Perturbation analysis of third-order tensor eigenvalue problem based on tensor-tensor multiplication

Changxin Mo *  Weiyang Ding †

June 8, 2021

Abstract

Perturbation analysis has been primarily considered to be one of the main issues in many fields and considerable progress, especially getting involved with matrices, has been made from then to now. In this paper, we pay our attention to the perturbation analysis subject on tensor eigenvalues under tensor-tensor multiplication sense; and also $\varepsilon$-pseudospectra theory for third-order tensors. The definition of the generalized $T$-eigenvalue of third-order tensors is given. Several classical results, such as the Bauer-Fike theorem and its general case, Gershgorin circle theorem and Kahan theorem, are extended from matrix to tensor case. The study on $\varepsilon$-pseudospectra of tensors is presented, together with various pseudospectra properties and numerical examples which show the boundaries of the $\varepsilon$-pseudospectra of certain tensors under different levels.

Keywords: perturbation theory, tensor-tensor multiplication, tensor eigenvalues, Bauer-Fike theorem, $\varepsilon$-pseudospectra theory, Gershgorin circle theorem, Kahan theorem, multidimensional ordinary differential equation, multi-linear time invariant system

AMS subject classifications: 15A18, 15A69

*E-mail: cxmo16@fudan.edu.cn. School of Mathematical Sciences, Chongqing Normal University, Chongqing, 400047, P. R. of China.

†Corresponding author. E-mail: dingwy@fudan.edu.cn. Institute of Science and Technology for Brain-Inspired Intelligence, Fudan University, Shanghai, China; Shanghai Center for Brain Science and Brain-Inspired Technology, Shanghai, China; Key Laboratory of Computational Neuroscience and Brain-Inspired Intelligence (Fudan University), Ministry of Education, China; MOE Frontiers Center for Brain Science, Fudan University, Shanghai, China; Zhangjiang Fudan International Innovation Center. W. Ding’s research is supported by the National Natural Science Foundation of China under Grant 11801479, Shanghai Municipal Science and Technology Major Project (No. 2018SHZDZX01), ZJ Lab, and Shanghai Center for Brain Science and Brain-Inspired Technology, the 111 Project (No. B18015).
1 Introduction

Perturbation theory, which has been studied for more than ninety years, stems from the ideas of Rayleigh [16] and Schrödinger [18] when they studied the eigenvalue problems in vibrating system and quantum mechanics, separately. Since then, extensive researches have been conducted by many researchers and those pioneering works can be found in the celebrated monographs: Perturbation theory of eigenvalue problems by Rellich [17], Perturbation Theory for Linear Operator by Kato [7] and Perturbation Analysis of Matrix by Sun [20], to name a few.

Tensor, which can be regarded as a generalization of the matrix to higher-order case, has been attracted considerable attention in different fields of science recently. As one of the most important and basic operations like matrix multiplication, tensor multiplication has been attended greatly by researchers. In 2008, a new type of tensor multiplication, that allows a third-order tensor to be written as a product of third-order tensors also, have been proposed by Kilmer et al. [10] when they considered the problem of generalizing the matrix SVD to tensor case, and it is termed as tensor-tensor multiplication. In the definition of this new multiplication, three operators are involved closely. For a third-tensor $A \in \mathbb{C}^{n_1 \times n_2 \times n_3}$, we get its frontal slices, denoted by $A_{i:k}$ or $A_k$ for short, by fixing the last index. Therefore it has $n_3$ frontal slices, i.e., $A_1, \ldots, A_{n_3}$, which are matrices with size $n_1 \times n_2$. The first operation we will introduce is about creating a block circulant matrix from the frontal slices of a tensor. That is,

$$\text{bcirc}(A) = \begin{bmatrix} A_1 & A_{n_3} & A_{n_3-1} & \cdots & A_2 \\ A_2 & A_1 & A_{n_3} & \cdots & A_3 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ A_{n_3} & A_{n_3-1} & \cdots & A_2 & A_1 \end{bmatrix}$$

with size $(n_1 n_3) \times (n_2 n_3)$. The other two operations could be regarded as the “inverse” of each other. They are the unfold and fold commands defined as follows,

$$\text{unfold}(A) = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_{n_3} \end{bmatrix}, \quad \text{fold}(\text{unfold}(A)) = A.$$

On the basis of the above three operations, the tensor-tensor multiplication of any two tensors $B \in \mathbb{C}^{n_1 \times p \times n_3}$ and $C \in \mathbb{C}^{p \times n_2 \times n_3}$ is defined as

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

One can easily check that $A \in \mathbb{C}^{n_1 \times n_2 \times n_3}$. By the size of the three tensors involved in (1.2), we can see that special attention should be paid to the first two indices of $B$ and
\( C \) since their frontal slices need to be consistent with the multiplication of matrices. The tensor-tensor multiplication (1.2) has demonstrated its usefulness in many areas, including, but not limited to, image processing (such as image deblurring and compression, object and facial recognition), tensor principal component analysis, tensor completion, pattern recognition; see [8, 11, 15, 22, 23] and the references therein.

On the basis of this important tensor-tensor multiplication, many researchers therefore considered the functions of multidimensional arrays. It can be regarded as a generalization of the functions of matrices, and many nice properties have been given; more details are discussed in the articles [12, 13, 14]. Specially, one basic but also important concept, \( T \)-eigenvalue, has been proposed by Miao et al. [14], and then the stability of \( T \)-eigenvalues has been introduced in [6] for studying the tensor Lyapunov equation that appears in spatially invariant systems. Moreover, some results, such as Weyl’s and Cauchy's interlacing theorems, from the matrix case to the tensor case have been given [6].

Motivated by these researches mentioned above, we pay our attention to the perturbation analysis of third-order tensors under the important tensor-tensor multiplication (1.2) in this paper. Many classical results of the matrix case will be generalized to the tensor case. Moreover, the pseudospectra theory for third-order tensors also has been considered.

The remainder of this paper is organized as follows. We describe some notations that often used in the following and revisit several basic concepts as well as fundamental results in section 2. Section 3 is one of the main parts which focus on the perturbation analysis results on third-order tensors. Several classical theorems on matrices are extended to tensor case. The issue on \( \epsilon \)-pseudospectra theory, which is the other main part, of tensors is studied in section 4. Section 5 presents the multidimensional ordinary differential equation and also investigate the close relationship of \( T \)-eigenvalue with it. Some properties of multidimensional ordinary differential equation are also given. This paper culminates with conclusions and remarks.

## 2 Preliminaries

In this section, we introduce the notations used throughout the paper and also review the basic concepts of tensors, such as identity tensor, transpose of a tensor, \( F \)-diagonal tensor, and orthogonal tensor.

Generally, scalars are denoted by lowercase letters, e.g., \( a \). Vectors and matrices are denoted by boldface lowercase letters and capital letters, respectively, e.g., \( \mathbf{v} \) and \( A \). Euler script letters are used to denote the higher-order tensors, e.g., \( A \). Frontal slices of a tensor \( \mathcal{T} \in \mathbb{C}^{n_1 \times n_2 \times n_3} \) are denoted by \( T_1, \ldots, T_{n_3} \).

Some basic concepts of tensors are revisited next.

**Definition 2.1.** ([9, Definition 3.14]). Let \( \mathcal{A} \in \mathbb{C}^{n_1 \times n_2 \times n_3} \), then the transposed tensor \( \mathcal{A}^\top \in \mathbb{C}^{n_2 \times n_1 \times n_3} \) (conjugate transposed tensor \( \mathcal{A}^H \in \mathbb{C}^{n_2 \times n_1 \times n_3} \)) is obtained by taking the
transpose (conjugate transpose) of each frontal slices and then reversing the order of transposed frontal slices 2 through \( n_3 \).

Unlike the transposed tensor which is well-defined for any \( n_1 \times n_2 \times n_3 \) tensor, the identity tensor, orthogonal tensor and the inverse of a tensor are only applicable for the tensors with square frontal slices.

**Definition 2.2.** ([9, Definition 3.4], identity tensor). Let \( I_{mm\ell} \in \mathbb{C}^{m \times m \times \ell} \). If its frontal slice \( I_1 \) is the identity matrix of size \( m \times m \), and whose other frontal slices \( I_2, \ldots, I_\ell \) are all zeros, then we call \( I_{mm\ell} \) an identity tensor.

**Definition 2.3.** ([9, Definition 3.5], inverse of a tensor). Let \( A \in \mathbb{C}^{m \times m \times \ell} \). We call tensor \( B \) an inverse of \( A \) if it satisfies the following two qualities

\[
A \ast B = I_{mm\ell}, \quad \text{and} \quad B \ast A = I_{mm\ell}.
\]

**Definition 2.4.** ([9, Definition 3.18], orthogonal and unitary tensor). Let \( Q \in \mathbb{R}^{m \times m \times \ell} \). We call \( Q \) an orthogonal tensor provided that \( Q \ast Q = Q \ast Q = I_{mm\ell} \). If \( Q \in \mathbb{C}^{m \times m \times \ell} \) and \( Q^H \ast Q = Q \ast Q^H = I_{mm\ell} \), then we call it an unitary tensor.

Based on the above definition, a tensor \( A \in \mathbb{C}^{m \times m \times n} \) is said to be symmetric if \( A = A^\top \), or Hermitian if \( A = A^H \) [6].

We call a third-order tensor \( D \in \mathbb{C}^{m \times m \times \ell} \) an \( F \)-diagonal tensor if all its frontal slices \( D_1, \ldots, D_\ell \) are diagonal matrices.

**Definition 2.5.** ([12, 14], \( F \)-diagonalizable tensor). Assume that \( A \in \mathbb{C}^{m \times m \times \ell} \) such that

\[
A = P \ast D \ast P^{-1},
\]

then we call \( A \) an \( F \)-diagonalizable tensor if \( D \) is an \( F \)-diagonal tensor.

Some useful lemmas are recalled as follows.

**Lemma 2.1.** ([6, 12, 13]). The following results hold for third-order tensors \( A \in \mathbb{C}^{m \times n \times p} \):

(a) The operator \( bcirc \) defined in (1.1) is a linear operator, i.e.,

\[
bcirc(\alpha A + \beta B) = \alpha bcirc(A) + \beta bcirc(B)
\]

where \( B \) has the same size as \( A \) and \( \alpha, \beta \) are constants.

(b) \( bcirc(\alpha A \ast B) = bcirc(\alpha A) bcirc(B) \) where \( B \in \mathbb{C}^{n \times n \times p} \).

(c) \( bcirc(A^\top) = (bcirc(A))^\top \), and \( bcirc(A^H) = (bcirc(A))^H \).

(d) If \( A \) is invertible, then its inverse tensor is unique and \( bcirc(A^{-1}) = (bcirc(A))^{-1} \).

**Lemma 2.2.** ([6, Theorems 2.7 and 2.8]). Let \( A \in \mathbb{C}^{m \times m \times n} \). Some fundamental results involving Hermitian or symmetric tensors are:

(a) The tensor \( A \) is symmetric if and only if \( bcirc(A) = (bcirc(A))^\top \).

(b) The tensor \( A \) is Hermitian if and only if \( bcirc(A) = (bcirc(A))^H \).

(c) All \( T \)-eigenvalues (cf. (3.3)) of a Hermitian tensor \( A \) are real.
Lemma 2.3. ([14, Lemma 4]). Suppose $A_1, \ldots, A_p, B_1, \ldots, B_p \in \mathbb{C}^{n \times n}$ are matrices satisfying

$$
\begin{bmatrix}
A_1 & A_p & A_{p-1} & \cdots & A_2 \\
A_2 & A_1 & A_p & & A_3 \\
\vdots & \vdots & \ddots & \ddots & \vdots \\
A_p & A_{p-1} & A_{p-2} & \cdots & A_1
\end{bmatrix} = (F_p \otimes I_n) \begin{bmatrix}
B_1 & & & \\
& B_2 & & \\
& & \ddots & \\
& & & B_p
\end{bmatrix} (F_p^H \otimes I_n)
$$

where $F_p$ is the normalized discrete Fourier matrix of size $p \times p$. Then, $B_1, \ldots, B_p$ are diagonal (sub-diagonal, upper-triangular, lower-triangular) matrices if and only if $A_1, \ldots, A_p$ are diagonal (sub-diagonal, upper-triangular, lower-triangular) matrices.

Similar as the matrix case, two tensors $A, B \in \mathbb{C}^{m \times m \times \ell}$ are said to commute if $A * B = B * A$. A tensor $A \in \mathbb{C}^{m \times m \times n}$ is normal if $A * A^H = A^H * A$, that is, if $A$ commutes with its conjugate transpose under tensor-tensor multiplication sense. Obviously, the symmetric tensor and Hermitian tensor are normal.

A normal tensor can always be $F$-diagonalizable by an unitary tensor, as the conclusion given next.

Lemma 2.4. ([14]). Let $A \in \mathbb{C}^{m \times m \times n}$ be a normal tensor, then there exists an unitary tensor $U \in \mathbb{C}^{m \times m \times \ell}$ such that

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

where $\mathcal{D}$ is an $F$-diagonal tensor.

3 Perturbation analysis on third-order tensors

In this section, we firstly give a definition of the generalized tensor-eigenvalue of third-order tensors. And then some classical results, such as the Gershgorin circle theorem [4], the Bauer-Fike theorem and its general case [1, 3, 19], and the Kahan theorem [20] which are well-known in matrix theory, are extended into the tensor case.

3.1 Tensor eigenvalues under tensor-tensor multiplication

Firstly, we give the definition of generalized tensor-eigenvalue for third-order tensors under tensor-tensor multiplication.

Definition 3.1. (Generalized $T$-eigenvalue of tensors). Let $A, B \in \mathbb{C}^{m \times m \times \ell}$. If there is a $\lambda \in \mathbb{C}$ and a nonzero tensor $\mathcal{X} \in \mathbb{C}^{m \times 1 \times \ell}$ such that

$$
A * \mathcal{X} = \lambda (B * \mathcal{X}),
$$

then $\lambda$ is called a generalized $T$-eigenvalue of $A$ relative to $B$ and $\mathcal{X}$ is a $T$-eigenvector associated to $\lambda$. 


Remark 3.1. By the definition of the tensor-tensor multiplication given in (1.2), we can see that the equality (3.1) is equivalent to
\[ \text{bcirc}(A) \text{ unfold}(\mathcal{X}) = \lambda \cdot (\text{bcirc}(B) \text{ unfold}(\mathcal{X})). \] (3.2)
And note that unfold(\mathcal{X}) is a vector with size \( m\ell \), thus this generalized eigenvalue problem based on tensor-tensor product has a close relationship with the classical generalized matrix eigenvalue problem. Hence, there are \( m\ell \) eigenvalues for the problem (3.1) if and only if
\[ \text{rank}(B) := \text{rank}(\text{bcirc}(B)) = m\ell. \]
If \( B \) is rank deficient, then the set of all the generalized T-eigenvalue of \( A \) relative to \( B \) may be finite, empty or infinite.

Remark 3.2. If we choose \( B \) as the identity tensor in (3.1), then we get
\[ A \ast \mathcal{X} = \lambda \mathcal{X}. \] (3.3)
This is the case that given in [6, Definition 2.5] which gives the definition of T-eigenvalue of a third-order tensor. An equivalent definition which based on the tensor singular value decomposition is also displayed in [14]. Therefore Definition 3.1 is the generalization of definitions given in [6, Definition 2.5] and also [14]. Similarly, as the equivalent form given in (3.2) for (3.1), we obtain that (3.3) can be transformed into
\[ \text{bcirc}(A) \text{ unfold}(\mathcal{X}) = \lambda \cdot \text{unfold}(\mathcal{X}). \]
That is to say, all T-eigenvalues of tensor \( A \) are actually eigenvalues of the circulant matrix \( \text{bcirc}(A) \), and vice versa.

3.2 Gershgorin circle theorem for tensors

Firstly, we consider the easy but fundamental Gershgorin circle theorem in this subsection.

Let \( A \in \mathbb{C}^{m \times m \times \ell} \). With the help of the normalized discrete Fourier transform matrix, \( \text{bcirc}(A) \) can be block-diagonalized. Furthermore, by a sequence of similarity transformation, it can be “more diagonal”. And in this case, we can use the diagonal entries to approximate the T-eigenvalues of original tensor \( A \).

Theorem 3.1. (Gershgorin circle theorem for tensors). Let \( A \in \mathbb{C}^{m \times m \times \ell} \) and assume
\[ X^{-1}(F_m \otimes I_n) \text{bcirc}(A)(F_m^H \otimes I_n)X = D + F \]
where \( X \) is the transformation matrix, \( D = \text{diag}(d_1,\ldots,d_{m\ell}) \) and \( F \) has zero diagonal entries. Then we have
\[ \Lambda(A) \subseteq \bigcup_{i=1}^{m\ell} \Theta_i \]
where $\Lambda(A)$ denotes the set of its $T$-eigenvalues and

$$\Theta_i = \left\{ z \in \mathbb{C} : |z - d_i| \leq \sum_{j=1}^{m\ell} |f_{ij}| \right\}, \ i = 1, \ldots, ml.$$

**Proof.** Without loss of generality, we assume that $\lambda \in \Lambda(A)$, and furthermore, we suppose that $\lambda \neq d_i$ for $i = 1, \ldots, m\ell$. Notice that $T$-eigenvalues are not affected by similarity transformations and the matrix $I - (\lambda I - D)^{-1}F$ is singular. Therefore,

$$1 \leq \|(D - \lambda I)^{-1}F\|_\infty \leq \frac{1}{|d_k - \lambda|} \sum_{j=1}^{m\ell} |f_{kj}|$$

for some $k$. The above inequalities imply that $|\lambda - d_k| \leq \sum_{j=1}^{m\ell} |f_{kj}|$ which further implies $\lambda \in \Theta_k$. Note that $\lambda$ is arbitrary and we complete the proof. \qed

It is noted that a different version of the Gershgorin circle theorem for tensors has been considered at the same time in [2]. In our theorem, the circles are based on the diagonal entries of the transformation form. However, in Theorem 5.2 of [2], the circles are based on the entries of the original tensor.

### 3.3 Bauer-Fike theorem for tensors

It is well-known that Bauer-Fike theorem is a classical result for a complex-valued diagonalizable matrix. It concerns the perturbation theory of the eigenvalue. More specifically, it states that an absolute upper bound for the deviation of one perturbed matrix eigenvalue from a properly chosen eigenvalue of the exact matrix can be estimated by the product of the condition number of the eigenvector matrix and the norm of the perturbation [1].

As for the case of non-diagonalizable matrices, the Bauer-Fike theorem has been generalized in [3, 4]. Moreover, this result on only part of the spectrum of a matrix was also considered in [3].

We generalize this celebrated result to third-order tensor case as follows. Noted that the Bauer-Fike Theorem considered in Theorem 5.3 of [2] at the same time is one special case of the next result. They only considered the 2-norm case.

**Theorem 3.2** (Bauer-Fike Theorem for Tensors). Let $A \in \mathbb{C}^{m \times n \times n}$ be an $F$-diagonalizable tensor. That is,

$$P^{-1} * A * P = D,$$

where $D$ is an $F$-diagonal tensor. Suppose that $\mu$ is a $T$-eigenvalue of $A + \delta A$ in which $\delta$ is a small number. Then, under the spectral norm or Frobenius norm case, there exists a $T$-eigenvalue $\lambda$ of $A$ such that

$$|\lambda - \mu| \leq \kappa_p(P)\|\delta A\|_p, \ p = 2, F.$$
Moreover, for the 1- and ∞-norms, we have

\[ |\lambda - \mu| \leq \kappa_p(P) \kappa_p(F_m \otimes I_n) \| \delta A \|_p, \ p = 1, \infty. \]

In the above expressions, \( F_m \) is the normalized discrete Fourier transform matrix; \( \kappa_p(P) = \| bcirc(P) \|_p \| bcirc(P^{-1}) \|_p \) is the condition number of \( P \), and \( \| A \|_p = \| bcirc(A) \|_p \).

**Proof.** By Lemma 2.1, we know that (3.4) is equivalent to

\[ (bcirc(P))^{-1} bcirc(A) bcirc(P) = bcirc(D). \]

The right-hand of the above equality is a block circulant matrix, thus it can be block-diagonalized by discrete Fourier transform matrix. That is to say,

\[ (F_m \otimes I_n) bcirc(D)(F_m^H \otimes I_n) = D = \begin{bmatrix} D^{(1)} & & \\ & \ddots & \\ & & D^{(n)} \end{bmatrix}. \]

By Lemma 2.3, we can see that \( D^{(1)}, D^{(2)}, \ldots, D^{(n)} \) are diagonal matrix since \( D \) is an \( F \)-diagonal matrix. It is not hard to see that all diagonal entries of those matrices \( D^{(1)}, D^{(2)}, \ldots, D^{(n)} \) are the \( T \)-eigenvalues of tensor \( A \). Also note that \( bcirc(A + \delta A) = bcirc(A) + bcirc(\delta A) \). And by the two arguments above, then we obtain

\[
(F_m \otimes I_n)[(bcirc(P))^{-1} bcirc(A + \delta A) bcirc(P)][F_m^H \otimes I_n]
= (F_m \otimes I_n)[(bcirc(P))^{-1} bcirc(A) bcirc(P)][F_m^H \otimes I_n]
+ (F_m \otimes I_n)[(bcirc(P))^{-1} bcirc(\delta A) bcirc(P)][F_m^H \otimes I_n]
= D + (F_m \otimes I_n)[(bcirc(P))^{-1} bcirc(\delta A) bcirc(P)][F_m^H \otimes I_n].
\]

(3.5)

Without loss of generality, we assume that \( \mu \notin \Lambda(A) \), otherwise the result is trivially true. Let \( \mu \) be a \( T \)-eigenvalue of \( A + \delta A \). Then \( \mu \) is an eigenvalue of \( bcirc(A + \delta A) \), and therefore \( \det(bcric(A + \delta A) - \mu I_{mn}) = 0 \). By the result of (3.5), one can find that

\[
0 = \det(bcric(A) + bcirc(\delta A) - \mu I_{mn})
= \det(F_m \otimes I_n) \cdot \det \left( (bcirc(P))^{-1} \right) \cdot \det(bcric(A) + bcirc(\delta A) - \mu I_{mn})
\cdot \det(bcric(P)) \cdot \det(F_m^H \otimes I_n)
= \det \left( D + (F_m \otimes I_n)[(bcirc(P))^{-1} bcirc(\delta A) bcirc(P)][F_m^H \otimes I_n] - \mu I_{mn} \right)
= \det(D - \mu I_{mn})
\cdot \det \left( (D - \mu I_{mn})^{-1} \left( (F_m \otimes I_n)[(bcirc(P))^{-1} bcirc(\delta A) bcirc(P)][F_m^H \otimes I_n] + I_{mn} \right) \right).
\]

(3.6)
The assumption that $\mu \notin \Lambda(\mathcal{A})$ implies that
\[
\det)((D - \mu I)^{-1}((F_m \otimes I_n)[(bcirc(\mathcal{P}))^{-1} bcirc(\delta \mathcal{A}) bcirc(\mathcal{P})](F_m^H \otimes I_n) + I_{mn})) = 0
\]
which shows that $-1$ is an eigenvalue of the matrix
\[
(D - \mu I)^{-1}((F_m \otimes I_n)[(bcirc(\mathcal{P}))^{-1} bcirc(\delta \mathcal{A}) bcirc(\mathcal{P})](F_m^H \otimes I_n)).
\]
Since all $p$-norms ($p = 1, 2, F, \infty$) are consistent matrix norms and thus we have
\[
| - 1| \leq \|(D - \mu I)^{-1}((F_m \otimes I_n)[(bcirc(\mathcal{P}))^{-1} bcirc(\delta \mathcal{A}) bcirc(\mathcal{P})](F_m^H \otimes I_n))\|_p
\]
\[
\leq \|(D - \mu I)^{-1}\|_p \|F_m \otimes I_n\|_p\|(bcirc(\mathcal{P}))^{-1}\|_p\| bcirc(\delta \mathcal{A})\|_p\| bcirc(\mathcal{P})\|_p\|F_m^H \otimes I_n\|_p
\]
\[
= \|(D - \mu I)^{-1}\|_p \kappa_p(bcirc(\mathcal{P}))\kappa_p(F_m \otimes I_n)\|bcirc(\delta \mathcal{A})\|_p
\]
\[
= \|(D - \mu I)^{-1}\|_p \kappa_p(\mathcal{P})\kappa_p(F_m \otimes I_n)\|bcirc(\delta \mathcal{A})\|_p.
\]
Notice that $(D - \mu I)^{-1}$ is a diagonal matrix, then for $p = 1, 2, \infty$ we have
\[
\|(D - \mu I)^{-1}\|_p = \max_{\|\mathbf{x}\|_p \neq 0} \frac{\|(D - \mu I)^{-1}\mathbf{x}\|_p}{\|\mathbf{x}\|_p} = \max_{\lambda \in \Lambda(\mathcal{A})} \frac{1}{|\lambda - \mu|} = \frac{1}{\min_{\lambda \in \Lambda(\mathcal{A})} |\lambda - \mu|}. \tag{3.7}
\]
Therefore
\[
\min_{\lambda \in \Lambda(\mathcal{A})} |\lambda - \mu| \leq \|(D - \mu I)^{-1}\|_p \kappa_p(\mathcal{P})\kappa_p(F_m \otimes I_n)\|bcirc(\delta \mathcal{A})\|_p. \tag{3.8}
\]
Finally, for the 2-norm case, we obtain
\[
\kappa_2(F_m \otimes I_n) = \|F_m \otimes I_n\|_2\|F_m^H \otimes I_n\|_2 = 1
\]
since $F_m \otimes I_n$ and $F_m^H \otimes I_n$ are unitary. The result for Frobenius norm is trivial since spectral norm of a matrix is not larger than its Frobenius norm. The proof is completed. \qed

The following conclusion describes the relationship between variation of $T$-spectrum and difference of two tensors. We omit the proof since it can be easily get by the above theorem.

**Corollary 3.1.** Let $\mathcal{A}, \mathcal{B} \in \mathbb{C}^{m \times m \times n}$ and $\mathcal{A}$ is an $F$-diagonalizable tensor with decomposition as
\[
\mathcal{P}^{-1} \ast \mathcal{A} \ast \mathcal{P} = \mathcal{D}.
\]
Let $s_\mathcal{A}(\mathcal{B})$ be the distance of two $T$-spectral sets $\Lambda_\mathcal{A} = \{\lambda_i\}_{i=1}^{mn}$ and $\Lambda_\mathcal{B} = \{\mu_i\}_{i=1}^{mn}$, and it is defined by
\[
s_\mathcal{A}(\mathcal{B}) = \max_{1 \leq j \leq mn} \left\{ \min_{1 \leq i \leq mn} |\lambda_i - \mu_j| \right\}.
\]
Then we have
\[
s_\mathcal{A}(\mathcal{B}) \leq \kappa_2(\mathcal{P})\|\mathcal{B} - \mathcal{A}\|_2.
\]
If $\mathcal{A}$ is a normal, then
\[
s_\mathcal{A}(\mathcal{B}) \leq \|\mathcal{B} - \mathcal{A}\|_2.
\]
Generally, most tensors are not $F$-diagonalizable which means that (3.4) is not satisfied. Therefore, we cannot get an $F$-diagonal tensor by a transformation under tensor-tensor multiplication. In this case, we have the following decomposition.

**Lemma 3.1.** [14] (T-Schur decomposition). Let $\mathcal{A} \in \mathbb{C}^{m \times m \times n}$, then there exists an unitary tensor $\mathcal{Q}$ such that

$$Q^{-1} \ast A \ast Q = T = D + N.$$  

(3.9)

where $D$ is an $F$-diagonal tensor and each frontal slice of $N$ is strictly upper triangular.

In the following, we present two general case of Bauer-Fike theorem for tensors (i.e., Theorem 3.2). They are based on $T$-Schur decomposition and can be viewed as the generalization of the matrix cases for non-diagonalizable matrices that given in [3, 4].

**Theorem 3.3.** (Generalization of Bauer-Fike theorem). Let $Q^{-1} \ast A \ast Q = D + N$ be a $T$-Schur decomposition of $\mathcal{A} \in \mathbb{C}^{m \times m \times n}$ as given in (3.9). The tensor $B \in \mathbb{C}^{m \times m \times n}$ and $\epsilon$ is a small scalar. If $\mu$ is a $T$-eigenvalue of $\mathcal{A} + \epsilon B$ and $q$ is the smallest positive number such that $|N|^q = 0$ where

$$N := \begin{bmatrix} N^{(1)} &  &  \\ & N^{(2)} &  \\ &  & \ddots \\  &  &  & N^{(n)} \end{bmatrix} = (F_m \otimes I_n) \cdot \text{bcirc}(N) \cdot (F_m^H \otimes I_n)$$

and $|N| = (|N_{ij}|)$ denotes the absolute of a matrix element-wisely, then for spectral and Frobenius norms we have

$$\min_{\lambda \in \Lambda(\mathcal{A})} |\lambda - \mu| \leq \max \{ \theta, \theta^{1/q} \}$$  

(3.10)

in which

$$\theta = \| \text{bcirc}(\epsilon B) \|_p \sum_{k=0}^{q-1} \| N \|^k_p, \ p = 2, F.$$  

(3.11)

For the 1- and $\infty$-norms, we get

$$\min_{\lambda \in \Lambda(\mathcal{A})} |\lambda - \mu| \leq \max \{ \theta_p, \theta_p^{1/q} \}$$  

(3.12)

where

$$\theta_p = \| \text{bcirc}(\epsilon B) \|_{p,k} \kappa_p(\mathcal{Q}) \kappa_p(F_m \otimes I_n) \sum_{k=0}^{q-1} \| N \|^k_2, \ p = 1, \infty.$$  

**Proof.** The theorem is clearly true if $\mu \in \Lambda(\mathcal{A})$, as the left-hand sides of (3.10) and (3.12) vanish. Therefore we assume that $\mu \notin \Lambda(\mathcal{A})$. By Lemma 2.1, we can see that

$$\mu I_{mn} - \text{bcirc}(\mathcal{A} + \epsilon B) = \mu I_{mn} - \text{bcirc}(\mathcal{A}) - \text{bcirc}(\epsilon B),$$

10
and moreover it is singular. This means that
\[
(F_m \otimes I_n) \text{bcirc}(Q)^{-1} \left[ \mu I_{mn} - \text{bcirc}(A) - \text{bcirc}(\epsilon B) \right] \text{bcirc}(Q) (F_m^H \otimes I_n) \tag{3.13}
\]
is also singular since the matrices multiplied on the left and right sides are nonsingular. Notice that
\[
(F_m \otimes I_n) \cdot \text{bcirc}(Q)^{-1} \text{bcirc}(A) \text{bcirc}(Q) \cdot (F_m^H \otimes I_n) = (F_m \otimes I_n) \cdot (\text{bcirc}(D) + \text{bcirc}(N)) \cdot (F_m^H \otimes I_n)
\]
(3.13) can be rewritten as
\[
\mu I_{mn} - D - N = (F_m \otimes I_n) \cdot \text{bcirc}(Q)^{-1} \text{bcirc}(A) \text{bcirc}(Q) \cdot (F_m^H \otimes I_n),
\]
and then the following matrix
\[
I_{mn} - (\mu I_{mn} - D - N)^{-1} (F_m \otimes I_n) \cdot \text{bcirc}(Q)^{-1} \text{bcirc}(A) \text{bcirc}(Q) \cdot (F_m^H \otimes I_n) \tag{3.14}
\]
is singular.

By the assumption that \(|N|^q = 0\) and note that \(\mu I_{mn} - D\) is diagonal, it follows that \(((\mu I_{mn} - D)^{-1} N)^q = 0\). Hence,
\[
((\mu I_{mn} - D) - N)^{-1} = \sum_{k=0}^{q-1} ((\mu I_{mn} - D)^{-1} N)^k (\mu I_{mn} - D)^{-1}
\]
and
\[
\|((\mu I_{mn} - D) - N)^{-1}\| \leq \frac{1}{\min_{\lambda \in \Lambda(A)} |\lambda - \mu|} \sum_{k=0}^{q-1} \left( \frac{\|N\|}{\min_{\lambda \in \Lambda(A)} |\lambda - \mu|} \right)^k
\]
under the 1-, 2- and \(\infty\)-norms cases.

If \(\min_{\lambda \in \Lambda(A)} |\lambda - \mu| \geq 1\), then
\[
\|((\mu I_{mn} - D) - N)^{-1}\| \leq \frac{1}{\min_{\lambda \in \Lambda(A)} |\lambda - \mu|} \sum_{k=0}^{q-1} \|N\|^k,
\]
and if \(\min_{\lambda \in \Lambda(A)} |\lambda - \mu| < 1\), then
\[
\|((\mu I_{mn} - D) - N)^{-1}\| \leq \frac{1}{\left( \min_{\lambda \in \Lambda(A)} |\lambda - \mu| \right)^q} \sum_{k=0}^{q-1} \|N\|^k.
\]
By (3.14), we obtain
\[
1 \leq \|(\mu I_{mn} - D - N)^{-1}(F_m \otimes I_n) \cdot \text{bcirc}(Q)^{-1} \text{bcirc}(\mathcal{B}) \text{bcirc}(Q) \cdot (F_m^H \otimes I_n)\|
\]

\[
= \|(\mu I_{mn} - D - N)^{-1}\| \|\text{bcirc}(\mathcal{B})\| \|F_m \otimes I_n\| \|F_m^H \otimes I_n\| \|\text{bcirc}(Q)^{-1}\| \|\text{bcirc}(Q)\|
\]

and under the spectral norm case, we get
\[
\min_{\lambda \in \Lambda(A)} |\lambda - \mu| \leq \|\text{bcirc}(\mathcal{B})\|_2 \sum_{k=0}^{q-1} \|N\|_2^k \quad \text{or} \quad \bigg( \min_{\lambda \in \Lambda(A)} |\lambda - \mu| \bigg)^q \leq \|\text{bcirc}(\mathcal{B})\|_2 \sum_{k=0}^{q-1} \|N\|_2^k
\]

for \(\min_{\lambda \in \Lambda(A)} |\lambda - \mu| \geq 1\) or \(\min_{\lambda \in \Lambda(A)} |\lambda - \mu| < 1\), respectively. Let \(\theta = \|\text{bcirc}(\mathcal{B})\|_2 \sum_{k=0}^{q-1} \|N\|_2^k\). Then we get the result (3.10) for spectral norm. The Frobenius norm case can be get easily.

For the 1- and \(\infty\)-norms, by using (3.14) again we get
\[
\min_{\lambda \in \Lambda(A)} |\lambda - \mu| \leq \max \{\theta_p, \theta_1^{1/q}\}
\]

where
\[
\theta_p = \|\text{bcirc}(\mathcal{B})\|_p \kappa_p(Q) \kappa_p(F_m \otimes I_n) \sum_{k=0}^{q-1} \|N\|_2^k
\]

and the proof is completed.

Now, we give one more general result of Theorem 3.2. Different from the above theorem which involving with an \(F\)-diagonal tensor, in the next result, we consider block-diagonal case.

Let \(A \in \mathbb{C}^{m \times m \times n}\). Notice that \(\text{bcirc}(A)\) can be block-diagonalized as follows,
\[
(F_m \otimes I_n) \text{bcirc}(A)(F_m^H \otimes I_n) = \begin{bmatrix}
A^{(1)} & & \\
& A^{(2)} & \\
& & \ddots \\
& & & A^{(n)}
\end{bmatrix}.
\]

By Lemma 2.3, we know that \(A^{(i)}\) where \(i = 1, \ldots, n\) may not diagonal since generally tensor \(A\) is not \(F\)-diagonal. For each matrix \(A^{(i)}\), let \(X^{(i)}\) be a transformation matrix such that \((X^{(i)})^{-1}A^{(i)}X^{(i)} = \text{diag}(A^{(i)}_{k_i})\) where \(A^{(i)}_{k_i}\) is in triangular Schur form with
\[
A^{(i)}_{k_i} = D^{(i)}_{k_i} + N^{(i)}_{k_i}, \quad k_i = 1, \ldots, \ell_i.
\]
Denote $\text{bcirc}(X) = (F_m^H \otimes I_n) \text{diag}(X^{(1)}, X^{(2)}, \ldots, X^{(n)})(F_m \otimes I_n)$. Then we get

$$(F_m \otimes I_n) \cdot \text{bcirc}(X)^{-1} \text{bcirc}(A) \cdot \text{bcirc}(X) \cdot (F_m^H \otimes I_n)$$

$$= \begin{bmatrix}
\text{diag}(A_{k_1}^{(1)}) & \text{diag}(A_{k_2}^{(2)}) & \cdots & \text{diag}(A_{k_n}^{(n)})
\end{bmatrix} = \begin{bmatrix}
A_1^{(1)} & \cdots & A_1^{(n)} \\
& \cdots & \\
& & A_\ell_1^{(1)} & \cdots & A_\ell_1^{(n)}
\end{bmatrix}
$$

$$= \begin{bmatrix}
D_1^{(1)} & \cdots & D_1^{(n)} \\
& \cdots & \\
& & D_\ell_1^{(1)} & \cdots & D_\ell_1^{(n)}
\end{bmatrix} + \begin{bmatrix}
N_1^{(1)} & \cdots & N_1^{(n)} \\
& \cdots & \\
& & N_{\ell_1}^{(1)} & \cdots & N_{\ell_1}^{(n)}
\end{bmatrix}
$$

$$:= \tilde{D} + \tilde{N}.$$  

By the above analysis, we have the following conclusion.

**Theorem 3.4.** If $\mu$ is a $T$-eigenvalue of $A + \epsilon B$ and $q$ is the dimension of $A_{k_1}^{(i)}$, then we have

$$\min_{\lambda \in \Lambda(A)} |\lambda - \mu| \leq \max\{\theta_1, \theta_1^{1/q}\}$$

where

$$\theta_1 = C\epsilon\|\text{bcirc}(B)\|p,k_p(\mathcal{X}), \ p = 2, F.$$  

and $C = \sum_{k=0}^{q-1} \|N^i\|^p_2$ provided that $\max_j \left\| (A_j^{(i)} - \mu I)^{-1} \right\|$ occurring at $j = k_i$.

For $1$- and $\infty$-norms, under the above condition, we get

$$\min_{\lambda \in \Lambda(A)} |\lambda - \mu| \leq \max\{\theta_2, \theta_2^{1/q}\}$$

where

$$\theta_2 = C\epsilon\|\text{bcirc}(B)\|p,k_p(\mathcal{X})k_p(F_m \otimes I_n).$$

**Proof.** We only need to consider the case that $\mu$ is not an $T$-eigenvalue of $A$. Hence $\mu I_{mn} - \tilde{D} - \tilde{N}$ is nonsingular. Similar as (3.14), the matrix

$$I_{mn} - (\mu I_{mn} - \tilde{D} - \tilde{N})^{-1}(F_m \otimes I_n) \cdot \text{bcirc}(X)^{-1} \text{bcirc}(A) \cdot \text{bcirc}(X) \cdot (F_m^H \otimes I_n)$$

is singular. By similar proof process as Theorem 3.3, we could get the conclusion. \qed
3.4 Kahan theorem for tensors

The result on a Hermite tensor that is perturbed by a Hermite tensor is studied in [6]. Next, we give a result that a Hermite tensor is perturbed by any tensors.

Theorem 3.5. (Kahan theorem for tensors). Let $A \in \mathbb{R}^{m \times m \times n}$ be a Hermite tensor. Suppose that its $T$-eigenvalues set is denoted by $\Lambda_A = \{\lambda_i\}_{i=1}^{mn}$ such that its $T$-eigenvalues are arranged in a non-increasing order:

$$\lambda_{\text{max}} = \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{mn-1} \geq \lambda_{mn} = \lambda_{\text{min}}.$$ (3.15)

Suppose that $B = A + \mathcal{E}$ and let $\Lambda_B = \{\beta_k + i\gamma_k\}_{k=1}^{mn}$ such that $\beta_1 \geq \beta_2 \geq \cdots \geq \beta_{mn-1} \geq \beta_{mn}$. Let

$$E_y = \frac{\text{bcirc} (\mathcal{E}) - \text{bcirc} (\mathcal{E})^H}{2i}$$

and

$$\sigma_k = \{\beta + i\gamma \in \mathbb{C} : |\beta + i\gamma - \lambda_k| \leq \|\mathcal{E}\|_2, |\gamma| \leq \|E_y\|_2\}.$$ Then

$$\Lambda_B \subset \bigcup_{k=1}^{mn} \sigma_k.$$ 

Proof. On one hand, by Lemma 2.4, a Hermite tensor is $F$-diagonalizable by a unitary tensor. According to Corollary 3.1, we know that there exists one $\lambda_k$ such that $|\beta + i\gamma - \lambda_k| \leq \|\mathcal{E}\|_2$ for any given $T$-eigenvalue $\beta + i\gamma$ of $B$.

On the other hand, suppose that $B \ast \mathcal{X} = (\beta + i\gamma)\mathcal{X}$, then by the definition of tensor-tensor multiplication, it is equivalent with

$$\text{bcirc}(B) \text{ unfold}(\mathcal{X}) = (\beta + i\gamma) \text{ unfold}(\mathcal{X}).$$

Thus it is reasonable to assume that $\|\text{unfold}(\mathcal{X})\|_2 = 1$ since $\mathcal{X}$ is nonzero which implies that the vector $\text{unfold}(\mathcal{X})$ is also nonzero. Therefore,

$$(\text{unfold}(\mathcal{X}))^H \text{bcirc}(B) \text{ unfold}(\mathcal{X}) = \beta + i\gamma \quad \text{and} \quad (\text{unfold}(\mathcal{X}))^H \text{bcirc}(B)^H \text{ unfold}(\mathcal{X}) = \beta - i\gamma$$

which implies that

$$\gamma = \frac{(\text{unfold}(\mathcal{X}))^H [\text{bcirc}(B) - \text{bcirc}(B)^H] \text{ unfold}(\mathcal{X})}{2i} = \text{ unfold}(\mathcal{X})^H E_y \text{ unfold}(\mathcal{X}).$$

Thus $|\gamma| \leq \|E_y\|_2$.

Our conclusion follows by combining this two parts. \qed

14
4 Pseudospectra of third-order tensors

Pseudospectra of finite-dimensional matrices has been thoroughly investigated in the classical book by Trefethen [21]. Three definitions of pseudospectra for any norm and one definition for spectral norm are given, and those definitions are equivalent under certain conditions. Many properties and representative numerical results are presented by many pictures.

In this section, we study the pseudospectra theory of third-order tensors under tensor-tensor multiplication sense. The definition of pseudospectra on the basis of $T$-eigenvalue defined in Definition 3.1 is given first. Some properties based on the definition are given soon afterwards.

4.1 Pseudospectra of third-order tensors under tensor-tensor multiplication

First, we give the definition of $\varepsilon$-pseudospectra of an $m \times m \times n$ tensor $A$. If the norm $\| \cdot \|_p$ is not specified, we take the convention that $p = 1, 2, \infty$.

**Definition 4.1 ($\varepsilon$-pseudospectra of a Tensor).** Let $A \in \mathbb{C}^{m \times m \times n}$. Then the block circulant matrix $\text{bcirc}(A)$ generated by the tensor $A$ can be factored as follows,

$$\text{bcirc}(A) = (F^H_m \otimes I_n) \cdot \begin{bmatrix} A^{(1)} & & \\ & A^{(2)} & \\ & & \ddots \\ & & & A^{(n)} \end{bmatrix} \cdot (F_m \otimes I_n). \quad (4.1)$$

For each square matrix $A^{(i)}$ where $i \in [n] := \{1, \ldots, n\}$, we have

$$\Lambda_\varepsilon(A^{(i)}) := \{ z \in \mathbb{C} : \|(zI_m - A^{(i)})^{-1}\| \geq \varepsilon^{-1} \}$$

where $\varepsilon$ is a positive scalar. If there are some $i$ such that $zI_m - A^{(i)}$ are singular, we define $\|(zI_m - A^{(i)})^{-1}\| = \infty$. In this definition, we denote the block diagonal matrix in (4.1) as $\text{diag}(A^{(1)}, \ldots, A^{(n)}) := A$.

(I) We call

$$\Lambda_\varepsilon(A) := \bigg\{ z \in \mathbb{C} : \max_{i \in [n]} \|(zI_m - A^{(i)})^{-1}\| \geq \varepsilon^{-1} \bigg\}$$

as the $\varepsilon$-pseudospectra of tensor $A$.

(II) $\Lambda_\varepsilon(A) = \{ z \in \mathbb{C} : z \in \Lambda(A + E) \text{ for some } E \in \mathbb{C}^{mn \times mn} \text{ with } \|E\| \leq \varepsilon \}$.

(III) $\Lambda_\varepsilon(A) = \{ z \in \mathbb{C} : \text{there exists } v \in \mathbb{C}^{mn} \text{ with } \|v\| = 1 \text{ such that } \|(A - zI_{mn})v\| \leq \varepsilon \}.$

**Theorem 4.1.** The three definitions (I), (II) and (III) are equivalent.
For a block diagonal matrix, we find that $\|A\| = \max_{i \in [n]} \|A^{(i)}\|$ and the inverse of $A$ can be get by computing the inverse of each $A^{(i)}$. Therefore,

$$\max_{i \in [n]} \| (zI_n - A^{(i)})^{-1} \| = \| (zI_m - A)^{-1} \|$$

and thus

$$\Lambda_{\varepsilon}(A) = \{ z \in \mathbb{C} : \| (zI_n - A)^{-1} \| \geq \varepsilon^{-1} \}.$$

(4.2)

The equivalence of (4.2) and (II), (III) can be easily got by the matrix case [21].

\begin{remark}
Notice that $z \in \mathbb{C}$ and it is variable, while $A^{(i)}$ is stationary when the tensor $A$ is given. Therefore by definition (I), we can also see that

$$\Lambda_{\varepsilon}(A) = \bigcup_{i=1}^{n} \Lambda_{\varepsilon}(A^{(i)}) = \bigcup_{i=1}^{n} \{ z \in \mathbb{C} : \| (zI_n - A^{(i)})^{-1} \| \geq \varepsilon^{-1} \}.$$

Under the case of the spectral norm, we give the following definition.

\begin{definition}
Let $A \in \mathbb{C}^{m \times m \times n}$. Then the block-circulant matrix $\text{bcirc}(A)$ generated by the tensor $A$ can be factored as (4.1). For each square matrix $A^{(i)}$ where $i \in [n]$, we have

$$\Lambda_{\varepsilon}(A^{(i)}) = \{ z \in \mathbb{C} : \sigma_{\text{min}}(zI_n - A^{(i)}) \leq \varepsilon \}$$

where $\varepsilon$ is a positive scalar and $\sigma_{\text{min}}(\cdot)$ denotes the minimum singular value. We call

$$\Lambda_{\varepsilon}(A) = \bigcup_{i=1}^{n} \Lambda_{\varepsilon}(A^{(i)})$$

as the $\varepsilon$-pseudospectra of tensor $A$.

By Remark 4.1, we get the following conclusion.

\begin{theorem}
The three definitions given in Definition 4.1 and the one in Definition 4.2 are also equivalent.
\end{theorem}

\subsection{Properties of pseudospectra of tensors}

We study the properties of pseudospectra for third-order tensors in this subsection. Many fundamental results are given in the next theorem.

\begin{theorem}
Let tensor $A \in \mathbb{C}^{m \times m \times n}$ and suppose that the positive scalar $\varepsilon$ is given arbitrarily.

1. The set $\Lambda_{\varepsilon}(A)$ is nonempty, open, and bounded. Moreover, there are at most $nm$ connected components, each containing one or more $T$-eigenvalues of $A$.

2. For any $c \in \mathbb{C}$, we have $\Lambda_{\varepsilon}(A + c) = c + \Lambda_{\varepsilon}(A)$ where $A + c$ is shorthand for $A + cI$ and $I$ is the identity matrix with the same size as $A$.

3. For any nonzero $c \in \mathbb{C}$, we have $\Lambda_{\varepsilon}(cA) = c\Lambda_{\varepsilon}(A)$.

4. If the spectral norm is applied, then $\Lambda_{\varepsilon}(A^H) = \overline{\Lambda_{\varepsilon}(A)}$.
\end{theorem}
PROOF. To prove the assertion of (1), one can use the fact that each \( \Lambda_\varepsilon (A^{(i)}) \) of the matrix \( A^{(i)} \) has the properties that nonempty, open, and bounded, with at most \( m \) connected components, each containing one or more eigenvalues of \( A^{(i)} \). Then same properties are also hold for the given tensor \( \mathcal{A} \). Moreover, there are at most \( nm \) connected components since

\[
\Lambda_\varepsilon (\mathcal{A}) = \bigcup_{i=1}^{n} \Lambda_\varepsilon (D_i).
\]

We now come to the matter of part (2). First, note that

\[
\text{bcirc}(A + cI) = \begin{bmatrix}
A_1 + cI_m & A_n & A_{n-1} & \cdots & A_2 \\
A_2 & A_1 + cI_m & A_n & \cdots & A_3 \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
A_n & A_{n-1} & \cdots & A_2 & A_1 + cI_m
\end{bmatrix}
\]

\[
= \text{bcirc}(A) + c\text{bcirc}(I)
\]

\[
= (F_m^H \otimes I_n) \cdot A \cdot (F_m \otimes I_n) + c \left( (F_m^H \otimes I_m) \cdot I_{mn} \cdot (F_m \otimes I_n) \right)
\]

\[
= (F_m^H \otimes I_n) \cdot (A + cI_{mn}) \cdot (F_m \otimes I_n)
\]

\[
= (F_m^H \otimes I_n) \cdot \begin{bmatrix}
A^{(1)} + cI_m \\
& A^{(2)} + cI_m \\
& & \ddots \\
& & & A^{(n)} + cI_m
\end{bmatrix} \cdot (F_m \otimes I_n).
\]

Therefore,

\[
\Lambda_\varepsilon (\mathcal{A} + c) = \bigcup_{i=1}^{n} \Lambda_\varepsilon (A^{(i)} + cI_m) = \bigcup_{i=1}^{n} \left[ \Lambda_\varepsilon (A^{(i)}) + c \right] = c + \Lambda_\varepsilon (\mathcal{A})
\]

for any \( c \in \mathbb{C} \) and we complete the proof of this part.

For the part (3), by Lemma 2.1, we know that

\[
\text{bcirc}(cA) = c\text{bcirc}(A) = c \left[ (F_m^H \otimes I_n) \cdot A \cdot (F_m \otimes I_n) \right] = (F_m^H \otimes I_n) \cdot (cA) \cdot (F_m \otimes I_n)
\]

which implies that

\[
\Lambda_{|c|\varepsilon} (cA) = \bigcup_{i=1}^{n} \Lambda_{|c|\varepsilon} (cD_i) = \bigcup_{i=1}^{n} c\Lambda_{\varepsilon} (D_i) = c \bigcup_{i=1}^{n} \Lambda_{\varepsilon} (D_i)
\]

since for any nonzero \( c \in \mathbb{C} \) and matrix \( A \in \mathbb{C}^{m \times m} \), the following equality

\[
\Lambda_{|c|\varepsilon} (cA) = c\Lambda_{\varepsilon} (A)
\]
holds. Thus we get the result that $\Lambda_{|c|\varepsilon}(cA) = c\Lambda_{\varepsilon}(A)$ for any nonzero $c \in \mathbb{C}$.

Now, we prove the last part of this theorem. By Lemma 2.1, we know that

$$\text{bcirc} \left( A^H \right) = \left( F_m^H \otimes I_n \right) \cdot A^H \cdot (F_m \otimes I_n).$$

Therefore,

$$\Lambda_{\varepsilon} \left( A^H \right) = \bigcup_{i=1}^n \Lambda_{\varepsilon}((A^{(i)})^H) = \bigcup_{i=1}^n \Lambda_{\varepsilon}(A^{(i)}) = \bigcup_{i=1}^n \Lambda_{\varepsilon}(A)$$

where the conclusion $\Lambda_{\varepsilon}(A^H) = \overline{\Lambda_{\varepsilon}(A)}$ under the two-norm for any matrix $A \in \mathbb{C}^{m \times m}$ is applied in the second equality.

**Remark 4.2.** By the results above, we can see that the function of pseudospectra on tensor $A$ is linear.

The properties of pseudospectra on normal tensors are given next.

**Theorem 4.4.** (Pseudospectra of a normal tensor). Let $\Delta_{\varepsilon}$ be an open $\varepsilon$-ball; that is, $\Delta_{\varepsilon} = \{ z \in \mathbb{C} : |z| < \varepsilon \}$. A sum of sets is defined as

$$\sigma(A) + \Delta_{\varepsilon} = \{ z : z = z_1 + z_2, z_1 \in \sigma(A), z_2 \in \Delta_{\varepsilon} \}$$

where $\Lambda(A)$ is the $T$-spectrum (sets of $T$-eigenvalues) of the tensor $A$. Then for any tensor $A \in \mathbb{C}^{m \times m \times n}$, we have

$$\Lambda_{\varepsilon}(A) \supseteq \Lambda(A) + \Delta_{\varepsilon} \quad \forall \varepsilon > 0. \quad (4.3)$$

Moreover if $A$ is normal and $\| \cdot \| = \| \cdot \|_2$, then

$$\Lambda_{\varepsilon}(A) = \Lambda(A) + \Delta_{\varepsilon} \quad \forall \varepsilon > 0. \quad (4.4)$$

**Proof.** If $\lambda$ is an $T$-eigenvalue of tensor $A$, then it is an eigenvalue of the matrix $\text{bcirc}(A)$. Therefore $\lambda + \mu$ is an eigenvalue of $\text{bcirc}(A) + \mu I$ for any $\mu \in \mathbb{C}$. Note that $\| \mu I \| = |\mu|$, then by the definition of pseudospectra on tensors, we obtain $\lambda + \mu \in \Lambda_{\varepsilon}(A)$ for any $|\mu| < \varepsilon$. The proof (4.3) is finished.

For the normal tensor case, by Lemma 2.4 and Lemma 2.1, we have

$$\text{bcirc}(U) \text{bcirc}(A)(\text{bcirc}(U))^H = \text{bcirc}(D)$$

and

$$(F_m \otimes I_n) \text{bcirc}(D)(F_m^H \otimes I_n) = \begin{bmatrix} D^{(1)} & & \\ & D^{(2)} & \\ & & \ddots \\ & & & D^{(n)} \end{bmatrix} := D \quad (4.5)$$
in which $D^{(i)}$ is diagonal for $i = 1, \ldots, n$ by Lemma 2.3. Also note that $\| \cdot \| = \| \cdot \|_2$, we may assume directly that $A$ is $F$-diagonal. Therefore, the diagonal entries of $\text{bcirc}(A)$ are equal to the $T$-eigenvalues. As we all know, the $\varepsilon$-pseudospectrum is just the union of the open $\varepsilon$-balls about the points of the spectrum for any normal matrix; equivalently, we have

$$\|(z - \text{bcirc}(A))^{-1}\|_2 = \frac{1}{\text{dist}(z, \Lambda(\text{bcirc}(A)))}$$

which implies

$$\text{dist}(z, \Lambda(\text{bcirc}(A))) < \varepsilon$$

by the $\varepsilon$-pseudospectrum of tensors. We get the conclusion since $\Lambda(A) + \Delta_\varepsilon$ is the same as $\{z : \text{dist}(z, \Lambda(A)) < \varepsilon\}$. \hfill \qed

**Theorem 4.5.** (Bauer-Fike Theorem). Suppose tensor $A \in \mathbb{C}^{m \times m \times n}$ is $F$-diagonalizable, i.e., it has decomposition (3.4). If the spectral norm is applied, then for each positive scalar $\varepsilon$, we have

$$\Lambda(A) + \Delta_\varepsilon \subseteq \Lambda_{\varepsilon}(A) \subseteq \Lambda(A) + \Delta_{\varepsilon\kappa_2(P)}.$$  

**Proof.** We only need to prove the second inclusion. By the definition of pseudospectra, Lemma 2.1 and decompositions (3.4) and (4.5), it is not hard to find that

$$\|(zI_{mn} - A)^{-1}\| = \| \text{bcirc}(P) (zI_{mn} - \text{bcirc}(D))^{-1} \text{bcirc}(P)^{-1}\|$$

$$\leq \frac{\kappa(P)}{\text{dist}(z, \Lambda(\text{bcirc}(D)))} = \frac{\kappa(P)}{\text{dist}(z, \Lambda(D))}$$

$$= \frac{\kappa(P)}{\text{dist}(z, \Lambda(A))}.$$  

Similar as the matrix case, we get our conclusion. \hfill \qed

### 4.3 An examples of the $\varepsilon$-pseudospectrum

**Example 4.1.** Let $A$ be a third-order tensor with three frontal faces $A_1, A_2$ and $A_3$. Firstly, we consider an example that all frontal faces are the same and each is a tridiagonal Toeplitz matrix, that is, $A_1 = A_2 = A_3 = T_{pz}$ where

$$T_{pz} = \begin{pmatrix}
0 & 1 & 0 & 1 & 0 \\
\frac{1}{4} & 0 & 1 & \ddots & \ddots \\
& \ddots & \ddots & \ddots & 0 \\
& & \frac{1}{4} & 0 & 1 \\
& & & \frac{1}{4} & 0
\end{pmatrix} \in \mathbb{R}^{N \times N}.$$  

We denote this tensor as $A_0$ and we can see that the size of $\text{bcirc}(A_0)$ is $3N \times 3N$. However, it is non-symmetrical. Notice that $T_{pz}$ can be symmetrized by the diagonal similarity transformation

$$DT_{pz}D^{-1} = S$$
where $D = \text{diag} \left( 2, 4, \ldots, 2^N \right)$ and

$$S = \begin{pmatrix}
0 & \frac{1}{2} & \frac{1}{2} & \cdots & \cdots & \frac{1}{2} \\
\frac{1}{2} & 0 & \frac{1}{2} & \cdots & \cdots & \frac{1}{2} \\
\frac{1}{2} & \frac{1}{2} & 0 & \cdots & \cdots & \frac{1}{2} \\
\cdot & \cdot & \cdot & \ddots & \ddots & \cdot \\
\frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \cdots & 0 & \frac{1}{2} \\
\frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \cdots & \frac{1}{2} & 0
\end{pmatrix} \in \mathbb{R}^{N \times N}.$$

Let $D_b$ be a block diagonal matrix such that $D_b = \text{diag} \left( D, D, D \right)$, then

$$D_b \text{bcirc}(A_0) D_b^{-1} = S_b$$

is symmetrical and thus all the eigenvalues are real which means that all the $T$-eigenvalues of the given tensor $A_0$ are real and only appear in the real axis. This agrees with the results Lemma 2.2.

Figure 1: Boundaries of pseudospectra $\Lambda_{\varepsilon}(A_0)$ for $\varepsilon = 10^{-1}, 10^{-2}, \ldots, 10^{-10}$ under the condition that $A_1 = A_2 = A_3 = T_{pz}$. The $T$-eigenvalues of tensor $A_0$ are plotted as crosses ‘×’. (Left: CONTOUR only; Right: SURF and CONTOUR)

However, the pseudospectra of $A_0$ lie far from the real axis. We consider a case where $N = 20$ and the results are given in Figure 1 in which the $T$-eigenvalues of tensor $A_0$ are plotted as crosses ‘×’. The spectral norm is chosen here and we set $\varepsilon = 10^{-1}, 10^{-2}, \ldots, 10^{-10}$. By the results we can see that the boundaries of $\Lambda_{\varepsilon}(A_0)$ lie far beyond the real axis.
Many $T$-eigenvalues of the given tensor $A_0$ are zero. Another tensor $A_1$ where $A_1 = T_{pz}$ and $A_2 = A_3 = 2T_{pz}$ is considered next. The corresponding matrix $bcirc(A_1)$ is non-symmetrical but all its eigenvalues are real. Similar results are given in Figure 2.

A case that the tensor has complex $T$-eigenvalues is considered next. By setting $A_3 = 2A_2 = 4A_1 = T_{pz}$, we get a new tensor $A_2$. The results are given in Figure 3.

At last, we consider an example where $A_1 = T_{pz}$, $A_2 = 10T_{pz}$ and $A_3 = eye(N)$ and we
denote this tensor as $A_3$. With similar parameters setting, we get the boundaries results in Figure 4.

![Figure 4](image1.png)

(a) CONTOUR only. (b) PCOLOR and CONTOUR.

Figure 4: Boundaries of pseudospectra $\Lambda_\varepsilon(A_3)$ for $\varepsilon = 10^{-1}, 10^{-2}, \ldots, 10^{-10}$ under the condition that $A_1 = T_{pz}, A_2 = 10T_{pz}$ and $A_3 = \text{eye}(N)$. The $T$-eigenvalues of tensor $A_2$ are plotted as crosses ‘×’. (Left: CONTOUR only; Right: PCOLOR and CONTOUR)

5 Multidimensional ordinary differential equation

We consider the first-order multidimensional ordinary differential equation in this section. During the studying process, we investigate the relationship between differential equation and $T$-eigenvalues.

Before studying the differential equation, we first suppose that the coefficient tensor $A \in \mathbb{C}^{m \times m \times n}$ has square frontal faces. The unknown function $\mathcal{Y}$ is defined as follows,

$$\mathcal{Y} : [0, \infty) \to \mathbb{C}^{m \times s \times n}. \quad (5.1)$$

Let the operator $\frac{d}{dt}$ be acting elementwise. Then, under the above assumptions, we consider the following differential equation \[12,\]

$$\frac{d\mathcal{Y}}{dt}(t) = A \ast \mathcal{Y}(t) \quad (5.2)$$

with initial condition $\mathcal{Y}(0) = \mathcal{Y}_0$ being given.

By unfolding both sides of (5.1) then we could get a equivalent matrix form,

$$\frac{d}{dt} \begin{bmatrix} Y_1(t) \\ \vdots \\ Y_n(t) \end{bmatrix} = \text{bcirc}(A) \begin{bmatrix} Y_1(t) \\ \vdots \\ Y_n(t) \end{bmatrix}. \quad (5.3)$$
Furthermore, according to the result in the theory of ordinary differential equation and by
the definition of tensor $T$-function [12], the solution of (5.2) can be represented as
\[
\mathcal{Y}(t) = \text{fold}(\exp(t \text{bcirc}(\mathcal{A}))) \cdot \text{unfold}((\mathcal{Y}(0))) = \exp(\mathcal{A}t) * \mathcal{Y}_0.
\] (5.4)

Consider the equation (5.4), the matrix which made up of $Y_1, \ldots, Y_n$ is of size $mn \times s$. For simplicity, we only consider one column case, that is,
\[
\frac{d\mathcal{X}}{dt}(t) = \mathcal{A} * \mathcal{X}(t), \quad \text{where } \mathcal{X} : [0, \infty) \to \mathbb{C}^{m \times 1 \times n}.
\] (5.5)

Therefore, we have
\[
\frac{d}{dt} \begin{bmatrix} X_1(t) \\ \vdots \\ X_n(t) \end{bmatrix} = \text{bcirc}(\mathcal{A}) \begin{bmatrix} X_1(t) \\ \vdots \\ X_n(t) \end{bmatrix}.
\] (5.6)

With initial value $\mathcal{X}(0)$ and by the theory of differential equations, we find that
\[
\begin{bmatrix} X_1(t) \\ \vdots \\ X_n(t) \end{bmatrix} = \exp(t \text{bcirc}(\mathcal{A})) \begin{bmatrix} X_1(0) \\ \vdots \\ X_n(0) \end{bmatrix}.
\] (5.7)

Notice that the left-hand side of (5.7) is an vector with length $np$. Therefore it is resealable to assume that solution of (5.7) has the form $\xi \exp(\lambda t)$ where the exponent $\lambda$ and vector $\xi$ are to be determined. Substituting this solution form for both sides of (5.6), we get
\[
(b \text{circ}(\mathcal{A}) - \lambda I_{mn})\xi = 0
\] (5.8)
and then
\[
\mathcal{A} * \text{fold}(\xi) = \lambda \cdot \text{fold}(\xi),
\]
where $\text{fold}(\xi) \in \mathbb{C}^{m \times 1 \times n}$, which shows that, by Definition 3.1, $\lambda$ is a $T$-eigenvalue of the given tensor $\mathcal{A}$ with $T$-eigenvector $\text{fold}(\xi)$.

Based on the above analysis, we conclude that the investigation concerning $T$-eigenvalues and eigenvectors plays an important part in analyzing the solution of the multidimensional ordinary differential equation (5.5). Furthermore, by the theories of matrix functions [5] and tensor $T$-functions [12, 13], we can see that $T$-eigenvalues defined in Definition 3.1 are also crucial and basic parts in studying theses matrix or tensor functions.

More generally, the differential equation (5.5) considered is a special case of
\[
\frac{d\mathcal{X}}{dt}(t) = \mathcal{A}(t) * \mathcal{X}(t) + \mathcal{G}(t)
\] (5.9)
in which $\mathcal{G}$ has the same size as $\mathcal{X}$. Then, similarly as the above process, we get
\[
\frac{d}{dt} \begin{bmatrix} X_1(t) \\ \vdots \\ X_n(t) \end{bmatrix} = \text{bcirc}(\mathcal{A}(t)) \begin{bmatrix} X_1(t) \\ \vdots \\ X_n(t) \end{bmatrix} + \begin{bmatrix} G_1(t) \\ \vdots \\ G_n(t) \end{bmatrix}.
\] (5.10)
Theorem 5.1. Let the mnm functions $A(i, j, k)(t)$ and mn functions $G(i, 1, k)(t)$ where $i, j \in [m]$ and $k \in [n]$ be continuous on an interval $(t_s, t_e)$. Suppose that the initial value is $X(t_0) = X_0$ in which $t_0$ is a specified value in $(t_s, t_e)$. Then there exists a unique solution $X(t) = (X(i, 1, k)(t))$ of system (5.9) that satisfies the initial conditions given above. Moreover, the solution exists throughout the assumed interval $(t_s, t_e)$.

Instead of considering the general (5.9) differential equation directly, we mainly pay our attention to

$$\frac{dX}{dt}(t) = A(t) \ast X(t) \quad (5.11)$$

obtained by setting $G(t) = 0$ in (5.9). And throughout this section, we assume that the continuous and initial conditions in Theorem 5.1 are satisfied.

Suppose that $X_1(t)$ and $X_2(t)$ are two solutions of equation (5.11); in other words, we have

$$\frac{dX_1}{dt}(t) = A(t) \ast X_1(t), \quad \frac{dX_2}{dt}(t) = A(t) \ast X_2(t).$$

Then, similar as the method in matrix case, we can generate more solutions by forming linear combinations of $X_1(t)$ and $X_2(t)$. As stated in the following theorem.

Theorem 5.2. (Principle of Superposition). If the tensor-valued functions $X_1(t)$ and $X_2(t)$ are two solutions of equation (5.11), then, for any scalars $c_1$ and $c_2$, the linear combination $c_1X_1(t) + c_2X_2(t)$ is also a solution.

Proof. Notice that

$$\frac{d(c_1X_1(t) + c_2X_2(t))}{dt} = c_1 \frac{dX_1}{dt}(t) + c_2 \frac{dX_2}{dt}(t)$$

$$= c_1 \text{fold}(bcirc(A(t)) \cdot \text{unfold}(X_1(t))) + c_2 \text{fold}(bcirc(A(t)) \cdot \text{unfold}(X_2(t)))$$

$$= \text{fold}(bcirc(A(t)) \cdot (c_1 \text{unfold}(X_1(t)) + c_2 \text{unfold}(X_2(t))))$$

$$= A(t) \ast (c_1X_1(t) + c_2X_2(t)),$$

and we get the result. □

By repeated application of the above theorem, it is easy to find that each finite linear combination of solutions of equation (5.11) is also a solution. Moreover, by analogy with the matrix case, we can prove that each solution of (5.11) is a combination of some solutions.

Theorem 5.3. Suppose that $X_1, \ldots, X_{mn}$ are solutions of (5.11) and under the unfold operation the mn vectors unfold($X_1), \ldots, \text{unfold}(X_{mn})$ are linearly independent for each $t \in (t_s, t_e)$. Then each solution of (5.11) can be expressed as a linear combination of the above mn solutions in exactly one way, that is,

$$X(t) = c_1X_1(t) + \cdots + c_{mn}X_{mn}(t). \quad (5.12)$$
Proof. We first suppose that \( t_0 \in (t_s, t_e) \) and let \( X_0 = X(t_0) \). If we can prove that there exist scalars \( c_1, \ldots, c_{mn} \) such that (5.12) holds for \( t = t_0 \),

\[
c_1X(t_0) + \cdots + c_{mn}X_{mn}(t_0) = X_0,
\]

then by Theorem 5.1 we can get the result.

By unfolding both sides of (5.13), we could get \( np \) equations in scalar form and the coefficients can be denoted as, at \( t = t_0 \),

\[
M[X_1, \ldots, X_{mn}] = \begin{bmatrix}
(X_1)_1(t) & \cdots & (X_{mn})_1(t) \\
\vdots & \ddots & \vdots \\
(X_1)_n(t) & \cdots & (X_{mn})_n(t)
\end{bmatrix}
\]

(5.14)

which is of size \( mn \times mn \). By the assumption of linear independence, we can see that the existence of \( c_1, \ldots, c_{mn} \) is guaranteed, and furthermore in only one way.

From the above analysis, the columns of matrix \( M \) are linearly independent at a given \( t \) if and only if \( \det(M) \neq 0 \). We call this determinant, denoted by \( W[X_1, \ldots, X_{mn}] \) as the Wronskian of the given \( mn \) solutions. Therefore we conclude that if the Wronskian of \( X_1, \ldots, X_{mn} \) is nonzero, then (5.12) holds.

For equation (5.11) that considered, if any set of solutions \( \{X_1, \ldots, X_{mn}\} \) with nonzero Wronskian \( W[X_1, \ldots, X_{mn}] \) at each time \( t \in [t_s, t_e] \), then it is said to be a fundamental set of solutions in \([t_s, t_e]\). If the scalars \( c_1, \ldots, c_{mn} \) can be chosen arbitrarily in (5.12), then we call it as a general solution.

By the uniqueness property of Theorem 5.1, actually the Wronskian only has two possible cases on the interval \([t_s, t_e]\). It is either identically zero or always nonzero.

**Theorem 5.4.** Let \( \{X_1, \ldots, X_{mn}\} \) be a set of solutions of (5.11) on the \([t_s, t_e]\). If the Wronskian \( W[X_1, \ldots, X_{mn}] \) is nonzero (zero) at arbitrary \( t_0 \in [t_s, t_e] \), then it is nonzero (zero) at the whole interval.

Proof. Since \( \text{unfold}(X_k) \) is a vector for each \( k = 1, \ldots, mn \), then for simplicity it is convenient to rewrite the Wronskian as

\[
W[X_1, \ldots, X_{mn}] = |\text{unfold}(X_1), \ldots, \text{unfold}(X_{mn})| = \begin{vmatrix}
(X_1)_1 & \cdots & (X_{mn})_1 \\
\vdots & \ddots & \vdots \\
(X_1)_n & \cdots & (X_{mn})_n
\end{vmatrix}
\]

(5.15)

25
Therefore by the properties of determinant, we get
\[
\frac{dW}{dt} = \left| \begin{array}{cccc}
\frac{d(X_1)_1}{dt} & \cdots & \frac{d(X_m)_1}{dt} & \\
\vdots & \ddots & \vdots & \\
\frac{d(X_1)_n}{dt} & \cdots & \frac{d(X_m)_n}{dt} & \\
\end{array} \right| + \cdots + \left| \begin{array}{cccc}
(X_1)_1 & \cdots & (X_m)_1 & \\
\vdots & \ddots & \vdots & \\
(X_1)_n & \cdots & (X_m)_n & \\
\end{array} \right|
\]
(5.16)
\[
= (X_1)_1 W + (X_2)_2 W + \cdots + (X_m)_n W
\]
\[
= ((X_1)_1 + (X_2)_2 + \cdots + (X_m)_n) W.
\]

Hence our conclusion follows since
\[
W(t) = c \exp \left( \int [(X_1)_1 + (X_2)_2 + \cdots + (X_m)_n] dt \right)
\]
(5.17)
in which the constant \(c\) is arbitrary.

The above theorem provides an easy criterion to distinguish the fundamental set of solutions from all sets that contains \(mn\) solutions of (5.11) by computing the value of their Wronskian at any point in \([t_s, t_e]\). Moreover, by this result, we can prove that the system we considered always has a fundamental set of solutions.

**Theorem 5.5.** Let \(\{X_1, \ldots, X_{mn}\}\) be a set of solutions of (5.11) on the \([t_s, t_e]\) that satisfies the following initial condition,
\[
X_i(t_0) = e_i^\top = (0, \ldots, 1, \ldots, 0), \quad i = 1, \ldots, mn, \quad t_0 \in [t_s, t_e].
\]
Then they form a fundamental set of solutions.

Now we give a result involving with complex-valued solution in the equation (5.11) whose coefficients are all real-valued.

**Theorem 5.6.** Let the coefficient \(A(t)\) in (5.11) be a real-valued continuous function. If \(X_1(t) + iX_2(t)\) is a is a complex-valued solution, then both \(X_1(t)\) and \(X_2(t)\) are real-valued solutions of (5.11).

**Proof.** By substituting \(X_1(t) + iX_2(t)\) for \(X(t)\) in (5.11) and notice that \(A(t)\) is real-valued, this result can be get.

### 6 Conclusions and Remarks

Based on the well-known tensor-tensor multiplication, the generalized \(T\)-eigenvalue and eigenvector are defined firstly in this paper, and then we focus our attention on the perturbation theory on \(T\)-eigenvalues. On one hand, we considered many classical perturbation results of matrix in the third-order tensor case, many conclusions, such as Bauer-Fike theorem, Kahan theorem and Gershgorin circle theorem, are presented. On
the other hand, the ε-pseudospectra of tensor is studied. Many equivalent definitions are given, and then we considered the properties. Some numerical experiments are also presented. In the last section, we investigate the close relationship between T-eigenvalues and multidimensional ordinary differential equations. Some properties of the differential equations are also presented.

However, there are still many perturbation results not considered. Also, giving an estimate of the T-eigenvalues of tensor $A$ is also meaningful. Future research directions including the theory of multi-linear time invariant system specified for third-order tensors under tensor-tensor multiplication as well as the important role of T-eigenvalue plays on in this system and also some algorithms for computation. This system is given by

$$\begin{cases} \frac{d\mathcal{X}(t)}{dt} = A \ast \mathcal{X}(t) + B \ast U(t) \\ \mathcal{Y}(t) = C \ast \mathcal{X}(t) \end{cases} \quad (6.1)$$

where, $\mathcal{X}(t) \in \mathbb{R}^{m \times 1 \times n}$ is the latent state space tensor, $U(t) \in \mathbb{R}^{\ell \times 1 \times n}$ is the control tensor and $\mathcal{Y}(t) \in \mathbb{R}^{s \times 1 \times n}$ is the output tensor; $A \in \mathbb{R}^{m \times m \times n}, B \in \mathbb{R}^{m \times \ell \times n}$ and $C \in \mathbb{R}^{s \times m \times n}$ are real valued tensors. Notions such as stability, reachability and observability of (6.1) can also be generalized from the classical control theory. Moreover, computational framework and also application-driven study will be given high priority in the future. Besides, the numerical range of third-order tensors and algorithms for computation are also under consideration.

Acknowledgments

The authors would like to express their appreciation to Prof. Yimin Wei for many useful discussions and valuable comments.

References

[1] F. L. Bauer and C. T. Fike, *Norms and exclusion theorems*, Numerische Mathematik, 2 (1960), pp. 137–141.

[2] Z. Cao and P. Xie, *On some tensor inequalities based on the t-product*, 2021. arXiv.

[3] K.-W. E. Chu, *Generalization of the Bauer-Fike theorem*, Numerische Mathematik, 49 (1986), pp. 685–691.

[4] G. H. Golub and C. F. Van Loan, *Matrix computations*, The Johns Hopkins University Press, Baltimore, 2013.

[5] N. J. Higham, *Functions of matrices: Theory and computation*, SIAM, Philadelphia, 2008.

[6] W. Hui Liu and X. Qing Jin, *A study on T-eigenvalues of third-order tensors*, Linear Algebra Appl., 612 (2021), pp. 357–374.
[7] T. Kato, *Perturbation theory for linear operators*, vol. 132, Springer Science & Business Media, 2013.

[8] 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.

[9] M. E. Kilmer and C. D. Martin, *Factorization strategies for third-order tensors*, Linear Algebra Appl., 435 (2011), pp. 641–658.

[10] M. E. Kilmer, C. D. Martin, and L. Perrone, *A third-order generalization of the matrix svd as a product of third-order tensors*, Tufts University, Department of Computer Science, Tech. Rep. TR-2008-4, (2008).

[11] Y. Liu, L. Chen, and C. Zhu, *Improved robust tensor principal component analysis via low-rank core matrix*, IEEE Journal of Selected Topics in Signal Processing, 12 (2018), pp. 1378–1389.

[12] K. Lund, *The tensor t-function: a definition for functions of third-order tensors*, Numer. Linear Algebra Appl., 27 (2020), pp. e2288, 17.

[13] Y. Miao, L. Qi, and Y. Wei, *Generalized tensor function via the tensor singular value decomposition based on the t-product*, Linear Algebra and its Applications, 590 (2020), pp. 258–303.

[14] ———, *T-Jordan canonical form and T-Drazin inverse based on the T-product*, Communications on Applied Mathematics and Computation, (2021), pp. 201–220.

[15] E. Newman and M. E. Kilmer, *Nonnegative tensor patch dictionary approaches for image compression and deblurring applications*, SIAM Journal on Imaging Sciences, 13 (2020), pp. 1084–1112.

[16] L. Rayleigh, *The theory of sound. Vol. I*, London, 1927.

[17] F. Rellich and J. Berkowitz, *Perturbation theory of eigenvalue problems*, CRC Press, 1969.

[18] E. Schrödinger, *Quantisierung als Eigenwertproblem*, Annalen Phys., 386 (1926), pp. 109–139.

[19] X. Shi and Y. Wei, *A sharp version of Bauer–Fikes theorem*, Journal of Computational and Applied Mathematics, 236 (2012), pp. 3218–3227.

[20] J. Sun, *Perturbation Analysis of Matrix*, Science Press, Beijing, 1987.

[21] L. N. Trefethen and M. Embree, *Spectra and pseudospectra: the behavior of nonnormal matrices and operators*, Princeton University Press, 2005.

[22] T. Wu, *Graph regularized low-rank representation for submodule clustering*, Pattern Recognition, 100 (2020), p. 107145.
[23] X.-L. Zhao, W.-H. Xu, T.-X. Jiang, Y. Wang, and M. K. Ng, *Deep plug-and-play prior for low-rank tensor completion*, Neurocomputing, 400 (2020), pp. 137 – 149.