GENERALIZED LOG-MAJORIZATION
AND MULTIVARIATE TRACE INEQUALITIES

FUMIO HIAI, ROBERT KÖNIG, AND MARCO TOMAMICHEL

ABSTRACT. We show that recent multivariate generalizations of the Araki-Lieb-Thirring inequality and the Golden-Thompson inequality [Sutter, Berta, and Tomamichel, *Comm. Math. Phys.* (2017)] for Schatten norms hold more generally for all unitarily invariant norms and certain variations thereof. The main technical contribution is a generalization of the concept of log-majorization which allows us to treat majorization with regards to logarithmic integral averages of vectors of singular values.

1. INTRODUCTION

Majorization and log-majorization are powerful and versatile tools for proving trace and norm inequalities (see, e.g., [1, 15, 10] for overviews on the topic). A fundamental property of unitarily invariant norms (including Schatten p-norms and the trace norm) can roughly be stated as follows (see [5, Thm. IV.2.2] or [9, Prop. 4.4.13]):

> *For two matrices A and B, the singular values of A are weakly majorized by the singular values of B if and only if $\|A\| \leq \|B\|$ for every unitarily invariant norm $\|\cdot\|$.* $\quad$ ($\star$)

A natural approach to prove norm inequalities for general unitarily invariant norms then proceeds as follows: First, the desired inequality is shown for the operator norm where such inequalities often boil down to operator inequalities and are easier to prove. Next, it is shown using antisymmetric tensor power calculus that the operator norm inequality implies log-majorization and thus weak majorization of the eigenvalues. Consequently, the desired inequalities follow directly from ($\star$).

Let us illustrate this approach with an example (the reader unfamiliar with the notation is referred to Section 2). For two positive definite operators A_1, A_2 and any $\theta \in (0, 1)$, consider the following special case of the Araki-Lieb-Thirring inequality [14, 4]:

$$\left\| \left(A_1^{\frac{\theta}{2}} A_2^\theta A_1^{\frac{\theta}{2}} \right)^{\frac{1}{\theta}} \right\| \leq \left\| A_1^{\frac{1}{2}} A_2 A_1^{\frac{1}{2}} \right\|, \tag{1}$$

where $\|\cdot\|$ denotes the operator norm. Its elementary proof, using operator monotonicity of $t \mapsto t^\theta$ as its main ingredient, simply argues that

$$A_1^{\frac{1}{2}} A_2 A_1^{\frac{1}{2}} \leq I \implies A_2 \leq A_1^{-1} \implies A_2^\theta \leq A_1^{-\theta} \implies A_1^{\frac{\theta}{2}} A_2^\theta A_1^{\frac{\theta}{2}} \leq I. \tag{2}$$

Inequality (1) for positive semi-definite operators follows by continuity. Applying (1) to the antisymmetric powers $\wedge^k A_1$ and $\wedge^k A_2$, using Properties (b),

(c) and (d) of the antisymmetric tensor power discussed in the next section, we find

$$\boldsymbol{\lambda}\left(\left(A_1^{\frac{\theta}{2}}A_2^{\theta}A_1^{\frac{\theta}{2}}\right)^{\frac{1}{\theta}}\right) \prec_{\log} \boldsymbol{\lambda}\left(A_1^{\frac{1}{2}}A_2A_1^{\frac{1}{2}}\right). \tag{3}$$

Here $\boldsymbol{\lambda}(A)$ is a vector comprising the eigenvalues of A in decreasing order counting multiplicities and $\prec_{\log}$ denotes log-majorization. Since log-majorization implies weak majorization, the relation $(\star)$ allows us to lift (1) to arbitrary unitarily invariant norms, including the trace. In fact, log-majorization is stronger than weak majorization and thus allows us to derive stronger norm inequalities (see, e.g., [9]).

The use of the antisymmetric tensor power approach has so far been restricted to matrix functions made from operations of products, absolute values and powers (see, e.g., [4] (explained above) and [3, 2]). In this work we extend the approach to settings with a logarithmic integral average so that it can be applied to recent multivariate trace inequalities [16]. For example, [16, Thm. 3.2] specialized to the operator norm and three positive semi-definite matrices A_1, A_2, A_3 and $\theta \in (0,1)$ generalizes the Araki-Lieb-Thirring inequality and reads

$$\log\left\|\left|A_1^{\theta}A_2^{\theta}A_3^{\theta}\right|^{\frac{1}{\theta}}\right\| \leq \int_{-\infty}^{\infty} \log\left\|A_1A_2^{1+\mathrm{i}t}A_3\right\| \, \mathrm{d}\beta_{\theta}(t), \tag{4}$$

where $\mathrm{d}\beta_{\theta}(t)$ is some probability measure on $\mathbb{R}$. Using the antisymmetric tensor power technique this can be transformed into a log-majorization relation, namely[1]

$$\boldsymbol{\lambda}\left(\left|A_1^{\theta}A_2^{\theta}A_3^{\theta}\right|^{\frac{1}{\theta}}\right) \prec_{\log} \exp\int_{-\infty}^{\infty}\log\boldsymbol{\lambda}\left(\left|A_1A_2^{1+\mathrm{i}t}A_3\right|\right)\,\mathrm{d}\beta_{\theta}(t). \tag{5}$$

However, known results for majorization or log-majorization in the spirit of $(\star)$ do not apply to (5) due to the integral average of vectors on the right-hand side.

In Sections 3 we extend $(\star)$ to the case of weak majorization relations where the right-hand side contains an integral average of vectors. Our first main result, Theorem 7 in Section 4, deals with weak log-majorization relations of the form (5). It establishes that the weak log-majorization relation is equivalent to two other conditions involving unitarily invariant norms, and in particular implies that (4) holds for all unitarily invariant norms and certain variations thereof. Our second main result, split into Theorems 10 and 14 in Section 5, proves a similar characterization directly for the log-majorization relation in (5) and implies even stronger inequalities for unitarily invariant norms. For the special case where no average is present, Propositions 8 and 13 imply new characterizations of weak log-majorization and log-majorization, respectively. The implications for multivariate trace

---

[1]The details of this derivation are given in Section 6.

inequalities are discussed in Section 6. There we present multivariate generalizations of the Araki-Lieb-Thirring inequality, the Golden-Thompson inequality [8, 17] and Lieb's triple matrix inequality [13] — beyond the generalizations recently established in [16]. We also provide a simplified proof of (4) for arbitrarily many matrices and the operator norm in Appendix A.

2. Preliminaries

Majorization. Let $\mathcal{H}$ be a Hilbert space of dimension $d := \dim \mathcal{H} < \infty$, $\mathcal{L}(\mathcal{H})$ the set of linear operators on $\mathcal{H}$, $\mathcal{P}(\mathcal{H})$ be the set of all positive semi-definite operators in $\mathcal{L}(\mathcal{H})$, and $\mathcal{P}_+(\mathcal{H})$ the set of all invertible (positive definite) operators in $\mathcal{P}(\mathcal{H})$. For self-adjoint $A, B \in \mathcal{L}(\mathcal{H})$, we write $A \geq B$ to indicate that $A - B \in \mathcal{P}(\mathcal{H})$.

We use bold font $\boldsymbol{a} = (a_1, \ldots, a_d) \in \mathbb{R}^d$ to denote vectors. Let $\mathbb{R}_+^d := \{\boldsymbol{a} \in \mathbb{R}^d : a_1, \ldots, a_d \geq 0\}$. For $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^d$ such that $a_1 \geq \cdots \geq a_d$ and $b_1 \geq \cdots \geq b_d$, *weak majorization*, denoted by $\boldsymbol{a} \prec_w \boldsymbol{b}$, is the relation

$$\sum_{i=1}^k a_i \leq \sum_{i=1}^k b_i, \qquad k \in [d], \tag{6}$$

where we used $[d] := \{1, 2, \ldots, d\}$. *Majorization*, $\boldsymbol{a} \prec \boldsymbol{b}$, additionally requires that equality holds for $k = d$. For $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}_+^d$ such that $a_1 \geq \cdots \geq a_d$ and $b_1 \geq \cdots \geq b_d$, *weak log-majorization*, $\boldsymbol{a} \prec_{w\log} \boldsymbol{b}$, is the relation

$$\prod_{i=1}^k a_i \leq \prod_{i=1}^k b_i, \qquad k \in [d], \tag{7}$$

and *log-majorization*, $\boldsymbol{a} \prec_{\log} \boldsymbol{b}$, additionally requires equality for $k = d$. For any function f on $\mathbb{R}_+$ we write $f(\boldsymbol{a}) = (f(a_1), \ldots, f(a_d))$ with conventions $\log 0 := -\infty$ and $e^{-\infty} = 0$. Moreover, weak majorization $\prec_w$ makes sense even for vectors having entries $-\infty$. With these conventions, it is evident that $\boldsymbol{a} \prec_{w\log} \boldsymbol{b}$ if and only if $\log \boldsymbol{a} \prec_w \log \boldsymbol{b}$. The following relation takes a prominent role (see [5, Thm. II.3.3] and [9, Prop. 4.1.4]):

Lemma 1. *For any convex function $f : [0, \infty) \to [0, \infty)$, we have $\boldsymbol{a} \prec \boldsymbol{b} \implies f(\boldsymbol{a}) \prec_w f(\boldsymbol{b})$. Moreover, if f is also non-decreasing, then $\boldsymbol{a} \prec_w \boldsymbol{b} \implies f(\boldsymbol{a}) \prec_w f(\boldsymbol{b})$.*

As a direct consequence when applied to the exponential function, weak log-majorization implies weak majorization.

Unitarily invariant norms. Let us denote the eigenvalues of a self-adjoint operator $A \in \mathcal{L}(\mathcal{H})$ in decreasing order counting multiplicities by the vector $\boldsymbol{\lambda}(A) = (\lambda_1(A), \ldots, \lambda_d(A))$. Moreover, let $\|\cdot\|_\Phi$ be a unitarily invariant norm on $\mathcal{P}(\mathcal{H})$ and $\Phi : \mathbb{R}_+^d \to \mathbb{R}_+$ the corresponding gauge function so that

$$\|L\|_\Phi = \||L|\|_\Phi = \Phi(\boldsymbol{\lambda}(|L|)). \tag{8}$$

(We refer the reader to [5, Sec. IV] for an introduction to unitarily invariant norms. The bijective correspondence between symmetric gauge functions on $\mathbb{R}_+^d$ and unitarily invariant norms on $\mathcal{P}(\mathcal{H})$ is due to von Neumann [18].)

Of particular interest here are Ky Fan norms. For $k \in [d]$, the Ky Fan k-norm, $\| \cdot \|_{(k)} : \mathcal{L}(\mathcal{H}) \to \mathbb{R}_+$, is defined as

$$L \mapsto \|L\|_{(k)} := \sum_{i=1}^{k} \lambda_i(|L|). \tag{9}$$

In particular, $\| \cdot \|_{(1)}$ is the operator norm $\| \cdot \|$. Another important and familiar one is the Schatten p-norm $\|L\|_p := (\mathrm{tr}\,|L|^p)^{1/p}$ for $p \geq 1$. In particular, $\| \cdot \|_1$ is the trace norm. The definition of $\| \cdot \|_p$ makes sense even for $0 < p < 1$ as a quasi-norm.

The following lemma is a Hölder inequality for the gauge function Φ and follows from [5, Thm. IV.1.6].

Lemma 2. *For $l \in [m]$ let $\boldsymbol{a}_l = (a_{l1}, \ldots, a_{ld}) \in \mathbb{R}_+^d$ and $\beta_l > 0$ with $\sum_{l=1}^m \beta_l = 1$. Then*

$$\Phi\left(\prod_{l=1}^m \boldsymbol{a}_l^{\beta_l}\right) \leq \prod_{l=1}^m \Phi^{\beta_l}(\boldsymbol{a}_l), \tag{10}$$

where

$$\prod_{l=1}^m \boldsymbol{a}_l^{\beta_l} := \left(\prod_{l=1}^m a_{l1}^{\beta_l}, \ldots, \prod_{l=1}^m a_{ld}^{\beta_l}\right). \tag{11}$$

Proof. The lemma for $m = 2$ is [5, Thm. IV.1.6]. The case $m = 3$ is shown as

$$\Phi\left(\boldsymbol{a}_1^{\beta_1} \boldsymbol{a}_2^{\beta_2} \boldsymbol{a}_3^{\beta_3}\right) = \Phi\left(\left(\boldsymbol{a}_1^{\frac{\beta_1}{\beta_1+\beta_2}} \boldsymbol{a}_2^{\frac{\beta_2}{\beta_1+\beta_2}}\right)^{\beta_1+\beta_2} \boldsymbol{a}_3^{\beta_3}\right) \tag{12}$$

$$\leq \Phi^{\beta_1+\beta_2}\left(\boldsymbol{a}_1^{\frac{\beta_1}{\beta_1+\beta_2}} \boldsymbol{a}_2^{\frac{\beta_2}{\beta_1+\beta_2}}\right) \Phi^{\beta_3}(\boldsymbol{a}_3) \leq \Phi^{\beta_1}(\boldsymbol{a}_1) \Phi^{\beta_2}(\boldsymbol{a}_2) \Phi^{\beta_3}(\boldsymbol{a}_3). \tag{13}$$

The general case can be shown similarly by induction. $\square$

Antisymmetric tensor product. For $k \in [d]$, let $\mathcal{H}^{\otimes k}$ be the kth tensor power of $\mathcal{H}$ and let $\mathcal{H}^{\wedge k}$ denote the antisymmetric subspace of $\mathcal{H}^{\otimes k}$. The kth *antisymmetric tensor power*, $\wedge^k : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H}^{\wedge k})$, maps any linear operator L to the restriction of $L^{\otimes k} \in \mathcal{L}(\mathcal{H}^{\otimes k})$ to the antisymmetric subspace $\mathcal{H}^{\wedge k}$ of $\mathcal{H}^{\otimes k}$. It satisfies the following rules (see, e.g., [5, Sec. I.5 and p. 18]):

Lemma 3. *Let $L, K \in \mathcal{L}(\mathcal{H})$ and $A \in \mathcal{P}(\mathcal{H})$. For any $k \in [d]$, we have:*

(a) $(\wedge^k L)^\dagger = \wedge^k(L^\dagger)$,
(b) $(\wedge^k L)(\wedge^k K) = \wedge^k(LK)$,
(c) $(\wedge^k A)^z = \wedge^k(A^z)$ for all $z \in \mathbb{C}$, and
(d) $\| \wedge^k L \| = \prod_{i=1}^k \lambda_i(|L|)$.

In particular, we note that if $L \in \mathcal{L}(\mathcal{H})$ is positive semi-definite, so is its antisymmetric tensor power $\wedge^k L \in \mathcal{L}(\mathcal{H}^{\wedge k})$.

3. (Weak) majorization with integral average

Let Ξ be a σ-compact metric space and ν a probability measure on the Borel σ-field of Ξ. Let $A \in \mathcal{L}(\mathcal{H})$ and $\xi \in \Xi \mapsto B_\xi \in \mathcal{L}(\mathcal{H})$ be a continuous function such that A and B_ξ for all $\xi \in \Xi$ are self-adjoint and $\sup\{\|B_\xi\| : \xi \in \Xi\} < \infty$. We use the convention

$$\int_\Xi \boldsymbol{\lambda}(B_\xi) \, d\nu(\xi) := \left(\int_\Xi \lambda_1(B_\xi) \, d\nu(\xi), \dots, \int_\Xi \lambda_d(B_\xi) \, d\nu(\xi) \right). \qquad (14)$$

The following two theorems are characterizations of weak majorization and majorization in the setting with integral average. They will be used in Sections 4 and 5.

Theorem 4. *With Ξ, ν, and self-adjoint $A, B_\xi \in \mathcal{L}(\mathcal{H})$ given as above, the following statements are equivalent:*
(a) $\boldsymbol{\lambda}(A) \prec_w \int_\Xi \boldsymbol{\lambda}(B_\xi) d\nu(\xi)$;
(b) for every non-decreasing convex function $f : \mathbb{R} \to [0, \infty)$ and for every unitarily invariant norm $\|\cdot\|$,

$$\|f(A)\| \leq \int_\Xi \|f(B_\xi)\| \, d\nu(\xi). \qquad (15)$$

Proof. Assume (a) and let f be as in (b). We have

$$\boldsymbol{\lambda}(f(A)) = f(\boldsymbol{\lambda}(A)) \prec_w f\left(\int_\Xi \boldsymbol{\lambda}(B_\xi) d\nu(\xi) \right) \qquad (16)$$

thanks to [9, Prop. 4.1.4(2)]. Since

$$f\left(\int_\Xi \boldsymbol{\lambda}(B_\xi) \, d\nu(\xi) \right) \leq \int_\Xi f(\boldsymbol{\lambda}(B_\xi)) \, d\nu(\xi) = \int_\Xi \boldsymbol{\lambda}(f(B_\xi)) \, d\nu(\xi), \qquad (17)$$

we have $\boldsymbol{\lambda}(f(A)) \prec_w \int_\Xi \boldsymbol{\lambda}(f(B_\xi)) \, d\nu(\xi)$. Since both sides of this relation are non-negative vectors, applying the gauge function Φ to them yields (see [9, Lemma 4.4.2])

$$\|f(A)\|_\Phi \leq \Phi\left(\int_\Xi \boldsymbol{\lambda}(f(B_\xi)) \, d\nu(\xi) \right) \qquad (18)$$

$$\leq \int_\Xi \Phi(\boldsymbol{\lambda}(f(B_\xi))) \, d\nu(\xi) = \int_\Xi \|f(B_\xi)\|_\Phi \, d\nu(\xi). \qquad (19)$$

Hence (b) holds.

To prove the converse, assume (b). Since B_ξ is uniformly bounded, there is an $\alpha > 0$ such that $A + \alpha I \geq 0$ and $B_\xi + \alpha I \geq 0$ for all $\xi \in \Xi$. We evaluate inequality (15) for $f(x) := \max\{x + \alpha, 0\}$ and the Ky Fan norm $\|\cdot\| = \|\cdot\|_{(k)}$ to get

$$\sum_{i=1}^k (\lambda_i(A) + \alpha) \leq \sum_{i=1}^k \int_\Xi (\lambda_i(B_\xi) + \alpha) \, d\nu(\xi). \qquad (20)$$

Therefore, $\sum_{i=1}^k \lambda_i(A) \leq \sum_{i=1}^k \int_\Xi \lambda_i(B_\xi) \, d\nu(\xi)$, which implies (a). $\qquad \square$

Remark 1. *In the case where $A, B_\xi \in \mathcal{P}(\mathcal{H})$ for all $\xi \in \Xi$, conditions (a) and (b) are also equivalent to the statement*

(c) for every unitarily invariant norm $\| \cdot \|$,

$$\|A\| \leq \int_\Xi \|B_\xi\| \, d\nu(\xi) \,. \tag{21}$$

Indeed, (b) $\implies$ (c) is obvious by letting $f(x) := \max\{x, 0\}$ in (b), and (c) $\implies$ (a) is seen by evaluating (c) for the Ky Fan norm $\| \cdot \| = \| \cdot \|_{(k)}$. The assumption $B_\xi \in \mathcal{P}(\mathcal{H})$ is essential for the latter implication. Statements (a) and (c) constitute a generalization of $(\star)$ when applied to $|A|$ and $|B_\xi|$ for $A, B_\xi \in \mathcal{L}(\mathcal{H})$.

Theorem 5. *With Ξ, ν, and self-adjoint $A, B_\xi \in \mathcal{L}(\mathcal{H})$ given as above, the following statements are equivalent:*

(d) $\boldsymbol{\lambda}(A) \prec \int_\Xi \boldsymbol{\lambda}(B_\xi) d\nu(\xi)$;
(e) inequality (15) holds for every convex function $f : \mathbb{R} \to [0, \infty)$ and for every unitarily invariant norm $\| \cdot \|$.

Proof. Assume (d) and let f be as in (e). It is obvious that $\boldsymbol{\lambda}(f(A)) \approx f(\boldsymbol{\lambda}(A))$, where for $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^d$, $\boldsymbol{a} \approx \boldsymbol{b}$ means that the entries of $\boldsymbol{a}$ coincide with those of $\boldsymbol{b}$ up to a permutation. Since Lemma 1 gives

$$f(\boldsymbol{\lambda}(A)) \prec_w f\left(\int_\Xi \boldsymbol{\lambda}(B_\xi) \, d\nu(\xi)\right) \leq \int_\Xi f(\boldsymbol{\lambda}(B_\xi)) \, d\nu(\xi) \,, \tag{22}$$

we have

$$\|f(A)\|_\Phi \leq \Phi\left(\int_\Xi f(\boldsymbol{\lambda}(B_\xi)) \, d\nu(\xi)\right) \tag{23}$$

$$\leq \int_\Xi \Phi(f(\boldsymbol{\lambda}(B_\xi))) \, d\nu(\xi) = \int_\Xi \|f(B_\xi)\|_\Phi \, d\nu(\xi) \,. \tag{24}$$

Hence (e) holds.

Conversely, if (e) is satisfied, then by Theorem 4 we have $\boldsymbol{\lambda}(A) \prec_w \int_\Xi \boldsymbol{\lambda}(B_\xi) \, d\nu(\xi)$. Hence, to prove (d), it suffices to show that $\operatorname{tr} A \geq \int_\Xi \operatorname{tr} B_\xi \, d\nu(\xi)$. Choose an $\alpha > 0$ such that $A \leq \alpha I$ and $B_\xi \leq \alpha I$ for all $\xi \in \Xi$. For $f(x) := \max\{\alpha - x, 0\}$ and the trace norm $\| \cdot \| = \| \cdot \|_1$, condition (e) implies that

$$\operatorname{tr}(\alpha I - A) \leq \int_\Xi \operatorname{tr}(\alpha I - B_\xi) \, d\nu(\xi) \,, \tag{25}$$

giving the desired inequality. $\qquad\square$

4. Weak log-majorization with integral average

In the following, we assume that $A, B_\xi \in \mathcal{P}(\mathcal{H})$ for all $\xi \in \Xi$. The above equivalent conditions (a), (b) of Theorem 4 and (c) of Remark 1 correspond to weak majorization, and (d), (e) of Theorem 5 correspond to majorization. We now consider stronger conditions than those, corresponding to (weak) log-majorization. We have the following chain of implications, where the last condition (4) is (b) of Theorem 4 and we use as in (14) the convention

$$\int_\Xi \log \boldsymbol{\lambda}(B_\xi) \, d\nu(\xi) := \left(\int_\Xi \log \lambda_1(B_\xi) \, d\nu(\xi), \dots, \int_\Xi \log \lambda_d(B_\xi) \, d\nu(\xi)\right). \tag{26}$$

Proposition 6. *With* Ξ, ν, A *and* B_ξ *given as above, consider the following statements:*

(1) $\boldsymbol{\lambda}(A) \prec_{\log} \exp \int_{\Xi} \log \boldsymbol{\lambda}(B_\xi) \, d\nu(\xi)$
(2) $\boldsymbol{\lambda}(A) \prec_{w\log} \exp \int_{\Xi} \log \boldsymbol{\lambda}(B_\xi) \, d\nu(\xi)$
(3) $\boldsymbol{\lambda}(A) \prec_{w\log} \int_{\Xi} \boldsymbol{\lambda}(B_\xi) \, d\nu(\xi)$
(4) $\boldsymbol{\lambda}(A) \prec_{w} \int_{\Xi} \boldsymbol{\lambda}(B_\xi) \, d\nu(\xi)$.
Then $(1) \implies (2) \implies (3) \implies (4)$.

Proof. The implication $(1) \implies (2)$ is trivial. Implication $(2) \implies (3)$ follows by Jensen's inequality, and $(3) \implies (4)$ follows by Lemma 1. □

The following theorem constitutes part of our main results and characterizes the second condition in this chain. We will give a similar characterization of the first condition in Theorems 10 and 14 below.

Theorem 7. *With* Ξ, ν, A *and* B_ξ *given as above, the following statements are equivalent:*

(i) $\boldsymbol{\lambda}(A) \prec_{w\log} \exp \int_{\Xi} \log \boldsymbol{\lambda}(B_\xi) \, d\nu(\xi)$, *i.e.,*
$\log \boldsymbol{\lambda}(A) \prec_{w} \int_{\Xi} \log \boldsymbol{\lambda}(B_\xi) \, d\nu(\xi)$;
(ii) for every continuous non-decreasing function $f : [0, \infty) \to [0, \infty)$ *such that* $x \mapsto \log f(e^x)$ *is convex on* $\mathbb{R}$, *and for every unitarily invariant norm* $\|\!|\cdot|\!\|$,

$$\|\!| f(A) |\!\| \leq \exp \int_{\Xi} \log \|\!| f(B_\xi) |\!\| \, d\nu(\xi) \, ; \tag{27}$$

(iii) for every continuous non-decreasing function $g : [0, \infty) \to [0, \infty)$ *such that* $x \mapsto g(e^x)$ *is convex on* $\mathbb{R}$, *and for every unitarily invariant norm* $\|\!|\cdot|\!\|$,

$$\|\!| g(A) |\!\| \leq \int_{\Xi} \|\!| g(B_\xi) |\!\| \, d\nu(\xi) \, . \tag{28}$$

When Ξ is a one-point set, Theorem 7 reduces to [9, Prop. 4.4.13], except condition (ii). The proof of (ii) $\implies$ (i) given below implies the next proposition, which appears to be a new characterization of weak log-majorization.

Proposition 8. *For* $A, B \in \mathcal{P}(\mathcal{H})$, $\boldsymbol{\lambda}(A) \prec_{w\log} \boldsymbol{\lambda}(B)$ *if and only if*
$$\|A^p\|_{(k)} \leq \|B^p\|_{(k)}, \qquad p > 0, \quad k \in [d]. \tag{29}$$

Before presenting the proof of Theorem 7, we discuss the convexity conditions appearing in (ii) and (iii). We will use the following properties in the proof of Theorem 7 and again in Section 5.

Lemma 9. *Let* $f : (0, \infty) \to [0, \infty)$ *be a continuous function such that* $x \mapsto \log f(e^x)$ *is convex on* $\mathbb{R}$. *Then:*

(1) $f(x) > 0$ *for any* $x > 0$ *unless* $f \equiv 0$.
(2) $f(0^+) := \lim_{x \searrow 0} f(x)$ *exists in* $[0, \infty]$, *and if* $f(0^+) < \infty$ *then* f *is non-decreasing on* $(0, \infty)$.

Similarly, if $g : (0, \infty) \to [0, \infty)$ *is a function such that* $x \mapsto g(e^x)$ *is convex on* $\mathbb{R}$, *then* g *is automatically continuous on* $(0, \infty)$, $g(0^+) := \lim_{x \searrow 0} g(x)$ *exists in* $[0, \infty]$, *and if* $g(0^+) < \infty$ *then* g *is non-decreasing on* $(0, \infty)$.

In particular, (2) implies that a non-decreasing continuous function $f :$ $(0, \infty) \to [0, \infty)$ with the property that $x \mapsto \log f(e^x)$ is convex on $\mathbb{R}$ extends to a continuous function $f : [0, \infty) \to [0, \infty)$. This corresponds to (ii) of Theorem 7. Analogously, if $g : (0, \infty) \to [0, \infty)$ is continuous and non-decreasing, and $x \mapsto g(e^x)$ is convex on $\mathbb{R}$, then g extends to a continuous function $g : [0, \infty) \to [0, \infty)$. Such functions appear in (iii) of Theorem 7. In Section 5, we will drop the assumption that f, g are non-decreasing and instead consider majorization instead of weak majorization.

For instance, for any $\alpha \geq 0$ and any $p > 0$, $f(x) := (\alpha + x)^p$ is an increasing function on $[0, \infty)$ such that $\log f(e^x)$ is convex on $\mathbb{R}$. For any $p > 0$, $g(x) := \log(1 + x^p)$ is an increasing function on $[0, \infty)$ such that $g(e^x)$ is convex on $\mathbb{R}$ but $\log g(e^x)$ is concave on $\mathbb{R}$.

Proof. (1) Let $h(x) := \log f(e^x)$, $x \in \mathbb{R}$. If the conclusion is not true, then there is an $\alpha > 0$ such that $f(\alpha) = 0$ and $f(x) > 0$ for $x \in (\alpha - \delta, \alpha)$ or $x \in (\alpha, \alpha + \delta)$ for some $\delta > 0$. Since $\lim_{x \to \log \alpha} h(x) = \log f(\alpha) = -\infty$, h cannot be convex around $\log \alpha$.

(2) From the convexity of h on $\mathbb{R}$ it follows that $h(-\infty) := \lim_{x \to -\infty} h(x)$ exists in $[-\infty, \infty]$. This implies that $f(0^+)$ exists in $[0, \infty]$. Unless h is non-decreasing on $\mathbb{R}$, the convexity of h implies that $h(-\infty) = \infty$ and so $f(0^+) = \infty$.

The proof of the statements for g is similar and omitted here. $\qquad\square$

Proof of Theorem 7. In comparing conditions (ii) and (iii), the convexity of $\log f(e^x)$ in (ii) is stronger than the convexity of $f(e^x)$ in (iii). Correspondingly, the conclusion of (ii) is stronger than that of (iii). Hence it is not clear how to pass directly between (ii) and (iii). The proof is thus split into four parts, corresponding to the implications (i) $\implies$ (ii), (ii) $\implies$ (i), (i) $\implies$ (iii), and (iii) $\implies$ (i).

Proof of (i) $\implies$ (ii). First, assume that $A, B_\xi \in \mathcal{P}_+(\mathcal{H})$ and $B_\xi \geq \varepsilon I$ for all $\xi \in \Xi$ with some $\varepsilon > 0$. Because f is non-increasing on $[0, \infty)$, we have

$$\log \boldsymbol{\lambda}(f(A)) = \log f(e^{\log \boldsymbol{\lambda}(A)}) . \tag{30}$$

Since $\log f(e^x)$ is convex on $\mathbb{R}$, Lemma 1 yields

$$\log f(e^{\log \boldsymbol{\lambda}(A)}) \prec_w \log f\left(\exp \int_\Xi \log \boldsymbol{\lambda}(B_\xi) \, d\nu(\xi)\right) \tag{31}$$

$$\leq \int_\Xi \log f(\boldsymbol{\lambda}(B_\xi)) \, d\nu(\xi) \tag{32}$$

from condition (i). Therefore, we have with (30)

$$\boldsymbol{\lambda}(f(A)) \prec_w \exp \int_\Xi \log f(\boldsymbol{\lambda}(B_\xi)) \, d\nu(\xi) \tag{33}$$

so that

$$\|f(A)\|_\Phi = \Phi(\boldsymbol{\lambda}(f(A))) \leq \Phi\left(\exp \int_\Xi \log f(\boldsymbol{\lambda}(B_\xi)) \, d\nu(\xi)\right). \tag{34}$$

By Lemma 9 (1) we may assume that $f(x) > 0$ for any $x > 0$, so the function $\xi \mapsto \log f(\lambda_i(B_\xi))$, $i \in [d]$, as well as $\xi \mapsto \log \|f(B_\xi)\|_\Phi$ are bounded and continuous on Ξ. Since, moreover, $\nu(\Xi) = 1 = \sup\{\nu(K) : K \subset \Xi \text{ compact}\}$

due to the σ-compactness of Ξ, a standard compactness argument shows that there are $\xi_l^{(m)} \in \Xi$ and $\beta_l^{(m)} > 0$ for $l \in [m]$ and $m \in \mathbb{N}$ with $\sum_{l=1}^{m} \beta_l^{(m)} = 1$ such that

$$\int_\Xi \log f(\lambda_i(B_\xi)) \, d\nu(\xi) = \lim_{m\to\infty} \sum_{l=1}^{m} \beta_l^{(m)} \log f\left(\lambda_i\left(B_{\xi_l^{(m)}}\right)\right), \qquad i \in [d], \tag{35}$$

$$\int_\Xi \log \|f(B_\xi)\|_\Phi \, d\nu(\xi) = \lim_{m\to\infty} \sum_{l=1}^{m} \beta_l^{(m)} \log \left\|f\left(B_{\xi_l^{(m)}}\right)\right\|_\Phi. \tag{36}$$

Therefore,

$$\Phi\left(\exp \int_\Xi \log f(\boldsymbol{\lambda}(B_\xi)) \, d\nu(\xi)\right) = \lim_{m\to\infty} \Phi\left(\prod_{l=1}^{m} f\left(\boldsymbol{\lambda}\left(B_{\xi_l^{(m)}}\right)\right)^{\beta_l^{(m)}}\right), \tag{37}$$

$$\exp \int_\Xi \log \|f(B_\xi)\|_\Phi \, d\nu(\xi) = \lim_{m\to\infty} \prod_{l=1}^{m} \left\|f\left(B_{\xi_l^{(m)}}\right)\right\|_\Phi^{\beta_l^{(m)}}. \tag{38}$$

Since Lemma 2 implies that

$$\Phi\left(\prod_{l=1}^{m} f\left(\boldsymbol{\lambda}\left(B_{\xi_l^{(m)}}\right)\right)^{\beta_l^{(m)}}\right) \leq \prod_{l=1}^{m} \Phi^{\beta_l^{(m)}}\left(f\left(\boldsymbol{\lambda}\left(B_{\xi_l^{(m)}}\right)\right)\right) \tag{39}$$

$$= \prod_{l=1}^{m} \Phi^{\beta_l^{(m)}}\left(\boldsymbol{\lambda}\left(f\left(B_{\xi_l^{(m)}}\right)\right)\right) \tag{40}$$

$$= \prod_{l=1}^{m} \left\|f\left(B_{\xi_l^{(m)}}\right)\right\|_\Phi^{\beta_l^{(m)}}, \tag{41}$$

we have

$$\Phi\left(\exp \int_\Xi \log f(\boldsymbol{\lambda}(B_\xi)) \, d\nu(\xi)\right) \leq \exp \int_\Xi \log \|f(B_\xi)\|_\Phi \, d\nu(\xi). \tag{42}$$

Combining (34) and (42) gives inequality (27).

Next, consider the general case where $A, B_\xi \in \mathcal{P}(\mathcal{H})$. For any $\varepsilon > 0$, since

$$\prod_{i=1}^{k} \lambda_i(A) \leq \prod_{i=1}^{k} \exp \int_\Xi \log \lambda_i(B_\xi) \, d\nu(\xi) \tag{43}$$

$$< \prod_{i=1}^{k} \exp \int_\Xi \log(\lambda_i(B_\xi) + \varepsilon) \, d\nu(\xi), \qquad k \in [d], \tag{44}$$

one can choose a $\delta_\varepsilon \in (0, \varepsilon)$ such that

$$\prod_{i=1}^{k} (\lambda_i(A) + \delta_\varepsilon) \leq \prod_{i=1}^{k} \exp \int_\Xi \log(\lambda_i(B_\xi) + \varepsilon) \, d\nu(\xi), \qquad k \in [d], \tag{45}$$

i.e., we have the weak log-majorization $\boldsymbol{\lambda}(A + \delta_\varepsilon I) \prec_{w\log} \exp \int_\Xi \log \boldsymbol{\lambda}(B_\xi + \varepsilon I) \, d\nu(\xi)$. By applying the first case to $A + \delta_\varepsilon I$ and $B_\xi + \varepsilon I$, we have

$$\|f(A + \delta_\varepsilon I)\|_\Phi \leq \exp \int_\Xi \log \|f(B_\xi + \varepsilon I)\|_\Phi \, d\nu(\xi). \tag{46}$$

Since $\log \|f(B_\xi + \varepsilon I)\|_\Phi \searrow \log \|f(B_\xi)\|_\Phi$ for every $\xi \in \Xi$ as $\varepsilon \searrow 0$, the monotone convergence theorem gives

$$\int_\Xi \log \|f(B_\xi + \varepsilon I)\|_\Phi \, d\nu(\xi) \searrow \int_\Xi \log \|f(B_\xi)\|_\Phi \, d\nu(\xi). \tag{47}$$

Hence letting $\varepsilon \searrow 0$ in (46) gives the desired inequality. $\qquad\square$

Proof of (ii) $\implies$ (i). First, assume that $B_\xi \in \mathcal{P}_+(\mathcal{H})$ for all $\xi \in \Xi$. Assume (ii), and for every $k \in [d]$ we prove that

$$\prod_{i=1}^k \lambda_i(A) \le \prod_{i=1}^k \exp \int_\Xi \log \lambda_i(B_\xi) \, d\nu(\xi). \tag{48}$$

Since (48) is obvious if $\lambda_k(A) = 0$, we may assume that $\lambda_k(A) > 0$. Applying inequality (27) in (ii) to $\|\cdot\| = \|\cdot\|_{(k)}$, and $f(x) = x^p$ for each $p > 0$ (which obviously satisfies the condition in (ii)), we have

$$\|A^p\|_{(k)} \le \exp \int_\Xi \log \|B_\xi^p\|_{(k)} \, d\nu(\xi), \tag{49}$$

i.e.,

$$\log \sum_{i=1}^k \lambda_i^p(A) \le \int_\Xi \log \sum_{i=1}^k \lambda_i^p(B_\xi) \, d\nu(\xi). \tag{50}$$

Therefore,

$$\frac{1}{p} \log \left(\frac{1}{k} \sum_{i=1}^k \lambda_i^p(A) \right) \le \int_\Xi \frac{1}{p} \log \left(\frac{1}{k} \sum_{i=1}^k \lambda_i^p(B_\xi) \right) d\nu(\xi). \tag{51}$$

Since, for $a_i > 0$, the function $p > 0 \mapsto \log\left(\frac{1}{k}\sum_{i=1}^k a_i^p\right)$ is convex, we find that, as $p \searrow 0$,

$$\frac{1}{p} \log \left(\frac{1}{k} \sum_{i=1}^k \lambda_i^p(A) \right) \searrow \frac{1}{k} \sum_{i=1}^k \log \lambda_i(A), \tag{52}$$

$$\frac{1}{p} \log \left(\frac{1}{k} \sum_{i=1}^k \lambda_i^p(B_\xi) \right) \searrow \frac{1}{k} \sum_{i=1}^k \log \lambda_i(B_\xi) \tag{53}$$

for every $\xi \in \Xi$. This relies on the fact that $x \mapsto f(x)/x$ is non-decreasing if f is convex and $f(0) = 0$. Hence the monotone convergence theorem yields

$$\int_\Xi \frac{1}{p} \log \left(\frac{1}{k} \sum_{i=1}^k \lambda_i^p(B_\xi) \right) d\nu(\xi) \searrow \frac{1}{k} \int_\Xi \sum_{i=1}^k \log \lambda_i(B_\xi) \, d\nu(\xi) \qquad \text{as } p \searrow 0. \tag{54}$$

Therefore, by letting $p \searrow 0$ in (51) we have

$$\sum_{i=1}^k \log \lambda_i(A) \le \int_\Xi \sum_{i=1}^k \log \lambda_i(B_\xi) \, d\nu(\xi), \tag{55}$$

implying (48).

Next, consider the general case where $B_\xi \in \mathcal{P}(\mathcal{H})$ for all $\xi \in \Xi$. Since the inequality (27) in (ii) holds with $B_\xi + \varepsilon I$ instead of B_ξ for any $\varepsilon > 0$, the above case implies that

$$\prod_{i=1}^k \lambda_i(A) \le \prod_{i=1}^k \exp \int_\Xi \log(\lambda_i(B_\xi) + \varepsilon) \, d\nu(\xi), \qquad k \in [d]. \tag{56}$$

Letting $\varepsilon \searrow 0$ gives (48) so that (i) follows. $\qquad\square$

Proof of (i) $\implies$ (iii). Assume first that $A, B_\xi \in \mathcal{P}_+(\mathcal{H})$ and $B_\xi \ge \varepsilon I$ for all $\xi \in \Xi$ with some $\varepsilon > 0$. Since (i) means that

$$\boldsymbol{\lambda}(\log A) = \log \boldsymbol{\lambda}(A) \prec_w \int_\Xi \log \boldsymbol{\lambda}(B_\xi) \, d\nu(\xi) = \int_\Xi \boldsymbol{\lambda}(\log B_\xi) \, d\nu(\xi), \tag{57}$$

one can apply (a) $\implies$ (b) of Theorem 4 to $\log A$, $\log B_\xi$ and $f(x) := g(e^x)$, where g is as in (iii). Inequality (28) then immediately follows. For the general case where $A, B_\xi \in \mathcal{P}(\mathcal{H})$, for any $\varepsilon > 0$ choose a $\delta_\varepsilon \in (0, \varepsilon)$ satisfying (45). Since the above case gives $\|g(A + \delta_\varepsilon I)\| \le \int_\Xi \|g(B_\xi + \varepsilon I)\| \, d\nu(\xi)$, we have (28) by letting $\varepsilon \searrow 0$.

$\qquad\square$

Proof of (iii) $\implies$ (i). For $k \in [d]$ let $\|\!\cdot\!\| = \|\!\cdot\!\|_{(k)}$ and $g(x) := \log(1 + \varepsilon^{-1} x)$ where $\varepsilon > 0$; then g satisfies the condition in (iii). Since

$$\|g(A)\|_{(k)} = \sum_{i=1}^k \log(\varepsilon + \lambda_i(A)) - k \log \varepsilon, \tag{58}$$

$$\int_\Xi \|g(B_\xi)\|_{(k)} \, d\nu(\xi) = \sum_{i=1}^k \int_\Xi \log(\varepsilon + \lambda_i(B_\xi)) \, d\nu(\xi) - k \log \varepsilon, \tag{59}$$

inequality (28) implies that

$$\sum_{i=1}^k \log(\varepsilon + \lambda_i(A)) \le \sum_{i=1}^k \int_\Xi \log(\varepsilon + \lambda_i(B_\xi)) \, d\nu(\xi). \tag{60}$$

Letting $\varepsilon \searrow 0$ gives

$$\sum_{i=1}^k \log \lambda_i(A) \le \sum_{i=1}^k \int_\Xi \log \lambda_i(B_\xi) \, d\nu(\xi), \tag{61}$$

and hence (i) follows. $\qquad\square$

5. Log-majorization with integral average

Consider the strongest condition (1) in the chain of implications in Proposition 6. Our first main result concerning this condition is the following.

Theorem 10. *With Ξ, ν, A and B_ξ given as above, the following statements are equivalent:*

(I) $\boldsymbol{\lambda}(A) \prec_{\log} \exp \int_\Xi \log \boldsymbol{\lambda}(B_\xi) \, d\nu(\xi)$, i.e., $\log \boldsymbol{\lambda}(A) \prec \int_\Xi \log \boldsymbol{\lambda}(B_\xi) \, d\nu(\xi)$;

(II) for every continuous function $f : (0, \infty) \to [0, \infty)$ such that $x \mapsto \log f(e^x)$ is convex on $\mathbb{R}$, and for every unitarily invariant norm $\|\cdot\|$,

$$\|f(A)\| \leq \exp \int_\Xi \log \|f(B_\xi)\| \, d\nu(\xi) \,. \tag{62}$$

In this statement, we extend f to $[0, \infty)$ by continuity and for any unitarily invariant norm $\|\cdot\|$ use the convention $\|f(A)\| = \infty$ when $f(0^+) = \infty$ and $A \in \mathcal{P}(\mathcal{H})$ is not invertible.

The proof requires a few auxiliary results. We first show that the right-hand side of (62) is well-defined.

Lemma 11. *Let $f : (0, \infty) \to [0, \infty)$ be a continuous function such that $x \mapsto \log f(e^x)$ is convex on $\mathbb{R}$. Then (with the extension and convention as stated in Theorem 10) $\int_\Xi \log \|f(B_\xi)\| \, d\nu(\xi)$ exists in $[-\infty, \infty]$.*

Proof. If $f(0^+) < \infty$, then by Lemma 9 (2) (and the uniform boundedness of B_ξ) we have $\sup_\xi \|f(B_\xi)\| < \infty$, so the integral $\int_\Xi \log \|f(B_\xi)\| \, d\nu(\xi)$ exists in $[-\infty, \infty)$. If $f(0^+) = \infty$, then one can choose $a > 0$ and $b \in \mathbb{R}$ such that $\log f(e^x) \geq -ax + b$ on $\mathbb{R}$ by the convexity assumption. This in turn implies that $f(x) \geq e^b x^{-a}$ on $(0, \infty)$. Hence we have $\inf_\xi \|f(B_\xi)\| > 0$, so the integral exists in $(-\infty, \infty]$. $\qquad\square$

Before proving the theorem we give another lemma.

Lemma 12. *Let $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^d_+$ be such that $a_1 \geq \cdots \geq a_d$ and $b_1 \geq \cdots \geq b_d$, and assume that $\boldsymbol{a} \prec_{\log} \boldsymbol{b}$. Furthermore, let $\boldsymbol{b}^{(m)} \in \mathbb{R}^d_+$, $m \in \mathbb{N}$, be such that $b_1^{(m)} \geq \cdots \geq b_d^{(m)} > 0$ and $\boldsymbol{b}^{(m)} \searrow \boldsymbol{b}$ as $m \to \infty$. Then there exist $m_0 \in \mathbb{N}$ and $\boldsymbol{a}^{(m)} \in \mathbb{R}^d_+$ for $m \geq m_0$ such that $a_1^{(m)} \geq \cdots \geq a_d^{(m)} > 0$, $\boldsymbol{a}^{(m)} \to \boldsymbol{a}$ as $m \to \infty$, and*

$$\boldsymbol{a} \leq \boldsymbol{a}^{(m)} \prec_{\log} \boldsymbol{b}^{(m)}, \qquad m \geq m_0 \,.$$

Proof. Assume that $\boldsymbol{a} \prec_{\log} \boldsymbol{b}$. The proof is divided into two cases. First, assume that $a_d > 0$ (hence $b_d > 0$ as well). For each $m \in \mathbb{N}$, since $\log \boldsymbol{a}, \log \boldsymbol{b} \in \mathbb{R}^d$ and $\log \boldsymbol{a} \prec \log \boldsymbol{b} \leq \log \boldsymbol{b}^{(m)}$ so that $\log \boldsymbol{a} \prec_w \log \boldsymbol{b}^{(m)}$, it follows from [9, Proposition 4.1.3] that there exists a $\boldsymbol{c}^{(m)} \in \mathbb{R}^d$ such that

$$\log \boldsymbol{a} \leq \boldsymbol{c}^{(m)} \prec \log \boldsymbol{b}^{(m)} \,. \tag{63}$$

Now define $\boldsymbol{a}^{(m)} := \exp \boldsymbol{c}^{(m)}$; then $\boldsymbol{a} \leq \boldsymbol{a}^{(m)} \prec_{\log} \boldsymbol{b}^{(m)}$. It remains to prove that $\boldsymbol{a}^{(m)} \to \boldsymbol{a}$. For this, note that

$$\sum_{i=1}^d \bigl(c_i^{(m)} - \log a_i\bigr) = \sum_{i=1}^d \bigl(\log b_i^{(m)} - \log a_i\bigr) \longrightarrow \sum_{i=1}^d (\log b_i - \log a_i) = 0 \tag{64}$$

as $m \to \infty$. Therefore, we have $c_i^{(m)} \to \log a_i$ so that $a_i^{(m)} \to a_i$ for all $i \in [d]$.

Secondly, assume that $a_d = 0$ (hence $b_d = 0$ as well). Assume that

$$a_1 \geq \cdots \geq a_r > 0 = a_{r+1} = \cdots = a_d \,,$$
$$b_1 \geq \cdots \geq b_s > 0 = b_{s+1} = \cdots = b_d \,.$$

Since $0 < \prod_{i=1}^{r} a_i \leq \prod_{i=1}^{r} b_i$, we have $r \leq s$. For each $m \in \mathbb{N}$ define

$$\alpha^{(m)} := \left(\frac{\prod_{i=1}^{s+1} b_i^{(m)}}{\prod_{i=1}^{r} a_i} \right)^{\frac{1}{s-r+1}} > 0. \tag{65}$$

Since $b_{s+1}^{(m)} \to b_{s+1} = 0$ and so $\alpha^{(m)} \to 0$ as $m \to \infty$, choose an m_0 such that $\alpha^{(m)} \leq \min\{a_r, b_s\}$ for all $m \geq m_0$. Define for $m \geq m_0$,

$$\boldsymbol{a}^{(m)} := \left(a_1, \ldots, a_r, \underbrace{\alpha^{(m)}, \ldots, \alpha^{(m)}}_{s-r+1}, b_{s+2}^{(m)}, \ldots, b_d^{(m)} \right). \tag{66}$$

Since

$$\alpha^{(m)} = \frac{\prod_{i=1}^{s+1} b_i^{(m)}}{\prod_{i=1}^{r} a_i \cdot \left(\alpha^{(m)} \right)^{s-r}} \geq \frac{\prod_{i=1}^{r} b_i}{\prod_{i=1}^{r} a_i} \left(\frac{b_s}{\alpha^{(m)}} \right)^{s-r} b_{s+1}^{(m)} \geq b_{s+1}^{(m)}, \tag{67}$$

we find that $\boldsymbol{a}^{(m)}$ is in decreasing order. We furthermore have $\boldsymbol{a}^{(m)} \to \boldsymbol{a}$ and

$$\prod_{i=1}^{r} a_i \cdot (\alpha^{(m)})^k \leq \prod_{i=1}^{r} b_i \cdot b_s^k \leq \prod_{i=1}^{r+k} b_i^{(m)}, \qquad 1 \leq k \leq s-r, \tag{68}$$

$$\prod_{i=1}^{r} a_i \cdot (\alpha^{(m)})^{s-r+1} = \prod_{i=1}^{s+1} b_i^{(m)}, \tag{69}$$

so that $\boldsymbol{a} \leq \boldsymbol{a}^{(m)} \prec_{\log} \boldsymbol{b}^{(m)}$ follows. $\qquad\square$

Proof of (I) $\implies$ (II). First, assume that $A, B_\xi \in \mathcal{P}_+(\mathcal{H})$ and $B_\xi \geq \varepsilon I$ for all $\xi \in \Xi$ with some $\varepsilon > 0$. Since $\boldsymbol{\lambda}(f(A)) \approx f(\boldsymbol{\lambda}(A))$, the corresponding part of the proof of Theorem 7 can be adopted with the slight modification that

$$\log \boldsymbol{\lambda}(f(A)) \approx \log f(e^{\log \boldsymbol{\lambda}(A)}) \tag{70}$$

instead of (30) because the assumption that f is non-decreasing has been dropped.

Next, consider the general case where $A, B_\xi \in \mathcal{P}(\mathcal{H})$. With $0 < \varepsilon_m \searrow 0$, we have

$$\prod_{i=1}^{k} \lambda_i(A) \leq \prod_{i=1}^{k} \exp \int_\Xi \log \lambda_i(B_\xi) \, d\nu(\xi) \tag{71}$$

$$< \prod_{i=1}^{k} \exp \int_\Xi \log(\lambda_i(B_\xi) + \varepsilon_m) \, d\nu(\xi), \qquad k \in [d]. \tag{72}$$

Since $\int_\Xi \log(\boldsymbol{\lambda}(B_\xi) + \varepsilon_m) \, d\nu(\xi) \searrow \int_\Xi \log \boldsymbol{\lambda}(B_\xi) \, d\nu(\xi)$ as $m \to \infty$ by the monotone convergence theorem, by Lemma 12 one can choose $\boldsymbol{a}^{(m)}$, $m \geq m_0$, such that $a_1^{(m)} \geq \cdots \geq a_d^{(m)} > 0$, $\boldsymbol{a}^{(m)} \to \boldsymbol{\lambda}(A)$ and

$$\boldsymbol{a}^{(m)} \prec_{\log} \exp \int_\Xi \log(\boldsymbol{\lambda}(B_\xi + \varepsilon_m I)) \, d\nu(\xi). \tag{73}$$

Choosing $A^{(m)} \in \mathcal{P}_+(\mathcal{H})$ with $\boldsymbol{\lambda}(A^{(m)}) = \boldsymbol{a}^{(m)}$ and applying the first case to $A^{(m)}$ and $B_\xi + \varepsilon_m I$, we have

$$\|f(A^{(m)})\|_\Phi \leq \exp \int_\Xi \log \|f(B_\xi + \varepsilon_m I)\|_\Phi \, \mathrm{d}\nu(\xi), \qquad m \geq m_0. \qquad (74)$$

When $f(0^+) < \infty$ and hence f is non-decreasing on $(0, \infty)$ by Lemma 9 (2), note that

$$\|f(A^{(m)})\|_\Phi = \Phi(f(\boldsymbol{a}^{(m)})) \longrightarrow \Phi(f(\boldsymbol{\lambda}(A))) = \|f(A)\|_\Phi \qquad (75)$$

and similarly $\|f(B_\xi + \varepsilon_m I)\|_\Phi \to \|f(B_\xi)\|_\Phi$ for every $\xi \in \Xi$ as $m \to \infty$. Since $\xi \mapsto \|f(B_\xi + \varepsilon_m I)\|_\Phi$ is uniformly bounded above (so $-\log \|f(B_\xi + \varepsilon_m I)\|_\Phi$ is uniformly bounded below), Fatou's lemma yields

$$\limsup_{m\to\infty} \int_\Xi \log \|f(B_\xi + \varepsilon_m I)\|_\Phi \, \mathrm{d}\nu(\xi) \leq \int_\Xi \log \|f(B_\xi)\|_\Phi \, \mathrm{d}\nu(\xi), \qquad (76)$$

and therefore, letting $m \to \infty$ in (74) gives inequality (62). Finally, when $f(0^+) = \infty$, we may assume that $\int_\Xi \log \|f(B_\xi)\|_\Phi \, \mathrm{d}\nu(\xi) < \infty$. In this case, f is decreasing on $(0, \delta)$ for some $\delta > 0$. We will argue below that there are constants $\alpha, \beta > 0$ such that

$$\alpha \leq \|f(B_\xi + \varepsilon_m I)\|_\Phi \leq \|f(B_\xi)\|_\Phi + \beta \qquad (77)$$

for all $\xi \in \Xi$ and $m \geq m_0$. Since $\int_\Xi \log(\|f(B_\xi)\|_\Phi + \beta) \, \mathrm{d}\nu(\xi) < \infty$, the Lebesgue convergence theorem can be used to get (62) by taking the limit of (74).

It remains to show (77). By the uniform boundedness of the operators B_ξ, there is a constant $\gamma > 0$ such that

$$0 < B_\xi + \varepsilon_m I \leq \gamma I, \qquad \xi \in \Xi, m \geq m_0 . \qquad (78)$$

Because f is decreasing on $(0, \delta)$ and $f(x) > 0$ for all $x > 0$ (see Lemma 9 (1)), this implies that $\xi \mapsto \|f(B_\xi + \varepsilon_m I)\|_\Phi$ is uniformly bounded from below, as claimed in (77). Observe that the upper bound in (77) is trivial for $B_\xi \in \mathcal{P}(\mathcal{H}) \backslash \mathcal{P}_+(\mathcal{H})$ since $\|f(B_\xi)\|_\Phi = \infty$ when B_ξ is not invertible. Hence assume that $B_\xi \in \mathcal{P}_+(\mathcal{H})$. Using the spectral decomposition $B_\xi = \sum_{\lambda\in\mathsf{spec}(B_\xi)} \lambda P_\lambda$, where $\mathsf{spec}(B_\xi)$ is the set of eigenvalues of B_ξ, we then have

$$f(B_\xi + \varepsilon_m I) = \sum_{\substack{\lambda\in\mathsf{spec}(B_\xi) \\ \lambda+\varepsilon_m<\delta}} f(\lambda + \varepsilon_m)P_\lambda + \sum_{\substack{\lambda\in\mathsf{spec}(B_\xi) \\ \lambda+\varepsilon_m\geq\delta}} f(\lambda + \varepsilon_m)P_\lambda \qquad (79)$$

$$\leq \sum_{\substack{\lambda\in\mathsf{spec}(B_\xi) \\ \lambda+\varepsilon_m<\delta}} f(\lambda)P_\lambda + \sum_{\substack{\lambda\in\mathsf{spec}(B_\xi) \\ \lambda+\varepsilon_m\geq\delta}} f(\lambda + \varepsilon_m)P_\lambda \qquad (80)$$

$$\leq f(B_\xi) + \sum_{\substack{\lambda\in\mathsf{spec}(B_\xi) \\ \lambda+\varepsilon_m\geq\delta}} f(\lambda + \varepsilon_m)P_\lambda \qquad (81)$$

The claim then follows by the triangle inequality for $\|\cdot\|_\Phi$ and the fact that $f(\lambda + \varepsilon_m) \leq \sup_{\delta\leq x\leq\gamma} f(x) < \infty$ for all $\lambda \in \mathsf{spec}(B_\xi)$ with $\lambda + \varepsilon_m \geq \delta$ and for all $\xi \in \Xi$. The last fact is immediately seen from (78) and the continuity of f.

□

Proof of (II) $\implies$ (I). The weak majorization relation

$$\prod_{i=1}^{k} \lambda_i(A) \leq \prod_{i=1}^{k} \exp \int_{\Xi} \log \lambda_i(B_\xi) \, d\nu(\xi) \ . \tag{82}$$

is obvious from (ii) $\implies$ (i) in Theorem 7 since condition (II) is stronger than (ii). It remains to prove that equality holds in (82) when $k = d$. It suffices to prove that

$$\log(\det A) \geq \int_{\Xi} \log\bigl(\det B_\xi\bigr) \, d\nu(\xi) \, . \tag{83}$$

For this, we may assume that $\int_{\Xi} \log\bigl(\det B_\xi\bigr) \, d\nu(\xi) > -\infty$ and so $B_\xi \in \mathcal{P}_+(\mathcal{H})$ for ν-a.e. $\xi \in \Xi$. So we may assume that $B_\xi \in \mathcal{P}_+(\mathcal{H})$ for all $\xi \in \Xi$. Moreover, replacing A, B_ξ with αA, αB_ξ for some $\alpha > 0$, we may assume that $B_\xi \leq I$ and so $\lambda_i(B_\xi) \leq 1$ for all $\xi \in \Xi$ and $i \in [d]$. For every $p > 0$, since

$$\frac{1}{d}\bigl\|B_\xi^{-p}\bigr\|_1 \leq \lambda_d(B_\xi)^{-p} \leq \bigl(\det B_\xi\bigr)^{-p} \, , \tag{84}$$

we find that

$$\frac{1}{p} \log\Bigl(\frac{1}{d}\bigl\|B_\xi^{-p}\bigr\|_1\Bigr) \leq -\log\bigl(\det B_\xi\bigr) \, . \tag{85}$$

Applying inequality (62) to $\|\cdot\| = \|\cdot\|_1$ and $f(x) = x^{-p}$ for any $p > 0$ we get

$$\log \|A^{-p}\|_1 \leq \int_{\Xi} \log \bigl\|B_\xi^{-p}\bigr\|_1 \, d\nu(\xi) \, . \tag{86}$$

This means that

$$\frac{1}{p} \log\Bigl(\frac{1}{d}\|A^{-p}\|_1\Bigr) \leq \int_{\Xi} \frac{1}{p} \log\Bigl(\frac{1}{d}\bigl\|B_\xi^{-p}\bigr\|_1\Bigr) \, d\nu(\xi) \, . \tag{87}$$

Similarly to (52) and (53) we find that, as $p \searrow 0$,

$$\frac{1}{p} \log\Bigl(\frac{1}{d}\|A^{-p}\|_1\Bigr) \searrow -\frac{1}{d} \log(\det A) \, , \tag{88}$$

$$\frac{1}{p} \log\Bigl(\frac{1}{d}\bigl\|B_\xi^{-p}\bigr\|_1\Bigr) \searrow -\frac{1}{d} \log\bigl(\det B_\xi\bigr) \tag{89}$$

for every $\xi \in \Xi$. Thanks to (85) the Lebesgue convergence theorem yields

$$\lim_{p \searrow 0} \int_{\Xi} \frac{1}{p} \log\Bigl(\frac{1}{d}\bigl\|B_\xi^{-p}\bigr\|_1\Bigr) \, d\nu(\xi) = -\frac{1}{d} \int_{\Xi} \log\bigl(\det B_\xi\bigr) \, d\nu(\xi) \, . \tag{90}$$

Therefore, letting $p \searrow 0$ in (87) gives (83), as desired. $\square$

When Ξ is a one-point set, Proposition 8 and Theorem 10 with the above proof of the implication (II) $\implies$ (I) yield a new characterization of log-majorization.

Proposition 13. *For $A, B \in \mathcal{P}(\mathcal{H})$, $\boldsymbol{\lambda}(A) \prec_{\log} \boldsymbol{\lambda}(B)$ if and only if*

$$\|A^p\|_{(k)} \leq \|B^p\|_{(k)}, \qquad p \in \mathbb{R} \setminus \{0\}, \quad k \in [d]. \tag{91}$$

It is natural to wonder whether the generalized log-majorization condition (1) of Proposition 6 is equivalent to a third condition analogous to (ii) of Theorem 7. The following theorem shows that this is the case under one additional technical assumption. In the statement, we use the same convention for $\|g(A)\|$ as introduced in Theorem 10.

Theorem 14. *With Ξ, ν, A and B_ξ given as above, consider the additional statement*

(III) for every continuous function $g : (0,\infty) \to [0,\infty)$ such that $x \mapsto g(e^x)$ is convex on $\mathbb{R}$, and for every unitarily invariant norm $\|\cdot\|$,

$$\|g(A)\| \le \int_\Xi \|g(B_\xi)\| \, d\nu(\xi) \ . \tag{92}$$

Then (I) $\implies$ (III), *and if* $\int_\Xi \left\|B_\xi^{-p}\right\|_1 d\nu(\xi) < \infty$ *for some* $p > 0$, *then* (III) $\implies$ (I).

Remark 2. The integrability assumption is essential in the proof of the implication (III) $\implies$ (I) with use of test functions x^{-p} for $p > 0$. Indeed, it is easy to provide an example of Ξ, ν and B_ξ such that $-\int_\Xi \log\left(\det B_\xi\right) d\nu(\xi) < \infty$ but $\int_\Xi \left\|B_\xi^{-p}\right\|_1 d\nu(\xi) = \infty$ for all $p > 0$. Since there is no good test function other than x^{-p}, it seems difficult to remove or relax the integrability assumption.

Proof of (I) $\implies$ (III). Assume first that $A, B_\xi \in \mathcal{P}_+(\mathcal{H})$ and $B_\xi \ge \varepsilon I$ for all $\xi \in \Xi$ with some $\varepsilon > 0$. Since (I) means that $\boldsymbol{\lambda}(\log A) \prec \int_\Xi \boldsymbol{\lambda}(\log B_\xi) \, d\nu(\xi)$, one can apply (d) $\implies$ (e) of Theorem 5 to $\log A$, $\log B_\xi$ and $f(x) := g(e^x)$ for g as in (III). Inequality (III) then follows. For general $A, B_\xi \in \mathcal{P}(\mathcal{H})$ satisfying (I), with $0 < \varepsilon_m \searrow 0$ choose $\boldsymbol{a}^{(m)}$ and $A^{(m)}$, $m \ge m_0$, as in the proof of (I) $\implies$ (II). By the first case we have

$$\|g(A^{(m)})\|_\Phi \le \int_\Xi \|g(B_\xi + \varepsilon_m I)\|_\Phi \, d\nu(\xi). \tag{93}$$

When $g(0^+) < \infty$, letting $m \to \infty$ in (93) gives inequality (92) immediately. When $g(0^+) = \infty$, the proof is similar to the last part of the proof (I) $\implies$ (II) by noting that there is a constant $\beta > 0$ such that $\|g(B_\xi + \varepsilon_m I)\|_\Phi \le \|g(B_\xi)\|_\Phi + \beta$ for all $\xi \in \Xi$ and $m \ge m_0$. $\qquad\square$

Proof of (III) $\implies$ (I) *under the integrability assumption.* The weak majorization relation

$$\sum_{i=1}^k \log \lambda_i(A) \le \sum_{i=1}^k \int_\Xi \log \lambda_i(B_\xi) \, d\nu(\xi) \ , \qquad k \in [d] \tag{94}$$

is obvious from (iii) $\implies$ (i) in Theorem 7 since condition (III) is stronger than (iii). It remains to prove that equality holds in (94) when $k = d$. Here, we use the assumption that $\int_\Xi \left\|B_\xi^{-p_0}\right\|_1 d\nu(\xi) < \infty$ for some $p_0 > 0$. Inequality (92) in (III) is applied to $\|\cdot\| = \|\cdot\|_1$ and $g(x) = x^{-p}$ for any $p > 0$, so that we have $\|A^{-p}\|_1 \le \int_\Xi \left\|B_\xi^{-p}\right\|_1 d\nu(\xi)$. Therefore,

$$\frac{1}{p} \log\left(\frac{1}{d}\|A^{-p}\|_1\right) \le \frac{1}{p} \log \int_\Xi \frac{1}{d}\left\|B_\xi^{-p}\right\|_1 d\nu(\xi). \tag{95}$$

Thanks to Lemma 15 separately shown below (and (88) as well), letting $p \searrow 0$ in (94) yields

$$-\frac{1}{d}\log(\det A) \leq -\frac{1}{d}\int_\Xi \log\big(\det B_\xi\big)\,\mathrm{d}\nu(\xi)\,, \qquad (96)$$

which gives the desired equality.

$\square$

Lemma 15. *Let* Ξ, ν *and* B_ξ *be as above. If* $\int_\Xi \big\|B_\xi^{-p_0}\big\|_1\,\mathrm{d}\nu(\xi) < \infty$ *for some* $p_0 > 0$, *then*

$$\lim_{p\searrow 0}\left(\frac{1}{p}\log\int_\Xi \frac{1}{d}\big\|B_\xi^{-p}\big\|_1\,\mathrm{d}\nu(\xi)\right) = -\frac{1}{d}\int_\Xi \log\big(\det B_\xi\big)\,\mathrm{d}\nu(\xi)\,. \qquad (97)$$

Proof. The following proof is similar to that of [6, Lemma 6.12]. The assumption implies that $B_\xi \in \mathcal{P}_+(\mathcal{H})$ for ν-a.e. $\xi \in \Xi$. So we may assume that $B_\xi \in \mathcal{P}_+(\mathcal{H})$ for all $\xi \in \Xi$. Moreover, replacing B_ξ with αB_ξ for some $\alpha > 0$, we may assume that $B_\xi \leq I$. Let $\widetilde{\nu} := \nu \otimes \mu$ be the product measure of ν and the uniform probability measure μ on $[d]$. Define

$$\phi(\xi, i, p) := \lambda_i(B_\xi)^{-p}, \qquad \xi \in \Xi, \quad i \in [d], \quad p > 0. \qquad (98)$$

It is clear that

$$\int_\Xi \frac{1}{d}\big\|B_\xi^{-p}\big\|_1\,\mathrm{d}\nu(\xi) = \int_{\Xi\times[d]} \phi(\xi, i, p)\,\mathrm{d}\widetilde{\nu}(\xi, i)\,, \qquad p > 0. \qquad (99)$$

Hence $(\xi, i) \mapsto \phi(\xi, i, p_0)$ is integrable with respect to $\widetilde{\nu}$. According to the mean value theorem applied to the function $p \mapsto \lambda_i(B_\xi)^p$, we have

$$\frac{\phi(\xi, i, p) - \phi(\xi, i, 0)}{p} = -\lambda_i(B_\xi)^{-\theta p}\log\lambda_i(B_\xi) \leq -\lambda_i(B_\xi)^{-p}\log\lambda_i(B_\xi)\,, \qquad (100)$$

for some $\theta \in (0, 1)$ depending on ξ, i, p, and

$$\lim_{p\searrow 0}\frac{\phi(\xi, i, p) - \phi(\xi, i, 0)}{p} = -\log\lambda_i(B_\xi)\,. \qquad (101)$$

Furthermore, when $0 < p < p_1 < p_0$, we have

$$-\lambda_i(B_\xi)^{-p}\log\lambda_i(B_\xi) \leq -\lambda_i(B_\xi)^{-p_1}\log\lambda_i(B_\xi) \qquad (102)$$

$$= \lambda_i(B_\xi)^{-p_0}\big\{-\lambda_i(B_\xi)^{p_0-p_1}\log\lambda_i(B_\xi)\big\}\,. \qquad (103)$$

Since $\sup_{0<\lambda\leq 1}\big(-\lambda^{p_0-p_1}\log\lambda\big) < \infty$, we find that

$$(\xi, i) \mapsto -\lambda_i(B_\xi)^{-p_1}\log\lambda_i(B_\xi)$$

is integrable with respect to $\widetilde{\nu}$. Hence the Lebesgue convergence theorem yields

$$\frac{\mathrm{d}}{\mathrm{d}p} \int_{\Xi \times [d]} \phi(\xi, i, p) \, \mathrm{d}\widetilde{\nu}(\xi, i) \Big|_{p=0^+} = \lim_{p \searrow 0} \int_{\Xi \times [d]} \frac{\phi(\xi, i, p) - \phi(\xi, i, 0)}{p} \, \mathrm{d}\widetilde{\nu}(\xi, i) \tag{104}$$

$$= -\int_{\Xi \times [d]} \log \lambda_i(B_\xi) \, \mathrm{d}\widetilde{\nu}(\xi, i) \tag{105}$$

$$= -\int_{\Xi} \frac{1}{d} \sum_{i=1}^{d} \log \lambda_i(B_\xi) \, \mathrm{d}\nu(\xi) \tag{106}$$

$$= -\frac{1}{d} \int_{\Xi} \log\big(\det B_\xi\big) \, \mathrm{d}\nu(\xi) \,, \tag{107}$$

where $\frac{\mathrm{d}}{\mathrm{d}p}(\cdot)\big|_{p=0^+}$ means the right derivative at $p = 0$. Now we obtain the desired equality since

$$\lim_{p \searrow 0} \left(\frac{1}{p} \log \int_{\Xi} \frac{1}{d} \big\| B_\xi^{-p} \big\|_1 \, \mathrm{d}\nu(\xi) \right) = \frac{\mathrm{d}}{\mathrm{d}p} \int_{\Xi \times [d]} \phi(\xi, i, p) \, \mathrm{d}\widetilde{\nu}(\xi, i) \Big|_{p=0^+} \tag{108}$$

as easily seen from (99). $\qquad\square$

6. Application to multivariate norm inequalities

We recall the inequality [16, Thm. 3.2] specialized to the operator norm. For $A_\ell \in \mathcal{P}(\mathcal{H})$, $\ell \in [n]$ and $\theta \in (0, 1]$, we have

$$\log \left\| \left| \prod_{\ell=1}^{n} A_\ell^\theta \right|^{\frac{1}{\theta}} \right\| \leq \int_{-\infty}^{\infty} \log \left\| \prod_{\ell=1}^{n} A_\ell^{1+it} \right\| \, \mathrm{d}\beta_\theta(t) \,, \tag{109}$$

where

$$\mathrm{d}\beta_\theta(t) := \frac{\sin(\pi\theta)}{2\theta\big(\cos(\pi t) + \cos(\pi\theta)\big)} \, \mathrm{d}t \,, \tag{110}$$

and the functional calculus A_ℓ^z for any $z \in \mathbb{C}$ is defined with the convention that $0^z = 0$. A concise proof of this special case is given in Appendix A. Using the rules of antisymmetric tensor power calculus presented in Lemma 3, we find

$$\left| \prod_{\ell=1}^{n} (\wedge^k A_\ell)^\theta \right|^{\frac{1}{\theta}} = \wedge^k \left| \prod_{\ell=1}^{n} A_\ell^\theta \right|^{\frac{1}{\theta}} \quad \text{and} \quad \left| \prod_{\ell=1}^{n} \left(\wedge^k A_\ell \right)^{1+it} \right| = \wedge^k \left| \prod_{\ell=1}^{n} A_\ell^{1+it} \right| . \tag{111}$$

The inequality (109) applied to the matrices $\wedge^k A_\ell$ for all $k \in [d]$ thus immediately yields the log-majorization relation

$$\log \boldsymbol{\lambda} \left(\left| \prod_{\ell=1}^{n} A_\ell^\theta \right|^{\frac{1}{\theta}} \right) \prec \int_{-\infty}^{\infty} \log \boldsymbol{\lambda} \left(\left| \prod_{\ell=1}^{n} A_\ell^{1+it} \right| \right) \mathrm{d}\beta_\theta(t) \,, \tag{112}$$

where in particular the equality condition for log-majorization is satisfied since

$$\det \left| \prod_{\ell=1}^{n} A_\ell^\theta \right|^{\frac{1}{\theta}} = \det \left| \prod_{\ell=1}^{n} A_\ell^{1+it} \right| = \prod_{\ell=1}^{n} \det A_\ell \,. \tag{113}$$

Hence we arrive at the following application of Theorem 10 and Theorem 14. Here we again use the continuous extension and convention of Theorem 10.

Corollary 16. *Let $A_\ell \in \mathcal{P}(\mathcal{H})$ for $\ell \in [n]$, $\theta \in (0,1]$ and $\|\cdot\|$ a unitarily invariant norm. Then, for any continuous function $f : (0,\infty) \to [0,\infty)$ such that $x \mapsto \log f(e^x)$ is convex on $\mathbb{R}$, we have*

$$\log \left\| f\left(\left| \prod_{\ell=1}^{n} A_\ell^\theta \right|^{\frac{1}{\theta}} \right) \right\| \le \int_{-\infty}^{\infty} \log \left\| f\left(\left| \prod_{\ell=1}^{n} A_\ell^{1+it} \right| \right) \right\| \mathrm{d}\beta_\theta(t) \,. \tag{114}$$

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

$$\left\| g\left(\left| \prod_{\ell=1}^{n} A_\ell^\theta \right|^{\frac{1}{\theta}} \right) \right\| \le \int_{-\infty}^{\infty} \left\| g\left(\left| \prod_{\ell=1}^{n} A_\ell^{1+it} \right| \right) \right\| \mathrm{d}\beta_\theta(t) \,. \tag{115}$$

These inequalities generalize and strengthen the results in [16]. For example, consider the function $f : x \mapsto x^q$ for $q \in \mathbb{R} \setminus \{0\}$ and the trace norm to find

Corollary 17. *Let $A_\ell \in \mathcal{P}(\mathcal{H})$ for $\ell \in [n]$ and $\theta \in (0,1]$. Then we have*

$$\log \operatorname{tr} \left| \prod_{\ell=1}^{n} A_\ell^\theta \right|^{\frac{q}{\theta}} \le \int_{-\infty}^{\infty} \log \operatorname{tr} \left| \prod_{\ell=1}^{n} A_\ell^{1+it} \right|^q \mathrm{d}\beta_\theta(t) \,, \tag{116}$$

for any $q \in \mathbb{R}\setminus\{0\}$.

Indeed, this is a strengthening of both [16, Thm. 3.2] (which establishes (116) for $q \ge 1$) and [16, Thm. 2.3] (which establishes a looser bound for $q > 0$ where the integration on the right-hand side of (116) is replaced by a supremum over t) to the case of arbitrary non-zero $q \in \mathbb{R}$.

Finally note that if $A_\ell \in \mathcal{P}_+(\mathcal{H})$, all of these inequalities remain valid in the limit $\theta \to 0$, where the Lie-Trotter product formula asserts that

$$\left| \prod_{\ell=1}^{n} A_\ell^\theta \right|^{\frac{1}{\theta}} \longrightarrow \exp\left(\sum_{\ell=1}^{n} \log A_\ell \right) \,. \tag{117}$$

Equations (114) and (115) thus hold with this substitution and $\theta = 0$.

Corollary 18. *Let $A_\ell \in \mathcal{P}_+(\mathcal{H})$ for $\ell \in [n]$. With $\|\cdot\|$, f and g given as in Corollary 16,*

$$\log \left\| f\left(\exp\left(\sum_{\ell=1}^{n} \log A_\ell \right) \right) \right\| \le \int_{-\infty}^{\infty} \log \left\| f\left(\left| \prod_{\ell=1}^{n} A_\ell^{1+it} \right| \right) \right\| \mathrm{d}\beta_0(t) \,, \tag{118}$$

$$\left\| g\left(\exp\left(\sum_{\ell=1}^{n} \log A_\ell \right) \right) \right\| \le \int_{-\infty}^{\infty} \left\| g\left(\left| \prod_{\ell=1}^{n} A_\ell^{1+it} \right| \right) \right\| \mathrm{d}\beta_0(t) \,. \tag{119}$$

These inequalities generalize [16, Cor. 3.3], where the result was shown for the norms $\| \cdot \|_p$ with $p \geq 1$ and f and g equal to the identity function. Using this inequality with $n = 4$ and $p = 2$, the authors of [16] obtained the best currently known lower bound on the remainder term in the strong subadditivity inequality involving the universal rotated Petz recovery map introduced in [12]. (The first such remainder terms involving recovery maps were recently presented in [7].) It remains an open problem whether this application to quantum information can be extended using the strengthened inequalities obtained here.

APPENDIX A. A SHORT PROOF OF (109)

Hirschman's strengthening of Hadamard's three line theorem [11] reads:

Lemma 19 (Hirschman). *Let $S := \{z \in \mathbb{C} : 0 \leq \mathrm{Re}(z) \leq 1\}$ and let $g(z)$ be uniformly bounded on S, holomorphic on the interior of S and continuous up to the boundary. Then for $\theta \in [0, 1]$, we have*

$$\log |g(\theta)| \leq \int_{-\infty}^{\infty} \log |g(\mathrm{i}t)|^{1-\theta} \, \mathrm{d}\beta_{1-\theta}(t) + \int_{-\infty}^{\infty} \log |g(1 + \mathrm{i}t)|^{\theta} \, \mathrm{d}\beta_{\theta}(t) . \quad (120)$$

Now let $G(z)$ be a uniformly bounded holomorphic function with values in $\mathbb{C}^{d \times d}$. Fix $\theta \in (0, 1)$ and let $u, v \in \mathbb{C}^d$ be normalized vectors such that $\langle u, G(\theta)v \rangle = \|G(\theta)\|$. Consequently, $g(z) := \langle u, G(z)v \rangle$ can be bounded as $|g(z)| \leq \|G(z)\|$ for all $z \in S$. It satisfies the assumptions of Hirschman's theorem, yielding

$$\log \|G(\theta)\| \leq \int_{-\infty}^{\infty} \log \|G(\mathrm{i}t)\|^{1-\theta} \, \mathrm{d}\beta_{1-\theta}(t) + \int_{-\infty}^{\infty} \log \|G(1 + \mathrm{i}t)\|^{\theta} \, \mathrm{d}\beta_{\theta}(t) . \quad (121)$$

As in [16, Thm. 3.2], consider now a set of n matrices $A_\ell \in \mathcal{P}(\mathcal{H})$, $\ell \in [n]$ and set $G(z) = \prod_{k=1}^{n} A_\ell^z$. Since $G(\mathrm{i}t)$ is a product of isometries, the first term in the right-hand side of (121) is non-positive and after dividing by θ we find

$$\log \left\| \left| \prod_{\ell=1}^{n} A_\ell^{\theta} \right|^{\frac{1}{\theta}} \right\| \leq \int_{-\infty}^{\infty} \log \left\| \prod_{\ell=1}^{n} A_\ell^{1+\mathrm{i}t} \right\| \, \mathrm{d}\beta_{\theta}(t) . \quad (122)$$

References

[1] Tsuyoshi Ando. Majorization, Doubly Stochastic Matrices, and Comparison of Eigenvalues. *Linear Algebra and its Applications*, 118:163–248, 1989.

[2] Tsuyoshi Ando. Majorizations and inequalities in matrix theory. *Linear Algebra and its Applications, Special Issue Honoring Ingram Olkin*, 199:17 – 67, 1994.

[3] Tsuyoshi Ando and Fumio Hiai. Log majorization and complementary Golden-Thompson type inequalities. *Linear Algebra and its Applications*, 197/198:113–131, 1994.

[4] Huzihiro Araki. On an inequality of Lieb and Thirring. *Letters in Mathematical Physics*, 19(2):167–170, 1990.

[5] Rajendra Bhatia. *Matrix Analysis*. Graduate Texts in Mathematics. Springer, 1997.

[6] Jean-Christophe Bourin and Fumio Hiai. Anti-norms on finite von Neumann algebras. *Publ. Res. Inst. Math. Sci.*, 51(2):207–235, 2015.

[7] Omar Fawzi and Renato Renner Quantum Conditional Mutual Information and Approximate Markov Chains. *Communications in Mathematical Physics* 340(2):575611, 2015.

[8] Sidney Golden. Lower Bounds for the Helmholtz Function. *Physical Review*, 137(4B):B1127–B1128, 1965.

[9] Fumio Hiai. Matrix Analysis: Matrix Monotone Functions, Matrix Means, and Majorization. *Interdisciplinary Information Sciences*, 16:139–248, 2010.

[10] Fumio Hiai and Dénes Petz. *Introduction to Matrix Analysis and Applications*. Springer, 2014.

[11] Isidore Isaac Hirschman. A Convexity Theorem for Certain Groups of Transformations. *Journal d'Analyse Mathématique*, 2(2):209–218, 1953.

[12] Marius Junge, Renato Renner, David Sutter, Mark M. Wilde, and Andreas Winter. Universal recovery from a decrease of quantum relative entropy, 2016. *arXiv:1509.07127*.

[13] Elliott H. Lieb. Convex Trace Functions and the Wigner-Yanase-Dyson Conjecture. *Advances in Mathematics*, 11(3):267–288, 1973.

[14] Elliott H. Lieb and Walter E. Thirring. *Inequalities for the Moments of the Eigenvalues of the Schrödinger Hamiltonian and Their Relation to Sobolev Inequalities*. Studies in Mathematical Physics. Princeton University Press, Berlin/Heidelberg, 1976.

[15] Albert W. Marshall, Ingram Olkin, and Barry C. Arnold. *Inequalities: Theory of Majorization and Its Applications*. Springer Series in Statistics. Springer New York, New York, NY, 2011.

[16] David Sutter, Mario Berta, and Marco Tomamichel. Multivariate trace inequalities. *Communications in Mathematical Physics* (accepted for publication, August 11, 2016).

[17] Colin J. Thompson. Inequality with Applications in Statistical Mechanics. *Journal of Mathematical Physics*, 6(11):1812, 1965.

[18] J. von Neumann. Some matrix-inequalities and metrization of matric-space. *Tomsk University Review* 1:286–300, 1937, reprinted in Collected Works, volume IV, pages 205–218.

Tohoku University (Emeritus)
E-mail address: hiai.fumio@gmail.com

Institute for Advanced Study &, Zentrum Mathematik, Technische Universität München, Garching, Germany
E-mail address: robert.koenig@tum.de

School of Physics, The University of Sydney, Sydney, Australia
E-mail address: marcotom.ch@gmail.com