

1-1-2021

Tensor multivariate trace inequalities and their applications

Shih Yu Chang

San Jose State University, shihyu.chang@sjsu.edu

Hsiao Chun Wu

Louisiana State University

Follow this and additional works at: https://scholarworks.sjsu.edu/faculty_rsca

Recommended Citation

Shih Yu Chang and Hsiao Chun Wu. "Tensor multivariate trace inequalities and their applications"
Mathematics and Statistics (2021): 394-410. <https://doi.org/10.13189/ms.2021.090322>

This Article is brought to you for free and open access by SJSU ScholarWorks. It has been accepted for inclusion in Faculty Research, Scholarly, and Creative Activity by an authorized administrator of SJSU ScholarWorks. For more information, please contact scholarworks@sjsu.edu.

Tensor Multivariate Trace Inequalities and Their Applications

Shih Yu Chang^{1,*}, Hsiao-Chun Wu²

¹Department of Applied Data Science, San Jose State University, San Jose, CA 95192, USA

²Division of Electrical & Computer Engineering, Louisiana State University, Baton Rouge, LA 70803 USA

*Corresponding Author: shihyu.chang@sjsu.edu, hwu1@lsu.edu

Received February 2, 2021; Revised March 28, 2021; Accepted April 18, 2021

Cite This Paper in the following Citation Styles

(a): [1] Shih Yu Chang, Hsiao-ChunWu, "Tensor Multivariate Trace Inequalities and Their Applications," *Mathematics and Statistics*, Vol.9, No.3, pp. 394-410, 2021. DOI: 10.13189/ms.2021.090322

(b): Shih Yu Chang, Hsiao-ChunWu, (2021). *Tensor Multivariate Trace Inequalities and Their Applications*. *Mathematics and Statistics*, 9(3), 394-410. DOI: 10.13189/ms.2021.090322

Abstract In linear algebra, the trace of a square matrix is defined as the sum of elements on the main diagonal. The trace of a matrix is the sum of its eigenvalues (counted with multiplicities), and it is invariant under the change of basis. This characterization can be used to define the trace of a tensor in general. Trace inequalities are mathematical relations between different multivariate trace functionals involving linear operators. These relations are straightforward equalities if the involved linear operators commute, however, they can be difficult to prove when the non-commuting linear operators are involved. Given two Hermitian tensors \mathcal{H}_1 and \mathcal{H}_2 that do not commute. Does there exist a method to transform one of the two tensors such that they commute without completely destroying the structure of the original tensor? The spectral pinching method is a tool to resolve this problem. In this work, we will apply such spectral pinching method to prove several trace inequalities that extend the Araki–Lieb–Thirring (ALT) inequality, Golden–Thompson(GT) inequality and logarithmic trace inequality to arbitrary many tensors. Our approaches rely on complex interpolation theory as well as asymptotic spectral pinching, providing a transparent mechanism to treat generic tensor multivariate trace inequalities. As an example application of our tensor extension of the Golden–Thompson inequality, we give the tail bound for the independent sum of tensors. Such bound will play a fundamental role in high-dimensional probability and statistical data analysis.

Keywords Tensor, Multivariate, Trace, Golden–Thompson Inequality, Araki–Lieb–Thirring Inequality

AMS 15A69,46B28,47H60,47A30

1 Introduction

Trace inequalities are mathematical relations between different multivariate trace functionals involving linear operators. These relations are straightforward equalities if the involved linear operators commute, however, they can be difficult to prove when the non-commuting linear operators are involved [4].

One of the most important trace inequalities is the famous Golden-Thompson inequality [8]. For any two Hermitian matrices \mathbf{H}_1 and \mathbf{H}_2 , we have

$$\text{Tr exp}(\mathbf{H}_1 + \mathbf{H}_2) \leq \text{Tr exp}(\mathbf{H}_1) \text{exp}(\mathbf{H}_2). \quad (1)$$

It is easy to see that the Eq. (1) becomes an identity if two Hermitian matrices \mathbf{H}_1 and \mathbf{H}_2 are commute. The inequality in Eq. (1) has been generalized to several situations. For example, it has been demonstrated that it remains valid by replacing the trace with any unitarily invariant norm [14, 24]. The Golden-Thompson inequality has been applied to many various fields ranging from quantum information processing [16, 17], statistical physics [26, 29], and random matrix theory [1, 27].

The Golden-Thompson inequality can be seen as a limiting case of the more general Araki–Lieb–Thirring (ALT) inequality [3, 18]. For any two positive semi-definite matrices \mathbf{A}_1 and \mathbf{A}_2 with $r \in (0, 1]$ and $q > 0$, ALT states that

$$\text{Tr} \left(\mathbf{A}_1^{\frac{r}{2}} \mathbf{A}_2^r \mathbf{A}_1^{\frac{r}{2}} \right)^{\frac{q}{r}} \leq \text{Tr} \left(\mathbf{A}_1^{\frac{1}{2}} \mathbf{A}_2 \mathbf{A}_1^{\frac{1}{2}} \right)^q. \tag{2}$$

The Golden-Thompson inequality for Schatten p -norm is obtained by the Lie-Trotter product formula by taking limit $r \rightarrow 0$. The ALT inequality has also been expanded to various directions [2, 13, 28].

The following theorem is about logarithmic trace inequality which can be used to bound quantum information divergence [2, 9]. For any two positive semi-definite matrices \mathbf{A}_1 and \mathbf{A}_2 with $r \in (0, 1]$ and $p > 0$, logarithmic trace inequality for matrix is

$$\frac{1}{p} \text{Tr} \mathbf{A}_1 \log \mathbf{A}_2^{\frac{p}{2}} \mathbf{A}_1^p \mathbf{A}_2^{\frac{p}{2}} \leq \text{Tr} \mathbf{A}_1 (\log \mathbf{A}_1 + \log \mathbf{A}_2) \leq \frac{1}{p} \text{Tr} \mathbf{A}_1 \log \mathbf{A}_1^{\frac{p}{2}} \mathbf{A}_2^p \mathbf{A}_1^{\frac{p}{2}}. \tag{3}$$

The paper is organized as follows. Preliminaries of tensors are given in Section 2. In Section 3, the method of pinching and complex interpolation theory will be introduced. Three useful matrix-based trace inequalities are extended to multivariate tensors in Section 4. The new Golden-Thompson inequality is applied to provide tail bounds for sums of random tensors in Section 5. Finally, the conclusions are given in Section 6.

2 Tensors Preliminaries

Throughout this paper, we denote scalars by lower-case letters (e.g., d, e, f, \dots), vectors by boldfaced lower-case letters (e.g., $\mathbf{d}, \mathbf{e}, \mathbf{f}, \dots$), matrices by boldfaced capitalized letters (e.g., $\mathbf{D}, \mathbf{E}, \mathbf{F}, \dots$), and tensors by calligraphic letters (e.g., $\mathcal{D}, \mathcal{E}, \mathcal{F}, \dots$), respectively. Tensors are referred to as multiarrays of values which can be deemed high-dimensional generalizations from vectors and matrices. Given a positive integer N , let $[N] \stackrel{\text{def}}{=} 1, 2, \dots, N$. An *order- N tensor* (or *N -th order tensor*) is represented by $\mathcal{A} \stackrel{\text{def}}{=} (a_{i_1, i_2, \dots, i_N})$, where $1 \leq i_j \leq I_j$ for $j \in [N]$, is a multidimensional array containing $I_1 \times I_2 \times \dots \times I_N$ entries. Let $\mathbb{C}^{I_1 \times \dots \times I_N}$ and $\mathbb{R}^{I_1 \times \dots \times I_N}$ be the sets of order- N $I_1 \times \dots \times I_N$ tensors over the complex field \mathbb{C} and the real field \mathbb{R} , respectively. For example, $\mathcal{A} \in \mathbb{C}^{I_1 \times \dots \times I_N}$ is an order- N multiarray, where the first, second, ..., and N -th orders have I_1, I_2, \dots , and I_N entries, respectively. Thus, each entry of \mathcal{A} can be represented by a_{i_1, \dots, i_N} . For example, when $N = 4$, $\mathcal{A} \in \mathbb{C}^{I_1 \times I_2 \times I_3 \times I_4}$ is a fourth-order tensor containing entries a_{i_1, i_2, i_3, i_4} , where $i_j \in [I_j]$ for $j \in [4]$.

Without loss of generality, one can partition the dimensions of a tensor into two groups, say M and N dimensions, separately. Therefore, for two order- $(M+N)$ tensors: $\mathcal{A} \stackrel{\text{def}}{=} (a_{i_1, \dots, i_M, j_1, \dots, j_N}) \in \mathbb{C}^{I_1 \times \dots \times I_M \times J_1 \times \dots \times J_N}$ and $\mathcal{B} \stackrel{\text{def}}{=} (b_{i_1, \dots, i_M, j_1, \dots, j_N}) \in \mathbb{C}^{I_1 \times \dots \times I_M \times J_1 \times \dots \times J_N}$, according to [15], the *tensor addition* $\mathcal{A} + \mathcal{B} \in \mathbb{C}^{I_1 \times \dots \times I_M \times J_1 \times \dots \times J_N}$ is given by

$$(\mathcal{A} + \mathcal{B})_{i_1, \dots, i_M, j_1 \times \dots \times j_N} \stackrel{\text{def}}{=} a_{i_1, \dots, i_M, j_1 \times \dots \times j_N} + b_{i_1, \dots, i_M, j_1 \times \dots \times j_N}. \tag{4}$$

On the other hand, for tensors $\mathcal{A} = (a_{i_1, \dots, i_M, j_1, \dots, j_N}) \in \mathbb{C}^{I_1 \times \dots \times I_M \times J_1 \times \dots \times J_N}$ and $\mathcal{B} = (b_{j_1, \dots, j_N, k_1, \dots, k_L}) \in \mathbb{C}^{J_1 \times \dots \times J_N \times K_1 \times \dots \times K_L}$, according to [15], the *Einstein product* (or simply referred to as *tensor product* in this work) $\mathcal{A} \star_N \mathcal{B} \in \mathbb{C}^{I_1 \times \dots \times I_M \times K_1 \times \dots \times K_L}$ is given by

$$(\mathcal{A} \star_N \mathcal{B})_{i_1, \dots, i_M, k_1 \times \dots \times k_L} \stackrel{\text{def}}{=} \sum_{j_1, \dots, j_N} a_{i_1, \dots, i_M, j_1, \dots, j_N} b_{j_1, \dots, j_N, k_1, \dots, k_L}. \tag{5}$$

This tensor product will be reduced to the standard matrix multiplication as $L = M = N = 1$. Other simplified situations can also be extended as *tensor–vector product* ($M > 1, N = 1$, and $L = 0$) and *tensor–matrix product* ($M > 1$ and $N = L = 1$). In analogy to matrix analysis, we define some typical tensors and elementary tensor-operations as follows.

Definition 1. A tensor whose entries are all zero is called a zero tensor, denoted by \mathcal{O} .

Definition 2. An identity tensor $\mathcal{I} \in \mathbb{C}^{I_1 \times \dots \times I_N \times J_1 \times \dots \times J_N}$ is defined by

$$(\mathcal{I})_{i_1 \times \dots \times i_N \times j_1 \times \dots \times j_N} \stackrel{\text{def}}{=} \prod_{k=1}^N \delta_{i_k, j_k}, \tag{6}$$

where $\delta_{i_k, j_k} \stackrel{\text{def}}{=} 1$ if $i_k = j_k$; otherwise $\delta_{i_k, j_k} \stackrel{\text{def}}{=} 0$.

In order to define the *Hermitian* tensor, the *conjugate transpose operation* (or *Hermitian adjoint*) of a tensor is specified as follows.

Definition 3. Given a tensor $\mathcal{A} \stackrel{\text{def}}{=} (a_{i_1, \dots, i_M, j_1, \dots, j_N}) \in \mathbb{C}^{I_1 \times \dots \times I_M \times J_1 \times \dots \times J_N}$, its conjugate transpose, denoted by \mathcal{A}^H , is defined as

$$(\mathcal{A}^H)_{j_1, \dots, j_N, i_1, \dots, i_M} \stackrel{\text{def}}{=} \overline{a_{i_1, \dots, i_M, j_1, \dots, j_N}}, \quad (7)$$

where the overline notation indicates the complex conjugate of the complex number $a_{i_1, \dots, i_M, j_1, \dots, j_N}$. If a tensor \mathcal{A} satisfying $\mathcal{A}^H = \mathcal{A}$, then \mathcal{A} is a Hermitian tensor.

Definition 4. Given a tensor $\mathcal{A} \stackrel{\text{def}}{=} (a_{i_1, \dots, i_M, j_1, \dots, j_M}) \in \mathbb{C}^{I_1 \times \dots \times I_M \times J_1 \times \dots \times J_M}$, if

$$\mathcal{A}^H \star_M \mathcal{A} = \mathcal{A} \star_M \mathcal{A}^H = \mathcal{I} \in \mathbb{C}^{I_1 \times \dots \times I_M \times J_1 \times \dots \times J_M}, \quad (8)$$

then \mathcal{A} is a unitary tensor.

Definition 5. Given a square tensor $\mathcal{A} \in \mathbb{C}^{I_1 \times \dots \times I_M \times I_1 \times \dots \times I_M}$, if there exists $\mathcal{X} \in \mathbb{C}^{I_1 \times \dots \times I_M \times I_1 \times \dots \times I_M}$ such that

$$\mathcal{A} \star_M \mathcal{X} = \mathcal{X} \star_M \mathcal{A} = \mathcal{I}, \quad (9)$$

then \mathcal{X} is the inverse of \mathcal{A} . We usually write $\mathcal{X} \stackrel{\text{def}}{=} \mathcal{A}^{-1}$ thereby.

We also list other necessary tensor operations here. The *trace* of a tensor is equivalent to the summation of all diagonal entries such that

$$\text{Tr}(\mathcal{A}) \stackrel{\text{def}}{=} \sum_{1 \leq i_j \leq I_j, j \in [M]} \mathcal{A}_{i_1, \dots, i_M, i_1, \dots, i_M}. \quad (10)$$

The *inner product* of two tensors $\mathcal{A}, \mathcal{B} \in \mathbb{C}^{I_1 \times \dots \times I_M \times J_1 \times \dots \times J_N}$ is given by

$$\langle \mathcal{A}, \mathcal{B} \rangle \stackrel{\text{def}}{=} \text{Tr}(\mathcal{A}^H \star_M \mathcal{B}). \quad (11)$$

According to Eq. (11), the *Frobenius norm* of a tensor \mathcal{A} is defined by

$$\|\mathcal{A}\| \stackrel{\text{def}}{=} \sqrt{\langle \mathcal{A}, \mathcal{A} \rangle}. \quad (12)$$

Definition 6. Given a square tensor $\mathcal{A} \in \mathbb{C}^{I_1 \times \dots \times I_M \times I_1 \times \dots \times I_M}$, the tensor exponential of the tensor \mathcal{A} is defined as

$$e^{\mathcal{A}} \stackrel{\text{def}}{=} \sum_{k=0}^{\infty} \frac{\mathcal{A}^k}{k!}, \quad (13)$$

where \mathcal{A}^0 is defined as the identity tensor $\mathcal{I} \in \mathbb{C}^{I_1 \times \dots \times I_M \times I_1 \times \dots \times I_M}$ and

$$\mathcal{A}^k = \underbrace{\mathcal{A} \star_M \mathcal{A} \star_M \dots \star_M \mathcal{A}}_{k \text{ terms of } \mathcal{A}}.$$

Given a tensor \mathcal{B} , the tensor \mathcal{A} is said to be a tensor logarithm of \mathcal{B} if $e^{\mathcal{A}} = \mathcal{B}$.

Following definitions are about the Kronecker product and the sum of two tensors.

Definition 7. Given two tensors $\mathcal{A} \in \mathbb{C}^{I_1 \times \dots \times I_M \times J_1 \times \dots \times J_N}$ and $\mathcal{B} \in \mathbb{C}^{K_1 \times \dots \times K_P \times L_1 \times \dots \times L_Q}$, we define the Kronecker product of two tensors $\mathcal{A} \otimes \mathcal{B}$ as

$$\mathcal{A} \otimes \mathcal{B} \stackrel{\text{def}}{=} (a_{i_1, \dots, i_M, j_1, \dots, j_N} \mathcal{B})_{i_1, \dots, i_M, j_1, \dots, j_N}. \quad (14)$$

Definition 8. Given two square tensors $\mathcal{A} \in \mathbb{C}^{I_1 \times \dots \times I_M \times I_1 \times \dots \times I_M}$ and $\mathcal{B} \in \mathbb{C}^{J_1 \times \dots \times J_P \times J_1 \times \dots \times J_P}$, we define the Kronecker sum of two tensors $\mathcal{A} \oplus \mathcal{B}$ as

$$\mathcal{A} \oplus \mathcal{B} \stackrel{\text{def}}{=} \mathcal{A} \otimes \mathcal{I}_{J_1 \times \dots \times J_P} + \mathcal{I}_{I_1 \times \dots \times I_M} \otimes \mathcal{B}, \quad (15)$$

where $\mathcal{I}_{I_1 \times \dots \times I_M} \in \mathbb{C}^{I_1 \times \dots \times I_M \times I_1 \times \dots \times I_M}$ and $\mathcal{I}_{J_1 \times \dots \times J_P} \in \mathbb{C}^{J_1 \times \dots \times J_P \times J_1 \times \dots \times J_P}$ are identity tensors.

We require the following two lemmas about Kronecker product which will be used for later proof in Theorem 1.

Lemma 1. Given tensors \mathcal{A}_1 and \mathcal{A}_2 which act on spaces S_1 and S_2 , respectively, we have following identities :

$$\text{Tr}(\mathcal{A}_1 \otimes \mathcal{A}_2) = \text{Tr}(\mathcal{A}_1)\text{Tr}(\mathcal{A}_2); \tag{16}$$

and

$$\exp(\mathcal{A}_1) \otimes \exp(\mathcal{A}_2) = \exp(\mathcal{A}_1 \otimes \mathcal{I}_{S_2} + \mathcal{I}_{S_1} \otimes \mathcal{A}_2), \tag{17}$$

where \mathcal{I}_{S_1} and \mathcal{I}_{S_2} are identity tensors acts on spaces S_1 and S_2 , respectively.

Proof:

We prove Eq. (16) first. Suppose the tensor $\mathcal{A}_1 \in \mathbb{C}^{I_1 \times \dots \times I_M \times I_1 \times \dots \times I_M}$, then its entries will be $(a_{i_1, \dots, i_M, j_1, \dots, j_M})$. By definition of the Kronecker product, we have

$$\begin{aligned} \text{Tr}(\mathcal{A}_1 \otimes \mathcal{A}_2) &= \sum_{i_1, \dots, i_M} \text{Tr}(a_{i_1, \dots, i_M, i_1, \dots, i_M} \mathcal{A}_2) \\ &= \sum_{i_1, \dots, i_M} a_{i_1, \dots, i_M, i_1, \dots, i_M} \text{Tr}(\mathcal{A}_2) = \text{Tr}(\mathcal{A}_1)\text{Tr}(\mathcal{A}_2). \end{aligned} \tag{18}$$

Next, we will verify the relation provided by Eq. (17). Because we have

$$\exp(\mathcal{A}_1) \otimes \exp(\mathcal{A}_2) =_1 \exp(\mathcal{A}_1 \oplus \mathcal{A}_2) = \exp(\mathcal{A}_1 \otimes \mathcal{I}_{S_2} + \mathcal{I}_{S_1} \otimes \mathcal{A}_2), \tag{19}$$

where the equality $=_1$ comes from Theorems 2 and 3 in [19], and the last equality is provided by Definition 8. \square

Lemma 2. Given positive tensors \mathcal{A}_1 and \mathcal{A}_2 which act on spaces S_1 and S_2 , respectively, we have following identity :

$$\log(\mathcal{A}_1 \otimes \mathcal{A}_2) = (\log \mathcal{A}_1) \otimes \mathcal{I}_{S_2} + \mathcal{I}_{S_1} \otimes (\log \mathcal{A}_2), \tag{20}$$

where \mathcal{I}_{S_1} and \mathcal{I}_{S_2} are identity tensors acts on spaces S_1 and S_2 , respectively.

Proof:

From the relation (17) and set $\mathcal{B}_1 = \log(\mathcal{A}_1)$, $\mathcal{B}_2 = \log(\mathcal{A}_2)$, we have

$$\begin{aligned} \exp(\mathcal{B}_1) \otimes \exp(\mathcal{B}_2) &= \exp(\mathcal{B}_1 \otimes \mathcal{I}_{S_2} + \mathcal{I}_{S_1} \otimes \mathcal{B}_2) \\ \Leftrightarrow \mathcal{A}_1 \otimes \mathcal{A}_2 &= \exp(\log(\mathcal{A}_1) \otimes \mathcal{I}_{S_2} + \mathcal{I}_{S_1} \otimes \log(\mathcal{A}_2)) \end{aligned} \tag{21}$$

By taking log at both sides, we have desired result

$$\log(\mathcal{A}_1 \otimes \mathcal{A}_2) = (\log \mathcal{A}_1) \otimes \mathcal{I}_{S_2} + \mathcal{I}_{S_1} \otimes (\log \mathcal{A}_2). \tag{22}$$

\square

3 Tools for Hermitian Tensors

In this section, we will introduce two main techniques used to prove multivariate trace inequalities for tensors. Spectrum pinching method is discussed in Section 3.1, and complex interpolation theory is presented in Section 3.2.

3.1 Pinching Map

The purpose for studying the pinching method arises from the following problem: Given two Hermitian tensors \mathcal{H}_1 and \mathcal{H}_2 that do not commute. Does there exist a method to transform one of the two tensors such that they commute without completely destroying the structure of the original tensor? The spectral pinching method is a tool to resolve this problem. Before discussing this method in detail we have to introduce the pinching map.

Given a Hermitian tensor $\mathcal{H} \in \mathbb{C}^{I_1 \times \dots \times I_M \times I_1 \times \dots \times I_M}$, we have spectral decomposition as [20]:

$$\mathcal{H} = \sum_{\lambda \in \text{sp}(\mathcal{H})} \lambda \mathcal{U}_\lambda, \tag{23}$$

where $\lambda \in \text{sp}(\mathcal{H}) \in \mathbb{R}$ and $\mathcal{U}_\lambda \in \mathbb{C}^{I_1 \times \dots \times I_M \times I_1 \times \dots \times I_M}$ are mutually orthogonal tensors. The *pinching map* with respect to \mathcal{H} is defined as

$$\mathfrak{P}_{\mathcal{H}} : \mathcal{X} \rightarrow \sum_{\lambda \in \text{sp}(\mathcal{H})} \mathcal{U}_\lambda \star_M \mathcal{X} \star_M \mathcal{U}_\lambda, \tag{24}$$

where $\mathcal{X} \in \mathbb{C}^{I_1 \times \dots \times I_M \times I_1 \times \dots \times I_M}$ is a Hermitian tensor. The pinching map possesses various nice properties that will be discussed at this section. For example, $\mathfrak{P}_{\mathcal{H}}(\mathcal{X})$ always commutes with \mathcal{H} for any nonnegative tensor \mathcal{X} . Two lemmas are introduced first which will be used to prove several useful properties about pinching maps.

Lemma 3. *Let the tensor $\mathcal{A} \in \mathbb{C}^{I_1 \times \dots \times I_M}$, where $|I_i| = N$ for $i \in [M]$. The number of distinct eigenvalues of $\mathcal{A}^{\otimes m}$, where \otimes is the Kronecker product defined in Definition 7, grows polynomially with m .*

Proof: Let us use the symbol $\text{sp}(\mathcal{A}^{\otimes m})$ to represent the spectrum space, i.e., the space of eigenvalues. Because the number of distinct eigenvalues of $\mathcal{A}^{\otimes m}$, denoted as $|\text{sp}(\mathcal{A}^{\otimes m})|$, is bounded by the number of different types of sequences of $N(M - 1)^{N-1} \stackrel{\text{def}}{=} e_{\mathcal{A}}$ symbols of length m from methods of types [5], then we have

$$|\text{sp}(\mathcal{A}^{\otimes m})| \leq \binom{m + e_{\mathcal{A}} - 1}{e_{\mathcal{A}} - 1} \leq \frac{(m + e_{\mathcal{A}} - 1)^{e_{\mathcal{A}} - 1}}{(e_{\mathcal{A}} - 1)!} = O(\text{poly}(m)), \tag{25}$$

where $O(\text{poly}(m))$ represents a function that grows with m polynomially. When $m = 1$, the number of $|\text{sp}(\mathcal{A}^{\otimes})|$ is upper bounded by $e_{\mathcal{A}}$, which is the number of eigenvalues of \mathcal{A} , see Theorem 1.1 in [22]. \square

For any probability measure μ be a probability measure on a measurable space (X, Σ) and consider a sequence of nonnegative tensors $\{\mathcal{A}_x\}_{x \in X}$, we have following triangle inequality:

$$\left\| \int \mathcal{A}_x \mu(dx) \right\|_p \leq \int \|\mathcal{A}_x\|_p \mu(dx), \tag{26}$$

due to the convexity of p -norm for $p \geq 1$. Quasi-norms with $p \in (0, 1)$ are no longer convex. However, we demonstrate in Lemma 4 that these quasi-norms still satisfy an asymptotic convexity property for Kronecker products of tensors in the sense of allowing an extra term associated with the number of tensors involving the Kronecker product.

Lemma 4. *Let $p \in (0, 1)$, μ be a probability measure on a measurable space (X, Σ) , and consider a sequence of nonnegative tensors $\{\mathcal{A}_x\}_{x \in X}$ with $\mathcal{A}_x \in \mathbb{C}^{I_1 \times \dots \times I_M}$ having Canonical Polyadic (CP) decomposition, i.e., each \mathcal{A}_x can be expressed as $\mathcal{A}_x = \sum_{k_x} \lambda_{k_x} \mathbf{a}_{1,k_x} \otimes \mathbf{a}_{2,k_x} \otimes \dots \otimes \mathbf{a}_{M,k_x}$, where $\lambda_{k_x} \geq 0$, and $\mathbf{a}_{i,k_x} \in \mathbb{C}^{I_i}$ for $i \in [M]$. Then we have*

$$\frac{1}{m} \log \left\| \int_X \mathcal{A}_x^{\otimes m} \mu(dx) \right\|_p \leq \frac{1}{m} \log \left(\int_X \|\mathcal{A}_x^{\otimes m}\|_p \mu(dx) \right) + \frac{\log \text{poly}(m)}{m}. \tag{27}$$

Proof: Let \mathfrak{H} be the Hilbert space where the tensor \mathcal{A}_x acts on. For any $x \in X$, consider the CP decomposition $\mathcal{A}_x = \sum_{k_x} \lambda_{k_x} \mathbf{a}_{1,k_x} \otimes \mathbf{a}_{2,k_x} \otimes \dots \otimes \mathbf{a}_{M,k_x}$. By introducing an isometric space \mathfrak{H}' to \mathfrak{H} , we define the vector $\mathbf{v}_{i,k_x} \in \mathfrak{H} \otimes \mathfrak{H}'$ by $\mathbf{v}_{i,k_x} = \sum_{k_x} \lambda_{k_x}^{\frac{1}{M}} \mathbf{a}_{i,k_x} \otimes \mathbf{a}_{i,k_x}$ for $i \in [M]$, i.e., the purification of \mathcal{A}_x indicating that $\text{Tr}_{\mathfrak{H}'}(\sum_{k_x} \lambda_{k_x} \mathbf{v}_{1,k_x} \otimes \mathbf{v}_{2,k_x} \otimes \dots \otimes \mathbf{v}_{M,k_x}) = \mathcal{A}_x$ [12]. Note that the projectors $(\sum_{k_x} \lambda_{k_x} \mathbf{v}_{1,k_x} \otimes \mathbf{v}_{2,k_x} \otimes \dots \otimes \mathbf{v}_{M,k_x})^{\otimes m}$ lie in the symmetric subspace of $(\mathfrak{H} \otimes \mathfrak{H}')^{\otimes m}$ whose dimension grows with $\text{poly}(m)$ from Lemma 3. Then, we have

$$\int_X \mathcal{A}_x^{\otimes m} \mu(dx) = \int_X \text{Tr}_{\mathfrak{H}'^{\otimes m}} \left(\sum_{k_x} \lambda_{k_x} \mathbf{v}_{1,k_x} \otimes \mathbf{v}_{2,k_x} \otimes \dots \otimes \mathbf{v}_{M,k_x} \right)^{\otimes m} \mu(dx). \tag{28}$$

From Caratheodory theorem (see Theorem 18 in [7]), there exists a discrete probability measure $Pr(x)$, where $x \in X_d$ and $X_d \in X$ is the discrete set with the cardinality as $\text{poly}(m)$ such that

$$\int_X \mathcal{A}_x^{\otimes m} \mu(dx) = \sum_{x \in X_d} Pr(x) \mathcal{A}_x^{\otimes m}, \text{ and } \int_X \|\mathcal{A}_x^{\otimes m}\|_p \mu(dx) = \sum_{x \in X_d} Pr(x) \|\mathcal{A}_x^{\otimes m}\|_p.$$

Therefore, we can get

$$\frac{1}{m} \log \left\| \int_X \mathcal{A}_x^{\otimes m} \mu(dx) \right\|_p = \frac{1}{m} \log \left\| \sum_{x \in X_d} Pr(x) \mathcal{A}_x^{\otimes m} \right\|_p. \tag{29}$$

When $p \in (0, 1)$, the Schatten p -norm satisfies following triangle inequality for tensors (see [11])

$$\left\| \sum_{i=1}^n \mathcal{A}_i \right\|_p^p \leq \sum_{i=1}^n \|\mathcal{A}_i\|_p^p, \tag{30}$$

and from Eq. (29), we obtain

$$\begin{aligned} \frac{1}{m} \log \left\| \int_X \mathcal{A}_x^{\otimes m} \mu(dx) \right\|_p &\leq \frac{1}{m} \log \left(\sum_{x \in X_d} \|Pr(x) \mathcal{A}_x^{\otimes m}\|_p^p \right)^{\frac{1}{p}} \\ &= \frac{1}{m} \log \left(|X_d|^{\frac{1}{p}} \left(\frac{1}{|X_d|} \sum_{x \in X_d} \|Pr(x) \mathcal{A}_x^{\otimes m}\|_p^p \right)^{\frac{1}{p}} \right). \end{aligned} \tag{31}$$

Since the map $s \rightarrow s^{\frac{1}{p}}$ is convex for $p \in (0, 1)$, we have

$$\begin{aligned} \frac{1}{m} \log \left\| \int_X \mathcal{A}_x^{\otimes m} \mu(dx) \right\|_p &\leq \frac{1}{m} \log \left(|X_d|^{\frac{1}{p}-1} \sum_{x \in X_d} \|Pr(x) \mathcal{A}_x^{\otimes m}\|_p \right) \\ &= \frac{1}{m} \log \left(\sum_{x \in X_d} Pr(x) \|\mathcal{A}_x^{\otimes m}\|_p \right) + \frac{1-p}{mp} \log |X_d| \\ &= \frac{1}{m} \log \left(\int_X \|\mathcal{A}_x^{\otimes m}\|_p \mu(dx) \right) + \frac{\log \text{poly}(m)}{m}, \end{aligned} \tag{32}$$

where $|X_d| = \text{poly}(m)$ is applied at the last step. □

From Eq. (26) and from Lemma 4, we also have

$$\frac{1}{m} \log \left\| \int_X \mathcal{A}_x^{\otimes m} \mu(dx) \right\|_p \leq \log \sup_{x \in X} \|\mathcal{A}_x\|_p + \frac{\log \text{poly}(m)}{m}. \tag{33}$$

for all $p > 0$.

We need the following definition about a family of probability distribution to represent a pinching map with integration.

Definition 9. We define a family of probability distribution on \mathbb{R} , named as $\mu_\Delta(x)$, which satisfies following properties:

- $\tilde{\mu}_\Delta(0) = 1$, where $\tilde{\mu}_\Delta$ is the Fourier transform of the distribution function μ_Δ .
- $\tilde{\mu}_\Delta(\omega) = 0$ if and only if $|\omega| \geq \Delta$.

Following lemma will provide an integral representation of the pinching map.

Lemma 5. [Integral Representation of Pinching Map]

Let $\mathcal{H}, \mathcal{X} \in \mathbb{C}^{I_1 \times \dots \times I_M \times I_1 \times \dots \times I_M}$ be Hermitian tensors with same dimensions and $\mu_{\Delta_{\mathcal{H}}}$ is a probability measure with properties given in Definition 9. The term $\Delta_{\mathcal{H}}$ is defined as $\Delta_{\mathcal{H}} \stackrel{\text{def}}{=} \min\{|\lambda_j - \lambda_k| : \lambda_j \neq \lambda_k\}$ where λ_j, λ_k are two distinct eigenvalues in the spectral decomposition of the tensor \mathcal{H} given by Eq. (23), then we have following integral representation for a pinching map

$$\mathfrak{P}_{\mathcal{H}}(\mathcal{X}) = \int_{-\infty}^{\infty} e^{\iota s \mathcal{H}} \star_M \mathcal{X} \star_M e^{-\iota s \mathcal{H}} \mu_{\Delta_{\mathcal{H}}}(s) ds, \tag{34}$$

where ι is $\sqrt{-1}$.

Proof:

Because the spectral decomposition of tensor \mathcal{H} is

$$\mathcal{H} = \sum_{\lambda \in \text{sp}(\mathcal{H})} \lambda \mathcal{U}_\lambda, \tag{35}$$

where $\lambda \in \text{sp}(\mathcal{H}) \in \mathbb{R}$ and \mathcal{U}_λ are mutually orthogonal tensors. For any $s \in \mathbb{R}$, we then have

$$e^{\iota s \mathcal{H}} = \sum_{\lambda \in \text{sp}(\mathcal{H})} e^{\iota s \lambda} \mathcal{U}_\lambda, \tag{36}$$

and

$$e^{\iota s \mathcal{H}} \star_M \mathcal{X} \star_M e^{-\iota s \mathcal{H}} = \sum_{\lambda, \lambda' \in \text{sp}(\mathcal{H})} e^{\iota s(\lambda - \lambda')} \mathcal{U}_\lambda \star_M \mathcal{X} \star_M \mathcal{U}_{\lambda'}. \quad (37)$$

If we integrate both sides of Eq. (35) with respect to measure $\mu_{\Delta_{\mathcal{H}}}$, we obtain

$$\begin{aligned} & \int_{-\infty}^{\infty} e^{\iota s \mathcal{H}} \star_M \mathcal{X} \star_M e^{-\iota s \mathcal{H}} \mu_{\Delta_{\mathcal{H}}}(s) ds \\ &= \int_{-\infty}^{\infty} \left(\sum_{\lambda, \lambda' \in \text{sp}(\mathcal{H})} e^{\iota s(\lambda - \lambda')} \mathcal{U}_\lambda \star_M \mathcal{X} \star_M \mathcal{U}_{\lambda'} \right) \mu_{\Delta_{\mathcal{H}}}(s) ds \\ &= \sum_{\lambda, \lambda' \in \text{sp}(\mathcal{H})} \mathcal{U}_\lambda \star_M \mathcal{X} \star_M \mathcal{U}_{\lambda'} \tilde{\mu}_{\Delta_{\mathcal{H}}}(\lambda - \lambda'). \end{aligned} \quad (38)$$

By applying the properties in Definition 9 and the definition of the spectral gap $\Delta_{\mathcal{H}}$, we finally obtain

$$\int_{-\infty}^{\infty} e^{\iota s \mathcal{H}} \star_M \mathcal{X} \star_M e^{-\iota s \mathcal{H}} \mu_{\Delta_{\mathcal{H}}}(s) ds = \sum_{\lambda \in \text{sp}(\mathcal{H})} \mathcal{U}_\lambda \star_M \mathcal{X} \star_M \mathcal{U}_\lambda = \mathfrak{P}_{\mathcal{H}}(\mathcal{X}), \quad (39)$$

which asserts this Lemma. \square

Following lemmas are introduced for those nice properties about pinching maps.

Lemma 6. [commutativity of pinching map] Given a Hermitian tensor $\mathcal{H} \in \mathbb{C}^{I_1 \times \dots \times I_M \times I_1 \times \dots \times I_M}$ and any nonnegative tensors $\mathcal{X} \in \mathbb{C}^{I_1 \times \dots \times I_M \times I_1 \times \dots \times I_M}$, we have

$$\mathfrak{P}_{\mathcal{H}}(\mathcal{X}) \star_M \mathcal{H} = \mathcal{H} \star_M \mathfrak{P}_{\mathcal{H}}(\mathcal{X}). \quad (40)$$

Proof:

Because we have

$$\begin{aligned} \mathfrak{P}_{\mathcal{H}}(\mathcal{X}) \star_M \mathcal{H} &= \sum_{\lambda, \lambda' \in \text{sp}(\mathcal{H})} \mathcal{U}_\lambda \star_M \mathcal{X} \star_M \mathcal{U}_{\lambda'} \star_M \mathcal{H} = \sum_{\lambda \in \text{sp}(\mathcal{H})} \lambda \mathcal{U}_\lambda \star_M \mathcal{X} \star_M \mathcal{U}_\lambda \\ &= \sum_{\lambda, \lambda' \in \text{sp}(\mathcal{H})} \lambda' \mathcal{U}_{\lambda'} \star_M \mathcal{U}_\lambda \star_M \mathcal{X} \star_M \mathcal{U}_\lambda = \mathcal{H} \star_M \mathfrak{P}_{\mathcal{H}}(\mathcal{X}). \end{aligned} \quad (41)$$

\square

Lemma 7. [trace identity of pinching map] Given a Hermitian tensor $\mathcal{H} \in \mathbb{C}^{I_1 \times \dots \times I_M \times I_1 \times \dots \times I_M}$ and any nonnegative tensors $\mathcal{X} \in \mathbb{C}^{I_1 \times \dots \times I_M \times I_1 \times \dots \times I_M}$, we have

$$\text{Tr}(\mathfrak{P}_{\mathcal{H}}(\mathcal{X}) \star_M \mathcal{H}) = \text{Tr}(\mathcal{X} \star_M \mathcal{H}). \quad (42)$$

Proof:

From linearity and cyclic properties of the trace, Lemma 5 and the fact that the tensor $e^{-\iota s \mathcal{H}}$ commutes with the tensor \mathcal{H} for all $s \in \mathbb{R}$, then we have

$$\begin{aligned} \text{Tr}(\mathfrak{P}_{\mathcal{H}}(\mathcal{X}) \star_M \mathcal{H}) &= \int_{-\infty}^{\infty} \text{Tr}(e^{\iota s \mathcal{H}} \star_M \mathcal{X} \star_M e^{-\iota s \mathcal{H}} \star_M \mathcal{H}) \mu_{\Delta_{\mathcal{H}}}(s) ds \\ &= \int_{-\infty}^{\infty} \text{Tr}(\mathcal{X} \star_M \mathcal{H}) \mu_{\Delta_{\mathcal{H}}}(s) ds = \text{Tr}(\mathcal{X} \star_M \mathcal{H}). \end{aligned} \quad (43)$$

\square

Lemma 8. [Pinching Inequality] Let \geq_{L_o} be Loewner order for two positive semi-definite tensors, i.e., we say that tensors $\mathcal{A} \geq_{L_o} \mathcal{B}$ if $\mathcal{A} - \mathcal{B}$ is a positive semi-definite tensor. Given a Hermitian tensor $\mathcal{H} \in \mathbb{C}^{I_1 \times \dots \times I_M \times I_1 \times \dots \times I_M}$ and any nonnegative tensors $\mathcal{X} \in \mathbb{C}^{I_1 \times \dots \times I_M \times I_1 \times \dots \times I_M}$, we have

$$\mathfrak{P}_{\mathcal{H}}(\mathcal{X}) \geq_{L_o} \frac{\mathcal{X}}{|\text{sp}(\mathcal{H})|}, \quad (44)$$

where $|\text{sp}(\mathcal{H})|$ is the cardinality for the eigenvalues in the space $\text{sp}(\mathcal{H})$.

Proof:

We first define the tensor \mathcal{V}_k as following:

$$\mathcal{V}_k \stackrel{\text{def}}{=} \sum_{j=1}^{|\text{sp}(\mathcal{H})|} \exp\left(\frac{\iota 2\pi j k}{|\text{sp}(\mathcal{H})|}\right) \mathcal{U}_{\lambda_j}. \tag{45}$$

Then, the pinching map $\mathfrak{P}_{\mathcal{H}}(\mathcal{X})$ can be expressed as

$$\mathfrak{P}_{\mathcal{H}}(\mathcal{X}) = \sum_{\lambda \in \text{sp}(\mathcal{H})} \mathcal{U}_{\lambda} \mathcal{X} \mathcal{U}_{\lambda} =_1 \frac{1}{|\text{sp}(\mathcal{H})|} \sum_{k=1}^{|\text{sp}(\mathcal{H})|} \mathcal{V}_k \mathcal{X} \mathcal{V}_k^H \geq_{Lo} \frac{\mathcal{X}}{|\text{sp}(\mathcal{H})|}, \tag{46}$$

where we use following fact in the equality $=_1$:

$$\sum_{k=1}^{|\text{sp}(\mathcal{H})|} \exp\left(\frac{\iota 2k(j-j')\pi}{|\text{sp}(\mathcal{H})|}\right) = |\text{sp}(\mathcal{H})| \mathbf{1}(j=j'). \tag{47}$$

For the inequality \geq_{Lo} , we use following relations

$$\mathcal{V}_k \mathcal{X} \mathcal{V}_k^H \geq_{Lo} \mathcal{O}, \tag{48}$$

and

$$\mathcal{V}_{|\text{sp}(\mathcal{H})|} = \mathcal{I}, \tag{49}$$

where the zero tensor \mathcal{O} and the identity tensor \mathcal{I} both are with the same dimensions as \mathcal{X} (or \mathcal{H}). □

Theorem 1 (Golden-Thompson inequality for tensors). *Given two Hermitian tensors $\mathcal{H}_1 \in \mathbb{C}^{I_1 \times \dots \times I_M \times I_1 \times \dots \times I_M}$ and $\mathcal{H}_2 \in \mathbb{C}^{I_1 \times \dots \times I_M \times I_1 \times \dots \times I_M}$, we have*

$$\text{Tr}(\exp^{\mathcal{H}_1 + \mathcal{H}_2}) \leq \text{Tr}(\exp^{\mathcal{H}_1} \star_M \exp^{\mathcal{H}_2}). \tag{50}$$

Proof:

Let $\mathcal{A}_1 = \exp(\mathcal{H}_1)$ and $\mathcal{A}_2 = \exp(\mathcal{H}_2)$, we have

$$\begin{aligned} \log \text{Tr}(\exp(\log \mathcal{A}_1 + \log \mathcal{A}_2)) &= \frac{1}{m} \log \text{Tr}(\exp(\log \mathcal{A}_1^{\otimes m} + \log \mathcal{A}_2^{\otimes m})) \\ &\leq_2 \frac{1}{m} \log \text{Tr}\left(\exp\left(\log \mathfrak{P}_{\mathcal{A}_2^{\otimes m}}(\mathcal{A}_1^{\otimes m}) + \log \mathcal{A}_2^{\otimes m}\right)\right) \\ &\quad + \frac{\log \text{poly}(m)}{m} \\ &= \frac{1}{m} \log\left(\text{Tr}\left(\mathfrak{P}_{\mathcal{A}_2^{\otimes m}}(\mathcal{A}_1^{\otimes m}) \star_M \mathcal{A}_2^{\otimes m}\right)\right) + \frac{\log \text{poly}(m)}{m} \\ &= \log \text{Tr}(\mathcal{A}_1 \star_M \mathcal{A}_2) + \frac{\log \text{poly}(m)}{m}. \end{aligned} \tag{51}$$

The equality $=_1$ comes from Lemmas 1 and 2. The inequality \leq_2 follows from pinching inequality (Lemma 8), the monotone of \log and $\text{Tr} \exp(\cdot)$ functions, and the number of eigenvalues of $\mathcal{A}_2^{\otimes m}$ growing polynomially with m (Lemma 3). The equality $=_3$ utilizes the commutativity property for tensors $\mathfrak{P}_{\mathcal{A}_2^{\otimes m}}(\mathcal{A}_1^{\otimes m})$ and $\mathcal{A}_2^{\otimes m}$ based on Lemma 6. Finally, the equality $=_4$ applies trace properties from Lemmas 1, 2, and Lemma 7. If $m \rightarrow \infty$, the result of this theorem is established. □

Theorem 2 (Araki-Lieb-Thirring for tensors). *Given two positive semi-definite tensors $\mathcal{A}_1 \in \mathbb{C}^{I_1 \times \dots \times I_M \times I_1 \times \dots \times I_M}$ and $\mathcal{A}_2 \in \mathbb{C}^{I_1 \times \dots \times I_M \times I_1 \times \dots \times I_M}$, and $q > 0$, then*

$$\text{Tr}\left(\left(\mathcal{A}_1^{\frac{r}{2}} \mathcal{A}_2^r \mathcal{A}_1^{\frac{r}{2}}\right)^{\frac{q}{r}}\right) \leq \text{Tr}\left(\left(\mathcal{A}_1^{\frac{1}{2}} \mathcal{A}_2 \mathcal{A}_1^{\frac{1}{2}}\right)^q\right), \quad \text{if } r \in (0, 1], \tag{52}$$

with equality if and only if $\mathcal{A}_1 \star_M \mathcal{A}_2 = \mathcal{A}_2 \star_M \mathcal{A}_1$. This inequality holds in the opposite direction for $r \geq 1$.

Proof:

Since we have

$$\begin{aligned}
\log \operatorname{Tr} \left(\left(\mathcal{A}_1^{\frac{r}{2}} \mathcal{A}_2^r \mathcal{A}_1^{\frac{r}{2}} \right)^{\frac{q}{r}} \right) &= {}_1 \frac{1}{m} \log \operatorname{Tr} \left(\left((\mathcal{A}_1^{\frac{r}{2}})^{\otimes m} (\mathcal{A}_2^r)^{\otimes m} (\mathcal{A}_1^{\frac{r}{2}})^{\otimes m} \right)^{\frac{q}{r}} \right) \\
&\leq {}_2 \frac{1}{m} \log \operatorname{Tr} \left(\left((\mathcal{A}_1^{\frac{r}{2}})^{\otimes m} \mathfrak{P}_{\mathcal{A}_1^{\otimes m}} ((\mathcal{A}_2^r)^{\otimes m}) (\mathcal{A}_1^{\frac{r}{2}})^{\otimes m} \right)^{\frac{q}{r}} \right) + \frac{\log \operatorname{poly}(m)}{m} \\
&= {}_3 \frac{1}{m} \log \operatorname{Tr} \left(\left(\mathfrak{P}_{\mathcal{A}_1^{\otimes m}} \left((\mathcal{A}_1^{\frac{1}{2}})^{\otimes m} (\mathcal{A}_2)^{\otimes m} (\mathcal{A}_1^{\frac{1}{2}})^{\otimes m} \right) \right)^q \right) + \frac{\log \operatorname{poly}(m)}{m} \\
&\leq {}_4 \frac{1}{m} \log \operatorname{Tr} \left(\left((\mathcal{A}_1^{\frac{1}{2}})^{\otimes m} (\mathcal{A}_2)^{\otimes m} (\mathcal{A}_1^{\frac{1}{2}})^{\otimes m} \right)^q \right) + \frac{\log \operatorname{poly}(m)}{m} \\
&= {}_5 \log \operatorname{Tr} \left(\mathcal{A}_1^{\frac{1}{2}} \mathcal{A}_2 \mathcal{A}_1^{\frac{1}{2}} \right)^q + \frac{\log \operatorname{poly}(m)}{m}. \tag{53}
\end{aligned}$$

The equality $=_1$ comes from Lemmas 1 and 2. The inequality \leq_2 follows from pinching inequality (Lemma 8), the monotone of $\mathcal{X} \rightarrow \operatorname{Tr}(\mathcal{X}^\alpha)$ function for $\alpha \geq 0$, and the number of eigenvalues of $\mathcal{A}_1^{\otimes m}$ growing polynomially with m (Lemma 3). The equality $=_3$ utilizes the commutativity property for tensors $\mathfrak{P}_{\mathcal{A}_1^{\otimes m}}(\mathcal{A}_2^{\otimes m})$ and $\mathcal{A}_1^{\otimes m}$ based on Lemma 6. The inequality \leq_4 utilizes Lemma 4, integral representation of the pinching map (see Lemma 5) and the fact that p -norms are unitary invariant for $p \geq 0$. Finally, the equality $=_5$ applies trace properties from Lemmas 1 and 2. If $m \rightarrow \infty$, the result of this theorem is established.

For case $r \geq 1$, if we perform following replacements $\mathcal{A}_1^r \leftarrow \mathcal{A}_1$, $\mathcal{A}_2^r \leftarrow \mathcal{A}_2$, $\frac{q}{r} \leftarrow q$, and $\frac{1}{r} \leftarrow r$, the inequality in this theorem will be reversed. \square

3.2 Complex Interpolation Theory

In this section, we will mention those definitions and theorems about complex interpolation theory which will be used to prove multivariate tensor trace inequalities in Sec. 4. Complex interpolation theory enable us to control the behaviors of the complex function defined on the strip $S \stackrel{\text{def}}{=} \{z \in \mathbb{C} : 0 \leq \Re(z) \leq 1\}$ by its boundary values, $\Re(z) = 0$ and $\Re(z) = 1$. We define a family of probability measure on \mathbb{R} as

$$\rho_\theta(s) \stackrel{\text{def}}{=} \frac{\sin(\pi\theta)}{2\theta(\cosh(\pi s) + \cos(\pi\theta))} \quad \text{for } \theta \in (0, 1). \tag{54}$$

Moreover, we have following limiting behaviors for ρ_θ :

$$\rho_0(s) \stackrel{\text{def}}{=} \lim_{\theta \searrow 0} \rho_\theta(s) = \frac{\pi}{2} (\cosh(\pi s) + 1)^{-1}, \tag{55}$$

and

$$\rho_1(s) \stackrel{\text{def}}{=} \lim_{\theta \nearrow 1} \rho_\theta(s) = \delta(s), \tag{56}$$

where δ is the Dirac δ -distribution.

We will introduce Stein-Hirschman theorem [10, 25] about complex interpolation theory.

Theorem 3. Let $p_0, p_1 \in [1, \infty]$, $\theta \in (0, 1)$, $\rho_\theta(s)$ defined by Eq. (54), define p_θ by $\frac{1}{p_\theta} = \frac{1-\theta}{p_0} + \frac{\theta}{p_1}$, and $S \stackrel{\text{def}}{=} \{z \in \mathbb{C} : 0 \leq \Re(z) \leq 1\}$. Let F be a map from S to bounded linear operators on a separable Hilbert space that is holomorphic in the interior of S and continuous on the boundary. If $z \rightarrow \|F(z)\|_{p, \Re(z)}$ is uniformly bounded on S , we have

$$\log \|F(\theta)\|_{p, (\theta)} \leq \int_{-\infty}^{\infty} \left(\rho_{1-\theta}(s) \log \|F(\iota s)\|_{p_0}^{1-\theta} + \rho_\theta(s) \log \|F(1 + \iota s)\|_{p_1}^\theta \right) ds \tag{57}$$

4 Multivariate Tensor Trace Inequalities

In order to extend Theorems 1 and 2 involving two tensors to multiple tensors, we require the following lemma about *Lie product formula* for tensors.

Lemma 9. Let $m \in \mathbb{N}$ and $(\mathcal{L}_k)_{k=1}^m$ be a finite sequence of bounded tensors with dimensions $\mathcal{L}_k \in \mathbb{C}^{I_1 \times \dots \times I_M \times I_1 \times \dots \times I_M}$, then we have

$$\lim_{n \rightarrow \infty} \left(\prod_{k=1}^m \exp\left(\frac{\mathcal{L}_k}{n}\right) \right)^n = \exp\left(\sum_{k=1}^m \mathcal{L}_k\right) \tag{58}$$

Proof:

We will prove the case for $m = 2$, and the general value of m can be obtained by mathematical induction. Let $\mathcal{L}_1, \mathcal{L}_2$ be bounded tensors act on some Hilbert space. Define $\mathcal{C} \stackrel{\text{def}}{=} \exp((\mathcal{L}_1 + \mathcal{L}_2)/n)$, and $\mathcal{D} \stackrel{\text{def}}{=} \exp(\mathcal{L}_1/n) \star_M \exp(\mathcal{L}_2/n)$. Note we have following estimates for the norm of tensors \mathcal{C}, \mathcal{D} :

$$\|\mathcal{C}\|, \|\mathcal{D}\| \leq \exp\left(\frac{\|\mathcal{L}_1\| + \|\mathcal{L}_2\|}{n}\right) = [\exp(\|\mathcal{L}_1\| + \|\mathcal{L}_2\|)]^{1/n}. \tag{59}$$

From the Cauchy-Product formula, the tensor \mathcal{D} can be expressed as:

$$\begin{aligned} \mathcal{D} &= \exp(\mathcal{L}_1/n) \star_M \exp(\mathcal{L}_2/n) = \sum_{i=0}^{\infty} \frac{(\mathcal{L}_1/n)^i}{i!} \star_M \sum_{j=0}^{\infty} \frac{(\mathcal{L}_2/n)^j}{j!} \\ &= \sum_{l=0}^{\infty} n^{-l} \sum_{i=0}^l \frac{\mathcal{L}_1^i}{i!} \star_M \frac{\mathcal{L}_2^{l-i}}{(l-i)!}, \end{aligned} \tag{60}$$

then we can bound the norm of $\mathcal{C} - \mathcal{D}$ as

$$\begin{aligned} \|\mathcal{C} - \mathcal{D}\| &= \left\| \sum_{i=0}^{\infty} \frac{([\mathcal{L}_1 + \mathcal{L}_2]/n)^i}{i!} - \sum_{l=0}^{\infty} n^{-l} \sum_{i=0}^l \frac{\mathcal{L}_1^i}{i!} \star_M \frac{\mathcal{L}_2^{l-i}}{(l-i)!} \right\| \\ &= \left\| \sum_{i=2}^{\infty} k^{-i} \frac{([\mathcal{L}_1 + \mathcal{L}_2])^i}{i!} - \sum_{m=l}^{\infty} n^{-l} \sum_{i=0}^l \frac{\mathcal{L}_1^i}{i!} \star_M \frac{\mathcal{L}_2^{l-i}}{(l-i)!} \right\| \\ &\leq \frac{1}{k^2} \left[\exp(\|\mathcal{L}_1\| + \|\mathcal{L}_2\|) + \sum_{l=2}^{\infty} n^{-l} \sum_{i=0}^l \frac{\|\mathcal{L}_1\|^i}{i!} \cdot \frac{\|\mathcal{L}_2\|^{l-i}}{(l-i)!} \right] \\ &= \frac{1}{n^2} \left[\exp(\|\mathcal{L}_1\| + \|\mathcal{L}_2\|) + \sum_{l=2}^{\infty} n^{-l} \frac{(\|\mathcal{L}_1\| + \|\mathcal{L}_2\|)^l}{l!} \right] \\ &\leq \frac{2 \exp(\|\mathcal{L}_1\| + \|\mathcal{L}_2\|)}{n^2}. \end{aligned} \tag{61}$$

For the difference between the higher power of \mathcal{C} and \mathcal{D} , we can bound them as

$$\begin{aligned} \|\mathcal{C}^n - \mathcal{D}^n\| &= \left\| \sum_{l=0}^{n-1} \mathcal{C}^l (\mathcal{C} - \mathcal{D}) \mathcal{C}^{n-l-1} \right\| \\ &\leq_1 \exp(\|\mathcal{L}_1\| + \|\mathcal{L}_2\|) \cdot n \cdot \|\mathcal{L}_1 - \mathcal{L}_2\|, \end{aligned} \tag{62}$$

where the inequality \leq_1 uses the following fact

$$\|\mathcal{C}\|^l \|\mathcal{D}\|^{n-l-1} \leq \exp(\|\mathcal{L}_1\| + \|\mathcal{L}_2\|)^{\frac{n-1}{n}} \leq \exp(\|\mathcal{L}_1\| + \|\mathcal{L}_2\|), \tag{63}$$

based on Eq. (59). By combining with Eq. (61), we have the following bound

$$\|\mathcal{C}^n - \mathcal{D}^n\| \leq \frac{2 \exp(2\|\mathcal{L}_1\| + 2\|\mathcal{L}_2\|)}{n}. \tag{64}$$

Then this lemma is proved when n goes to infinity. □

4.1 Multivariate Araki-Lieb-Thirring Inequality

In this section, we will provide a theorem for multivariate Araki-Lieb-Thirring (ALT) inequality for tensors.

Theorem 4. Let $p \geq 1, \theta \in (0, 1]$, probability distribution ρ_θ defined by (54), $n \in \mathbb{N}$, and consider a finite sequence $(\mathcal{A})_{k=1}^n$ of positive semi-definite tensors. Then, we have

$$\log \left\| \left\| \prod_{k=1}^n \mathcal{A}_k^\theta \right\| \right\|_p^{\frac{1}{\theta}} \leq \int_{-\infty}^{\infty} \log \left\| \prod_{k=1}^n \mathcal{A}_k^{1+\iota s} \right\|_p \rho_\theta(s) ds. \tag{65}$$

Proof:

For $\theta = 1$, the both sides of Eq. (65) are equal to $\log \|\prod_{k=1}^n \mathcal{A}_k\|_p$. We will prove the cases for $\theta \in (0, 1)$. We prove the result for strictly positive definite tensors and note that the generalization to positive semi-definite tensors follows by continuity. We define the function $F(z) \stackrel{\text{def}}{=} \prod_{k=1}^n \mathcal{A}_k^z = \prod_{k=1}^n \exp(z \log \mathcal{A}_k)$ which satisfies the conditions of Theorem 3. By selecting $p_0 = \infty, p_1 = p$, and $p_\theta = \frac{p}{\theta}$ in Theorem 3, one can obtain

$$\log \|F(1 + \iota s)\|_{p_1}^\theta = \theta \log \left\| \prod_{k=1}^n \mathcal{A}_k^{1+\iota s} \right\|_p, \tag{66}$$

and

$$\log \|F(\iota s)\|_{p_0}^{1-\theta} = (1 - \theta) \log \left\| \prod_{k=1}^n \mathcal{A}_k^{\iota s} \right\|_\infty = 0, \tag{67}$$

since tensors $\mathcal{A}_k^{\iota s}$ are unitary. We also have

$$\log \|F(\theta)\|_{p_\theta} = \log \left\| \prod_{k=1}^n \mathcal{A}_k^\theta \right\|_{\frac{p}{\theta}} = \theta \log \left\| \prod_{k=1}^n \mathcal{A}_k^\theta \right\|_p^{\frac{1}{\theta}}, \tag{68}$$

and this theorem is proved by putting Eqs. (66), (67), (68) into Eq. (57). □

4.2 Multivariate Golden-Thompson Inequality

In this section, we will provide a theorem for multivariate Golden-Thompson (GT) inequality for tensors.

Theorem 5. *Let $p \geq 1$, probability distribution ρ_0 defined by (55), $n \in \mathbb{N}$, and consider a finite sequence $(\mathcal{H})_{k=1}^n$ of Hermitian tensors. Then, we have*

$$\log \left\| \exp \left(\sum_{k=1}^n \mathcal{H}_k \right) \right\|_p \leq \int_{-\infty}^{\infty} \log \left\| \prod_{k=1}^n \exp((1 + \iota s)\mathcal{H}_k) \right\|_p \rho_0(s) ds. \tag{69}$$

Proof: From Theorem 9 and Lie product formula given by Lemma 9, this theorem is proved by taking $\theta \rightarrow 0$ in Eq. (65). □

The multivariate Golden-Thompson inequality provided by Theorem 5 is only true for Hermitian tensors. The following theorem generalizes Theorem 5 to general tensors.

Theorem 6. *Let $p \geq 1$, probability distribution ρ_0 defined by (55), $n \in \mathbb{N}$, and consider a finite sequence $(\mathcal{A})_{k=1}^n$ of tensors. Then, we have*

$$\log \left\| \exp \left(\sum_{k=1}^n \mathcal{A}_k \right) \right\|_p \leq \int_{-\infty}^{\infty} \log \left\| \prod_{k=1}^n \exp((1 + \iota s)\Re(\mathcal{A}_k)) \right\|_p \rho_0(s) ds, \tag{70}$$

where $\Re(\mathcal{A}_k)$ is the real part of the tensor \mathcal{A}_k defined as $\Re(\mathcal{A}_k) \stackrel{\text{def}}{=} \frac{1}{2}(\mathcal{A}_k + \mathcal{A}_k^H)$.

Proof: We also define the imaginary part of the tensor \mathcal{A}_k as $\Im(\mathcal{A}_k) \stackrel{\text{def}}{=} \frac{1}{2i}(\mathcal{A}_k - \mathcal{A}_k^H)$ and note that the both $\Re(\mathcal{A}_k)$ and $\Im(\mathcal{A}_k)$ are Hermitian tensors. We define the function $F(z) \stackrel{\text{def}}{=} \prod_{k=1}^n \exp(z\Re(\mathcal{A}_k) + \iota\theta\Im(\mathcal{A}_k))$ which satisfies the conditions of Theorem 3. By selecting $p_0 = \infty, p_1 = p$, and $p_\theta = \frac{p}{\theta}$ in Theorem 3, one can obtain

$$\begin{aligned} \log \left\| \exp \left(\theta \sum_{k=1}^n \mathcal{A}_k \right) \right\|_p^{\frac{1}{\theta}} &= \log \|F(\theta)\|_{p_\theta} \\ &\leq \int_{-\infty}^{\infty} \log \|F(1 + \iota s)\|_{p_\theta}^\theta \rho_\theta(s) ds \\ &= \theta \int_{-\infty}^{\infty} \log \left\| \prod_{k=1}^n \exp((1 + \iota s)\Re(\mathcal{A}_k) + \iota\theta\Im(\mathcal{A}_k)) \right\|_p \rho_\theta(s) ds \end{aligned} \tag{71}$$

where we used that $\log \|F(\iota s)\|_\infty = 0$ since $F(\iota s)$ is unitary in the inequality step. By dividing θ at both sides of Eq. (71) and taking $\theta \rightarrow 0$, the theorem is proved by applying Lie product formula given by Lemma 9 again. □

4.3 Multivariate Logarithmic Trace Inequality for Tensors

In this section, we will apply Theorem 5 to prove multivariate logarithmic trace inequality. We have to define relative entropy between two tensors first.

Definition 10. Given two positive definite tensors $\mathcal{A} \in \mathbb{C}^{I_1 \times \dots \times I_M \times I_1 \times \dots \times I_M}$ and tensor $\mathcal{B} \in \mathbb{C}^{I_1 \times \dots \times I_M \times I_1 \times \dots \times I_M}$, where the tensor \mathcal{A} has the trace equal to one. The relative entropy between tensors \mathcal{A} and \mathcal{B} is defined as

$$D(\mathcal{A} \parallel \mathcal{B}) \stackrel{\text{def}}{=} \text{Tr} \mathcal{A} \star_M (\log \mathcal{A} - \log \mathcal{B}). \tag{72}$$

Based on this relative entropy definition, we have the following lemma about variational expression of relative entropy.

Lemma 10. Given two positive definite tensors $\mathcal{A} \in \mathbb{C}^{I_1 \times \dots \times I_M \times I_1 \times \dots \times I_M}$ and tensor $\mathcal{B} \in \mathbb{C}^{I_1 \times \dots \times I_M \times I_1 \times \dots \times I_M}$, where the tensor \mathcal{A} has the trace equal to one. Then, we have

$$D(\mathcal{A} \parallel \mathcal{B}) = \sup_{\mathcal{X}} (\text{Tr} (\mathcal{A} \star_M \log \mathcal{X}) - \log \text{Tr} (\exp(\log \mathcal{B} + \log \mathcal{X}))), \tag{73}$$

and

$$D(\mathcal{A} \parallel \mathcal{B}) = \sup_{\mathcal{X}} (\text{Tr} (\mathcal{A} \star_M \log \mathcal{X}) + 1 - \text{Tr} (\exp(\log \mathcal{B} + \log \mathcal{X}))), \tag{74}$$

where $\mathcal{X} \in \mathbb{C}^{I_1 \times \dots \times I_M \times I_1 \times \dots \times I_M}$ is a positive definite tensor.

Proof:

For any Hermitian \mathcal{H} tensor with dimensions $\mathcal{H} \in \mathbb{C}^{I_1 \times \dots \times I_M \times I_1 \times \dots \times I_M}$, we first show that

$$\log \text{Tr} (e^{\mathcal{H} + \log \mathcal{B}}) = \sup_{\mathcal{A}} (\text{Tr} (\mathcal{A} \mathcal{H}) - D(\mathcal{A} \parallel \mathcal{B})). \tag{75}$$

We define a function with tensor argument as $g(\mathcal{A}) = \text{Tr} (\mathcal{A} \mathcal{H}) - D(\mathcal{A} \parallel \mathcal{B})$, then let $\mathcal{A} = \sum_{\lambda \in \text{sp}(\mathcal{A})} \lambda \mathcal{U}_\lambda$ denote the spectrum decomposition of \mathcal{A} . Because the trace of \mathcal{A} is one, we have

$$g \left(\sum_{\lambda \in \text{sp}(\mathcal{A})} \lambda \mathcal{U}_\lambda \right) = \sum_{\lambda \in \text{sp}(\mathcal{A})} (\lambda \text{Tr} \mathcal{U}_\lambda \mathcal{H} + \lambda \text{Tr} \mathcal{U}_\lambda \log \mathcal{B} - \lambda \log \lambda). \tag{76}$$

By taking derivative with respect to λ for Eq. (76), we have

$$\left. \frac{\partial}{\partial \lambda} g \left(\sum_{\lambda \in \text{sp}(\mathcal{A})} \lambda \mathcal{U}_\lambda \right) \right|_{\lambda=0} = \infty, \tag{77}$$

this shows that the minimizer for Eq. (75) is a strictly positive tensors $\tilde{\mathcal{A}}$ with $\text{Tr} \tilde{\mathcal{A}} = 1$. For any Hermitian tensor \mathcal{Y} with $\text{Tr} \tilde{\mathcal{Y}} = 0$, we have

$$0 = \left. \frac{dg(\tilde{\mathcal{A}} + t\mathcal{Y})}{dt} \right|_{t=0} = \text{Tr} [\mathcal{Y} (\mathcal{H} + \log \mathcal{B} - \log \tilde{\mathcal{A}})]. \tag{78}$$

This indicates that $\mathcal{H} + \log \mathcal{B} - \log \tilde{\mathcal{A}}$ is proportional to the identity tensor. Then, we will have

$$\tilde{\mathcal{A}} = \frac{e^{\mathcal{H} + \log \mathcal{B}}}{\text{Tr} e^{\mathcal{H} + \log \mathcal{B}}} \text{ and } g(\tilde{\mathcal{A}}) = \log \text{Tr} e^{\mathcal{H} + \log \mathcal{B}}, \tag{79}$$

which proves Eq. (75).

We first prove Eq. (73) based on Eq. (75). Because Eq. (75) implies that the functional $\mathcal{H} \rightarrow \log \text{Tr} e^{\mathcal{H} + \log \mathcal{B}}$ is convex, then let $\tilde{\mathcal{H}} = \log \mathcal{A} - \log \mathcal{B}$, we can have a concave function

$$f(\mathcal{H}) \stackrel{\text{def}}{=} \text{Tr} \mathcal{A} \mathcal{H} - \log \text{Tr} e^{\mathcal{H} + \log \mathcal{B}}. \tag{80}$$

For any Hermitian tensor \mathcal{Y} , we have

$$\left. \frac{df(\tilde{\mathcal{H}} + t\mathcal{Y})}{dt} \right|_{t=0} = 0, \tag{81}$$

because $\text{Tr}\mathcal{A} = 1$ and $\frac{d\text{Tr}e^{\log \mathcal{A} + t\mathcal{Y}}}{dt}\Big|_{t=0} = \text{Tr}\mathcal{A}\mathcal{Y}$. Therefore, the tensor $\tilde{\mathcal{H}}$ is the maximizer of function f and

$$f(\tilde{\mathcal{H}}) = \text{Tr}\mathcal{A}(\log \mathcal{A} - \log \mathcal{B}) = D(\mathcal{A} \parallel \mathcal{B}). \tag{82}$$

Since for any any Hermitian tensor \mathcal{H} can be expressed as $\mathcal{H} = \log \mathcal{X}$ for some positive semi-definite tensor, we proved Eq. (73).

Now, we are ready to prove Eq. (74). From $\log x \leq x - 1$ for $x \geq 0$, we have $\log \text{Tr}e^{\log \mathcal{B} + \log \mathcal{X}} \leq \text{Tr}e^{\log \mathcal{B} + \log \mathcal{X}} - 1$. Hence, we have

$$\sup_{\mathcal{X}} (\text{Tr}\mathcal{A} \log \mathcal{X} - \log \text{Tr}e^{\log \mathcal{B} + \log \mathcal{X}}) \geq \sup_{\mathcal{X}} (\text{Tr}\mathcal{A} \log \mathcal{X} + 1 - \text{Tr}e^{\log \mathcal{B} + \log \mathcal{X}}). \tag{83}$$

Because $\text{Tr}\mathcal{A} \log \mathcal{X} - \log \text{Tr}e^{\log \mathcal{B} + \log \mathcal{X}}$ is invariant under the scaling transform from \mathcal{X} to $\gamma\mathcal{X}$ for $\gamma \in \mathbb{R}_+$, we can assume that $\text{Tr}e^{\log \mathcal{B} + \log \mathcal{X}} = 1$. Then, we have

$$\begin{aligned} \sup_{\mathcal{X}} (\text{Tr}\mathcal{A} \log \mathcal{X} - \log \text{Tr}e^{\log \mathcal{B} + \log \mathcal{X}}) &= \sup_{\mathcal{X}} (\text{Tr}\mathcal{A} \log \mathcal{X} - \log \text{Tr}e^{\log \mathcal{B} + \log \mathcal{X}} : \\ &\quad \text{Tr}e^{\log \mathcal{B} + \log \mathcal{X}} = 1) \\ &\leq \sup_{\mathcal{X}} (\text{Tr}\mathcal{A} \log \mathcal{X} - 1 + \text{Tr}e^{\log \mathcal{B} + \log \mathcal{X}}). \end{aligned} \tag{84}$$

From both Eqs (83) and (84), we prove Eq. (74). □

We are ready to present multivariate logarithmic trace inequality for tensors by the following theorem.

Theorem 7. *Let $0 < q \leq 1$, probability distribution ρ_0 defined by (55), $n \in \mathbb{N}$, and consider a finite sequence $(\mathcal{A})_{k=1}^n$ of positive semi-definite tensors. Then, we have*

$$\begin{aligned} &\sum_{k=1}^n \text{Tr}\mathcal{A}_1 \log \mathcal{A}_k \geq \\ &\frac{1}{q} \int_{-\infty}^{\infty} \left(\text{Tr}\mathcal{A}_1 \log \mathcal{A}_n^{\frac{q(1+\iota s)}{2}} \cdots \mathcal{A}_3^{\frac{q(1+\iota s)}{2}} \mathcal{A}_2^{\frac{q}{2}} \mathcal{A}_1^q \mathcal{A}_2^{\frac{q}{2}} \mathcal{A}_3^{\frac{q(1-\iota s)}{2}} \cdots \mathcal{A}_n^{\frac{q(1-\iota s)}{2}} \right) \rho_0(s) ds, \end{aligned} \tag{85}$$

which the equality will be valid when $q \rightarrow 0$.

Proof: Because the inequality given by Eq. (85) is invariant under multiplication of the tensors $\mathcal{A}_1, \mathcal{A}_2, \dots, \mathcal{A}_n$ with positive numbers a_1, a_2, \dots, a_n , we can add constraints on the norms of tensors without loss of generality. We assume that the $\text{Tr}\mathcal{A}_1 = 1$.

From the relative entropy in Definition 10, we have

$$\begin{aligned} \sum_{k=1}^n \text{Tr}\mathcal{A}_1 \log \mathcal{A}_k &= D(\mathcal{A}_1 \parallel \exp(\sum_{k=2}^n \log \mathcal{A}_k^{-1})) \\ &= \sup_{\mathcal{X}} \left(\text{Tr}\mathcal{A}_1 \log \mathcal{X} + 1 - \text{Tr} \exp \left(\log \mathcal{X} - \sum_{k=2}^n \log \mathcal{A}_k \right) \right), \end{aligned} \tag{86}$$

where we apply Lemma 10. From Theorem 6 and set $\mathcal{H}_k = \log \mathcal{A}_k^q$, we have

$$\begin{aligned} &\text{Tr} \exp \left(\sum_{k=1}^n \log \mathcal{A}_k \right) \leq \\ &\int_{-\infty}^{\infty} \text{Tr} \left(\mathcal{A}_n^{\frac{q(1+\iota s)}{2}} \cdots \mathcal{A}_3^{\frac{q(1+\iota s)}{2}} \mathcal{A}_2^{\frac{q(1+\iota s)}{2}} \mathcal{A}_1^q \mathcal{A}_2^{\frac{q(1-\iota s)}{2}} \mathcal{A}_3^{\frac{q(1-\iota s)}{2}} \cdots \mathcal{A}_n^{\frac{q(1-\iota s)}{2}} \right)^{\frac{1}{q}} \rho_0(s) ds \end{aligned} \tag{87}$$

using the concavity of the logarithm and Jensen's inequality. Applying Eq. (87) to Eq. (86), we get

$$\begin{aligned} \sum_{k=1}^n \text{Tr}\mathcal{A}_1 \log \mathcal{A}_k &\geq \sup_{\mathcal{X}} \left(\int_{-\infty}^{\infty} (\text{Tr}\mathcal{A}_1 \log \mathcal{X}) \rho_0(s) ds + 1 \right. \\ &\quad \left. - \text{Tr} \left(\mathcal{A}_2^{\frac{-q}{2}} \mathcal{A}_3^{\frac{-q(1+\iota s)}{2}} \cdots \mathcal{A}_n^{\frac{-q(1+\iota s)}{2}} \mathcal{X}^q \mathcal{A}_n^{\frac{-q(1-\iota s)}{2}} \cdots \mathcal{A}_3^{\frac{-q(1-\iota s)}{2}} \mathcal{A}_2^{\frac{-q}{2}} \right)^{\frac{1}{q}} \right). \end{aligned} \tag{88}$$

If we set the tensor \mathcal{X} as

$$\mathcal{X} \stackrel{\text{def}}{=} \left(\mathcal{A}_n^{\frac{q(1+\iota s)}{2}} \cdots \mathcal{A}_3^{\frac{q(1+\iota s)}{2}} \mathcal{A}_2^{\frac{q}{2}} \mathcal{A}_1^q \mathcal{A}_2^{\frac{q}{2}} \mathcal{A}_3^{\frac{q(1-\iota s)}{2}} \cdots \mathcal{A}_n^{\frac{q(1-\iota s)}{2}} \right)^{\frac{1}{q}}, \tag{89}$$

the tensor \mathcal{X} becomes a positive semi-definite tensor. Substituting Eq. (89) into Eq. (88), this theorem is proved for $0 < q \leq 1$.

For $q \rightarrow 0$, we wish to prove the equality at Eq. (85). Because $\log \mathcal{X} \geq_{L_o} \mathcal{I} - \mathcal{X}^{-1}$ for any positive tensor \mathcal{X} , we have

$$\begin{aligned} & \text{Tr} \mathcal{A}_1 \log \mathcal{A}_n^{\frac{q(1+\iota s)}{2}} \cdots \mathcal{A}_3^{\frac{q(1+\iota s)}{2}} \mathcal{A}_2^{\frac{q}{2}} \mathcal{A}_1^q \mathcal{A}_2^{\frac{q}{2}} \mathcal{A}_3^{\frac{q(1-\iota s)}{2}} \cdots \mathcal{A}_n^{\frac{q(1-\iota s)}{2}} \geq \\ & \text{Tr} \mathcal{A}_1 \left(\mathcal{I} - \mathcal{A}_n^{-\frac{q(1-\iota s)}{2}} \cdots \mathcal{A}_3^{-\frac{q(1-\iota s)}{2}} \mathcal{A}_2^{-\frac{q}{2}} \mathcal{A}_1^{-q} \mathcal{A}_2^{-\frac{q}{2}} \mathcal{A}_3^{-\frac{q(1+\iota s)}{2}} \cdots \mathcal{A}_n^{-\frac{q(1+\iota s)}{2}} \right) \\ & \stackrel{\text{def}}{=} h_q(s), \end{aligned} \tag{90}$$

and we can assume that $h_q(s) \geq 0$ since we can scale each tensor \mathcal{A}_k by a positive number for $k \in [n]$. By Fatou’s lemma, we have

$$\liminf_{q \rightarrow 0} \int_{-\infty}^{\infty} \frac{h_q(s)}{q} \rho_0(s) ds \geq \int_{-\infty}^{\infty} \liminf_{q \rightarrow 0} \frac{h_q(s)}{q} \rho_0(s) ds. \tag{91}$$

We also have $h_0(s) = 0$ and

$$\liminf_{q \rightarrow 0} \frac{h_q(s)}{q} = \sum_{k=1}^n \text{Tr} \mathcal{A}_1 \log \mathcal{A}_k. \tag{92}$$

By Eqs (90) and (91), we have the equality at Eq (85) as $q \rightarrow 0$. □

5 Applications: Random Tensors

This section will apply multivariate Golden-Thompson inequality from Theorem 5 to form the tail bound for independent sum of random tensors.

Consider a random Hermitian tensor $\mathcal{X} \in \mathbb{C}^{I_1 \times \cdots \times I_M \times I_1 \times \cdots \times I_M}$, i.e., each entry in this tensor is an independent random variable with $x_{i_1, \dots, i_M, j_1, \dots, j_M} = \overline{x_{j_1, \dots, j_M, i_1, \dots, i_M}}$. We assume that the random tensor \mathcal{X} has moments of all order n . We can construct tensor extensions of the moment generating function (MGF), and the cumulant generating function (CGF):

$$M(t) \stackrel{\text{def}}{=} \mathbb{E} e^{t\mathcal{X}}, \text{ and } C(t) \stackrel{\text{def}}{=} \log \mathbb{E} e^{t\mathcal{X}}, \tag{93}$$

where $t \in \mathbb{R}$. The tensor MGF and CGF can be expressed as power series expansions:

$$M(t) = \mathcal{I} + \sum_{n=1}^{\infty} \frac{t^n \mathbb{E}(\mathcal{X}^n)}{n!}, \text{ and } C(t) = \sum_{n=1}^{\infty} \frac{t^n \Phi_n}{n!}, \tag{94}$$

where the coefficients $\mathbb{E}(\mathcal{X}^n)$ are called *tensor moments*, and Φ_n are named as *tensor cumulants*. The tensor cumulant Φ_n has a formal expression as a noncommutative polynomial in the tensor moments up to order n . For example, the first cumulant is the mean and the second cumulant is the variance:

$$\Phi_1 = \mathbb{E}\mathcal{X}, \text{ and } \Phi_2 = \mathbb{E}(\mathcal{X}^2) - (\mathbb{E}\mathcal{X})^2. \tag{95}$$

5.1 Laplace Transform Method for Tensors

We will apply Laplace transform bound to bound the maximum eigenvalue of a random Hermitian tensor by following lemma. Lemma ?? help us to control tail probabilities for the maximum eigenvalue of a random tensor by producing a bound for the trace of the tensor MGF defined in Eq. (93).

Lemma 11. *Let $\mathcal{Y} \in \mathbb{C}^{I_1 \times \cdots \times I_M \times I_1 \times \cdots \times I_M}$ be a random Hermitian tensor and assume that $|I_i| = N$ for $1 \leq i \leq M$. For $\zeta \in \mathbb{R}$, we have*

$$\mathbb{P}(\lambda_{\max}(\mathcal{Y}) \geq \zeta) \leq (2M - 1)^N \inf_{t > 0} (e^{-\zeta t} \mathbb{E} \text{Tr} e^{t\mathcal{Y}}) \tag{96}$$

Proof:

Given a fix value t , we have

$$\mathbb{P}(\lambda_{\max}(\mathcal{Y}) \geq \zeta) = \mathbb{P}(\lambda_{\max}(t\mathcal{Y}) \geq t\zeta) = \mathbb{P}(e^{\lambda_{\max}(t\mathcal{Y})} \geq e^{t\zeta}) \leq e^{-t\zeta} \mathbb{E} e^{\lambda_{\max}(t\mathcal{Y})}. \tag{97}$$

The first equality uses the homogeneity of the maximum eigenvalue map, the second equality comes from the monotonicity of the scalar exponential function, and the last relation is Markov’s inequality. Because we have

$$e^{\lambda_{\max}(t\mathcal{Y})} = \lambda_{\max}(e^{t\mathcal{Y}}) \leq (2M - 1)^N \text{Tr} e^{t\mathcal{Y}}, \tag{98}$$

where the first equality used the spectral mapping theorem, and the inequality holds because the exponential of an Hermitian tensor is positive definite and the maximum eigenvalue of a positive definite tensor is dominated by the trace [21]. From Eqs (97) and (98), this lemma is established. □

5.2 Tail Bounds for Independent Sum

This section contains abstract tail bounds for the sum of independent random tensors. This general inequality can serve as the progenitor of other random tensors majorization inequality.

Theorem 8. Consider n independent random Hermitian tensors $\mathcal{X}_k \in \mathbb{C}^{I_1 \times \dots \times I_M \times I_1 \times \dots \times I_M}$ for $k \in [n]$, for all $\zeta \in \mathbb{R}$, we have

$$\mathbb{P}(\lambda_{\max} \left(\sum_{k=1}^n \mathcal{X}_k \right) \geq \zeta) \leq (2M-1)^N \inf_{t>0} \left\{ e^{-t\zeta} \int_{-\infty}^{\infty} \text{Tr} \left[\mathbb{E}(e^{t\mathcal{X}_1}) \star_M \left(\prod_{k=2}^{n-1} \mathbb{E}(e^{\frac{(1+\iota s)t\mathcal{X}_k}{2}}) \right) \star_M \mathbb{E}(e^{t\mathcal{X}_n}) \star_M \left(\prod_{k=n-1}^2 \mathbb{E}(e^{\frac{(1-\iota s)t\mathcal{X}_k}{2}}) \right) \right] \rho_0(s) ds \right\} \quad (99)$$

Proof:

By setting $\mathcal{H}_k = t\mathcal{X}_k$, $p = 2$ in Theorem 5, we will have

$$\text{Tr} e^{\left(\sum_{k=1}^n t\mathcal{X}_k \right)} \leq \int_{-\infty}^{\infty} \text{Tr} \left[e^{t\mathcal{X}_1} \star_M \left(\prod_{k=2}^{n-1} e^{\frac{(1+\iota s)t\mathcal{X}_k}{2}} \right) \star_M e^{t\mathcal{X}_n} \star_M \left(\prod_{k=n-1}^2 e^{\frac{(1-\iota s)t\mathcal{X}_k}{2}} \right) \right] \rho_0(s) ds. \quad (100)$$

By taking the expectation of both sides and applying the independence property for all random tensors \mathcal{X}_k , we obtain

$$\mathbb{E} \text{Tr} e^{\left(\sum_{k=1}^n t\mathcal{X}_k \right)} \leq \int_{-\infty}^{\infty} \text{Tr} \left[\mathbb{E}(e^{t\mathcal{X}_1}) \star_M \left(\prod_{k=2}^{n-1} \mathbb{E}(e^{\frac{(1+\iota s)t\mathcal{X}_k}{2}}) \right) \star_M \mathbb{E}(e^{t\mathcal{X}_n}) \star_M \left(\prod_{k=n-1}^2 \mathbb{E}(e^{\frac{(1-\iota s)t\mathcal{X}_k}{2}}) \right) \right] \rho_0(s) ds. \quad (101)$$

By combining Eq. (101) with Lemma 11, the theorem is proved. \square

6 Conclusions

In this work, we extend Araki–Lieb–Thirring (ALT) inequality, Golden–Thompson (GT) inequality and logarithmic trace inequality to arbitrary many tensors. Our proofs utilize complex interpolation theory and asymptotic spectral pinching, providing a powerful mechanism to deal with multivariate trace inequalities for tensors. We then apply tensor Golden–Thompson inequality to provide the tail bound for the independent sum of tensors and this bound will play a crucial role in high-dimensional probability and statistical data analysis. We believe our work can be applied to various science and engineering problems involving multivariate variables [23, 6].

Acknowledgments

The helpful comments of the referees are gratefully acknowledged.

REFERENCES

- [1] R. AHLWEDE AND A. WINTER, *Strong converse for identification via quantum channels*, IEEE Transactions on Information Theory, 48 (2002), pp. 569–579.
- [2] T. ANDO AND F. HIAI, *Log majorization and complementary golden-thompson type inequalities*, Linear Algebra and its Applications, 197 (1994), pp. 113–131.
- [3] H. ARAKI, *On an inequality of lieb and thirring*, LMaPh, 19 (1990), pp. 167–170.

- [4] E. CARLEN, *Trace inequalities and quantum entropy: an introductory course*, Entropy and the quantum, 529 (2010), pp. 73–140.
- [5] I. CSISZÁR, *The method of types [information theory]*, IEEE Transactions on Information Theory, 44 (1998), pp. 2505–2523.
- [6] S. A. DILDABAYEV AND G. K. ZAKIRYANOVA, *Green's tensor and 2dimensional bie method in static of massif with arbitrary anisotropy*, Mathematics and Statistics, 7 (2019), pp. 158–171, DOI: 10.13189/ms.2019.070502.
- [7] H. G. EGGLESTON, *Convexity*, no. 47, CUP Archive, 1958.
- [8] P. J. FORRESTER AND C. J. THOMPSON, *The golden-thompson inequality: Historical aspects and random matrix applications*, Journal of Mathematical Physics, 55 (2014), p. 023503.
- [9] F. HIAI AND D. PETZ, *The golden-thompson trace inequality is complemented*, Linear algebra and its applications, 181 (1993), pp. 153–185.
- [10] I. I. HIRSCHMAN, *A convexity theorem for certain groups of transformations*, Journal d'Analyse Mathématique, 2 (1952), pp. 209–218.
- [11] F. KITTANEH, *Norm inequalities for certain operator sums*, Journal of Functional Analysis, 143 (1997), pp. 337–348.
- [12] M. KLEINMANN, H. KAMPERMANN, T. MEYER, AND D. BRUSS, *Physical purification of quantum states*, Physical Review A, 73 (2006), p. 062309.
- [13] H. KOSAKI, *An inequality of araki-lieb-thirring (von neumann algebra case)*, Proceedings of the American Mathematical Society, 114 (1992), pp. 477–481.
- [14] A. LENARD AND P. HALMOS, *Generalization of the Golden-Thompson inequality*, Indiana University Mathematics Journal, (1971), pp. 457–467.
- [15] M. LIANG AND B. ZHENG, *Further results on moore–penrose inverses of tensors with application to tensor nearness problems*, Computers & Mathematics with Applications, 77 (2019), pp. 1282–1293.
- [16] E. H. LIEB AND M. B. RUSKAI, *A fundamental property of quantum-mechanical entropy*, Physical Review Letters, 30 (1973), p. 434.
- [17] E. H. LIEB AND M. B. RUSKAI, *Proof of the strong subadditivity of quantum-mechanical entropy*, Les rencontres physiciens-mathématiciens de Strasbourg-RCP25, 19 (1973), pp. 36–55.
- [18] E. H. LIEB AND W. E. THIRRING, *Inequalities for the moments of the eigenvalues of the schrodinger hamiltonian and their relation to sobolev inequalities*, in The Stability of Matter: From Atoms to Stars, Springer, 1991, pp. 135–169.
- [19] H. NEUDECKER, *A note on kronecker matrix products and matrix equation systems*, SIAM Journal on Applied Mathematics, 17 (1969), pp. 603–606.
- [20] G. NI, *Hermitian tensor and quantum mixed state*, arXiv preprint arXiv:1902.02640, (2019).
- [21] L. QI, *Eigenvalues of a real supersymmetric tensor*, Journal of Symbolic Computation, 40 (2005), pp. 1302–1324.
- [22] L. QI, H. CHEN, AND Y. CHEN, *Tensor eigenvalues and their applications*, vol. 39, Springer, 2018.
- [23] S. SASIWANNAPONG, S. SUKPARUNGSEE, P. BUSABABODHIN, AND Y. AREEPONG, *Construction of bivariate copulas on a multivariate exponentially weighted moving average control chart*, Mathematics and Statistics, 8 (2020), pp. 520–526, DOI: 10.13189/ms.2020.080504.
- [24] I. SEGAL, *Notes towards the construction of nonlinear relativistic quantum fields. iii: Properties of the c^* -dynamics for a certain class of interactions*, Bulletin of the American Mathematical Society, 75 (1969), pp. 1390–1395.
- [25] E. M. STEIN, *Interpolation of linear operators*, Transactions of the American Mathematical Society, 83 (1956), pp. 482–492.
- [26] C. J. THOMPSON, *Inequality with applications in statistical mechanics*, Journal of Mathematical Physics, 6 (1965), pp. 1812–1813.

- [27] J. A. TROPP, *User-friendly tail bounds for sums of random matrices*, Foundations of computational mathematics, 12 (2012), pp. 389–434.
- [28] B.-Y. WANG AND F. ZHANG, *Trace and eigenvalue inequalities for ordinary and hadamard products of positive semidefinite hermitian matrices*, SIAM journal on matrix analysis and applications, 16 (1995), pp. 1173–1183.
- [29] A. WEHRL, *General properties of entropy*, Reviews of Modern Physics, 50 (1978), p. 221.