Downloaded 02/25/21 to 141.214.17.120. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

SIAM J. CONTROL OPTIM.
Vol. 59, No. 1, pp. 749–776
© 2021 Society for Industrial and Applied Mathematics

MULTILINEAR CONTROL SYSTEMS THEORY*

CAN CHEN†, AMIT SURANA‡, ANTHONY M. BLOCH§, AND INDIKA RAJAPAKSE¶

Abstract. In this paper, we provide a system theoretic treatment of a new class of multilinear time-invariant (MLTI) systems in which the states, inputs, and outputs are tensors, and the system evolution is governed by multilinear operators. The MLTI system representation is based on the Einstein product and even-order paired tensors. There is a particular tensor unfolding which gives rise to an isomorphism from this tensor space to the general linear group, i.e., the group of invertible matrices. By leveraging this unfolding operation, one can extend classical linear time-invariant (LTI) system notions, including stability, reachability, and observability, to MLTI systems. While the unfolding-based formulation is a powerful theoretical construct, the computational advantages of MLTI systems can only be fully realized while working with the tensor form, where hidden patterns/structures can be exploited for efficient representations and computations. Along these lines, we establish new results which enable one to express tensor unfolding–based stability, reachability, and observability criteria in terms of more standard notions of tensor ranks/decompositions. In addition, we develop a generalized CANDECOMP/PARAFAC decomposition– and tensor train decomposition–based model reduction framework, which can significantly reduce the number of MLTI system parameters. We demonstrate our framework with numerical examples.

Key words. multilinear time-invariant systems, stability, reachability, observability, model reduction, tensor unfolding, tensor ranks/decompositions, block tensors

AMS subject classifications. 15A69, 93B05, 93B07, 93B40, 93C05, 93D20

DOI. 10.1137/19M1262589

1. Introduction.

Controlling high-dimensional systems remains an extremely challenging task as many control strategies do not scale well with the dimension of the systems. Of particular interest in this paper are complex biological and engineering systems in which structure, function, and dynamics are highly coupled. Such interactions can be naturally and compactly captured by tensors. Tensors are multidimensional arrays generalized from vectors and matrices and have wide applications in many domains such as social networks, biology, cognitive science, applied mechanics, scientific computation, and signal processing [8, 12, 17, 22, 24]. For example, the organization of the interphase nucleus in the human genome reflects a dynamical interaction between 3D genome structure, its function, and its relationship to phenotype, a concept known as the 4D Nucleome (4DN) [8]. 4DN research requires a comprehensive view of genome-wide structure, gene expression, and the proteome and the phenotype, which fits naturally into a tensorial representation [42, 51]. In

*Received by the editors May 20, 2019; accepted for publication (in revised form) December 11, 2020; published electronically February 25, 2021. This is an extended version of a conference paper presented in [6].
https://doi.org/10.1137/19M1262589
Funding: This work was supported in part under AFOSR Award FA9550-18-1-0028, NSF grant DMS 1613819, the Smale Institute, and the Lifelong Learning Machines program from DARPA/MTO.
†Department of Mathematics and Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI 48109 USA (canc@umich.edu).
‡Raytheon Technologies Research Center, East Hartford, CT 06108 USA (amit.surana@rtx.com).
§Department of Mathematics, University of Michigan, Ann Arbor, MI 48109 USA (abloch@umich.edu).
¶Department of Computational Medicine & Bioinformatics, Medical School, and Department of Mathematics, University of Michigan, Ann Arbor, MI 48109 USA (indikar@umich.edu).

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

order to apply the standard system and control framework in applications such as these, tensors need to be vectorized, leading to an extremely high-dimensional system representation in which the number of states/parameters scales exponentially with the number of dimensions of the tensors involved [51]. Moreover, with the vectorization of tensors, hidden patterns/structures, e.g., redundancy/correlations, can get lost, and thus one cannot exploit such inherent structures for efficient representations and computations.

In order to take advantage of tensor algebraic computations, recently a new class of multilinear time-invariant (MLTI) system has been introduced [43, 51], in which the states and outputs are preserved as tensors. The system evolution is generated by the action of multilinear operators which are formed using Tucker products of matrices. By using tensor unfolding, an operation that transforms a tensor into a matrix, Rogers, Li, and Russell [43] and Surana, Patterson, and Rajapakse [51] developed methods for model identification/reduction from tensor time series data. An application of such tensor-based representation/identification for skeleton-based human behavior recognition from videos demonstrated significant improvements in classification accuracy compared to standard linear time-invariant (LTI) based approaches [12]. However, the MLTI system representation is limited because it assumes the multilinear operators are formed from the Tucker products of matrices (and thus precludes more general tensorial representations) and does not incorporate control inputs.

The role of tensor algebra has also been explored for modeling and simulation of nonlinear dynamics, where the vector field is a multilinear function of states [26]. Tensor decomposition techniques such as CANDECOMP/PARAFAC decomposition (CPD) and tensor train decomposition (TTD) can reduce system size, thus reducing memory usage and enabling efficient computation during simulations. Note that in contrast to the MLTI systems framework of [43, 51], in this application, tensor algebra is applied to the system represented in conventional vector form. The author in [17] exploits tensor decompositions to compute numerical solutions of master equations associated with Markov processes on extremely large state spaces. The Einstein product and even-order paired tensors, along with TTD, were utilized for developing tensor representations for operators based on nearest-neighbor interactions, construction of pseudoinverses for dimensionality reduction methods, and the approximation of transfer operators of dynamical systems.

Similarly, using the Einstein product and even-order paired tensors, Chen et al. [6] generalized the notion of MLTI systems introduced in [43, 51] and also incorporated control inputs. The Einstein product is a tensor contraction used quite often in tensor calculus and has profound applications in the study of continuum mechanics and the field of relativity theory [15, 28]. Moreover, the space of even-order tensors with the Einstein product has many desirable properties. Brazell et al. [3] discovered that one particular tensor unfolding gives rise to an isomorphism from this tensor space (of even-order tensors equipped with the Einstein product) to the general linear group, i.e., the group of invertible matrices. This isomorphism enables one to define matrix equivalent concepts for tensors, including tensor inverse, positive definiteness, and eigenvalue decomposition. Using these tensor constructs, Chen et al. [6] developed tensor algebraic conditions for stability, reachability, and observability for generalized input/output MLTI systems. A new notion of block tensors was also introduced which enables one to express these conditions in a compact fashion. Interestingly, these conditions look analogous to the classical conditions for stability, reachability, and observability in LTI systems, and reduce to them as a special case.

Downloaded 02/25/21 to 141.214.17.120. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/25/21 to 141.214.17.120. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

This paper is an extended version of the introductory paper [6], and in addition to providing various technical details, we also present several new results. The key contributions of this paper are as follows:

1. In [6], the reachability and observability conditions for MLTI systems were stated in terms of the unfolding rank, which requires matricization of the reachability/observability tensors. Here we establish new results relating the unfolding rank to other more standard notions of tensor ranks, including multilinear ranks, CP rank, and TT-ranks. Using such relations, we provide criteria for reachability and observability which do not require tensor unfolding and can be computed using efficient tensor algebraic methods. Similarly, we express MLTI system stability conditions using higher-order singular value decomposition (HOSVD), CPD, and TTD.
2. Using generalized CPD/TTD, we develop a framework for model reduction of MLTI systems. This approach takes advantage of tensor decompositions which otherwise cannot be exploited after unfolding the MLTI systems to obtain a standard LTI form. It also successfully realizes the tensor decomposition–based criteria for stability, reachability, and observability. Furthermore, we establish new stability results by utilizing the factor matrices from tensor decompositions for this reduced model with lesser computational costs.
3. We provide computational and memory complexity analysis for the CPD- and TTD-based methods in comparison to unfolding-based matrix methods and demonstrate our framework in four numerical examples.

The paper is organized into nine sections. In section 2, we review tensor preliminaries, including various notions of tensor products, tensor unfolding, and properties of even-order paired tensors. Section 3 introduces the MLTI system representation using the Einstein product and even-order paired tensors in detail. In section 4, we discuss notions of block tensors and tensor ranks/decompositions. We also build new results relating the unfolding rank of a tensor to other more standard notions of tensor ranks. We establish stability, reachability, and observability conditions for MLTI systems in section 5. The application of generalized CPD/TTD for model reduction is discussed in section 6. Four numerical examples are presented in section 7. Finally, we summarize different numerical approaches associated with MLTI systems in section 8 and conclude in section 9 with future research directions.

2. Tensor preliminaries.

We take most of the concepts and notation for tensor algebra from the comprehensive works of Kolda et al. [24, 25] and Ragnarsson and Van Loan [40, 41]. A tensor is a multidimensional array. The order of a tensor is the number of its dimensions, and each dimension is called a mode. An Nth order tensor usually is denoted by $\mathsf{X} \in \mathbb{R}^{J_1 \times J_2 \times \cdots \times J_N}$. The sets of indexed indices and size of X are denoted by $\mathbf{j} = \{j_1, j_2, \dots, j_N\}$ and $\mathcal{J} = \{J_1, J_2, \dots, J_N\}$, respectively. Let $\Pi_{\mathcal{J}}$ represent the product of all elements in $\mathcal{J}$, i.e., $\Pi_{\mathcal{J}} = \prod_{n=1}^{N} J_n$. It is therefore reasonable to consider scalars $x \in \mathbb{R}$ as zero-order tensors, vectors $\mathbf{v} \in \mathbb{R}^{J}$ as first-order tensors, and matrices $\mathbf{A} \in \mathbb{R}^{J \times I}$ as second-order tensors.

2.1. Tensor products.

By extending the notion of vector outer product, the outer product of two tensors $\mathsf{X} \in \mathbb{R}^{J_1 \times J_2 \times \cdots \times J_N}$ and $\mathsf{Y} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_M}$ is defined as

$$(\mathsf{X} \circ \mathsf{Y})_{j_1 j_2 \dots j_N i_1 i_2 \dots i_M} = \mathsf{X}_{j_1 j_2 \dots j_N} \mathsf{Y}_{i_1 i_2 \dots i_M}.$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

In contrast, the inner product of two tensors $\mathsf{X}, \mathsf{Y} \in \mathbb{R}^{J_1 \times J_2 \times \cdots \times J_N}$ is defined as

$$\langle \mathsf{X}, \mathsf{Y} \rangle = \sum_{j_1=1}^{J_1} \cdots \sum_{j_N=1}^{J_N} \mathsf{X}_{j_1 j_2 \dots j_N} \mathsf{Y}_{j_1 j_2 \dots j_N},$$

leading to the tensor Frobenius norm $\|\mathsf{X}\|^2 = \langle \mathsf{X}, \mathsf{X} \rangle$. We say two tensors X and Y are orthogonal if the inner product $\langle \mathsf{X}, \mathsf{Y} \rangle = 0$. The matrix tensor multiplication $\mathsf{X} \times_n \mathbf{A}$ along mode n for a matrix $\mathbf{A} \in \mathbb{R}^{I \times J_n}$ is defined by $(\mathsf{X} \times_n \mathbf{A})_{j_1 j_2 \dots j_{n-1} i j_{n+1} \dots j_N} = \sum_{j_n=1}^{J_n} \mathsf{X}_{j_1 j_2 \dots j_n \dots j_N} \mathbf{A}_{i j_n}$. This product can be generalized to what is known as the Tucker product, for $\mathbf{A}_n \in \mathbb{R}^{I_n \times J_n}$,

$$\mathsf{X} \times_1 \mathbf{A}_1 \times_2 \mathbf{A}_2 \times_3 \cdots \times_N \mathbf{A}_N = \mathsf{X} \times \{\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_N\} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}. \tag{2.1}$$

2.2. Tensor unfolding. Tensor unfolding is considered as a critical operation in tensor computations [24, 25, 40]. In order to unfold a tensor $\mathsf{X} \in \mathbb{R}^{J_1 \times J_2 \times \cdots \times J_N}$ into a vector or a matrix, we use an index mapping function $ivec(\cdot, \mathcal{J}) : \mathbb{Z}^+ \times \mathbb{Z}^+ \times \overset{N}{\cdots} \times \mathbb{Z}^+ \to \mathbb{Z}^+$ as defined by Ragnarsson and Van Loan [40, 41], which is given by

$$ivec(\mathbf{j}, \mathcal{J}) = j_1 + \sum_{k=2}^{N} (j_k - 1) \prod_{l=1}^{k-1} J_l.$$

The index mapping function $ivec$ returns the index for tensor vectorization, i.e., $\mathbf{x} \in \mathbb{R}^{\Pi \mathcal{J}}$ is the vectorization of X such that $\mathbf{x}_{ivec(\mathbf{j},\mathcal{J})} = \mathsf{X}_{j_1 j_2 \dots j_N}$. If $N = 2$, $ivec$ will stack all the columns of X.

For tensor matricization, let z be an integer such that $1 \leq z < N$, and let $\mathbb{S}$ be a permutation of $\{1, 2, \dots, N\}$. If $\mathbf{r} = \{\mathbb{S}(1), \mathbb{S}(2), \dots, \mathbb{S}(z)\}$ and $\mathbf{c} = \{\mathbb{S}(z+1), \mathbb{S}(z+2), \dots, \mathbb{S}(N)\}$ with $\mathcal{P} = \{J_{\mathbb{S}(1)}, J_{\mathbb{S}(2)}, \dots, J_{\mathbb{S}(z)}\}$ and $\mathcal{Q} = \{J_{\mathbb{S}(z+1)}, J_{\mathbb{S}(z+2)}, \dots, J_{\mathbb{S}(N)}\}$, respectively, the $\mathbf{rc}$-unfolding matrix of X, denoted by $\mathbf{X}_{(\mathbf{rc})} \in \mathbb{R}^{\Pi \mathcal{P} \times \Pi \mathcal{Q}}$, is given by

$$(\mathbf{X}_{(\mathbf{rc})})_{pq} = \mathsf{X}^{\mathbb{S}}_{p_1 p_2 \dots p_z q_1 q_2 \dots q_{N-z}}, \tag{2.2}$$

where $p = ivec(\mathbf{p}, \mathcal{P})$, $q = ivec(\mathbf{q}, \mathcal{Q})$, and $\mathsf{X}^{\mathbb{S}}$ is the $\mathbb{S}$-transpose of X defined as

$$\mathsf{X}^{\mathbb{S}}_{j_{\mathbb{S}(1)} j_{\mathbb{S}(2)} \dots j_{\mathbb{S}(N)}} = \mathsf{X}_{j_1 j_2 \dots j_N}.$$

When $z = 1$ and $\mathbb{S} = \begin{pmatrix} 1 & 2 & \dots & n & n+1 & \dots & N \\ n & 1 & \dots & n-1 & n+1 & \dots & N \end{pmatrix}$, the tensor unfolding is called the n-mode matricization, denoted by $\mathbf{X}_{(n)}$.

2.3. Even-order paired tensors. Here we discuss the notion of even-order paired tensors and the Einstein product, which will play an important role in developing the MLTI systems theory.

DEFINITION 2.1. *Even-order paired tensors are $2N$th order tensors with elements specified using a pairwise index notation, i.e., $\mathsf{A}_{j_1 i_1 \dots j_N i_N}$ for $\mathsf{A} \in \mathbb{R}^{J_1 \times I_1 \times \cdots \times J_N \times I_N}$.*

DEFINITION 2.2. *Given an even-order paired tensor $\mathsf{A} \in \mathbb{R}^{J_1 \times I_1 \times \cdots \times J_N \times I_N}$, the Einstein product between A and an Nth order tensor $\mathsf{X} \in \mathbb{R}^{I_1 \times I_2 \times \cdots \times I_N}$ is the contraction along the second index in each pair from A, i.e.,*

$$(\mathsf{A} * \mathsf{X})_{j_1 j_2 \dots j_N} = \sum_{i_1=1}^{I_1} \cdots \sum_{i_N=1}^{I_N} \mathsf{A}_{j_1 i_1 \dots j_N i_N} \mathsf{X}_{i_1 i_2 \dots i_N}. \tag{2.3}$$

Downloaded 02/25/21 to 141.214.17.120. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/25/21 to 141.214.17.120. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

We use the pairwise index notation for even-order tensors because it is convenient for defining the unfolding transformation φ (see Definition 2.4) and for representing core matrices/tensors in tensor decompositions (see subsection 6.1). Note that even-order paired tensors and the Einstein product (2.3) can be viewed as multidimensional generalizations of matrices and the standard matrix-vector product, respectively [17]. Similar to the standard matrix-matrix product, one can also define a generalized form of the Einstein product between two even-order paired tensors. We will see later that the Einstein product can be efficiently computed using tensor decompositions of even-order paired tensors; see Proposition 6.3.

DEFINITION 2.3. *Given two even-order paired tensors* $\mathsf{A} \in \mathbb{R}^{J_1\times K_1\times\cdots\times J_N\times K_N}$ *and* $\mathsf{B} \in \mathbb{R}^{K_1\times I_1\times\cdots\times K_N\times I_N}$*, the Einstein product* $\mathsf{A} * \mathsf{B} \in \mathbb{R}^{J_1\times I_1\times\cdots\times J_N\times I_N}$ *is defined by*

$$(\mathsf{A} * \mathsf{B})_{j_1 i_1 \dots j_N i_N} = \sum_{k_1=1}^{K_1} \cdots \sum_{k_N=1}^{K_N} \mathsf{A}_{j_1 k_1 \dots j_N k_N} \mathsf{B}_{k_1 i_1 \dots k_N i_N}. \tag{2.4}$$

Brazell et al. [3] investigated properties of even-order tensors under the Einstein product (different from (2.4)) through construction of an isomorphism to $\mathrm{GL}(n,\mathbb{R})$, i.e., the set of $n\times n$ real-valued invertible matrices. The existence of the isomorphism enables one to generalize several matrix concepts, such as invertibility and eigenvalue decomposition, to the tensor case [3, 9, 18, 30, 50]. We can establish an analogous isomorphism for even-order paired tensors by a permutation of indices.

DEFINITION 2.4. *Define the map* φ*:* $\mathbb{T}_{J_1 I_1 \dots J_N I_N}(\mathbb{R}) \to \mathbb{M}_{\Pi_{\mathcal{J}}\Pi_{\mathcal{I}}}(\mathbb{R})$ *with* $\varphi(\mathsf{A}) = \boldsymbol{A}$ *defined componentwise as*

$$\mathsf{A}_{j_1 i_1 \dots j_N i_N} \xrightarrow{\varphi} \boldsymbol{A}_{ivec(\boldsymbol{j},\mathcal{J})\, ivec(\boldsymbol{i},\mathcal{I})}, \tag{2.5}$$

where $\mathbb{T}_{J_1 I_1 \dots J_N I_N}(\mathbb{R})$ *is the set of all real* $J_1 \times I_1 \times \cdots \times J_N \times I_N$ *even-order paired tensors, and* $\mathbb{M}_{\Pi_{\mathcal{J}}\Pi_{\mathcal{I}}}(\mathbb{R})$ *is set of all real* $\Pi_{\mathcal{J}} \times \Pi_{\mathcal{I}}$ *matrices.*

The map φ can be viewed as a tensor unfolding discussed in (2.2) with $z = N$ and $\mathbb{S} = \begin{pmatrix} 1 & 2 & \dots & N & N+1 & N+2 & \dots & 2N \\ 1 & 3 & \dots & 2N-1 & 2 & 4 & \dots & 2N \end{pmatrix}$, so the Frobenius norm is preserved through φ, i.e., $\|\mathsf{A}\| = \|\varphi(\mathsf{A})\|$. More significantly, φ is bijective, and the restriction of φ^{-1} on the general linear group produces a group isomorphism.

COROLLARY 2.5. *Let* $J_n = I_n$ *for all* n *and* $\mathbb{G}_{J_1 J_1 \dots J_N J_N}(\mathbb{R}) = \varphi^{-1}(GL(\Pi_{\mathcal{J}}, \mathbb{R}))$*, i.e.,* $\mathbb{G}_{J_1 J_1 \dots J_N J_N}$ *is the space of all even-order paired tensors which maps to the general linear group under* φ*. Then* $\mathbb{G}_{J_1 J_1 \dots J_N J_N}(\mathbb{R})$ *is a group equipped with the Einstein product* (2.4)*, and* φ *is a group isomorphism.*

Detailed proofs can be found in [3, 18]. Based on the unfolding property, we can define some tensor notation analogous to matrices as follows:

1. For an even-order paired tensor $\mathsf{A} \in \mathbb{R}^{J_1\times I_1\times\cdots\times J_N\times I_N}$, $\mathsf{T} \in \mathbb{R}^{I_1\times J_1\times\cdots\times I_N\times J_N}$ is called the U-transpose of A if $\mathsf{T}_{i_1 j_1 \dots i_N j_N} = \mathsf{A}_{j_1 i_1 \dots j_N i_N}$, and is denoted by $\mathsf{A}^\top$. We refer to an even-order paired tensor that is identical to its U-transpose as weakly symmetric.
2. For an even-order paired tensor $\mathsf{A} \in \mathbb{R}^{J_1\times I_1\times\cdots\times J_N\times I_N}$, the unfolding rank of A is defined as $\mathrm{rank}_U(\mathsf{A}) = \mathrm{rank}(\varphi(\mathsf{A}))$ [30].
3. An even-order "square" tensor $\mathsf{D} \in \mathbb{R}^{J_1\times J_1\times\cdots\times J_N\times J_N}$ is called the U-diagonal tensor if all its entries are zeros except for $\mathsf{D}_{j_1 j_1 \dots j_N j_N}$. If all the diagonal entries $\mathsf{D}_{j_1 j_1 \dots j_N j_N} = 1$, then D is the U-identity tensor, denoted by I.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/25/21 to 141.214.17.120. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

4. For an even-order square tensor $\mathsf{A} \in \mathbb{R}^{J_1\times J_1\times\cdots\times J_N\times J_N}$, if there exists a tensor $\mathsf{B} \in \mathbb{R}^{J_1\times J_1\times\cdots\times J_N\times J_N}$ such that $\mathsf{A} * \mathsf{B} = \mathsf{B} * \mathsf{A} = \mathsf{I}$, then B is called the U-inverse of A, denoted by A^{-1}.
5. An even-order square tensor $\mathsf{A} \in \mathbb{R}^{J_1\times J_1\times\cdots\times J_N\times J_N}$ is called U-positive definite if $\mathsf{X}^\top * \mathsf{A} * \mathsf{X} > 0$ for any nonzero tensor $\mathsf{X} \in \mathbb{R}^{J_1\times J_2\times\cdots\times J_N}$.
6. For an even-order square tensor $\mathsf{A} \in \mathbb{R}^{J_1\times J_1\times\cdots\times J_N\times J_N}$, the unfolding determinant of A is defined as $\det_U(\mathsf{A}) = \det(\varphi(\mathsf{A}))$ [30].

In Appendix A.1, we show that the notion of U-positive definiteness is a generalization of M-positive definiteness and rank-one positive definiteness proposed in [19, 39] for the even-order elasticity tensors.

3. MLTI system representation.

To describe the evolution of tensor time series, the authors in [43, 51] introduced an MLTI system using the Tucker product, which can be generalized by incorporating control inputs as follows:

$$
\begin{cases} \mathsf{X}_{t+1} = \mathsf{X}_t \times \{\mathbf{A}_1, \dots, \mathbf{A}_N\} + \mathsf{U}_t \times \{\mathbf{B}_1, \dots, \mathbf{B}_N\}, \\ \mathsf{Y}_t = \mathsf{X}_t \times \{\mathbf{C}_1, \dots, \mathbf{C}_N\}, \end{cases} \tag{3.1}
$$

where $\mathsf{X}_t \in \mathbb{R}^{J_1\times J_2\times\cdots\times J_N}$ is the latent state space tensor, $\mathsf{Y}_t \in \mathbb{R}^{I_1\times I_2\times\cdots\times I_N}$ is the output tensor, and $\mathsf{U}_t \in \mathbb{R}^{K_1\times K_2\times\cdots\times K_N}$ is the control tensor. $\mathbf{A}_n \in \mathbb{R}^{J_n\times J_n}$, $\mathbf{B}_n \in \mathbb{R}^{J_n\times K_n}$, and $\mathbf{C}_n \in \mathbb{R}^{I_n\times J_n}$ are real-valued matrices for $n = 1, 2, \dots, N$. The Tucker product provides a suitable way to deal with MLTI systems because it allows one to exploit matrix computations. However, we find that (3.1) can be replaced by a more general representation using the notion of even-order paired tensors and the Einstein product. Moreover, the new representation is more concise and systematic compared to the tensor-based linear system proposed in [12].

Definition 3.1. *A more general representation of an MLTI system is given by*

$$
\begin{cases} \mathsf{X}_{t+1} = \mathsf{A} * \mathsf{X}_t + \mathsf{B} * \mathsf{U}_t, \\ \mathsf{Y}_t = \mathsf{C} * \mathsf{X}_t, \end{cases} \tag{3.2}
$$

where $\mathsf{A} \in \mathbb{R}^{J_1\times J_1\times\cdots\times J_N\times J_N}$, $\mathsf{B} \in \mathbb{R}^{J_1\times K_1\times\cdots\times J_N\times K_N}$, *and* $\mathsf{C} \in \mathbb{R}^{I_1\times J_1\times\cdots\times I_N\times J_N}$ *are even-order paired tensors.*

Lemma 3.2. *Let* $\mathsf{A} \in \mathbb{R}^{J_1\times I_1\times\cdots\times J_N\times I_N}$ *be an even-order paired tensor. Then the product* $\mathsf{A} \times \{\boldsymbol{U}_1, \boldsymbol{V}_1, \dots, \boldsymbol{U}_N, \boldsymbol{V}_N\} = \mathsf{U} * \mathsf{A} * \mathsf{V}^\top \in \mathbb{R}^{K_1\times L_1\times\cdots\times K_N\times L_N}$ *for* $\mathsf{U} = \boldsymbol{U}_1 \circ \boldsymbol{U}_2 \circ \cdots \circ \boldsymbol{U}_N$ *and* $\mathsf{V} = \boldsymbol{V}_1 \circ \boldsymbol{V}_2 \circ \cdots \circ \boldsymbol{V}_N$, *where* $\boldsymbol{U}_n \in \mathbb{R}^{K_n\times J_n}$ *and* $\boldsymbol{V}_n \in \mathbb{R}^{L_n\times I_n}$.

Proof. This follows from the definitions of the Tucker and Einstein products. ◻

Proposition 3.3. *The governing equations* (3.2) *can be obtained from* (3.1) *by setting* A, B, *and* C *to be the outer products of component matrices* $\{\boldsymbol{A}_1, \boldsymbol{A}_2, \dots, \boldsymbol{A}_N\}$, $\{\boldsymbol{B}_1, \boldsymbol{B}_2, \dots, \boldsymbol{B}_N\}$, *and* $\{\boldsymbol{C}_1, \boldsymbol{C}_2, \dots, \boldsymbol{C}_N\}$, *respectively.*

Proof. The result follows from Lemma 3.2 with $I_n = 1$ and $\mathbf{V}_n = 1$ for all n. ◻

The main advantages of the MLTI system (3.2) are as follows:

1. The Einstein product representation (3.2) of MLTI systems is indeed the generalization of (3.1). While Proposition 3.3 shows that MLTI systems in the form of (3.1) can always be transformed into the form of (3.2), the converse is not always true; see (6.4), for example. It is true only when $R_1 = R_2 = R_3 = 1$.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/25/21 to 141.214.17.120. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

2. The MLTI system (3.2) takes a form similar to the standard LTI system model with matrix product replaced with the Einstein product, so the representation is more natural for developing the MLTI systems theory including notions of stability, reachability, and observability. Moreover, the concept of transfer functions, which is commonly used in modern control theory, can be extended for MLTI systems; see Definition 3.4.
3. We can exploit tensor decompositions (see subsection 4.2) of the even-order paired tensors A, B, and C to accelerate computations in MLTI systems theory. In particular, if A, B, and C possess low tensor rank structures, we can obtain a low-parameter MLTI representation. In addition, many operations such as the Einstein product and unfolding rank can be achieved efficiently in the tensor decomposition format compared to unfolding-based matrix methods; see the remarks in sections 5 and 6.
4. Traditional LTI model reduction and identification techniques such as balanced truncation and eigensystem realization algorithm can be extended using the form of (3.2).

DEFINITION 3.4. *The transfer function* $\mathsf{G}(z)$ *of* (3.2) *is given by*

$$\mathsf{G}(z) = \mathsf{C} * (z\mathsf{I} - \mathsf{A})^{-1} * \mathsf{B}, \tag{3.3}$$

where z is a complex variable.

We first investigate the elementary solution to the MLTI system (3.2), which is crucial in the analysis of stability, reachability, and observability.

PROPOSITION 3.5. *For an unforced MLTI system* $\mathsf{X}_{t+1} = \mathsf{A} * \mathsf{X}_t$, *the solution for* X *at time k, given initial condition* X_0, *is* $\mathsf{X}_k = \mathsf{A}^k * \mathsf{X}_0$, *where* $\mathsf{A}^k = \mathsf{A} * \mathsf{A} * \overset{k}{\cdots} * \mathsf{A}$.

The proof is straightforward using the notion of even-order paired tensors and the Einstein product. Applying Proposition 3.5, we can write down the explicit solution of (3.2), which takes an analogous form to the LTI system

$$\mathsf{X}_k = \mathsf{A}^k * \mathsf{X}_0 + \sum_{j=0}^{k-1} \mathsf{A}^{k-j-1} * \mathsf{B} * \mathsf{U}_j. \tag{3.4}$$

Lastly, we want to note that one can always transform the MLTI system (3.2) into an LTI system using φ, i.e., $\mathbf{x}_{t+1} = \varphi(\mathsf{A})\mathbf{x}_t + \varphi(\mathsf{B})\mathbf{u}_t$, and determine the stability, reachability, and observability using classical matrix techniques.

4. Tensor algebra continued.

We next discuss notions of block tensors and tensor decompositions which will form the basis for developing tensor algebra–based concepts of stability, reachability, and observability of the MLTI system (3.2).

4.1. Block tensors.

Analogously to block matrices, one can define the notion of block tensors. For tensors of the same size, we propose a block tensor construction (first appeared in [6]) which does not introduce any wasteful zeros compared to the block tensors proposed in [50], and thus offers computational advantages.

DEFINITION 4.1. *Let* $\mathsf{A}, \mathsf{B} \in \mathbb{R}^{J_1 \times I_1 \times \cdots \times J_N \times I_N}$ *be two even-order paired tensors. The n-mode row block tensor* $\begin{vmatrix} \mathsf{A} & \mathsf{B} \end{vmatrix}_n \in \mathbb{R}^{J_1 \times I_1 \times \cdots \times J_n \times 2I_n \times \cdots \times J_N \times I_N}$ *is defined by*

$$(\begin{vmatrix} \mathsf{A} & \mathsf{B} \end{vmatrix}_n)_{j_1 l_1 \dots j_N l_N} = \begin{cases} \mathsf{A}_{j_1 l_1 \dots j_N l_N}, & j_k = 1, \dots, J_k,\ l_k = 1, \dots, I_k\ \forall k, \\ \mathsf{B}_{j_1 l_1 \dots j_N l_N}, & j_k = 1, \dots, J_k\ \forall k, l_k = 1, \dots, I_k \textit{ for } k \neq n, \\ & \textit{and } l_k = I_k + 1, \dots, 2I_k \textit{ for } k = n. \end{cases}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/25/21 to 141.214.17.120. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

For example, if $\mathsf{A}, \mathsf{B} \in \mathbb{R}^{2\times 2\times 2\times 2}$, then the 1-mode row block tensor is given by $\begin{vmatrix}\mathsf{A} & \mathsf{B}\end{vmatrix}_1 \in \mathbb{R}^{2\times 4\times 2\times 2}$ such that $(\begin{vmatrix}\mathsf{A} & \mathsf{B}\end{vmatrix}_1)_{:i_1::} = \mathsf{A}$ for $i_1 = 1, 2$ and $(\begin{vmatrix}\mathsf{A} & \mathsf{B}\end{vmatrix}_1)_{:i_1::} = \mathsf{B}$ for $i_1 = 3, 4$. Similarly for $\begin{vmatrix}\mathsf{A} & \mathsf{B}\end{vmatrix}_2 \in \mathbb{R}^{2\times 2\times 2\times 4}$. Detailed explanations of the MATLAB colon operation ":" can be found in Appendix D.1. When $N = 1$, it reduces to the row block matrices. The n-mode column block tensor

$$\begin{vmatrix}\mathsf{A} \\ \mathsf{B}\end{vmatrix}_n \in \mathbb{R}^{J_1\times I_1\times\cdots\times 2J_n\times I_n\times\cdots\times J_N\times I_N}$$

can be defined in a similar manner. The n-mode block tensors exhibit many properties analogous to block matrix computations, e.g., the Einstein product can distribute over block tensors, and the blocks of n-mode row block tensors map to contiguous blocks under φ up to some permutations [40]; see details in Appendix A.2. Therefore, rank is preserved in the block tensor unfolding, i.e., $\operatorname{rank}_U(\begin{vmatrix}\mathsf{A} & \mathsf{B}\end{vmatrix}_n) = \operatorname{rank}(\begin{bmatrix}\varphi(\mathsf{A}) & \varphi(\mathsf{B})\end{bmatrix})$, where $[\cdot]$ denotes the block matrix operation.

Given K even-order paired tensors $\mathsf{X}_n \in \mathbb{R}^{J_1\times I_1\times\cdots\times J_N\times I_N}$, one can apply Definition 4.1 successively to create a $J_1 \times I_1 \times \cdots \times J_n \times I_nK \times \cdots \times J_N \times I_N$ even-order n-mode row block tensor. However, a more general concatenation approach can be defined for multiple blocks.

DEFINITION 4.2. *Given K even-order paired tensors $\mathsf{X}_n \in \mathbb{R}^{J_1\times I_1\times\cdots\times J_N\times I_N}$, if $K = K_1K_2\ldots K_N$, the $J_1 \times I_1K_1 \times \cdots \times J_N \times I_NK_N$ even-order mode row block tensor Y can be constructed in the following way:*

1. *Compute the 1-mode row block tensor concatenation over $\{\mathsf{X}_1, \ldots, \mathsf{X}_{K_1}\}$, $\{\mathsf{X}_{K_1+1}, \ldots, \mathsf{X}_{2K_1}\}$, and so on to obtain $K_2K_3\ldots K_N$ block tensors denoted by $\mathsf{X}^{(1)}_1, \mathsf{X}^{(1)}_2, \ldots, \mathsf{X}^{(1)}_{K_2K_3\ldots K_N}$.*
2. *Compute the 2-mode row block tensors concatenation over $\{\mathsf{X}^{(1)}_1, \ldots, \mathsf{X}^{(1)}_{K_2}\}$, $\{\mathsf{X}^{(1)}_{K_2+1}, \ldots, \mathsf{X}^{(1)}_{2K_2}\}$, and so on to obtain $K_3K_4\ldots K_N$ block tensors denoted by $\mathsf{X}^{(2)}_1, \mathsf{X}^{(2)}_2, \ldots, \mathsf{X}^{(2)}_{K_3K_4\ldots K_N}$.*
3. *Keep repeating the process until the last N-mode row block tensor is obtained.*

We denote the mode row block tensor as $\mathsf{Y} = \begin{vmatrix}\mathsf{X}_1 & \mathsf{X}_2 & \ldots & \mathsf{X}_K\end{vmatrix}$.

For example, suppose that $\mathsf{X}_n \in \mathbb{R}^{2\times 2\times 2\times 2\times 2\times 2}$ for $n = 1, 2, \ldots, K$ and $K = 8$. Let $K = K_1K_2K_3$ with $K_1 = K_2 = K_3 = 2$. Given this factorization of K, the mode row block tensor $\mathsf{Y} \in \mathbb{R}^{2\times 4\times 2\times 4\times 2\times 4}$ is constructed in the manner shown in Figure 1, in which $\mathsf{X}^{(1)}_n \in \mathbb{R}^{2\times 4\times 2\times 2\times 2\times 2}$ and $\mathsf{X}^{(2)}_n \in \mathbb{R}^{2\times 4\times 2\times 4\times 2\times 2}$. Another factorization with $K_1 = 2$, $K_2 = 4$, and $K_3 = 1$ would return $\mathsf{Y} \in \mathbb{R}^{2\times 4\times 2\times 8\times 2\times 2}$. The generalized mode column block tensors with multiple blocks can be constructed in a similar manner. When $I_n = 1$ for all n, the above generalized mode row block tensor maps exactly to contiguous blocks in its unfolding under φ, which could be beneficial in many block tensor applications. Furthermore, the choices of K_n may affect the structure of mode block tensors, which can be significant in tensor ranks/decompositions [7].

4.2. Tensor ranks and decompositions. There are several definitions of tensor ranks [11, 24, 25], which are intimately related to different notions of tensor decompositions. The multilinear ranks or the n-ranks of X are the ranks of the n-mode matricizations, denoted by $\operatorname{rank}_n(\mathsf{X})$. The multilinear ranks are related to the so-called higher-order singular value decomposition (HOSVD), a multilinear generalization of the matrix singular value decomposition (SVD) [2, 10].

THEOREM 4.3 (HOSVD). *A tensor $\mathsf{X} \in \mathbb{R}^{J_1\times J_2\times\cdots\times J_N}$ can be written as*

$$\mathsf{X} = \mathsf{S} \times_1 \boldsymbol{U}_1 \times_2 \cdots \times_N \boldsymbol{U}_N, \tag{4.1}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

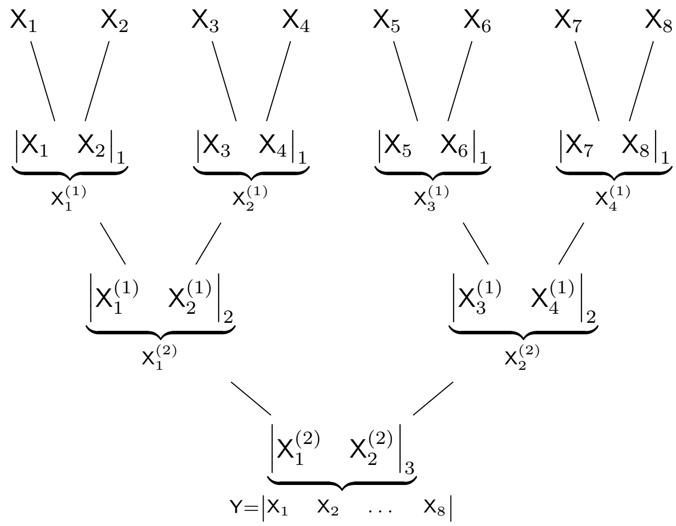

FIG. 1. *An example of mode row block tensor.*

where $\boldsymbol{U}_n \in \mathbb{R}^{J_n \times J_n}$ are orthogonal matrices, and $\mathsf{S} \in \mathbb{R}^{J_1 \times J_2 \times \cdots \times J_N}$ is a tensor of which the subtensors $\mathsf{S}_{j_n=\alpha}$ obtained by fixing the nth index to α have the following properties:

1. *All-orthogonality: two subtensors $\mathsf{S}_{j_n=\alpha}$ and $\mathsf{S}_{j_n=\beta}$ are orthogonal for all possible values of n, α, and β subject to $\alpha \neq \beta$.*
2. *Ordering: $\|\mathsf{S}_{j_n=1}\| \geq \cdots \geq \|\mathsf{S}_{j_n=J_n}\| \geq 0$ for all possible values of n.*

The Frobenius norms $\|\mathsf{S}_{j_n=j}\|$, denoted by $\gamma_j^{(n)}$, are the n-mode singular values of X.

De Lathauwer, De Moor, and Vandewalle [10] showed that the number of nonvanishing n-mode singular values from the HOSVD of a tensor is equal to its n-mode multilinear rank. In addition, the error bound of the low multilinear rank approximation is provided in [10]. Unlike the matrix SVD, the approximation fails to obtain the best rank approximation of X. Nevertheless, it still can provide a "good" estimate with appropriate n-mode singular values truncated [10].

Analogous to rank-one matrices, a tensor X is rank-one if it can be written as the outer product of N vectors, i.e., $\mathsf{X} = \mathbf{a}^{(1)} \circ \mathbf{a}^{(2)} \circ \cdots \circ \mathbf{a}^{(N)}$. The CANDECOMP/PARAFAC Decomposition (CPD) decomposes a tensor $\mathsf{X} \in \mathbb{R}^{J_1 \times J_2 \times \cdots \times J_N}$ into a sum of rank-one tensors in the form of outer products. It is often useful to normalize all the vectors and have weights $\lambda_r > 0$ in descending order in front:

$$\mathsf{X} = \sum_{r=1}^{R} \lambda_r \mathbf{a}_r^{(1)} \circ \mathbf{a}_r^{(2)} \circ \cdots \circ \mathbf{a}_r^{(N)}, \tag{4.2}$$

where $\mathbf{a}_r^{(n)} \in \mathbb{R}^{J_n}$ have unit length, and R is called the CP rank of X if it is the minimum integer that achieves (4.2). The factor matrices $\mathbf{A}^{(n)} \in \mathbb{R}^{J_n \times R}$ are the combination of the vectors from the rank-one components for $n = 1, 2, \dots, N$, i.e., $\mathbf{A}^{(n)} = \begin{bmatrix}\mathbf{a}_1^{(n)} & \mathbf{a}_2^{(n)} & \dots & \mathbf{a}_R^{(n)}\end{bmatrix}$. The CPD is unique up to scaling and permutation under a weak condition: for $N \geq 2$ and $R \geq 2$, $\sum_{n=1}^{N} k_{\mathbf{A}^{(n)}} \geq 2R + (N-1)$, where $k_{\mathbf{A}^{(n)}}$, called the k-rank of a matrix, is the maximum number of columns of $\mathbf{A}^{(n)}$ that are linearly independent of each other [27, 47, 49].

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/25/21 to 141.214.17.120. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

Downloaded 02/25/21 to 141.214.17.120. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

The CP rank of a tensor is always greater than or equal to its multilinear ranks [11]. In fact, it is greater than or equal to any unfolding matrix rank [37] (which can be used in unfolding rank and TT-ranks defined later too). The best CP rank approximation is ill-posed [11], but carefully truncating the CP rank will yield a good estimate of the original tensor. Both CPD and HOSVD are special cases of Tucker decomposition, which decomposes a tensor into the form of Tucker product (2.1), i.e., $\mathsf{Y} = \mathsf{X} \times \{\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_N\}$ [26].

The tensor train decomposition (TTD) of $\mathsf{X} \in \mathbb{R}^{J_1 \times J_2 \times \cdots \times J_N}$ is given by

$$\mathsf{X} = \sum_{r_0=1}^{R_0} \cdots \sum_{r_N=1}^{R_N} \mathsf{X}^{(1)}_{r_0:r_1} \circ \mathsf{X}^{(2)}_{r_1:r_2} \circ \cdots \circ \mathsf{X}^{(N)}_{r_{N-1}:r_N}, \tag{4.3}$$

where $\{R_0, R_1, \dots, R_N\}$ is the set of TT-ranks with $R_0 = R_N = 1$, and $\mathsf{X}^{(n)} \in \mathbb{R}^{R_{n-1} \times J_n \times R_n}$ are called the core tensors of the TTD [35]. Here we have used ":" for brevity of notation; for the full definition see Appendix D.1. Standard TTD algorithms, such as Algorithm 3 in [23], with zero truncation will return the optimal TT-ranks

$$R_n = \text{rank}\left(\texttt{reshape}\left(\mathsf{X}, \prod_{i=1}^{n} J_i, \prod_{i=n+1}^{N} J_i\right)\right)$$

for $n = 1, 2, \dots, N-1$. A core tensor $\mathsf{X}^{(n)}$ is left-orthonormal if $(\bar{\mathbf{X}}^{(n)})^\top \bar{\mathbf{X}}^{(n)} = \mathbf{I} \in \mathbb{R}^{R_n \times R_n}$, and is right-orthonormal if $\underline{\mathbf{X}}^{(n)} (\underline{\mathbf{X}}^{(n)})^\top = \mathbf{I} \in \mathbb{R}^{R_{n-1} \times R_{n-1}}$, where $\bar{\mathbf{X}}^{(n)} = \texttt{reshape}(\mathsf{X}^{(n)}, R_{n-1}J_n, R_n)$ and $\underline{\mathbf{X}}^{(n)} = \texttt{reshape}(\mathsf{X}^{(n)}, R_{n-1}, J_n R_n)$, respectively [14, 23]. Here `reshape` refers to the reshape operation in MATLAB; see details in Appendix D.2. Detailed algorithms for left- and right-orthonormalization can be found in [23]. TTD is advantageous in that it provides better compression, i.e., truncating the TT-ranks results in a quasi-optimal approximation of X, and is computationally more robust [35].

Eigenvalue problems for tensors were first explored by Qi [38] and Lim [29] independently. Brazell et al. [3] formulated a new tensor eigenvalue problem through the isomorphism φ for fourth-order tensors, and Cui et al. [9] extended the tensor eigenvalue problem to even-order tensors.

DEFINITION 4.4. *Let $\mathsf{A} \in \mathbb{R}^{J_1 \times J_1 \times \cdots \times J_N \times J_N}$ be an even-order square tensor. If $\mathsf{X} \in \mathbb{C}^{J_1 \times J_2 \cdots \times J_N}$ is a nonzero Nth order tensor, $\lambda \in \mathbb{C}$, and X and λ satisfy $\mathsf{A} * \mathsf{X} = \lambda \mathsf{X}$, then we refer to λ and X as the U-eigenvalue and U-eigentensor of A, respectively.*

The algebraic and geometric multiplicities of U-eigenvalues can be defined as for matrices. The generalization of the Caley–Hamilton theorem for the tensor case can be obtained by the isomorphism property, i.e., an even-order square tensor A satisfies its own characteristic polynomial $p(\lambda) = \det_U(\lambda \mathsf{I} - \mathsf{A})$. Moreover, it can be shown that the notion of U-eigenvalues is a generalization of Z-eigenvalues and M-eigenvalues as proposed in [19, 29, 38]. Detailed proofs are omitted in this paper.

PROPOSITION 4.5. *The tensor eigenvalue problem in Definition 4.4 can be represented by $\mathsf{A} = \mathsf{V} * \mathsf{D} * \mathsf{V}^{-1}$, where $\mathsf{D} \in \mathbb{R}^{J_1 \times J_1 \times \cdots \times J_N \times J_N}$ is a U-diagonal tensor with U-eigenvalues on its diagonal, and $\mathsf{V} \in \mathbb{R}^{J_1 \times J_1 \times \cdots \times J_N \times J_N}$ is a mode row block tensor consisting of all the U-eigentensors, i.e., $\mathsf{V} = \begin{vmatrix} \mathsf{X}_1 & \mathsf{X}_2 & \dots & \mathsf{X}_{\Pi_{\mathcal{J}}} \end{vmatrix}$. We have chosen $K_n = J_n$ in applying the mode row block tensor operation which enables us to express the tensor eigenvalue decomposition in a form analogous to the matrix case.*

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Proof. The proof follows immediately from Proposition A.3. ∎

4.3. Rank relations. We establish new results relating the unfolding rank of an even-order paired tensor to its multilinear ranks, CP rank, and TT-ranks. These relationships are useful for checking multilinear generalizations of reachability and observability rank conditions.

PROPOSITION 4.6. *Let* $\mathsf{A} \in \mathbb{R}^{J_1\times I_1\times\cdots\times J_N\times I_N}$ *be an even-order paired tensor. If* $\text{rank}_U(\mathsf{A}) = \Pi_{\mathcal{J}}$ *(or* $\text{rank}_U(\mathsf{A}) = \Pi_{\mathcal{I}}$*), then* $\text{rank}_{2n-1}(\mathsf{A}) = J_n$ *(or* $\text{rank}_{2n}(\mathsf{A}) = I_n$*) for* $n = 1, 2, \dots, N$*.*

PROPOSITION 4.7. *Let* $\mathsf{A} \in \mathbb{R}^{J_1\times I_1\times\cdots\times J_N\times I_N}$ *be an even-order paired tensor given in the CPD format* (4.2) *with CP rank equal to R. If the conditions*

$$\sum_{n=1:2}^{2N} k_{\boldsymbol{A}^{(n)}} \geq R + N - 1, \qquad \sum_{n=2:2}^{2N} k_{\boldsymbol{A}^{(n)}} \geq R + N - 1 \tag{4.4}$$

are satisfied for every $k_{\boldsymbol{A}^{(n)}} \geq 1$*, then* $\text{rank}_U(\mathsf{A}) = R$*.*

The notations $\sum_{n=1:2}^{2N}$ and $\sum_{n=2:2}^{2N}$ represent the sums of all odd indices and all even indices, respectively. The detailed proofs of Propositions 4.6 and 4.7 can be found in Appendices B.1 and B.2, respectively.

PROPOSITION 4.8. *Let* $\mathsf{A} \in \mathbb{R}^{J_1\times I_1\times\cdots\times J_N\times I_N}$ *be an even-order paired tensor. Then* $\text{rank}_U(\mathsf{A}) = \tilde{R}_N$*, where* $\tilde{R}_N$ *is the Nth optimal TT-rank of* $\tilde{\mathsf{A}}$*, the* $\mathbb{S}$*-transpose of* A *with* $\mathbb{S} = \begin{pmatrix} 1 & 2 & \dots & N & N+1 & N+2 & \dots & 2N \\ 1 & 3 & \dots & 2N-1 & 2 & 4 & \dots & 2N \end{pmatrix}$.

Proof. The result follows from the definition of optimal TT-ranks. ∎

Remark. Given the TTD of A, the TTD of $\tilde{\mathsf{A}} \in \mathbb{R}^{J_1\times\cdots\times J_N\times I_1\times\cdots\times I_N}$ can be constructed by manipulating the core tensors $\mathsf{A}^{(n)}$ without converting back to the full format. Assume that $J_n = I_n = J$ for all n, and R is the average of the TT-ranks of A. If R remains unchanged or decreases during this conversion, the computational complexity is estimated to be at most $\mathcal{O}(N^2J^3R^3)$.[1] A detailed algorithm for the TTD-based permutation can be found in [7].

5. MLTI systems theory. We now introduce the concepts of stability, reachability, and observability for MLTI systems. Note that some preliminary results have appeared in our introductory paper [6].

5.1. Stability. There are many notions of stability for dynamical systems [4, 21, 46]. For LTI systems, it is conventional to investigate so-called internal stability. Generalizing from LTI systems, the equilibrium point $\mathsf{X} = \mathsf{O}$ (O denotes the zero tensors) of an unforced MLTI system is called stable if $\|\mathsf{X}_t\| \leq \gamma\|\mathsf{X}_0\|$ for some $\gamma > 0$, asymptotically stable if $\|\mathsf{X}_t\| \to 0$ as $t \to \infty$, and unstable if it is not stable.

PROPOSITION 5.1. *Let* λ_j *be the U-eigenvalues of* A *for* $j = ivec(\boldsymbol{j}, \mathcal{J})$*. For an unforced MLTI system, the equilibrium point* $\mathsf{X} = \mathsf{O}$ *is*

1. *stable if and only if* $|\lambda_j| \leq 1$ *for all* $j = 1, 2, \dots, \Pi_{\mathcal{J}}$*; for those equal to* 1*, its algebraic and geometry multiplicities must be equal;*
2. *asymptotically stable if* $|\lambda_j| < 1$ *for all* $j = 1, 2, \dots, \Pi_{\mathcal{J}}$*;*
3. *unstable if* $|\lambda_j| > 1$ *for some* $j = 1, 2, \dots, \Pi_{\mathcal{J}}$*.*

[1]Big O notation: $f(x) = \mathcal{O}(g(x))$ as $x \to \infty$ if and only if there exist a positive real number M and a real number x_0 such that $|f(x)| \leq Mg(x)$ for all $x \geq x_0$.

Downloaded 02/25/21 to 141.214.17.120. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/25/21 to 141.214.17.120. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

Proof. We only focus on the case when A has a full set of U-eigentensors. It follows from Propositions 3.5 and 4.5 that $\mathsf{A}^k = \sum_{j_1=1}^{J_1}\cdots\sum_{j_N=1}^{J_N}\lambda_j^k\mathsf{W}_{j_1j_1\dots j_Nj_N}$ for some even-order square tensors $\mathsf{W}_{j_1j_1\dots j_Nj_N}$. Then the results follow immediately. ◻

COROLLARY 5.2. *Suppose that the HOSVD of* A *is provided with* n*-mode singular values. For an unforced MLTI system, the equilibrium point* $\mathsf{X}=\mathsf{O}$ *is asymptotically stable if the sum of the* n*-mode singular values square is less than one for any* n*.*

Proof. Without loss of generality, suppose that $n=1$. Based on Property 8 in [10], $\sum_{j=1}^{J_1}(\gamma_j^{(1)})^2 = \|\mathsf{A}\|^2 = \|\varphi(\mathsf{A})\|^2$. In addition, we know that the magnitude of the maximal eigenvalue of a matrix is less than or equal to its Frobenius norm. Hence, the proof follows immediately from Proposition 5.1. ◻

COROLLARY 5.3. *Suppose that the CPD of* A *is provided and its factor matrices* $\boldsymbol{A}^{(n)}$ *and* $\boldsymbol{A}^{(m)}$ *have all the column vectors orthonormal for at least one odd* n *and even* m*. For an unforced MLTI system, the equilibrium point* $\mathsf{X}=\mathsf{O}$ *is asymptotically stable if the first weight element* $\lambda_1<1$*.*

The proof of Corollary 5.3 is presented in Appendix B.3.

COROLLARY 5.4. *Suppose that the TTD of* $\tilde{\mathsf{A}} \in \mathbb{R}^{J_1\times\cdots\times J_N\times J_1\times\cdots\times J_N}$*, defined in Proposition* 4.8*, is provided with the first* $N-1$ *core tensors left-orthonormal and the last* N *core tensors right-orthonormal. For an unforced MLTI system, the equilibrium point* $\mathsf{X}=\mathsf{O}$ *is asymptotically stable if the largest singular value of* $\bar{\tilde{\boldsymbol{A}}}^{(N)}$ *is less than one, where* $\bar{\tilde{\boldsymbol{A}}}^{(N)} = \texttt{reshape}(\tilde{\mathsf{A}}^{(N)}, R_{N-1}J_N, R_N)$*.*

Proof. Based on the results of [23], the singular values of $\bar{\tilde{\mathbf{A}}}^{(N)}$ are the singular values of $\varphi(\mathsf{A})$. In addition, we know that the magnitude of the maximal eigenvalue of a matrix is less than or equal to its largest singular value. Hence, the proof follows immediately from Proposition 5.1. ◻

Remark. Although Proposition 5.1 offers strong stability results for unforced MLTI systems, computing U-eigenvalues usually requires an order of $\mathcal{O}(\Pi_{\mathcal{J}}^3)$ number of operations through tensor unfolding and matrix eigenvalue decomposition. To the contrary, Corollaries 5.2 to 5.4 can be used to determine the stability of MLTI systems much faster. In particular, if the TTD of $\tilde{\mathsf{A}}$ is provided, the time complexity of left- and right-orthonormalization is about $\mathcal{O}(NJR^3)$, assuming that $J_n=J$ for all n, and R is the average of the TT-ranks of $\tilde{\mathsf{A}}$ [35]. Moreover, truncating the TT-rank $\tilde{R}_N$ of $\tilde{\mathsf{A}}$ would not alter the largest singular values of $\bar{\tilde{\mathbf{A}}}^{(N)}$. Therefore, setting $\tilde{R}_N=1$ and computing the vector 2-norm of $\bar{\tilde{\mathbf{A}}}^{(N)}$ will return the largest singular value of $\varphi(\mathsf{A})$.

5.2. Reachability. In this and the following subsections, we introduce the definitions of reachability and observability for MLTI systems which are similar to analogous concepts for the LTI systems [4, 21, 46]. We then establish sufficient and necessary conditions for reachability and observability for MLTI systems.

DEFINITION 5.5. *The MLTI system* (3.2) *is said to be reachable on* $[t_0,t_1]$ *if, given any initial condition* X_0 *and any final state* X_1*, there exists a sequence of inputs* U_t *that steers the state of the system from* $\mathsf{X}_{t_0}=\mathsf{X}_0$ *to* $\mathsf{X}_{t_1}=\mathsf{X}_1$*.*

THEOREM 5.6. *The pair* (A,B) *is reachable on* $[t_0,t_1]$ *if and only if the reachability Gramian*

$$\mathsf{W}_r(t_0,t_1) = \sum_{t=t_0}^{t_1-1}\mathsf{A}^{t_1-t-1} * \mathsf{B} * \mathsf{B}^\top * (\mathsf{A}^\top)^{t_1-t-1}, \tag{5.1}$$

which is a weakly symmetric even-order square tensor, is U-positive definite.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/25/21 to 141.214.17.120. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

Proof. Suppose $\mathsf{W}_r(t_0,t_1)$ is U-positive definite, and let X_0 be the initial state and X_1 be the desired final state. Choose $\mathsf{U}_t = \mathsf{B}^\top * (\mathsf{A}^\top)^{t_1-t-1} * \mathsf{W}_r^{-1}(t_0,t_1) * \mathsf{V}$ for some constant tensor V. It follows from the solution of the system (3.2) that $\mathsf{X}_{t_1} = \mathsf{A}^{t_1} * \mathsf{X}_0 + \sum_{j=0}^{t_1-1} \mathsf{A}^{t_1-j-1} * \mathsf{B} * \mathsf{U}_t = \mathsf{A}^{t_1} * \mathsf{X}_0 + \mathsf{W}_r(t_0,t_1) * \mathsf{W}_r^{-1}(t_0,t_1) * \mathsf{V} = \mathsf{A}^{t_1} * \mathsf{X}_0 + \mathsf{V}$. Taking $\mathsf{V} = -\mathsf{A}^{t_1} * \mathsf{X}_0 + \mathsf{X}_1$, we have $\mathsf{X}_{t_1} = \mathsf{X}_1$.

We show the converse by contradiction. Suppose $\mathsf{W}_r(t_0,t_1)$ is not U-positive definite. Then there exists $\mathsf{X}_a \neq \mathsf{O}$ such that $\mathsf{X}_a^\top * \mathsf{A}^{t_1-t-1} * \mathsf{B} = \mathsf{O}$ for any t. Take $\mathsf{X}_1 = \mathsf{X}_a + \mathsf{A}^{t_1} * \mathsf{X}_0$, and it follows that $\mathsf{X}_a + \mathsf{A}^{t_1} * \mathsf{X}_0 = \mathsf{A}^{t_1} * \mathsf{X}_0 + \sum_{j=t_0}^{t_1-1} \mathsf{A}^{t_1-j-1} * \mathsf{B} * \mathsf{U}_j$. Multiplying from the left by $\mathsf{X}_a^\top$ yields $\mathsf{X}_a^\top * \mathsf{X}_a = \sum_{j=t_0}^{t_1-1} \mathsf{X}_a^\top * \mathsf{A}^{t_1-j-1} * \mathsf{B} * \mathsf{U}_j = 0$, which implies that $\mathsf{X}_a = \mathsf{O}$, a contradiction. ∎

Corollary 5.7. *If the reachability Gramian $\mathsf{W}_r(t_0,t_1)$ is not M-positive definite, the pair (A,B) is not reachable on $[t_0,t_1]$.*

Proof. The proof follows immediately from Proposition A.2 and Theorem 5.6. ∎

The reachability Gramian assesses to what degree each state is affected by an input [45]. The infinite horizon reachability Gramian can be computed from the tensor Lyapunov equation, which is defined by

$$\mathsf{W}_r - \mathsf{A} * \mathsf{W}_r * \mathsf{A}^\top = \mathsf{B} * \mathsf{B}^\top. \tag{5.2}$$

By the unfolding property, if the pair (A,B) is reachable over an infinite horizon and all the U-eigenvalues of A have magnitude less than one, one can show that there exists a unique weakly symmetric U-positive definite solution W_r. Solving the infinite horizon reachability Gramian from the tensor Lyapunov equation may be computationally intensive, so a tensor version of the Kalman rank condition is also provided.

Proposition 5.8. *The pair (A,B) is reachable if and only if the $J_1 \times J_1K_1 \times \cdots \times J_N \times J_NK_N$ even-order reachability tensor*

$$\mathcal{R} = \begin{vmatrix} \mathsf{B} & \mathsf{A} * \mathsf{B} & \dots & \mathsf{A}^{\Pi_{\mathcal{J}}-1} * \mathsf{B} \end{vmatrix} \tag{5.3}$$

spans $\mathbb{R}^{J_1\times J_2\times\cdots\times J_N}$. In other words, $\mathrm{rank}_U(\mathcal{R}) = \Pi_{\mathcal{J}}$.

Proof. The proof follows from Proposition A.3 and the generalized Cayley–Hamilton theorem discussed in the tensor eigenvalue problem. ∎

First, any choice of construction for the mode row block tensor works for the reachability tensor. Second, when $N = 1$, Proposition 5.8 simplifies to the famous Kalman rank condition for reachability of LTI systems. The following corollaries involving with HOSVD (multilinear ranks), CPD (CP rank), and TTD (TT-ranks) provide useful necessary or sufficient conditions for reachability of MLTI systems if the reachability tensor $\mathcal{R}$ is given in the HOSVD, CPD, or TTD format.

Corollary 5.9. *Given the reachability tensor $\mathcal{R}$ in* (5.3)*, if* $\mathrm{rank}_{2n-1}(\mathcal{R}) \neq J_n$ *for some n, the pair (A,B) is not reachable.*

Proof. The proof follows immediately from Propositions 4.6 and 5.8. ∎

Corollary 5.10. *Given the reachability tensor $\mathcal{R}$ in* (5.3)*, if the set of n-mode singular values of $\mathcal{R}$ obtained from the HOSVD contains zero for odd n, the pair (A,B) is not reachable.*

Proof. We know that the number of nonvanishing n-mode singular values is equal to its corresponding n-mode multilinear rank. Hence, the result follows immediately from Proposition 5.8 and Corollary 5.9. ∎

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/25/21 to 141.214.17.120. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

COROLLARY 5.11. *Given the reachability tensor* $\mathcal{R}$ *in* (5.3)*, if the CPD of* $\mathcal{R}$ *satisfies* (4.4) *with CP rank equal to* $\Pi_{\mathcal{J}}$*, the pair* (A,B) *is reachable. Conversely, if the pair* (A,B) *is reachable, then the CP rank of* $\mathcal{R}$ *is greater than or equal to* $\Pi_{\mathcal{J}}$*.*

Proof. The first part of the proof follows immediately from Propositions 4.7 and 5.8. The second part of the proof follows from the fact that the CP rank of a tensor is greater than or equal to its unfolding rank. ◻

COROLLARY 5.12. *Given the reachability tensor* $\mathcal{R}$ *in* (5.3)*, the pair* (A,B) *is reachable if and only if the* N*th optimal TT-rank of* $\tilde{\mathcal{R}} \in \mathbb{R}^{J_1\times\cdots\times J_N\times J_1K_1\times\cdots\times J_NK_N}$*, defined in Proposition* 4.8*, is equal to* $\Pi_{\mathcal{J}}$*.*

Proof. The proof follows immediately from Propositions 4.8 and 5.8. ◻

Remark. Finding the unfolding rank of the reachability tensor $\mathcal{R}$ through tensor unfolding and matrix QR decomposition is computationally expensive and has a $\mathcal{O}(\Pi_{\mathcal{J}}^3\Pi_{\mathcal{K}})$ time complexity. However, if the reachability tensor $\mathcal{R}$ is already given in the tensor decomposition format, computing the unfolding rank can be achieved efficiently based on Corollaries 5.10 to 5.12. Particularly, if the TTD of $\tilde{\mathcal{R}}$ is provided, we do not need any additional computation to obtain the unfolding rank.

5.3. Observability. The results of observability can be simply obtained by the duality principle, similarly to LTI systems.

DEFINITION 5.13. *The MLTI system* (3.1) *is said to be observable on* $[t_0,t_1]$ *if any initial state* $\mathsf{X}_{t_0}=\mathsf{X}_0$ *can be uniquely determined by* Y_t *on* $[t_0,t_1]$*.*

THEOREM 5.14. *The pair* (A,C) *is observable on* $[t_0,t_1]$ *if and only if the observability Gramian*

$$\mathsf{W}_o(t_0,t_1)=\sum_{t=t_0}^{t_1-1}(\mathsf{A}^\top)^{t-t_0}*\mathsf{C}^\top*\mathsf{C}*\mathsf{A}^{t-t_0}, \tag{5.4}$$

which is a weakly symmetric even-order square tensor, is U-positive definite.

Proof. Suppose that $\mathsf{W}_o(t_0,t_1)$ is U-positive definite, and let X_0 be the initial state such that $\mathsf{Y}_t=\mathsf{C}*\mathsf{X}_t=\mathsf{C}*\mathsf{A}^{t-t_0}*\mathsf{X}_0$ for any $t\in[t_0,t_1]$. Multiplying from the left by $(\mathsf{A}^\top)^{t-t_0}*\mathsf{C}^\top$ yields $(\mathsf{A}^\top)^{t-t_0}*\mathsf{C}^\top*\mathsf{Y}_t=(\mathsf{A}^\top)^{t-t_0}*\mathsf{C}^\top*\mathsf{C}*\mathsf{A}^{t-t_0}*\mathsf{X}_0$, which implies that $\sum_{t=t_0}^{t_1-1}(\mathsf{A}^\top)^{t-t_0}*\mathsf{C}^\top*\mathsf{Y}_t=\sum_{t=t_0}^{t_1-1}(\mathsf{A}^\top)^{t-t_0}*\mathsf{C}^\top*\mathsf{C}*\mathsf{A}^{t-t_0}*\mathsf{X}_0=\mathsf{W}_o(t_0,t_1)*\mathsf{X}_0$. Since $\mathsf{W}_o(t_0,t_1)$ is U-invertible, this equation has a unique solution $\mathsf{X}_0=\mathsf{W}_o^{-1}(t_0,t_1)\sum_{t=t_0}^{t_1-1}(\mathsf{A}^\top)^{t-t_0}*\mathsf{C}^\top*\mathsf{Y}_t$. Hence, (A,C) is observable on $[t_0,t_1]$.

Again, we show the converse by contradiction. Suppose that $\mathsf{W}_o(t_0,t_1)$ is not U-positive definite. Then there exists $\mathsf{X}_a\neq\mathsf{O}$ such that $\mathsf{C}*\mathsf{A}^{t-t_0}*\mathsf{X}_a=\mathsf{O}$ for any t. Take $\mathsf{X}_{t_0}=\mathsf{X}_0+\mathsf{X}_a$ for some initial state X_0. Then $\mathsf{Y}_t=\mathsf{C}*\mathsf{A}^{t-t_0}*\mathsf{X}_0+\mathsf{C}*\mathsf{A}^{t-t_0}*\mathsf{X}_a=\mathsf{C}*\mathsf{A}^{t-t_0}*\mathsf{X}_0$ for any $t\in[t_0,t_1]$. The initial states X_0 and $\mathsf{X}_0+\mathsf{X}_a$ produce the same output, which implies that (A,C) is not observable on $[t_0,t_1]$, a contradiction. ◻

COROLLARY 5.15. *If the observability Gramian* $\mathsf{W}_o(t_0,t_1)$ *is not M-positive definite, the pair* (A,C) *is not observable on* $[t_0,t_1]$*.*

The observability Gramian assesses to what degree each state affects future outputs [45]. The infinite horizon observability Gramian can be computed from the tensor Lyapunov equation defined by

$$\mathsf{A}^\top*\mathsf{W}_o*\mathsf{A}-\mathsf{W}_o=-\mathsf{C}^\top*\mathsf{C}. \tag{5.5}$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

If the pair (A, C) is observable and all the U-eigenvalues of A have magnitude less than one, then there exists a unique weakly symmetric U-positive definite solution W_o.

The following results can be proved similarly to those in subsection 5.2.

PROPOSITION 5.16. *The pair* (A, C) *is observable if and only if the* $I_1J_1 \times J_1 \times \cdots \times I_NJ_N \times J_N$ *even-order observability tensor*

$$\mathcal{O} = \begin{vmatrix} \mathsf{C} & \mathsf{C} * \mathsf{A} & \dots & \mathsf{C} * \mathsf{A}^{\Pi_{\mathcal{J}}-1} \end{vmatrix}^{\top} \tag{5.6}$$

spans $\mathbb{R}^{J_1 \times J_2 \times \cdots \times J_N}$*. In other words,* $\mathrm{rank}_U(\mathcal{O}) = \Pi_{\mathcal{J}}$*.*

COROLLARY 5.17. *Given the observability tensor* $\mathcal{O}$ *in* (5.6)*, if* $\mathrm{rank}_{2n}(\mathcal{O}) \neq J_n$ *for some* n*, the pair* (A, C) *is not observable.*

COROLLARY 5.18. *Given the observability tensor* $\mathcal{O}$ *in* (5.6)*, if the set of* n*-mode singular values of* $\mathcal{O}$ *obtained from the HOSVD contains zero for even* n*, the pair* (A, C) *is not observable.*

COROLLARY 5.19. *Given the observability tensor* $\mathcal{O}$ *in* (5.6)*, if the CPD of* $\mathcal{O}$ *satisfies* (4.4) *with CP rank equal to* $\Pi_{\mathcal{J}}$*, the pair* (A, C) *is observable. Conversely, if the pair* (A, C) *is observable, then the CP rank of* $\mathcal{O}$ *is greater than or equal to* $\Pi_{\mathcal{J}}$*.*

COROLLARY 5.20. *Given the observability tensor* $\mathcal{O}$ *in* (5.6)*, the pair* (A, C) *is observable if and only if the* N*th optimal TT-rank of* $\tilde{\mathcal{O}} \in \mathbb{R}^{I_1J_1 \times \cdots \times I_NJ_N \times J_1 \times \cdots \times J_N}$*, defined in Proposition* 4.8*, is equal to* $\Pi_{\mathcal{J}}$*.*

6. Model reduction for MLTI systems.

Based on the observations in section 5, it is more natural to manipulate MLTI systems in the tensor decomposition format so that all the computational advantages can be realized. This may also result in a more compressed representation.

6.1. Generalized CPD/TTD.

We first introduce the notion of generalized CPD/TTD for even-order paired tensors described in [17], in which the generalized CPD can also be viewed as the extension of the Kronecker rank approximation proposed by Van Loan [52]. Generalized CPD and TTD share a similar format and possess many analogous properties.

DEFINITION 6.1. *Given an even-order paired tensor* $\mathsf{A} \in \mathbb{R}^{J_1 \times I_1 \times \cdots \times J_N \times I_N}$*, the generalized CPD of* A *is given by*

$$\mathsf{A} = \sum_{r=1}^{R} \mathsf{A}^{(1)}_{r::} \circ \mathsf{A}^{(2)}_{r::} \circ \cdots \circ \mathsf{A}^{(N)}_{r::}, \tag{6.1}$$

where $\mathsf{A}^{(n)} \in \mathbb{R}^{R \times J_n \times I_n}$*. Extending Van Loan's definition* [52]*, we call the smallest* R *that achieves* (6.1) *the Kronecker rank of* A*.*

DEFINITION 6.2. *Given an even-order paired tensor* $\mathsf{A} \in \mathbb{R}^{J_1 \times I_1 \times \cdots \times J_N \times I_N}$*, the generalized TTD of* A *is given by*

$$\mathsf{A} = \sum_{r_0=1}^{R_0} \cdots \sum_{r_N=1}^{R_N} \mathsf{A}^{(1)}_{r_0::r_1} \circ \mathsf{A}^{(2)}_{r_1::r_2} \circ \cdots \circ \mathsf{A}^{(N)}_{r_{N-1}::r_N}, \tag{6.2}$$

where $\mathsf{A}^{(n)} \in \mathbb{R}^{R_{n-1} \times J_n \times I_n \times R_n}$*, and* $\{R_0, R_1, \dots, R_N\}$ *is the set of TT-ranks with* $R_0 = R_N = 1$*.*

Downloaded 02/25/21 to 141.214.17.120. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/25/21 to 141.214.17.120. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

Please refer to Appendix D.1 for the use of ":" notation. Given two even-order paired tensors in the generalized CPD/TTD format, the Einstein product (2.4) between the two can be computed without having to reconstruct the full tensors, i.e., keeping the original format [17]. The following proposition states the case for generalized CPD, which also applies to generalized TTD.

PROPOSITION 6.3. *Given two even-order paired tensors* $\mathsf{A} \in \mathbb{R}^{J_1\times K_1\times\cdots\times J_N\times K_N}$ *and* $\mathsf{B} \in \mathbb{R}^{K_1\times I_1\times\cdots\times K_N\times I_N}$ *in the format of* (6.1) *with Kronecker ranks R and S, respectively, the Einstein product* $\mathsf{A} * \mathsf{B}$ *is given by*

$$\mathsf{A} * \mathsf{B} = \sum_{t=1}^{T} \mathsf{E}_{t::}^{(1)} \circ \mathsf{E}_{t::}^{(2)} \circ \cdots \circ \mathsf{E}_{t::}^{(N)}, \tag{6.3}$$

where $\mathsf{E}_{t::}^{(n)} = \mathsf{A}_{r::}^{(n)} \mathsf{B}_{s::}^{(n)} \in \mathbb{R}^{J_n\times I_n}$, *and* $t = ivec(\{r,s\},\{R,S\})$ *with* $T = RS$.

Remark. The computational complexity of the Einstein product (6.3) is about $\mathcal{O}(NJ^3R^2)$, assuming that $J_n = I_n = K_n = J$ and $R = S$, which is much lower than $\mathcal{O}(J^{3N})$ from the Einstein product (2.4) if R is small.

The generalized CPD can be recovered from the standard CPD, and similarly for generalized TTD (see Algorithm C.1). The algorithm below is extended from the results by Van Loan [52] about the Kronecker rank approximation. Thus, one can easily obtain generalized CPD by using any technique for computing the standard CPD, including alternating least square (ALS) and modified ALS methods [24, 25].

Algorithm 6.1 Generalized CPD.

1: Given an even-order paired tensor $\mathsf{A} \in \mathbb{R}^{J_1\times I_1\times\cdots\times J_N\times I_N}$
2: Set $\check{\mathsf{A}} = \texttt{reshape}(\mathsf{A}, J_1I_1, J_2I_2, \dots, J_NI_N)$
3: Apply CPD algorithms on $\check{\mathsf{A}}$ such that $\check{\mathsf{A}} = \sum_{r=1}^{R} \lambda_r \mathbf{a}_1^{(r)} \circ \mathbf{a}_2^{(r)} \circ \cdots \circ \mathbf{a}_N^{(r)}$
4: Set $\mathsf{A}_{r::}^{(n)} = \lambda_r^{\frac{1}{N}} \texttt{reshape}(\mathbf{a}_n^{(r)}, J_n, I_n)$ for $n = 1, 2, \dots, N$
5: **return** Component tensors $\mathsf{A}^{(n)}$ for $n = 1, 2, \dots, N$

6.2. MLTI model reduction. The problem of model reduction has been studied heavily in the framework of classical control [13, 16, 34]. Methods including proper orthogonal decomposition (POD), scale-separation and averaging, and balanced truncation are applied in many engineering applications when dealing with high-dimensional linear/nonlinear systems [33]. As mentioned in section 3, using generalized CPD/TTD, we propose a new MLTI representation with fewer parameters. Note that we omit colons in each component tensor in this and the following subsections for simplicity (e.g., $\mathsf{A}_r^{(n)} = \mathsf{A}_{r::}^{(n)}$).

PROPOSITION 6.4. *The MLTI system* (3.2) *is equivalent to*

$$\begin{cases} \mathsf{X}_{t+1} = \displaystyle\sum_{r=1}^{R_1} \mathsf{X}_t \times \{\mathsf{A}_r^{(1)}, \dots, \mathsf{A}_r^{(N)}\} + \sum_{r=1}^{R_2} \mathsf{U}_t \times \{\mathsf{B}_r^{(1)}, \dots, \mathsf{B}_r^{(N)}\}, \\ \mathsf{Y}_t = \displaystyle\sum_{r=1}^{R_3} \mathsf{X}_t \times \{\mathsf{C}_r^{(1)}, \dots, \mathsf{C}_r^{(N)}\}, \end{cases} \tag{6.4}$$

where R_1, R_2, R_3 are the Kronecker ranks of the system, and $\mathsf{A}^{(n)} \in \mathbb{R}^{R_1\times J_n\times J_n}$, $\mathsf{B}^{(n)} \in \mathbb{R}^{R_2\times J_n\times K_n}$, *and* $\mathsf{C}^{(n)} \in \mathbb{R}^{R_3\times I_n\times J_n}$.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/25/21 to 141.214.17.120. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

Proof. The proof follows from Definition 6.1 and Proposition 3.3. ∎

Remark. The number of parameters of the MLTI system representation (6.4) is $R_1 \sum_{n=1}^{N} J_n^2 + R_2 \sum_{n=1}^{N} J_n K_n + R_3 \sum_{n=1}^{N} I_n J_n$. If the Kronecker ranks R_1, R_2, R_3 are relatively small, the total number of parameters is much less than that of the MLTI system model (3.2), which is given by $\prod_{n=1}^{N} J_n^2 + \prod_{n=1}^{N} J_n K_n + \prod_{n=1}^{N} I_n J_n$.

The MLTI system representation (6.4) is attractive for systems captured by sparse tensors or tensors with low Kronecker ranks where the two advantages, model reduction and computational efficiency, can be exploited. In particular, if A, B, and C are fourth-order paired tensors, the generalized CPDs are reduced to matrix SVD problems; see section 9.2 in [52]. However, there are two major drawbacks. First, for $N > 2$, there is no exact method to compute the Kronecker rank of a tensor [25], and truncating the rank does not ensure a good estimate. Second, current CPD algorithms are not numerically stable, which could result in ill-conditioning during the tensor decomposition and low rank approximation. One way to fix these issues is to replace generalized CPD by generalized TTD in (6.4), which takes a similar form. The algorithms for computing generalized TTD are numerically stable with unique optimal TT-ranks [35]. Most importantly, the TTD-based results obtained in section 5 can be realized in the form of (6.4). For example, we can determine the stability of MLTI systems from the TTD of $\tilde{\mathsf{A}}$ defined in Proposition 4.8, which can be obtained from the generalized TTD of A efficiently (similar to the remark in subsection 4.3).

Recall from section 3 that one can always convert the MLTI system (3.2) to an equivalent LTI form and then apply traditional model reduction approaches, e.g., balanced truncation. However, after converting to a matrix form, the low tensor rank structure exploited in the form of (6.4) may not be preserved, and thus low memory requirements cannot be achieved; see subsection 7.3. Furthermore, as shown in [7], the MLTI system (6.4) can be used to further develop a higher-order balanced truncation framework directly in the TTD format, which can provide additional computation and memory benefits over unfolding-based model reduction methods.

6.3. Explicit solution and stability. In addition to using tensor decompositions, we can exploit matrix calculations of the factor matrices $\mathsf{A}_r^{(n)}$ to develop notions including explicit solution and stability for the MLTI system (6.4) which also have lower computational costs compared to unfolding based methods.

PROPOSITION 6.5 (solution). *For an unforced MLTI system* $\mathsf{X}_{t+1} = \sum_{r=1}^{R_1} \mathsf{X}_t \times \{\mathsf{A}_r^{(1)}, \mathsf{A}_r^{(2)}, \dots, \mathsf{A}_r^{(N)}\}$, *the solution for* X *at time* k, *given initial condition* X_0, *is*

$$\mathsf{X}_k = \sum_{r=1}^{R_1^k} \mathsf{X}_0 \times \{\bar{\mathsf{A}}_r^{(1)}, \bar{\mathsf{A}}_r^{(2)}, \dots, \bar{\mathsf{A}}_r^{(N)}\}, \tag{6.5}$$

where $\bar{\mathsf{A}}_r^{(n)} = \mathsf{A}_{r_1}^{(n)} \mathsf{A}_{r_2}^{(n)} \dots \mathsf{A}_{r_k}^{(n)}$ *for* $r = ivec(\{r_1, r_2, \dots, r_k\}, \{R_1, R_1, \overset{k}{\dots}, R_1\})$.

Proof. The result follows immediately from Propositions 3.3 and 6.3. ∎

If the Kronecker rank R_1 is small, computing the explicit solution using (6.5) can be faster than using the Einstein product (2.4). Additionally, we can assess the stability of the unforced MLTI system of (6.4) based upon the Lyapunov approach.

PROPOSITION 6.6 (stability). *For the unforced MLTI system of* (6.4), *the equilibrium point* $\mathsf{X} = \mathsf{O}$ *is*

1. *stable (in the sense of Lyapunov) if* $\sum_{r=1}^{R_1} \prod_{n=1}^{N} \alpha_r^{(n)} = 1$;

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

2. *asymptotically stable (in the sense of Lyapunov) if* $\sum_{r=1}^{R_1} \prod_{n=1}^{N} \alpha_r^{(n)} < 1$, *where* $\alpha_r^{(n)}$ *denote the largest singular values of* $\mathsf{A}_r^{(n)}$.

Proof. Let's consider $V(\mathsf{X}) = \|\mathsf{X}\|$ as the Lyapunov function candidate, and let $f(\mathsf{X}) = \sum_{r=1}^{R_1} \mathsf{X} \times \{\mathsf{A}_r^{(1)}, \mathsf{A}_r^{(2)}, \dots, \mathsf{A}_r^{(N)}\}$. Then it follows that $V(f(\mathsf{X})) - V(\mathsf{X}) = \|\sum_{r=1}^{R_1} \mathsf{X} \times \{\mathsf{A}_r^{(1)}, \mathsf{A}_r^{(2)}, \dots, \mathsf{A}_r^{(N)}\}\| - \|\mathsf{X}\| \leq \sum_{r=1}^{R_1} \|\mathsf{X} \times \{\mathsf{A}_r^{(1)}, \mathsf{A}_r^{(2)}, \dots, \mathsf{A}_r^{(N)}\}\| - \|\mathsf{X}\| \leq (\sum_{r=1}^{R_1} \prod_{n=1}^{N} \alpha_r^{(n)} - 1)\|\mathsf{X}\|$, where the last inequality is based on Theorem 6 in [20]. Then the results follow immediately. ∎

Remark. The computational complexity of finding the matrix SVDs of the factor matrices can be estimated as $\mathcal{O}(NJ^3R_1)$, assuming that $J_n = J$ for all n.

When all the Kronecker ranks of the system $R_1 = R_2 = R_3 = 1$, the MLTI system (6.4) reduces to the Tucker product representation proposed by Surana, Patterson, and Rajapakse [51], which provides a more direct way to see that the Tucker-based MLTI model is only a special case of the MLTI system (3.2). Additionally, we can obtain stronger stability conditions for the unforced MLTI system in this case.

PROPOSITION 6.7 (stability). *Suppose that* $R_1 = 1$ *in* (6.4), *and* $\rho^{(n)}$ *are the spectral radii of* $\mathsf{A}_1^{(n)}$. *Then the unforced MLTI system of* (6.4) *is*

1. *stable if and only if* $\prod_{n=1}^{N} \rho^{(n)} \leq 1$, *and when* $\prod_{n=1}^{N} \rho^{(n)} = 1$, *their corresponding eigenvalues must have equal algebraic and geometric multiplicity;*
2. *asymptotically stable if* $\prod_{n=1}^{N} \rho^{(n)} < 1$;
3. *unstable if* $\prod_{n=1}^{N} \rho^{(n)} > 1$.

Proof. Based on (2.25) in [40], $\varphi(\mathsf{A}) = \mathsf{A}_1^{(N)} \otimes \mathsf{A}_1^{(N-1)} \otimes \cdots \otimes \mathsf{A}_1^{(1)}$, where the operation $\otimes$ denotes the Kronecker product. Moreover, the U-eigenvalues of A are equal to the products of eigenvalues of these component matrices $\mathsf{A}_1^{(n)}$, and the U-eigenvalues have equal algebraic and geometric multiplicities if and only if the factor eigenvalues have equal multiplicities [5]. Then the results follow immediately from Proposition 5.1. ∎

The above results including Propositions 6.4 to 6.6 can be reformulated by replacing the Kronecker rank summation by a series of TT-rank summations if A, B, and C are given in the generalized TTD format. Finally, the Kronecker product can be used to unfold the MLTI system (6.4) into an LTI system, i.e.,

$$\varphi(\mathsf{A}) = \sum_{r=1}^{R_1} \mathsf{A}_r^{(N)} \otimes \mathsf{A}_r^{(N-1)} \otimes \cdots \otimes \mathsf{A}_r^{(1)},$$

and similarly for $\varphi(\mathsf{B})$ and $\varphi(\mathsf{C})$. Hence, one can apply traditional control theory techniques to determine the MLTI system properties.

7. Numerical examples.

We provide four examples to illustrate the MLTI systems theory and model reduction using the techniques developed above. All the numerical examples presented were performed on a Linux machine with 8 GB RAM and a 2.4 GHz Intel Core i5 processor and were conducted in MATLAB R2018a with Tensor Toolbox 2.6 [1] and the TT toolbox [36].

7.1. Reachability and observability tensors.

In this example, we consider a simple single-input and single-output (SISO) system that is given by (3.1) with

$$\mathbf{A}_1 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0.2 & 0.5 & 0.8 \end{bmatrix}, \ \mathbf{A}_2 = \begin{bmatrix} 0 & 1 \\ 0.5 & 0 \end{bmatrix},$$

Downloaded 02/25/21 to 141.214.17.120. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

$$\mathbf{B}_1 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix},\ \mathbf{B}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$
$$\mathbf{C}_1 = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix},\ \mathbf{C}_2 = \begin{bmatrix} 1 & 0 \end{bmatrix},$$

and the states $\mathsf{X}_t \in \mathbb{R}^{3\times 2}$ are second-order tensors, i.e., matrices. The product of the two spectral radii of $\mathbf{A}_1$ and $\mathbf{A}_2$ is 0.9207, which implies that the system is asymptotically stable. In addition, the reachability and observability tensors based on (5.3) and (5.6) are given by

$$\mathcal{R}_{::11} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0.8 & 0 \end{bmatrix}, \qquad \mathcal{R}_{::21} = \begin{bmatrix} 0 & 0 & 0.5 \\ 0 & 0 & 0.4 \\ 1 & 0 & 0.57 \end{bmatrix},$$
$$\mathcal{R}_{::12} = \begin{bmatrix} 0.4 & 0 & 0.378 \\ 0.57 & 0 & 0.4849 \\ 0.756 & 0 & 0.6339 \end{bmatrix}, \quad \mathcal{R}_{::22} = \begin{bmatrix} 0 & 0.285 & 0 \\ 0 & 0.378 & 0 \\ 0 & 0.4849 & 0 \end{bmatrix}$$

and

$$\mathcal{O}_{::11} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0.5 \end{bmatrix}, \quad \mathcal{O}_{::21} = \begin{bmatrix} 0 & 0 & 0 \\ 0.04 & 0.15 & 0.285 \\ 0 & 0 & 0 \end{bmatrix},$$
$$\mathcal{O}_{::12} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad \mathcal{O}_{::22} = \begin{bmatrix} 0.1 & 0.25 & 0.4 \\ 0 & 0 & 0 \\ 0.057 & 0.1825 & 0.378 \end{bmatrix},$$

respectively. We compute the TTDs of the permuted tensors $\tilde{\mathcal{R}}$ and $\tilde{\mathcal{O}}$, respectively, and observe that $\text{rank}_U(\mathcal{R}) = 6$ and $\text{rank}_U(\mathcal{O}) = 6$. The system therefore is both reachable and observable.

7.2. Kronecker rank/TT-rank approximation. In this example, we consider a SISO MLTI system (3.2) with random sparse tensors $\mathsf{A} \in \mathbb{R}^{3\times 3\times 3\times 3\times 3\times 3}$, $\mathsf{B} \in \mathbb{R}^{3\times 3\times 3}$, and $\mathsf{C} \in \mathbb{R}^{3\times 3\times 3}$. According to Algorithm 6.1, we compute the generalized CPDs of A, B, and C using the tensor toolbox function `cp_als` with estimated Kronecker ranks $R_1 = 49$, $R_2 = 2$, and $R_3 = 2$, respectively; see "Generalized CPD" in Table 1. Note that the number of parameters in the system with full Kronecker ranks could be greater than that for the original system. We then fix R_2 and R_3 and gradually truncate R_1, since R_1 is most critical in determining the number of parameters of the reduced system. As we can see in the table, the number of parameters decreases dramatically as R_1 decreases. In order to assess the approximation error resulting from this truncation, we compute the relative error using the $\mathcal{H}$-infinity norm $\|\cdot\|_\infty$ between the full system and reduced system transfer functions based on (3.3). In particular, we find that when $R_1 = 10$, the reduced MLTI system is still close to the original system with $\mathcal{H}$-infinity norm relative error of 0.0888.

We repeat a similar process for TT-rank approximation through generalized TTD (see Algorithm C.1). The results are shown in the same table. We find that both generalized CPD and TTD can achieve efficient model reduction while keeping the approximation errors low. Generalized TTD in particular achieves better accuracy for a similar number of reduced parameters as compared to generalized CPD, but the latter can maintain a reasonable approximation error with an even lower number of parameters. The Bode diagrams for the reduced MLTI systems are shown in Figure 2.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/25/21 to 141.214.17.120. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

Downloaded 02/25/21 to 141.214.17.120. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

TABLE 1
Kronecker rank/TT-rank approximations of the MLTI system. We omit the first and last trivial TT-ranks in the generalized TTDs of A, B, *and* C.

	Reduced ranks	# Parameters	$\frac{\|\mathsf{G}_{\text{full}}-\mathsf{G}_{\text{red}}\|_\infty}{\|\mathsf{G}_{\text{full}}\|_\infty}$
Full system	-	783	-
Generalized CPD	$49, 2, 2$	1359	1.58×10^{-10}
	$20, 2, 2$	576	0.0223
	$10, 2, 2$	306	0.0888
Generalized TTD	$\{7, 8\}, \{1, 2\}, \{2, 2\}$	678	4.39×10^{-15}
	$\{7, 6\}, \{1, 2\}, \{2, 2\}$	534	0.0099
	$\{7, 5\}, \{1, 2\}, \{2, 2\}$	462	0.4911

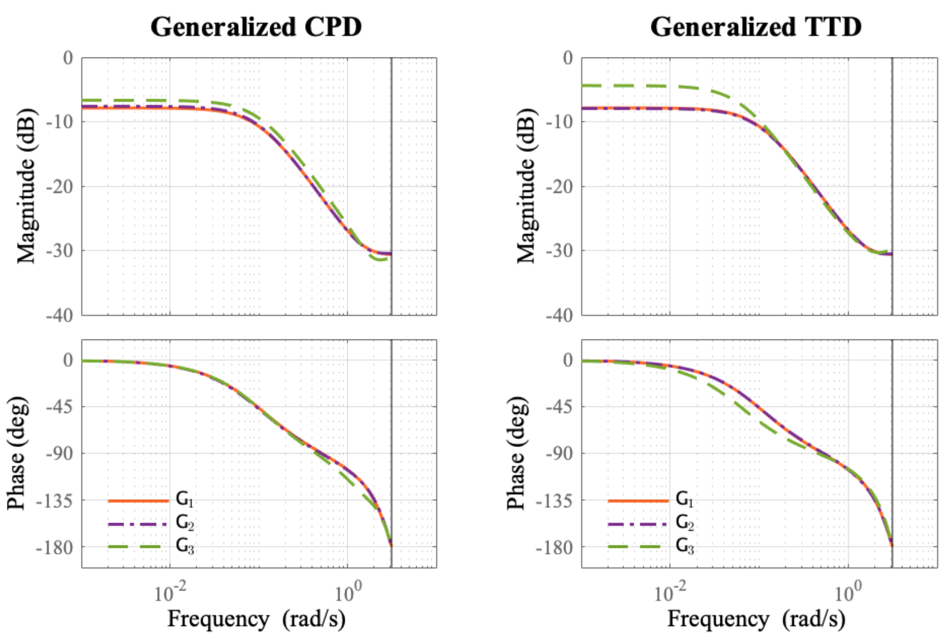

FIG. 2. Bode diagrams. G_1, G_2, *and* G_3 *are the transfer functions for the three reduced MLTI systems corresponding to Table* 1, *respectively. One may view* G_1 *as the transfer function of the original system. Since the function* `cp_als` *is not numerically stable, the results may not be exactly consistent with Table* 1 *for those obtained by generalized CPD.*

Note that in this example, we manually selected the truncation to study the tradeoff between number of parameters in the reduced system and the approximation error.

7.3. Memory consumption comparison. In this example, we consider a multiple-input and multiple-output (MIMO) MLTI system (3.2) with random even-order paired tensors $\mathsf{A}, \mathsf{B}, \mathsf{C} \in \mathbb{R}^{6\times 6\times 6\times 6\times 6\times 6}$ that possess low TT-ranks. We compare the memory consumptions of the generalized TTD-based representation (6.4) with the reduced models obtained from the unfolding-based balanced truncation. The results are shown in Table 2. One can clearly see that if the MLTI systems possess low TT-rank structure, the generalized TTD-based approach achieves much better accuracy for a similar number of parameters as compared to balanced truncation.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/25/21 to 141.214.17.120. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

TABLE 2

Memory consumption comparison between methods based on generalized TTD and balanced truncation methods. We reported the TT-ranks of A, B, *and* C *(ignoring the first and last trivial TT-ranks) and the number of singular values retained in the Hankel matrix during the balanced truncation.*

	Ranks	# Parameters	$\frac{\\|\mathsf{G}_{\text{full}}-\mathsf{G}_{\text{red}}\\|_\infty}{\\|\mathsf{G}_{\text{full}}\\|_\infty}$
Full system	-	139968	-
Generalized TTD	{6,6}, {6,6}, {6,6}	5184	3.98×10^{-15}
Balanced truncation	200	120000	0.0169
	100	30000	0.1001
	40	4800	0.2360

TABLE 3

Run time comparison between the TTD- and SVD-based methods in finding the largest singular value of $\varphi(\mathsf{A})$. *For the TTD-based method, the reported computational time includes conversion from the generalized TTD of* A *to the TTD of* $\tilde{\mathsf{A}}$ *and left- and right-orthonormalization.*

n	TTD(s)	SVD(s)	$\sigma_{\max}$	Relative error	Stability
6	0.0399	6.8551×10^{-4}	0.8082	1.3738×10^{-16}	asy. stable
8	0.0491	0.0439	0.9626	4.1523×10^{-15}	asy. stable
10	0.0591	0.4979	0.8645	3.8527×10^{-15}	asy. stable
12	0.0909	30.7663	0.8485	5.7573×10^{-15}	asy. stable
14	0.2623	2115.1	0.9984	1.3566×10^{-14}	asy. stable

7.4. Computational time comparison. In this example, we consider unforced MLTI systems (6.4) with random sparse even-order paired tensors $\mathsf{A} \in \mathbb{R}^{2\times 2\times \overset{n}{\cdots} \times 2\times 2}$ in the generalized TTD format such that $\varphi(\mathsf{A}) \in \mathbb{R}^{2^n \times 2^n}$. We compare the run time of Corollary 5.4 with the matrix SVD of $\varphi(\mathsf{A})$ for determining the stability of the systems. The results are shown in Table 3. When $n \geq 10$, the TTD-based method for finding the largest singular value of $\varphi(\mathsf{A})$ exhibits a significant time advantage compared to the matrix SVD–based method for which the time increases exponentially.

8. Discussion. While tensor unfolding to a matrix form provides the advantage of leveraging highly optimized matrix algebra libraries, in doing so, however, one may not be able to exploit the higher-order hidden patterns/structures, e.g., redundancy/correlations, present in the tensor. For instance, in the context of solving PDEs, Brazell et al. [3] found that higher-order tensor representations preserve low bandwidth, thereby keeping the computational cost and memory requirement low. As shown in subsections 7.3 and 7.4, TTD-based methods are more efficient in terms of computational speed and memory requirements compared to unfolding-based methods when the MLTI systems have low TT-rank structure. Although CPD typically offers better compression than TTD, the computation of CP rank is NP-hard, and the lower rank approximations can be ill-posed [11]. TTD is more suitable for numerical computations with well-developed TT-algebra [35]. Basic tensor operations such as addition, the Einstein product, the Frobenius norm, block tensor, solution to multilinear equations, and tensor pseudoinverse can be computed and maintained in the TTD format, without requiring full tensor representation. This can provide significant computational advantages in finding the reachability/observability tensors and associated unfolding ranks according to Corollaries 5.12 and 5.20, and in obtaining the solution of the tensor Lyapunov equations. For details, we refer the reader to [7] and the references therein.

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/25/21 to 141.214.17.120. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

Another line of approach is to exploit the isomorphism property to build algorithms directly in the full tensor format from existing methods. For example, Brazell et al. [3] proposed the higher-order biconjugate gradient (HOBG) method for solving multilinear systems which can be used for computing U-inverses and MLTI system transfer functions. Analogously, one can generalize the matrix-based Rayleigh quotient iteration method for computing U-eigenvalues (which can be used for determining MLTI system stability) directly in tensor form; see Algorithm C.2. However, the computational efficiency of this type of method remains to be investigated. Finally, one can explore hybrid methods by combining tensor algebra–based and matrix-based methods to provide the advantages of both approaches; see some examples in [7] in the context of MLTI model reduction. In the future, it would be worthwhile to systematically explore which of the above-mentioned approaches or combination thereof is best given the problem structure.

9. Conclusion. In this paper, we provided a comprehensive treatment of a newly introduced MLTI system representation using even-order paired tensors and the Einstein product. We established new results which enable one to express tensor unfolding–based stability, reachability, and observability criteria in terms of more standard notions of tensor ranks/decompositions. We introduced a generalized CPD/TTD-based model reduction framework which can significantly reduce the number of MLTI system parameters and realize the tensor decomposition–based methods. We also presented computational complexity analysis of our proposed framework and illustrated the benefits through numerical examples. In particular, TTD offers several computational advantages over CPD and HOSVD and provides a good representational choice for facilitating numerical computations associated with MLTI systems.

As mentioned in section 8, more work is required to fully realize the potential of tensor algebra–based computations for MLTI systems. It will also be worthwhile to develop theoretical and computational frameworks for observer and feedback control design for MLTI systems, and to apply these techniques in real world complex systems. One particular application we plan to investigate is that of cellular reprogramming, which involves introducing transcription factors as a control mechanism to transform one cell type to another. These systems naturally have matrix or tensor state spaces describing their genome-wide structure and gene expression [31, 44]. Such applications would also need to account for nonlinearity and stochasticity in tensor-based dynamical system representation and analysis framework and is an important direction for future research.

Appendix A. Additional tensor algebra.

A.1. M-positive definiteness/rank-one positive definiteness.

DEFINITION A.1. *An even-order square tensor* $\mathsf{A} \in \mathbb{R}^{J_1\times J_1\times\cdots\times J_N\times J_N}$ *is called M-positive definite if the multilinear functional*

$$\mathsf{A} \times \{\boldsymbol{x}_1^\top, \boldsymbol{x}_1^\top, \dots, \boldsymbol{x}_N^\top, \boldsymbol{x}_N^\top\} > 0 \tag{A.1}$$

for any nonzero vector $\boldsymbol{x}_n$*. If all* $\boldsymbol{x}_n$ *are equal,* A *is called rank-one positive definite.*

PROPOSITION A.2. *If an even-order square tensor* $\mathsf{A} \in \mathbb{R}^{J_1\times J_1\times\cdots\times J_N\times J_N}$ *is U-positive definite, it is M-positive definite. Moreover, if* $J_1 = J_2 = \cdots = J_N$*, U-positive definiteness also implies rank-one positive definiteness.*

Proof. By Lemma 3.2, it follows that $\mathsf{A} \times \{\mathbf{x}_1^\top, \mathbf{x}_1^\top, \dots, \mathbf{x}_N^\top, \mathbf{x}_N^\top\} = \mathsf{X}^\top * \mathsf{A} * \mathsf{X}$ for $\mathsf{X} = \mathbf{x}_1 \circ \mathbf{x}_2 \circ \cdots \circ \mathbf{x}_N$, i.e., X is a rank-one tensor. Moreover, if $J_1 = J_2 = \cdots = J_N$, M-

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/25/21 to 141.214.17.120. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

positive definiteness implies rank-one positive definiteness [39]. Therefore, the results follow immediately. ∎

A.2. Block tensor properties.

PROPOSITION A.3. *Let* $\mathsf{A}, \mathsf{B} \in \mathbb{R}^{J_1 \times I_1 \times \cdots \times J_N \times I_N}$ *and* $\mathsf{C}, \mathsf{D} \in \mathbb{R}^{I_1 \times K_1 \times \cdots \times I_N \times K_N}$. *Then the following properties hold:*

1. $\mathcal{P} * \begin{vmatrix} \mathsf{A} & \mathsf{B} \end{vmatrix}_n = \begin{vmatrix} \mathsf{P} * \mathsf{A} & \mathsf{P} * \mathsf{B} \end{vmatrix}_n$ *for any* $\mathsf{P} \in \mathbb{R}^{L_1 \times J_1 \times \cdots \times L_N \times J_N}$;
2. $\begin{vmatrix} \mathsf{C} \\ \mathsf{D} \end{vmatrix}_n * \mathsf{Q} = \begin{vmatrix} \mathsf{C} * \mathsf{Q} \\ \mathsf{D} * \mathsf{Q} \end{vmatrix}_n$ *for any* $\mathsf{Q} \in \mathbb{R}^{K_1 \times R_1 \times \cdots \times K_N \times R_N}$;
3. $\begin{vmatrix} \mathsf{A} & \mathsf{B} \end{vmatrix}_n * \begin{vmatrix} \mathsf{C} \\ \mathsf{D} \end{vmatrix}_n = \mathsf{A} * \mathsf{C} + \mathsf{B} * \mathsf{D}$.

Proof. The proof follows immediately from the definition of n-mode row/column block tensors and the Einstein product. ∎

PROPOSITION A.4. *Let* $\mathsf{A}, \mathsf{B} \in \mathbb{R}^{J_1 \times I_1 \times \cdots \times J_N \times I_N}$ *be two even-order paired tensors. Then* $\varphi(\begin{vmatrix} \mathsf{A} & \mathsf{B} \end{vmatrix}_n) = \begin{bmatrix} \varphi(\mathsf{A}) & \varphi(\mathsf{B}) \end{bmatrix} \boldsymbol{P}$, *where* $\boldsymbol{P}$ *is a column permutation matrix. In particular, when* $I_n = 1$ *for all* n *or* $n = N$, $\boldsymbol{P}$ *is the identity matrix.*

Proof. We consider the case for $N = 2$. Since the size of the odd modes of the block tensor remains the same, we only need to consider the even modes' unfolding transformation. When $n = 1$, the index mapping function for the even modes is

$$ivec(\mathbf{i}, \mathcal{I}) = i_1 + 2(i_2 - 1)I_1$$

for $i_1 = 1, 2, \ldots, 2I_1$. Based on the definition of n-mode row block tensors, the first I_1 columns of $\varphi(\begin{vmatrix} \mathsf{A} & \mathsf{B} \end{vmatrix}_1)$ are the vectorizations of $\mathsf{A}_{:i_1:i_2}$ for $i_1 = 1, 2, \ldots, I_1$ and $i_2 = 1$, and the second I_1 columns are the vectorizations of $\mathsf{B}_{:i_1:i_2}$ for $i_1 = I_1+1, I_1+2, \ldots, 2I_1$ and $i_2 = 1$. The alternating pattern continues for all I_2 pairs of I_1 columns. Hence, $\varphi(\begin{vmatrix} \mathsf{A} & \mathsf{B} \end{vmatrix}_1) = \begin{bmatrix} \varphi(\mathsf{A}) & \varphi(\mathsf{B}) \end{bmatrix} \mathbf{P}$ for some column permutation matrix $\mathbf{P}$. When $n = 2$, the index mapping function for the even modes is given by

$$ivec(\mathbf{i}, \mathcal{I}) = i_1 + (i_2 - 1)I_1$$

for $i_2 = 1, 2, \ldots, 2I_2$. Similarly, the first $I_1 I_2$ columns of $\varphi(\begin{vmatrix} \mathsf{A} & \mathsf{B} \end{vmatrix}_2)$ are the vectorizations of $\mathsf{A}_{:i_1:i_2}$ for $i_1 = 1, 2, \ldots, I_1$ and $i_2 = 1, 2, \ldots, I_2$, and the second $I_1 I_2$ columns are the vectorizations of $\mathsf{B}_{:i_1:i_2}$ for $i_1 = 1, 2, \ldots, I_1$ and $i_2 = I_2 + 1, I_2 + 2, \ldots, 2I_2$. Hence, $\varphi(\begin{vmatrix} \mathsf{A} & \mathsf{B} \end{vmatrix}_2) = \begin{bmatrix} \varphi(\mathsf{A}) & \varphi(\mathsf{B}) \end{bmatrix}$. A similar analysis can be used to prove the case for $N > 2$. Moreover, when $I_n = 1$ for all n, $\varphi(\mathsf{A})$ and $\varphi(\mathsf{B})$ are vectors, so no permutation needs to be considered. The proposition can be considered as a special case of Theorem 3.3 in [40]. ∎

Appendix B. Tensor ranks/decompositions proofs.

B.1. Proof of Proposition 4.6. Without loss of generality, assume that $\Pi_{\mathcal{I}} \leq \Pi_{\mathcal{J}}$ and $\text{rank}_U(\mathsf{A}) = \Pi_{\mathcal{I}}$. Then $\varphi(\mathsf{A})$ has $\Pi_{\mathcal{I}}$ linearly independent columns. The goal here is to construct a transformation from $\varphi(\mathsf{A})$ to $\mathbf{A}_{(2n)}^\top$, which can be easily visualized through the representation $(z, \mathbb{S})$ defined in (2.2). Let

$$\mathbb{S}_1 = \begin{pmatrix} 1 & 2 & \ldots & N & N+1 & N+2 & \ldots & 2N \\ 1 & 3 & \ldots & 2N-1 & 2 & 4 & \ldots & 2N \end{pmatrix},$$

$$\mathbb{S}_2 = \begin{pmatrix} 1 & 2 & \ldots & 2n-1 & 2n & \ldots & 2N-1 & 2N \\ 1 & 2 & \ldots & 2n-1 & 2n+1 & \ldots & 2N & 2n \end{pmatrix},$$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

$$\mathbb{S}_3 = \begin{pmatrix} 1 & 2 & \dots & N & N+1 & N+2 & \dots & N+n & N+n+1 & \dots & 2N \\ 1 & 3 & \dots & 2N-1 & 2n & 2 & \dots & 2n-2 & 2n+2 & \dots & 2N \end{pmatrix},$$

$$\mathbb{S}_4 = \begin{pmatrix} 1 & 2 & \dots & N & N+1 & \dots & N+n-1 & N+n & \dots & 2N-1 & 2N \\ 1 & 3 & \dots & 2N-1 & 2 & \dots & 2n-2 & 2n+2 & \dots & 2N & 2n \end{pmatrix}.$$

Clearly, $\varphi(\mathsf{A})$ and $\mathbf{A}_{(2n)}^\top$ can be represented by $(N, \mathbb{S}_1)$ and $(2N-1, \mathbb{S}_2)$, respectively. According to the definition of the index mapping function $ivec(\mathbf{i}, \mathcal{I})$, we first require a column permutation matrix $\mathbf{P}$ such that $\varphi(\mathsf{A})\mathbf{P}$ is represented by $(N, \mathbb{S}_3)$. Every I_n columns of $\varphi(\mathsf{A})\mathbf{P}$ correspond to the columns of $\mathbf{A}_{(2n)}^\top$. Collect each set of I_n columns of $\varphi(\mathsf{A})\mathbf{P}$ and stack them vertically to form a tall matrix $\tilde{\mathbf{A}}$ with the representation $(2N-1, \mathbb{S}_4)$. Since the columns of $\varphi(\mathsf{A})\mathbf{P}$ are linearly independent, $\operatorname{rank}(\tilde{\mathbf{A}}) = I_n$. Finally, according to the definition of the index mapping function $ivec(\mathbf{j}, \mathcal{J})$, we require a row permutation matrix $\mathbf{Q}$ such that $\mathbf{Q}\tilde{\mathbf{A}} = \mathbf{A}_{(2n)}^\top$. Hence, $\operatorname{rank}_{2n}(\mathsf{A}) = \operatorname{rank}(\mathbf{A}_{(2n)}^\top) = I_n$. Note that the converse of the statement is incorrect.

B.2. Proof of Proposition 4.7. In order to prove Proposition 4.7, we need to introduce the concept of the Khatri–Rao product.

DEFINITION B.1. *Given two matrices $\boldsymbol{A} \in \mathbb{R}^{J\times I}$ and $\boldsymbol{B} \in \mathbb{R}^{K\times I}$, the Khatri–Rao product, denoted by $\boldsymbol{A} \odot \boldsymbol{B}$, results in a $JK \times I$ matrix:*

$$\boldsymbol{A} \odot \boldsymbol{B} = \begin{bmatrix} \boldsymbol{a}_1 \otimes \boldsymbol{b}_1 & \boldsymbol{a}_2 \otimes \boldsymbol{b}_2 & \dots & \boldsymbol{a}_I \otimes \boldsymbol{b}_I \end{bmatrix},$$

where $\otimes$ denotes the Kronecker product, and $\boldsymbol{a}_n$ and $\boldsymbol{b}_n$ are the column vectors of $\boldsymbol{A}$ and $\boldsymbol{B}$, respectively.

The following lemma, provided by Sidiropoulos et al. [48, 49] gives some properties of rank and k-rank of the Khatri–Rao product $\mathbf{A} \odot \mathbf{B}$.

LEMMA B.2. *Given two matrices $\boldsymbol{A} \in \mathbb{R}^{J\times R}, \boldsymbol{B} \in \mathbb{R}^{I\times R}$, the Khatri–Rao product $\boldsymbol{A} \odot \boldsymbol{B}$ has column rank R if $k_{\boldsymbol{A}} + k_{\boldsymbol{B}} \geq R+1$ for $k_{\boldsymbol{A}}, k_{\boldsymbol{B}} \geq 1$. Moreover, $k_{\boldsymbol{A}\odot\boldsymbol{B}} \geq \min\{k_{\boldsymbol{A}} + k_{\boldsymbol{B}} - 1, R\}$.*

PROPOSITION B.3. *Given matrices $\boldsymbol{A}^{(n)} \in \mathbb{R}^{J_n\times R}$, the Khatri–Rao product $\boldsymbol{A}^{(1)} \odot \boldsymbol{A}^{(2)} \odot \cdots \odot \boldsymbol{A}^{(N)}$ has column rank R if $\sum_{n=1}^N k_{\boldsymbol{A}^{(n)}} \geq R + N - 1$ for $k_{\boldsymbol{A}^{(n)}} \geq 1$.*

Proof. Suppose that $N = 3$. By Lemma B.2, the Khatri–Rao product $\mathbf{A}^{(1)} \odot \mathbf{A}^{(2)} \odot \mathbf{A}^{(3)}$ has full column rank R if $k_{\mathbf{A}^{(1)}\odot\mathbf{A}^{(2)}} + k_{\mathbf{A}^{(3)}} \geq R+1$. Since we know that $k_{\mathbf{A}\odot\mathbf{B}} \geq \min\{k_{\mathbf{A}} + k_{\mathbf{B}} - 1, R\}$, the above inequality can be satisfied if

$$\min\{k_{\mathbf{A}^{(1)}} + k_{\mathbf{A}^{(2)}} - 1, R\} + k_{\mathbf{A}^{(3)}} \geq R + 1.$$

When $k_{\mathbf{A}^{(1)}} + k_{\mathbf{A}^{(2)}} > R+1$, the condition is reduced to $k_{\mathbf{A}^{(3)}} \geq 1$, and when $k_{\mathbf{A}^{(1)}} + k_{\mathbf{A}^{(2)}} \leq R+1$, the condition becomes $k_{\mathbf{A}^{(1)}} + k_{\mathbf{A}^{(2)}} + k_{\mathbf{A}^{(3)}} \geq R+2$. Therefore, the Khatri–Rao product $\mathbf{A}^{(1)} \odot \mathbf{A}^{(2)} \odot \mathbf{A}^{(3)}$ has full column rank R if $k_{\mathbf{A}^{(1)}} + k_{\mathbf{A}^{(2)}} + k_{\mathbf{A}^{(3)}} \geq R+2$. The result can be easily extended to $n = N$ using the same approach. ∎

Now, we can prove Proposition 4.7. Suppose that A has the CPD format (4.2) with CP rank equal to R. Applying the unfolding transformation φ yields

$$\varphi(\mathsf{A}) = (\mathbf{A}^{(2N-1)} \odot \cdots \odot \mathbf{A}^{(1)})\mathbf{S}(\mathbf{A}^{(2N)} \odot \cdots \odot \mathbf{A}^{(2)})^\top,$$

where $\mathbf{S} \in \mathbb{R}^{R\times R}$ is a diagonal matrix containing the weights of the CPD on its diagonal. By Proposition B.3, the two Khatri–Rao products $\mathbf{A}^{(2N-1)} \odot \cdots \odot \mathbf{A}^{(1)}$

Downloaded 02/25/21 to 141.214.17.120. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Downloaded 02/25/21 to 141.214.17.120. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

and $\mathbf{A}^{(2N)} \odot \cdots \odot \mathbf{A}^{(2)}$ have full column rank R if the two conditions $\sum_{n=1:2}^{2N} k_{\mathbf{A}^{(n)}} \geq R+N-1$ and $\sum_{n=2:2}^{2N} k_{\mathbf{A}^{(n)}} \geq R+N-1$ are satisfied. Hence, $\operatorname{rank}_U(\mathsf{A}) = R$. Note that we do not require the CPD of A to be unique in the statement.

B.3. Proof of Corollary 5.3. The proof is formulated similarly to the one above. We need to use the properties of the Khatri–Rao product.

LEMMA B.4. *Given matrices $\boldsymbol{A}^{(n)} \in \mathbb{R}^{J_n \times R}$, the Khatri–Rao product $\boldsymbol{A}^{(1)} \odot \boldsymbol{A}^{(2)} \odot \cdots \odot \boldsymbol{A}^{(N)}$ has all the column vectors orthogonal if at least one of $\boldsymbol{A}^{(n)}$ has all the column vectors orthogonal for $n = 1, 2, \ldots, N$.*

Proof. Suppose that $N = 2$. Based on the properties of the Kronecker product, for any $1 \leq n, m \leq R$, the inner product between $\mathbf{a}_n^{(1)} \otimes \mathbf{a}_n^{(2)}$ and $\mathbf{a}_m^{(1)} \otimes \mathbf{a}_m^{(2)}$ is given by

$$(\mathbf{a}_n^{(1)} \otimes \mathbf{a}_n^{(2)})^\top (\mathbf{a}_m^{(1)} \otimes \mathbf{a}_m^{(2)}) = ((\mathbf{a}_n^{(1)})^\top \mathbf{a}_m^{(1)})) \otimes ((\mathbf{a}_n^{(2)})^\top \mathbf{a}_m^{(2)})).$$

Therefore, if $\mathbf{A}^{(1)}$ or $\mathbf{A}^{(2)}$ has all column vectors orthogonal, then the inner product between $\mathbf{a}_n^{(1)} \otimes \mathbf{a}_n^{(2)}$ and $\mathbf{a}_m^{(1)} \otimes \mathbf{a}_m^{(2)}$ is zero for any n, m. ∎

Now we can prove Corollary 5.3. Suppose that A has the CPD format (4.2). Applying the unfolding transformation φ yields

$$\varphi(\mathsf{A}) = (\mathbf{A}^{(2N-1)} \odot \cdots \odot \mathbf{A}^{(1)}) \mathbf{S} (\mathbf{A}^{(2N)} \odot \cdots \odot \mathbf{A}^{(2)})^\top,$$

where $\mathbf{S} \in \mathbb{R}^{R \times R}$ is a diagonal matrix containing the weights of the CPD on its diagonal. By Lemma B.4, the two Khatri–Rao products $\mathbf{A}^{(2N-1)} \odot \cdots \odot \mathbf{A}^{(1)}$ and $\mathbf{A}^{(2N)} \odot \cdots \odot \mathbf{A}^{(2)}$ have all the column vectors orthonormal if $\mathbf{A}^{(n)}$ and $\mathbf{A}^{(m)}$ have all the column vectors orthonormal for at least one odd n and even m. Thus, λ_1 will be the largest singular value of $\varphi(\mathsf{A})$. In addition, we know that the magnitude of the maximal eigenvalue of a matrix is less than or equal to its largest singular value. Hence, the proof follows immediately from Proposition 5.1. Note that there is one special case when the CPD uniqueness condition fails, i.e., $\sum_{n=1}^{2N} k_{\mathbf{A}^{(n)}} = 2R+2N-2$. However, different CPDs, satisfying the orthonormal condition, correspond to the same matrix SVD under φ up to some orthogonal transformations.

Appendix C. Numerical algorithms.

Algorithm C.1 Generalized TTD.

1: Given an even-order paired tensor $\mathsf{A} \in \mathbb{R}^{J_1 \times I_1 \times \cdots \times J_N \times I_N}$
2: Set $\check{\mathsf{A}} = \texttt{reshape}(\mathsf{A}, J_1 I_1, J_2 I_2, \ldots, J_N I_N)$
3: Apply the standard TTD algorithm on $\check{\mathsf{A}}$ such that

$$\check{\mathsf{A}} = \sum_{r_0=1}^{R_0} \cdots \sum_{r_N=1}^{R_N} \check{\mathsf{A}}^{(1)}_{r_0:r_1} \circ \check{\mathsf{A}}^{(2)}_{r_1:r_2} \circ \cdots \circ \check{\mathsf{A}}^{(N)}_{r_{N-1}:r_N}$$

4: Set $\mathsf{A}^{(n)}_{r_{n-1}::r_n} = \texttt{reshape}(\check{\mathsf{A}}^{(n)}_{r_{n-1}:r_n}, J_n, I_n)$ for $n = 1, 2, \ldots, N$
5: **return** Component tensors $\mathsf{A}^{(n)}$ for $n = 1, 2, \ldots, N$

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

Algorithm C.2 Higher-Order Rayleigh Quotient Iteration.

Given an even-order square tensor $\mathsf{A} \in \mathbb{R}^{J_1 \times J_1 \times \cdots \times J_N \times J_N}$
Initialize $\mathsf{X}_0 \in \mathbb{R}^{J_1 \times J_2 \times \cdots \times J_N}$ with $\|\mathsf{X}_0\| = 1$
Compute $\lambda_0 = \mathsf{X}_0^\top * \mathsf{A} * \mathsf{X}_0$
for $k = 1, 2, \ldots$ **do**
  Solve $(\mathsf{A} - \lambda_{k-1}\mathsf{I}) * \mathsf{Y} = \mathsf{X}_{k-1}$ using HOBG proposed in [3]
  Set $\mathsf{X}_k = \frac{\mathsf{Y}}{\|\mathsf{Y}\|}$
  Compute $\lambda_k = \mathsf{X}_k^\top * \mathsf{A} * \mathsf{X}_k$
end for
return U-eigenvalue λ and U-eigentensor X

Appendix D. MATLAB functions.

D.1. The colon operator. The colon is one of the most useful operators in MATLAB and can create vectors and subscript arrays and specify for iterations. For our purpose, it acts as shorthand to include all subscripts in a particular array dimension [32]. For example, $\mathbf{A}_{:i}$ is equivalent to $\mathbf{A}_{ji}$ for all j. In the following, we represent TTD and generalized CPD and TTD in their componentwise forms:

1. (4.3) $\Leftrightarrow \mathsf{X}_{j_1 j_2 \ldots j_N} = \sum_{r_0=1}^{R_0} \cdots \sum_{r_N=1}^{R_N} \mathsf{X}^{(1)}_{r_0 j_1 r_1} \mathsf{X}^{(2)}_{r_1 j_2 r_2} \ldots \mathsf{X}^{(N)}_{r_{N-1} j_N r_N}$.
2. (6.1) $\Leftrightarrow \mathsf{A}_{j_1 i_1 \ldots j_N i_N} = \sum_{r=1}^{R} \mathsf{A}^{(1)}_{r j_1 i_1} \mathsf{A}^{(2)}_{r j_2 i_2} \ldots \mathsf{A}^{(N)}_{r j_N i_N}$.
3. (6.2) $\Leftrightarrow \mathsf{A}_{j_1 i_1 \ldots j_N i_N} = \sum_{r_0=1}^{R_0} \cdots \sum_{r_N=1}^{R_N} \mathsf{A}^{(1)}_{r_0 j_1 i_1 r_1} \mathsf{A}^{(2)}_{r_1 j_2 i_2 r_2} \ldots \mathsf{A}^{(N)}_{r_{N-1} j_N i_N r_N}$.

D.2. The `reshape` operator. The command $\mathsf{B} = \texttt{reshape}(\mathsf{A}, J_1, J_2, \ldots, J_N)$ reshapes a tensor A into a $J_1 \times J_2 \times \cdots \times J_N$ order tensor such that the number of elements in B matches the number of elements in A [32].

Acknowledgments. We would like to thank Dr. Frederick Leve at the Air Force Office of Scientific Research (AFOSR) for support and encouragement. We would also like to thank the two referees for their constructive comments, which led to a significant improvement of the paper.

REFERENCES

[1] B. W. Bader and T. G. Kolda, *MATLAB Tensor Toolbox*, Version 2.6, 2015.
[2] G. Bergqvist and E. G. Larsson, *The higher-order singular value decomposition: Theory and an application [lecture notes]*, IEEE Signal Processing Mag., 27 (2010), pp. 151–154.
[3] M. Brazell, N. Li, C. Navasca, and C. Tamon, *Solving multilinear systems via tensor inversion*, SIAM J. Matrix Anal. Appl., 34 (2013), pp. 542–570, https://doi.org/10.1137/100804577.
[4] R. W. Brockett, *Finite Dimensional Linear Systems*, SIAM, Philadelphia, 2015, https://doi.org/10.1137/1.9781611973884.
[5] B. J. Broxson, *The Kronecker Product*, Ph.D. thesis, University of North Florida, Jacksonville, FL, 2006.
[6] C. Chen, A. Surana, A. Bloch, and I. Rajapakse, *Multilinear time invariant system theory*, in 2019 Proceedings of the Conference on Control and Its Applications, SIAM, Philadelphia, pp. 118–125, https://doi.org/10.1137/1.9781611975758.18.
[7] C. Chen, A. Surana, A. Bloch, and I. Rajapakse, *Data-Driven Model Reduction for Multilinear Control Systems via Tensor Trains*, preprint, https://arxiv.org/abs/1912.03569, 2019.
[8] H. Chen, J. Chen, L. A. Muir, S. Ronquist, W. Meixner, M. Ljungman, T. Ried, S. Smale, and I. Rajapakse, *Functional organization of the human 4D nucleome*, Proc. Natl. Acad. Sci. USA, 112 (2015), pp. 8002–8007.

Downloaded 02/25/21 to 141.214.17.120. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

[9] L.-B. Cui, C. Chen, W. Li, and M. K. Ng, *An eigenvalue problem for even order tensors with its applications*, Linear Multilinear Algebra, 64 (2016), pp. 602–621.
[10] L. De Lathauwer, B. De Moor, and J. Vandewalle, *A multilinear singular value decomposition*, SIAM J. Matrix Anal. Appl., 21 (2000), pp. 1253–1278, https://doi.org/10.1137/S0895479896305696.
[11] V. de Silva and L.-H. Lim, *Tensor rank and the ill-posedness of the best low-rank approximation problem*, SIAM J. Matrix Anal. Appl., 30 (2008), pp. 1084–1127, https://doi.org/10.1137/06066518X.
[12] W. Ding, K. Liu, E. Belyaev, and F. Cheng, *Tensor-based linear dynamical systems for action recognition from 3D skeletons*, Pattern Recognit., 77 (2018), pp. 75–86.
[13] Y. Dolgin and E. Zeheb, *Model reduction of uncertain systems retaining the uncertainty structure*, Systems Control Lett., 54 (2005), pp. 771–779.
[14] S. Dolgov, B. Khoromskij, I. Oseledets, and D. Savostyanov, *Computation of extreme eigenvalues in higher dimensions using block tensor train format*, Comput. Phys. Commun., 185 (2014), pp. 1207–1216.
[15] A. Einstein, *The Foundation of the General Theory of Relativity*, Princeton University Press, Princeton, NJ, 2007.
[16] L. Fortuna, G. Nunnari, and A. Gallo, *Model Order Reduction Techniques with Applications in Electrical Engineering*, Springer, London, 2012.
[17] P. Gelß, *The Tensor-train Format and Its Applications: Modeling and Analysis of Chemical Reaction Networks, Catalytic Processes, Fluid Flows, and Brownian Dynamics*, Ph.D. Thesis, Freie Universität Berlin, Berlin, Germany, 2017.
[18] Z.-H. He, C. Navasca, and Q.-W. Wang, *Tensor Decompositions and Tensor Equations over Quaternion Algebra*, preprint, https://arxiv.org/abs/1710.07552, 2017.
[19] Z. Huang and L. Qi, *Positive definiteness of paired symmetric tensors and elasticity tensors*, J. Comput. Appl. Math., 338 (2017), pp. 22–43.
[20] B. Jiang, F. Yang, and S. Zhang, *Tensor and its Tucker core: The invariance relationships*, Numer. Linear Algebra Appl., 24 (2017), e2086.
[21] T. Kailath, *Linear Systems*, Inform. Syst. Sci. Ser., Prentice Hall, Upper Saddle River, NJ, 1980.
[22] B. Khoromskij, *Tensor Numerical Methods in Scientific Computing*, De Gruyter, Berlin, Germany, 2018.
[23] S. Klus, P. Gelß, S. Peitz, and C. Schütte, *Tensor-based dynamic mode decomposition*, Nonlinearity, 31 (2018), pp. 3359–3380.
[24] T. G. Kolda, *Multilinear Operators for Higher-order Decompositions*, Sandia National Laboratories, 2006.
[25] T. G. Kolda and B. W. Bader, *Tensor decompositions and applications*, SIAM Rev., 51 (2009), pp. 455–500, https://doi.org/10.1137/07070111X.
[26] K. Kruppa, *Comparison of tensor decomposition methods for simulation of multilinear time-invariant systems with the MTI Toolbox*, IFAC-PapersOnLine, 50 (2017), pp. 5610–5615.
[27] J. B. Kruskal, *Three-way arrays: Rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics*, Linear Algebra Appl., 18 (1977), pp. 95–138.
[28] W. Lai, D. Rubin, E. Krempl, and D. Rubin, *Introduction to Continuum Mechanics*, Elsevier Science, Amsterdam, Netherlands, 2009.
[29] L.-H. Lim, *Singular values and eigenvalues of tensors: A variational approach*, in Proceedings of the 1st IEEE Int. Workshop Comput. Adv. Multi-Sensor Adapt. Process., 2005, pp. 129–132.
[30] M. Liang, B. Zheng, and R. Zhao, *Tensor inversion and its application to the tensor equations with Einstein product*, Linear Multilinear Algebra, 67 (2019), pp. 843–870.
[31] S. Liu, H. Chen, S. Ronquist, L. Seaman, N. Ceglia, W. Meixner, P.-Y. Chen, G. Higgins, P. Baldi, S. Smale, A. Hero, L. A. Muir, and I. Rajapakse, *Genome architecture mediates transcriptional control of human myogenic reprogramming*, Science, 6 (2018), pp. 232–246.
[32] MATLAB, 9.7.0.1190202 *(R2019b)*, The MathWorks Inc., Natick, MA, 2018.
[33] I. Mezic, *Spectral properties of dynamical systems, model reduction and decompositions*, Nonlinear Dyn., 41 (2005), pp. 309–325.
[34] G. Obinata and B. Anderson, *Model Reduction for Control System Design*, Commun. Control Engrg., Springer London, 2012.
[35] I. Oseledets, *Tensor-train decomposition*, SIAM J. Sci. Comput., 33 (2011), pp. 2295–2317, https://doi.org/10.1137/090752286.
[36] I. Oseledets, S. Dolgov, V. Kazeev, O. Lebedeva, and T. Mach, *TT-Toolbox*, version 2.2.2, 2014.

Downloaded 02/25/21 to 141.214.17.120. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.

[37] I. Oseledets and E. Tyrtyshnikov, *Breaking the curse of dimensionality, or how to use SVD in many dimensions*, SIAM J. Sci. Comput., 31 (2009), pp. 3744–3759, https://doi.org/10.1137/090748330.
[38] L. Qi, *Eigenvalues of a real supersymmetric tensor*, J. Symbolic Comput., 40 (2005), pp. 1302–1324.
[39] L. Qi, H.-H. Dai, and D. Han, *Conditions for strong ellipticity and M-eigenvalues*, Front. Math. China, 4 (2009), pp. 349–364.
[40] S. Ragnarsson and C. F. Van Loan, *Block tensor unfoldings*, SIAM J. Matrix Anal. Appl., 33 (2012), pp. 149–169, https://doi.org/10.1137/110820609.
[41] S. Ragnarsson and C. F. Van Loan, *Block tensors and symmetric embeddings*, Linear Algebra Appl., 438 (2013), pp. 853–874.
[42] I. Rajapakse and M. Groudine, *On emerging nuclear order*, J. Cell Biol., 192 (2011), pp. 711–721.
[43] M. Rogers, L. Li, and S. J. Russell, *Multilinear dynamical systems for tensor time series*, in Advances in Neural Information Processing Systems, Adv. Neural Inf. Process. Syst. 26, Curran Associates, Inc., 2013, pp. 2634–2642.
[44] S. Ronquist, G. Patterson, L. A. Muir, S. Lindsly, H. Chen, M. Brown, M. S. Wicha, A. Bloch, R. Brockett, and I. Rajapakse, *Algorithm for cellular reprogramming*, Proc. Natl. Acad. Sci. USA, 114 (2017), pp. 11832–11837.
[45] C. W. Rowley, *Model reduction for fluids, using balanced proper orthogonal decomposition*, Internat. J. Bifur. Chaos, 15 (2005), pp. 997–1013.
[46] W. J. Rugh, *Linear System Theory*, 2nd ed., Prentice Hall, Upper Saddle River, NJ, 1996.
[47] N. D. Sidiropoulos and R. Bro, *On the uniqueness of multilinear decomposition of N-way arrays*, J. Chemometrics, 14 (2000), pp. 229–239.
[48] N. D. Sidiropoulos, R. Bro, and G. B. Giannakis, *Parallel factor analysis in sensor array processing*, IEEE Trans. Signal Process., 48 (2000), pp. 2377–2388.
[49] A. Stegeman and N. D. Sidiropoulos, *On Kruskal's uniqueness condition for the CANDECOMP/PARAFAC decomposition*, Linear Algebra Appl., 420 (2007), pp. 540–552.
[50] L. Sun, B. Zheng, C. Bu, and Y. Wei, *Moore-Penrose inverse of tensors via Einstein product*, Linear Multilinear Algebra, 64 (2016), pp. 686–698.
[51] A. Surana, G. Patterson, and I. Rajapakse, *Dynamic tensor time series modeling and analysis*, in IEEE 55th Conf. Decis. Control, 2016, pp. 1637–1642.
[52] C. F. Van Loan, *Structured Matrix Problems from Tensors*, in Exploiting Hidden Structure in Matrix Computations: Algorithms and Applications, Lecture Notes in Math. 2173, Springer, Cham, 2016, pp. 1–63.

Downloaded 02/25/21 to 141.214.17.120. Redistribution subject to SIAM license or copyright; see https://epubs.siam.org/page/terms

Copyright © by SIAM. Unauthorized reproduction of this article is prohibited.