ \vdots \\ I(n^{k-1} : n^{k-1} : n^k, :) \end{pmatrix},$$

where I is the $n^k \times n^k$ identity matrix and Matlab colon notation is used to denote submatrices.

It is easily verified that for a symmetric k-way tensor $\mathcal{A}$, we have that S vec($\mathcal{A}$) = vec($\mathcal{A}$). We now turn this reasoning on its head and define a symmetric tensor as any tensor $\mathcal{A}$ that satisfies S vec($\mathcal{A}$) = vec($\mathcal{A}$). The perfect shuffle matrix S reduces to the matrix defined in Van Loan[2, p. 86] for the case $k = 2$. In this sense, Definition 2 generalizes the notion of a perfect shuffle matrix to tensors.

In what follows, we will apply the TKPSVD algorithm to construct the $\tilde{\mathcal{A}}$ tensor and see how this affects the equation S vec($\mathcal{A}$) = vec($\mathcal{A}$). In order to illustrate this process, we will consider the three-way example tensor from Section 4 and suppose that it is symmetric. This symmetry implies that

$$\mathcal{A}_{[i_1 i_4 i_7][i_2 i_5 i_8][i_3 i_6 i_9]} = \mathcal{A}_{[i_2 i_5 i_8][i_3 i_6 i_9][i_1 i_4 i_7]} = \cdots = \mathcal{A}_{[i_3 i_6 i_9][i_1 i_4 i_7][i_2 i_5 i_8]}. \tag{14}$$

In other words, the symmetry of $\mathcal{A}$ is equivalent with swapping i_1 with either i_2 or i_3, i_4 with either i_5 or i_6, and i_7 with either i_8 or i_9. The TKPSVD algorithm reshapes and permutes the symmetric tensor $\mathcal{A}$ into the tensor $\tilde{\mathcal{A}}$, with entries $\tilde{\mathcal{A}}_{[i_1 i_2 i_3][i_4 i_5 i_6][i_7 i_8 i_9]}$. Although $\tilde{\mathcal{A}}$ is not symmetric, the symmetry of $\mathcal{A}$ still allows us to swap the indices as indicated in Equation 14 such that

$$\tilde{\mathcal{A}}_{[i_1 i_2 i_3][i_4 i_5 i_6][i_7 i_8 i_9]} = \tilde{\mathcal{A}}_{[i_4 i_5 i_6][i_7 i_8 i_9][i_1 i_2 i_3]} = \cdots = \tilde{\mathcal{A}}_{[i_7 i_8 i_9][i_1 i_2 i_3][i_4 i_5 i_6]},$$

which can be rewritten as

$$\tilde{\mathcal{A}} \times_1 S_1 \times_2 S_2 \times_3 S_3 = \tilde{\mathcal{A}}, \tag{15}$$

where all S_i matrices are perfect shuffle matrices. By using Equation 4, Equation 15 can be rewritten as

$$(S_3 \otimes S_2 \otimes S_1)\ \mathrm{vec}(\tilde{\mathcal{A}}) = \mathrm{vec}(\tilde{\mathcal{A}}), \tag{16}$$

which is nothing else but a reformulation of the symmetry $S\ \mathrm{vec}(\mathcal{A}) = \mathrm{vec}(\mathcal{A})$ in terms of $\mathrm{vec}(\tilde{\mathcal{A}})$. If Q denotes the permutation matrix such that $Q\ \mathrm{vec}(A) = \mathrm{vec}(\tilde{\mathcal{A}})$, then from

$$\begin{aligned} &(S_3 \otimes S_2 \otimes S_1)\ \mathrm{vec}(\tilde{\mathcal{A}}) = \mathrm{vec}(\tilde{\mathcal{A}}),\\ \iff &(S_3 \otimes S_2 \otimes S_1)\ Q\ \mathrm{vec}(A) = Q\ \mathrm{vec}(A),\\ \iff &Q^T\ (S_3 \otimes S_2 \otimes S_1)\ Q\ \mathrm{vec}(A) = \mathrm{vec}(A), \end{aligned}$$

we infer that $S = Q^T(S_3 \otimes S_2 \otimes S_1)Q$. This can be interpreted as the perfect shuffle matrix S being "decomposed" into a KP of smaller perfect shuffle matrices. Another way of seeing this equality is that S and $S_3 \otimes S_2 \otimes S_1$ are permutation similar.

5.2 | Centrosymmetry

Another interesting and useful permutation of the entries of $\mathrm{vec}(\mathcal{A})$ is the exchange matrix J, which is the $n^k \times n^k$ column-reversed identity matrix. This permutation maps each index i_j of $\mathrm{vec}(\mathcal{A})$ to $n - i_j + 1$, for example, for the three-way tensor $\mathcal{A}$ from Section 4, the entry $\mathrm{vec}(\mathcal{A})_{[i_1 i_4 i_7 i_2 i_5 i_8 i_3 i_6 i_9]}$ is mapped to

$$\mathrm{vec}(\mathcal{A})_{[n-i_1+1\ n-i_4+1\ n-i_7+1\ n-i_2+1\ n-i_5+1\ n-i_8+1\ n-i_3+1\ n-i_6+1\ n-i_9+1]}$$

and vice-versa. A k-way cubical tensor $\mathcal{A}$ is defined to be centrosymmetric when

$$J\ \mathrm{vec}(\mathcal{A}) = \mathrm{vec}(\mathcal{A}).$$

Our definition of centrosymmetric tensors is equivalent with the alternative definitions given in Chen et al[20] and Zhao and Yang.[21] The "decomposition" argument of J is completely analogous to the decomposition of the perfect shuffle matrix S for symmetric tensors. Following the TKPSVD reshapings and permutations leads to the expression $J = Q^T(J_3 \otimes J_2 \otimes J_1)Q$, where J_1, J_2, J_3 are exchange matrices and Q is the same permutation as described in Section 5.1.

5.3 | Persymmetry

Given the definition of the perfect shuffle matrix S and exchange matrix J, we now define a k-way cubical tensor $\mathcal{A}$ to be persymmetric when

$$S\ J\ \mathrm{vec}(\mathcal{A}) = \mathrm{vec}(\mathcal{A})$$

applies. Using similar arguments as in the symmetric and centrosymmetric cases, we can write the following decomposition $SJ = Q^T(S_3J_3 \otimes S_2J_2 \otimes S_1J_1)Q$. Using the mixed-product property of the KP, we can rewrite the permutation decomposition as $SJ = Q^T(S_3 \otimes S_2 \otimes S_1)(J_3 \otimes J_2 \otimes J_1)Q$.

5.4 | General symmetric tensor

We now define general symmetric tensors by generalizing the previous three examples of particular symmetries.

Definition 3. A k-way cubical tensor $\mathcal{A}$ is a general symmetric tensor if

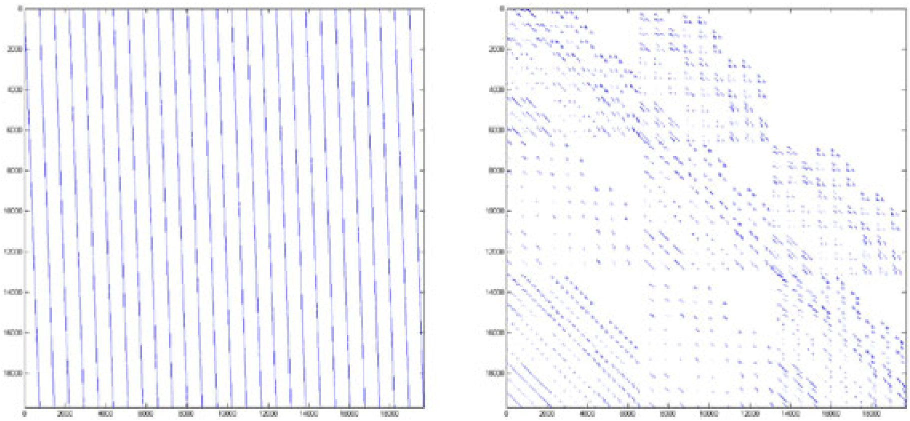
$$P\ \mathrm{vec}(\mathcal{A}) = \mathrm{vec}(\mathcal{A}),$$

where the permutation matrix P can be written for any arbitrary degree d and dimensions $n_r^{(i)}$ into a KP of smaller permutation matrices $P_1, \ldots, P_d$ as

$$P = Q^T\ (P_d \otimes \cdots \otimes P_2 \otimes P_1)\ Q, \tag{17}$$

where Q is the permutation matrix such that $Q\ \mathrm{vec}(\mathcal{A}) = \mathrm{vec}(\tilde{\mathcal{A}})$ and the dimension of each of the permutation matrices P_i is $\prod_{r=1}^{k} n_r^{(i)}$.

General skew-symmetric tensors are defined similarly as in Definition 3 where now $P\ \mathrm{vec}(\mathcal{A}) = -\mathrm{vec}(\mathcal{A})$ needs to hold. The permutation matrices $P_1, \ldots, P_k$ do not necessarily need to be of the same dimension. In fact, the definition requires that Equation 17 holds for any arbitrary degree d and dimensions $n_r^{(i)}$ of the KP factors. For example, there are several ways in which a general symmetric $12 \times 12 \times 12$ tensor $\mathcal{A}$ can be decomposed into a TKPSVD. There are three different decompositions when $d = 3$, depending on the order of the $3 \times 3 \times 3, 2 \times 2 \times 2$, and $2 \times 2 \times 2$ tensors. Likewise, when $d = 2$, there are different orderings of the $3 \times 3 \times 3, 4 \times 4 \times 4$ or $6 \times 6 \times 6, 2 \times 2 \times 2$ tensors. Each of these TKPSVDs is characterized by different P_k and Q permutation matrices; nevertheless, Equation 17 needs to hold for all of them for $\mathcal{A}$ to be general symmetric.



(a) Perfect shuffle matrix S (b) Hankel permutation matrix P

FIGURE 4 Permutation matrices for a $27 \times 27 \times 27$ Hankel tensor

5.5 | Shifted-index structure

Within the set of general symmetric tensors, there are other interesting, more restrictive structures that we call shifted-index structures. These are tensors whose entries do not change when at least one index is "shifted."

Definition 4. A k-way cubical tensor $\mathcal{A}$ has a shifted-index structure if

$$\mathcal{A}_{[i_1][i_2]\cdots[i_k]} = \mathcal{A}_{[i_1+\Delta_1][i_2+\Delta_2]\cdots[i_k+\Delta_k]},$$

where at least one of the integer shifts $\Delta_1, \ldots, \Delta_k$ is nonzero. For any two nonzero shifts Δ_i, Δ_j, either $\Delta_i = \Delta_j$ or $\Delta_i = -\Delta_j$ must be satisfied.

Of course, none of the shifted indices $i_1 + \Delta_1, \ldots, i_k + \Delta_k$ are allowed to go "out of bounds." The case where $\Delta_1 = \Delta_2 = \cdots = \Delta_k$ is called a Toeplitz tensor and is a special case of a persymmetric tensor. Similarly, a symmetric tensor for which all shifts are zero except for one arbitrary pair $\Delta_i = -\Delta_j$ is called a Hankel tensor. A tensor for which all shifts are zero except Δ_r has constant entries along the r fibres. It is straightforward to show that any shifted-index structure is also a general symmetry by writing down its corresponding permutation matrix and showing that Definition 3 applies.

Example 1. Consider a $27 \times 27 \times 27$ Hankel tensor $\mathcal{H}$. The Hankel structure means that the tensor is also symmetric, which implies that $S\ \mathrm{vec}(\mathcal{H}) = \mathrm{vec}(\mathcal{H})$ with S the 19683×19683 perfect shuffle matrix. Now consider a degree-3 TKPSVD, where each KP factor is a $3 \times 3 \times 3$ tensor. We can then retrieve S from the 27×27 perfect shuffle matrix $S_1 = S_2 = S_3$ as $S = Q^T(S_3 \otimes S_2 \otimes S_1)Q$, where Q is the permutation matrix in $\mathrm{vec}(\tilde{\mathcal{H}}) = Q\ \mathrm{vec}(\mathcal{H})$. The 27×27 permutation matrices $P_1 = P_2 = P_3$ that define a $3 \times 3 \times 3$ Hankel tensor $\mathcal{A}$ are completely specified by the vector of indices

$$i = [1, 4, 5, 10, 7, 8, 11, 12, 15, 2, 13, 14, 19, 16, 17, 20, 21, 24, 3, 22, 23, 6, 25, 26, 9, 18, 27],$$

because $\mathrm{vec}(\mathcal{A})(i) = \mathrm{vec}(\mathcal{A})$, where $\mathrm{vec}(\mathcal{A})(i)$ is Matlab notation to denote $P_3\ \mathrm{vec}(\mathcal{A})$. If we now set $P = Q^T(P_3 \otimes P_2 \otimes P_1)Q$, then indeed $P\ \mathrm{vec}(\mathcal{H}) = \mathrm{vec}(\mathcal{H})$ is satisfied. Figure 4 shows the nonzero pattern of both S and P. Observe that although the Hankel permutation P is a special case of a symmetry, the nonzero pattern is very different from that of S.

6 | PRESERVATION OF STRUCTURES

It is quite a remarkable fact that all general symmetries, including the shifted-index structures, are preserved in the cubical KP factors $\mathcal{A}_j^{(i)}$ when they are computed according to Algorithm 4.1 . The orthogonality of the rank-1 terms in the PD plays a crucial role in this. Another critical element are the scalar coefficients σ_j's of the TKPSVD, which are required to be distinct. We now show how this comes about.

6.1 | General symmetry

In order to prove general symmetry preservation in the KP factors, we need the following useful lemma.

Lemma 6.1. Suppose $a = \text{vec}(\mathcal{A}) \in \mathbb{R}^{n^k \times 1}$ with $a^T a = 1$ and P is a permutation matrix that corresponds with a general symmetry. Then the cubical tensor $\mathcal{A}$ obtained from reshaping a is general symmetric if and only if $a^T P a = 1$ or general skew symmetric if and only if $a^T P a = -1$.

Proof. We first prove $Pa = a \Rightarrow a^T\ P\ a = 1$. Because a has unit norm, we can write $a^T a = 1$ and substitution of a by Pa then results in $a^T P a = 1$. The proof for $a^T\ P\ a = 1 \Rightarrow Pa = a$ goes as follows. Let $b = Pa$, then we have that $||b||_2 = 1$ and $a^T\ b = \cos\alpha$. Because $\cos\alpha = 1$, it follows that $\alpha = 0$ and a is a multiple of b, but because $||a||_2 = ||b||_2 = 1$, it follows that $a = b = Pa$. The proof for the skew symmetry of $\mathcal{A}$ follows the same logic. □

The general symmetry of $\mathcal{A}$ implies that

$$\tilde{\mathcal{A}} = \tilde{\mathcal{A}} \times_1 P_1 \times_2 P_2 \times_3 \cdots \times_d P_d, \tag{18}$$

where all P_i's are permutation matrices. We now substitute $\tilde{\mathcal{A}}$ in both sides of Equation 18 by its PD and obtain

$$\sum_{j=1}^{R} \sigma_j\ a_j^{(1)} \circ a_j^{(2)} \circ \cdots \circ a_j^{(d)} = \sum_{i=1}^{R} \sigma_j\ P_1 a_j^{(1)} \circ P_2 a_j^{(2)} \circ \cdots \circ P_d a_j^{(d)}. \tag{19}$$

The orthogonality of each rank-1 term and $||a_j^{(i)}||_2 = 1$ implies that the mode products of both sides of Equation 19 with $(a_k^{(1)})^T, \ldots, (a_k^{(d)})^T$ along the modes $1, 2, \ldots, d$, respectively, for any $k \in \{1, \ldots, R\} \subset \mathbb{N}$ results in

$$\sigma_k = \sum_{j=1}^{R} \sigma_j \left(a_k^{(1)}\right)^T P_1\ a_j^{(1)} \left(a_k^{(2)}\right)^T P_2\ a_j^{(2)} \cdots \left(a_k^{(d)}\right)^T P_d\ a_j^{(d)}. \tag{20}$$

We have that each of the $(a_k^{(i)})^T P_i a_j^{(i)}$ scalars lies in the real interval $[-1, 1]$, because $||a_j^{(i)}||_2 = 1$ for all j and P is a permutation matrix. We now assume that the following condition holds.

Condition 1. All σ_j's in the PD of $\tilde{\mathcal{A}}$ are distinct and all terms on the right-hand side of Equation 20 except for the one corresponding with σ_k vanish.

The equality in Equation 20 holds under Condition 1 when $(a_k^{(i)})^T\ P_i\ a_j^{(i)} = 0$ for at least one particular i when $k \neq j$ and when

$$\prod_{i=1}^{d} \left(a_k^{(i)}\right)^T\ P_i\ a_k^{(i)} = 1. \tag{21}$$

The only possible way for Equation 21 to be true under the constraint that each of the $(a_k^{(i)})^T P_i a_k^{(i)}$ scalars lies in the interval $[-1, 1]$ is when

$$\left(a_k^{(i)}\right)^T\ P_i\ a_k^{(i)} = \pm 1. \tag{22}$$

From Lemma 6.1, we know that if the right-hand side of Equation 22 is 1, then $\mathcal{A}_k^{(i)}$ is general symmetric, otherwise $\mathcal{A}_k^{(i)}$ is general skew symmetric. In addtion, Equation 21 implies that there are either zero or an even number of general skew-symmetric KP factors in each term. This proves the main theorem on general symmetry preservation in the TKPSVD.

Theorem 2. Let $\mathcal{A}$ be a general symmetric tensor with a dth-degree TKPSVD into cubical KP factors $A_j^{(i)}$. If Condition 1 holds, then each of the $A_j^{(i)}$ factors in the TKPSVD is either a general symmetric or a general skew-symmetric tensor. There are always either zero or an even number of skew-symmetric factors in each term of Equation 1.

In practice, as soon as all σ_j's in the PD of $\tilde{\mathcal{A}}$ are distinct, then each of the rank-1 terms will be general symmetric. Let us see what happens when Condition 1 is not satisfied. Suppose that $\sigma_k = \sigma_{k+1}$ and that the terms corresponding with the other σ_j's vanish. We can then combine the σ_k and σ_{k+1} terms such that now

$$\prod_{i=1}^{d} \left(a_k^{(i)}\right)^T\ P_i\ a_k^{(i)} + \prod_{i=1}^{d} \left(a_{k+1}^{(i)}\right)^T\ P_i\ a_{k+1}^{(i)} = 1 \tag{23}$$

needs to hold. Contrary to the case of distinct σ_j's, there are now multiple ways that Equation 23 can be satisfied without $(a_k^{(i)})^T\ P_i\ a_k^{(i)} = \pm 1$ being true, which implies that the terms corresponding with σ_k, σ_{k+1} will not necessarily have general symmetric factors. A very particular case where Condition 1 is not satisfied for all terms is for symmetric tensors of order $k > 2$. We discuss this case, along with other general symmetries, in Section 7.

TABLE 1 Comparison of CPD, HOSVD, and TTr1SVD for the TKPSVD

Method	# of terms	Storage (kB)	Runtime (s)	Relative error
Orthogonal CPD	8	3.14	768.07	0.947
HOSVD	6,912	146.19	0.18	2.21e-15
TTr1SVD	216	154.06	0.21	2.39e-15

Note. CPD = canonical polyadic decomposition; HOSVD = higher order singular value decomposition; TKPSVD = tensor-based Kronecker product singular value decomposition; TTr1SVD = tensor-train rank-1 SVD.

7 | NUMERICAL EXPERIMENTS

In this section, we discuss numerical experiments that illustrate different aspects of the TKPSVD algorithm. We will demonstrate the structure preservation of general symmetries and shifted-index structures and discuss a curious observation for symmetric tensors. We also compare the use of the CPD with orthogonal matrix factors, the HOSVD and TTr1SVD in terms of runtime and storage. Finally, we illustrate how the KP structure can be interpreted as a multiresolution decomposition of images. All computations were done in Matlab on a 64-bit 4-core 3.3 GHz desktop computer with 16 GB RAM.

7.1 | General symmetric structure

As a first example, we demonstrate the use of the CPD with orthogonal factor matrices, the HOSVD and the TTr1SVD to compute the TKPSVD, together with the preservation of centrosymmetry in the KP factors.

Example 2. We construct a $24 \times 24 \times 24$ centrosymmetric tensor $\mathcal{A}$ with its distinct entries drawn from a standard normal distribution and compute a third degree decomposition with factor sizes $4 \times 4 \times 4, 3 \times 3 \times 3, 2 \times 2 \times 2$, respectively. The reshaping and permutation steps result in a $8 \times 27 \times 64$ tensor $\tilde{\mathcal{A}}$. Table 1 compares the use of the CPD with orthogonal factor matrices, the HOSVD and TTr1SVD for the PD of $\tilde{\mathcal{A}}$. We list the total number of rank-1 terms, the required memory for storage of the PD of $\tilde{\mathcal{A}}$, the total runtime to compute the TKPSVD (computed as the median over 100 runs), and the relative error $||\mathcal{A} - \sum_{j=1}^{R} \sigma_j \mathcal{A}_j^{(d)} \cdots \mathcal{A}_j^{(1)}||_F / ||\mathcal{A}||_F$. The total number of rank-1 terms for the orthogonal CPD is limited to only 8, because the first factor matrix will be an orthogonal 8×8 matrix. This results in a large relative error. The runtime for computing the orthogonal CPD is also very long compared to the HOSVD and TTr1SVD. The main difference between the HOSVD and TTr1SVD lies in the total number of rank-1 terms. But because the HOSVD reuses the mode vectors, this results in slightly less required memory. All rank-1 terms computed with both the HOSVD and TTr1SVD retain the centroysymmetric structure. The TTr1SVD method results in 56 terms that have two skew-centrosymmetric factors, The HOSVD has 1,792 such terms. Note that the HOSVD has a $8 \times 27 \times 64$ core tensor containing 6,912 nonzero entries.

Due to the limitations of the CPD with orthogonal factors, as demonstrated in Example 2, we will refrain from using it for the following examples in this section. We now demonstrate the occurrence of σ's with multiplicities for symmetric tensors. Consequently, KP terms that belong to the same σ will not inherit the general symmetry in their factors. This is true for when both the TTr1SVD and HOSVD are used. The TTr1SVD case however has a lot more regularity than the HOSVD. When the TKPSVD of a k-way symmetric tensor is computed with the TTr1SVD, then all multiple σ's have a multiplicity of $k - 1$.

Example 3. Consider a symmetric $8 \times 8 \times 8$ tensor with distinct entries drawn from a standard normal distribution. We compute its degree-3 TKPSVD using the TTr1SVD and obtain 56 KP terms. For this three-way tensor, each multiple σ has a multiplicity of $k - 1 = 2$. There are eight such pairs, which implies that 16 terms are not (skew)-symmetric. Next, we compute a degree-3 decomposition for an 8×8 symmetric matrix, which results in 14 KP terms. All the σ's are distinct, which implies that all terms in the decomposition are (skew)-symmetric. Finally, we compute the degree-3 TKPSVD of a four-way symmetric tensor. This decomposition consists of 230 terms. There are 20 three-tuples of multiple σ's, which means that 60 KP terms are not (skew)-symmetric.

7.2 | Shifted-index structure

Next, we investigate the dependence of the total number of KP terms and total runtime on the ordering of the KP factors for both the TTr1SVD and HOSVD.

Example 4. Consider a $64 \times 64 \times 64 \times 64$ Hankel tensor with its distinct entries drawn from a standard normal distribution. We compute its TKPSVD into three Kronecker factors with dimensions $2 \times 2 \times 2 \times 2, 4 \times 4 \times 4 \times 4$, and $8 \times 8 \times 8 \times 8$ over all possible orderings of the factors and investigate the total number of obtained KP terms for both the TTr1SVD and HOSVD, together with total runtimes. The results are shown in Table 2. The first thing to notice is that the total number of terms in the TKPSVD and total runtime are quite independent from the factor ordering when the HOSVD is used. On average, about 2,000 terms are needed and the computation takes a little over 4 min. The TTr1SVD needs about 20 times less terms and is for one particular ordering more than 80 times faster. It can also be seen from Table 2 that the runtime of the TTr1SVD method depends somewhat on the ordering of the dimensions of the tensor. The smallest runtime is achieved when the dimensions are sorted in ascending fashion from left to right. This is inherent to the TTr1SVD algorithm, as it progressively reshapes the tensor and computes matrix SVDs. For the cases 8,2,4 and 8,4,2, the first step of the TTr1SVD algorithm is to compute the economical SVD of a 4096×4096 matrix, and for the 2,4,8 case, the first step is the economical SVD of a 16×1048576 matrix. Note that although the HOSVD requires more terms, just like in Example 2, it will require less memory for storage due to the fact that it reuses the mode vectors for each KP term. All of the terms in every of the decompositions retained the Hankel structure.

7.3 | Multiresolution decomposition of images

An interesting illustration of the TKPSVD is in the multiresolution decomposition of a $n_1 \times n_2 \times 3$ colour image $\mathcal{A}$. This example gives an interpretation to two different aspects of the TKPSVD: truncation of the KP and truncation of the number of terms. Indeed, every pixel of the $n_1^{(d)} \times n_2^{(d)} \times 3$ image $\mathcal{A}^{(d)}$ is "blown up" by each KP in $\mathcal{A}^{(d)} \otimes \mathcal{A}^{(d-1)} \otimes \cdots \otimes \mathcal{A}^{(1)}$ until the resolution $n_1 \times n_2 \times 3$ is obtained. Truncating the KPs to only a few factors hence effectively reduces the resolution. Furthermore, compression can be achieved at different resolutions by retaining only a few terms. The compression rate achieved for retaining k factors and r number of terms in Equation 1 is defined as

$$\frac{\prod_{i=0}^{k-1} n_1^{(d-i)}\, n_2^{(d-i)}\, n_3^{(d-1)}}{r \sum_{i=0}^{k-1} n_1^{(d-i)}\, n_2^{(d-i)}\, n_3^{(d-1)}}.$$

TABLE 2 Number of KP terms and total runtime for TTr1SVD and HOSVD

	# of terms		Runtime (seconds)	
Ordering	**TTr1SVD**	**HOSVD**	**TTr1SVD**	**HOSVD**
2,4,8	65	1968	9.76	254.30
2,8,4	65	1973	2.47	251.70
4,2,8	65	1993	5.80	254.96
4,8,2	65	2146	5.75	256.41
8,2,4	145	2067	247.66	249.97
8,4,2	145	2105	239.95	254.16

Note. HOSVD = higher order singular value decomposition; KP = Kronecker product; TKPSVD = tensor-based Kronecker product singular value decomposition; TTr1SVD = tensor-train rank-1 SVD.

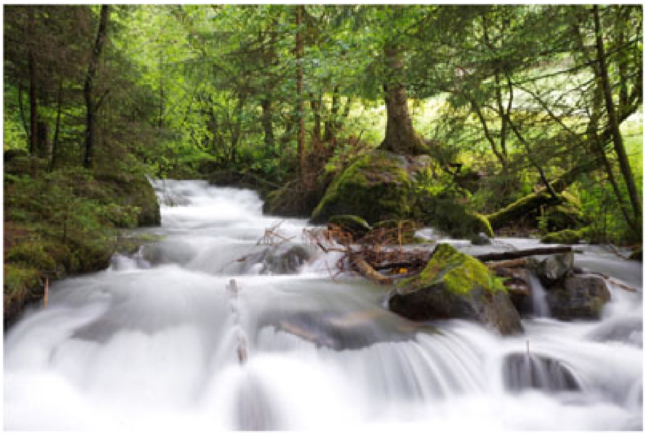


FIGURE 5 Original $4000 \times 6000 \times 3$ image

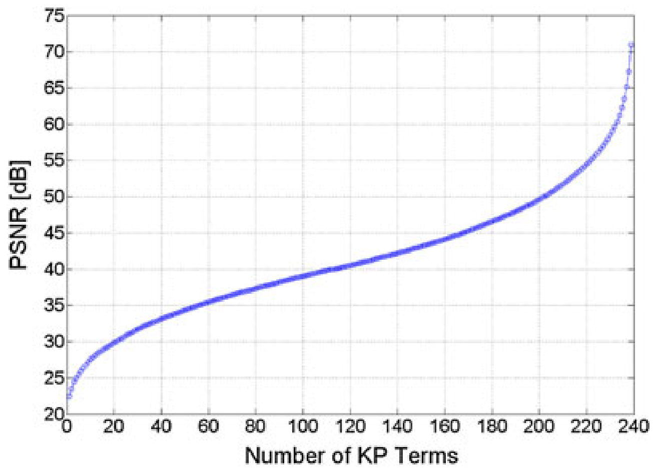
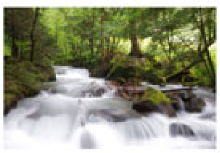
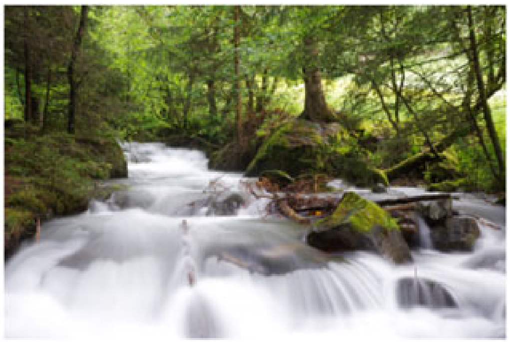


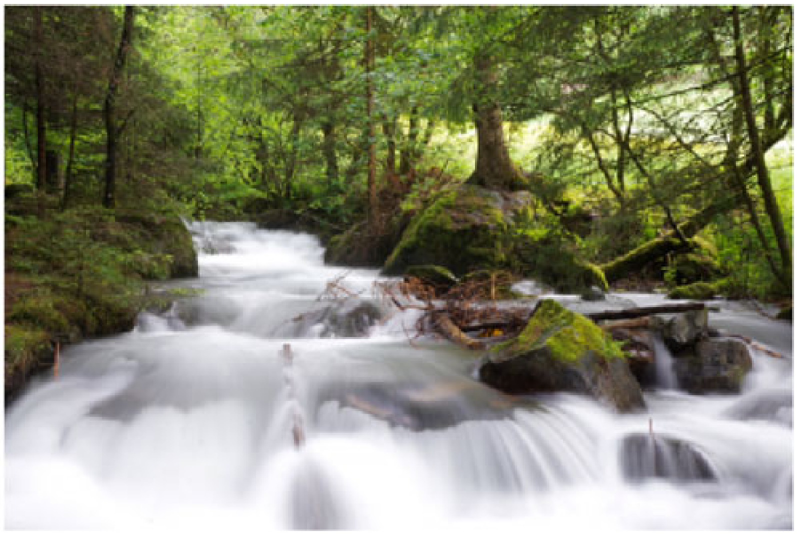
FIGURE 6 Increase of the PSNR as the number of TTr1SVD–KP terms grows for the $4000 \times 6000 \times 3$ resolution. KP=Kronecker product; PSNR=peak signal-to-noise ratio; TTr1SVD=tensor-train rank-1 singular value decomposition



(a) 250×375



(b) 500×750



(c) 1000×1500

FIGURE 7 First term of the Kronecker product decomposition for three different resolutions

This is illustrated with the $6000 \times 4000 \times 3$ colour image in Figure 5*. A degree-5 TKPSVD is computed with dimensions

$$(250 \times 375 \times 3) \otimes (2 \times 2 \times 1) \otimes (2 \times 2 \times 1) \otimes (2 \times 2 \times 1) \otimes (2 \times 2 \times 1).$$

The total runtime to compute the TKPSVD using the TTr1SVD and HOSVD was 18 and 30 s, respectively. In contrast, computing a standard SVD of one of the three slices $\mathcal{A}(:,:,i)$ takes about 1 min and consists of 4,000 rank-1 terms. The TTr1SVD needs 240 terms, and the HOSVD needs 65,536. For this particular example, the compression rate when retaining k factors and r terms can be approximated by

$$\frac{(250 \times 375 \times 3)\ (2 \times 2 \times 1)^{k-1}}{r(250 \times 375 \times 3 + (k-1)(2 \times 2 \times 1))} \approx \frac{(2 \times 2 \times 1)^{k-1}}{r} = \frac{4^{k-1}}{r}.$$

This implies that the maximal compression rate, when $r = 1$, is dependent on the resolution, namely, the number of KP factors k in each term. At the largest resolution ($k = 5$), the maximal compression rate is approximately 256, but at the smallest resolution, no compression is possible through truncation of KP terms. A common measure to quantify the quality of reconstruction of lossy compressed images is the peak signal-to-noise ratio (PSNR). The PSNR is defined as

$$PSNR = 20\log_{10}(\mathrm{MAX_I}) - 10\log_{10}(\mathrm{MSE}),$$

where $\mathrm{MAX_I}$ is the maximal possible pixel value, 255 in our case, and MSE is the mean squared error $||\mathcal{A} - \hat{\mathcal{A}}||_F^2/(n_1 \cdot n_2 \cdot 3)$. Figure 6 shows the PSNR as a function of the number of retained KP terms, computed from the TTr1SVD for the highest possible resolution. In this case, the PSNR can be completely determined from the σ's in the TKPSVD using Equation 3. Acceptable values of the PSNR are between 30 and 50 dB and are obtained from retaining the first 20 terms, which corresponds with a compression rate of about 12.8. Figure 7 displays one-term approximants for three different resolutions. For Figure 7(a-c), the PSNR is 58dB, 58dB, and 57dB, respectively.

8 | CONCLUSIONS

In this paper, we introduced the tensor KPSVD that decomposes a real k-way tensor $\mathcal{A}$ into a linear combination of tensor KP terms with an arbitrary number of d factors $\mathcal{A} = \sum_{j=1}^{R} \sigma_j\ \mathcal{A}_j^{(d)} \otimes \cdots \otimes \mathcal{A}_j^{(1)}$. This decomposition enables easy computation of a KP approximation and a very straightforward determination of the relative approximation error without explicit construction of the approximant. We proved that for many different structured tensors, the KP factors $\mathcal{A}_j^{(1)}, \ldots, \mathcal{A}_j^{(d)}$ are guaranteed to inherit this structure. In addition, we introduced the new framework of general symmetric tensors, which includes many different structures such as symmetric, persymmetric, centrosymmetric, Toeplitz, and Hankel tensors.

ACKNOWLEDGEMENTS

This work was supported by the Hong Kong Research Grants Council under General Research Fund (GRF) Project 17212315. Also, the authors would like to thank Martijn Boussé and Nico Vervliet for their invaluable help on computing a CPD with orthogonal factor matrices in Tensorlab.

REFERENCES

1. Collar AR. On centrosymmetric and centroskew matrices. Quart J Mech Appl Math. 1962;15(3):265–281.
2. Van Loan CF. The ubiquitous Kronecker product. J Comput Appl Math. 2000;123(1-2):85–100.
3. Hackbusch W, Khoromskij BN, Tyrtyshnikov EE. Hierarchical Kronecker tensor-product approximations. J Numer Math. 2005;13:119–156.
4. Van Loan CF, Pitsianis N. Approximation with Kronecker Products. Linear algebra for large scale and real time applications. Dordrecht, The Netherlands: Kluwer Publications, 1993; p. 293–314.
5. Kamm J, Nagy JG. Optimal Kronecker product approximation of block Toeplitz matrices. SIAM J Matrix Anal Appl. 2000;22(1):155–172.
6. Nagy JG, Ng MK, Perrone L. Kronecker product approximations for image restoration with reflexive boundary conditions. SIAM J Matrix Anal Appl. 2003;25(3):829–841.
7. Nagy JG, Kilmer ME. Kronecker product approximation for preconditioning in three-dimensional imaging applications. IEEE T Image Process. March 2006;15(3):604–613.
8. Rezghi M, Hosseini SM, Eldén L. Best Kronecker product approximation of the blurring operator in three dimensional image restoration problems. SIAM J Matrix Anal Appl. 2014;35(3):1086–1104.

*absfreepic.com/free-photos/download/water-nature-fall-6000x4000_90673.html

9. Carroll JD, Chang J-J. Analysis of individual differences in multidimensional scaling via an n-way generalization of “Eckart-Young” decomposition. Psychometrika. 1970;35(3):283–319.
10. Harshman RA. Foundations of the PARAFAC procedure: Models and conditions for an “explanatory” multi-modal factor analysis. UCLA Working Papers in Phonetics. 1970;16(1):84.
11. De Lathauwer L, De Moor B, Vandewalle J. A multilinear singular value decomposition. SIAM J Matrix Anal Appl. 2000;21(4):1253–1278.
12. Batselier K, Liu H, Wong N. A constructive algorithm for decomposing a tensor into a finite sum of orthonormal rank-1 terms. SIAM J Matrix Anal Appl. 2015;36(3):1315–1337.
13. Regalia PA, Sanjit MK. Kronecker products, unitary matrices and signal processing applications. SIAM Review. 1989;31(4):586–613.
14. Kolda TG, Bader BW. Tensor decompositions and applications. SIAM Rev. 2009;51(3):455–500.
15. Tucker L. Some mathematical notes on three-mode factor analysis. Psychometrika. 1966;31(3):279–311.
16. Tucker LR. The extension of factor analysis to three-dimensional matrices. In: Gulliksen H, Frederiksen N, editors. Contributions to mathematical psychology. New York: Holt, Rinehart and Winston, 1964; p. 110–127.
17. Salmi J, Richter A, Koivunen V. Sequential unfolding svd for low rank orthogonal tensor approximation. 2008 42nd Asilomar Conference on Signals, Systems and Computers, 2008; p. 1713–1717.
18. Phan AH, Cichocki A, Tichavský P, Mandic DP, Matsuoka K. On revealing replicating structures in multiway data: A novel tensor decomposition approach. International Conference on Latent variable analysis and signal separation; March 12-15, 2012; Tel Aviv, Israel. Berlin Heidelberg: Springer; 2012. p. 297–305.
19. Sorber L, Barel MV, de Lathauwer L. Tensorlab version 2.0. 2014. Available from: http://www.tensorlab.net/ [accessed December 2015].
20. Chen H, Chen Z, Qi L. Centrosymmetric, skew centrosymmetric and centrosymmetric cauchy tensors. 2014. ArXiv e-print 1406.7409.
21. Zhao X, Yang Q. The spectral radius of nonnegative centrosymmetric tensor. J High School Numer Math. 2014;36:58–66.

How to cite this article: Batselier K, Wong N. A constructive arbitrary-degree Kronecker product decomposition of tensors. *Numer Linear Algebra Appl*. 2017;24:e2097. https://doi.org/10.1002/nla.2097