THE TENSOR T-FUNCTION: A DEFINITION FOR FUNCTIONS OF THIRD-ORDER TENSORS *

KATHRYN LUND†

Abstract. A definition for functions of multidimensional arrays is presented. The definition is valid for third-order tensors in the tensor t-product formalism and is therefore referred to as the “tensor t-function.” By making use of its connection to block circulant matrices, the tensor t-function is shown to have similar properties as matrix functions in a number of fundamental scenarios. To demonstrate the definition’s potential in applications, the notion of network communicability is generalized to third-order tensors and computed for a small-scale example via block Krylov methods for matrix functions.

Key words. block Krylov methods, block GMRES, block FOM, restarts, tensors, t-product, multidimensional arrays, matrix functions, functions of operators, block circulant matrices, network analysis

1. Introduction. Functions of matrices— that is, $f(A)$, where $f$ is a scalar function, and $A$ a square matrix— have applications in a number of fields. They emerge as measures of centrality and communicability in networks [9] and as exponential integrators in differential equations [15]. As high-dimensional analogues of matrices, tensors also play crucial roles in network analysis [6] and multidimensional differential equations [17]. A variety of decompositions and algorithms have been developed over the years to extract and understand properties of tensors [20]. A natural question is whether the notion of functions of tensors, defined in analogy to functions of matrices as a scalar function taking a tensor $A$ as its argument, could prove to be yet another useful tool for studying multidimensional data.

Unfortunately, the definition of such a notion is not nearly as straightforward for tensors as it is for matrices. For matrices, the definitions of integration, polynomials, eigendecompositions (ED), and singular value decompositions (SVD) are unique and well established throughout linear algebra, and all of these notions serve as building blocks for definitions of matrix functions, reducing to the same object under reasonable circumstances [2, 14]. Classical decompositions such as Tucker and CANDECOMP/PARAFAC (CP) generalize the SVD in some sense; but many other generalizations of ED and SVD also exist for tensors [8, 18, 20, 21, 22, 24, 25, 26]. Each decomposition is based on maintaining or extracting some inherent structures, which are distinct in high-order settings. That is, a tensor function definition based on the Tucker decomposition would produce a fundamentally different object compared to one based on the CP decomposition.

As a starting point, we propose a definition for functions of tensors based on one of these paradigms, the tensor t-product [5, 18, 19]. The beauty of such a definition is that it reduces to the $f(A)B$ problem, i.e., a function of a matrix acting on a block vector, for which new methods have lately been developed [1, 11, 31]. One can think of this object in two ways: 1) as a new application of matrix function theory, especially the $f(A)B$ problem; and 2) as a generalization of such theory to higher-order arrays. The definition we propose also behaves similarly to matrix functions in that many expected properties can be derived in an intuitive and analogous fashion.

*This work was supported in part by the U.S. National Science Foundation under grant DMS-1418882 and the U.S. Department of Energy under grant DE-SC 0016578.
†Temple University and Bergische Universität Wuppertal, (katlund@temple.edu, https://math.temple.edu/~katlund/).
This report proceeds as follows. We recapitulate matrix function definitions and properties in Section 1.1. Section 2 restates the tensor t-product framework and poses a definition for a the tensor t-function, a new definition for a tensor function within this framework. We also present statements and proofs of t-function properties in analogy to the core properties of matrix functions. A possible application for the tensor t-exponential as a generalized communicability measure is discussed in Section 3. In Section 4, we discuss possible methods for computing the tensor t-function, in particular block Krylov methods for matrix functions, and demonstrate the efficacy of these methods for the tensor t-exponential. We make concluding remarks in Section 5.

We make a brief comment on syntax and disambiguation: the phrase "tensor function" already has an established meaning in physics; see, e.g., [3, 4, 32]. The most precise phrase for our object of interest would be “a function of a multidimensional array,” in analogy to “a function of a matrix.” However, since combinations of prepositional phrases can be cumbersome in English, we risk compounding literature searches by resorting to the term “tensor function.”

1.1. Definitions of matrix functions. Following [12, 14, 27], we concern ourselves with the three main matrix function definitions, based on the Jordan canonical form, Hermite interpolating polynomials, and the Cauchy-Stieltjes integral form. In each case, the validity of the definition boils down to the differentiability of $f$ on the spectrum of $A$. When $f$ is analytic on the spectrum of $A$, all the definitions are equivalent, and we can switch between them freely.

Let $A \in \mathbb{C}^{n \times n}$ be a matrix with spectrum $\text{spec}(A) := \{\lambda_j\}_{j=1}^N$, where $N \leq n$ and the $\lambda_j$ are distinct. An $m \times m$ Jordan block $J_m(\lambda)$ of an eigenvalue $\lambda$ has the form

$$J_m(\lambda) = \begin{bmatrix} \lambda & 1 \\ & \ddots & \ddots \\ & & \ddots & 1 \\ & & & \lambda \end{bmatrix} \in \mathbb{C}^{m \times m}.$$

Suppose that $A$ has Jordan canonical form

$$A = XJX^{-1} = X^{-1} \text{diag}(J_{m_1}(\lambda_{j_1}), \ldots, J_{m_p}(\lambda_{j_p}))X, \quad (1.1)$$

with $p$ blocks of sizes $m_i$ such that $\sum_{i=1}^p m_i = n$, and where the values $\{\lambda_{j_k}\}_{k=1}^p \in \text{spec}(A)$. Note that eigenvalues may be repeated in the sequence $\{\lambda_{j_k}\}_{k=1}^p$. Let $n_j$ denote the index of $\lambda_j$, or the size of the largest Jordan block associated to $\lambda_j$.

A function is defined on the spectrum of $A$ if all the following values exist:

$$f^{(k)}(\lambda_j), \quad k = 0, \ldots, n_j - 1, \quad j = 1, \ldots, N.$$

**Definition 1.1.** Suppose $A \in \mathbb{C}^{n \times n}$ has Jordan form (1.1) and that $f$ is defined on the spectrum of $A$. Then we define

$$f(A) := Xf(J)X^{-1},$$
where \( f(J) := \operatorname{diag}(f(J_{m_1}(\lambda_{j_1})), \ldots, f(J_{m_p}(\lambda_{j_p}))) \), and

\[
f(J_{m_i}(\lambda_{j_k})) := \begin{bmatrix}
f(\lambda_{j_k}) & f'(\lambda_{j_k}) & \frac{f''(\lambda_{j_k})}{2!} & \cdots & \frac{f^{(n_{jk}-1)}(\lambda_{j_k})}{(n_{jk}-1)!} \\
0 & f(\lambda_{j_k}) & f'(\lambda_{j_k}) & \cdots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
0 & \cdots & \cdots & 0 & f(\lambda_{j_k})
\end{bmatrix} \in \mathbb{C}^{m_i \times m_i},
\]

Note that when \( A \) is diagonalizable with \( \operatorname{spec}(A) = \{\lambda_j\}_{j=1}^n \) (possibly no longer distinct), Definition 1.1 reduces to

\[
f(A) = X \operatorname{diag}(f(\lambda_1), \ldots, f(\lambda_n))X^{-1}.
\]

Matrix powers are well defined, so a scalar polynomial evaluated on a matrix is naturally defined. Given \( p(z) = \sum_{k=0}^m z^k c_k \), for some \( \{c_k\}_{k=1}^m \subset \mathbb{C} \), we have that \( p(A) := \sum_{k=1}^m A^k c_k \). Based on this, we can define non-polynomial functions of matrices by using again derivatives as we did in Definition 1.1.

**Definition 1.2.** Suppose that \( f \) is defined on \( \operatorname{spec}(A) \), and let \( p \) with \( \deg p \leq \sum_{j=1}^n n_j \) be the unique Hermite interpolating polynomial satisfying

\[
p^{(k)}(\lambda_j) = f^{(k)}(\lambda_j), \text{ for all } k = 0, \ldots, n_j - 1, \quad j = 1, \ldots, N.
\]

We then define \( f(A) := p(A) \).

**Theorem 1.3** (Theorem 1.3 from [14]). For polynomials \( p \) and \( q \) and \( A \in \mathbb{C}^{n \times n} \), \( p(A) = q(A) \) if and only if \( p \) and \( q \) take the same values on the spectrum of \( A \).

The proof follows by noting that the minimal polynomial of \( A \) i.e., the polynomial \( \psi \) of least degree such that \( \psi(A) = 0 \) divides \( p - q \), and consequences thereof.

Crucial for our methods and analysis is the Cauchy-Stieltjes integral definition.

**Definition 1.4.** Let \( \mathbb{D} \subset \mathbb{C} \) be a region, and suppose that \( f : \mathbb{D} \to \mathbb{C} \) is analytic with integral representation

\[
f(z) = \int_{\Gamma} \frac{g(t)}{t - z} \, dt, \quad z \in \mathbb{D},
\]

with a path \( \Gamma \subset \mathbb{C} \setminus \mathbb{D} \) and function \( g : \Gamma \to \mathbb{C} \). Further suppose that the spectrum of \( A \) is contained in \( \mathbb{C} \setminus \mathbb{D} \). Then we define

\[
f(A) := \int_{\Gamma} g(t)(tI - A)^{-1} \, dt.
\]

When \( f \) is analytic, \( g = \frac{1}{2\pi i}f \), and \( \Gamma \) is a contour enclosing the spectrum of \( A \), then Definition 1.2 reduces to the usual Cauchy integral definition.

Various matrix function properties will prove useful throughout our analysis. Their proofs follow by examining the polynomial and Jordan form definitions of matrix functions.

**Theorem 1.5** (Theorem 1.13 in [14]). Let \( A \in \mathbb{C}^{n \times n} \) and let \( f \) be defined on the spectrum of \( A \). Then

(i) \( f(A)A = Af(A) \);
(ii) \( f(A^T) = f(A)^T \);
(iii) \( f(XAX^{-1}) = Xf(A)X^{-1} \); and
(iv) \( f(\lambda) \in \operatorname{spec}(f(A)) \) for all \( \lambda \in \operatorname{spec}(A) \).
Fig. 2.1: Different views of a third-order tensor $A \in \mathbb{C}^{n_1 \times n_2 \times n_3}$. (a) column fibers: $A(:, j, k)$; (b) row fibers: $A(i, ;, k)$; (c) tube fibers: $A(i, j, ;)$; (d) horizontal slices: $A(i, ;, ;)$; (e) lateral slices: $A(, j, ;)$; (f) frontal slices: $A(, , k)$

2. A definition for tensor functions. We direct the reader now to Figure 2.1 for different “views” of a third-order tensor, which will be useful in visualizing the forthcoming concepts. We also make use of some notions from block matrices. Define the standard block unit vectors $\hat{E}_{np}^{n \times n} := \hat{e}_p^k \otimes I_n^{\times n}$, where $\hat{e}_p^k \in \mathbb{C}^{p \times n}$ is the vector of all zeros except for the $k$th entry, and $I_n^{\times n}$ is the identity in $\mathbb{C}^{n \times n}$. When the dimensions are clear from context, we drop the superscripts. See (2.1) for various ways of expressing $\hat{E}_{np}^{n \times n}$.

$$\hat{E}_{np}^{n \times n} = \begin{bmatrix} I_{n \times n} \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \otimes I_n^{\times n} = \text{unfold}(I_{n \times n \times p}), \quad (2.1)$$

where $\text{unfold}$ is defined shortly.

In [5, 18, 19], a new concept is proposed for multiplying third-order tensors, based on viewing a tensor as a stack of frontal slices (as in Figure 2.1(f)). We consider a tensor $A$ of size $m \times n \times p$ and $B$ of size $n \times s \times p$ and denote their frontal faces respectively as $A^{(k)}$ and $B^{(k)}$, $k = 1, \ldots, p$. We also define the operations $\text{bcirc}$, $\text{unfold}$, $\text{fold}$, as

$\text{bcirc}(A) := \begin{bmatrix} A^{(1)} & A^{(p)} & A^{(p-1)} & \ldots & A^{(2)} \\ A^{(2)} & A^{(1)} & A^{(p)} & \ldots & A^{(3)} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ A^{(p)} & A^{(p-1)} & \ldots & A^{(2)} & A^{(1)} \end{bmatrix}$, \quad (2.2)

$\text{unfold}(A) := \begin{bmatrix} A^{(1)} \\ A^{(2)} \\ \vdots \\ A^{(p)} \end{bmatrix}$, and $\text{fold}(\text{unfold}(A)) := A$.

The $t$-product of two tensors $A$ and $B$ is then given as

$A \ast B := \text{fold}(\text{bcirc}(A)\text{unfold}(B))$.

Note that the operators $\text{fold}$, $\text{unfold}$, and $\text{bcirc}$ are linear.

The notion of transposition is defined face-wise, i.e., $A^*$ is the $n \times m \times p$ tensor obtained by taking the conjugate transpose of each frontal slice of $A$ and then reversing the order of the second through $p$th transposed slices.
For tensors with \( n \times n \) square faces, there is a tensor identity \( I_{n \times n \times p} \in \mathbb{C}^{n \times n \times p} \), whose first frontal slice is the \( n \times n \) identity matrix and whose remaining frontal slices are all the zero matrix. Recall from (2.3) that \( \hat{E}^{np \times n} = \hat{e}_{1}^p \otimes I_n \); it follows that

\[
\hat{E}^{np \times n} = \text{unfold}(I_{n \times n \times p}). \tag{2.3}
\]

With \( I_{n \times n \times p} \), one can then define the notion of an inverse with respect to the t-product. Namely, \( A, B \in \mathbb{C}^{n \times n \times p} \) are inverses of each other if \( A \ast B = I_{n \times n \times p} \) and \( B \ast A = I_{n \times n \times p} \). The t-product formalism further gives rise to its own notion of polynomials, with powers of tensors defined as \( A^{j} := \underbrace{A \ast \cdots \ast A}_{j \text{ times}} \).

Assuming that \( A \in \mathbb{C}^{n \times n \times p} \) has diagonalizable faces, we can also define a tensor eigendecomposition. That is, we have that \( A^{(k)} = X^{(k)} D^{(k)} (X^{(k)})^{-1} \), for all \( k = 1, \ldots, p \), and define \( X \) and \( D \) to be the tensors whose faces are \( X^{(k)} \) and \( D^{(k)} \), respectively. Then

\[
A = X \ast D \ast X^{-1} \quad \text{and} \quad A \ast \hat{X}_{i} = \hat{X}_{i} \ast d_i, \tag{2.4}
\]

where \( \hat{X}_{i} \) are the \( n \times 1 \times p \) lateral slices of \( X \) (see Figure 2.1(e)) and \( d_i \) are the \( 1 \times 1 \times p \) tubal fibers of \( D \) (see Figure 2.1). We say that \( D \) is f-diagonal, i.e., that each of its frontal faces is a diagonal matrix.

The eigenvalue decomposition (2.4) is not unique. See [13] for an alternative circulant-based interpretation of third-order tensors, as well as a deeper exploration of a unique canonical eigendecomposition for tensors. Uniqueness, while useful, is not necessary for our development.

2.1. The tensor t-exponential. As motivation, we consider the solution to a multidimensional ordinary differential equation. Suppose that \( A \) has square frontal faces, i.e., that \( A \in \mathbb{C}^{n \times n \times p} \) and let \( B : [0, \infty) \to \mathbb{C}^{n \times n \times p} \) be an unknown function with \( B(0) \) given. With \( \frac{d}{dt} \) acting element-wise, we consider the differential equation

\[
\frac{dB}{dt} (t) = A \ast B(t). \tag{2.5}
\]

Unfolding both sides leads to

\[
\frac{d}{dt} \begin{bmatrix}
B^{(1)}(t) \\
\vdots \\
B^{(n)}(t)
\end{bmatrix} = \text{bcirc}(A) \begin{bmatrix}
B^{(1)}(t) \\
\vdots \\
B^{(n)}(t)
\end{bmatrix},
\]

whose solution can be expressed in terms of the matrix exponential as

\[
\begin{bmatrix}
B^{(1)}(t) \\
\vdots \\
B^{(n)}(t)
\end{bmatrix} = \exp(\text{bcirc}(A)t) \begin{bmatrix}
B^{(1)}(0) \\
\vdots \\
B^{(n)}(0)
\end{bmatrix}.
\]

Folding both sides again leads to the tensor t-exponential,

\[
B(t) = \text{fold}(\exp(At)\text{unfold}(B(0))) =: \exp(At) \ast B(0). \tag{2.6}
\]
2.2. The tensor t-function. Using the tensor t-exponential as inspiration, we can define a more general notion for the scalar function $f$ of a tensor $A \in \mathbb{C}^{n \times n \times p}$ multiplied by a tensor $B \in \mathbb{C}^{n \times n \times p}$ as

$$f(A) \ast B := \text{fold}(f(\text{bcirc}(A)) \cdot \text{unfold}(B)),$$

which we call the tensor t-function. Note that $f(\text{bcirc}(A)) \cdot \text{unfold}(B)$ is merely a matrix function times a block vector. If $B = I_{n \times n \times p}$, then by equation (2.3) the definition for $f(A)$ reduces to

$$f(A) := \text{fold}\left(f(\text{bcirc}(A))\hat{E}_{1}^{np \times n}\right).$$

But does the definition (2.7) behave “as expected” in common scenarios? To answer this question, we require some results on block circulant matrices and the tensor t-product.

**Theorem 2.1** (Theorem 5.6.5 in [7]). Suppose $A, B \in \mathbb{C}^{np \times np}$ are block circulant matrices with $n \times n$ blocks. Let \{\(\alpha_j\)\}_{j=1}^k be scalars. Then $A^T$, $A^*$, $\alpha_1 A + \alpha_2 B$, $AB$, $q(A) = \sum_{j=1}^k \alpha_j A^j$, and $A^{-1}$ (when it exists) are also block circulant.

**Remark 2.2.** From (2.2), we can see that any block circulant matrix $C \in \mathbb{C}^{np \times np}$ can be represented by its first column $C \hat{E}_{1}^{np \times n}$. Let $C \in \mathbb{C}^{n \times n \times p}$ be a tensor whose frontal faces are the block entries of $C \hat{E}_{1}^{np \times n}$. Then $C = \text{fold}(C \hat{E}_{1}^{np \times n})$.

**Lemma 2.3.** Let $A \in \mathbb{C}^{n \times n \times p}$ and $B \in \mathbb{C}^{n \times n \times p}$. Then

- (i) $\text{unfold}(A) = \text{bcirc}(A)\hat{E}_{1}^{np \times n}$;
- (ii) $\text{bcirc}(\text{fold}(\text{bcirc}(A)\hat{E}_{1}^{np \times n})) = \text{bcirc}(A)$;
- (iii) $\text{bcirc}(A \ast B) = \text{bcirc}(A)\text{bcirc}(B)$;
- (iv) $\text{bcirc}(A)^j = \text{bcirc}(A^j)$, for all $j = 0, 1, \ldots$; and
- (v) $(A \ast B)^* = B^* \ast A^*$.

**Proof.** We drop the superscripts on $\hat{E}_{1}^{np \times n}$ for ease of presentation. Parts (i) and (ii) follow from Remark (2.2). To prove part (iii), we note by part (i) that

$$\text{bcirc}(A \ast B) = \text{bcirc}(\text{fold}(\text{bcirc}(A)\text{unfold}(B)))$$

$$= \text{bcirc}(\text{fold}(\text{bcirc}(A)\text{bcirc}(B)\hat{E}_{1})).$$

Note that $\text{bcirc}(A)\text{bcirc}(B)$ is a block circulant matrix by Theorem 2.1. Then by part (ii),

$$\text{bcirc}(\text{fold}(\text{bcirc}(A)\text{bcirc}(B)\hat{E}_{1})) = \text{bcirc}(A)\text{bcirc}(B).$$

Part (iv) follows by induction on part (iii). Part (v) is the same as [19, Lemma 3.16].

Let $D$ be an $n \times n \times p$ f-diagonal tensor, i.e., a tensor whose $n \times n$ frontal slices are diagonal matrices. Alternatively, one can think of such a tensor as an $n \times n$ matrix nonzero tube fibers on the diagonal, and zero tube fibers everywhere else. (Reference Figure 2.1(c).) The following theorem summarizes the relationship between the block circulant of $D$ and those of its tube fibers.

**Theorem 2.4.** Let $D \in \mathbb{C}^{n \times n \times p}$ be f-diagonal, and let $\{d_i\}_{i=1}^n \subset \mathbb{C}^{1 \times 1 \times p}$ denote its diagonal tube fibers. Then the spectrum of $\text{bcirc}(D)$ is identical to the union of the spectra of $\text{bcirc}(d_i)$, $i = 1, \ldots n$. 


Proof. We begin by deriving an expression for $\text{bcirc}(\mathcal{D})$ in terms of the $p \times p$ circulant matrices $\text{bcirc}(d_i)$. Denote each slice as $D^{(k)}$, $k = 1, \ldots, p$, with diagonal entries denoted as $d_i^{(k)}$, for $i = 1, \ldots, n$; i.e.,

$$D^{(k)} = \begin{bmatrix} d_1^{(k)} \\ & \ddots \\ & & d_n^{(k)} \end{bmatrix}.$$ 

Then we can express $\text{bcirc}(\mathcal{D})$ as follows:

$$\text{bcirc}(\mathcal{D}) = \begin{bmatrix} D^{(1)} & D^{(p)} & \cdots & D^{(2)} \\ D^{(2)} & D^{(1)} & \ddots & \vdots \\ \vdots & \ddots & \ddots & D^{(p)} \\ D^{(p)} & \cdots & D^{(2)} & D^{(1)} \end{bmatrix}.$$ 

Collecting the highlighted elements, note that the block circulant of the first tube fiber is given as

$$\text{bcirc}(d_1) = \begin{bmatrix} d_1^{(1)} & d_1^{(p)} & \cdots & d_1^{(2)} \\ d_2^{(1)} & \ddots & \vdots & \vdots \\ \vdots & \ddots & d_1^{(1)} & d_1^{(p)} \\ d_n^{(1)} & \cdots & d_1^{(2)} & d_1^{(1)} \end{bmatrix}.$$ 

Defining

$$\hat{I}_1 := \begin{bmatrix} 1 \\ \vdots \\ 0 \end{bmatrix} \in \mathbb{C}^{n \times n},$$
it holds that

\[
bcirc(d_1) \otimes \hat{I}_1 = \begin{bmatrix}
\vdots & \vdots & & \vdots \\
\ddots & 0 & & 0 \\
0 & 0 & & \ddots \\
\vdots & \ddots & \ddots & \vdots \\
0 & \cdots & 0 & 0
\end{bmatrix}
\]

Noting the same pattern for each \(i = 1, \ldots, n\), it is not hard to see that

\[
bcirc(D) = \sum_{i=1}^{n} bcirc(d_i) \otimes \hat{I}_i, \quad (2.9)
\]

where \(\hat{I}_i \in \mathbb{C}^{n \times n}\) is zero everywhere except for the \(i\)th entry, which is one.

Recall from [7, Section 3.2] that a circulant matrix is unitarily diagonalizable by the discrete Fourier transform (DFT). That is, for a \(p \times p\) circulant matrix \(C\), and with \(F_p\) denoting the \(p \times p\) DFT, \(F_p^*CF_p = \Lambda\), where \(\Lambda \in \mathbb{C}^{p \times p}\) is diagonal. Since each \(bcirc(d_i)\) is a \(p \times p\) circulant matrix, there exists for each \(i = 1, \ldots, n\) a diagonal \(\Lambda_i \in \mathbb{C}^{p \times p}\) such that

\[
F_p bcirc(d_i) F_p^* = \Lambda_i. \quad (2.10)
\]

Additionally, recall from [16, Lemma 4.2.10] the following useful property of the Kronecker product for matrices \(A, B, C,\) and \(D\) such that the products \(AC\) and \(BD\) exist:

\[
(A \otimes B)(C \otimes D) = (AB) \times (CD) \quad (2.11)
\]

We consequently have that

\[
(F_p \otimes I_{n \times n})bcirc(D)(F_p^* \otimes I_{n \times n}) = (F_p \otimes I_{n \times n}) \left( \sum_{i=1}^{n} bcirc(d_i) \otimes \hat{I}_i \right) (F_p^* \otimes I_{n \times n})
\]

\[
= \sum_{i=1}^{n} (F_p \otimes I_{n \times n})bcirc(d_i) \otimes \hat{I}_i (F_p^* \otimes I_{n \times n})
\]

\[
= \sum_{i=1}^{n} (F_p bcirc(d_i)F_p^*) \otimes (I_{n \times n} \hat{I}_i I_{n \times n}), \text{ by (2.11)}
\]

\[
= \sum_{i=1}^{n} \Lambda_i \otimes \hat{I}_i, \text{ by (2.10)}.
\]
Noting that $F_p \otimes I_{n \times n}$ is unitary and that the matrix $\Lambda := \sum_{i=1}^{n} \Lambda_i \otimes \hat{I}_i$ is a diagonal matrix whose entries are precisely the diagonal entries of all the $\Lambda_i$ concludes the proof.

**Corollary 2.5.** Let $\mathcal{D} \in \mathbb{C}^{n \times n \times p}$ be $f$-diagonal, and let $\{d_i\}_{i=1}^{n} \subset \mathbb{C}^1 \times 1 \times p$ denote its diagonal tube fibers. Then a function $f$ being defined on the spectrum of $\text{bcirc}(\mathcal{D})$ is equivalent to $f$ being defined on the union of the spectra of $\text{bcirc}(d_i)$, $i = 1, \ldots, n$.

An immediate consequence of Theorem 2.4 is that a function $f$ being defined on the spectrum of $\text{bcirc}(\mathcal{D})$ is equivalent to $f$ being defined on the spectra of $\text{bcirc}(d_i)$, $i = 1, \ldots, n$. The interpolating polynomials for $f(\text{bcirc}(\mathcal{D}))$ and $f(\text{bcirc}(d_i))$, $i = 1, \ldots, n$, are also related.

The following theorem ensures that definition (2.7) is well defined when $f$ is a polynomial, when $\mathcal{A}$ and $\mathcal{B}$ are second-order tensors (i.e., matrices), and when $f$ is the inverse function.

**Theorem 2.6.** Let $\mathcal{A} \in \mathbb{C}^{n \times n \times p}$ and $\mathcal{B} \in \mathbb{C}^{n \times s \times p}$.

(i) If $f \equiv q$, where $q$ is a polynomial, then the tensor t-function definition (2.7) matches the polynomial notion in the t-product formalism, i.e.,

$$\text{fold}(q(\text{bcirc}(\mathcal{A})) \cdot \text{unfold}(\mathcal{B})) = \text{fold}(\text{bcirc}(q(\mathcal{A})) \cdot \text{unfold}(\mathcal{B})).$$

(ii) Let $q$ be the scalar polynomial guaranteed by Definition 1.2 so that $f(\text{bcirc}(\mathcal{A})) = q(\text{bcirc}(\mathcal{A}))$. Then $f(\mathcal{A}) \ast \mathcal{B} = q(\mathcal{A}) \ast \mathcal{B}$.

(iii) If $\mathcal{A}$ is a matrix and $\mathcal{B}$ a block vector (i.e., if $p = 1$), then $f(\mathcal{A}) \ast \mathcal{B}$ reduces to the usual matrix function definition.

(iv) If $f(z) = z^{-1}$, then $f(\mathcal{A}) \ast \mathcal{A} = \mathcal{A} \ast f(\mathcal{A}) = I_{n \times n \times p}$.

**Proof.** For part (i), let $q(z) = \sum_{j=1}^{m} c_j z^j$. Then by Lemma 2.3(iv) and the linearity of fold, we have that

$$\text{fold}(q(\text{bcirc}(\mathcal{A})) \cdot \text{unfold}(\mathcal{B})) = \text{fold}\left(\sum_{j=1}^{m} c_j \text{bcirc}(\mathcal{A}^j) \cdot \text{unfold}(\mathcal{B})\right)$$

$$= \sum_{j=1}^{m} c_j \text{fold}(\text{bcirc}(\mathcal{A}^j) \cdot \text{unfold}(\mathcal{B}))$$

$$= \sum_{j=1}^{m} c_j \text{bcirc}(\mathcal{A}^j) \ast \mathcal{B}$$

$$= \text{fold}(\text{bcirc}(q(\mathcal{A})) \cdot \text{unfold}(\mathcal{B})).$$

Part (ii) is a special case of part (i). As for part (iii), since $p = 1$, we have that $\text{fold}(\mathcal{A}) = \text{bcirc}(\mathcal{A}) = \mathcal{A} = \text{unfold}(\mathcal{A})$, and similarly for $\mathcal{B}$. Then the definition of $f(\mathcal{A}) \ast \mathcal{B}$ reduces immediately to the matrix function case. Part (iv) follows by carefully unwrapping the definition of $f(\mathcal{A})$:

$$f(\mathcal{A}) \ast \mathcal{A} = \text{fold}\left(\text{bcirc}(\mathcal{A})^{-1} \text{unfold}(\mathcal{A})\right)$$

$$= \text{fold}\left(\text{bcirc}(\mathcal{A})^{-1} \text{bcirc}(\mathcal{A}) \mathcal{E}_1^{n \times p \times n}\right), \text{by Lemma 2.3(i)}$$

$$= \text{fold}\left(\mathcal{E}_1^{n \times p \times n}\right) = I_{n \times n \times p}.$$
Likewise with the other product:

\[
A \ast f(A) = \text{fold}(\text{bcirc}(A) \text{unfold}(\text{fold}(\text{bcirc}(A)^{-1} \text{unfold}(I_{n \times n \times p}))))
\]

\[
= \text{fold}(\text{bcirc}(A) \text{bcirc}(A)^{-1} \tilde{E}_{1}^{n \times n})
\]

\[
= \text{fold}(\tilde{E}_{1}^{n \times n} ) = I_{n \times n \times p}.
\]

The definition (2.7) possesses generalized versions of many of the core properties of matrix functions.

**Theorem 2.7.** Let \( A \in \mathbb{C}^{n \times n \times p} \), and let \( f : \mathbb{C} \to \mathbb{C} \) be defined on a region in the complex plane containing the spectrum of \( \text{bcirc}(A) \). For part (iv), assume that \( A \) has an eigendecomposition as in equation (2.4), with \( A \ast \tilde{X}_{i} = D \ast \tilde{X}_{i} = \tilde{X}_{i} \ast d_{i}, \) \( i = 1, \ldots, n \). Then it holds that

(i) \( f(A) \) commutes with \( A \);

(ii) \( f(A^{*}) = f(A)^{*} \);

(iii) \( f(\mathcal{X} \ast A \ast \mathcal{X}^{-1}) = \mathcal{X} f(A) \mathcal{X}^{-1} \); and

(iv) \( f(D) \ast \tilde{X}_{i} = \tilde{X}_{i} \ast f(d_{i}) \), for all \( i = 1, \ldots, n \).

**Proof.** For parts (i)-(iii), it suffices by Theorem 2.6(ii) to show that the statements hold for \( f(z) = \sum_{j=1}^{m} c_{j} z^{j} \). Part (i) then follows immediately. To prove part (ii), we need only show that \( (A^{j})^{*} = (A^{*})^{j} \) for all \( j = 0, 1, \ldots \), which follows by induction from Lemma 2.3.(v). Part (iii) also follows inductively. The base cases \( j = 0, 1 \) clearly hold. Assume for some \( j = k \), \( (\mathcal{X} \ast A \ast \mathcal{X}^{-1})^{k+1} = \mathcal{X} \ast (A)^{k+1} \ast \mathcal{X}^{-1} \), and then note that

\[
(\mathcal{X} \ast A \ast \mathcal{X}^{-1})^{k+1} = (\mathcal{X} \ast A \ast \mathcal{X}^{-1})^{k} \ast (\mathcal{X} \ast A \ast \mathcal{X}^{-1})
\]

\[
= \mathcal{X} \ast (A)^{k} \ast \mathcal{X}^{-1} \ast \mathcal{X} \ast A \ast \mathcal{X}^{-1} = \mathcal{X} \ast (A)^{k+1} \ast \mathcal{X}^{-1}.
\]

For part (iv), we fix \( i \in \{i, \ldots, n\} \). By Corollary 2.5, \( f \) being defined on \( \text{spec}(\text{bcirc}(D)) \) implies that it is also defined on \( \text{spec}(\text{bcirc}(d_{i})) \). Let \( q \) and \( q_{i} \) be the polynomials guaranteed by Theorem 1.2 such that \( f(\text{bcirc}(D)) = q(\text{bcirc}(D)) \) and \( f(\text{bcirc}(d_{i})) = q_{i}(\text{bcirc}(d_{i})) \). By Theorem 2.4, \( \text{spec}(\text{bcirc}(d_{i})) \subset \text{spec}(\text{bcirc}(D)) \), so by Theorem 2.4, it follows that \( q_{i}(\text{bcirc}(d_{i})) = q(\text{bcirc}(d_{i})) \). Then it suffices to prove part (iv) for \( D \), \( j = 1, 0, \ldots \). The cases \( j = 0, 1 \) clearly hold, and we assume the statement holds for some \( j = k \geq 1 \). Then

\[
D^{k+1} \ast \tilde{X}_{i} = D \ast (D^{k} \ast \tilde{X}_{i}) = D \ast \tilde{X}_{i} \ast d_{i}^{k} = \tilde{X}_{i} \ast d_{i}^{k+1}.
\]

**Remark 2.8.** When \( A \) has an eigendecomposition \( \mathcal{X} \ast D \ast \mathcal{X}^{-1} \) as in (2.4), then by Theorems 2.4 and 2.7, an equivalent definition for \( f(A) \) is given as

\[
f(A) = \mathcal{X} \ast \begin{bmatrix} f(d_{1}) & \cdots & \cdot \cdot \cdot \cdot \cdot \cdot \end{bmatrix} \ast \mathcal{X}^{-1},
\]

where the inner matrix should be regarded three-dimensionally, with its elements being tube fibers (cf. Figure 2.1(c)).

**2.3. Block diagonalization and the discrete Fourier transform.** Per recommendations for tensor computations in [18, 19], we can reduce the computational
effort of computing $f(A) * B$ by taking advantage of the fact that $bcirc(A)$ can be block diagonalized by the discrete Fourier transform (DFT) along the tubal fibers of $A$. Let $F_p$ denote the DFT of size $p \times p$. Then we have that

$$(F_p \otimes I_n)bcirc(A)(F_p^\ast \otimes I_n) = \begin{bmatrix} D_1 & & \\ & D_2 & \\ & & \ddots \\ & & & D_p \end{bmatrix} =: D,$$

where $D_k$ are $n \times n$ matrices. Then by Theorem 1.5(iii),

$$f(bcirc(A)) = (F_p^\ast \otimes I_n)f(D)(F_p \otimes I_n).$$

3. Centrality and communicability of a third-order network. More precisely, we use the term network to denote an undirected, unweighted graph with $n$ nodes. The graph, and by extension, the network, can be represented by its adjacency matrix $A \in \mathbb{R}^{n \times n}$. The $ij$th entry of $A$ is 1 if nodes $i$ and $j$ are connected, and 0 otherwise. As a rule, a node is not connected to itself, so $A_{ii} = 0$. The centrality of the $i$th node is defined as $\exp(A)_{ii}$, while the communicability between nodes $i$ and $j$ is defined as $\exp(A)_{ij}$.

These notions can be extended to higher-order situations. Suppose we are concerned instead about triplets, instead of pairs, of nodes. Then it is possible to construct an adjacency tensor $\mathcal{A}$, where a 1 at entry $\mathcal{A}_{ijk}$ indicates that distinct nodes $i$, $j$, and $k$ are connected and 0 otherwise. Alternatively, it is not hard to imagine a time-dependent network stored as a tensor, where each frontal face corresponds to a sampling of the network at discrete times. In either situation, we could compute the communicability of a triple as $\exp(\mathcal{A})_{ijk}$, where $\exp(\mathcal{A})$ is our tensor $t$-exponential. Centrality for a node $i$ would thus be defined as $\exp(\mathcal{A})_{ii}$.

4. Computing the tensor $t$-function. While the tensor $t$-function itself poses a number of interesting questions for multilinear algebra, we also want to demonstrate that this object has potential utility in real-life applications. We therefore need methods for approximating $f(A) * B$ numerically. The t-eigendecomposition and t-Krylov methods of [18] are viable options, but a full eigendecomposition may be expensive to compute for large tensors, and crafting Krylov methods for tensor functions remains an open problem. A preliminary exploration indicates that the theory behind t-Krylov methods for $f(A) * B$ would be analogous to the generalized block Krylov framework of [11], in the sense that one has to think of third-order tensors as matrices over vectors in the same way that block matrices are treated in [11] as matrices over matrices. However, we can also simply use the block Krylov methods for matrix functions of [11] and the forthcoming thesis [23], since $f(A) * B$ effectively reduces to $f(bcirc(A))$unfold($B$), which is a matrix function times a block vector.

4.1. A generalized block Krylov framework. We recount here the comprehensive block framework from [11] and [23]. Let $\mathcal{S} \subset \mathbb{C}^{s \times s}$ be a $*$-subalgebra with identity.

**Definition 4.1.** A mapping $\langle \cdot, \cdot \rangle_S$ from $\mathbb{C}^{n \times s} \times \mathbb{C}^{n \times s}$ to $\mathcal{S}$ is called a block inner product onto $\mathcal{S}$ if it satisfies the following conditions for all $X, Y, Z \in \mathbb{C}^{n \times s}$ and $C \in \mathcal{S}$:

(i) $\langle X + Y, Z \rangle_S = \langle X, Z \rangle_S + \langle Y, Z \rangle_S$ and $\langle X, YC \rangle_S = \langle X, Y \rangle_S C$;
(ii) $\langle X, Y \rangle_S = \langle Y, X \rangle_S^*$.
(iii) $\langle X, X \rangle_S$ is positive definite if $X$ has full rank, and $\langle X, X \rangle_S = 0$ if and only if $X = 0$.

**Definition 4.2.** A mapping $N$ which maps all $X \in \mathbb{C}^{n \times s}$ with full rank on a matrix $N(X) \in S$ is called a scaling quotient if for all such $X$ there exists $Y \in \mathbb{C}^{n \times s}$ such that $X = YN(X)$ and $\langle Y, Y \rangle_S = I$.

**Definition 4.3.** Let $X, Y \in \mathbb{C}^{n \times s}$.

(i) $X, Y$ are $\langle \cdot, \cdot \rangle_S$-orthogonal, if $\langle X, Y \rangle_S = 0$.

(ii) $X$ is $\langle \cdot, \cdot \rangle_S$-normalized if $N(X) = I$.

(iii) $\{X_1, \ldots, X_m\} \subset \mathbb{C}^{n \times s}$ is $\langle \cdot, \cdot \rangle_S$-orthonormal if $\langle X_i, X_j \rangle_S = \delta_{ij} I$, where $\delta_{ij}$ is the Kronecker delta.

We say that a set of vectors $\{X_j\}_{j=1}^m \subset \mathbb{C}^{n \times s}$ S-spans a space $\mathcal{K} \subset \mathbb{C}^{n \times s}$ and write $\mathcal{K} = \text{span}_S\{X_j\}_{j=1}^m$, where

$$\text{span}_S\{X_j\}_{j=1}^m := \left\{ \sum_{j=1}^m X_j \Gamma_j : \Gamma_j \in S \text{ for all } j = 1, \ldots, m \right\}.$$ 

The set $\{X_j\}_{j=1}^m$ constitutes an $\langle \cdot, \cdot \rangle_S$-orthonormal basis for $\mathcal{K}$ if $m$ is the dimension of $\mathcal{K}$, $\mathcal{K} = \text{span}_S\{X_j\}_{j=1}^m$, and $\{X_j\}_{j=1}^m$ are orthonormal.

We define the mth block Krylov subspace for $A$ and $B$ as

$$\mathcal{K}_m^S(A, B) = \text{span}_S\{B, AB, \ldots, A^{m-1}B\}.$$ 

There exist many choices for $S$, $\langle \cdot, \cdot \rangle_S$, and $N$. We consider the classical and global choices:

$$
\begin{array}{c|cc}
S & \text{classical} & \text{global} \\
\hline
\langle X, Y \rangle_S & I_S \otimes \mathbb{C} & \text{Cl}_S \\
N(X) & \text{diag}(X^*Y) & \frac{1}{s} \text{trace}(X^*Y) I_s \\
& R, \text{ where } X = QR & s \|X\|_F I_s \\
\end{array}
$$

Algorithm 4.1 is the generalization of the Arnoldi procedure within this framework. We assume that Algorithm 4.1 runs to completion without breaking down, i.e., that we obtain

(i) a $\langle \cdot, \cdot \rangle_S$-orthonormal basis $\{V_k\}_{k=1}^{m+1} \subset \mathbb{C}^{n \times s}$, such that each $V_k$ has full rank and $\mathcal{K}_m^S(A, B) = \text{span}_S\{V_k\}_{k=1}^m$, and

(ii) a block upper Hessenberg matrix $H_m \in S^{m \times m}$ and $H_{m+1, m} \in S$, all satisfying the block Arnoldi relation

$$AV_m = V_m H_m + V_{m+1} H_{m+1, m} \hat{E}_m^*, \quad (4.1)$$

where $V_m = [V_1 | \ldots | V_m] \in \mathbb{C}^{n \times ms}$, and $(H_m)_{ij} = H_{ij}$.

The paper [11] also establishes theory for a block full orthogonalization method for functions of matrices (B(FOM)$^2$) with restarts, given by Algorithm 4.2.

A restarted block harmonic method for matrix functions like that of [10] is also possible; see the thesis [23]. The main idea is to replace $H_m^{(k)}$ in Algorithm 4.2 with $H_m^{(k)} + M^{(k)}$, where $M^{(k)} := H_m^{(k)} E_m H_{m+1, m}^* H_{m+1, m}^* E_m^*$, and to derive a corresponding expression for the error function $\Delta_m^{(k)}$. Analysis for this method is based on the techniques of [10, 11, 29, 30] and is contained in the thesis [23]. The B(FOM)$^2$ are known to converge for Stieltjes functions $f$ of block Hermitian positive definite, and the block harmonic methods are known to converge for Stieltjes $f$ on a subset.
Algorithm 4.1 Block Arnoldi

1: Given: $A, B, S, \langle \cdot, \cdot \rangle_S, N, m$
2: Compute $B = N(B)$ and $V_1 = BB^{-1}$
3: for $k = 1, \ldots, m$ do
4: Compute $W = AV_k$
5: for $j = 1, \ldots, k$ do
6: $H_{j,k} = \langle V_j, W \rangle_S$
7: $W = W - V_j H_{j,k}$
8: end for
9: Compute $H_{k+1,k} = N(W)$ and $V_{k+1} = WH_{k+1,k}^{-1}$
10: end for
11: return $B, V_m = [V_1 \ldots V_m], H_m = (H_{j,k})_{j,k=1}^m, V_{m+1},$ and $H_{m+1,m}$

Algorithm 4.2 B(FOM)$^2(m)$: block full orthogonalization method for functions of matrices with restarts

1: Given $f, A, B, S, \langle \cdot, \cdot \rangle_S, N, m, t, \text{tol}$
2: Run Algorithm 4.1 with inputs $A, B, S, \langle \cdot, \cdot \rangle_S, N, m$ and store $V_{m+1,1}, H_{m+1,1}^{(1)}$ and $B_{m+1,1}^{(1)}$
3: Compute and store $F_m^{(1)} = V_m^{(1)} f (H_m^{(1)}) \tilde{E}_1 B$
4: Compute and store $C_m^{(1)}(t) = H_{m+1,m}^{(1)} \tilde{E}_m^{(1)} (H_m^{(1)} + tI)^{-1} \tilde{E}_1 B^{(1)}$ to define $\Delta_m^{(1)}(z)$
5: for $k = 1, 2, \ldots, \text{until convergence}$ do
6: Run Algorithm 4.1 with inputs $A, V_{m+1}^{(k)}, S, \langle \cdot, \cdot \rangle_S, N, m$ and store $V_{m+1}^{(k+1)}$ in place of the previous basis
7: Compute $D_m^{(k)} := V_{m+1}^{(k)} \Delta_m^{(k)} (H_m^{(k+1)}) \circ \tilde{E}_1$, where $\Delta_m^{(k)}(z)$ is evaluated via quadrature
8: Compute $F_m^{(k+1)} := F_m^{(k)} + D_m^{(k)}$ and replace $F_m^{(k)}$
9: Compute $C_m^{(k+1)}(t) = H_{m+1,m}^{(k+1)} \tilde{E}_m^{(k+1)} (H_m^{(k+1)} + tI)^{-1} \tilde{E}_1 B^{(k+1)} C_m^{(k)}(t)$ and replace $C_m^{(k)}(t)$
10: end for
11: return $F_m^{(k+1)}$

of block positive real matrices. Regardless, many numerical examples indicate the methods' stability and applicability for other functions with Cauchy-Stieltjes integral representation and matrices.

4.2. The tensor $t$-exponential on a small third-order network. We take $A \in \mathbb{C}^{n \times n \times p}$ to be a tensor whose $p$ frontal faces are each adjacency matrices for an undirected, unweighted network. More specifically, the frontal faces of $A$ are symmetric, and the entries are binary. The sparsity structure of this tensor is given in Figure 4.1 for $n = p = 50$. Note that we must actually compute $\exp(A) * I = \text{fold}(\exp(b\text{circ}(A)) \tilde{E}_1)$ (see Definition (2.8)). With $n = p = 40$, this leads to a $1600 \times 1600$ matrix function times a $1600 \times 40$ block vector. The sparsity patterns of $b\text{circ}(A)$ and $b\text{circ}(D)$, where $D$ is determined by folding a block eigendecomposition of $b\text{circ}(A)$, are shown in Figure 4.2. The block matrix $b\text{circ}(D)$ is determined by
applying MATLAB’s fast Fourier transform to \( \text{bcirc}(A) \). Note that \( \text{bcirc}(A) \) is not symmetric, but it has a nice banded structure. It should also be noted that while the blocks of \( \text{bcirc}(D) \) appear to be structurally identical, they are not numerically equal.

Fig. 4.1: Sparsity structure for \( A \). Blue indicates that a face is closer to the “front” and pink farther to the “back”; see Figure 2.1(f) for how the faces are oriented.

Fig. 4.2: Sparsity patterns for block circulants

We compute \( \exp(A) \ast I \) with the standard and harmonic versions of Algorithms 4.2, both with the classical and global block inner products. The convergence behavior of each version is displayed in Figure 4.3. The restart cycle length is \( m = 15 \), and the error tolerance is 1e-12. Despite the pathological behavior known to occur with FOM-like methods acting on circulant-type matrices [28], the BFOM methods do not suffer from the block circulant matrices here. In fact, the BFOM methods converge just as well as the block harmonic methods. The methods based on \( \text{bcirc}(D) \) (case (A)) are only a little less accurate than those based on \( \text{bcirc}(A) \) (case (B)), and they require the same number of iterations.

5. Conclusion. The main purpose of this report is to establish a first notion for functions of multidimensional arrays. Our definition for the tensor t-function \( f(A) \ast B \) shows versatility and consistency, and our numerical results indicate that block Krylov methods can compute \( f(A) \ast B \) with few iterations and still achieve high accuracy. In particular, global B(FOM)\(^2\) shows promise, since it is more accurate than the global harmonic method and requires the same number of iterations to converge as both classical methods, which are computationally more expensive per cycle.
The second aim of this report is to invite fellow researchers to pursue the many open problems posed by this new definition and the concept of tensor functions. Key problems include finding applications for $f(A) * B$ in real-life scenarios and comparing our definition of communicability for a third-order network to existing network analysis tools.

Acknowledgments. The author would like to thank Misha Kilmer for useful conversations and the images used in Figure 2.1, as well as Andreas Frommer and Daniel B. Szyld for comments on Theorem 2.7.

REFERENCES

[1] A. H. AL-MOHY AND N. J. HIGHAM, Computing the action of the matrix exponential with an application to exponential integrators, SIAM Journal on Scientific Computing, 33 (2011), pp. 488–511, https://doi.org/10.1137/100788860.
[2] F. ARRIGO, M. BENZI, AND C. FENU, Computation of generalized matrix functions, SIAM Journal on Matrix Analysis and Applications, 37 (2016), pp. 836–860, https://doi.org/10.1137/15M1049634, https://arxiv.org/abs/1512.01446.
[3] J. BETTEN, Creep mechanics, Springer, Berlin, 3rd ed., 2008.
[4] J. P. BOEHLER, ed., Applications of Tensor Functions in Solid Mechanics, Springer, Wien, 1987.
[5] K. BRAMAN, Third-order tensors as linear operators on a space of matrices, Linear Algebra and Its Applications, 433 (2010), pp. 1241–1253, https://doi.org/10.1016/j.laa.2010.05.025, http://dx.doi.org/10.1016/j.laa.2010.05.025.
[6] A. CICHOCKI, Era of Big Data Processing: A New Approach via Tensor Networks and Tensor Decompositions, Tech. Report arXiv:1403.2048v4, 2014, http://arxiv.org/abs/1403.2048.
[7] P. J. DAVIS, Circulant Matrices, AMS Chelsea Publishing, Providence, 2nd ed., 2012.
[8] L. DE LAIHAUWER, B. DE MOOR, AND J. VANDEWALLE, A multilinear singular value decomposition, SIAM Journal on Matrix Analysis and Applications, 21 (2000), pp. 1253–1278, https://doi.org/10.1137/S0895479896305696, http://epubs.siam.org/doi/abs/10.1137/S0895479896305696.
[9] E. ESTRADA AND D. J. HIGHAM, Network properties revealed through matrix functions, SIAM Review, 52 (2010), pp. 696–714, https://doi.org/10.1137/090764070, http://link.aip.org/link/SIREAD/v52/i4/p696/s1&Agg=doi.
[10] A. FROMMER, S. GÜTTEL, AND M. SCHWEITZER, Convergence of restarted Krylov subspace methods for Stieltjes functions of matrices, SIAM Journal on Matrix Analysis and Applications, 35 (2014), pp. 1602–1624, https://doi.org/10.1137/140973463.
[11] A. FROMMER, K. LUND, AND D. B. SZYLD, Block Krylov subspace methods for functions of
matrices}}, Electronic Transactions on Numerical Analysis, 47 (2017), pp. 100–126, http://etna.mcs.kent.edu/vol.47.2017/pp100-126.dir/pp100-126.pdf.

[12] A. Frommer and V. Simoncini, Matrix functions, in Model Order Reduction: Theory, Research Aspects and Applications, W. H. A. Schilders, H. A. van der Vorst, and J. Rommes, eds., vol. 13 of Mathematics in Industry, Berlin, 2008, Springer, pp. 275–304.

[13] D. F. Gleich, C. Greif, and J. M. Varah, The power and Arnoldi methods in an algebra of circulants, Numerical Linear Algebra with Applications, 20 (2013), pp. 809–831, https://doi.org/10.1002/nla, https://arxiv.org/abs/1112.5546v3.

[14] N. J. Higham, Functions of Matrices, SIAM, Philadelphia, 2008.

[15] R. A. Horn and C. R. Johnson, Topics in Matrix Analysis, Cambridge University Press, Cambridge, 1991.

[16] B. N. Khoromskij, Tensor numerical methods for multidimensional PDES: theoretical analysis and initial applications, ESAIM: Proceedings and Surveys, 48 (2015), pp. 1–28, https://doi.org/10.1051/proc/201448001, http://www.esaim-proc.org/10.1051/proc/201448001.

[17] T. G. Kolda and B. W. Bader, Tensor decompositions and applications, SIAM Review, 51 (2009), pp. 455–500, https://doi.org/10.1137/07070111X.

[18] T. G. Kolda and D. R. Mayo, Shifted power method for computing tensor eigenpairs, SIAM Journal on Matrix Analysis and Applications, 32 (2011), pp. 1095–1124.

[19] L.-H. Lim, Singular values and eigenvalues of tensors: a variational approach, in Proceedings of the IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP '05), vol. 3, 2005, pp. 129–132, http://ieeexplore.ieee.org/document/1574201/.

[20] M. E. Kilmer and C. D. Martin, Factorization strategies for third-order tensors, Linear Algebra and Its Applications, 435 (2011), pp. 641–658, https://doi.org/10.1016/j.laa.2010.09.020, http://dx.doi.org/10.1016/j.laa.2010.09.020.

[21] M. E. Kilmer, K. Braman, N. Hao, and R. C. Hoover, Third-order tensors as operators on matrices: a theoretical and computational framework with applications in imaging, SIAM Journal on Matrix Analysis and Applications, 34 (2013), pp. 148–172, https://doi.org/10.1137/110837711.

[22] M. E. Kilmer and C. D. Martin, Factorization strategies for third-order tensors, Linear Algebra and Its Applications, 435 (2011), pp. 641–658, https://doi.org/10.1016/j.laa.2010.09.020, http://dx.doi.org/10.1016/j.laa.2010.09.020.

[23] T. G. Kolda and B. W. Bader, Tensor decompositions and applications, SIAM Review, 51 (2009), pp. 455–500, https://doi.org/10.1137/07070111X.

[24] T. G. Kolda and D. R. Mayo, Shifted power method for computing tensor eigenpairs, SIAM Journal on Matrix Analysis and Applications, 32 (2011), pp. 1095–1124.

[25] L.-H. Lim, Singular values and eigenvalues of tensors: a variational approach, in Proceedings of the IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP '05), vol. 3, 2005, pp. 129–132, http://ieeexplore.ieee.org/document/1574201/.

[26] M. Schweitzer, Restarting and error estimation in polynomial and extended Krylov subspace methods for the approximation of matrix functions, PhD thesis, Fakult¨ at f¨ ur Mathematik und Naturwissenschaften, Bergische Universit¨ at Wuppertal, 2015.

[27] M. Schweitzer, Any finite convergence curve is possible in the initial iterations of restarted FOM, Electronic Transactions on Numerical Analysis, 45 (2016), pp. 133–145.

[28] V. Simoncini, Ritz and Pseudo-Ritz values using matrix polynomials, Linear Algebra and its Applications, 241-243 (1996), pp. 787–801.

[29] V. Simoncini and E. Gallopoulos, Convergence properties of block GMRES and matrix polynomials, Linear Algebra and its Applications, 247 (1996), pp. 97–119, https://doi.org/10.1016/0024-3795(95)00093-3.

[30] V. Simoncini and L. Lopez, Analysis of projection methods for rational function approximation to the matrix exponential, SIAM Journal on Numerical Analysis, 44 (2006), pp. 613–635, https://doi.org/10.1137/05062590.

[31] Q.-S. Zheng, Theory of representations for tensor functions– A unified invariant approach to constitutive equations, Applied Mechanics Reviews, 47 (1994), p. 545, https://doi.org/10.1115/1.3111066, http://appliedmechanicsreviews.asmedigitalcollection.asme.org/article.aspx?articleid=1395390.