TYPE Original Research
PUBLISHED 05 February 2024
DOI 10.3389/fams.2023.1259836

OPEN ACCESS

EDITED BY
Subhas Khajanchi,
Presidency University, India

REVIEWED BY
Debasisha Mishra,
National Institute of Technology Raipur, India
Ida Mascolo,
University of Naples Federico II, Italy

*CORRESPONDENCE
Harry Leib
✉ harry.leib@mcgill.ca

RECEIVED 16 July 2023
ACCEPTED 29 December 2023
PUBLISHED 05 February 2024

CITATION
Pandey D, Venugopal A and Leib H (2024) Linear to multi-linear algebra and systems using tensors. *Front. Appl. Math. Stat.* 9:1259836. doi: 10.3389/fams.2023.1259836

Linear to multi-linear algebra and systems using tensors

Divyanshu Pandey[1], Adithya Venugopal[2] and Harry Leib[3]*

[1]Department of Electrical and Computer Engineering, Rice University, Houston, TX, United States, [2]Independent Researcher, North York, ON, Canada, [3]Department of Electrical and Computer Engineering, McGill University, Montreal, QC, Canada

In the past few decades, multi-linear algebra also known as tensor algebra has been adapted and employed as a tool for various engineering applications. Recent developments in tensor algebra have indicated that several well-known concepts from linear algebra can be extended to a multi-linear setting with the help of a special form of tensor contracted product, known as the Einstein product. Thus, the tensor contracted product and its properties can be harnessed to define the notions of multi-linear system theory where the input, output signals, and the system are inherently multi-domain or multi-modal. This study provides an overview of tensor algebra tools which can be seen as an extension of linear algebra, at the same time highlighting the differences and advantages that the multi-linear setting brings forth. In particular, the notions of tensor inversion, tensor singular value, and tensor eigenvalue decomposition using the Einstein product are explained. In addition, this study also introduces the notion of contracted convolution for both discrete and continuous multi-linear system tensors. Tensor network representation of various tensor operations is also presented. In addition, application of tensor tools in developing transceiver schemes for multi-domain communication systems, with an example of MIMO CDMA system, is presented. This study provides a foundation for professionals whose research involves multi-domain or multi-modal signals and systems.

KEYWORDS

tensors, contracted product, Einstein product, contracted convolution, multi-linear systems

1 Introduction

Tensors are multi-way arrays that are indexed by multiple indices and the number of indices is called the *order* of the tensor [1]. Subsequently, matrices and vectors can be seen as order two and order one tensors respectively. Higher-order tensors are inherently capable of mathematically representing processes and systems with dependency on more than two indices. Hence, tensors are widely employed for several applications in many engineering and science disciplines. Tensors were initially introduced for applications in Physics during the early nineteenth century [2]. Later with the study of Tucker [3], tensors were used in Psychometrics in the 1960s for extending two-way data analysis to higher-order datasets and further in Chemometrics in the 1980s [4, 5]. The last few decades have witnessed a surge in their applications in areas such as data mining [6, 7], computer vision [8, 9], neuroscience [10], machine learning [11], signal processing [12–14], multi-domain communications [15, 16], and controls system theory [17, 18]. When appropriately employed, tensors can help in developing models that capture interactions between various parameters of multi-domain systems. Such tensor-based system representation can enhance the understanding of the mutual effects of various system domains.

Given the wide scope of applications that tensors support, there have been many recent publications summarizing the essential topics in tensor algebra. One such primary reference is Kolda and Bader [1] where the fundamental tensor decompositions such as Tucker, PARAFAC, and their variants are discussed in great detail with applications. Another useful reference is Comon [2] which presents tensors as a mapping from one linear space to another, along with a discussion on tensor ranks. A more signal processing oriented outlook on tensors is considered in Cichocki et al. [12], including applications such as Big Data storage and Compressed sensing. A more recent and exhaustive tutorial style study is Sidiropoulos et al. [11] which presents a detailed overview of up-to-date tensor decomposition algorithms, computations, and applications in machine learning. Similarly, Chen et al. [19] presents such an overview with applications in multiple-input multiple-output (MIMO) wireless communications. In addition, Kisil et al. [20] provides a detailed review of many tensor decompositions with a focus on the needs of Data Analytics community. However, all these studies do not consider in particular the notions of tensor contracted product and contracted convolution, which are the crux of this study. With the help of a specific form of contracted product, known as the Einstein product of tensors, various tensor decompositions and properties can be established which may be viewed as an intuitive and meaningful extension of the corresponding linear algebra concepts.

The most popular and widely used decompositions in the case of matrices are the singular value and eigenvalue decompositions. In order to consider their extensions to higher-order tensors, it is important to note there is no single generalization that preserves all the properties of the matrix case [21, 22]. The most commonly used generalization of the matrix singular value decomposition is known as a higher-order singular value decomposition (HOSVD) which is basically the same as Tucker decomposition for higher-order tensors [23]. Similarly, several definitions exist in the literature for tensor eigenvalues as a generalization of the matrix eigenvalues [24]. More recently, in order to solve a set of multi-linear equations using tensor inversion, a specific notion of tensor singular value decomposition and eigenvalue decomposition was introduced in Brazell et al. [25], which generalizes the matrix SVD and EVD to tensors through a fixed transformation of the tensors into matrices. The authors in Brazell et al. [25] establish the equivalence between the Einstein product of tensors and the matrix product of the transformed tensors, thereby proving that a tensor group endowed with the Einstein product is structurally similar or isomorphic to a general linear group of matrices. The notion of equivalence between the Einstein product of tensors and the corresponding matrix product of the transformed tensors is important and relevant as it helps in developing many tools and concepts from matrix theory such as matrix inverse, ranks, and determinants for tensors. Hence as a follow-up to Brazell et al. [25], several other studies explored different notions of linear algebra which can be extended to multi-linear algebra using the Einstein product [26–32].

The purpose of this study is 2-fold. First, we intend to present an overview of tensor algebra concepts developed in the past decade using the Einstein product. Since there is a natural way of extending linear algebra concepts to tensors, in this paper we present a summary of the most commonly used and relevant concepts which can equip the reader with tools to define and prove other properties more specific to their intended applications. Second, this study introduces the notion of contracted convolutions for both discrete and continuous system tensors. The theory of linear time invariant (LTI) systems has been an indispensable tool in various engineering applications such as communication systems, and controls. Now with the evolution of these subjects to multi-domain communication systems and multi-linear systems theory, there is a need to better understand the classical topics in a multi-domain setting. This study intends to provide such tools through a tutorial style presentation of the subject matter leading to a mechanism to develop more tools needed for research and applications in any multi-domain/multi-dimensional/multi-modal/multi-linear setting.

The organization of this study is as follows: In Section 2, we present basic tensor definitions and operations, including the concept of signal tensors and contracted convolutions. In Section 3, we present the tensor network representation of various tensor operations. Section 4 presents some tensor decompositions based on the Einstein product. Section 5 defines the notions of multi-linear system tensors and discusses their stability in both time and frequency domains. It also includes a detailed discussion on the application of tensors to multi-linear system representation with an example of MIMO CDMA system. The study is concluded in Section 6.

2 Fundamentals of tensors and notation

A tensor is a multi-way array whose elements are indexed by three or more indices. Each index may correspond to a different domain, dimension, or mode of the quantity being represented by the array. The order of the tensor is the number of such indices or domains or dimensions or modes. A vector is often referred to as a tensor of order-1, a matrix as a tensor of order-2 and tensors of order greater than 2 are known as higher-order tensors.

2.1 Notations

In this study, we use lowercase underline fonts to represent vectors, e.g., $\underline{x}$, uppercase fonts to represent matrices, e.g., X and uppercase calligraphic fonts to represent tensors, e.g., $\mathcal{X}$. The individual elements of a tensor are denoted by the indices in subscript, e.g., the (i_1, i_2, i_3)th element of a third-order tensor $\mathcal{X}$ is denoted by $\mathcal{X}_{i_1,i_2,i_3}$. A colon in subscript for a mode corresponds to every element of that mode corresponding to fixed other modes. For instance, $\mathcal{X}_{:,i_2,i_3}$ denotes every element of tensor $\mathcal{X}$ corresponding to i_2th second and i_3th third mode. The nth element in a sequence is denoted by a superscript in parentheses, e.g., $\mathcal{A}^{(n)}$ denotes the n^{th} tensor in a sequence of tensors. We use $\mathbb{C}$ and $\mathbb{C}_k$ to denote a set of complex numbers and a set of complex numbers which are a function of k, respectively.

2.2 Definitions and tensor operations

Definition 1. **Tensor linear space** : The set of all tensors of size $I_1 \times \cdots \times I_K$ over $\mathbb{C}$ forms a linear space, denoted as $\mathbb{T}_{I_1,\ldots,I_K}(\mathbb{C})$. For $\mathcal{A}, \mathcal{B} \in \mathbb{T}_{I_1,\ldots,I_K}(\mathbb{C})$ and $\alpha \in \mathbb{C}$, the sum $\mathcal{A} + \mathcal{B} = \mathcal{C} \in \mathbb{T}_{I_1,\ldots,I_K}(\mathbb{C})$ where $\mathcal{C}_{i_1,\ldots,i_k} = \mathcal{A}_{i_1,\ldots,i_k} + \mathcal{B}_{i_1,\ldots,i_k}$, and scalar multiplication $\alpha \cdot \mathcal{A} = \mathcal{D} \in \mathbb{T}_{I_1,\ldots,I_K}(\mathbb{C})$ where $\mathcal{D}_{i_1,\ldots,i_k} = \alpha \mathcal{A}_{i_1,\ldots,i_k}$ [14].

Definition 2. **Fiber**: Fiber is defined by fixing every index in a tensor but one. A matrix column is a mode-1 fiber, and a matrix row is a mode-2 fiber. Similarly, a third-order tensor has column (mode-1), row (mode-2), and tube (mode-3) fibers [1].

Definition 3. **Slices**: Slices are two-dimensional sections of a tensor defined by fixing all but two indices.

Definition 4. **Norm**: The $p-$norm of an order N tensor $\mathcal{X} \in \mathbb{C}^{I_1 \times I_2 \times \cdots \times I_N}$ is defined as

$$\|\mathcal{X}\|_p = \Big(\sum_{i_1=1}^{I_1} \sum_{i_2=1}^{I_2} \ldots\ldots \sum_{i_N=1}^{I_N} \mid \mathcal{X}_{i_1,i_2,\ldots,i_N} \mid^p \Big)^{1/p}. \tag{1}$$

Subsequently, the $2-$norm or the Frobenius norm of $\mathcal{X}$ is defined as the square root of the sum of the square of absolute values of all its elements:

$$\|\mathcal{X}\|_2 = \sqrt{\sum_{i_1=1}^{I_1} \sum_{i_2=1}^{I_2} \ldots\ldots \sum_{i_N=1}^{I_N} \mid \mathcal{X}_{i_1,i_2,\ldots,i_N} \mid^2}. \tag{2}$$

In addition, the $1-$norm and $\infty-$norm of a tensor are defined as

$$\|\mathcal{X}\|_1 = \sum_{i_1,\ldots,i_N} \mid \mathcal{X}_{i_1,i_2,\ldots,i_N} \mid, \tag{3}$$

$$\|\mathcal{X}\|_\infty = \max_{i_1,\ldots,i_N} \mid \mathcal{X}_{i_1,i_2,\ldots,i_N} \mid . \tag{4}$$

Definition 5. **Kronecker product of Matrices**: The Kronecker product of two matrices A of size $I \times J$ and B of size $K \times L$, denoted by $\mathrm{A} \otimes \mathrm{B}$, is a matrix of size $(IK) \times (JL)$ and is defined as

$$\mathrm{A} \otimes \mathrm{B} = \begin{bmatrix} \mathrm{A}_{1,1}\mathrm{B} & \mathrm{A}_{1,2}\mathrm{B} & \ldots & \mathrm{A}_{1,J}\mathrm{B} \\ \mathrm{A}_{2,1}\mathrm{B} & \mathrm{A}_{2,2}\mathrm{B} & \ldots & \mathrm{A}_{2,J}\mathrm{B} \\ \vdots & \vdots & \ddots & \vdots \\ \mathrm{A}_{I,1}\mathrm{B} & \mathrm{A}_{I,2}\mathrm{B} & \ldots & \mathrm{A}_{I,J}\mathrm{B} \end{bmatrix}. \tag{5}$$

Definition 6. **Matricization transformation** : Let us denote the linear space of $P \times Q$ matrices over $\mathbb{C}$ as $\mathbb{M}_{P,Q}(\mathbb{C})$. For an order $K = N + M$ tensor $\mathcal{A} \in \mathbb{C}^{I_1 \times \cdots \times I_N \times J_1 \times \cdots \times J_M}$, the transformation $f_{I_1,\ldots,I_N|J_1,\ldots,J_M} : \mathbb{T}_{I_1,\ldots,I_N,J_1,\ldots,J_M}(\mathbb{C}) \Rightarrow \mathbb{M}_{I_1 \cdot I_2 \cdots I_{N-1} \cdot I_N, J_1 \cdot J_2 \cdots J_{M-1} \cdot J_M}(\mathbb{C})$ with $f_{I_1,\ldots,I_N|J_1,\ldots,J_M}(\mathcal{A}) = \mathrm{A}$ is defined component-wise as [25]

$$\mathcal{A}_{i_1,i_2,\ldots,i_N,j_1,j_2,\ldots,j_M} \xrightarrow{f_{I_1,\ldots,I_N|J_1,\ldots,J_M}} \mathrm{A}_{i_1+\sum_{k=2}^{N}(i_k-1)\prod_{l=1}^{k-1} I_l, j_1+\sum_{k=2}^{M}(j_k-1)\prod_{l=1}^{k-1} J_l}. \tag{6}$$

This transformation is essentially a matrix unfolding of a tensor by partitioning its indices into two disjoint subsets corresponding to rows and columns [33]. The widely used vectorization operation as defined in [34] is a specific case of Equation (6) where $J_1 = \cdots = J_M = 1$. The bar notation in subscript of $f_{I_1,\ldots,I_N|J_1,\ldots,J_M}$ denotes the partitioning after N modes of an $N + M$ order tensor. The first N modes correspond to the rows, and the last M modes correspond to the columns of the representing matrix. This transformation is bijective [28], and it preserves addition and scalar multiplication operations, i.e., for $\mathcal{A}, \mathcal{B} \in \mathbb{T}_{I_1,\ldots,I_N,J_1,\ldots,J_M}(\mathbb{C})$ and any scalar $\alpha \in \mathbb{C}$, we have $f_{I_1,\ldots,I_N|J_1,\ldots,J_M}(\mathcal{A} + \mathcal{B}) = f_{I_1,\ldots,I_N|J_1,\ldots,J_M}(\mathcal{A}) + f_{I_1,\ldots,I_N|J_1,\ldots,J_M}(\mathcal{B})$ and $f_{I_1,\ldots,I_N|J_1,\ldots,J_M}(\alpha\mathcal{A}) = \alpha f_{I_1,\ldots,I_N|J_1,\ldots,J_M}(\mathcal{A})$. Hence, the linear spaces $\mathbb{T}_{I_1,\ldots,I_N,J_1,\ldots,J_M}(\mathbb{C})$ and $\mathbb{M}_{I_1 \cdot I_2 \cdots I_{N-1} \cdot I_N, J_1 \cdot J_2 \cdots J_{M-1} \cdot J_M}(\mathbb{C})$ are isomorphic, and the transformation $f_{I_1,\ldots,I_N|J_1,\ldots,J_M}$ is an isomorphism between the linear spaces. For a matrix, the transformation (6) does no change when $N = M = 1$, creates a column vector when $N = 2, M = 0$ and a row vector when $N = 0, M = 2$.

2.2.1 Tensor products

Tensors have multiple modes; hence, a product between two tensors can be defined in various ways. In this section, we present definitions of the most commonly used tensor products.

Definition 7. **Tensor Contracted product** [33]: Consider two tensors $\mathcal{X} \in \mathbb{C}^{I_1 \times I_2 \times \cdots \times I_M \times J_1 \times J_2 \times \cdots \times J_N}$ and $\mathcal{Y} \in \mathbb{C}^{I_1 \times I_2 \times \cdots \times I_M \times K_1 \times K_2 \times \cdots \times K_P}$. We can multiply both tensors along their common M modes, and the resulting tensor $\mathcal{Z} \in \mathbb{C}^{J_1 \times J_2 \times \cdots \times J_N \times K_1 \times K_2 \times \cdots \times K_P}$ is given by

$$\mathcal{Z} = \{\mathcal{X}, \mathcal{Y}\}_{\{1,\ldots,M;1,\ldots,M\}} \tag{7}$$

where

$$\mathcal{Z}_{j_1,\ldots,j_N,k_1,\ldots,k_P} = \sum_{i_1=1}^{I_1} \cdots \sum_{i_M=1}^{I_M} \mathcal{X}_{i_1,\ldots,i_M,j_1,\ldots,j_N} \mathcal{Y}_{i_1,\ldots,i_M,k_1,\ldots,k_P}. \tag{8}$$

It is important to note that the modes to be contracted need not be consecutive. However, the size of the corresponding dimensions must be equal. For example, tensors $\mathcal{A} \in \mathbb{C}^{K \times L \times M \times N}$ and $\mathcal{B} \in \mathbb{C}^{K \times M \times Q \times R}$ can be contracted along the first and third mode of $\mathcal{A}$ and first and second mode of $\mathcal{B}$ as $\mathcal{C} = \{\mathcal{A}, \mathcal{B}\}_{\{1,3;1,2\}}$ where $\mathcal{C} \in \mathbb{C}^{L \times N \times Q \times R}$. Matrix multiplication between $\mathrm{A} \in \mathbb{C}^{I \times J}$ and $\mathrm{B} \in \mathbb{C}^{J \times K}$ can be seen as a specific case of the contracted product as $\mathrm{A} \cdot \mathrm{B} = \{\mathrm{A}, \mathrm{B}\}_{\{2;1\}}$ where $\cdot$ represents usual matrix multiplication. Several other tensor products can be defined as specific cases of contracted products. One such commonly used tensor product is the Einstein product where the modes to be contracted are at a fixed location as defined next.

Definition 8. **Einstein product** : The Einstein product between tensors $\mathcal{A} \in \mathbb{C}^{I_1 \times \cdots \times I_P \times K_1 \times \cdots \times K_N}$ and $\mathcal{B} \in \mathbb{C}^{K_1 \times \cdots \times K_N \times J_1 \cdots \times J_M}$ is defined as a contraction between their N common modes, denoted by $*_N$, as [25]:

$$(\mathcal{A} *_N \mathcal{B})_{i_1,\ldots,i_P,j_1,\ldots,j_M} = \sum_{k_1,\ldots,k_N} \mathcal{A}_{i_1,i_2,\ldots,i_P,k_1,\ldots,k_N} \mathcal{B}_{k_1,\ldots k_N,j_1,j_2,\ldots,j_M}. \tag{9}$$

In Einstein product, contraction is over N consecutive modes and can also be written using the more general notation with contracted modes in subscript. For instance, for sixth-order tensors $\mathcal{H}, \mathcal{S} \in \mathbb{C}^{I\times J\times K\times I\times J\times K}$, we have

$$\begin{aligned}(\mathcal{H} *_3 \mathcal{S})_{i,j,k,\hat{i},\hat{j},\hat{k}} &= \sum_{u=1}^{I}\sum_{v=1}^{J}\sum_{w=1}^{K} \mathcal{H}_{i,j,k,u,v,w}\mathcal{S}_{u,v,w,\hat{i},\hat{j},\hat{k}} \\ &= (\{\mathcal{H}, \mathcal{S}\}_{\{4,5,6;1,2,3\}})_{i,j,k,\hat{i},\hat{j},\hat{k}}. \end{aligned} \tag{10}$$

Note that one can define Einstein product for several specific mode orderings. For instance, in [17], Einstein product is defined as contraction over N alternate modes and not consecutive modes. However, that would not change the concepts presented here, so far as we remain consistent with the definition.

Definition 9. **Inner product**: The inner product of two tensors $\mathcal{X} \in \mathbb{C}^{I_1\times I_2\times\cdots\times I_N}$ and $\mathcal{Y} \in \mathbb{C}^{I_1\times I_2\times\cdots\times I_N}$ of the same order N with all the dimensions of same length is given by

$$\langle \mathcal{X}, \mathcal{Y}\rangle = \sum_{i_1=1}^{I_1}\sum_{i_2=1}^{I_2}\cdots\sum_{i_N=1}^{I_N} \mathcal{X}_{i_1,i_2,\ldots,i_N}\mathcal{Y}_{i_1,i_2,\ldots,i_N}. \tag{11}$$

It can also be seen as the Einstein product of tensors where contraction is along all the dimensions, i.e., $\langle \mathcal{X}, \mathcal{Y}\rangle = \mathcal{X} *_N \mathcal{Y} = \mathcal{Y} *_N \mathcal{X}$.

Definition 10. **Outer product**: Consider two tensors $\mathcal{X} \in \mathbb{C}^{I_1\times I_2\times\cdots\times I_N}$ and $\mathcal{Y} \in \mathbb{C}^{J_1\times J_2\times\cdots\times J_M}$ of order N and M, respectively. The outer product between $\mathcal{X}$ and $\mathcal{Y}$ denoted by $\mathcal{X} \circ \mathcal{Y}$ is given by a tensor of size $I_1 \times I_2 \times \cdots \times I_N \times J_1 \times J_2 \times \cdots \times J_M$ with individual elements as

$$(\mathcal{X} \circ \mathcal{Y})_{i_1,i_2,\ldots,i_N,j_1,j_2,\ldots,j_M} = \mathcal{X}_{i_1,i_2,\ldots,i_N}\mathcal{Y}_{j_1,j_2,\ldots,j_M}. \tag{12}$$

It can also be seen as a special case of the Einstein product of tensors in Equation (9) with $N = 0$.

Definition 11. **n-mode product**: 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 $\mathrm{U} \in \mathbb{C}^{J\times I_n}$ is denoted by $\mathcal{A}\times_n\mathrm{U}$ and is defined as [23]:

$$(\mathcal{A}\times_n \mathrm{U})_{i_1,i_2,\ldots,i_{n-1},j,i_{n+1},\ldots,i_N} = \sum_{i_n=1}^{I_n} \mathcal{X}_{i_1,i_2,\ldots,i_N}\mathrm{U}_{j,i_n}. \tag{13}$$

Each mode-n fiber is multiplied by the matrix U. The result of n-mode product is a tensor of the same order but with a new nth mode of size J. The resulting tensor is of the size $I_1 \times I_2 \times \cdots \times I_{n-1} \times J \times I_{n+1} \times \ldots I_N$.

Definition 12. **Square tensors** : A tensor $\mathcal{A} \in \mathbb{C}^{I_1\times\cdots\times I_N\times J_1\times\cdots\times J_M}$ is called a square tensor if $N = M$ and $I_k = J_k$ for $k = 1, \ldots, N$ [28].

For order-4 tensors $\mathcal{A}, \mathcal{B}$ of size $I \times J \times I \times J$, it was shown in Brazell et al. [25] that $f_{I,J|I,J}(\mathcal{A} *_2 \mathcal{B}) = f_{I,J|I,J}(\mathcal{A}) \cdot f_{I,J|I,J}(\mathcal{B})$. This result was further extended to a tensor of any order and size in Wang and Xu [32] as the following lemma:

Lemma 1. For tensors $\mathcal{A} \in \mathbb{C}^{I_1\times\cdots\times I_N\times J_1\times\cdots\times J_M}$ and $\mathcal{B} \in \mathbb{C}^{J_1\times\cdots\times J_M\times K_1\times\cdots\times K_P}$ using the matrix unfolding from Equation (6), we get

$$f_{I_1,\ldots,I_N|K_1,\ldots,K_P}(\mathcal{A} *_M \mathcal{B}) = f_{I_1,\ldots,I_N|J_1,\ldots,J_M}(\mathcal{A}) \cdot f_{J_1,\ldots,J_M|K_1,\ldots,K_P}(\mathcal{B}). \tag{14}$$

Definition 13. **Pseudo-diagonal tensors** : Any tensor $\mathcal{D} \in \mathbb{C}^{I_1\times\cdots\times I_N\times J_1\times\cdots\times J_M}$ of order $N + M$ is called pseudo-diagonal if its transformation $\mathrm{D} = f_{I_1,\ldots,I_N|J_1,\ldots,J_M}(\mathcal{D})$ yields a diagonal matrix such that $\mathrm{D}_{i,j}$ is non-zero only when $i = j$ [35].

Since the mapping (6) is bijective, one can establish that a diagonal matrix $\mathrm{D} \in \mathbb{C}^{I\times J}$ under inverse transformation $f^{-1}_{I_1,\ldots,I_N|J_1,\ldots,J_M}(\mathrm{D})$ will yield a pseudo-diagonal tensor where $I = I_1 \cdots I_N$ and $J = J_1 \cdots J_M$. A square tensor $\mathcal{D} \in \mathbb{C}^{I_1\times\cdots\times I_N\times I_1\times\cdots\times I_N}$ is pseudo-diagonal if all its elements $\mathcal{D}_{i_1,\ldots,i_N,j_1,\ldots,j_N}$ are zero except when $i_1 = j_1, i_2 = j_2, \ldots, i_N = j_N$. Such a tensor is referred to as a diagonal tensor in Brazell et al. [25] and Sun et al. [27] and as a U-diagonal tensor in [17]. However, we define it as pseudo-diagonal in this study, so as to distinguish it from the diagonal tensor definition more widely found in the literature which states that a diagonal tensor is one where elements $\mathcal{D}_{i_1,\ldots,i_N}$ are zero except when $i_1 = i_2 = \cdots = i_N$ [1]. This can be seen as a stricter diagonal rule as non-zero elements exist only when all the modes have the same index whereas in a pseudo-diagonal tensor, say of order $2N$, elements are non-zero when every ith and $(i + N)$th mode have the same index for $i = 1, \ldots, N$. An illustration of order-4 tensor showing the difference between diagonal and pseudo-diagonal structures can be found in Pandey et al. [16]. For a matrix, which has just two modes, the diagonal and pseudo-diagonal structures are the same. The notion of pseudo-diagonality can be defined with respect to partition after any number of modes. For instance, for a third-order pseudo-diagonal tensor, it is important to specify whether the pseudo-diagonalilty is with respect to partition after the first mode or the second mode. For simplicity, in this study wherever we write a pseudo-diagonal tensor explicitly as order $N+M$ or $2N$, the pseudo-diagonality is with respect to partition after first N modes.

Definition 14. **Pseudo-triangular tensor** [15]: A tensor $\mathcal{A} \in \mathbb{C}^{I_1\times\cdots\times I_N\times I_1\times\cdots\times I_N}$ is defined to be pseudo-lower triangular if

$$\mathcal{A}_{i_1,\ldots,i_N,i'_1,\ldots,i'_N} = \begin{cases} 0 & \text{if } (i'_1 + \sum_{k=2}^{N}(i'_k - 1)\prod_{l=1}^{k-1} I_l) \geq (i_1 + \sum_{k=2}^{N}(i_k - 1)\prod_{l=1}^{k-1} I_l) \\ a_{i_1,\ldots,i_N,i'_1,\ldots,i'_N} & \text{otherwise} \end{cases} \tag{15}$$

where $a_{i_1,\ldots,i_N,i'_1,\ldots,i'_N}$ are arbitrary scalars. Similarly, the tensor is said to be pseudo-upper triangular if

$$\mathcal{A}_{i_1,\ldots,i_N,i'_1,\ldots,i'_N} = \begin{cases} 0 & \text{if } (i'_1 + \sum_{k=2}^{N}(i'_k - 1)\prod_{l=1}^{k-1} I_l) \leq (i_1 + \sum_{k=2}^{N}(i_k - 1)\prod_{l=1}^{k-1} I_l) \\ a_{i_1,\ldots,i_N,i'_1,\ldots,i'_N} & \text{otherwise.} \end{cases} \tag{16}$$

An illustration of an upper triangular tensor of size $J_1 \times J_2 \times I_1 \times I_2$ with $I_1 = I_2 = J_1 = J_2 = 3$ is presented in Figure 1 and its pseudo-upper triangular elements highlighted in gray along with its pseudo-diagonal elements shown in black. A similar illustration of a lower triangular tensor can be found in Venugopal and Leib [15]. It can be readily seen that a lower triangular tensor becomes a lower triangular matrix under the tensor-to-matrix transformation defined in Equation (6), and a pseudo-upper triangular tensor becomes an upper triangular matrix.

Definition 15. **Identity tensor** [15]: An identity tensor $\mathcal{I}_N \in \mathbb{C}^{I_1\times\cdots\times I_N\times I_1\times\cdots\times I_N}$ is a pseudo-diagonal tensor of order $2N$ such

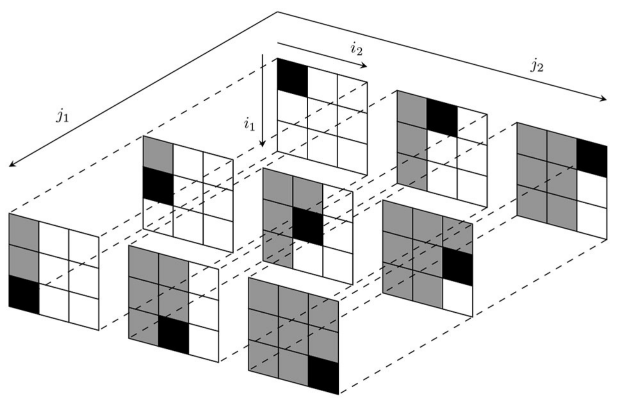

FIGURE 1
Pseudo-upper triangular tensor.

that for any tensor $\mathcal{A} \in \mathbb{C}^{I_1\times\cdots\times I_N\times I_1\times\cdots\times I_N}$, we have $\mathcal{A} *_N \mathcal{I}_N = \mathcal{I}_N *_N \mathcal{A} = \mathcal{A}$ and in which all non-zero entries are 1, i.e.,

$$(\mathcal{I}_N)_{i_1,i_2,\ldots,i_N,j_1,j_2,\ldots,j_N} = \prod_{k=1}^{N} \delta_{i_k j_k}, \quad \text{where } \delta_{pq} = \begin{cases} 1, & p = q \\ 0, & p \neq q. \end{cases} \tag{17}$$

2.2.2 Transpose, Hermitian, and inverse of a tensor

The transpose of a matrix is a permutation of its two indices corresponding to rows and columns. Since elements of a higher-order tensor are indexed by multiple indices, there are several permutations of such indices, and hence, there can be multiple ways to write the transpose or Hermitian of a tensor. Such permutation-dependent transpose of a tensor is defined in Pan [36].

Assume the set $S_N = \{1, 2, \ldots, N\}$ and σ is a permutation of S_N. We denote $\sigma(j) = i_j$ for $j = 1, 2, \ldots, N$ where $\{i_1, i_2, \ldots, i_N\} = \{1, 2, \ldots, N\} = S_N$. Since S_N is a finite set with N elements, it has $N!$ different permutations. Hence, discounting the identity permutation $\sigma(j) = [1, 2, \ldots, N]$, there are $N! - 1$ different transposes for a tensor with N dimensions or modes. For a tensor $\mathcal{A} \in \mathbb{C}^{I_1\times\cdots\times I_N}$, we define its transpose associated with a certain permutation σ as $\mathcal{A}^{T\sigma} \in \mathbb{C}^{I_{\sigma(1)}\times\cdots\times I_{\sigma(N)}}$ with entries

$$\mathcal{A}^{T\sigma}_{i_{\sigma(1)},i_{\sigma(2)},\ldots,i_{\sigma(N)}} = \mathcal{A}_{i_1,i_2,\ldots,i_N}. \tag{18}$$

Similarly, the Hermitian of a tensor $\mathcal{A} \in \mathbb{C}^{I_1\times\cdots\times I_N}$ associated with a permutation σ is defined as the conjugate of its transpose and is denoted as $\mathcal{A}^{H\sigma} = (\mathcal{A}^{T\sigma})^* \in \mathbb{C}^{I_{\sigma(1)}\times\cdots\times I_{\sigma(N)}}$ with entries

$$\mathcal{A}^{H\sigma}_{i_{\sigma(1)},i_{\sigma(2)},\ldots,i_{\sigma(N)}} = (\mathcal{A}^{T\sigma}_{i_{\sigma(1)},i_{\sigma(2)},\ldots,i_{\sigma(N)}})^* = (\mathcal{A}_{i_1,i_2,\ldots,i_N})^*. \tag{19}$$

For example, a transpose of a third-order tensor $\mathcal{X} \in \mathbb{C}^{I_1\times I_2\times I_3}$ such that its third mode is transposed with the first can be written as $\mathcal{X}^{T\sigma}$ where $\sigma = [3, 2, 1]$ with components $\mathcal{X}^{T\sigma}_{i_3,i_2,i_1} = \mathcal{X}_{i_1,i_2,i_3}$. For two tensors $\mathcal{A} \in \mathbb{C}^{I_1\times\cdots\times I_N}$ and $\mathcal{B} \in \mathbb{C}^{I_1\times\cdots\times I_N}$, we have [36]

$$\langle \mathcal{A}, \mathcal{B} \rangle = \langle \mathcal{A}^{T\sigma}, \mathcal{B}^{T\sigma} \rangle. \tag{20}$$

Consider a tensor $\mathcal{Y} \in \mathbb{C}^{I_1\times\cdots\times I_N\times J_1\times\cdots\times J_M}$ and its transposition $\mathcal{Y}^{T\sigma}$ where the last M modes are swapped with the first N modes. Such a permutation can be denoted as $\sigma = [(N+1), \ldots (N+M), 1, \ \ldots \ N]$ where $\mathcal{Y}^{T\sigma}_{j_1,\ldots,j_M,i_1,\ldots i_N} = \mathcal{Y}_{i_1,\ldots,i_N,j_1,\ldots,j_M}$. Since we will use tensors to define system theory elements with fixed order M output and order N input, the most often encountered case of transpose or Hermitian in this study would be after N modes of an $N+M$ or $2N$ tensor, i.e., $\sigma = [(N+1), \ldots (N+M), 1, \ \ldots \ N]$. Henceforth, in such a case we drop the superscript σ for ease of representation and represent such a transpose by $\mathcal{Y}^T$ and its conjugate by $\mathcal{Y}^H$.

Furthermore, a square tensor $\mathcal{U} \in \mathbb{C}^{I_1\times\cdots\times I_N\times I_1\times\cdots\times I_N}$ is called a *unitary* tensor if $\mathcal{U}^H *_N \mathcal{U} = \mathcal{U} *_N \mathcal{U}^H = \mathcal{I}_N$.

The tensor $\mathcal{A}^{-1} \in \mathbb{C}^{I_1\times\cdots\times I_N\times I_1\times\cdots\times I_N}$ is an *inverse* of a square tensor of same size, $\mathcal{A} \in \mathbb{C}^{I_1\times\cdots\times I_N\times I_1\times\cdots\times I_N}$ if $\mathcal{A} *_N \mathcal{A}^{-1} = \mathcal{A}^{-1} *_N \mathcal{A} = \mathcal{I}_N$ [28]. The inverse of a tensor exists if its transformation $f_{I_1,\ldots,I_N|I_1,\ldots,I_N}(\mathcal{A})$ is invertible [25]. Several algorithms using the Einstein product such as Higher-order Bi-conjugate Gradient method [25] or Newton's method [37] can be used to find tensor inverse without relying on actually transforming the tensor into a matrix.

As a generalization of the matrix Moore-Penrose inverse, the *Moore-Penrose inverse* of a tensor $\mathcal{A} \in \mathbb{C}^{I_1\times\cdots\times I_N\times J_1\times\cdots\times J_N}$ is defined as a tensor $\mathcal{A}^+ \in \mathbb{C}^{J_1\times\cdots\times J_N\times I_1\times\cdots\times I_N}$ that satisfies [27, 38]:

$$\begin{aligned}
&\mathcal{A} *_N \mathcal{A}^+ *_N \mathcal{A} = \mathcal{A}, \\
&\mathcal{A}^+ *_N \mathcal{A} *_N \mathcal{A}^+ = \mathcal{A}^+, \\
&(\mathcal{A} *_N \mathcal{A}^+)^H = \mathcal{A} *_N \mathcal{A}^+, \\
&(\mathcal{A}^+ *_N \mathcal{A})^H = \mathcal{A}^+ *_N \mathcal{A}.
\end{aligned}$$

For a tensor $\mathcal{A} \in \mathbb{C}^{I_1\times\cdots\times I_N\times J_1\times\times\cdots\times J_N}$, the Moore-Penrose inverse always exists and is unique [27].

Based on the definition of tensor inverse, Hermitian, and the Einstein product, several tensor algebra relations and properties can be derived. Here, we present a few properties that are often used and can be easily derived:

1. *Associativity*: For tensors $\mathcal{A} \in \mathbb{C}^{I_1\times\cdots\times I_P\times J_1\times\cdots\times J_N}$, $\mathcal{B} \in \mathbb{C}^{J_1\times\cdots\times J_N\times K_1\times\cdots\times K_M}$, and $\mathcal{C} \in \mathbb{C}^{K_1\times\cdots\times K_M\times T_1\times\cdots\times T_Q}$, we have
$$(\mathcal{A} *_N \mathcal{B}) *_M \mathcal{C} = \mathcal{A} *_N (\mathcal{B} *_M \mathcal{C}) \tag{21}$$
2. *Commutativity*: The Einstein product is not commutative in all cases. However, for the specific case where the contraction is taken over all the N modes of one of the tensors, say for tensors $\mathcal{A} \in \mathbb{C}^{I_1\times\cdots\times I_P\times J_1\times\cdots\times J_N}$ and $\mathcal{B} \in \mathbb{C}^{J_1\times\cdots\times J_N}$, we get
$$\mathcal{A} *_N \mathcal{B} = \mathcal{B} *_N \mathcal{A}^T. \tag{22}$$
3. *Distributivity*: For tensors, $\mathcal{A}, \mathcal{B} \in \mathbb{C}^{I_1\times\cdots\times I_P\times J_1\times\cdots\times J_N}$ and $\mathcal{C} \in \mathbb{C}^{J_1\times\cdots\times J_N\times K_1\times\cdots\times K_M}$, we have
$$(\mathcal{A} + \mathcal{B}) *_N \mathcal{C} = (\mathcal{A} *_N \mathcal{C}) + (\mathcal{B} *_N \mathcal{C}). \tag{23}$$
4. For tensors $\mathcal{A} \in \mathbb{C}^{I_1\times\cdots\times I_M\times J_1\times\cdots\times J_N}$ and $\mathcal{B} \in \mathbb{C}^{J_1\times\cdots\times J_N\times K_1\times\cdots\times K_P}$, we have
$$(\mathcal{A} *_N \mathcal{B})^H = \mathcal{B}^H *_N \mathcal{A}^H. \tag{24}$$
5. For square invertible tensors $\mathcal{A}$ and $\mathcal{B} \in \mathbb{C}^{I_1\times\cdots\times I_N\times I_1\times\cdots\times I_N}$, we have
$$(\mathcal{A} *_N \mathcal{B})^{-1} = \mathcal{B}^{-1} *_N \mathcal{A}^{-1}. \tag{25}$$

2.2.3 Function tensors

A *function tensor* $\mathcal{A}(x) \in \mathbb{C}_x^{I_1\times\cdots\times I_N}$ is an order N tensor whose components are functions of x. Using a third-order function tensor as an example, each component of $\mathcal{A}(x)$ is written as $\mathcal{A}_{i,j,k}(x)$. If x takes discrete values, we represent the function tensor using square bracket notation as $\mathcal{A}[x]$.

A generalization of the function tensor would be the *multivariate function tensor* $\mathcal{A}(x_1,\ldots,x_p) \in \mathbb{C}_{x_1,\ldots,x_p}^{I_1\times\cdots\times I_N}$, which is an order N tensor whose components are functions of the continuous variables $x_1,\ldots,x_p$. If the variables take discrete values, we denote the function tensor as $\mathcal{A}[x_1,\cdots,x_p]$. Using the same example of a third-order tensor, each component can be written as $\mathcal{A}_{i,j,k}(x_1,x_2,\ldots,x_p)$.

A linear system is often expressed as $\mathrm{A}\underline{\mathrm{x}} = \underline{\mathrm{b}}$ where $\mathrm{A} \in \mathbb{C}^{M\times N}$ is a matrix operating upon the vector $\underline{\mathrm{x}} \in \mathbb{C}^N$ to produce another vector $\underline{\mathrm{b}} \in \mathbb{C}^M$ [25]. Essentially, the matrix defines a linear operator $\mathcal{L}:\mathbb{C}^N \to \mathbb{C}^M$ between two vector linear spaces $\mathbb{C}^N$ and $\mathbb{C}^M$. A multi-linear system can be thus defined as a linear operator between two tensor linear spaces $\mathbb{C}^{I_1\times\cdots\times I_N}$ and $\mathbb{C}^{J_1\times\cdots\times J_M}$, i.e., $\mathcal{ML}:\mathbb{C}^{I_1\times\cdots\times I_N} \to \mathbb{C}^{J_1\times\cdots\times J_M}$. Multi-linear systems model several phenomena in various science and engineering applications. However, often in literature, a multi-linear system is degenerated into a linear system by mapping the tensor linear space $\mathbb{C}^{I_1\times\cdots\times I_N}$ into a vector linear space $\mathbb{C}^{I_1\cdots I_N}$ through vectorization. The vectorization process allows one to use tools from linear algebra for convenience but also leads to a representation where the distinction between different modes of the system is lost. Thus, possible hidden patterns, structures, and correlations cannot be explicitly identified in the vectorized tensor entities. With the help of tensor contracted product, one can develop signals and system representation without having to rely on vectorization, at the same time extending tools from linear to multi-linear setting intuitively.

2.3 Discrete time signal tensors

A *discrete time signal tensor* $\mathcal{X}[n] \in \mathbb{C}_n^{I_1\times\cdots\times I_N}$ is a function tensor whose components are functions of the sampled time index n. A discrete tensor signal can also be called a tensor sequence indexed by n.

A multi-linear time invariant *discrete system tensor* is an order $N+M$ tensor sequence $\mathcal{H}[k] \in \mathbb{C}_k^{J_1\times\cdots\times J_M\times I_1\times\cdots\times I_N}$ that couples an input tensor sequence $\mathcal{X}[k] \in \mathbb{C}_k^{I_1\times\cdots\times I_N}$ of order N with an output tensor sequence $\mathcal{Y}[k] \in \mathbb{C}_k^{J_1\times\cdots\times J_M}$ of order M through a discrete contracted convolution defined as

$$\mathcal{Y}[k] = \sum_n \{\mathcal{H}[n], \mathcal{X}[k-n]\}_{\{M+1,\ldots,M+N;1,\ldots,N\}}. \tag{26}$$

Most often the ordering of the modes while defining such system tensors is fixed, where the system tensor contracts over all the input modes. Hence for a more compact notation, we can define the contracted convolution using the Einstein product as

$$\mathcal{Y}[k] = \sum_n \mathcal{H}[n] *_N \mathcal{X}[k-n]. \tag{27}$$

In scalar signals and systems notations, a convolution between two functions is often represented using an asterisk ($*$). However, to make a distinction with the Einstein product notation which also uses the asterisk symbol, we denote the contracted convolution using the notation $\bullet_N$, i.e.,

$$\mathcal{Y}[k] = \mathcal{H}[k] \bullet_N \mathcal{X}[k] = \sum_n \mathcal{H}[n] *_N \mathcal{X}[k-n]. \tag{28}$$

The complex frequency domain representation of discrete signal tensors can be given using the z-transform of the signal tensors, as discussed next.

Definition 16. **z-transform of a Discrete Tensor Sequence:** The z-transform of $\mathcal{X}[n] \in \mathbb{C}_n^{I_1\times\ldots I_N}$ denoted by $\breve{\mathcal{X}}(z) = \mathcal{Z}(\mathcal{X}[k]) \in \mathbb{C}_z^{I_1\times\cdots\times I_N}$ is a tensor of the z-transform of its components defined as

$$\breve{\mathcal{X}}(z) = \mathcal{Z}(\mathcal{X}[k]) = \sum_n \mathcal{X}[n]z^{-n} \tag{29}$$

with components $\breve{\mathcal{X}}_{i_1,\ldots,i_N}(z) = \sum_n \mathcal{X}_{i_1,\ldots,i_N}[n]z^{-n}$.

The discrete time Fourier transform denoted by $\bar{\mathcal{X}}(\omega) = \mathcal{F}(\mathcal{X}[k])$ of a tensor sequence can be found by substituting $z = e^{j\omega}$ in its z-transform as

$$\bar{\mathcal{X}}(\omega) = \breve{\mathcal{X}}(z)\Big|_{z=e^{j\omega}} = \sum_n \mathcal{X}[n]z^{-n}\Big|_{z=e^{j\omega}} = \sum_n \mathcal{X}[n]e^{-j\omega n}. \tag{30}$$

Taking the z-transform of Equation (28), we get

$$\begin{aligned}
\breve{\mathcal{Y}}(z) &= \sum_k \mathcal{Y}[k]z^{-k} \\
&= \sum_k \left(\sum_n \mathcal{H}[n] *_N \mathcal{X}[k-n]\right)z^{-k} \\
&= \sum_k \left(\sum_n \mathcal{H}[n] *_N \mathcal{X}[k-n]\right)z^{n-k}z^{-n} \\
&= \sum_n \mathcal{H}[n] *_N \left(\sum_k \mathcal{X}[k-n]z^{n-k}\right)z^{-n} \\
&= \sum_n \mathcal{H}[n]z^{-n} *_N \breve{\mathcal{X}}(z) \\
&= \left(\sum_n \mathcal{H}[n]z^{-n}\right) *_N \breve{\mathcal{X}}(z) \\
&= \breve{\mathcal{H}}(z) *_N \breve{\mathcal{X}}(z).
\end{aligned} \tag{31}$$

which shows that the discrete contracted convolution between two tensors in the time domain as given by Equation (28) leads to the Einstein product between the tensors in the z-domain.

2.4 Continuous time signal tensors

A *continuous time signal tensor* $\mathcal{X}(t) \in \mathbb{C}_t^{I_1\times\cdots\times I_N}$ is a function tensor whose components are functions of the continuous time variable t.

A multi-linear time invariant *continuous system tensor* is an order $N+M$ tensor $\mathcal{H}(t) \in \mathbb{C}_t^{J_1\times\cdots\times J_M\times I_1\times\cdots\times I_N}$ that couples an order N input continuous tensor signal $\mathcal{X}(t) \in \mathbb{C}_t^{I_1\times\cdots\times I_N}$ with an order M output tensor signal $\mathcal{Y}(t) \in \mathbb{C}_t^{J_1\times\cdots\times J_M}$ through a continuous contracted convolution defined as

$$\mathcal{Y}(t) = \int \{\mathcal{H}(u), \mathcal{X}(t-u)\}_{\{M+1,\ldots,M+N;1,\ldots,N\}} du. \tag{32}$$

In cases where the mode sequence is fixed, similar to the discrete case, we can define a more compact notation using the Einstein product as:

$$\mathcal{Y}(t) = \mathcal{H}(t) \bullet_N \mathcal{X}(t) = \int \mathcal{H}(u) *_N \mathcal{X}(t-u) du. \tag{33}$$

The frequency domain representation of continuous signal tensors can be given using the Fourier transform of the signal tensor as defined next.

Definition 17. **Fourier transform**: The Fourier transform of $\mathcal{X}(t) \in \mathbb{C}_t^{I_1\times\cdots\times I_N}$ denoted by $\bar{\mathcal{X}}(\omega) = \mathcal{F}(\mathcal{X}(t))$ is a tensor of the Fourier transform of its components defined as

$$\bar{\mathcal{X}}(\omega) = \mathcal{F}(\mathcal{X}(\omega)) = \int \mathcal{X}(t) e^{-j\omega t} dt \tag{34}$$

with components $\bar{\mathcal{X}}_{i_1,\ldots,i_N}(\omega) = \int \mathcal{X}_{i_1,\ldots,i_N}(t) e^{-i\omega t} dt$.

Using similar line of derivation as for Equation (31), it can be shown that Equation (33) can be written in frequency domain as $\bar{\mathcal{Y}}(\omega) = \bar{\mathcal{H}}(\omega) *_N \bar{\mathcal{X}}(\omega)$.

3 Tensor networks

Tensor network (TN) diagrams are a graphical way of illustrating tensor operations [39]. A TN diagram uses a node to represent a tensor, and each outgoing edge from a node represents a mode of the tensor. As such, a vector can be represented through a node with a single edge, a matrix through a node with double edges, and an order N tensor through a node with N edges. This is illustrated in Figure 2. Any form of tensor contraction can be visually presented through a TN diagram.

3.1 Illustration of contracted products

A contraction between two modes of a tensor is represented in TN by connecting the edges corresponding to the modes that are to be contracted. Hence, the number of free edges represents the order of the resulting tensor. As such any contracted product can be illustrated through a TN diagram, and a few examples are shown in Figure 3. In Figure 3A, the mode-n product of a tensor $\mathcal{A} \in \mathbb{C}^{I_1\times\cdots\times I_N}$ with $\mathrm{U} \in \mathbb{C}^{J\times I_n}$, i.e., $\mathcal{A} \times_n \mathrm{U}$ from Equation (13) is depicted where the nth edge of $\mathcal{A}$ is connected with the second edge of U to represent the contraction of these modes of the same dimension. Figure 3B shows the inner product between two third-order tensors $\mathcal{A}, \mathcal{B} \in \mathbb{C}^{I\times J\times K}$ where all the edges of both the tensors are connected. Since there is no free edge remaining, the result is a scalar. In Figure 3C, a fifth-order tensor $\mathcal{A} \in \mathbb{C}^{I\times J\times K\times L\times M}$ contracts with a fourth-order tensor $\mathcal{B} \in \mathbb{C}^{J\times P\times L\times N}$ along its two common modes as $\{\mathcal{A}, \mathcal{B}\}_{\{2,4;1,3\}}$. The resulting tensor is an order-5 tensor as there are a total of five free edges in the diagram. Finally, Figure 3D shows the Einstein product between tensors from Equation (9) where the common N modes are connected and we have $P+M$ free edges.

3.2 Illustration of contracted convolutions

A contracted convolution is an operation between tensor functions. A function tensor in a TN diagram is represented by a node which is a function of a variable. To represent a function tensor in a TN, we use a rectangular node rather than a circular node. For a given value of the time index k, the contracted convolution from Equation (28) can be depicted using a TN diagram as shown in Figure 4. Note that each $\mathcal{Y}[k]$ is calculated by computing the Einstein product for all the values of n. Hence, the TN diagram contains a sequence of Einstein product representations between $\mathcal{H}[n]$ and $\mathcal{X}[k-n]$. We suggest a compact representation of the contracted convolution similar to the contracted product, where we connect the edges of the function tensor using a dashed line to represent a contracted convolution as shown in Figure 5. This creates a distinction between the two representations. If the corresponding edges are connected via solid lines, it represents a contracted product, and if they are connected via dashed lines, it represents a contracted convolution.

To the best of our knowledge, a diagrammatic representation of contracted convolution operation has not been proposed in the literature yet. However, we suggest this representation as it allows an easy way to illustrate multi-domain systems and capture their interactions visually. Note that various contracted products have already been widely depicted in the literature using TN for ease of illustration. Even though a contracted convolution can be seen as a sum of contracted products, depicting it using contracted products in a TN, such as in Figure 4, would lead to a rather complicated representation. Moreover, when considering multiple systems spanning multiple domains interacting with each other, such complicated illustrations would be challenging to interpret and analyze mathematically. Hence, our choice of illustration for contracted convolution in a TN provides an elegant and interpretable visualization of multi-domain systems interaction. In Section 5.3.2, we consider an example of a multi-domain communication system that can be modeled using contracted convolutions, represented through a TN, to illustrate the ease of system representation using our proposed method.

Very often, TN diagrams are used to represent tensor operations as they provide a better visual understanding and thereby aid in developing algorithms to compute tensor operations by making use of elements from graph theory and data structures. Furthermore, a TN diagram can also be used to illustrate how a tensor is formed from several other component tensors. Hence, most tensor decompositions studied in literature are often represented using a TN. In the next section, we discuss some tensor decompositions.

4 Tensor decompositions

Several tensor decompositions such as the Tucker decomposition, Canonical Polyadic (CP) or the Parallel Factor (PARAFAC) decomposition, Tensor Train decomposition, and many more have been extensively studied in the literature [1, 11, 20]. However, in the past decade with the help of Einstein product and its properties, a generalization of matrix SVD and EVD has been proposed in the literature which has found applications in solving multi-linear system of equations and

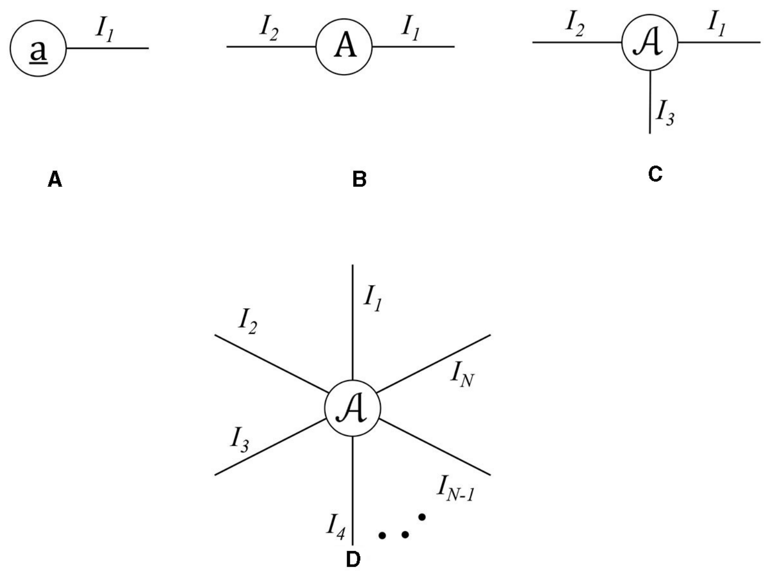

FIGURE 2
TN diagram representation of **(A)** vector of size I_1, **(B)** matrix of size $I_1 \times I_2$, **(C)** order-3 tensor of size $I_1 \times I_2 \times I_3$, and **(D)** order N tensor of size $I_1 \times \cdots \times I_N$.

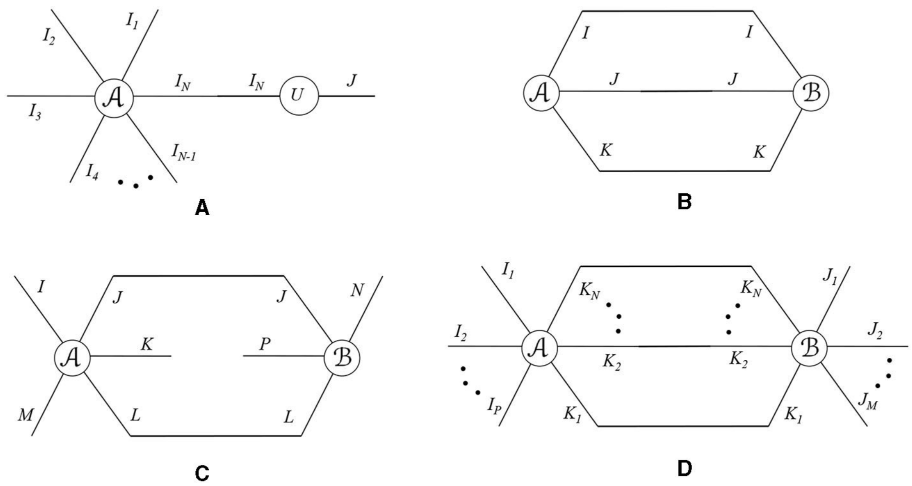

FIGURE 3
TN representation of contracted product **(A)** mode-n product between $\mathcal{A} \in \mathbb{C}^{I_1 \times \cdots \times I_N}$ and $\mathrm{U} \in \mathbb{C}^{J \times I_N}$, **(B)** between tensor $\mathcal{A}, \mathcal{B} \in \mathbb{C}^{I \times J \times K}$ over all the three modes (inner product), **(C)** between $\mathcal{A} \in \mathbb{C}^{I \times J \times K \times L \times M}$ and $\mathcal{B} \in \mathbb{C}^{J \times P \times L \times N}$ as $\{\mathcal{A}, \mathcal{B}\}_{\{2,4;1,3\}}$, and **(D)** Einstein product between tensors $\mathcal{A} \in \mathbb{C}^{I_1 \times \cdots \times I_P \times K_1 \times \cdots \times K_N}$ and $\mathcal{B} \in \mathbb{C}^{K_1 \times \cdots \times K_N \times J_1 \times \cdots \times J_M}$.

systems theory [18, 25, 26]. In this section, we present some such decompositions using the Einstein product of tensors.

4.1 Tensor singular value decomposition (SVD)

Tucker decomposition of a tensor can be seen as a higher-order SVD [23] and has found many applications, particularly in extracting low-rank structures in higher dimensional data [40]. A more specific version of tensor SVD is explored in Brazell et al. [25] as a tool for finding tensor inversion and solving multi-linear systems. Note that Brazell et al. [25] presents SVD for square tensors only. The idea of SVD from Brazell et al. [25] is further generalized for any even order tensor in Sun et al. [27]. However, it can be further extended for any arbitrary order and size of the tensor. We present a tensor SVD theorem here for any tensor of order $N + M$.

Theorem 1. For a tensor, $\mathcal{A} \in \mathbb{C}^{I_1 \times \cdots \times I_N \times J_1 \times \cdots \times J_M}$, the SVD of $\mathcal{A}$ has the form:

$$\mathcal{A} = \mathcal{U} *_N \mathcal{D} *_M \mathcal{V}^H \tag{35}$$

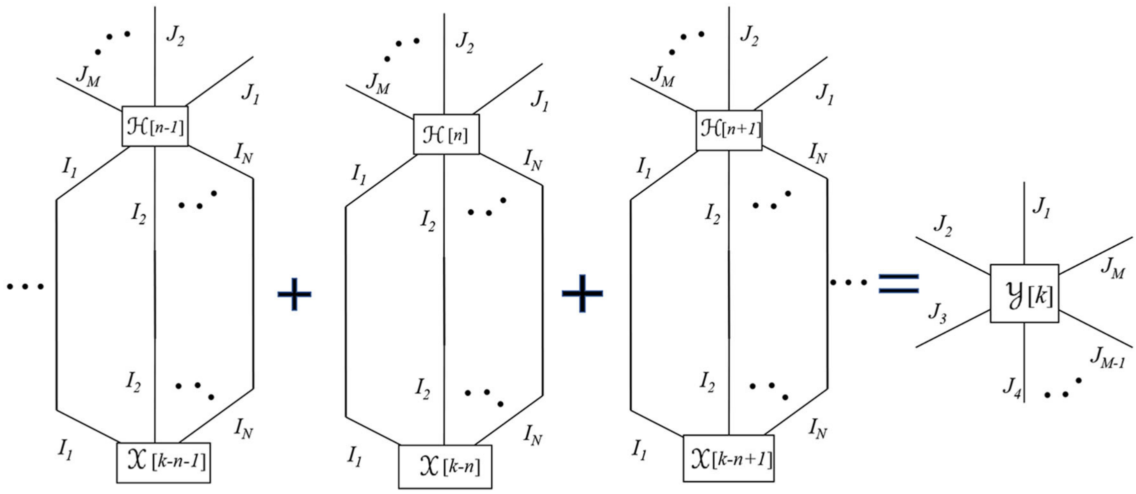

FIGURE 4
TN representation of contracted convolution from Equation (28).

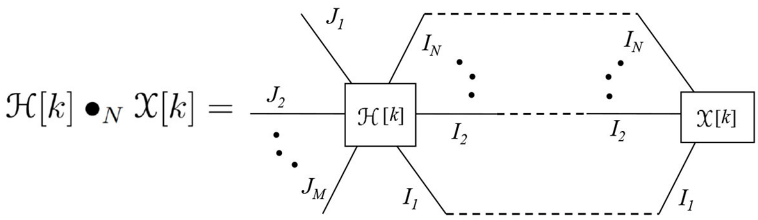

FIGURE 5
Compact TN representation of contracted convolution from Equation (28).

where $\mathcal{U} \in \mathbb{C}^{I_1\times\cdots\times I_N\times I_1\times\cdots\times I_N}$ and $\mathcal{V} \in \mathbb{C}^{J_1\times\cdots\times J_M\times J_1\times\cdots\times J_M}$ are unitary tensors, and $\mathcal{D} \in \mathbb{C}^{I_1\times\cdots\times I_N\times J_1\times\cdots\times J_M}$ is a pseudo-diagonal tensor whose non-zero values are the singular values of $\mathcal{A}$.

Proof. For tensors $\mathcal{A} \in \mathbb{C}^{I_1\times\cdots\times I_N\times J_1\times\cdots\times J_M}$ and $\mathcal{B} \in \mathbb{C}^{J_1\times\cdots\times J_M\times K_1\times\cdots\times K_P}$, from Equation (14), we get

$$\mathcal{A} *_M \mathcal{B} = f^{-1}_{I_1,\ldots,I_N|K_1,\ldots,K_P}[f_{I_1,\ldots,I_N|J_1,\ldots,J_M}(\mathcal{A}) \cdot f_{J_1,\ldots,J_M|K_1,\ldots,K_P}(\mathcal{B})]. \tag{36}$$

If $\mathrm{A} \in \mathbb{C}^{I_1I_2\cdots I_N\times J_1J_2\cdots J_M}$ and $\mathrm{B} \in \mathbb{C}^{J_1J_2\cdots J_M\times K_1K_2\cdots K_P}$ are transformed matrices from $\mathcal{A}$ and $\mathcal{B}$, respectively, then substituting $f_{I_1,\ldots,I_N|J_1,\ldots,J_M}(\mathcal{A}) = \mathrm{A}$ and $f_{J_1,\ldots,J_M|K_1,\ldots,K_P}(\mathcal{B}) = \mathrm{B}$ in Equation (36) gives us

$$f^{-1}_{I_1,\ldots,I_N|K_1,\ldots,K_P}(\mathrm{A}\cdot\mathrm{B}) = \mathcal{A} *_M \mathcal{B} = f^{-1}_{I_1,\ldots,I_N|J_1,\ldots,J_M}(\mathrm{A}) *_M f^{-1}_{J_1,\ldots,J_M|K_1,\ldots,K_P}(\mathrm{B}). \tag{37}$$

Hence if $\mathrm{A} = \mathrm{U}\cdot\mathrm{D}\cdot\mathrm{V}^H$ (obtained from matrix SVD), then based on Equation (37), for an order $N+M$ tensor $\mathcal{A} \in \mathbb{C}^{I_1\times\cdots\times I_N\times J_1\times\cdots\times J_M}$, we have

$$\begin{aligned}\mathcal{A} &= f^{-1}_{I_1,\ldots,I_N|J_1,\ldots,J_M}(\mathrm{A}) = f^{-1}_{I_1,\ldots,I_N|J_1,\ldots,J_M}(\mathrm{U}\cdot\mathrm{D}\cdot\mathrm{V}^H)\\ &= f^{-1}_{I_1,\ldots,I_N|I_1,\ldots,I_N}(\mathrm{U}) *_N f^{-1}_{I_1,\ldots,I_N|J_1,\ldots,J_M}(\mathrm{D}) *_M f^{-1}_{J_1,\ldots,J_M|J_1,\ldots,J_M}(\mathrm{V}^H) =\\ &\mathcal{U} *_N \mathcal{D} *_M \mathcal{V}^H.\end{aligned} \tag{38}$$

Note that for a tensor of order $N+M$, we will get different SVDs for different values of N and M, but for a given N and M, the SVD is unique which depends on the matrix SVD of $f_{I_1,\ldots,I_N|J_1,\ldots,J_M}(\mathcal{A}) = \mathrm{A}$. A proof of this theorem for $2N$ order tensors with $N = M$ using transformation defined in Equation (6) is provided in [25]. This SVD can be seen as a specific case of Tucker decomposition by expressing the unitary tensors in terms of the factor matrices obtained through Tucker decomposition. Let us consider an example of a fourth-order tensor. For a tensor $\mathcal{A} \in \mathbb{C}^{I_1\times I_2\times K_1\times K_2}$, the Tucker decomposition has the form:

$$\mathcal{A} = \mathcal{D}\times_1\mathrm{B}^{(1)}\times_2\mathrm{B}^{(2)}\times_3\mathrm{B}^{(3)}\times_4\mathrm{B}^{(4)} \tag{39}$$

where $\mathrm{B}^{(i)}$ are factor matrices along all four modes of the tensor and $\times_n$ denotes the n-mode product. Now, Equation (39) can be written in matrix form as follows [26]:

$$\mathrm{A} = \underbrace{(\mathrm{B}^{(1)}\otimes\mathrm{B}^{(2)})}_{\mathrm{U}}\cdot\mathrm{D}\cdot\underbrace{(\mathrm{B}^{(3)T}\otimes\mathrm{B}^{(4)T})}_{\mathrm{V}^H}. \tag{40}$$

Now using the transformation from Equation (6), we can map the elements of matrix U to a tensor $\mathcal{U}$ as

$$\mathcal{U}_{i,j,k,l} = \mathrm{U}_{i+(j-1)I_1,k+(l-1)K_1}. \tag{41}$$

Also since $\mathrm{U} = (\mathrm{B}^{(1)} \otimes \mathrm{B}^{(2)})$ and Kronecker product when written element-wise can be expressed as [12]

$$\begin{aligned} \mathrm{U}_{i+(j-1)I_1, k+(l-1)K_1} &= \mathrm{B}^{(1)}_{j,l} \cdot \mathrm{B}^{(2)}_{i,k} \\ \Rightarrow \mathcal{U}_{i,j,k,l} &= \mathrm{B}^{(1)}_{j,l} \cdot \mathrm{B}^{(2)}_{i,k}. \end{aligned} \tag{42}$$

This relation can be seen as the unitary tensor $\mathcal{U}$ being the outer product of matrices $\mathrm{B}^{(1)}$ and $\mathrm{B}^{(2)}$ [25] but with different mode permutation. Similar relation can be established for $\mathcal{V}$ in terms of $\mathrm{B}^{(3)}$ and $\mathrm{B}^{(4)}$.

4.2 Tensor eigenvalue decomposition (EVD)

As a generalization of matrix eigenvalues to tensors, several definitions exist in the literature for tensor eigenvalues [24]. But most of these definitions apply to super-symmetric tensors which are defined as a class of tensors that are invariant under any permutation of their indices [41]. Such an approach has applications in Physics and Mechanics [41], but there is no single generalization to a higher-order tensor case that preserves all the properties of matrix eigenvalues [21]. Here, we present a particular generalization from Liang et al.[28], Cui et al. [26], and Chen et al. [18] which can be seen as the extension of the matrix spectral decomposition theorem and has applications in multi-linear system theory.

Definition 18. Let $\mathcal{A} \in \mathbb{C}^{I_1\times\cdots\times I_N\times I_1\times\cdots\times I_N}$, $\mathcal{X} \in \mathbb{C}^{I_1\times\cdots\times I_N}$, $\lambda \in \mathbb{C}$, where $\mathcal{X}$ and λ satisfy $\mathcal{A} *_N \mathcal{X} = \lambda\mathcal{X}$, then we call $\mathcal{X}$ and λ as *eigentensor* and *eigenvalue* of $\mathcal{A}$, respectively [26].

Using the definition of Hermitian tensor and tensor eigenvalues, the following lemma can be readily established:

Lemma 2. The eigenvalues of a complex Hermitian tensor $\mathcal{A} \in \mathbb{C}^{I_1\times\cdots\times I_N\times I_1\times\cdots\times I_N}$ are always real.

Theorem 2. The EVD of a Hermitian tensor $\mathcal{A} \in \mathbb{C}^{I_1\times\cdots\times I_N\times I_1\times\cdots\times I_N}$ is given as [25]

$$\mathcal{A} = \mathcal{U} *_N \mathcal{D} *_N \mathcal{U}^H \tag{43}$$

where $\mathcal{U} \in \mathbb{C}^{I_1\times\cdots\times I_N\times I_1\times\cdots\times I_N}$ is a unitary tensor, and $\mathcal{D} \in \mathbb{C}^{I_1\times\cdots\times I_N\times I_1\times\cdots\times I_N}$ is a square pseudo-diagonal tensor, i.e., $\mathcal{D}_{i_1,\ldots,i_N,j_1,\ldots,j_N} = 0$ if $(i_1,\ldots,i_N) \neq (j_1,\ldots,j_N)$ with its non-zero values being the eigenvalues of $\mathcal{A}$ and $\mathcal{U}$ containing the eigentensors of $\mathcal{A}$.

This theorem can be proven using Lemma 1, and details are provided in Brazell et al., and Liang et al. [25, 28]. The eigenvalues of $\mathcal{A}$ are same as the eigenvalues of $f_{I_1,\ldots,I_N|I_1,\ldots,I_N}(\mathcal{A})$ [42]. We will refer to a tensor $\mathcal{A} \in \mathbb{C}^{I_1\times\cdots\times I_N\times I_1\times\cdots\times I_N}$ as *positive semi-definite*, denoted by $\mathcal{A} \succeq 0$ if all its eigenvalues are non-negative, which is same as $f_{I_1,\ldots,I_N|I_1,\ldots,I_N}(\mathcal{A})$ being a positive semi-definite matrix. A tensor is *positive definite*, $\mathcal{A} \succ 0$, if all its eigenvalues are strictly greater than zero. A positive semi-definite pseudo-diagonal tensor $\mathcal{D}$ will have all its components non-negative. Its square root can be denoted as $\mathcal{D}^{1/2}$ which is also pseudo-diagonal positive semi-definite whose elements are the square root of elements of $\mathcal{D}$ such that $\mathcal{D}^{1/2} *_N \mathcal{D}^{1/2} = \mathcal{D}$. Similarly, if $\mathcal{D}$ is positive definite, its inverse can be denoted as $\mathcal{D}^{-1}$ which is also pseudo-diagonal whose non-zero elements are the reciprocal of the corresponding elements of $\mathcal{D}$. Based on tensor EVD, we can also write the square root of any Hermitian positive semi-definite tensor as $\mathcal{A}^{1/2} = \mathcal{U} *_N \mathcal{D}^{1/2} *_N \mathcal{U}^H$ and inverse of any Hermitian positive definite tensor as $\mathcal{A}^{-1} = \mathcal{U} *_N \mathcal{D}^{-1} *_N \mathcal{U}^H$. It is straightforward to see that the singular values of a tensor $\mathcal{A} \in \mathbb{C}^{I_1\times\cdots\times I_N\times J_1\times\cdots\times J_M}$ are the square root of the eigenvalues of tensor $\mathcal{A}^H *_N \mathcal{A}$. From SVD, if $\mathcal{A} = \mathcal{U} *_N \mathcal{D} *_M \mathcal{V}^H$, then

$$\begin{aligned} \mathcal{A}^H *_N \mathcal{A} = (\mathcal{V} *_M \mathcal{D}^H *_N \mathcal{U}^H) *_N (\mathcal{U} *_N \mathcal{D} *_M \mathcal{V}^H) = \\ \mathcal{V} *_M \mathcal{D}^H *_N \mathcal{D} *_M \mathcal{V}^H \end{aligned} \tag{44}$$

where $\mathcal{D}^H *_N \mathcal{D}$ is the pseudo-diagonal tensor with eigenvalues of $\mathcal{A}^H *_N \mathcal{A}$ on its pseudo-diagonal which are square of the singular values obtained from the SVD of $\mathcal{A}$.

Definition 19. **Trace**: The trace of a tensor $\mathcal{A} \in \mathbb{C}^{I_1\times\cdots\times I_N\times I_1\times\cdots\times I_N}$ is defined as the sum of its pseudo-diagonal entries:

$$\mathrm{tr}(\mathcal{A}) = \sum_{i_1,\ldots,i_N} \mathcal{A}_{i_1,i_2,\ldots,i_N,i_1,i_2,\ldots,i_N}. \tag{45}$$

Definition 20. **Determinant**: The determinant of a tensor $\mathcal{A} \in \mathbb{C}^{I_1\times\cdots\times I_N\times I_1\times\cdots\times I_N}$ is defined as the product of its eigenvalues, i.e., if $\mathcal{A} = \mathcal{U} *_N \mathcal{D} *_N \mathcal{U}^H$, then

$$\det(\mathcal{A}) = \prod_{i_1,\ldots,i_N} \mathcal{D}_{i_1,i_2,\ldots,i_N,i_1,i_2,\ldots,i_N}. \tag{46}$$

The eigenvalues of $\mathcal{A}$ are the same as that of its matrix transformation, hence $\det(\mathcal{A}) = \det(f_{I_1,\ldots,I_N|I_1,\ldots,I_N}(\mathcal{A}))$. Note that there exist other definitions in the literature for determinants based on how one chooses to define the eigenvalues of tensors [43]. The definition presented here is the same as the unfolding determinant in [28].

4.2.1 Some properties of trace and determinant

The following properties can be easily shown by writing the tensors component-wise or using Lemma 1.

1. For two tensors $\mathcal{A} \in \mathbb{C}^{I_1\times\cdots\times I_N}$ and $\mathcal{B} \in \mathbb{C}^{I_1\times\cdots\times I_N}$ of same size and order N,

$$\mathcal{A} *_N \mathcal{B} = \mathcal{B} *_N \mathcal{A} = \mathrm{tr}(\mathcal{A} \circ \mathcal{B}) = \mathrm{tr}(\mathcal{B} \circ \mathcal{A}). \tag{47}$$

2. For tensors $\mathcal{A} \in \mathbb{C}^{I_1\times\cdots\times I_N\times J_1\times\cdots\times J_M}$ and $\mathcal{B} \in \mathbb{C}^{J_1\times\cdots\times J_M\times I_1\times\cdots\times I_N}$, we have

$$\mathrm{tr}(\mathcal{A} *_M \mathcal{B}) = \mathrm{tr}(\mathcal{B} *_N \mathcal{A}), \tag{48}$$

$$\det(\mathcal{I}_N + \mathcal{A} *_M \mathcal{B}) = \det(\mathcal{I}_M + \mathcal{B} *_N \mathcal{A}). \tag{49}$$

where $\mathcal{I}_N$ and $\mathcal{I}_M$ are identity tensors of order $2N$ and $2M$, respectively. To prove Equation (49), we can use Lemma 1 and Sylvester's matrix determinant identity [44].

3. For tensors $\mathcal{A}, \mathcal{B} \in \mathbb{C}^{I_1\times\cdots\times I_N\times I_1\times\cdots\times I_N}$, we have

$$\det(\mathcal{A} *_N \mathcal{B}) = \det(\mathcal{B} *_N \mathcal{A}) = \det(\mathcal{A}) \cdot \det(\mathcal{B}). \tag{50}$$

4. Trace of a square tensor is the sum of its eigenvalues.
5. The absolute value of the determinant of a unitary tensor is 1, and the determinant of a square pseudo-diagonal tensor is the product of its pseudo-diagonal entries.

4.3 Tensor LU decomposition

LU decomposition is a powerful tool in linear algebra that can be used for solving systems of equations. In order to solve systems of multi-linear equations, Liang et al. [28] proposed an LU decomposition form for tensors. For $\mathcal{A} \in \mathbb{C}^{I_1\times\cdots\times I_N\times I_1\times\cdots\times I_N}$, the LU factorization takes the form:

$$\mathcal{A} = \mathcal{L} *_N \mathcal{U} \tag{51}$$

where $\mathcal{L}, \mathcal{U} \in \mathbb{C}^{I_1\times\cdots\times I_N\times I_1\times\cdots\times I_N}$ are pseudo-lower and pseudo-upper triangular tensors, respectively. In order to solve a system of multi-linear equation $\mathcal{A} *_N \mathcal{X} = \mathcal{B}$ to find $\mathcal{X}$, LU decomposition of $\mathcal{A}$ can be used to break the equation into two pseudo-triangular equations $\mathcal{U} *_N \mathcal{X} = \mathcal{Y}$ and $\mathcal{L} *_N \mathcal{Y} = \mathcal{B}$. These two equations can be solved using forward and backward substitution algorithms proposed in Liang et al. [28]. When $\mathcal{B}$ is an identity tensor, this method can also be used for finding the inverse of a tensor. More details on computing LU decomposition and the required conditions for its existence can be found in Liang et al. [28].

Note that all these tensor decompositions represent a given tensor in terms of a contracted product between factor tensors. Hence, they all can be represented using tensor network diagrams. For example, we show the TN diagram corresponding to tensor SVD in Figure 6. A detailed TN representation of several other tensor decompositions such as Tucker, PARAFAC, and Tensor Train Decomposition is also presented in Cichocki [39].

5 Multi-linear tensor systems

Using the tools presented so far, we will now present the notions of multi-linear system theory using tensors.

5.1 Discrete time multi-linear tensor systems

A discrete time multi-linear tensor system is characterized by an order $N + M$ system tensor $\mathcal{H}[k] \in \mathbb{C}_k^{J_1\times\cdots\times J_M\times I_1\times\cdots\times I_N}$ which produces an order M output tensor sequence $\mathcal{Y}[k] \in \mathbb{C}_k^{J_1\times\cdots\times J_M}$ from an input tensor sequence $\mathcal{X}[k] \in \mathbb{C}_k^{I_1\times\cdots\times I_N}$ through a discrete contracted convolution as defined in Equation (28). The system tensor can be seen as an impulse response tensor whose $(j_1, \ldots, j_M, i_1, \ldots, i_N)$th entry is the impulse response from the $(i_1, \ldots, i_N)$th input to the $(j_1, \ldots, j_M)$th output.

A system tensor is considered $p-$stable if corresponding to every input of finite $p-$norm, the system produces an output which is also finite $p-$norm. When $p \to \infty$, this notion is known as Bounded Input Bounded Output (BIBO) stability. The $\infty-$norm of a signal tensor $\mathcal{X}$ is essentially its peak amplitude evaluated over all the tensor components and all times, i.e.,

$$\|\mathcal{X}\|_\infty = \sup_k \|\mathcal{X}[k]\|_\infty = \sup_k \max_{i_1,\ldots,i_N} \mid \mathcal{X}_{i_1,i_2,\ldots,i_N}[k] \mid . \tag{52}$$

Theorem 3. For a discrete multi-linear time invariant system with order N input and order M output with an order $N + M$ impulse response system tensor $\mathcal{H}[k] \in \mathbb{C}_k^{J_1\times\cdots\times J_M\times I_1\times\cdots\times I_N}$, is BIBO stable if and only if

$$\max_{j_1,\ldots,j_M} \sum_{i_1,\ldots,i_N} \sum_k \mid \mathcal{H}_{j_1,\ldots,j_M,i_1,\ldots,i_N}[k] \mid < \infty. \tag{53}$$

Proof. If the input signal tensor $\mathcal{X}[k]$ satisfies $\|\mathcal{X}\|_\infty < \infty$, then output is given via discrete contracted convolution as

$$\mathcal{Y}[k] = \sum_l \mathcal{H}[k-l] *_N \mathcal{X}[l] \tag{54}$$

and thus,

$$\begin{aligned}
&\max_{j_1,\ldots,j_M} \mid \mathcal{Y}_{j_1,\ldots,j_M}[k] \mid = \\
&\max_{j_1,\ldots,j_M} \mid \sum_l \sum_{i_1,\ldots,i_N} \mathcal{H}_{j_1,\ldots,j_M,i_1,\ldots,i_N}[k-l] \mathcal{X}_{i_1,\ldots,i_N}[l] \mid & (55)\\
&\le \Big(\max_{j_1,\ldots,j_M} \sum_l \sum_{i_1,\ldots,i_N} \mid \mathcal{H}_{j_1,\ldots,j_M,i_1,\ldots,i_N}[k-l] \mid \Big) & (56)\\
&\max_{i_1,\ldots,i_N} \sup_l \mid \mathcal{X}_{i_1,\ldots,i_N}[l] \mid \\
&= \Big(\max_{j_1,\ldots,j_M} \sum_l \sum_{i_1,\ldots,i_N} \mid \mathcal{H}_{j_1,\ldots,j_M,i_1,\ldots,i_N}[k-l] \mid \Big) \|\mathcal{X}\|_\infty . & (57)
\end{aligned}$$

Hence we get,

$$\begin{aligned}
\|\mathcal{Y}\|_\infty &= \sup_k \|\mathcal{Y}[k]\|_\infty = \sup_k \max_{j_1,\ldots,j_M} \mid \mathcal{Y}_{j_1,\ldots,j_M}[k] \mid & (58)\\
&\le \Big(\max_{j_1,\ldots,j_M} \sum_{i_1,\ldots,i_N} \sum_k \mid \mathcal{H}_{j_1,\ldots,j_M,i_1,\ldots,i_N}[k] \mid \Big) \|\mathcal{X}\|_\infty < \infty. & (59)
\end{aligned}$$

which proves that output is bounded if (53) is satisfied. To prove the converse of the theorem, it suffices to show any example where if (53) is not satisfied, there exists a bounded input which leads to an unbounded output. For this, we can simply consider the case where input and output are scalars which is a special case of the tensor formulation. Equation (53) in that case translates to BIBO condition for SISO LTI system, i.e., $\sum_k |h[k]| < \infty$. Hence, it can be readily verified that a signum input defined as $x[n] = sgn(h[-n])$ which is bounded will lead to an unbounded output if the impulse response sequence is not absolutely summable [45].

□

The BIBO stability condition for a MIMO LTI system requires that every element of the impulse response matrix must be absolutely summable. The condition from Equation (53) can be seen as an extension to the tensor case, where every element of the impulse response tensor must be absolutely summable.

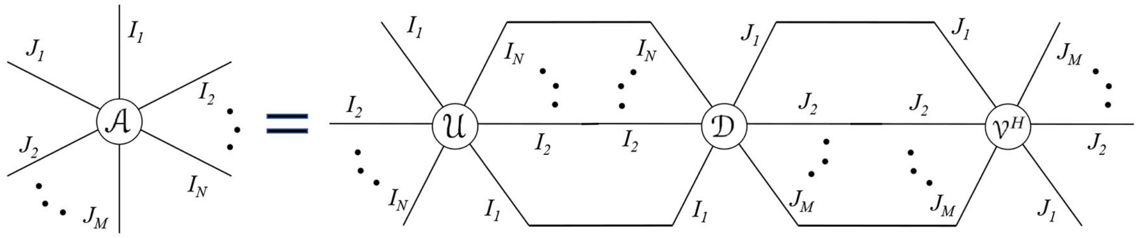

FIGURE 6
TN representation of tensor SVD from Equation (35).

Furthermore, we extend the definitions of poles and zeros from matrix-based systems to tensors. A matrix transfer function $\breve{\mathrm{H}}(z)$ has a pole at frequency ν if some entry of $\breve{\mathrm{H}}(z)$ has a pole at $z = \nu$ [46]. In addition, $\breve{\mathrm{H}}(z)$ has a zero at frequency γ if the rank of $\breve{\mathrm{H}}(z)$ drops at $z = \gamma$ [46]. Similarly, a tensor transfer function $\breve{\mathcal{H}}(z)$ has a pole at frequency ν if some entry of $\breve{\mathcal{H}}(z)$ has a pole at $z = \nu$. In addition, $\breve{\mathcal{H}}(z)$ has a zero at frequency γ if the rank of $f_{J_1,\ldots,J_M|I_1,\ldots,I_N}(\breve{\mathcal{H}}(z))$ drops at $z = \gamma$. Such a rank is also sometimes referred to as the unfolding rank of the tensor [18, 28].

A tensor system is BIBO stable if all its components are BIBO stable. This implies that every pole of every entry of its transfer function has a magnitude less than 1, i.e., all the poles lie within the unit circle on the z-plane.

5.2 Continuous time multi-linear tensor systems

A continuous time multi-linear tensor system is characterized by an order $N + M$ system tensor $\mathcal{H}(t) \in \mathbb{C}_t^{J_1\times\cdots\times J_M\times I_1\times\cdots\times I_N}$ which produces an order M output tensor signal $\mathcal{Y}(t) \in \mathbb{C}_t^{J_1\times\cdots\times J_M}$ from an input tensor signal $\mathcal{X}(t) \in \mathbb{C}_t^{I_1\times\cdots\times I_N}$ through a contracted convolution as defined in Equation (33). The system tensor can be seen as an impulse response tensor whose $(j_1,\ldots,j_M,i_1,\ldots,i_N)$th entry is the impulse response from the $(i_1,\ldots,i_N)$th input to the $(j_1,\ldots,j_M)$th output.

Similar to the discrete case, a continuous system tensor is considered $p-$stable if corresponding to every input of finite $p-$norm, the system produces an output which is also finite $p-$norm. This notion is known as Bounded Input Bounded Output (BIBO) stability if $p = \infty$. The $\infty-$norm of a continuous signal tensor $\mathcal{X}$ is essentially its peak amplitude evaluated over all the tensor components and all times, i.e.,

$$\|\mathcal{X}\|_\infty = \sup_t \|\mathcal{X}(t)\|_\infty = \sup_t \max_{i_1,\ldots,i_N} \mid \mathcal{X}_{i_1,i_2,\ldots,i_N}(t) \mid . \tag{60}$$

Theorem 4. For a continuous multi-linear time invariant system with order N input and order M output with an order $M + N$ impulse response system tensor $\mathcal{H}(t) \in \mathbb{C}_t^{J_1\times\cdots\times J_M\times I_1\times\cdots\times I_N}$, is BIBO stable if and only if

$$\max_{j_1,\ldots,j_M} \sum_{i_1,\ldots,i_N} \int \mid \mathcal{H}_{j_1,\ldots,j_M,i_1,\ldots,i_N}(t) \mid dt < \infty. \tag{61}$$

The condition from Equation (61) implies that every element of the impulse response tensor must be absolutely integrable. The proof of Theorem 4 follows the same line of proof as of Theorem 3. Furthermore, a continuous system tensor with transfer function $\breve{\mathcal{H}}(\omega)$ is BIBO stable if all its components are BIBO stable. This implies that every pole of every entry of its transfer function has a real part less than 0.

5.3 Applications of multi-linear tensor systems

A tensor multi-linear (TML) system can be used to model and represent processes where two tensor signals are coupled through a multi-linear functional. Among various other applications, the use of tensors is ubiquitous in modern communication systems where the signals and systems involved have an inherent multi-domain structure. The physical layer model of modern communication systems invariably spans more than one domain of transmission and reception such as space, time, and frequency to name a few. Consequently, the associated signal processing at the transmitter and receiver has to be cognizant of the multiple domains and their mutual effect on each other for efficient resource utilization. Hence, the signals and the systems involved are best represented using tensors.

5.3.1 System model for multi-domain communication systems

The domains of transmission and reception in modern communication systems depend on specific system configuration. A few examples of possible domains include time slots, sub-carriers, antennas, code sequences, propagation delays, and users. Thus, a generic communication system model which is agnostic to the physical interpretation of the domains, using contracted convolution, was proposed in Venugopal and Leib [15]. Furthermore, a discrete version of the model from Venugopal and Leib [15] using the Einstein product has been proposed in Pandey and Leib [35] where the capacity analysis of higher-order tensor channels is presented.

Consider a multi-domain communication system where the input signal is an order N tensor $\mathcal{X}(t) \in \mathbb{C}_t^{I_1\times\cdots\times I_N}$, which passes through an order $M + N$ multi-linear channel, $\mathcal{H}(t) \in \mathbb{C}_t^{J_1\times\cdots\times J_M\times I_1\times\cdots\times I_N}$. The received signal, or the channel output, in this case can be defined as an order M tensor obtained from the contracted convolution between the channel and the input as Venugopal and Leib [15]:

$$\mathcal{Y}(t) = \mathcal{H}(t) \bullet_N \mathcal{X}(t) + \mathcal{N}(t) \quad (62)$$

where $\mathcal{N}(t)$ denotes the additive noise tensor of the same size as $\mathcal{Y}(t)$. Based on the relation between time and frequency domain as discussed in Section 2.4, the frequency domain system model can be specified using the Einstein product as

$$\bar{\mathcal{Y}}(\omega) = \bar{\mathcal{H}}(\omega) *_N \bar{\mathcal{X}}(\omega) + \bar{\mathcal{N}}(\omega). \quad (63)$$

Note that such a system model is domain-agnostic and can be used to model several systems. For more illustrative examples, we direct readers to [15] which shows that several systems such as Multi-Input Mulitple-Output (MIMO) Orthogonal Frequency Division Multiplexing (OFDM), Generalized Frequency Division Multiplexing (GFDM), and Frequency Bank Multi-carrier (FBMC) can be modeled using Equation (62). A more detailed discussion on the multi-domain communication system modeling can also be found in Pandey et al.[16] and Pandey and Leib [35]. In this study, we present an example of MIMO Code Division Multiple Access (CDMA) system in Section 5.3.3 modeled using tensor contraction. However, first we show the TN representation of higher-order channels in the next section.

5.3.2 TN representation of TML systems

The channel is expressed as a TML system, and its coupling with the input in both frequency and time domain can be represented in a TN diagram as shown in Figure 7. Note that each edge of the input is connected with the common edge of the channel via a dashed line in the time domain and via a solid line in the frequency domain. Instead of a regular block diagram representation of such systems, the TN diagram has the advantage that it graphically details all the modes of the input and the channel. Thus, just by looking at the free edges of the overall TN diagram, one can determine the modes of the output. The linearity of the system is reflected in the fact that any given edge of the channel is connected with a single edge of the input. Thus, a TML system is easy to identify visually in a TN diagram by observing the presence of one-on-one edge connections between the system and the input.

Note that in a communication system, the input signal is often precoded before transmission by a transmit filter and the output signal is processed via a receive filter. The transmit and receiver filters can also be considered system tensors. Thus, the TML channel $\mathcal{H}(t)$ can be seen as a cascade of three system tensors. Let the transmit filter be represented by $\mathcal{H}_T(t) \in \mathbb{C}_t^{K_1\times\cdots\times K_Q\times I_1\times\cdots\times I_N}$ which transforms the order N input into an order Q transmit signal. The physical channel between the source and destination is modeled as an order $P + Q$ tensor $\mathcal{H}_C(t) \in \mathbb{C}_t^{L_1\times\cdots\times L_P\times K_1\times\cdots\times K_Q}$, and the receive filter is represented by $\mathcal{H}_R(t) \in \mathbb{C}_t^{J_1\times\cdots\times J_M\times L_1\times\cdots\times L_P}$. In this case, the equivalent channel $\mathcal{H}(t)$ is obtained via a cascade of the three system tensors as

$$\mathcal{H}(t) = \mathcal{H}_R(t) \bullet_P \mathcal{H}_C(t) \bullet_Q \mathcal{H}_T(t). \quad (64)$$

A detailed derivation of such a channel representation can be found in Venugopal and Leib [15]. A cascade of TML systems is conveniently represented in a TN diagram as shown in Figure 8 which illustrates the coupling of the receive filter, physical channel, and transmit filter system tensors in both time and frequency domains. Hence, the nodes for $\mathcal{H}(t)$ and $\bar{\mathcal{H}}(\omega)$ in Figure 7 can be broken down into component system tensors from Figure 8. A tensor system has multiple modes, and the contraction can be along various combinations of such modes. Hence, the TN representation becomes extremely useful as opposed to regular block diagrams, since it allows depiction of the state and coupling of each mode.

5.3.3 Example of Tensor Contraction for MIMO CDMA systems

Code Division Multiple Access (CDMA) is a spread spectrum technique used in communication systems where multiple transmitting users can send information simultaneously over a single communication channel, thereby enabling multiple access. Each user employs the entire bandwidth along with a distinct pseudo-random spreading code to transmit information which is used to distinguish the users at the receiver. More details on CDMA can be found in Proakis and Salehi [47].

Consider an uplink scenario where K users are transmitting information to a single base station (BS). Assume a simple additive white Gaussian Noise (AWGN) channel. Each user is assigned a distinct spreading sequence denoted by vector $\underline{s}^{(k)} \in \mathbb{C}^L$ of length L which transmits a symbol $x^{(k)}$ for user k. The received signal at the BS can be written as [48]:

$$\underline{y} = \sum_{k=1}^{K} x^{(k)} \underline{s}^{(k)} + \underline{z}, \quad (65)$$

where $\underline{z} \in \mathbb{C}^L$ represents the noise vector. Now consider the extension of such a system model in the presence of flat fading channel and multiple antennas. Assume K users each with N_T transmit antennas are transmitting simultaneously to a BS with N_R receive antennas. To allow multiple access, all the transmit antennas of all different users are assigned different spreading sequences of length L. Let $\underline{s}^{(k,i)} \in \mathbb{C}^L$ denotes the length L spreading vector for the data transmitted by the ith antenna of the kth user, $x^{(k,i)}$. Transmit symbols are assumed to have zero mean and energy $E_s = \mathbb{E}[|x^{(k,i)}|^2]$, and the transmit vector from each user and each antenna is generated as $x^{(k,i)}\underline{s}^{(k,i)}$. The MIMO communication channel between user k and the BS is defined as a matrix $\mathrm{H}^{(k)} \in \mathbb{C}^{N_R\times N_T}$ where the random channel matrix has independent and identically distributed (i.i.d.) zero mean circular symmetric complex Gaussian entries with variance $1/N_R$. The distribution is denoted as $\mathcal{CN}(0, 1/N_R)$. The received signal $\mathrm{Y} \in \mathbb{C}^{N_R\times L}$ can be written as [48, 49]:

$$\mathrm{Y} = \sum_{k=1}^{K} \mathrm{H}^{(k)}\mathrm{X}^{(k)}\mathrm{S}^{(k)} + \mathrm{Z}, \quad (66)$$

where $\mathrm{X}^{(k)}$ is an $N_T \times N_T$ diagonal matrix defined as $diag(x^{(k,1)}, x^{(k,2)}, \ldots, x^{(k,N_T)})$ and $\mathrm{S}^{(k)}$ is an $N_T \times L$ matrix defined as $(\underline{s}^{(k,1)T}, \underline{s}^{(k,2)T}, \ldots, \underline{s}^{(k,N_T)T})^T$. Also Z represents $N_R \times L$ noise matrix with i.i.d. components distributed as $\mathcal{CN}(0, N_0)$. In [48], a per-user matched filter receiver is considered for such a system by assuming the interference from other users as noise. It is shown in [48] that

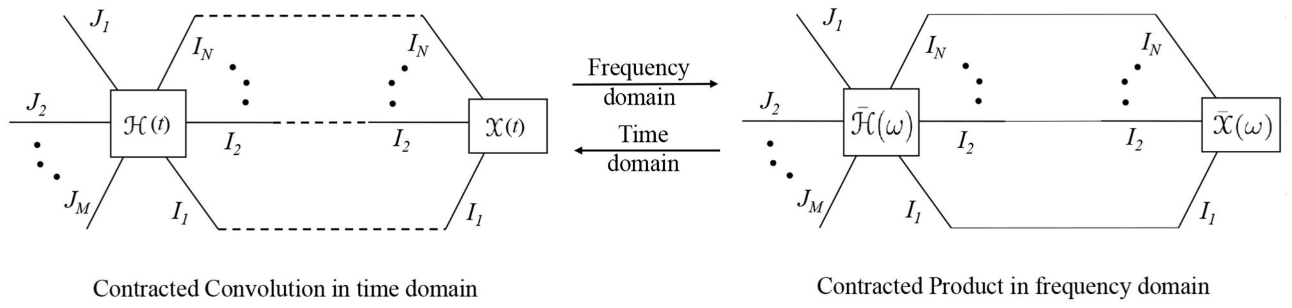

FIGURE 7
TN representation of TML system in time and frequency domain.

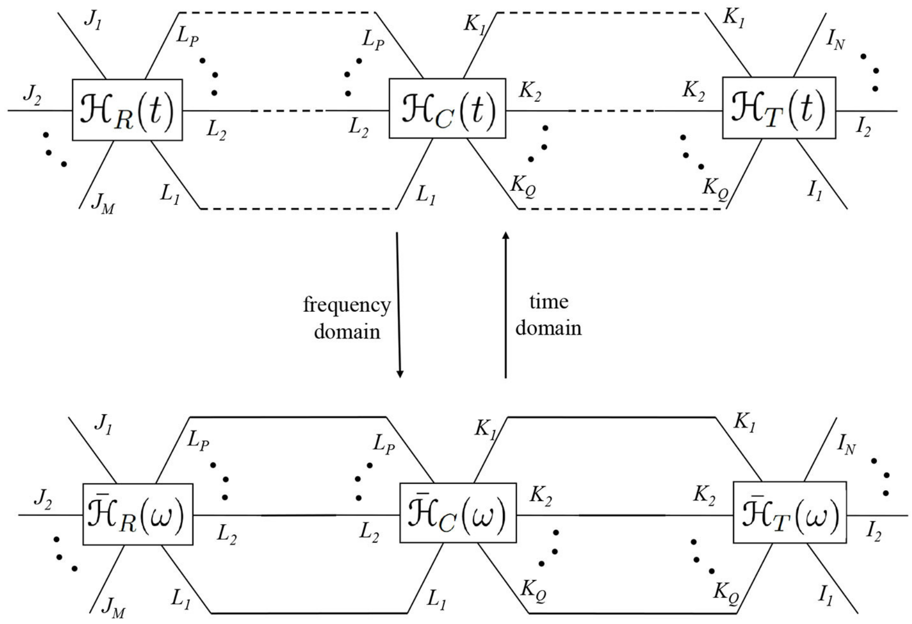

FIGURE 8
TN representation of equivalent multi-linear channel in time and frequency domain.

such a receiver underperforms as compared to a multi-user receiver which detects the transmit symbols for all the users together. Hence, several multi-user receivers are presented in Nordio and Taricco [48] by rewriting the system model from Equation (66) as

$$\mathrm{Y} = \mathrm{H}\bar{\mathrm{X}}\mathrm{S} + \mathrm{Z}, \tag{67}$$

where $\mathrm{H} = (\mathrm{H}^{(1)}, \ldots, \mathrm{H}^{(K)}) \in \mathbb{C}^{N_R \times K \cdot N_T}$, $\bar{\mathrm{X}} = diag(x^{(1,1)}, \ldots, x^{(1,N_T)}, \ldots, x^{(K,1)}, \ldots, x^{(K,N_T)}) \in \mathbb{C}^{K \cdot N_T \times K \cdot N_T}$, and $\mathrm{S} = (\mathrm{S}^{(1)T}, \ldots, \mathrm{S}^{(K)T})^T \in \mathbb{C}^{K \cdot N_T \times L}$. Based on this, a multi-user receiver that aims to mitigate the effects of H (spatial interference) and S (multiple access interference) is considered. The received signal is linearly processed in two stages as $\mathrm{Y} \rightarrow \mathrm{AY} \rightarrow \mathrm{AYB}$, and the transmit signal is decoded as [48]:

$$\hat{x}^{(k,i)} = \arg\max_{x \in \mathcal{D}} |(\mathrm{AYB})_{j,j} - x|^2, \quad \text{where } j = (k-1)N_T + i. \tag{68}$$

The set $\mathcal{D}$ denotes the set of symbols in the transmit constellation map. Essentially AYB represents an estimated version of matrix $\bar{\mathrm{X}}$, whose diagonal elements at index j are used to decode the transmitted symbols and map them back to index (k, i). The matrices A and B separately aim to mitigate the effects of spatial interference and multiple access interference on the received signal and are defined as

$$\mathrm{A} \triangleq \begin{cases} (\mathrm{H}^H\mathrm{H})^{-1}\mathrm{H}^H, & \text{ZF.} \\ (\mathrm{H}^H\mathrm{H} + \dfrac{N_0}{E_s}\mathrm{I}_{KN_T})^{-1}\mathrm{H}^H, & \text{LMMSE.} \end{cases} \tag{69}$$

and

$$\mathrm{B} \triangleq \begin{cases} \mathrm{S}^H(\mathrm{S}\mathrm{S}^H)^{-1}, & \text{DECOR.} \\ \mathrm{S}^H(\mathrm{S}\mathrm{S}^H + \dfrac{N_0}{E_s}\mathrm{I}_{KN_T})^{-1}, & \text{LMMSE.} \end{cases} \tag{70}$$

where I_{KN_T} is an identity matrix of size $K \cdot N_T \times K \cdot N_T$. The zero forcing (ZF) receiver ignores the impact of noise and only tries to counter the effect of the channel, while the linear minimum mean square error (LMMSE) receiver tries to reduce the noise while simultaneously aiming to mitigate the effect of channel. The DECOR choice represents a multi-user decorrelator receiver. At high SNR value, i.e., as $N_0/E_s \rightarrow 0$, the LMMSE option reduces to ZF and DECOR.

Such a receiver based on jointly processing all the users gives better performance than a per-user receiver [48]. However, it still has a drawback in that it tries to combat spatial interference and multiple access interference separately in two stages. Moreover, while the input in Equation (67) is represented as a matrix $\bar{\mathrm{X}}$, only its diagonal contains the transmit elements which is formed from the concatenation of various $x^{(k,i)}$. Thus, such a system model does not fully exploit the multi-linearity of the system and tries to force a linear structure by manipulating the entities involved in order to fit the vector-based well-known LMMSE, ZF, or DECOR solutions. In fact, the tensor framework can be ideally used to represent such a system model while keeping the natural structure of the system intact and developing a tensor multi-linear (TML) receiver.

Since the input symbol $x^{(k,i)}$ is indexed by two indices k and i, it is natural to represent the input as a matrix X of size $K \times N_T$ with elements $\mathrm{X}_{k,i} = x^{(k,i)}$. Furthermore, the input signal is transmitted as a vector $\underline{\mathrm{x}}^{(k,i)}$ of length L corresponding to each user index k and antenna index i. Hence, the transmitted signal through the channel can be represented as a third-order tensor $\mathcal{X}$ of size $K \times N_T \times L$ where $\mathcal{X}_{k,i,l} = \underline{\mathrm{x}}_l^{(k,i)}$. To generate $\mathcal{X}$ from X, we define the spreading sequences as an order-5 tensor $\mathcal{S} \in \mathbb{C}^{K\times N_T\times L\times K\times N_T}$ with elements $\mathcal{S}_{k,i,l,k',i'} = \underline{\mathrm{s}}_l^{(k,i)}$ when $k = k', i = i'$, for all l, and 0 elsewhere. Then, we have $\mathcal{X} = \mathcal{S} *_2 \mathrm{X}$. Note that here we assume that the elements of X are mapped one to one with a spreading sequence; hence, the entries of $\mathcal{S}$ corresponding to $k \neq k', i \neq i'$ are zero. In certain applications, a linear combination of the input symbols might be transmitted, in which case the structure of $\mathcal{S}$ which represents a transmit filtering operation will change accordingly. The channel matrices $\mathrm{H}^{(k)}$ corresponding to each user can be represented as a slice in a third-order tensor $\mathcal{H} \in \mathbb{C}^{N_R\times K\times N_T}$ where $\mathcal{H}_{:,k,:} = \mathrm{H}^{(k)}$. Thus, the system model can be given as

$$\mathrm{Y} = \underbrace{\mathcal{H} *_2 \mathcal{S}}_{\bar{\mathcal{H}}} *_2 \mathrm{X} + \mathrm{Z} \tag{71}$$

where $\bar{\mathcal{H}} \in \mathbb{C}^{N_R\times L\times K\times N_T}$ represents the equivalent fourth-order TML channel between the order two input X and order two output Y.

Note the advantage of modeling the system model through (71) is that all the associated entities retain their natural structure, and a joint TML receiver can be designed to combat the effect of all the interferences of all the users simultaneously. A multi-linear minimum mean square error receiver that acts across all the domains simultaneously can be represented through a tensor $\mathcal{R} \in \mathbb{C}^{K\times N_T\times N_R\times L}$ which produces an estimate of the input X by acting upon the received tensor Y as $\tilde{\mathrm{X}} = \mathcal{R} *_2 \mathrm{Y}$. Thus, each element of the estimated input at the receiver is a linear combination of all the elements of Y, where the coefficients of the linear combinations are encapsulated in $\mathcal{R}$. An optimal choice of $\mathcal{R}$ which minimizes the mean square error between X and $\tilde{\mathrm{X}}$, defined as $\mathbb{E}[||\mathrm{X} - \tilde{\mathrm{X}}||_2^2]$, is given as [14]:

$$\mathcal{R} = \bar{\mathcal{H}}^H *_2 (\bar{\mathcal{H}} *_2 \bar{\mathcal{H}}^H + \frac{N_0}{E_s}\mathcal{I})^{-1}, \tag{72}$$

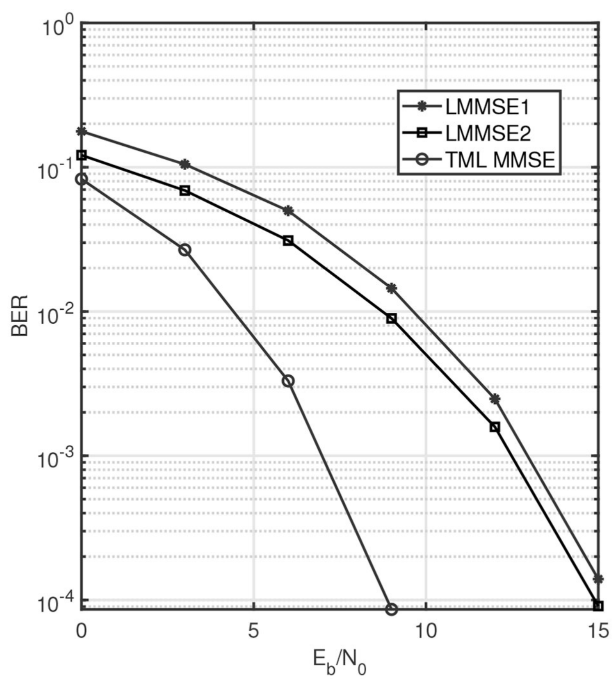

FIGURE 9
BER performance for different receivers against SNR for $L = 32, K = 4, N_R = 32$, and $N_T = 4$.

where $\mathcal{I}$ is an identity tensor of size $K \times N_T \times K \times N_T$. The estimated symbol $\tilde{\mathrm{X}}$ can be used to detect the transmit symbols as

$$\hat{x}^{(k,i)} = \underset{x\in\mathcal{D}}{\arg\max} |\tilde{\mathrm{X}}_{k,i} - x|^2. \tag{73}$$

We will refer to such a receiver as a TML MMSE receiver. Since a TML MMSE receiver jointly acts upon symbols across all domains, it aids in detecting the transmit symbol by exploiting the multi-domain interference terms. Through simulation results, we compare the performance of TML MMSE receiver with Equation (68). In Equation (68), we assume A to be the LMMSE matrix from Equation (69), and for B, we simulate both the DECOR and LMMSE matrices from Equation (70). Hence, we simulate LMMSE-DECOR and LMMSE-LMMSE cases from [48]. We will refer to the former as LMMSE1 and the latter as LMMSE2 for our discussion going forward. The simulation parameters used are the same as in Nordio and Taricco [48] where entries of $\mathrm{H}^{(k)}$ are i.i.d. which are $\mathcal{CN}(0, 1/N_R)$. It is assumed that the channel realizations are known at the receiver. We assume uncoded transmission with 4QAM modulation where symbols are normalized to have unit energy, i.e., $E_s = 1$. The spreading sequences are generated with i.i.d. symbols equiprobable over the set $\{\pm L^{-1/2}, \pm jL^{-1/2}\}$. We use bit error rate (BER) and normalized mean square error (NMSE) as performance measures. All the results are plotted against E_b/N_0 in dB where E_b is the energy per bit defined as $E_s/2$. Thus, E_b/N_0 represents the received SNR per bit. We perform Monte Carlo simulations where the results are averaged over 100 different channel realizations, and at least 100-bit errors were collected for each SNR to calculate BER. The mean square error is normalized with respect to the

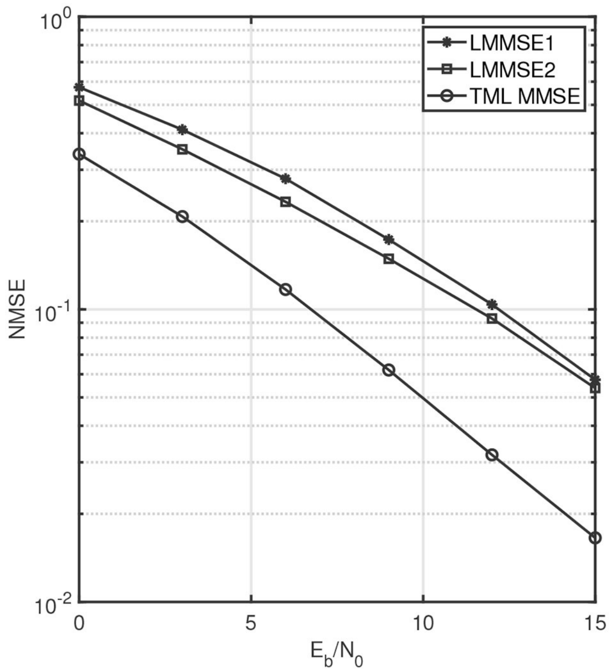

FIGURE 10
Normalized MSE for different receivers against SNR for $L = 32, K = 4, N_T = 32$, and $N_T = 4$.

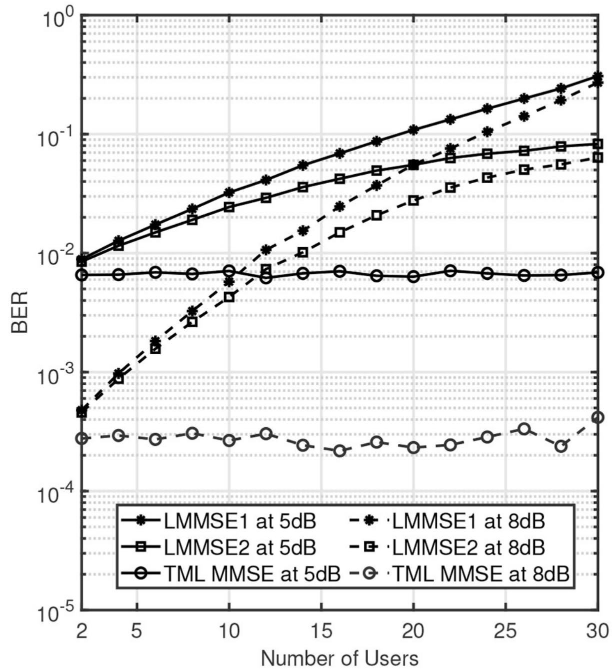

FIGURE 11
BER performance for different receivers against number of users.

number of elements in X and is thus defined as NMSE $= ||\text{X} - \tilde{\text{X}}||_F^2/(N_T \cdot K)$. To compare the performance difference between TML MMSE, LMMSE1, and LMMSE2 against SNR, we plot the BER and NMSE for the three receivers in Figures 9, 10, respectively. We take $L = 32, K = 4, N_T = 4, N_R = 32$. It can be clearly seen in both the figures that the BER and the NMSE decrease as SNR increases. In particular, the TML MMSE leads to a much lower BER and NMSE compared to the other two receivers as it exploits the multi-linearity of the equivalent channel to jointly combat interference across all the domains. Within LMMSE1 and LMMSE2, it can be observed that LMMSE2 performs better as the choice of DECOR for B from Equation (70) is sub-optimal as compared to LMMSE.

Furthermore, the advantage of TML MMSE can be clearly seen when BER and NMSE are observed for a fixed SNR per bit and a variable number of users. Consider $L = 64, N_R = 64, N_T = 2$ and number of users K is variable. Figures 11, 12 present BER and NMSE performance against K for two fixed values of SNR per bit. The solid lines correspond to a 5dB SNR per bit, and dashed lines correspond to an 8dB SNR per bit. It can be clearly seen that for a fixed SNR per bit, the BER and NMSE curves for TML MMSE case remain almost flat as the number of users increases. On the other hand, the performance of LMMSE1 and LMMSE2 significantly degrades with an increase in the number of users. As the number of users increases, the interference across domains also increases which is only efficiently utilized in the TML MMSE receiver.

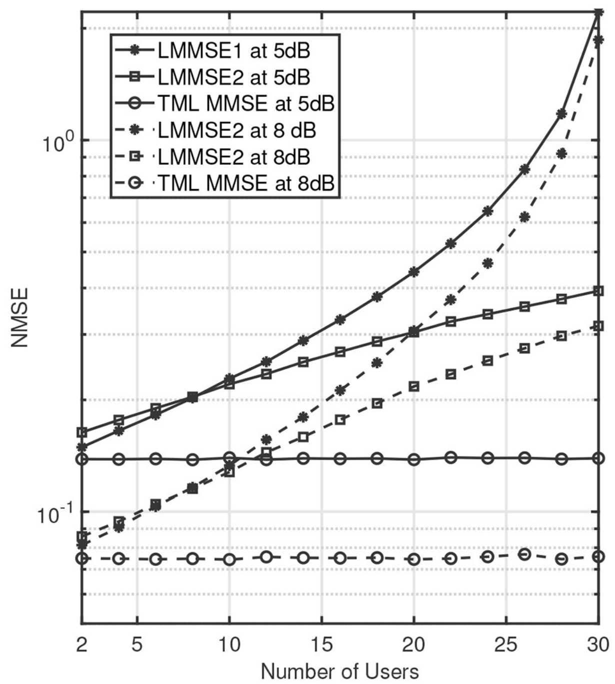

FIGURE 12
Normalized MSE for different receivers against number of users.

Note that Equation (71) can be re-written as a system of linear equations by using vectorization of the input, output, and noise, and considering the channel as a concatenated matrix $f_{N_R,L|K,N_T}(\bar{\mathcal{H}})$. Subsequently, a joint receiver can also be designed using the transformed matrix channel, as presented in Nordio and Taricco [48], which is conceptually equivalent to the TML MMSE approach presented in this study. However, the concatenation of various domains obscures the different domain representations (indices) in the system. Such an approach makes it difficult to incorporate domain-specific constraints at the transceiver. For instance, a common transmit constraint is the power budget which in most practical cases would be different for different users. Thus, designing a transmission scheme with per-user power constraints becomes important and can be achieved using the tensor framework as it maintains the identifiability of domains. Such a consideration has been presented in Pandey et al. [16] and Pandey and Leib [35].

5.3.4 Other examples

In Venugopal and Leib [15], several multi-domain communication systems such as MIMO OFDM, GFDM, and FBMC are represented using the tensor contracted convolution and contracted product. In addition, Venugopal and Leib [15] develops tensor-based receiver equalization methods which are used to combat interference in communication systems using the notion of tensor inversion. A tensor-based receiver and precoder are presented in Pandey and Leib [50] for a MIMO GFDM system where the channel is represented as a sixth-order tensor. The tensor EVD presented in this study is used to design transmit coding operations and perform an information theoretic analysis of the tensor channel [35] leading to the notion of multi-domain water filling power allocation method. In addition, the discrete multi-linear contracted convolution is used to design tensor partial response signaling (TPRS) systems to shape the spectrum and cross-spectrum of transmit signals [16]. The tensor inversion method can also be used to develop estimation techniques for various signal processing applications such as in big data or wireless communications as shown in Pandey and Leib [14]. Another example of the use of tensor Einstein product is for image restoration and reconstruction applications where the objective is to retrieve an image affected by noise, a focal-field distribution, and aperture function [26]. The image data are stored as a three-dimensional tensor, and an order-6 tensor acts as a channel obtained from the point spread function [26], such that output is given using the Einstein product between the input and channel. Another area where the Einstein product properties have been used is the multi-linear dynamical system theory [17, 18]. In Chen et al. [17], a generalized multi-linear time invariant system theory is developed using the Einstein product which can be applied to dynamical systems such as the human genome, social networks, and cognitive sciences. The notion of tensor eigenvalue decomposition presented in this study is used in Chen et al. [17] to derive conditions for stability, reachability, and observability for dynamical systems. The Einstein product has this distinct advantage that it lets us develop tensor algebra notions similar to linear algebra at the same time without disturbing or reshaping the structure of tensors. In addition, the more general tensor contracted product and contracted convolution can be used to model multi-domain systems with any mode ordering as well.

5.3.5 Discussion

The tensor algebra concepts presented in this study provide a structured, intuitive, and mathematically sound framework to characterize and analyze multi-linear systems. Traditionally, matrix-based methods have been used for this purpose by ignoring the signal variability across multiple domains and thereby converting the inherently higher-domain signals and systems into a single concatenated domain with no physical interpretation. Such obfuscated representations are heavily motivated by the ease of employing well-known tools from linear algebra. Several software packages and tools still employ matrix numerical methods for computation. Moreover, since matrix algebra is a standard topic in most engineering programs, engineers are better equipped to handle vectors and matrices than tensors. One primary objective of this study is to provide an easy transition from matrix algebra to tensor algebra for engineering students and researchers. With a proper understanding of the tools from multi-linear algebra, it is straightforward to see the inherent advantages in the tensor formalism. The tensor framework retains the natural structure of the signals and systems involved, thereby capturing the mutual effect of various domains in the system model. The proposed solutions developed with such a framework remain domain-aware and are interpretable.

In addition, sometimes matrix-based representations are offered as a low-complexity solution to tensor-based methods. For instance, in a multi-carrier communication system, a per-subcarrier receive processing using matrix methods is much more computationally efficient as opposed to a tensor-based receiver acting jointly across all users. However, it should be noted that a matrix-based method in itself would not reduce the computational complexity unless some additional assumptions on inter-domain interferences are assumed. For instance, a per-subcarrier receiver assumes zero inter-carrier interference. Without such an assumption, even the degenerate matrix-based formulation of the problem would lead to a large matrix-based model with similar complexity as the equivalent tensor model but with additional loss of distinction between domains. The tensor method provides a structured manner to incorporate all the inter-domain interferences in the system design without resorting to any restructuring. Moreover, the restructuring of data may not be possible in certain applications. With growing big data and IoT applications, even large vector-based data are often stored using tensorization for reducing its storage complexity [39] through tensor train decompositions. Thus, the tensor structure plays a crucial role, and all the mathematical operations for data analysis are expected to be performed while keeping the structure intact. This could be done by resorting to tools from tensor algebra as discussed in this study.

6 Summary and concluding remarks

This study presented a review of tensor algebra concepts developed using the contracted product, more specifically the Einstein product, extending the common notions in linear algebra to a multi-linear setting. In particular, the notion of tensor inverse, singular, and eigenvalue decompositions, LU decomposition were discussed. We also studied the tensor network representations of tensor contractions and convolutions. The notions of time invariant discrete and continuous multi-linear systems which can be defined using the contracted convolutions were also presented. We presented an application in a multi-domain communication system where the channel is modeled as a multi-linear system. The multi-linearity of the channel allowed us to develop a receiver that jointly combats interference across all the domains, thereby giving much better BER and MSE performance as compared to linear receivers which act on a specific domain at a time. The tensor algebra notions discussed in this study have extensive applications in various fields such as communications, signals and systems, controls, and image processing, to name a few. In the presence of several other tensor tutorial studies in literature, this study by no means intends to summarize all the multi-linear algebra concepts but provides an introduction to the main concepts from a signals and systems perspective in a tensor setting.

Author contributions

DP: Investigation, Methodology, Software, Writing—original draft, Writing—review & editing. AV: Investigation, Methodology, Writing—original draft. HL: Conceptualization, Funding acquisition, Investigation, Resources, Supervision, Writing—review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) for project titled "Tensor modulation for space-time-frequency communication systems" under Grant RGPIN-2016-03647.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

1. Kolda TG, Bader BW. Tensor decompositions and applications. *SIAM Review*. (2009) 51:455–500. doi: 10.1137/07070111X

2. Comon P. Tensors: a brief introduction. *IEEE Signal Process Mag*. (2014) 31:44–53. doi: 10.1109/MSP.2014.2298533

3. 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. 110–27.

4. Appellof CJ, Davidson ER. Strategies for analyzing data from video fluorometric monitoring of liquid chromatographic effluents. *Anal Chem*. (1981) 53:2053–6. doi: 10.1021/ac00236a025

5. Bro R. Review on multiway analysis in Chemistry 2000-2005. *Crit Rev Anal Chem*. (2006) 36:279–93. doi: 10.1080/10408340600969965

6. Li X, Ng MK, Ye Y. MultiComm: finding community structure in multi-dimensional networks. *IEEE Trans Knowl Data Eng*. (2014) 26:929–41. doi: 10.1109/TKDE.2013.48

7. Papalexakis EE, Faloutsos C, Sidiropoulos ND. Tensors for data mining and data fusion: Models, applications, and scalable algorithms. *ACM Trans Intell Syst*. (2017) 8:1–44. doi: 10.1145/2915921

8. Shashua A, Hazan T. Non-negative Tensor Factorization with Applications to Statistics and Computer Vision. In: *Proceedings of the 22nd International Conference on Machine Learning*. New York, NY: ACM (2005). p. 792–799.

9. Guerrero JJ, Murillo AC, Sags C. Localization and matching using the planar trifocal tensor with bearing-only data. *IEEE Transact Robot*. (2008) 24:494–501. doi: 10.1109/TRO.2008.918043

10. Latchoumane CFV, Vialatte F, Sol-Casals J, Maurice M, Wimalaratna S, Hudson N. et al. Multiway array decomposition analysis of EEGs in Alzheimer's disease. *J Neurosci Methods*. (2012) 207:41–50. doi: 10.1016/j.jneumeth.2012.03.005

11. Sidiropoulos ND, Lathauwer LD, Fu X, Huang K, Papalexakis EE, Faloutsos C. Tensor decomposition for signal processing and machine learning. *IEEE Transact Signal Process*. (2017) 65:3551–82. doi: 10.1109/TSP.2017.2690524

12. Cichocki A, Mandic D, De Lathauwer L, Zhou G, Zhao Q, Caiafa C, et al. Tensor Decompositions for Signal Processing Applications: From two-way to multiway component analysis. *IEEE Signal Process Mag*. (2015) 32:145–63. doi: 10.1109/MSP.2013.2297439

13. Sidiropoulos ND, Bro R, Giannakis GB. Parallel factor analysis in sensor array processing. *IEEE Transact Signal Process*. (2000) 48:2377–88. doi: 10.1109/78.852018

14. Pandey D, Leib H, A. Tensor framework for multi-linear complex MMSE estimation. *IEEE Open J Signal Process*. (2021) 2:336–58. doi: 10.1109/OJSP.2021.3084541

15. Venugopal A, Leib H, A. Tensor based framework for multi-domain communication systems. *IEEE Open J Commun Soc*. (2020) 1:606–33. doi: 10.1109/OJCOMS.2020.2987543

16. Pandey D, Venugopal A, Leib H. Multi-domain communication systems and networks: a tensor-based approach. *MDPI Network*. (2021) 1:50–74. doi: 10.3390/network1020005

17. Chen C, Surana A, Bloch A, Rajapakse I. Multilinear time invariant system theory. In: *2019 Proceedings of the Conference on Control and its Applications*. Society for Industrial and Applied Mathematics (SIAM) (2019). p. 118–125.

18. Chen C, Surana A, Bloch AM, Rajapakse I. Multilinear control systems theory. *SIAM J Control Optim*. (2021) 59:749–76. doi: 10.1137/19M1262589

19. Chen H, Ahmad F, Vorobyov S, Porikli F. Tensor decompositions in wireless communications and MIMO radar. *IEEE J Sel Top Signal Process*. (2021) 15:438–53. doi: 10.1109/JSTSP.2021.3061937

20. Kisil I, Calvi GG, Dees BS, Mandic DP. *Tensor Decompositions and Practical Applications: A Hands-on Tutorial. In: Recent Trends in Learning From Data*. Cham: Springer (2020). p. 69–97.

21. Lim LH. Singular values and Eigenvalues of tensors: A variational approach. In: *1st IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing*. Puerto Vallarta: IEEE (2005). p. 129–132.

22. Weiland S, Van Belzen F. Singular Value Decompositions and low rank approximations of tensors. *IEEE Trans Signal Process*. (2010) 58:1171. doi: 10.1109/TSP.2009.2034308

23. Lathauwer LD, Moor BD, Vandewalle J. A multilinear singular value decomposition. *SIAM J Matrix Anal Appl*. (2000) 21:1253–78. doi: 10.1137/S0895479896305696

24. Qi L. Eigenvalues of a real supersymmetric tensor. *J Symbolic Comp*. (2005) 40:1302–24. doi: 10.1016/j.jsc.2005.05.007

25. Brazell M, Li N, Navasca C, Tamon C. Solving multilinear systems via tensor inversion. *SIAM J Matrix Analy Appl*. (2013) 34:542–70. doi: 10.1137/100804577

26. Cui LB, Chen C, Li W, Ng MK. An Eigenvalue problem for even order tensors with its applications. *Linear Multili Algebra*. (2016) 64:602–21. doi: 10.1080/03081087.2015.1071311

27. Sun L, Zheng B, Bu C, Wei Y. Moore-Penrose inverse of tensors via Einstein product. *Linear Multili Algebra*. (2016) 64:686–98. doi: 10.1080/03081087.2015.1083933

28. lin Liang M, Zheng B, juan Zhao R. Tensor inversion and its application to the tensor equations with Einstein product. *Linear Multili Algebra*. (2019) 67:843–70. doi: 10.1080/03081087.2018.1500993

29. Huang B, Li W. Numerical subspace algorithms for solving the tensor equations involving Einstein product. *Num Linear Algebra Appl*. (2021) 28:e2351. doi: 10.1002/nla.2351

30. Huang B. Numerical study on Moore-Penrose inverse of tensors via Einstein product. *Num Algor*. (2021) 87:1767–97. doi: 10.1007/s11075-021-01074-0

31. Huang B, Ma C. An iterative algorithm to solve the generalized Sylvester tensor equations. *Linear Multili Algebra*. (2018) 2023:1–26. doi: 10.1080/03081087.2023.2176416

32. Wang QW, Xu X. Iterative algorithms for solving some tensor equations. *Linear Multilinear Algebra*. (2019) 67:1325–49. doi: 10.1080/03081087.2018.1452889

33. Bader BW, Kolda TG. Algorithm 862: MATLAB tensor classes for fast algorithm prototyping. *ACM Trans Mathemat Softw*. (2006) 32:635–53. doi: 10.1145/1186785.1186794

34. De Lathauwer L, Castaing J, Cardoso JF. Fourth-order cumulant-based blind identification of underdetermined mixtures. *IEEE Trans Signal Process*. (2007) 55:2965–73. doi: 10.1109/TSP.2007.893943

35. Pandey D, Leib H. The tensor multi-linear channel and its Shannon capacity. *IEEE Access*. (2022) 10:34907–44. doi: 10.1109/ACCESS.2022.3160187

36. Pan R. Tensor transpose and its properties. *arXiv [Preprint]. arXiv*:1411.1503. (2014). Available online at: https://arxiv.org/abs/1411.1503

37. Pandey D, Leib H. Tensor multi-linear MMSE estimation using the Einstein product. In: *Advances in Information and Communication (FICC 2021)*. Cham: Springer International Publishing (2021). p. 47–64.

38. Panigrahy K, Mishra D. Extension of Moore Penrose inverse of tensor via Einstein product. *Linear Multilinear Algebra*. (2020) 0:1–24. doi: 10.1080/03081087.2020.1748848

39. Cichocki A. Era of big data processing: A new approach via tensor networks and tensor decompositions. *arXiv [Preprint]. arXiv*:1403.2048. (2014). Available online at: https://arxiv.org/abs/1403.2048

40. Zhang A, Xia D. Tensor SVD: Statistical and computational limits. *IEEE Trans Inform Theory*. (2018) 11: 7311–38. doi: 10.1109/TIT.2018.2841377

41. Qi L, Chen H, Chen Y. Tensor Eigenvalues and their applications. Cham: Springer (2018).

42. Luo Z, Qi L, Toint PL. Bernstein concentration inequalities for tensors via Einstein products. *arXiv [Preprint]. arXiv*:1902.03056. (2019). Available online at: https://arxiv.org/abs/1902.03056

43. Hu S, Huang ZH, Ling C, Qi L. On Determinants and Eigenvalue theory of tensors. *J Symbolic Comp*. (2013) 50:508–31. doi: 10.1016/j.jsc.2012.10.001

44. Petersen KB, Pedersen MS. *The Matrix Cookbook*. (2012). Available online at: http://www2.compute.dtu.dk/pubdb/pubs/3274-full.html

45. Mitra SK, Kuo Y. Digital signal processing: a computer-based approach. New York: McGraw-Hill New York. (2006).

46. Dahleh M, Dahleh MA, Verghese G. *Chapter 27: Poles and Zeros of MIMO Systems. In: Lecture Notes on Dynamic Systems and Control*. Cambridge MA: MIT OpenCourseWare (2011). Available online at: https://ocw.mit.edu/courses/electrical-engineering-and-computer-science/6-241j-dynamic-systems-and-control-spring-2011/readings/MIT6_241JS11_chap27.pdf

47. Proakis JG, Salehi M. *Fundamentals of Communication Systems*. Bangalore: Pearson Education India. (2007).

48. Nordio A, Taricco G. Linear receivers for the multiple-input multiple-output multiple-access channel. *IEEE Trans Commun*. (2006) 54:1446–56. doi: 10.1109/TCOMM.2006.878831

49. Takeuchi K, Tanaka T, Yano T. Asymptotic analysis of general multiuser detectors in MIMO DS-CDMA channels. *IEEE J Selected Areas Commun*. (2008) 26:486–96. doi: 10.1109/JSAC.2008.080407

50. Pandey D, Leib H. A tensor based Precoder and Receiver for MIMO GFDM systems. In: *ICC 2021 - IEEE International Conference on Communications (held virtually between 14-23 June 2021)*. Montreal, QC: IEEE (2021). p. 1-6.