Generalized T-Product Tensor Bernstein Bounds

Shih Yu Chang¹ and Yimin Wei^{2,*}

 ¹ Department of Applied Data Science, San Jose State University, San Jose, CA 95192, USA
 ² School of Mathematical Sciences and Shanghai Key Laboratory of Contemporary Applied Mathematics, Fudan University, Shanghai 200433, China

Received 20 November 2021; Accepted (in revised version) 6 January 2022

Abstract. Since Kilmer et al. introduced the new multiplication method between two third-order tensors around 2008 and third-order tensors with such multiplication structure are also called as T-product tensors, T-product tensors have been applied to many fields in science and engineering, such as low-rank tensor approximation, signal processing, image feature extraction, machine learning, computer vision, and the multi-view clustering problem, etc. However, there are very few works dedicated to exploring the behavior of random T-product tensors. This work considers the problem about the tail behavior of the unitarily invariant norm for the summation of random symmetric T-product tensors. Majorization and antisymmetric Kronecker product tools are main techniques utilized to establish inequalities for unitarily norms of multivariate T-product tensors. The Laplace transform method is integrated with these inequalities for unitarily norms of multivariate T-product tensors to provide us with Bernstein bound estimation of Ky Fan k-norm for functions of the symmetric random T-product tensors summation. Finally, we also apply T-product Bernstein inequality to bound Ky Fan norm of covariance T-product tensor induced by hypergraph signal processing.

AMS subject classifications: 15B52, 60B20, 11M50, 15A69

Key words: T-product tensors, T-eigenvalues, T-singular values, Bernstein bound, Courant-Fischer theorem for T-product tensors.

^{*}Corresponding author.

Emails: shihyu.chang@sjsu.edu (S. Chang), ymwei@fudan.edu.cn (Y. Wei)

1 Introduction

Since Kilmer et al. introduced the new multiplication method between two thirdorder tensors (T-product tensors), many new algebraic properties about such new multiplication rule between two third-order tensors are investigated [9,11]. These useful algebraic properties of T-product tensors have been discovered as powerful tools in many science and engineering fields [21, 28]. Although T-product tensors have attracted many practical applications, all of these applications of T-product tensors assume that T-product tensors under consideration are deterministic. This assumption is not practical in general scientific and engineering applications based on T-product tensors. In [4,5], the authors have tried to establish several new tail bounds for sums of random T-product tensors. These probability bounds characterize large-deviation behavior of the extreme T-eigenvalue of the sums of random T-product tensors (definitions about T-eigenvalues and T-singular values associated to T-product tensors are given in Section 2.1). The authors first apply Lapalace transform method and Lieb's concavity theorem for T-product tensors obtained from the work [4] to build several inequalities based on random T-product tensors, then utilize these inequalities to generalize the classical bounds associated with the names Chernoff, and Bernstein from the scalar to the T-product tensor setting. Tail bounds for the norm of a sum of random rectangular T-product tensors are also derived from corollaries of random symmetric T-product tensors cases. The proof mechanism is also applied to T-product tensor valued martingales and T-product tensor-based Azuma, Hoeffding and McDiarmid inequalities are also derived [5]. The random tensor and its applications in MRI and the tensor normal distribution can be found in [1, 23].

In this work, we will apply majorization techniques to establish new Bernstein bounds based on the summation of random symmetric T-product tensors. Compared to the previous work studied in [4, 5, 14], we make following generalizations: (1) besides bounds related to extreme values of T-eigenvalues, we consider more general unitarily invariant norm for T-product tensors; (2) the bounds derived in [5] can only be applied to the identity map for the summation of random symmetric T-product tensors, this work can derive new bounds for any polynomial function raised by any power greater or equal than one for the summation of random symmetric T-product tensors. In order to drive these new bounds, we also establish Courant-Fischer minmax theorem for T-product tensors in Theorem 2.1 and marjoization relation for T-singular values in Lemma 4.1. Our main theorem is provided below:

Theorem 1.1 (Generalized T-product tensor Bernstein bound). Consider a sequence $\{\mathcal{X}_j \in \mathbb{R}^{m \times m \times p}\}$ of independent, random symmetric T-product tensors with random structure defined by Definition 4.1. Let g be a polynomial function with degree n and nonnegative coefficients a_0, a_1, \dots, a_n raised by power $s \ge 1$, i.e., $g(x) = (a_0+a_1x+\dots+a_nx^n)^s$ with $s \ge 1$. Suppose following condition is satisfied:

$$g\left(\exp\left(t\sum_{j=1}^{M}\mathcal{X}_{j}\right)\right) \succeq \exp\left(tg\left(\sum_{j=1}^{M}\mathcal{X}_{j}\right)\right) \quad almost \ surely, \tag{1.1}$$

where t > 0, and we also have

$$\mathcal{X}_{j}^{p} \leq \frac{p!\mathcal{A}^{2}}{2} \quad almost \ surely \ for \ p = 2, 3, 4, \cdots.$$
 (1.2)

Then we have following inequality:

$$\Pr\left(\left\|g\left(\sum_{j=1}^{M} \mathcal{X}_{j}\right)\right\|_{(k)} \ge \theta\right)$$

$$\leq (n+1)^{s-1} \inf_{t>0} e^{-\theta t} k \left\{a_{0}^{s} + \sum_{l=1}^{n} a_{l}^{ls} \left[1 + Mlst\Phi(m,d_{1},d_{2}) + \frac{(Mlst)^{2}\sigma_{1}(\mathcal{A}^{2})}{2(1 - Mlst)}\right]\right\}. \quad (1.3)$$

The rest of this paper is organized as follows. In Section 2, we review T-product tensors basic concepts and introduce a powerful scheme about antisymmetric Kronecker product for T-product tensors. In Section 3, we apply a majorization technique to prove T-product tensor norm inequalities. We then apply new derived T-product tensor norm inequalities to obtain random T-product tensor Bernstein bounds for the extreme T-eigenvalues and Ky Fan k-norm in Section 4. Finally, concluding remarks are given by Section 6.

2 T-product tensors

In this section, we will introduce fundamental facts about T-product tensors in Section 2.1. Several unitarily invariant norms about a T-product tensor are defined in Section 2.2. A powerful scheme about antisymmetric Kronecker product for T-product tensors will be provided by Section 2.3.

2.1 T-product tensor fundamental facts

All third order tensors considered in this work will adopt T-product between two third order tensors multiplication. Basic definitions like identity, symmetric Tproduct tenor, inner product and computations can be found at [2, 10, 13, 17, 18, 24, 26]. Notions about trace, T-positive definite (TPD) tensor, T-positive semidefinite (TPSD) tensor, SVD of a symmetric T-product tensor are provided by [27].

If a T-product tensor $\mathcal{C} \in \mathbb{R}^{m \times m \times p}$ can be diagonalized as

$$\operatorname{bcirc}(\mathcal{C}) = \left(\mathbf{F}_{p}^{\mathrm{H}} \otimes \mathbf{I}_{m}\right) \operatorname{Diag}(\mathbf{C}_{i} : i \in \{1, \cdots, m\}) \left(\mathbf{F}_{p} \otimes \mathbf{I}_{m}\right),$$
(2.1)

the *j*-th eigenvalue of the matrix \mathbf{C}_i is called a T-eigenvalue [12,19], denoted by $\lambda_{i,j}$. We then define the *determinant* of a T-product tensor $\mathcal{C} \in \mathbb{R}^{m \times m \times p}$, represented by $\det(\mathcal{C})$, as

$$\det(\mathcal{C}) = \prod_{i=1, j=1}^{i=m, j=p} \lambda_{i,j}.$$
(2.2)

Similarly, singular values of each matrix C_i are T-singular values of the tensor C.

If a symmetric T-product tensor $C \in \mathbb{R}^{m \times m \times p}$ can be expressed as the format shown by Eq. (2.1), the T-eigenvalues of C with respect to the matrix \mathbf{C}_i are denoted as λ_{i,k_i} , where $1 \leq k_i \leq m$, and we assume that $\lambda_{i,1} \geq \lambda_{i,2} \geq \cdots \geq \lambda_{i,m}$ (including multiplicities). Then, λ_{i,k_i} is the k_i -th largest T-eigenvalue associated to the matrix \mathbf{C}_i . If we sort all T-eigenvalues of C from the largest one to the smallest one, we use \tilde{k} , a smallest integer between 1 to $m \times p$ (inclusive) associated with p given positive integers k_1, k_2, \cdots, k_p that satisfies

$$\lambda_{\tilde{k}} = \min_{1 \le i \le m} \lambda_{i,k_i} \quad \text{and} \quad \lambda_{\tilde{k}} \ge \lambda_{i,k_i+1}, \tag{2.3}$$

and we set \tilde{i} from $\lambda_{\tilde{k}}$ as

$$\tilde{i} = \arg\min_{i} \{\lambda_{\tilde{k}} = \lambda_{i,k_i}\}.$$
(2.4)

Then, we will have the following Courant-Fischer theorem for T-product tensors.

Theorem 2.1. Given a symmetric T-product tensor $C \in \mathbb{R}^{m \times m \times p}$ and p positive integers k_1, k_2, \dots, k_p with $1 \le k_i \le m$, then we have

$$\lambda_{\tilde{k}} = \max_{\substack{S \in \mathbb{R}^{m \times 1 \times p} \\ \dim(S) = \{k_1, \cdots, k_p\}}} \min_{\mathcal{X} \in S} \frac{\langle \mathcal{X}, \mathcal{C} \star \mathcal{X} \rangle}{\langle \mathcal{X}, \mathcal{X} \rangle}$$
$$= \min_{\substack{T \in \mathbb{R}^{m \times 1 \times p} \\ \dim(T) = \{n - k_1, \cdots, n - k_{\tilde{i}-1}, n - k_{\tilde{i}+1}, n - k_{\tilde{i}+1}, \cdots, n - k_p\}}} \max_{\mathcal{X} \in T} \frac{\langle \mathcal{X}, \mathcal{C} \star \mathcal{X} \rangle}{\langle \mathcal{X}, \mathcal{X} \rangle}, \qquad (2.5)$$

where $\lambda_{\tilde{k}}$ and \tilde{i} are defined by Eqs. (2.3) and (2.4).

Proof. First, we have to express $\langle \mathcal{X}, \mathcal{C} \star \mathcal{X} \rangle$ by matrices of \mathbf{C}_i and \mathbf{X}_i through the representation shown by Eq. (2.1). It is

$$\langle \mathcal{X}, \mathcal{C} \star \mathcal{X} \rangle = \frac{1}{p} \langle \operatorname{bcirc}(\mathcal{X}), \operatorname{bcirc}(\mathcal{C}) \operatorname{bcirc}(\mathcal{X}) \rangle$$

$$= \frac{1}{p} \operatorname{Tr} \left(\operatorname{bcirc}(\mathcal{X})^{\mathrm{H}} \operatorname{bcirc}(\mathcal{C}) \operatorname{bcirc}(\mathcal{X}) \right)$$

$$= \frac{1}{p} \operatorname{Tr} \left(\mathbf{F}_{p}^{\mathrm{H}} \operatorname{Diag} \left(\mathbf{x}_{i}^{\mathrm{H}} \mathbf{A}_{i} \mathbf{x}_{i} : i \in \{1, \cdots, p\} \right) \mathbf{F}_{p} \right)$$

$$= \frac{1}{p} \operatorname{Tr} \left(\operatorname{Diag} \left(\mathbf{x}_{i}^{\mathrm{H}} \mathbf{A}_{i} \mathbf{x}_{i} : i \in \{1, \cdots, p\} \right) \right) = \frac{1}{p} \sum_{i=1}^{p} \mathbf{x}_{i}^{\mathrm{H}} \mathbf{A}_{i} \mathbf{x}_{i}.$$
(2.6)

We will just verify the first characterization of $\lambda_{\tilde{k}}$. The other is similar. Let S_i be the projection of S to the space with dimension k_i spanned by $\mathbf{v}_{i,1}, \dots, \mathbf{v}_{i,k_i}$, for every $\mathbf{x}_i \in S_i$, we can write

$$\mathbf{x}_i = \sum_{j=1}^{k_i} c_{i,j} \mathbf{v}_{i,j}$$

To show that the value $\lambda_{\tilde{k}}$ is achievable, note that

$$\frac{\langle \mathcal{X}, \mathcal{C} \star \mathcal{X} \rangle}{\langle \mathcal{X}, \mathcal{X} \rangle} = \frac{\frac{1}{p} \sum_{i=1}^{p} \mathbf{x}_{i}^{\mathrm{H}} \mathbf{A}_{i} \mathbf{x}_{i}}{\frac{1}{p} \sum_{i=1}^{p} \mathbf{x}_{i}^{\mathrm{H}} \mathbf{x}_{i}} = \frac{\sum_{i=1}^{p} \sum_{j=1}^{k_{i}} \lambda_{i,j} c_{i,j}^{*} c_{i,j}}{\sum_{i=1}^{p} \sum_{j=1}^{k_{i}} c_{i,j}^{*} c_{i,j}}$$
$$\geq \frac{\sum_{i=1}^{p} \sum_{j=1}^{k_{i}} \lambda_{\tilde{k}} c_{i,j}^{*} c_{i,j}}{\sum_{i=1}^{p} \sum_{j=1}^{k_{i}} c_{i,j}^{*} c_{i,j}} = \lambda_{\tilde{k}}.$$
(2.7)

To verify that this is the maximum, let $T_{\tilde{i}}$ be the projection of T to the space with dimension $k_{\tilde{i}}$ with dimension $n-k_{\tilde{i}}+1$, then the intersection of S and $T_{\tilde{i}}$ is not empty. We have

$$\min_{\mathcal{X}\in S} \frac{\langle \mathcal{X}, \mathcal{C}\star\mathcal{X} \rangle}{\langle \mathcal{X}, \mathcal{X} \rangle} \leq \min_{\mathcal{X}\in S\cap T} \frac{\langle \mathcal{X}, \mathcal{C}\star\mathcal{X} \rangle}{\langle \mathcal{X}, \mathcal{X} \rangle}.$$
(2.8)

Any such $\mathbf{x}_{\tilde{i}} \in S \cap T_{\tilde{i}}$ can be expressed as $\mathbf{x}_{\tilde{i}} = \sum_{j=k_{\tilde{i}}}^{m} c_{\tilde{i},j} \mathbf{v}_{\tilde{i}j}$, and any *i* for $i \neq \tilde{i}$, we have

 $\mathbf{x}_i \in S \cap T_i$ expressed as $\mathbf{x}_i = \sum_{j=k_i+1}^m c_{i,j} \mathbf{v}_{i,j}$. Then, we have

$$\frac{\langle \mathcal{X}, \mathcal{C} \star \mathcal{X} \rangle}{\langle \mathcal{X}, \mathcal{X} \rangle} = \frac{\frac{1}{p} \sum_{i=1}^{p} \mathbf{x}_{i}^{\mathrm{H}} \mathbf{A}_{i} \mathbf{x}_{i}}{\frac{1}{p} \sum_{i=1}^{p} \mathbf{x}_{i}^{\mathrm{H}} \mathbf{x}_{i}} = \frac{\sum_{i=1}^{p} \sum_{j=k_{i}+1; i \neq \tilde{i}}^{m} \lambda_{i,j} c_{i,j}^{*} c_{i,j}}{\sum_{j=k_{\tilde{i}}; i=\tilde{i}}^{p} \sum_{j=k_{\tilde{i}}; i=\tilde{i}}^{m} c_{i,j}^{*} c_{i,j}} \\
\leq \frac{\sum_{i=1}^{p} \sum_{j=k_{i}+1; i \neq \tilde{i}}^{m} \lambda_{\tilde{k}} c_{i,j}^{*} c_{i,j}}{\sum_{i=1}^{p} \sum_{j=k_{\tilde{i}}; i=\tilde{i}}^{m} c_{i,j}^{*} c_{i,j}} = \lambda_{\tilde{k}}.$$
(2.9)

Therefore, for all subspaces S of dimensions $\{k_1, \dots, k_p\}$, we have

$$\min_{\mathcal{X}\in S} \frac{\langle \mathcal{X}, \mathcal{C} \star \mathcal{X} \rangle}{\langle \mathcal{X}, \mathcal{X} \rangle} \leq \lambda_{\tilde{k}}.$$

Thus, we complete the proof.

Given a symmetric T-product tensor C with associated matrices \mathbf{C}_i provided by Eq. (2.1), next theorem is the representation of the summation of all the largest k_i T-eigenvalues of \mathbf{C}_i and the summation of all the smallest k_i T-eigenvalues of \mathbf{C}_i .

Theorem 2.2. Let $C \in \mathbb{R}^{m \times m \times p}$ be a symmetric *T*-product tensor with associated matrices \mathbf{C}_i provided by Eq. (2.1), and we sort *T*-eigenvalues of the matrix \mathbf{C}_i as $\lambda_{i,1} \geq \lambda_{i,2} \geq \cdots \geq \lambda_{i,k_i}$. Then, we have

$$\sum_{i=1}^{p} \max_{\mathbf{U}_{i}\mathbf{U}_{i}^{\mathrm{H}}=\mathbf{I}_{k_{i}}} \operatorname{Tr}\left(\mathbf{U}_{i}\mathbf{C}_{i}\mathbf{U}_{i}^{\mathrm{H}}\right) = \sum_{i=1}^{p} \sum_{j=1}^{k_{i}} \lambda_{i,j}(\mathbf{C}_{i})$$
(2.10)

and

$$\sum_{i=1}^{p} \min_{\mathbf{U}_{i}\mathbf{U}_{i}^{\mathrm{H}}=\mathbf{I}_{k_{i}}} \operatorname{Tr}\left(\mathbf{U}_{i}\mathbf{C}_{i}\mathbf{U}_{i}^{\mathrm{H}}\right) = \sum_{i=1}^{p} \sum_{j=1}^{k_{i}} \lambda_{i,m-j+1}(\mathbf{C}_{i}), \qquad (2.11)$$

where \mathbf{U}_i are $k_i \times m$ complex matrices.

Proof. From Theorem 5 in [27], we may assume that C_i are diagonal matrices, denoted as D_i , since C_i are symmetric T-product matrices. Therefore, we have the expression

$$\mathbf{C}_i = \mathbf{V} \mathbf{D}_i \mathbf{V}^{\mathrm{H}}.$$

Then, we have

$$\operatorname{Tr}\left(\mathbf{U}_{i}\mathbf{D}_{i}\mathbf{U}_{i}^{\mathrm{H}}\right) = \sum_{j=1}^{k_{i}} \sum_{l=1}^{m} u_{j,l}^{*} u_{j,l} \lambda_{i,l}(\mathbf{C}_{i}) = \sum_{j=1}^{k_{i}} \sum_{l=1}^{m} p_{j,l} \lambda_{i,l}(\mathbf{C}_{i})$$
$$= \begin{bmatrix} k_{i} \text{ terms} \\ 1, 1, \cdots, 1 \end{bmatrix} \mathbf{P} \begin{bmatrix} \lambda_{i,1} \\ \lambda_{i,2} \\ \vdots \\ \lambda_{i,m} \end{bmatrix}, \qquad (2.12)$$

where $\mathbf{P} = (p_{j,l})$ is a $k_i \times m$ stochastic matrix. Then, we can concatenate an $(m - k_i) \times m$ matrix \mathbf{Q} to the matrix \mathbf{P} to make the following matrix $\begin{bmatrix} \mathbf{P} \\ \mathbf{Q} \end{bmatrix}$ as doubly stochastic from 2.C.1(4) from [16]. Then, Eq. (2.12) can be expressed as

$$\operatorname{Tr}\left(\mathbf{U}_{i}\mathbf{D}_{i}\mathbf{U}_{i}^{\mathrm{H}}\right) = \begin{bmatrix} \lambda_{i} \text{ terms} & m-k_{i} \text{ terms} \\ 1,1,\cdots,1, & 0,0,\cdots,0 \end{bmatrix} \begin{bmatrix} \mathbf{P} \\ \mathbf{Q} \end{bmatrix} \begin{bmatrix} \lambda_{i,1} \\ \lambda_{i,2} \\ \vdots \\ \lambda_{i,m} \end{bmatrix}.$$
(2.13)

Given two lists of real numbers, $[a_1, \dots, a_n]$ and $[b_1, \dots, b_n]$, we use $[a_1, \dots, a_n] \prec [b_1, \dots, b_n]$ to represent the following relationships:

$$\sum_{i=1}^{k} a_i \le \sum_{i=1}^{k} b_i \tag{2.14}$$

holds for any k between 1 and n. From Eq. (2.13), we have

$$[\lambda_{i,1},\cdots,\lambda_{i,m}][\mathbf{P}^{\mathrm{T}},\mathbf{Q}^{\mathrm{T}}]\prec[\lambda_{i,1},\cdots,\lambda_{i,m}]$$

and 3.H.2.b from [16], we have

$$\operatorname{Tr}\left(\mathbf{U}_{i}\mathbf{C}_{i}\mathbf{U}_{i}^{\mathrm{H}}\right) \leq \sum_{j=1}^{k_{i}} \lambda_{i,j}(\mathbf{C}_{i})$$

$$(2.15)$$

and

$$\operatorname{Tr}\left(\mathbf{U}_{i}\mathbf{C}_{i}\mathbf{U}_{i}^{\mathrm{H}}\right) \geq \sum_{j=1}^{k_{i}} \lambda_{i,m-j+1}(\mathbf{C}_{i}).$$

$$(2.16)$$

Finally, this theorem is proved by applying $\sum_{i=1}^{p}$ to both sides of Eqs. (2.15) and (2.16) with respect to the index *i*, and note that

$$\mathbf{U}_i \mathbf{V}_i = (\mathbf{I}_{k_i}, \mathbf{O})$$
 and $\mathbf{U}_i \mathbf{V}_i = (\mathbf{O}, \mathbf{I}_{k_i}),$

respectively.

2.2 Unitarily invariant T-product tensor norms

Let us represent the T-eigenvalues of a symmetric T-product tensor $\mathcal{H} \in \mathbb{R}^{m \times m \times p}$ in decreasing order by the vector

$$\vec{\lambda}(\mathcal{H}) = (\lambda_1(\mathcal{H}), \cdots, \lambda_{m \times p}(\mathcal{H})),$$

where $m \times p$ is the total number of T-eigenvalues. We use $\mathbb{R}_{\geq 0}(\mathbb{R}_{>0})$ to represent a set of nonnegative (positive) real numbers. Let $\|\cdot\|_{\rho}$ be a unitarily invariant tensor norm, i.e.,

$$\left\|\mathcal{H}\star\mathcal{U}\right\|_{
ho} = \left\|\mathcal{U}\star\mathcal{H}\right\|_{
ho} = \left\|\mathcal{H}\right\|_{
ho},$$

where \mathcal{U} is any unitary tensor. Let $\rho: \mathbb{R}_{\geq 0}^{m \times p} \to \mathbb{R}_{\geq 0}$ be the corresponding gauge function that satisfies Hölder's inequality so that

$$\|\mathcal{H}\|_{\rho} = \||\mathcal{H}|\|_{\rho} = \rho(\vec{\lambda}(|\mathcal{H}|)), \qquad (2.17)$$

where $|\mathcal{H}| = \sqrt{\mathcal{H}^H \star \mathcal{H}}$. The bijective correspondence between symmetric gauge functions on $\mathbb{R}^{m \times p}_{>0}$ and unitarily invariant norms is due to von Neumann [6].

Several popular norms can be treated as special cases of unitarily invariant tensor norm. The first one is Ky Fan like k-norm [6] for tensors. For $k \in \{1, 2, \dots, m \times p\}$, the Ky Fan k-norm [6] for tensors $\mathcal{H} \in \mathbb{R}^{m \times m \times p}$, denoted as $\|\mathcal{H}\|_{(k)}$, is defined as:

$$\left\|\mathcal{H}\right\|_{(k)} = \sum_{i=1}^{k} \lambda_i(|\mathcal{H}|).$$
(2.18)

If k=1, the Ky Fan k-norm for tensors is the tensor operator norm, denoted as $\|\mathcal{H}\|$. The second one is Schatten *p*-norm for tensors, denoted as $\|\mathcal{H}\|_p$, is defined as:

$$\left\|\mathcal{H}\right\|_{p} = (\mathrm{Tr}|\mathcal{H}|^{p})^{\frac{1}{p}},\tag{2.19}$$

where $p \ge 1$. If p = 1, it is the trace norm.

32

2.3 Antisymmetric Kronecker product for T-product tensors

In this section, we will discuss a machinery of antisymmetric Kronecker product for T-product tensors and this scheme will be used later for log-majorization results. Let \mathfrak{H} be an $m \times p$ -dimensional Hilbert space. For each $k \in \mathbb{N}$, let $\mathfrak{H}^{\otimes k}$ denote the k-fold Kronecker product of \mathfrak{H} , which is the $(m \times p)^k$ -dimensional Hilbert space with respect to the inner product defined by

$$\langle \mathbf{X}_1 \otimes \cdots \otimes \mathbf{X}_k, \mathbf{Y}_1 \otimes \cdots \otimes \mathbf{Y}_k \rangle = \prod_{i=1}^k \langle \mathbf{X}_i, \mathbf{Y}_i \rangle.$$
 (2.20)

For $\mathbf{X}_1, \cdots, \mathbf{X}_k \in \mathfrak{H}$, we define $\mathbf{X}_1 \wedge \cdots \wedge \mathbf{X}_k \in \mathfrak{H}^{\otimes k}$ by

$$\mathbf{X}_{1} \wedge \cdots \wedge \mathbf{X}_{k} = \frac{1}{\sqrt{k!}} \sum_{\sigma} (\operatorname{sgn}\sigma) \mathbf{X}_{\sigma(1)} \otimes \cdots \otimes \mathbf{X}_{\sigma(k)}, \qquad (2.21)$$

where σ runs over all permutations on $\{1, 2, \dots, k\}$ and $\operatorname{sgn} \sigma = \pm 1$ depending on σ is even or odd. The subspace of $\mathfrak{H}^{\otimes k}$ spanned by $\{\mathbf{X}_1 \wedge \dots \wedge \mathbf{X}_k\}$, where $\mathbf{X}_i \in \mathfrak{H}$, is named as k-fold antisymmetric Kronecker product of \mathfrak{H} and represented by $\mathfrak{H}^{\wedge k}$.

For each $C \in \mathbb{R}^{m \times m \times p}$ and $k \in \mathbb{N}$, the k-fold Kronecker product $C^{\otimes k} \in \mathbb{R}^{m^k \times m^k \times p^k}$ is given by

$$\mathcal{C}^{\otimes k} \star (\mathbf{X}_1 \otimes \cdots \otimes \mathbf{X}_k) = (\mathcal{C} \star \mathbf{X}_1) \otimes \cdots \otimes (\mathcal{C} \star \mathbf{X}_k).$$
(2.22)

Because $\mathfrak{H}^{\wedge k}$ is invariant for $\mathcal{C}^{\otimes k}$, the antisymmetric Kronecker product of $\mathcal{C}^{\wedge k}$ of \mathcal{C} can be defined as $\mathcal{C}^{\wedge k} = \mathcal{C}^{\otimes}|_{\mathfrak{H}^{\wedge k}}$, then we have

$$\mathcal{C}^{\wedge k} \star (\mathbf{X}_1 \wedge \dots \wedge \mathbf{X}_k) = (\mathcal{C} \star \mathbf{X}_1) \wedge \dots \wedge (\mathcal{C} \star \mathbf{X}_k).$$
(2.23)

We will provide the following lemmas about antisymmetric Kronecker product.

Lemma 2.1. Let $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{E} \in \mathbb{R}^{m \times m \times p}$ be *T*-product tensors, for any $k \in \{1, 2, \dots, m \times p\}$, we have

1. $(\mathcal{A}^{\wedge k})^{\mathrm{T}} = (\mathcal{A}^{\mathrm{T}})^{\wedge k}$.

2.
$$(\mathcal{A}^{\wedge k}) \star (\mathcal{B}^{\wedge k}) = (\mathcal{A} \star \mathcal{B})^{\wedge k}$$
.

- 3. If $\lim_{i\to\infty} \|\mathcal{A}_i \mathcal{A}\| \to 0$, then $\lim_{i\to\infty} \|\mathcal{A}_i^{\wedge k} \mathcal{A}^{\wedge k}\| \to 0$.
- 4. If $\mathcal{C} \succeq \mathcal{O}$ (zero tensor), then $\mathcal{C}^{\wedge k} \succeq \mathcal{O}$ and $(\mathcal{C}^p)^{\wedge k} = (\mathcal{C}^{\wedge k})^p$ for all $p \in \mathbb{R}_{>0}$.

- 5. $|\mathcal{A}|^{\wedge k} = |\mathcal{A}^{\wedge k}|.$
- 6. If $\mathcal{E} \succeq \mathcal{O}$ and \mathcal{E} is invertible, $(\mathcal{E}^z)^{\wedge k} = (\mathcal{E}^{\wedge k})^z$ for all $z \in \mathbb{E}$.

$$\gamma. \|\mathcal{E}^{\wedge k}\| = \prod_{i=1}^{k} \lambda_i(|\mathcal{E}|).$$

The proof can be found at Lemma 3 in [3].

3 Multivariate T-product tensor norm inequalities

In this section, the majorization with integral average and log-majorization with integral average will be introduced in Section 3.1 and Section 3.2. These majorization results will be used to prove T-product tensor norm inequalities in Section 3.3. The basic concept about majorization and its applications can refer to [16].

3.1 Majorization with integral average

Let Ω be a σ -compact metric space and ν a probability measure on the Borel σ -field of Ω . Let $\mathcal{C}, \mathcal{D}_{\tau} \in \mathbb{R}^{m \times m \times p}$ be symmetric T-product tensors. We further assume that tensors $\mathcal{C}, \mathcal{D}_{\tau}$ are uniformly bounded in their norm for $\tau \in \Omega$. Let $\tau \in \Omega \to \mathcal{D}_{\tau}$ be a continuous function such that $\sup\{\|\mathcal{D}_{\tau}\|: \tau \in \Omega\} < \infty$. For notational convenience, we define the following relation:

$$\left[\int_{\Omega} \lambda_1(\mathcal{D}_{\tau}) d\nu(\tau), \cdots, \int_{\Omega} \lambda_{m \times p}(\mathcal{D}_{\tau}) d\nu(\tau)\right] = \int_{\Omega^{m \times p}} \vec{\lambda}(\mathcal{D}_{\tau}) d\nu^{m \times p}(\tau).$$
(3.1)

If f is a single variable function, the notation $f(\mathcal{C})$ represents a tensor function with respect to the tensor \mathcal{C} .

Theorem 3.1. Let Ω , ν , C, \mathcal{D}_{τ} be defined as the beginning part of Section 3.1, and $f: \mathbb{R} \to [0, \infty)$ be a non-decreasing convex function, we have following two equivalent statements:

$$\vec{\lambda}(\mathcal{C}) \prec_w \int_{\Omega^{m \times p}} \vec{\lambda}(\mathcal{D}_\tau) d\nu^{m \times p}(\tau) \iff \|f(\mathcal{C})\|_{\rho} \leq \int_{\Omega} \|f(\mathcal{D}_\tau)\|_{\rho} d\nu(\tau), \tag{3.2}$$

where $\|\cdot\|_{o}$ is the unitarily invariant norm defined in Eq. (2.17).

Proof. We assume that the left statement of Eq. (3.2) is true and the function f is a non-decreasing convex function. From Lemma 1 in [8], we have

$$\vec{\lambda}(f(\mathcal{C})) = f(\vec{\lambda}(\mathcal{C})) \prec_w f\left(\int_{\Omega^{m \times p}} \vec{\lambda}(\mathcal{D}_\tau) d\nu^{m \times p}(\tau)\right).$$
(3.3)

From the convexity of f, we also have

$$f\left(\int_{\Omega^{m\times p}} \vec{\lambda}(\mathcal{D}_{\tau}) d\nu^{m\times p}(\tau)\right) \leq \int_{\Omega^{m\times p}} f(\vec{\lambda}(\mathcal{D}_{\tau})) d\nu^{m\times p}(\tau)$$
$$= \int_{\Omega^{m\times p}} \vec{\lambda}(f(\mathcal{D}_{\tau})) d\nu^{m\times p}(\tau). \tag{3.4}$$

Then, we obtain

$$\vec{\lambda}(f(\mathcal{C})) \prec_w = \int_{\Omega^{m \times p}} \vec{\lambda}(f(\mathcal{D}_{\tau})) d\nu^{m \times p}(\tau)$$

By applying Lemma 4.4.2 in [7] to both sides of

$$\vec{\lambda}(f(\mathcal{C})) \prec_w = \int_{\Omega^{m \times p}} \vec{\lambda}(f(\mathcal{D}_{\tau})) d\nu^{m \times p}(\tau)$$

with gauge function ρ , we obtain

$$\|f(\mathcal{C})\|_{\rho} \leq \rho \left(\int_{\Omega^{m \times p}} \vec{\lambda}(f(\mathcal{D}_{\tau})) d\nu^{m \times p}(\tau) \right)$$

$$\leq \int_{\Omega} \rho(\vec{\lambda}(f(\mathcal{D}_{\tau}))) d\nu(\tau) = \int_{\Omega} \|f(\mathcal{D}_{\tau})\|_{\rho} d\nu(\tau).$$
(3.5)

Therefore, the right statement of Eq. (3.2) is true from the left statement.

On the other hand, if the right statement of Eq. (3.2) is true, we select a function $f = \max\{x+c,0\}$, where c is a positive real constant satisfying $\mathcal{C}+c\mathcal{I} \geq \mathcal{O}$, $\mathcal{D}_{\tau}+c\mathcal{I} \geq \mathcal{O}$ for all $\tau \in \Omega$, and tensors $\mathcal{C}+c\mathcal{I}$, $\mathcal{D}_{\tau}+c\mathcal{I}$. If the Ky Fan k-norm at the right statement of Eq. (3.2) is applied, we have

$$\sum_{i=1}^{k} (\lambda_i(\mathcal{C}) + c) \leq \sum_{i=1}^{k} \int_{\Omega} (\lambda_i(\mathcal{D}_{\tau}) + c) d\nu(\tau).$$
(3.6)

Hence,

$$\sum_{i=1}^{k} \lambda_i(\mathcal{C}) \leq \sum_{i=1}^{k} \int_{\Omega} \lambda_i(\mathcal{D}_{\tau}) d\nu(\tau)$$

this is the left statement of Eq. (3.2).

Next theorem will provide a stronger version of Theorem 3.1 by removing weak majorization conditions.

Theorem 3.2. Let Ω , ν , C, \mathcal{D}_{τ} be defined as the beginning part of Section 3.1, and $f: \mathbb{R} \to [0,\infty)$ be a convex function, we have following two equivalent statements:

$$\vec{\lambda}(\mathcal{C}) \prec \int_{\Omega^{m \times p}} \vec{\lambda}(\mathcal{D}_{\tau}) d\nu^{m \times p}(\tau) \iff \|f(\mathcal{C})\|_{\rho} \leq \int_{\Omega} \|f(\mathcal{D}_{\tau})\|_{\rho} d\nu(\tau), \tag{3.7}$$

where $\|\cdot\|_{\rho}$ is the unitarily invariant norm defined in Eq. (2.17).

Proof. We assume that the left statement of Eq. (3.7) is true and the function f is a convex function. Again, from Lemma 1 in [8], we have

$$\vec{\lambda}(f(\mathcal{A})) = f(\vec{\lambda}(\mathcal{A})) \prec_{w} f\left(\left(\int_{\Omega^{m \times p}} \vec{\lambda}(\mathcal{D}_{\tau}) d\nu^{m \times p}(\tau)\right)\right)$$
$$\leq \int_{\Omega^{m \times p}} f(\vec{\lambda}(\mathcal{D}_{\tau})) d\nu^{m \times p}(\tau), \tag{3.8}$$

then,

$$\|f(\mathcal{A})\|_{\rho} \leq \rho \left(\int_{\Omega^{m \times p}} f(\vec{\lambda}(\mathcal{D}_{\tau})) d\nu^{m \times p}(\tau) \right)$$

$$\leq \int_{\Omega} \rho \left(f(\vec{\lambda}(\mathcal{D}_{\tau})) \right) d\nu(\tau) = \int_{\Omega} \|f(\mathcal{D}_{\tau})\|_{\rho} d\nu(\tau).$$
(3.9)

This proves the right statement of Eq. (3.7).

Now, we assume that the right statement of Eq. (3.7) is true. From Theorem 3.1, we already have

$$\vec{\lambda}(\mathcal{C}) \prec_w \int_{\Omega^{m \times p}} \vec{\lambda}(\mathcal{D}_{\tau}) d\nu^{m \times p}(\tau).$$

It is enough to prove

$$\sum_{i=1}^{m \times p} \lambda_i(\mathcal{C}) \ge \int_{\Omega} \sum_{i=1}^{m \times p} \lambda_i(\mathcal{D}_{\tau}) d\nu(\tau)$$

We define a function $f = \max\{c-x, 0\}$, where c is a positive real constant satisfying $C \leq c\mathcal{I}$, $\mathcal{D}_{\tau} \leq c\mathcal{I}$ for all $\tau \in \Omega$ and tensors $c\mathcal{I} - \mathcal{C}$, $c\mathcal{I} - \mathcal{D}_{\tau}$. If the trace norm is applied, i.e., the sum of the absolute value of all eigenvalues of a symmetric T-product tensor, then the right statement of Eq. (3.7) becomes

$$\sum_{i=1}^{m \times p} \lambda_i (c\mathcal{I} - \mathcal{C}) \leq \int_{\Omega} \sum_{i=1}^{m \times p} \lambda_i (c\mathcal{I} - \mathcal{D}_{\tau}) d\nu(\tau).$$
(3.10)

The desired inequality

$$\sum_{i=1}^{m \times p} \lambda_i(\mathcal{C}) \ge \int_{\Omega} \sum_{i=1}^{m \times p} \lambda_i(\mathcal{D}_{\tau}) d\nu(\tau)$$

is established.

3.2 Log-majorization with integral average

The purpose of this section is to consider log-majorization issues for unitarily invariant norms of TPSD T-product tensors. In this section, let $C, \mathcal{D}_{\tau} \in \mathbb{R}^{m \times m \times p}$ be TPSD T-product tensors with $m \times p$ nonnegative T-eigenvalues by keeping notations with the same definitions as at the beginning of the Section 3.1. For notational convenience, we define the following relation for logarithm vector:

$$\left[\int_{\Omega} \log \lambda_1(\mathcal{D}_{\tau}) d\nu(\tau), \cdots, \int_{\Omega} \log \lambda_{m \times p}(\mathcal{D}_{\tau}) d\nu(\tau)\right] = \int_{\Omega^{m \times p}} \log \vec{\lambda}(\mathcal{D}_{\tau}) d\nu^{m \times p}(\tau).$$

Theorem 3.3. Let C, \mathcal{D}_{τ} be TPSD T-product tensors, $f:(0,\infty) \to [0,\infty)$ be a continuous function such that the mapping $x \to \log f(e^x)$ is convex on \mathbb{R} , and $g:(0,\infty) \to [0,\infty)$ be a continuous function such that the mapping $x \to g(e^x)$ is convex on \mathbb{R} , then we have following three equivalent statements:

$$\vec{\lambda}(\mathcal{C}) \prec_{w\log} \exp \int_{\Omega^{m \times p}} \log \vec{\lambda}(\mathcal{D}_{\tau}) d\nu^{m \times p}(\tau),$$
 (3.11a)

$$\|f(\mathcal{C})\|_{\rho} \leq \exp \int_{\Omega} \log \|f(\mathcal{D}_{\tau})\|_{\rho} d\nu(\tau), \qquad (3.11b)$$

$$\|g(\mathcal{C})\|_{\rho} \leq \int_{\Omega} \|g(\mathcal{D}_{\tau})\|_{\rho} d\nu(\tau).$$
(3.11c)

Proof. The roadmap of this proof is to prove equivalent statements between Eq. (3.11a) and Eq. (3.11b) first, followed by equivalent statements between Eq. (3.11a) and Eq. (3.11c).

Eq. (3.11a) \implies Eq. (3.11b).

There are two cases to be discussed in this part of proof: C, D_{τ} are TPD tensors, and C, D_{τ} are TPSD T-product tensors. At the beginning, we consider the case that C, D_{τ} are TPD tensors.

37

Since \mathcal{D}_{τ} are positive, we can find $\varepsilon > 0$ such that $\mathcal{D}_{\tau} \ge \varepsilon \mathcal{I}$ for all $\tau \in \Omega$. From Eq. (3.11a), the convexity of $\log f(e^x)$ and Lemma 1 in [8], we have

$$\vec{\lambda}(f(\mathcal{C})) = f\left(\exp\left(\log\vec{\lambda}(\mathcal{C})\right)\right) \prec_{w} f\left(\exp\int_{\Omega^{m\times p}} \vec{\lambda}(\mathcal{D}_{\tau}) d\nu^{m\times p}(\tau)\right)$$
$$\leq \exp\left(\int_{\Omega^{m\times p}} \log f\left(\vec{\lambda}(\mathcal{D}_{\tau})\right) d\nu^{m\times p}(\tau)\right).$$
(3.12)

Then, from Eq. (2.17), we obtain

$$\|f(\mathcal{C})\|_{\rho} \leq \rho \left(\exp\left(\int_{\Omega^{m \times p}} \log f\left(\vec{\lambda}(\mathcal{D}_{\tau})\right) d\nu^{m \times p}(\tau)\right) \right).$$
(3.13)

From the function f properties, we can assume that f(x) > 0 for any x > 0. Then, we have following bounded and continuous maps on Ω : $\tau \to \log f(\lambda_i(\mathcal{D}_{\tau}))$ for $i \in \{1, 2, \dots, m \times p\}$, and $\tau \to \log \|f(\mathcal{D}_{\tau})\|_{\rho}$. Because we have $\nu(\Omega) = 1$ and σ -compactness of Ω , we have $\tau_k^{(n)} \in \Omega$ and $\alpha_k^{(n)}$ for $k \in \{1, 2, \dots, n\}$ and $n \in \mathbb{N}$ with $\sum_{k=1}^n \alpha_k^{(n)} = 1$ such that

$$\int_{\Omega} \log f(\lambda_i(\mathcal{D}_{\tau})) d\nu(\tau) = \lim_{n \to \infty} \sum_{k=1}^n \alpha_k^{(n)} \log f(\lambda_i(\mathcal{D}_{\tau_k^{(n)}})),$$
(3.14)

where $i \in \{1, 2, \dots, m \times p\}$. Moreover,

$$\int_{\Omega} \log \|f(\mathcal{D}_{\tau})\|_{\rho} d\nu(\tau) = \lim_{n \to \infty} \sum_{k=1}^{n} \alpha_k^{(n)} \log \left\|f(\mathcal{D}_{\tau_k^{(n)}})\right\|_{\rho}.$$
(3.15)

By taking the exponential at both sides of Eq. (3.14) and apply the gauge function ρ , we have

$$\rho\left(\exp\int_{\Omega^{m\times p}}\log f(\vec{\lambda}(\mathcal{D}_{\tau}))d\nu^{m\times p}(\tau)\right) = \lim_{n\to\infty}\rho\left(\prod_{k=1}^{n}f\left(\vec{\lambda}\left(\mathcal{D}_{\tau_{k}^{(n)}}\right)\right)^{\alpha_{k}^{(n)}}\right).$$
 (3.16)

Similarly, by taking the exponential at both sides of Eq. (3.15), we have

$$\exp\left(\int_{\Omega} \log \|f(\mathcal{D}_{\tau})\|_{\rho} d\nu(\tau)\right) = \lim_{n \to \infty} \prod_{k=1}^{n} \left\|f\left(\mathcal{D}_{\tau_{k}^{(n)}}\right)\right\|_{\rho}^{\alpha_{k}^{(n)}}.$$
(3.17)

From Lemma 2 in [8], we have

$$\rho\left(\prod_{k=1}^{n} f\left(\vec{\lambda}\left(\mathcal{D}_{\tau_{k}^{(n)}}\right)\right)^{\alpha_{k}^{(n)}}\right) \leq \prod_{k=1}^{n} \rho^{\alpha_{k}^{(n)}} \left(f\left(\vec{\lambda}\left(\mathcal{D}_{\tau_{k}^{(n)}}\right)\right)\right) \\
= \prod_{k=1}^{n} \rho^{\alpha_{k}^{(n)}} \left(\vec{\lambda}\left(f\left(\mathcal{D}_{\tau_{k}^{(n)}}\right)\right)\right) = \prod_{k=1}^{n} \left\|f\left(\mathcal{D}_{\tau_{k}^{(n)}}\right)\right\|_{\rho}^{\alpha_{k}^{(n)}}.$$
(3.18)

S. Chang and Y. Wei / Ann. Appl. Math., 38 (2022), pp. 25-61

From Eqs. (3.16), (3.17) and (3.18), we have

$$\rho\left(\exp\int_{\Omega^{m\times p}}\log f(\vec{\lambda}(\mathcal{D}_{\tau}))d\nu^{m\times p}(\tau)\right) \leq \exp\int_{\Omega}\log\|f(\mathcal{D}_{\tau})\|_{\rho}d\nu(\tau). \tag{3.19}$$

Then, Eq. (3.11b) is proved from Eqs. (3.13) and (3.19).

Next, we consider that C, \mathcal{D}_{τ} are TPSD T-product tensors. For any $\delta > 0$, we have following log-majorization relation:

$$\prod_{i=1}^{k} (\lambda_i(\mathcal{C}) + \epsilon_{\delta}) \leq \prod_{i=1}^{k} \exp \int_{\Omega} \log (\lambda_i(\mathcal{D}_{\tau}) + \delta) d\nu(\tau),$$
(3.20)

where $\epsilon_{\delta} > 0$ and $k \in \{1, 2, \dots r\}$. Then, we can apply the previous case result about TPD tensors to TPD tensors $C + \epsilon_{\delta} \mathcal{I}$ and $\mathcal{D}_{\tau} + \delta \mathcal{I}$, and get

$$\|f(\mathcal{C}) + \epsilon_{\delta} \mathcal{I}\|_{\rho} \leq \exp \int_{\Omega} \log \|f(\mathcal{D}_{\tau}) + \delta \mathcal{I}\|_{\rho} d\nu(\tau).$$
(3.21)

As $\delta \rightarrow 0$, Eq. (3.21) will give us Eq. (3.11b) for TPSD T-product tensors.

Eq. (3.11a) \Leftarrow **Eq.** (3.11b).

We consider TPD tensors at first phase by assuming that \mathcal{D}_{τ} are TPD T-product tensors for all $\tau \in \Omega$. We may also assume that the tensor \mathcal{C} is a TPD T-product tensor. Since if this is a TPSD T-product tensor, i.e., some $\lambda_i = 0$, we always have following inequality valid:

$$\prod_{i=1}^{k} \lambda_i(\mathcal{C}) \le \prod_{i=1}^{k} \exp \int_{\Omega} \log \lambda_i(\mathcal{D}_{\tau}) d\nu(\tau).$$
(3.22)

If we apply $f(x) = x^p$ for p > 0 and $\|\cdot\|_{\rho}$ as Ky Fan k-norm in Eq. (3.11b), we have

$$\log \sum_{i=1}^{k} \lambda_{i}^{p}(\mathcal{C}) \leq \int_{\Omega} \log \sum_{i=1}^{k} \lambda_{i}^{p}(\mathcal{D}_{\tau}) d\nu(\tau).$$
(3.23)

If we add $\log \frac{1}{k}$ and multiply $\frac{1}{p}$ at both sides of Eq. (3.23), we have

$$\frac{1}{p}\log\left(\frac{1}{k}\sum_{i=1}^{k}\lambda_{i}^{p}(\mathcal{C})\right) \leq \int_{\Omega}\frac{1}{p}\log\left(\frac{1}{k}\sum_{i=1}^{k}\lambda_{i}^{p}(\mathcal{D}_{\tau})\right)d\nu(\tau).$$
(3.24)

From L'Hopital's Rule, if $p \rightarrow 0$, we have

$$\frac{1}{p} \log \left(\frac{1}{k} \sum_{i=1}^{k} \lambda_i^p(\mathcal{C}) \right) \rightarrow \frac{1}{k} \sum_{i=1}^{k} \log \lambda_i(\mathcal{C})$$
(3.25)

and

$$\frac{1}{p} \log \left(\frac{1}{k} \sum_{i=1}^{k} \lambda_i^p(\mathcal{D}_{\tau}) \right) \rightarrow \frac{1}{k} \sum_{i=1}^{k} \log \lambda_i(\mathcal{D}_{\tau}), \qquad (3.26)$$

where $\tau \in \Omega$. Appling Eqs. (3.25) and (3.26) into Eq. (3.24) and taking $p \to 0$, we have

$$\sum_{i=1}^{k} \lambda_i(\mathcal{C}) \leq \int_{\Omega} \sum_{i=1}^{k} \log \lambda_i(\mathcal{D}_{\tau}) d\nu(\tau).$$
(3.27)

Therefore, Eq. (3.11a) is true for TPD tensors.

For TPSD T-product tensors \mathcal{D}_{τ} , since Eq. (3.11b) is valid for $\mathcal{D}_{\tau} + \delta \mathcal{I}$ for any $\delta > 0$, we can apply the previous case result about TPD tensors to $\mathcal{D}_{\tau} + \delta \mathcal{I}$ and obtain

$$\prod_{i=1}^{k} \lambda_i(\mathcal{C}) \le \prod_{i=1}^{k} \exp \int_{\Omega} \log(\lambda_i(\mathcal{D}_{\tau}) + \delta) d\nu(\tau), \qquad (3.28)$$

where $k \in \{1, 2, \dots, r\}$. Eq. (3.11a) is still true for TPSD T-product tensors as $\delta \to 0$. Eq. (3.11a) \implies Eq. (3.11c).

If \mathcal{C} , \mathcal{D}_{τ} are TPD tensors, and $\mathcal{D}_{\tau} \geq \delta \mathcal{I}$ for all $\tau \in \Omega$. From Eq. (3.11a), we have

$$\vec{\lambda}(\log \mathcal{C}) = \log \vec{\lambda}(\mathcal{C}) \prec_w \int_{\Omega^{m \times p}} \log \vec{\lambda}(\mathcal{D}_{\tau}) d\nu^{m \times p}(\tau)$$
$$= \int_{\Omega^{m \times p}} \vec{\lambda}(\log \mathcal{D}_{\tau}) d\nu^{m \times p}(\tau).$$
(3.29)

If we apply Theorem 3.1 to $\log C$, $\log D_{\tau}$ with function $f(x) = g(e^x)$, where g is used in Eq. (3.11c), Eq. (3.11c) is implied.

If \mathcal{C} , \mathcal{D}_{τ} are TPSD T-product tensors and any $\delta > 0$, we can find $\epsilon_{\delta} \in (0, \delta)$ to satisfy following:

$$\prod_{i=1}^{k} (\lambda_i(\mathcal{C}) + \epsilon_{\delta}) \leq \prod_{i=1}^{k} \exp \int_{\Omega} \log(\lambda_i(\mathcal{D}_{\tau}) + \delta) d\nu(\tau).$$
(3.30)

Then, from TPD T-product tensor case, we have

$$\|g(\mathcal{C}+\epsilon_{\delta}\mathcal{I})\|_{\rho} \leq \int_{\Omega} \|g(\mathcal{D}_{\tau}+\delta\mathcal{I})\|_{\rho} d\nu(\tau).$$
(3.31)

Eq. (3.11c) is obtained by taking $\delta \rightarrow 0$ in Eq. (3.31).

Eq. (3.11a) \Leftarrow **Eq.** (3.11c).

For $k \in \{1, 2, \dots, r\}$, if we apply $g(x) = \log(\delta + x)$, where $\delta > 0$, and Ky Fan k-norm in Eq. (3.11c), we have

$$\sum_{i=1}^{k} \log(\delta + \lambda_i(\mathcal{C})) \le \sum_{i=1}^{k} \int_{\Omega} \log(\delta + \lambda_i(\mathcal{D}_{\tau})) d\nu(\tau).$$
(3.32)

Then, we have following relation as $\delta \rightarrow 0$:

$$\sum_{i=1}^{k} \log \lambda_i(\mathcal{C}) \leq \sum_{i=1}^{k} \int_{\Omega} \log \lambda_i(\mathcal{D}_{\tau}) d\nu(\tau).$$
(3.33)

Therefore, Eq. (3.11a) can be derived from Eq. (3.11c).

Next theorem will extend Theorem 3.3 to non-weak version.

Theorem 3.4. Let C, D_{τ} be TPSD T-product tensors with

$$\int_{\Omega} \left\| \mathcal{D}_{\tau}^{-p} \right\|_{\rho} d\nu(\tau) < \infty$$

for any p > 0, $f: (0,\infty) \to [0,\infty)$ be a continuous function such that the mapping $x \to \log f(e^x)$ is convex on \mathbb{R} , and $g: (0,\infty) \to [0,\infty)$ be a continuous function such that the mapping $x \to g(e^x)$ is convex on \mathbb{R} , then we have following three equivalent statements:

$$\vec{\lambda}(\mathcal{C}) \prec_{\log} \exp \int_{\Omega^{m \times p}} \log \vec{\lambda}(\mathcal{D}_{\tau}) d\nu^{m \times p}(\tau), \qquad (3.34a)$$

$$\|f(\mathcal{C})\|_{\rho} \leq \exp \int_{\Omega} \log \|f(\mathcal{D}_{\tau})\|_{\rho} d\nu(\tau), \qquad (3.34b)$$

$$\|g(\mathcal{C})\|_{\rho} \leq \int_{\Omega} \|g(\mathcal{D}_{\tau})\|_{\rho} d\nu(\tau).$$
(3.34c)

Proof. The proof plan is similar to the proof in Theorem 3.3. We prove the equivalence between Eq. (3.34a) and Eq. (3.34b) first, then prove the equivalence between Eq. (3.34a) and Eq. (3.34c).

Eq. (3.34a) \implies Eq. (3.34b).

First, we assume that C, \mathcal{D}_{τ} are TPD T-product tensors with $\mathcal{D}_{\tau} \geq \delta \mathcal{I}$ for all $\tau \in \Omega$. The corresponding part of the proof in Theorem 3.3 about TPD tensors C, \mathcal{D}_{τ} can be applied here.

For case that \mathcal{C} , \mathcal{D}_{τ} are TPSD T-product tensors, we have

$$\prod_{i=1}^{k} \lambda_i(\mathcal{C}) \le \prod_{i=1}^{k} \exp \int_{\Omega} \log(\lambda_i(\mathcal{D}_{\tau}) + \delta_n) d\nu(\tau), \qquad (3.35)$$

where $\delta_n > 0$ and $\delta_n \to 0$. Because we have

$$\int_{\Omega^{m \times p}} \log \left(\vec{\lambda}(\mathcal{D}_{\tau}) + \delta_n \right) d\nu^{m \times p}(\tau) \quad \to \quad \int_{\Omega^{m \times p}} \log \vec{\lambda}(\mathcal{D}_{\tau}) d\nu^{m \times p}(\tau), \tag{3.36}$$

as $n \to \infty$; from Lemma 12 in [8], we can find $\mathbf{a}^{(n)}$ with $n \ge n_0$ such that $a_1^{(n)} \ge \cdots \ge a_r^{(n)} > 0$, $\mathbf{a}^{(n)} \to \vec{\lambda}(\mathcal{C})$ and

$$\mathbf{a}^{(n)} \prec_{\log} \exp \int_{\Omega^{m \times p}} \log \vec{\lambda} (\mathcal{D}_{\tau} + \delta_n \mathcal{I}) d\nu^{m \times p}(\tau).$$

Selecting $\mathcal{C}^{(n)}$ with $\vec{\lambda}(\mathcal{C}^{(n)}) = \mathbf{a}^{(n)}$ and applying TPD tensors case to $\mathcal{C}^{(n)}$ and $\mathcal{D}_{\tau} + \delta_n \mathcal{I}$, we obtain

$$\left\| f(\mathcal{C}^{(n)}) \right\|_{\rho} \leq \exp \int_{\Omega} \log \left\| f(\mathcal{D}_{\tau} + \delta_n \mathcal{I}) \right\|_{\rho} d\nu(\tau), \qquad (3.37)$$

where $n \ge n_0$.

There are two situations for the function f near 0: $f(0^+) < \infty$ and $f(0^+) = \infty$. For the case with $f(0^+) < \infty$, we have

$$\left\| f(\mathcal{C}^{(n)}) \right\|_{\rho} = \rho(f(\mathbf{a}^{(n)})) \rightarrow \rho(f(\vec{\lambda}(\mathcal{C}))) = \left\| f(\mathcal{C}) \right\|_{\rho}$$
(3.38)

and

$$\|f(\mathcal{D}_{\tau} + \delta_n \mathcal{I})\|_{\rho} \to \|f(\mathcal{D}_{\tau})\|_{\rho}, \qquad (3.39)$$

where $\tau \in \Omega$ and $n \to \infty$. From Fatou–Lebesgue theorem, we then have

$$\limsup_{n \to \infty} \int_{\Omega} \log \| f(\mathcal{D}_{\tau} + \delta_n \mathcal{I}) \|_{\rho} d\nu(\tau) \leq \int_{\Omega} \log \| f(\mathcal{D}_{\tau}) \|_{\rho}.$$
(3.40)

By taking $n \to \infty$ in Eq. (3.37) and using Eqs. (3.38), (3.39), (3.40), we have Eq. (3.34b) for case that $f(0^+) < \infty$.

For the case with $f(0^+) = \infty$, we assume that $\int_{\Omega} \log \|f(\mathcal{D}_{\tau})\|_{\rho} d\nu(\tau) < \infty$, (since the inequality in Eq. (3.34b) is always true for $\int_{\Omega} \log \|f(\mathcal{D}_{\tau})\|_{\rho} d\nu(\tau) = \infty$). Since f

is decreasing on $(0,\epsilon)$ for some $\epsilon > 0$. We claim that the following relation is valid: there are two constants a, b > 0 such that

$$a \leq \|f(\mathcal{D}_{\tau} + \delta_n \mathcal{I})\|_{\rho} \leq \|f(\mathcal{D}_{\tau})\|_{\rho} + b \tag{3.41}$$

for all $\tau \in \Omega$ and $n \ge n_0$. If Eq. (3.41) is valid and

$$\int_{\Omega} \log \|f(\mathcal{D}_{\tau})\|_{\rho} d\nu(\tau) < \infty,$$

from Lebesgue's dominated convergence theorem, we also have Eq. (3.34b) for case that $f(0^+) = \infty$ by taking $n \to \infty$ in Eq. (3.37).

Below, we will prove the claim stated by Eq. (3.41). By the uniform boundedness of tensors \mathcal{D}_{τ} , there is a constant $\kappa > 0$ such that

$$0 < \mathcal{D}_{\tau} + \delta_n \mathcal{I} \le \kappa \mathcal{I}, \tag{3.42}$$

where $\tau \in \Omega$ and $n \ge n_0$. We may assume that \mathcal{D}_{τ} is TPD tensors because $||f(\mathcal{D}_{\tau})||_{\rho} = \infty$, i.e., Eq. (3.41) being true automatically, when \mathcal{D}_{τ} is TPSD T-product tensors. From SVD of symmetric T-product tensors, we have

$$f(\mathcal{D}_{\tau} + \delta_{n}\mathcal{I}) = \sum_{i', \text{ s.t. } \lambda_{i'}(\mathcal{D}_{\tau}) + \delta_{n} < \epsilon} f(\lambda_{i'}(\mathcal{D}_{\tau}) + \delta_{n})\mathcal{U}_{i'} \star \mathcal{U}_{i'}^{\mathrm{H}} + \sum_{j', \text{ s.t. } \lambda_{j'}(\mathcal{D}_{\tau}) + \delta_{n} \ge \epsilon} f(\lambda_{j'}(\mathcal{D}_{\tau}) + \delta_{n})\mathcal{U}_{j'} \star \mathcal{U}_{j'}^{\mathrm{H}} \leq \sum_{i', \text{ s.t. } \lambda_{i'}(\mathcal{D}_{\tau}) + \delta_{n} < \epsilon} f(\lambda_{i'}(\mathcal{D}_{\tau}))\mathcal{U}_{i'} \star \mathcal{U}_{i'}^{\mathrm{H}} + \sum_{j', \text{ s.t. } \lambda_{j'}(\mathcal{D}_{\tau}) + \delta_{n} \ge \epsilon} f(\lambda_{j'}(\mathcal{D}_{\tau}) + \delta_{n})\mathcal{U}_{j'} \star \mathcal{U}_{j'}^{\mathrm{H}} \leq f(\mathcal{D}_{\tau}) + \sum_{j', \text{ s.t. } \lambda_{j'}(\mathcal{D}_{\tau}) + \delta_{n} \ge \epsilon} f(\lambda_{j'}(\mathcal{D}_{\tau}) + \delta_{n})\mathcal{U}_{j'} \star \mathcal{U}_{j'}^{\mathrm{H}}.$$
(3.43)

Therefore, the claim in Eq. (3.41) follows by the triangle inequality for $\|\cdot\|_{\rho}$ and $f(\lambda_{j'}(\mathcal{D}_{\tau})+\delta_n)<\infty$ for $\lambda_{j'}(\mathcal{D}_{\tau})+\delta_n\geq\epsilon$.

Eq. (3.34a) \Leftarrow **Eq.** (3.34b).

The weak majorization relation

$$\prod_{i=1}^{k} \lambda_i(\mathcal{C}) \le \prod_{i=1}^{k} \exp \int_{\Omega} \log \lambda_i(\mathcal{D}_{\tau}) d\nu(\tau)$$
(3.44)

is valid for $k < m \times p$ from Eq. (3.11a) \implies Eq. (3.11b) in Theorem 3.3. We wish to prove that Eq. (3.44) becomes equal for $k = m \times p$. It is equivalent to prove that

$$\log \det(\mathcal{C}) \ge \int_{\Omega} \log \det(\mathcal{D}_{\tau}) d\nu(\tau),$$
 (3.45)

where $det(\cdot)$ is defined by Eq. (2.2). We can assume that

$$\int_{\Omega} \log \det(\mathcal{D}_{\tau}) d\nu(\tau) \ge -\infty$$

since Eq. (3.45) is true for

$$\int_{\Omega} \log \det(\mathcal{D}_{\tau}) d\nu(\tau) = -\infty.$$

Then, \mathcal{D}_{τ} are TPD tensors.

If we scale tensors C, D_{τ} as aC, aD_{τ} by some a > 0, we can assume $D_{\tau} \leq \mathcal{I}$ and $\lambda_i(D_{\tau}) \leq 1$ for all $\tau \in \Omega$ and $i \in \{1, 2, \dots, m \times p\}$. Then for any p > 0, we have

$$\frac{1}{m \times p} \left\| \mathcal{D}_{\tau}^{-\varrho} \right\|_{1} \leq \lambda_{r}^{-\varrho} (\mathcal{D}_{\tau}) \leq (\det(\mathcal{D}_{\tau}))^{-\varrho}, \qquad (3.46)$$

and

$$\frac{1}{\varrho} \log\left(\frac{\left\|\mathcal{D}_{\tau}^{-\varrho}\right\|_{1}}{m \times p}\right) \leq -\log \det(\mathcal{D}_{\tau}).$$
(3.47)

If we use tensor trace norm, represented by $\|\cdot\|_1$, as unitarily invariant tensor norm and $f(x) = x^{-\varrho}$ for any $\varrho > 0$ in Eq. (3.34b), we obtain

$$\log \left\| \mathcal{C}^{-\varrho} \right\|_1 \leq \int_{\Omega} \log \left\| \mathcal{D}_{\tau}^{-\varrho} \right\|_1 d\nu(\tau).$$
(3.48)

By adding $\log \frac{1}{m \times p}$ and multiplying $\frac{1}{\rho}$ for both sides of Eq. (3.48), we have

$$\frac{1}{\varrho} \log\left(\frac{\|\mathcal{C}^{-\varrho}\|_1}{m \times p}\right) \le \int_{\Omega} \frac{1}{\varrho} \log\left(\frac{\|\mathcal{D}_{\tau}^{-\varrho}\|_1}{m \times p}\right) d\nu(\tau).$$
(3.49)

Similar to Eqs. (3.25) and (3.26), we have following two relations as $\rho \rightarrow 0$:

$$\frac{1}{\varrho} \log\left(\frac{\|\mathcal{C}^{-\varrho}\|_1}{m \times p}\right) \to \frac{-1}{m \times p} \operatorname{logdet}(\mathcal{C})$$
(3.50)

and

$$\frac{1}{\varrho} \log\left(\frac{\|\mathcal{D}_{\tau}^{-\varrho}\|_{1}}{m \times p}\right) \to \frac{-1}{m \times p} \operatorname{logdet}(\mathcal{D}_{\tau}).$$
(3.51)

From Eq. (3.47) and Lebesgue's dominated convergence theorem, we have

$$\lim_{\varrho \to 0} \int_{\Omega} \frac{1}{\varrho} \log\left(\frac{\|\mathcal{D}_{\tau}^{-\varrho}\|_{1}}{m \times p}\right) d\nu(\tau) = \frac{-1}{m \times p} \int_{\Omega} \log\det(\mathcal{D}_{\tau})\nu(\tau).$$
(3.52)

Finally, we have Eq. (3.45) from Eqs. (3.49) and (3.52).

Eq. (3.34a) \implies Eq. (3.34c).

First, we assume that \mathcal{C} , \mathcal{D}_{τ} are TPD tensors and $\mathcal{D}_{\tau} \geq \delta \mathcal{I}$ for $\tau \in \Omega$. From Eq. (3.34a), we can apply Theorem 3.2 to $\log \mathcal{C}$, $\log \mathcal{D}_{\tau}$ and $f(x) = g(e^x)$ to obtain Eq. (3.34c).

For \mathcal{C} , \mathcal{D}_{τ} are TPSD T-product tensors, we can choose $\mathbf{a}^{(n)}$ and corresponding $\mathcal{C}^{(n)}$ for $n \ge n_0$ given $\delta_n \to 0$ with $\delta_n > 0$ as the proof in Eq. (3.34a) \Longrightarrow Eq. (3.34b). Since tensors $\mathcal{C}^{(n)}$, $\mathcal{D}_{\tau} + \delta_n \mathcal{I}$ are TPD T-product tensors, we then have

$$\left\|g(\mathcal{C}^{(n)})\right\|_{\rho} \leq \int_{\Omega} \left\|g(\mathcal{D}_{\tau} + \delta_n \mathcal{I})\right\|_{\rho} d\nu(\tau).$$
(3.53)

If $g(0^+) < \infty$, Eq. (3.34c) is obtained from Eq. (3.53) by taking $n \to \infty$. On the other hand, if $g(0^+) = \infty$, we can apply the argument similar to the portion about $f(0^+) = \infty$ in the proof for Eq. (3.34a) \Longrightarrow Eq. (3.34b) to get a, b > 0 such that

$$a \leq \|g(\mathcal{D}_{\tau} + \delta_n \mathcal{I})\|_{\rho} \leq \|g(\mathcal{D}_{\tau})\|_{\rho} + b, \qquad (3.54)$$

for all $\tau \in \Omega$ and $n \ge n_0$. Since the case that

$$\int_{\Omega} \|g(\mathcal{D}_{\tau})\|_{\rho} d\nu(\tau) = \infty$$

will have Eq. (3.34c), we only consider the case that

$$\int_{\Omega} \|g(\mathcal{D}_{\tau})\|_{\rho} d\nu(\tau) < \infty.$$

Then, we have Eq. (3.34c) from Eqs. (3.53), (3.54) and Lebesgue's dominated convergence theorem.

Eq. (3.34a) \Leftarrow **Eq.** (3.34c).

The weak majorization relation

$$\sum_{i=1}^{k} \log \lambda_i(\mathcal{C}) \le \sum_{i=1}^{k} \int_{\Omega} \log \lambda_i(\mathcal{D}_{\tau}) d\nu(\tau)$$
(3.55)

is true from the implication from Eq. (3.11a) to Eq. (3.11c) in Theorem 3.3. We have to show that this relation becomes identity for $k=m\times p$. If we apply $\|\cdot\|_{\rho}=\|\cdot\|_1$ and $g(x)=x^{-\varrho}$ for any $\varrho>0$ in Eq. (3.34c), we have

$$\frac{1}{\varrho} \log\left(\frac{\|\mathcal{C}^{-\varrho}\|_1}{m \times p}\right) \leq \frac{1}{\varrho} \log\left(\int_{\Omega} \frac{\|\mathcal{D}_{\tau}^{-\varrho}\|_1}{m \times p} d\nu(\tau)\right).$$
(3.56)

Then, we will get

$$\frac{-\log\det(\mathcal{C})}{m \times p} = \lim_{\varrho \to 0} \frac{1}{\varrho} \log\left(\frac{\|\mathcal{C}^{-\varrho}\|_{1}}{m \times p}\right) \\
\leq \lim_{\varrho \to 0} \frac{1}{p} \log\left(\int_{\Omega} \frac{\|\mathcal{D}_{\tau}^{-\varrho}\|_{1}}{m \times p} d\nu(\tau)\right) =_{1} \frac{-\int_{\Omega} \log\det(\mathcal{D}_{\tau}) d\nu(\tau)}{m \times p}, \quad (3.57)$$

which will prove the identity for Eq. (3.55) when $k = m \times p$. The equality in $=_1$ will be proved by the following Lemma 3.1.

Lemma 3.1. Let \mathcal{D}_{τ} be TPSD T-product tensors with $\int_{\Omega} \|\mathcal{D}_{\tau}^{-p}\|_{\rho} d\nu(\tau) < \infty$ for any p > 0, then we have

$$\lim_{p \to 0} \left(\frac{1}{p} \log \int_{\Omega} \frac{\|\mathcal{D}_{\tau}^{-p}\|_{1}}{m \times p} d\nu(\tau) \right) = -\frac{1}{m \times p} \int_{\Omega} \log \det(\mathcal{D}_{\tau}) d\nu(\tau).$$
(3.58)

The proof can be found at Lemma 6 in [3].

3.3 T-product tensor norm inequalities by majorization

In this section, we will apply derived majorization inequalities for T-product tensors to multivariate T-product tensor norm inequalities which will be used to bound random T-product tensor concentration inequalities in later sections. We will begin to present a Lie-Trotter product formula for tensors.

Lemma 3.2. Let $m \in \mathbb{N}$ and $(\mathcal{L}_k)_{k=1}^M$ be a finite sequence of bounded T-product tensors with dimensions $\mathcal{L}_k \in \mathbb{R}^{m \times m \times p}$, then we have

$$\lim_{n \to \infty} \left(\prod_{k=1}^{M} \exp\left(\frac{\mathcal{L}_k}{n}\right) \right)^n = \exp\left(\sum_{k=1}^{M} \mathcal{L}_k \right).$$
(3.59)

S. Chang and Y. Wei / Ann. Appl. Math., 38 (2022), pp. 25-61

The proof of this lemma can be found in Lemma 7 in [3].

Below, new multivariate norm inequalities for T-product tensors are provided according to previous majorization theorems.

Theorem 3.5. Let $C_i \in \mathbb{R}^{m \times m \times p}$ be TPD tensors, where $1 \le i \le n$, $\|\cdot\|_{\rho}$ be a unitarily invariant norm with corresponding gauge function ρ . For any continuous function $f:(0,\infty) \to [0,\infty)$ such that $x \to \log f(e^x)$ is convex on \mathbb{R} , we have

$$\left\| f\left(\exp\left(\sum_{i=1}^{n} \log \mathcal{C}_{i}\right) \right) \right\|_{\rho} \leq \exp \int_{-\infty}^{\infty} \log \left\| f\left(\left| \prod_{i=1}^{n} \mathcal{C}_{i}^{1+\iota t} \right| \right) \right\|_{\rho} \beta_{0}(t) dt, \quad (3.60)$$

where $\beta_0(t) = \frac{\pi}{2(\cosh(\pi t)+1)}$.

For any continuous function $g(0,\infty) \to [0,\infty)$ such that $x \to g(e^x)$ is convex on \mathbb{R} , we have

$$\left\| g\left(\exp\left(\sum_{i=1}^{n} \log \mathcal{C}_{i}\right) \right) \right\|_{\rho} \leq \int_{-\infty}^{\infty} \left\| g\left(\left| \prod_{i=1}^{n} \mathcal{C}_{i}^{1+\iota t} \right| \right) \right\|_{\rho} \beta_{0}(t) dt.$$
(3.61)

Proof. From Hirschman interpolation theorem [22] and $\theta \in [0,1]$, we have

$$\log|h(\theta)| \le \int_{-\infty}^{\infty} \log|h(\iota t)|^{1-\theta} \beta_{1-\theta}(t) dt + \int_{-\infty}^{\infty} \log|h(1+\iota t)|^{\theta} \beta_{\theta}(t) dt,$$
(3.62)

where h(z) be uniformly bounded on $S = \{z \in \mathbb{C} : 0 \leq \Re(z) \leq 1\}$ and holomorphic on S. The term $d\beta_{\theta}(t)$ is defined as :

$$\beta_{\theta}(t) = \frac{\sin(\pi\theta)}{2\theta(\cos(\pi t) + \cos(\pi\theta))}.$$
(3.63)

Let H(z) be a uniformly bounded holomorphic function with values in $\mathbb{C}^{m \times m \times p}$. Fix some $\theta \in [0,1]$ and let $\mathcal{U}, \mathcal{V} \in \mathbb{C}^{m \times m \times p}$ be normalized tensors such that

$$\langle \mathcal{U}, \mathcal{H}(\theta) \star \mathcal{V} \rangle = \| H(\theta) \|.$$

If we define h(z) as $h(z) = \langle \mathcal{U}, \mathcal{H}(z) \star \mathcal{V} \rangle$, we have following bound: $|h(z)| \leq ||H(z)||$ for all $z \in S$. From Hirschman interpolation theorem, we then have following interpolation theorem for tensor-valued function:

$$\log \|H(\theta)\| \le \int_{-\infty}^{\infty} \log \|H(\iota t)\|^{1-\theta} \beta_{1-\theta}(t) dt + \int_{-\infty}^{\infty} \log \|H(1+\iota t)\|^{\theta} \beta_{\theta}(t) dt.$$
(3.64)

Let

$$H(z) = \prod_{i=1}^{n} \mathcal{C}_{i}^{z}.$$

Then the first term in the R.H.S. of Eq. (3.64) is zero since $H(\iota t)$ is a product of unitary tensors. Then we have

$$\log \left\| \left| \prod_{i=1}^{n} \mathcal{C}_{i}^{\theta} \right|^{\frac{1}{\theta}} \right\| \leq \int_{-\infty}^{\infty} \log \left\| \prod_{i=1}^{n} \mathcal{C}_{i}^{1+\iota t} \right\| \beta_{\theta}(t) dt.$$
(3.65)

From Lemma 2.1, we have following relations:

$$\left|\prod_{i=1}^{n} \left(\wedge^{k} \mathcal{C}_{i}\right)^{\theta}\right|^{\frac{1}{\theta}} = \wedge^{k} \left|\prod_{i=1}^{n} \mathcal{C}_{i}^{\theta}\right|^{\frac{1}{\theta}}$$
(3.66)

and

$$\left|\prod_{i=1}^{n} \left(\wedge^{k} \mathcal{C}_{i}\right)^{1+\iota t}\right| = \wedge^{k} \left|\prod_{i=1}^{n} \mathcal{C}_{i}^{1+\iota t}\right|.$$
(3.67)

If Eq. (3.65) is applied to $\wedge^k C_i$ for $1 \le k \le r$, we have following log-majorization relation from Eqs. (3.66) and (3.67):

$$\log \vec{\lambda} \left(\left| \prod_{i=1}^{n} \mathcal{C}_{i}^{\theta} \right|^{\frac{1}{\theta}} \right) \prec \int_{-\infty}^{\infty} \log \vec{\lambda} \left| \prod_{i=1}^{n} \mathcal{C}_{i}^{1+\iota t} \right|^{\frac{1}{\theta}} \beta_{\theta}(t) dt.$$
(3.68)

Moreover, we have the equality condition in Eq. (3.68) for k = r due to following identies:

$$\det \left| \prod_{i=1}^{n} \mathcal{C}_{i}^{\theta} \right|^{\frac{1}{\theta}} = \det \left| \prod_{i=1}^{n} \mathcal{C}_{i}^{1+\iota t} \right| = \prod_{i=1}^{n} \det \mathcal{C}_{i}.$$
(3.69)

At this stage, we are ready to apply Theorem 3.4 for the log-majorization provided by Eq. (3.68) to get following facts:

$$\left\| f\left(\left| \prod_{i=1}^{n} \mathcal{C}_{i}^{\theta} \right|^{\frac{1}{\theta}} \right) \right\|_{\rho} \leq \exp \int_{-\infty}^{\infty} \log \left\| f\left(\left| \prod_{i=1}^{n} \mathcal{C}_{i}^{1+\iota t} \right| \right) \right\|_{\rho} \beta_{\theta}(t) dt$$
(3.70)

48

and

$$\left\| g\left(\left| \prod_{i=1}^{n} \mathcal{C}_{i}^{\theta} \right|^{\frac{1}{\theta}} \right) \right\|_{\rho} \leq \int_{-\infty}^{\infty} \left\| g\left(\left| \prod_{i=1}^{n} \mathcal{C}_{i}^{1+\iota t} \right| \right) \right\|_{\rho} \beta_{\theta}(t) dt.$$
(3.71)

From Lie product formula for tensors given by Lemma 3.2, we have

$$\left|\prod_{i=1}^{n} \mathcal{C}_{i}^{\theta}\right|^{\frac{1}{\theta}} \to \exp\left(\sum_{i=1}^{n} \log \mathcal{C}_{i}\right).$$
(3.72)

By setting $\theta \to 0$ in Eqs. (3.70), (3.71) and using Lie product formula given by Eq. (3.72), we will get Eqs. (3.60) and (3.61).

4 Applications of T-product tensor norm inequalities

The purpose of this section is to apply new derived T-product tensor norm inequalities to obtain random symmetric T-product tensor Bernstein bounds. In Section 4.1, Ky Fan k-norm inequalities for T-product tensors will be provided and such Ky Fan k-norm inequalities will be utilized to establish T-product tensor Bernstein bounds in Section 4.2 and Section 4.3.

4.1 Ky fan k-norm tail bounds

We will present several lemmas required to prove Ky Fan k-norm tail bounds.

We have following lemma about the majorization relation of T-singular values among T-product tensors summation.

Lemma 4.1. Given two symmetric T-product tensors $C, D \in \mathbb{R}^{m \times m \times p}$. We have following majorization relation about T-singular values:

$$\sigma(\mathcal{C} + \mathcal{D}) \prec_w \sigma(\mathcal{C}) + \sigma(\mathcal{D}). \tag{4.1}$$

Proof. Since we have

$$\sum_{i=1}^{p} \sum_{j=1}^{k_{i}} \sigma_{i,j}(\mathcal{C}+\mathcal{D}) =_{1} \max_{\mathbf{U}_{i}\mathbf{U}_{i}^{\mathrm{H}}=_{1}\mathbf{I}_{k_{i}}} \Re\left(\sum_{i=1}^{p} \operatorname{Tr}\mathbf{U}_{i}(\mathbf{C}_{i}+\mathbf{D}_{i})\mathbf{U}_{i}^{\mathrm{H}}\right)$$

$$\leq \max_{\mathbf{U}_{i}\mathbf{U}_{i}^{\mathrm{H}}=_{1}\mathbf{I}_{k_{i}}} \Re\left(\sum_{i=1}^{p} \operatorname{Tr}\mathbf{U}_{i}\mathbf{C}_{i}\mathbf{U}_{i}^{\mathrm{H}}\right) + \max_{\mathbf{U}_{i}\mathbf{U}_{i}^{\mathrm{H}}=_{1}\mathbf{I}_{k_{i}}} \Re\left(\sum_{i=1}^{p} \operatorname{Tr}\mathbf{U}_{i}\mathbf{D}_{i}\mathbf{U}_{i}^{\mathrm{H}}\right)$$

$$=_{2}\sum_{i=1}^{p} \sum_{j=1}^{k_{i}} \sigma_{i,j}(\mathcal{C}) + \sum_{i=1}^{p} \sum_{j=1}^{k_{i}} \sigma_{i,j}(\mathcal{D}), \qquad (4.2)$$

where \Re is the operation to take the real part, and the equalities $=_1$ and $=_2$ come from Theorem 2.2.

We are ready to introduce the following two lemmas about Ky Fan k-norm inequalities for the product of tensors (Lemma 4.2) and the summation of tensors (Lemma 4.3).

Lemma 4.2. Let $C_i \in \mathbb{R}^{m \times m \times p}$ be symmetric *T*-product tensors and let p_i be positive real numbers satisfying $\sum_{i=1}^{M} \frac{1}{p_i} = 1$. Then, we have

$$\left\| \left\| \prod_{i=1}^{M} \mathcal{C}_{i} \right\|_{(k)} \leq \prod_{i=1}^{M} \left(\left\| \left| \mathcal{C}_{i} \right|^{sp_{i}} \right\|_{(k)} \right)^{\frac{1}{p_{i}}} \leq \sum_{i=1}^{M} \frac{\left\| \left| \mathcal{C}_{i} \right|^{sp_{i}} \right\|_{(k)}}{p_{i}},$$
(4.3)

where $s \ge 1$ and $k \in \{1, 2, \cdots, m \times p\}$.

The proof can be found at Lemma 10 in [3].

Lemma 4.3. Let $C_i \in \mathbb{R}^{m \times m \times p}$ be symmetric *T*-product tensors, then we have

$$\left\| \left\| \sum_{i=1}^{M} \mathcal{C}_{i} \right\|_{(k)}^{s} \le m^{s-1} \sum_{i=1}^{M} \left\| \left| \mathcal{C}_{i} \right|^{s} \right\|_{(k)},$$
(4.4)

where $s \ge 1$ and $k \in \{1, 2, \cdots, m \times p\}$.

The proof can be found at Lemma 11 in [3].

Now, we are ready to present our main theorem about Ky Fan k-norm probability bound for a function of tensors summation.

Theorem 4.1. Consider a sequence $\{\mathcal{X}_j \in \mathbb{R}^{m \times m \times p}\}$ of independent, random, symmetric *T*-product tensors. Let g(x) be a polynomial function with degree *n* and non-negative coefficients a_0, a_1, \dots, a_n raised by power $s \ge 1$, then g(x) can be expressed as:

$$g(x) = (a_0 + a_1 x + \dots + a_n x^n)^s.$$
(4.5)

Suppose following condition is satisfied:

$$g\left(\exp\left(t\sum_{j=1}^{M}\mathcal{X}_{j}\right)\right) \succeq \exp\left(tg\left(\sum_{j=1}^{M}\mathcal{X}_{j}\right)\right) \quad almost \; surely,$$
 (4.6)

where t > 0. Then, we have

$$\Pr\left(\left\|g\left(\sum_{j=1}^{M} \mathcal{X}_{j}\right)\right\|_{(k)} \ge \theta\right)$$

$$\leq (n+1)^{s-1} \inf_{t, \ p_{j}} \exp\left(-\theta t\right) \cdot \left(ka_{0}^{s} + \sum_{l=1}^{n} \sum_{j=1}^{M} \frac{a_{l}^{ls} \mathbb{E} \|\exp\left(p_{j} lst \mathcal{X}_{j}\right)\|_{(k)}}{p_{j}}\right), \qquad (4.7)$$

where $\sum_{j=1}^{M} \frac{1}{p_j} = 1$ and $p_j > 0$.

Proof. Let t > 0 be a parameter to be chosen later. Then

$$\Pr\left(\left\|g\left(\sum_{j=1}^{M} \mathcal{X}_{j}\right)\right\|_{(k)} \ge \theta\right)$$
$$=\Pr\left(\left\|\exp\left(tg\left(\sum_{j=1}^{M} \mathcal{X}_{j}\right)\right)\right\|_{(k)} \ge \exp(\theta t)\right)$$
$$\le_{1}\exp(-\theta t)\mathbb{E}\left(\left\|\exp\left(tg\left(\sum_{j=1}^{M} \mathcal{X}_{j}\right)\right)\right\|_{(k)}\right)$$
$$\le_{2}\exp(-\theta t)\mathbb{E}\left(\left\|g\left(\exp\left(t\sum_{j=1}^{M} \mathcal{X}_{j}\right)\right)\right\|_{(k)}\right), \tag{4.8}$$

where \leq_1 uses Markov's inequality, \leq_2 requires conditions provided by Eq. (4.6).

We can further bound the expectation term in Eq. (4.7) as

$$\mathbb{E}\left(\left\|g\left(\exp\left(t\sum_{j=1}^{M}\mathcal{X}_{j}\right)\right)\right\|_{(k)}\right)\right\|_{(k)}\right) \leq 3\mathbb{E}\int_{-\infty}^{\infty}\left\|g\left(\left|\prod_{j=1}^{M}e^{(1+\iota\tau)t\mathcal{X}_{j}}\right|\right)\right\|_{(k)}\beta_{0}(\tau)d\tau\right) \leq 4(n+1)^{s-1}\left(ka_{0}^{s}+\sum_{l=1}^{n}a_{l}^{ls}\mathbb{E}\int_{-\infty}^{\infty}\left\|\left|\prod_{j=1}^{M}e^{(1+\iota\tau)t\mathcal{X}_{j}}\right|^{ls}\right\|_{(k)}\beta_{0}(\tau)d\tau\right), \quad (4.9)$$

where \leq_3 from Eq. (3.61) in Theorem 3.5, \leq_4 is obtained from function g definition and Lemma 4.3. Again, the expectation term in Eq. (4.9) can be further bounded by Lemma 4.2 as

$$\mathbb{E} \int_{-\infty}^{\infty} \left\| \left\| \prod_{j=1}^{M} e^{(1+\iota\tau)t\mathcal{X}_{j}} \right\|_{(k)}^{ls} \right\|_{(k)} \beta_{0}(\tau) d\tau \\
\leq \mathbb{E} \int_{-\infty}^{\infty} \sum_{j=1}^{M} \frac{\left\| \left| e^{t\mathcal{X}_{j}} \right|^{p_{j}ls} \right\|_{(k)}}{p_{j}} \beta_{0}(\tau) d\tau = \sum_{j=1}^{M} \frac{\mathbb{E} \left\| e^{p_{j}lst\mathcal{X}_{j}} \right\|_{(k)}}{p_{j}}.$$
(4.10)

Note that the final equality is obtained due to that the integrand is independent of the variable τ and $\int_{-\infty}^{\infty} \beta_0(\tau) d\tau = 1$.

Finally, this theorem is established from Eqs. (4.8), (4.9), and (4.10).

Remark 4.1. The condition provided by Eq. (4.6) can be achieved by normalizing tensors \mathcal{X}_j through scaling.

4.2 T-product tensor Bernstein bound

In this section, we will present a tensor Bernstein bound for the maximum and the minimum T-eigenvalue for summation of random symmetric T-product tensors. We will provide the following definition to define a random structure for the T-product tensor $\mathcal{X} \in \mathbb{R}^{m \times m \times p}$.

Definition 4.1. Random structure for random symmetric T-product tensor $\mathcal{X} \in \mathbb{R}^{m \times m \times p}$

- S. Chang and Y. Wei / Ann. Appl. Math., 38 (2022), pp. 25-61
 - 1. There are p Hermitian matrices with size $m \times m$, denoted as $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_p$, obtained from Eq. (2.1). The entries for the matrix \mathbf{X}_i are denoted by $(x_{i_{j,k}})$, where $x_{i_{j,k}}$ is a complex number.
 - 2. For each \mathbf{X}_i , the random variables $x_{i_{j,j}}$, $\Re x_{i_{j,k}}$ for j < k, and $\Im x_{i_{j,k}}$ for j < k, are independent.
 - 3. For each \mathbf{X}_i , the random variables $x_{i_{j,j}}$ follow Gaussian distribution with zero mean and variance as $\frac{1}{m}$.
 - 4. For each \mathbf{X}_i , the random variables $\Re x_{i_{j,k}}$ for j < k, and $\Im x_{i_{j,k}}$ for j < k, follow Gaussian distribution with zero mean and variance as $\frac{1}{2m}$.

Following lemma is about the expectation of the largest T-eigenvalue of symmetric T-product tensor $\exp(\gamma \mathcal{X})$, where γ is a real number.

Lemma 4.4. Given a random symmetric T-product tensor $\mathcal{X} \in \mathbb{R}^{m \times m \times p}$ satisfying Definition 4.1 and any real number γ , we have

$$\mathbb{E}\lambda_{1}(\exp(\gamma\mathcal{X})) \leq \frac{3mc_{1}c_{2}}{2} \int_{-\infty}^{\infty} (y-2)^{1/2} \exp\left[\gamma y + c_{2}m(y-2)^{3/2}\right] dy$$

= $\Psi(m,\gamma,c_{1},c_{2}),$ (4.11)

where λ_1 is the largest T-eigenvalue, and c_1 , c_2 are constants related to the bound of cumulative distribution function of the largest eigenvalue of the random Hermitian matrix **X**.

The proof can be found at Lemma 12 in [3].

We are ready to present our theorem about the maximum and the minimum of T-eigenvalue for the summation of random symmetric T-product tensors.

Theorem 4.2 (T-product tensor Bernstein bound for T-eigenvalue). Consider a sequence $\{\mathcal{X}_j \in \mathbb{R}^{m \times m \times p}\}$ of independent, random, symmetric T-product tensors with random structure defined by Definition 4.1. Then we have following inequalities: given $\theta_1 > 0$, we have

$$\Pr\left(\lambda_{\max}\left(\sum_{j=1}^{M} \mathcal{X}_{j}\right) \ge \theta_{1}\right) \le \inf_{t>0} [\exp(-\theta_{1}t)\Psi(m, Mt, c_{1}, c_{2})], \quad (4.12)$$

and, given $\theta_2 < 0$, we have

$$\Pr\left(\lambda_{\min}\left(\sum_{j=1}^{M} \mathcal{X}_{j}\right) \leq \theta_{2}\right) \leq \inf_{t>0} [\exp(\theta_{2}t)\Psi(m, -Mt, c_{1}, c_{2})].$$
(4.13)

The Ψ function is defined by Eq. (4.11).

Proof. Since we have

$$\Pr\left(\lambda_{\max}\left(\sum_{j=1}^{M} \mathcal{X}_{j}\right) \geq \theta_{1}\right)$$

$$=_{1}\Pr\left(\sigma_{\max}\left(\sum_{j=1}^{M} \mathcal{X}_{j}\right) \geq \theta_{1}\right)$$

$$\leq_{2}\inf_{t,p_{j}}\exp(-\theta_{1}t)\left(\sum_{j=1}^{M} \frac{\mathbb{E}\sigma_{\max}(\exp(p_{j}t\mathcal{X}_{j}))}{p_{j}}\right)$$

$$\leq_{3}\inf_{t,p_{j}}\exp(-\theta_{1}t)\left(\sum_{j=1}^{M} \frac{\Psi(m,p_{j}t,c_{1},c_{2})}{p_{j}}\right)$$

$$\leq_{4}\inf_{t>0}\exp(-\theta_{1}t)\left(\sum_{j=1}^{M} \frac{\Psi(m,Mt,c_{1},c_{2})}{M}\right)$$

$$=\inf_{t>0}[\exp(-\theta_{1}t)\Psi(m,Mt,c_{1},c_{2})], \qquad (4.14)$$

where $=_1$ comes from that maximum singular value equals to the maximum absolute value of an T-eigenvalue and the maximum and the minimum of T-eigenvalue has same distribution due to the symmetry of random structure given by Definition 4.1; the inequality \leq_2 comes from Theorem 4.1 when g is the identity function; the equality \leq_3 comes from Lemma 4.4 and

$$\sigma_{\max}(\exp(p_j t \mathcal{X}_j)) = \lambda_{\max}(\exp(p_j t \mathcal{X}_j))$$

due to TPD of $\exp(p_j t \mathcal{X}_j)$; the inequality \leq_4 is obtained by selecting $p_j = M$. Therefore, we have Eq. (4.12).

For the minimum T-eigenvalue, we also have

$$\Pr\left(\lambda_{\min}\left(\sum_{j=1}^{M} \mathcal{X}_{j}\right) \leq \theta_{2}\right)$$

=_1 \Pr\left(\lambda_{\max}\left(\sum_{j=1}^{M} - \mathcal{X}_{j}\right) \geq -\theta_{2}\right)
=_2 \Pr\left(\sigma_{\pmax}\left(\sum_{j=1}^{M} - \mathcal{X}_{j}\right) \ge - \theta_{2}\right)
\left(\lambda_{\pmax}\left(\text{exp}(-p_{j}t\mathcal{X}_{j})\right) \ge - \theta_{2}\right)

S. Chang and Y. Wei / Ann. Appl. Math., 38 (2022), pp. 25-61

$$\leq_{4} \inf_{t, p_{j}} \exp(\theta_{2}t) \left(\sum_{j=1}^{M} \frac{\Psi(m, -p_{j}t, c_{1}, c_{2})}{p_{j}} \right)$$

$$\leq_{5} \inf_{t>0} [\exp(\theta_{2}t) \Psi(m, -Mt, c_{1}, c_{2})], \qquad (4.15)$$

where $=_1$ comes from Theorem 2.1; $=_2$ is true since the maximum singular value equals to the maximum absolute value of an T-eigenvalue and the maximum and the minimum of T-eigenvalue has same distribution due to the symmetry of random structure given by Definition 4.1; the inequality \leq_3 comes from Theorem 4.1 again when g is an identity map; the equality \leq_4 comes from Lemma 4.4 and

$$\sigma_{\max}(\exp(p_j t \mathcal{X}_j)) = \lambda_{\max}(\exp(p_j t \mathcal{X}_j))$$

due to TPD of $\exp(p_j t \mathcal{X}_j)$; the inequality \leq_5 is obtained by selecting $p_j = M$. Hence, we have Eq. (4.13).

4.3 Generalized T-product tensor Bernstein bound

In this section, we will present a generalized tensor Bernstein bound for Ky Fan k-norm, and we will begin with a lemma to bound exponential of a random T-product tensor.

Lemma 4.5. Suppose that $\mathcal{X} \in \mathbb{R}^{m \times m \times p}$ is a random symmetric T-product tensor that satisfies

$$\mathcal{X}^{p} \leq \frac{p!\mathcal{A}^{2}}{2}$$
 almost surely for $p = 2, 3, 4, \cdots,$ (4.16)

where \mathcal{A} is a fixed TPD tensor. Then, we have

$$e^{t\mathcal{X}} \leq \mathcal{I} + t\mathcal{X} + \frac{t^2 \mathcal{A}^2}{2(1-t)}$$
 almost surely, (4.17)

where 0 < t < 1.

The proof can be found at Lemma 13 in [3].

Lemma 4.6. Given a random symmetric T-product tensor $\mathcal{X} \in \mathbb{R}^{m \times m \times p}$ satisfying Definition 4.1, we have

$$\mathbb{E}\sigma_1(\mathcal{X}) \le \int_{-2}^{\infty} d_1 \exp(-d_2 m z^{3/2}) dz = \Phi(m, d_1, d_2), \tag{4.18}$$

where σ_1 is the largest T-singular value, and d_1 , d_2 are constants related to the upper bound of the largest eigenvalue of the random Hermitian matrix **X**. The proof can be found at Lemma 14 in [3].

Following lemma is about Ky Fan k-norm bound for the exponential of a random T-product tensor with subexponential constraints.

Lemma 4.7. Given a symmetric random T-product tensor $\mathcal{X} \in \mathbb{R}^{m \times m \times p}$ with random structure defined by Definition 4.1 and

$$\mathcal{X}^p \leq \frac{p!\mathcal{A}^2}{2}$$
 almost surely for $p = 2, 3, 4, \cdots,$ (4.19)

where \mathcal{A} is a TPD T-product tensor. Then, we have following bound about the expectation value of Ky Fan k-norm for the random T-product tensor $\exp(\theta \mathcal{X})$

$$\mathbb{E} \| \exp(\theta \mathcal{X}) \|_{(k)} \leq k \left[1 + \theta \Phi(m, d_1, d_2) + \frac{\theta^2}{2(1-\theta)} \sigma_1(\mathcal{A}^2) \right].$$
(4.20)

Proof. From Lemma 4.5, we have

$$\mathbb{E} \|\exp(\theta \mathcal{X})\|_{(k)} = \sum_{l=1}^{k} \mathbb{E}\sigma_{l}(\exp(\theta \mathcal{X}))$$

$$\leq \sum_{l=1}^{k} \mathbb{E}\sigma_{l}\left(\mathcal{I} + \theta \mathcal{X} + \frac{\theta^{2} \mathcal{A}^{2}}{2(1-\theta)}\right)$$

$$\leq k \mathbb{E}\sigma_{1}\left(\mathcal{I} + \theta \mathcal{X} + \frac{\theta^{2} \mathcal{A}^{2}}{2(1-\theta)}\right), \qquad (4.21)$$

where $\sigma_l(\cdot)$ is the *l*-th largest T-singular value.

From Lemma 4.1, we have

$$\sigma_1(\mathcal{A} + \mathcal{B}) \leq \sigma_1(\mathcal{A}) + \sigma_1(\mathcal{B})$$

for two symmetric T-product tensors \mathcal{A} and \mathcal{B} . Then, we can bound $\mathbb{E}\sigma_1(\mathcal{I}+\theta\mathcal{X}+\frac{\theta^2\mathcal{A}^2}{2(1-\theta)})$ as

$$\mathbb{E}\sigma_{1}\left(\mathcal{I}+\theta\mathcal{X}+\frac{\theta^{2}\mathcal{A}^{2}}{2(1-\theta)}\right)$$

$$\leq 1+\theta\mathbb{E}\sigma_{1}(\mathcal{X})+\frac{\theta^{2}}{2(1-\theta)}\sigma_{1}\left(\mathcal{A}^{2}\right)$$

$$\leq 1+\theta\mathbb{E}\Phi(m,d_{1},d_{2})+\frac{\theta^{2}}{2(1-\theta)}\sigma_{1}\left(\mathcal{A}^{2}\right),$$
(4.22)

where we use $\Phi(m, d_1, d_2)$ from Lemma 4.6 to bound $\mathbb{E}\sigma_1(\mathcal{X})$ in the last inequality. This Lemma is proved by multiplying k at Eq. (4.22). We are ready to prove our main theorem, Theorem 1.1, about the generalized T-product tensor Bernstein bound.

Proof. Since we have

$$\begin{aligned} &\Pr\left(\left\|g\left(\sum_{j=1}^{M}\mathcal{X}_{j}\right)\right\|_{(k)} \geq \theta\right) \\ \leq_{1}(n+1)^{s-1} \inf_{t,p_{j}} e^{-\theta t} \left(ka_{0}^{s} + \sum_{l=1}^{n} \sum_{j=1}^{M} \frac{a_{l}^{ls} \mathbb{E} \left\|\exp\left(p_{j} lst \mathcal{X}_{j}\right)\right\|_{(k)}}{p_{j}}\right) \\ \leq_{2}(n+1)^{s-1} \inf_{t,p_{j}} e^{-\theta t} \left\{ka_{0}^{s} + \sum_{l=1}^{n} \sum_{j=1}^{M} \frac{a_{l}^{ls} k \left[1 + p_{j} lst \Phi(m, d_{1}, d_{2}) + \frac{(p_{j} lst)^{2} \sigma_{1}(\mathcal{A}^{2})}{2(1 - p_{j} lst)}\right]}{p_{j}}\right\} \\ \leq_{3}(n+1)^{s-1} \inf_{t>0} e^{-\theta t} k \left\{a_{0}^{s} + \sum_{l=1}^{n} a_{l}^{ls} \left[1 + M lst \Phi(m, d_{1}, d_{2}) + \frac{(M lst)^{2} \sigma_{1}(\mathcal{A}^{2})}{2(1 - M lst)}\right]\right\}, \end{aligned}$$

where the inequality \leq_1 comes from Theorem 4.1; the inequality \leq_2 comes from Lemma 4.7; the inequality \leq_3 is obtained by setting $p_j = M$.

5 Covariance T-product tensor characterization by generalized T-product Bernstein bound

In this section, we will try to apply generalized T-product Bernstein Bound derived in Section 4.3 to bound Ky Fan norm of covariance T-product tensor induced by hypergraph signal processing. In [15], Marques et al. provide a comprehensive introduction to the spectral analysis and estimation of graph stationary processes based on graph signal processing (GSP). We extend their settings from vectors/matrices used in traditional GSP to hypergraph signal processing, where T-product tensors are applied to characterize hypergraph 3-uniform signals, i.e., signals are represented by three (3) dimensional data array [25].

Let $\mathfrak{G} = (\mathfrak{N}, \mathfrak{E})$ be a directed hypergraph with nodes set \mathfrak{N} and directed edges set \mathfrak{E} such that if there exists a hyperedge among a set of M nodes $(i, j, k) \in \mathfrak{E}$. We associate \mathfrak{G} with the hypergraph shift operator (HGSO) S, defined as an square T-product tensor with dimensions $m \times m \times p$ whose entry $s_{(i,j,k)} \neq 0$ if $(i, j, k) \in \mathfrak{E}$. We introduce a hypergraph filter $\mathcal{H}: \mathbb{R}^{m \times 1 \times p} \to \mathbb{R}^{m \times 1 \times p}$, defined as a linear hypergraph signal operator with the form

$$\mathcal{H} = \sum_{k=0}^{K-1} h_k \mathcal{S}^k, \tag{5.1}$$

where h_k are scalar coefficients. The covariance tensor of output signals $\mathbf{X} \in \mathbb{R}^{m \times 1 \times p}$ after filtering white input signals by hypergraph filter shown in Eq. (5.1) will be expressed as

$$\mathcal{C}_{\mathbf{X}}(\mathbf{h}) = \mathcal{H}^{\mathrm{H}} \star \mathcal{H} = \sum_{k=0, k'=0}^{K-1} h_k h_{k'} \mathcal{S}^k \star \left(\mathcal{S}^{\mathrm{H}}\right)^k$$
$$= \sum_{1}^{2(K-1)} \gamma_k \mathcal{S}^k, \qquad (5.2)$$

where $=_1$ is true if HGSO S is a symmetric T-product tensor. The coefficients $\gamma_k = \sum_{k'+k''=k} h_{k'} h_{k''}$.

It is shown by the work [20] that although the correlation information of signal is given by the *dense* tensor, the actual relation is easier to be described by the more sparse tensor S. Examples about relationships between the HGSO and the covariance tensor $C_{\mathbf{X}}(\mathbf{h})$ include

- $C_{\mathbf{X}}(\mathbf{h}) = \sum_{k=0}^{2(K-1)} \gamma_k \mathcal{S}^k$, as in graph filtering;
- $C_{\mathbf{X}}(\mathbf{h}) = S^{-1}$, as in in conditionally independent Markov random fields;
- $C_{\mathbf{X}}(\mathbf{h}) = (\mathcal{I} \mathcal{S})^{-2}$, as in symmetric structural equation models with white exogenous inputs.

In the sequel, we will bound the Ky Fan norm for the covariance tensor $C_{\mathbf{X}}(\mathbf{h})$ when $\mathbf{h} = [h_0, h_1]$. In random environment, suppose HGSO S is obtained by sample average as

$$\mathcal{S} = \frac{1}{M} \sum_{j=1}^{M} \mathcal{X}_j = \sum_{j=1}^{M} \mathcal{X}'_j, \qquad (5.3)$$

where $\mathcal{X}' = \frac{\mathcal{X}}{M}$. Since the graph filter coefficients are $\mathbf{h} = [h_0, h_1]$, from Eq. (5.2), the corresponding polynomial relation between $\mathcal{C}_{\mathbf{X}}(\mathbf{h})$ and \mathcal{S} is

$$\mathcal{C}_{\mathbf{X}}([h_0, h_1]) = h_0^2 + 2h_0 h_1 \mathcal{S} + h_1^2 \mathcal{S}^2, \qquad (5.4)$$

S. Chang and Y. Wei / Ann. Appl. Math., 38 (2022), pp. 25-61

which is the polynomial function

$$g(x) = (a_0 + a_1x + a_2x^2)^1 = h_0^2 + 2h_0h_1x + h_1^2x^2$$

in Theorem 1.1. We assume that random sampled tensors \mathcal{X}'_j are identical distributed as \mathcal{X}' are satisfy Eq. (1.1) and Eq. (1.2). Then we have following bound of Ky Fan norm for the covaraince $\mathcal{C}_{\mathbf{X}}[h_0,h_1]$ from Theorem 1.1:

$$\Pr\left(\left\|\mathcal{C}_{\mathbf{X}}([h_{0},h_{1}])\right\|_{(k)} \ge \theta\right)$$

$$\leq \inf_{t>0} ke^{-\theta t} \cdot \left\{a_{0} + \sum_{l=1}^{2} a_{l}^{l} \left[1 + Mlt\Phi(m,d_{1},d_{2}) + \frac{(Mlt)^{2}\sigma_{1}(\mathcal{A}^{2})}{2(1 - Mlt)}\right]\right\},$$
(5.5)

where $a_0 = h_0^2$, $a_1 = 2h_0h_1$, and $a_2 = h_1^2$. The usefulness of Eq. (5.5) is that we can control Ky Fan norm for the covariance $C_{\mathbf{X}}[h_0, h_1]$ via graph filter parameters h_0, h_1 , and this controllability is crucial in GSP system design.

6 Conclusions

This work extend previous work in [5] by making following generalizations via majorization techniques: (1) besides bounds related to extreme values of T-eigenvalues, this works considers more general unitarily invariant norm for T-product tensors; (2) this work derives new bounds for any polynomial function raised by any power greater or equal than one for the summation of random symmetric T-product tensors. We also establish the Courant-Fischer min-max theorem for T-product tensors and marjoization relation for T-singular values which are by-products of our procedure to prove the generalized random T-product Bernstein bounds. Eventually, we apply T-product Bernstein inequality to bound Ky Fan norm of covariance Tproduct tensor induced by hypergraph signal processing.

Possible future work about this research is to consider tail bounds behaviors for the summation of random symmetric T-product tensors equipped with other random structures different from random structure provided by Definition 4.1.

Acknowledgements

The helpful comments of the handling editor and the referee are gratefully acknowledged. Prof. Wei is supported by Innovation Program of Shanghai Municipal Education Commission and the National Natural Science Foundation of China under grant No. 11771099.

References

- P. J. Basser and S. Pajevic, A normal distribution for tensor-valued random variables: applications to diffusion tensor MRI, IEEE Trans. Med. Imag., 22 (2003), pp. 785– 794.
- [2] K. Braman, Third-order tensors as linear operators on a space of matrices, Linear Algebra Appl., 433 (2010), pp. 1241–1253.
- [3] S. Y. Chang and Y. Wei, Generalized T-product tensor Bernstein bounds, arXiv preprint arXiv: 2109.10880, (2021).
- [4] S. Y. Chang and Y. Wei, T-product tensors part I: Inequalities, Comput. Appl. Math., to appear, 2022.
- [5] S. Y. Chang and Y. Wei, T product tensors part II: Tail bounds for sums of random t product tensors, arXiv preprint arXiv:2107.06224, (2021).
- [6] K. Fan and A. J. Hoffman, Some metric inequalities in the space of matrices, Proc. Am. Math. Soc., 6 (1955), pp. 111–116.
- [7] F. Hiai, Matrix analysis: matrix monotone functions, matrix means, and majorization, Interdisciplinary Information Sciences, 16 (2010), pp. 139–248.
- [8] F. Hiai, R. König, and M. Tomamichel, Generalized Log-Majorization and Multivariate Trace Inequalities, Annales Henri Poincaré, vol. 18, Springer, 2017, pp. 2499–2521.
- [9] 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 J. Matrix Anal. Appl., 34 (2013), pp. 148–172.
- [10] M. E. Kilmer and C. D. Martin, Factorization strategies for third-order tensors, Linear Algebra Appl., 435 (2011), pp. 641–658.
- [11] 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).
- [12] W. Liu and X. Jin, A study on T-eigenvalues of third-order tensors, Linear Algebra Appl., 612 (2021), pp. 357–374.
- [13] K. Lund, The tensor T-function: A definition for functions of third-order tensors, Numer. Linear Algebra Appl., 27 (2020), e2288.
- [14] Z. Luo, L. Qi, and P. L. Toint, Tensor Bernstein concentration inequalities with an application to sample estimators for high-order moments, Front. Math. China, 15 (2020), pp. 367–384.
- [15] A. G. Marques, S. Segarra, G. Leus, and A. Ribeiro, Stationary graph processes and spectral estimation, IEEE Trans. Signal Proc., 65 (2017), pp. 5911–5926.
- [16] A. W. Marshall, I. Olkin, and B. C. Arnold, Inequalities: Theory of Majorization and Its Applications, vol. 143, Springer, 2011.
- [17] C. D. Martin, R. Shafer, and B. LaRue, An order-p tensor factorization with applications in imaging, SIAM J. Sci. Comput., 35 (2013), pp. A474–A490.
- [18] Y. Miao, L. Qi, and Y. Wei, Generalized tensor function via the tensor singular value decomposition based on the T-product, Linear Algebra Appl., 590 (2020), pp. 258–

303.

- [19] Y. Miao, L. Qi, and Y. Wei, T-Jordan canonical form and T-Drazin inverse based on the T-product, Commun. Appl. Math. Comput., 3 (2021), pp. 201–220.
- [20] M. Navarro, Y. Wang, A. G. Marques, C. Uhler, and S. Segarra, Joint inference of multiple graphs from matrix polynomials, arXiv preprint arXiv:2010.08120, (2020).
- [21] L. Qi and G. Yu, T-singular values and T-sketching for third order tensors, arXiv preprint arXiv: 2103.00976, (2021).
- [22] D. Sutter, M. Berta, and M. Tomamichel, Multivariate trace inequalities, Commun. Math. Phys., 352 (2017), pp. 37–58.
- [23] F. Szczepankiewicz, C.-F. Westin, and M. Nilsson, Gradient waveform design for tensor-valued encoding in diffusion MRI, J. Neurosci. Methods, 348 (2021), 109007.
- [24] J. Zhang, A. K. Saibaba, M. E. Kilmer, and S. Aeron, A randomized tensor singular value decomposition based on the t-product, Numer. Linear Algebra Appl., 25 (2018), e2179.
- [25] S. Zhang, Z. Ding, and S. Cui, Introducing hypergraph signal processing: Theoretical foundation and practical applications, IEEE Int. Things J., 7 (2019), pp. 639–660.
- [26] Z. Zhang and S. Aeron, Exact tensor completion using t-SVD, IEEE Trans. Signal Proc., 65 (2016), pp. 1511–1526.
- [27] M.-M. Zheng, Z.-H. Huang, and Y. Wang, T-positive semidefiniteness of thirdorder symmetric tensors and T-semidefinite programming, Comput. Optim. Appl., 78 (2021), pp. 239–272.
- [28] P. Zhou, C. Lu, Z. Lin, and C. Zhang, Tensor factorization for low-rank tensor completion, IEEE Trans. Image Proc., 27 (2017), pp. 1152–1163.