Matrix limit theorems of Kato type related to positive linear maps and operator means

Fumio Hiai

1 Tohoku University (Emeritus), Hakusan 3-8-16-303, Abiko 270-1154, Japan

Abstract

We obtain limit theorems for \(\Phi(A^p)^{1/p}\) and \((A^p \sigma B)^{1/p}\) as \(p \to \infty\) for positive matrices \(A, B\), where \(\Phi\) is a positive linear map between matrix algebras (in particular, \(\Phi(A) = KAK^*\)) and \(\sigma\) is an operator mean (in particular, the weighted geometric mean), which are considered as certain reciprocal Lie-Trotter formulas and also a generalization of Kato’s limit to the supremum \(A \vee B\) with respect to the spectral order.

2010 Mathematics Subject Classification: Primary 15A45, 15A42, 47A64

Key words and phrases: positive semidefinite matrix, Lie-Trotter formula, positive linear map, operator mean, operator monotone function, geometric mean, antisymmetric tensor power, Rényi relative entropy

1 Introduction

For any matrices \(X\) and \(Y\), the well-known Lie-Trotter formula is the convergence

\[
\lim_{n \to \infty} \left( e^{X/n} e^{Y/n} \right)^n = e^{X+Y}.
\]

The symmetric form with a continuous parameter is also well-known for positive semidefinite matrices \(A, B \geq 0\) as

\[
\lim_{p \to 0} (A^{p/2} B^p A^{p/2})^{1/p} = P_0 \exp(\log A + \log B),
\]

where \(P_0\) is the orthogonal projection onto the intersection of the supports of \(A, B\) and \(\log A + \log B\) is defined as \(P_0(\log A)P_0 + P_0(\log B)P_0\). When \(\sigma\) is an operator mean corresponding to an operator monotone function \(f\) on \((0, \infty)\) such that \(\alpha := f'(1)\) is in \((0, 1)\), the operator mean version of the Lie-Trotter formula is also known to hold as

\[
\lim_{p \to 0} (A^p \sigma B^p)^{1/p} = P_0 \exp((1 - \alpha) \log A + \alpha \log B)
\]
for matrices $A, B \geq 0$. In particular, when $\sigma$ is the geometric mean $A\# B$ (introduced first in [15] and further discussed in [13]) corresponding to the operator monotone function $f(x) = x^{1/2}$, (1.2) yields

$$\lim_{p \downarrow 0} (A^p \# B^p)^{2/p} = P_0 \exp(\log A + \log B),$$

(1.3)

which has the same right-hand side as (1.1). Due to the Araki-Lieb-Thirring inequality and the Ando-Hiai log-majorization [4, 1], it is worthwhile to note that $(A^p/B^p \sigma B^p A^p/B^p)^{1/p}$ and $(A^p \# B^p)^{2/p}$ both tend to $P_0 \exp(\log A + \log B)$ as $p \downarrow 0$, with the former decreasing and the latter increasing in the log-majorization order (see [1] for details on log-majorization for matrices).

In the previous paper [6], under the name “reciprocal Lie-Trotter formula”, we considered the question complementary to (1.1) and (1.3), that is, about what happens to the limits of $(A^p/B^p \sigma B^p A^p/B^p)^{1/p}$ and $(A^p \# B^p)^{2/p}$ as $p \to \infty$ instead of 0. If $A$ and $B$ are commuting, then $(A^p/B^p \sigma B^p A^p/B^p)^{1/p} = (A^p \# B^p)^{2/p} = AB$, independently of $p > 0$. However, if $A$ and $B$ are not commuting, then the question is rather complicated. Indeed, although we can prove the existence of the limit $\lim_{p \to \infty} (A^p/B^p \sigma B^p A^p/B^p)^{1/p}$, the description of the limit has a rather complicated combinatorial nature. Moreover, it is unknown so far whether the limit of $(A^p \# B^p)^{2/p}$ as $p \to \infty$ exists or not. In the present paper, we consider a similar (but seemingly a bit simpler) question about what happens to the limits of $(BA^p B^p)^{1/p}$ and $(A^p \# B^p)^{1/p}$ as $p$ tends to $\infty$, the case where $B$ is fixed without the $p$-power, in certain more general settings.

The rest of the paper is organized as follows. In Section 2, we first prove the existence of the limit of $(KA^p K^*)^{1/p}$ as $p \to \infty$ and give the description of the limit in terms of the diagonalization (eigenvalues and eigenvectors) data of $A$ and the images of the eigenvectors by $K$. We then extend the result to the limit of $\Phi(A^p)^{1/p}$ as $p \to \infty$ for a positive linear map $\Phi$ between matrix algebras. For instance, this limit is applied to the map $\Phi(A \oplus B) := (A + B)/2$ to reformulate Kato’s limit theorem $((A^p + B^p)/2)^{1/p} \to A \lor B$ in [12]. Another application is given to find the limit formula as $\alpha \downarrow 0$ of the sandwiched $\alpha$-Rényi divergence [14, 18], a new relative entropy relevant to quantum information theory. In Section 3, we discuss the limit behavior of $(A^p \sigma B)^{1/p}$ as $p \to \infty$ for operator means $\sigma$. To do this, we may assume without loss of generality that $B$ is an orthogonal projection $E$. Under a certain condition on $\sigma$, we prove that $(A^p \sigma E)^{1/p}$ is decreasing as $1 \leq p \to \infty$, so that the limit as $p \to \infty$ exists. Furthermore, when $\sigma$ is the the weighted geometric mean, we obtain an explicit description of the limit in terms of $E$ and the spectral projections of $A$.

It is worth noting that a limit formula in the same vein as those in [12] and this paper was formerly given in [2, 3] for the spectral shorting operation.
2 \ \lim_{p \to \infty} \Phi(A^p)^{1/p} \text{ for positive linear maps } \Phi

For each \( n \in \mathbb{N} \) we write \( \mathbb{M}_n \) for the \( n \times n \) complex matrix algebra and \( \mathbb{M}_n^+ \) for the set of positive semidefinite matrices in \( \mathbb{M}_n \). When \( A \in \mathbb{M}_n \) is positive definite, we write \( A > 0 \). We denote by \( \text{Tr} \) the usual trace functional on \( \mathbb{M}_n \). For \( A \in \mathbb{M}_n^+ \), \( \lambda_1(A) \geq \cdots \geq \lambda_n(A) \) are the eigenvalues of \( A \) in decreasing order with multiplicities, and \( \text{ran} \ A \) is the range of \( A \).

Let \( A \in \mathbb{M}_n^+ \) be given, whose diagonalization is

\[
A = V \text{diag}(a_1, \ldots, a_n)V^* = \sum_{i=1}^{n} a_i |v_i\rangle \langle v_i|
\]

with the eigenvalues \( a_1 \geq \cdots \geq a_n \) and a unitary matrix \( V = [v_1 \cdots v_n] \) so that \( Av_i = a_i v_i \) for \( 1 \leq i \leq d \). Let \( K \in \mathbb{M}_n \) and assume that \( K \neq 0 \) (our problem below is trivial when \( K = 0 \)). Consider the sequence of vectors \( Kv_1, \ldots, Kv_n \) in \( \mathbb{C}^n \). Let \( 1 \leq l_1 < l_2 < \cdots < l_m \) be chosen so that if \( l_{k-1} < i < l_k \) for \( 1 \leq k \leq m \), then \( Kv_i \) is in \( \text{span}\{Kv_{l_1}, \ldots, Kv_{l_{k-1}}\} \) (this means, in particular, \( Kv_i = 0 \) if \( i < l_1 \)). Then \( \{Kv_{l_1}, \ldots, Kv_{l_m}\} \) is a linearly independent subset of \( \{Kv_1, \ldots, Kv_n\} \), so we perform the Gram-Schmidt orthogonalization to obtain an orthonormal vectors \( u_1, \ldots, u_m \) from \( Kv_{l_1}, \ldots, Kv_{l_m} \). In particular, \( u_1 = Kv_{l_1} / \|Kv_{l_1}\| \). The next theorem is our first limit theorem. This implicitly says that the right-hand side of (2.2) is independent of the expression of (2.1) (note that \( v_i \)'s are not unique for degenerate eigenvalues \( a_i \)).

**Theorem 2.1.** We have

\[
\lim_{p \to \infty} (KA^pK^*)^{1/p} = \sum_{k=1}^{m} a_{l_k} |u_k\rangle \langle u_k|,
\]

and in particular,

\[
\lim_{p \to \infty} \lambda_k((KA^pK^*)^{1/p}) = \begin{cases} a_{l_k}, & 1 \leq k \leq m, \\ 0, & m < k \leq n. \end{cases}
\]

**Proof.** Write \( Z_p := (KA^pK^*)^{1/p} \) and \( \lambda_i(p) := \lambda_i(Z_p) \) for \( p > 0 \) and \( 1 \leq i \leq n \). First we prove (2.3). Note that

\[
Z_p^p = KA^pK^* = KV \text{diag}(a_1^p, \ldots, a_n^p)V^* K^* = [a_1^p Kv_1 \ a_2^p Kv_2 \ \cdots \ a_n^p Kv_n] [Kv_1 \ Kv_2 \ \cdots \ Kv_n]^*.
\]

Since

\[
\lambda_i(p)^p \leq \text{Tr} \ Z_p^p = \text{Tr} [Kv_1 \ \cdots \ Kv_n]^* [a_1^p Kv_1 \ \cdots \ a_n^p Kv_n] = \sum_{i=1}^{n} a_i^p \langle Kv_i, Kv_i \rangle \leq a_1^p \sum_{i=1}^{n} \|Kv_i\|^2,
\]

for \( i \leq m \) and

\[
\lambda_i(p)^p \leq \text{Tr} \ Z_p^p = \text{Tr} [Kv_1 \ \cdots \ Kv_n]^* [a_1^p Kv_1 \ \cdots \ a_n^p Kv_n] = \sum_{i=m+1}^{n} a_i^p \langle Kv_i, Kv_i \rangle \leq a_m^p \sum_{i=m+1}^{n} \|Kv_i\|^2,
\]

for \( m < i \leq n \).
we have
\[ \limsup_{p \to \infty} \lambda_1(p) \leq a_1. \]

Moreover, since
\[ n\lambda_1(p)^p \geq \text{Tr} Z^p_p = \sum_{i=1}^n a_i^p \langle Kv_i, Kv_i \rangle \geq a_{l_1} \|Kv_i\|^2, \]
we have
\[ \liminf_{p \to \infty} \lambda_1(p) \geq a_{l_1}. \]

Therefore, (2.3) holds for \( k = 1 \).

To prove (2.3) for \( k > 1 \), we consider the antisymmetric tensor powers \( A^{\wedge k} \) and \( K^{\wedge k} \) for each \( k = 1, \ldots, n \). Note that
\[ Z^{\wedge k}_p = \left( (K^{\wedge k})(A^{\wedge k})^p (K^{\wedge k})^* \right)^{1/p} \]
and
\[ A^{\wedge k} = \sum_{1 \leq i_1 < \cdots < i_k \leq d} a_{i_1} \cdots a_{i_k} \langle v_{i_1} \wedge \cdots \wedge v_{i_k} \rangle \langle v_{i_1} \wedge \cdots \wedge v_{i_k} \rangle. \]

The above case applied to \( A^{\wedge k} \) and \( K^{\wedge k} \) yields that
\[ \lim_{p \to \infty} \lambda_1(p) \lambda_2(p) \cdots \lambda_k(p) = \lim_{p \to \infty} \lambda_1(Z^{\wedge k}_p) \]
\[ = \max \{ a_{i_1} \cdots a_{i_k} : 1 \leq i_1 < \cdots < i_k \leq d, \ K^{\wedge k}(v_{i_1} \wedge \cdots \wedge v_{i_k}) \neq 0 \}. \] (2.6)

Since \( K^{\wedge k}(v_{i_1} \wedge \cdots \wedge v_{i_k}) = K v_{i_1} \wedge \cdots \wedge K v_{i_k} \) is non-zero if and only if \( \{ K v_{i_1}, \ldots, K v_{i_k} \} \) is linearly independent, it is easy to see that (2.6) is equal to \( a_{i_1} \cdots a_{i_k} \) if \( k \leq m \) and equal to 0 if \( k > m \). Therefore,
\[ \lim_{p \to \infty} \lambda_1(p) \lambda_2(p) \cdots \lambda_k(p) = \begin{cases} a_{i_1} \cdots a_{i_k}, & 1 \leq k \leq m, \\ 0, & m < k \leq n, \end{cases} \]
which implies (2.3).

Now, for \( p > 0 \) choose an orthonormal basis \( \{ u_1(p), \ldots, u_n(p) \} \) of \( \mathbb{C}^n \) for which \( Z_p u_i(p) = \lambda_i(p) u_i(p) \) for \( 1 \leq i \leq n \). To prove (2.2), write \( \tilde{a}_k := a_{l_k} \) for \( 1 \leq k \leq m \). If \( \tilde{a}_1 = 0 \) then it is obvious that \( \lim_{p \to \infty} Z_p = 0 \). So assume that \( \tilde{a}_1 > 0 \) and furthermore \( \tilde{a}_1 > \tilde{a}_2 \), i.e., \( \lim_{p \to \infty} \lambda_1(p) > \lim_{p \to \infty} \lambda_2(p) \) at the moment. From (2.3) we have
\[ Z^p_p = \sum_{i=1}^n a_i^p \langle Kv_i \rangle \langle Kv_i \rangle \]
\[ = \sum_{i=l_1}^{l_2-1} a_i^p \langle Kv_i \rangle \langle Kv_i \rangle + \sum_{i=l_2}^n a_i^p \langle Kv_i \rangle \langle Kv_i \rangle \]

4
so that

$$
\left( \frac{Z_p}{\tilde{a}_1} \right)^p = \sum_{i=1}^{l_2-1} \left( \frac{a_i}{\tilde{a}_1} \right)^p \|Kv_i\|^2 |u_1\rangle \langle u_1| + \sum_{i=l_2}^{n} \left( \frac{a_i}{\tilde{a}_1} \right)^p |Kv_i\rangle \langle Kv_i| 
$$

as $p \to \infty$ for some $\alpha > 0$, since $a_i/\tilde{a}_1 \leq \tilde{a}_2/\tilde{a}_1 < 1$ for $i \geq l_2$. Hence, for any $p > 0$ sufficiently large, the largest eigenvalue of $(Z_p/\tilde{a}_1)^p$ is simple and the corresponding eigen projection converges to $|u_1\rangle \langle u_1|$ as $p \to \infty$. Since the eigen projection $E_1(p)$ of $Z_p$ corresponding to the largest eigenvalue $\lambda_1(p)$ (simple for any large $p > 0$) is the same as that of $(Z_p/\tilde{a}_1)^p$, we have

$$
E_1(p) = |u_1(p)\rangle \langle u_1(p)| \longrightarrow |u_1\rangle \langle u_1| \quad \text{as } p \to \infty.
$$

In the general situation, we assume that $\tilde{a}_1 \geq \cdots \geq \tilde{a}_k > \tilde{a}_{k+1}$ with $1 \leq k \leq m$, where $\tilde{a}_{k+1} = 0$ if $k = m$. From (2.3) note that

$$
\lim_{p \to \infty} \lambda_1(Z_p^{\wedge k}) = \tilde{a}_1 \cdots \tilde{a}_{k-1} \tilde{a}_k > \tilde{a}_1 \cdots \tilde{a}_{k-1} \tilde{a}_{k+1} = \lim_{p \to \infty} \lambda_2(Z_p^{\wedge k}).
$$

Hence, for any sufficiently large $p > 0$, the largest eigenvalue $\lambda_1(Z_p^{\wedge k}) = \lambda_1(p) \cdots \lambda_k(p)$ of $Z_p^{\wedge k}$ is simple, and from the above case applied to (2.5) it follows that

$$
|u_1(p) \wedge \cdots \wedge u_k(p)\rangle \langle u_1(p) \wedge \cdots \wedge u_k(p)| \longrightarrow |u_1 \wedge \cdots \wedge u_k\rangle \langle u_1 \wedge \cdots \wedge u_k| \quad \text{as } p \to \infty,
$$

since the vector in the present situation corresponding to $u_1 = Kv_{l_1}/\|Kv_{l_1}\|$ is

$$
\frac{K^{\wedge k}(v_{l_1} \wedge \cdots \wedge v_{l_k})}{\|K^{\wedge k}(v_{l_1} \wedge \cdots \wedge v_{l_k})\|} = \frac{Kv_{l_1} \wedge \cdots \wedge Kv_{l_k}}{\|Kv_{l_1} \wedge \cdots \wedge Kv_{l_k}\|} = u_1 \wedge \cdots \wedge u_k.
$$

(The last identity follows from the fact that, for linearly independent $w_1, \ldots, w_k$, $w_1 \wedge \cdots \wedge w_k/\|w_1 \wedge \cdots \wedge w_k\| = w'_1 \wedge \cdots \wedge w'_k$ if $w'_1, \ldots, w'_k$ are the Gram-Schmidt orthogonalization of $w_1, \ldots, w_k$. By [6] Lemma 2.4 we see that the orthogonal projection $E_k(p)$ onto span$\{u_1(p), \ldots, u_k(p)\}$ converges to the orthogonal projection $E_k$ of span$\{u_1, \ldots, u_k\}$.

Finally, let $0 = k_0 < k_1 < \cdots < k_{s-1} < k_s = m$ be such that

$$
\tilde{a}_1 = \cdots = \tilde{a}_{k_1} > \tilde{a}_{k_1+1} = \cdots = \tilde{a}_{k_2} > \cdots > \tilde{a}_{k_{s-1}+1} = \cdots = \tilde{a}_{k_s}.
$$

The above argument says that, for every $r = 1, \ldots, s-1$, the orthogonal projection $E_{kr}(p)$ onto span$\{u_1(p), \ldots, u_{kr}(p)\}$ converges to the orthogonal projection $E_{kr}$ onto span$\{u_1, \ldots, u_{kr}\}$. When $\tilde{a}_{k_r} > 0$, this holds for $r = s$ as well. Therefore, when $\tilde{a}_{k_s} > 0$, we have

$$
Z_p = \sum_{i=1}^{n} \lambda_i(p) |u_i(p)\rangle \langle u_i(p)|
$$
\[
\sum_{r=1}^{s-1} \sum_{i=k_{r-1}+1}^{k_r} \lambda_i(p) |u_i(p)| \langle u_i(p) \rangle + \sum_{i=k_s+1}^{n} \lambda_i(p) |u_i(p)| \langle u_i(p) \rangle
\]

\[
= \sum_{r=1}^{s} \sum_{i=k_{r-1}+1}^{k_r} (\lambda_i(p) - \bar{a}_i) |u_i(p)| \langle u_i(p) \rangle + \sum_{r=1}^{s} \bar{a}_{k_r}(E_{k_r}(p) - E_{k_{r-1}}(p))
\]

\[
+ \sum_{i=k_s+1}^{n} \lambda_i(p) |u_i(p)| \langle u_i(p) \rangle
\]

\[
\rightarrow \sum_{r=1}^{s} \bar{a}_{k_r}(E_{k_r} - E_{k_{r-1}}) = \sum_{i=1}^{m} \bar{a}_i |u_i| \langle u_i \rangle, \quad \text{where } E_0(p) = E_0 = 0.
\]

When \( \bar{a}_{k_s} = 0 \), we may modify the above estimate as

\[
Z_p = \sum_{r=1}^{s-1} \sum_{i=k_{r-1}+1}^{k_r} \lambda_i(p) |u_i(p)| \langle u_i(p) \rangle + \sum_{i=k_{s-1}+1}^{n} \lambda_i(p) |u_i(p)| \langle u_i(p) \rangle
\]

\[
\rightarrow \sum_{r=1}^{s-1} \bar{a}_{k_r}(E_{k_r} - E_{k_{r-1}}) = \sum_{i=k_{s-1}+1}^{k_s-1} \bar{a}_i |u_i| \langle u_i \rangle = \sum_{i=1}^{m} \bar{a}_i |u_i| \langle u_i \rangle.
\]

\[\square\]

The following corollary of Theorem 2.1 is an improvement of [7, Theorem 1.2].

**Corollary 2.2.** Let \( A \in \mathbb{M}_n \) be positive definite. We have \( \lim_{p \to \infty} \lambda_i((KA^pK^*)^{1/p}) = a_i \) for all \( i = 1, \ldots, n \) if and only if \( \{Kv_1, \ldots, Kv_n\} \) is linearly independent.

**Remark 2.3.** Note that Theorem 2.1 can easily extend to the case where \( K \) is a rectangle \( n' \times n \) matrix. In fact, when \( n' < n \) we may apply Theorem 2.1 to \( n \times n \) matrices \( \begin{bmatrix} K & O \\ O & \end{bmatrix} \) and \( A \), and when \( n' > n \) we may apply to \( n' \times n' \) matrices \( \begin{bmatrix} K & O \\ O & \end{bmatrix} \) and \( A \oplus O_{n'-n} \).

A linear map \( \Phi : \mathbb{M}_n \to \mathbb{M}_{n'} \) is said to be positive if \( \Phi(A) \in \mathbb{M}_{n'}^+ \) for all \( A \in \mathbb{M}_n^+ \), which is further said to be strictly positive if \( \Phi(I_n) > 0 \), that is, \( \Phi(A) > 0 \) for all \( A \in \mathbb{M}_n \), \( A > 0 \). The following is an extended and refined version of Theorem 2.1.

**Theorem 2.4.** Let \( \Phi : \mathbb{M}_n \to \mathbb{M}_{n'} \) be a positive linear map. Let \( A \in \mathbb{M}_n^+ \) be given as \( A = \sum_{i=1}^{n} a_i |v_i\rangle \langle v_i| \) with \( a_1 \geq \cdots \geq a_n \) and an orthonormal basis \( \{v_1, \ldots, v_n\} \) of \( \mathbb{C}^n \). Then \( \lim_{p \to \infty} \Phi(A^p)^{1/p} \) exists and

\[
\lim_{p \to \infty} \Phi(A^p)^{1/p} = \sum_{i=1}^{n} a_i P_{\mathcal{M}_i},
\]

where \( \mathcal{M}_1 := \text{ran } \Phi(\langle v_1 \rangle \langle v_1 \rangle) \).
\[ \mathcal{M}_i := \bigvee_{j=1}^{i} \operatorname{ran} \Phi(|v_0\rangle\langle v_0|) \oplus \bigvee_{j=1}^{i-1} \operatorname{ran} \Phi(|v_j\rangle\langle v_j|), \quad 2 \leq i \leq n, \]

and \( P_{\mathcal{M}_i} \) is the orthogonal projection onto \( \mathcal{M}_i \) for \( 1 \leq i \leq n \).

**Proof.** Let \( C^*(I, A) \) be the commutative \( C^* \)-subalgebra of \( M_n \) generated by \( I, A \). We can consider the composition of the conditional expectation from \( M_n \) onto \( C^*(I, A) \) with respect to \( \text{Tr} \) and \( \Phi|_{C^*(I, A)} : C^*(I, A) \to M_{\nu}^* \) instead of \( \Phi \), so we may assume that \( \Phi \) is completely positive. By the Stinespring representation there are a \( \nu \in \mathbb{N} \), a \( * \)-homomorphism \( \pi : M_n \to M_{\nu}^* \) and a linear map \( K : \mathbb{C}^n_{\nu} \to \mathbb{C}^{\nu^\prime} \) such that \( \Phi(X) = K\pi(X)K^* \) for all \( X \in M_n \). Moreover, since \( \pi : M_n \to M_{\nu}^* \) is represented, under a suitable change of an orthonormal basis of \( \mathbb{C}^{\nu^\prime} \), as \( \Phi(X) = I_{\nu} \otimes X \) for all \( X \in M_n \) under identification \( M_{\nu}^* = M_\nu \otimes M_n \), we can assume that \( \Phi \) is given (with a change of \( K \)) as

\[ \Phi(X) = K(I_{\nu} \otimes X)K^*, \quad X \in M_n. \]

We then write

\[ I_{\nu} \otimes A = (I_{\nu} \otimes V)\operatorname{diag}(a_1, \ldots, a_1, a_2, \ldots, a_2, \ldots, a_n, \ldots, a_n) (I_{\nu} \otimes V)^* = \sum_{i=1}^{n} a_i (|e_1 \otimes v_i\rangle\langle e_1 \otimes v_i| + \cdots + |e_\nu \otimes v_i\rangle\langle e_\nu \otimes v_i|). \]

Now, we consider the following sequence of \( n\nu \) vectors in \( \mathbb{C}^{\nu^\prime} \):

\[ K(e_1 \otimes v_1), \ldots, K(e_\nu \otimes v_1), K(e_1 \otimes v_2), \ldots, K(e_\nu \otimes v_2), \ldots, K(e_1 \otimes v_n), \ldots, K(e_\nu \otimes v_n), \]

and if \( K(e_j \otimes v_i) \) is a linear combination of the vectors in the sequence preceding it, then we remove it from the sequence. We write the resulting linearly independent subsequence as

\[ K(e_j \otimes v_{l_1}) \quad (j \in J_1), \quad K(e_j \otimes v_{l_2}) \quad (j \in J_2), \ldots, \quad K(e_j \otimes v_{l_m}) \quad (j \in J_m), \]

where \( 1 \leq l_1 < l_2 < \cdots < l_m \leq n \) and \( J_1, \ldots, J_m \subseteq \{1, \ldots, \nu\} \). Furthermore, by performing the Gram-Schmidt orthogonalization to this subsequence, we end up making an orthonormal sequence of vectors in \( \mathbb{C}^{\nu^\prime} \) as follows:

\[ u_j^{(l_1)} \quad (j \in J_1), \quad u_j^{(l_2)} \quad (j \in J_2), \ldots, \quad u_j^{(l_m)} \quad (j \in J_m). \]

Since

\[ \Phi(A^p)^{1/p} = K((I_{\nu} \otimes A)^p)^{1/p}K^*, \]

Theorem 2.1 and Remark 2.3 imply that \( \lim_{p \to \infty} \Phi(A^p)^{1/p} \) exists and

\[ \lim_{p \to \infty} \Phi(A^p)^{1/p} = \sum_{k=1}^{m} a_k \left( \sum_{j \in J_k} |u_j^{(l_k)}\rangle\langle u_j^{(l_k)}| \right), \]

7
Lemma 2.5. For any finite set \( \{w_1, \ldots, w_k\} \) in \( \mathbb{C}^n \), \( \text{span}\{w_1, \ldots, w_k\} \) is equal to the range of \( \sum |w_j\rangle \langle w_j| \). More generally, for every \( B_1, \ldots, B_k \in \mathbb{M}_{n'}^+ \), \( \bigvee_{j=1}^k \text{ran } B_j \) is equal to the range of \( B_1 + \cdots + B_k \). 

\[ \sum_{j \in J_k} |u_j^{(l_k)}\rangle \langle u_j^{(l_k)}| = 0 \text{ if } J_k = \emptyset. \]

The next step of the proof is to find what is \( \sum_{j \in J_k} |u_j^{(l_k)}\rangle \langle u_j^{(l_k)}| \) for \( 1 \leq k \leq m \). For this we first note that

\[
\sum_{j=1}^\nu |K(e_j \otimes v_i)\rangle \langle K(e_j \otimes v_i)| = K\left( \sum_{j=1}^\nu |e_j \otimes v_i\rangle \langle e_j \otimes v_i| \right) K^* = K(I_\nu \otimes |v_i\rangle \langle v_i|) K^* = \Phi(|v_i\rangle \langle v_i|).
\]

From Lemma 2.5 below this implies that

\[
\mathcal{R}_i := \text{ran } \Phi(|v_i\rangle \langle v_i|) = \text{span}\{K(e_j \otimes v_i) : 1 \leq j \leq \nu\}.
\]

Through the procedure of the Gram-Schmidt diagonalization we see that

\[
\begin{align*}
\mathcal{R}_i &= 0, \quad 1 \leq i < l_1, \\
\mathcal{R}_{l_1} &= \text{span}\{u_j^{(l_1)} : j \in J_1\}, \\
\mathcal{R}_i &\subset \mathcal{R}_{l_1}, \quad l_1 < i < l_2, \\
(\mathcal{R}_{l_1} \lor \mathcal{R}_{l_2}) \ominus \mathcal{R}_{l_1} &= \text{span}\{u_j^{(l_2)} : j \in J_2\}, \\
\mathcal{R}_i &\subset \mathcal{R}_{l_1} \lor \mathcal{R}_{l_2}, \quad l_2 < i < l_3, \\
(\mathcal{R}_{l_1} \lor \mathcal{R}_{l_2} \lor \mathcal{R}_{l_3}) \ominus (\mathcal{R}_{l_1} \lor \mathcal{R}_{l_2}) &= \text{span}\{u_j^{(l_3)} : j \in J_3\}, \\
&\vdots \\
\mathcal{R}_i &\subset \mathcal{R}_{l_1} \lor \cdots \lor \mathcal{R}_{l_{m-1}}, \quad l_{m-1} < i < l_m, \\
(\mathcal{R}_{l_1} \lor \cdots \lor \mathcal{R}_{l_m}) \ominus (\mathcal{R}_{l_1} \lor \cdots \lor \mathcal{R}_{l_{m-1}}) &= \text{span}\{u_j^{(l_m)} : j \in J_m\}, \\
\mathcal{R}_i &\subset \mathcal{R}_{l_1} \lor \cdots \lor \mathcal{R}_{l_m}, \quad l_m < i \leq n.
\end{align*}
\]

Now, let \( P_{M_i} \) be the orthogonal projections, respectively, onto the subspaces

\[
\mathcal{M}_1 := \mathcal{R}_1, \quad \mathcal{M}_i := (\mathcal{R}_1 \lor \cdots \lor \mathcal{R}_i) \ominus (\mathcal{R}_1 \lor \cdots \lor \mathcal{R}_{i-1}), \quad 2 \leq i \leq n,
\]

so that \( P_{M_i} = 0 \) if \( i \not\in \{l_1, \ldots, l_m\} \) and \( P_{M_{l_k}} \) is the orthogonal projection onto \( \text{span}\{u_j^{(l_k)} : j \in J_k\} \) for \( 1 \leq k \leq m \). Therefore, we have

\[
\lim_{p \to \infty} \Phi(A^p)^{1/p} = \sum_{k=1}^m a_{l_k} \left( \sum_{j \in J_k} |u_j^{(l_k)}\rangle \langle u_j^{(l_k)}| \right) = \sum_{k=1}^m a_{l_k} P_{M_{l_k}} = \sum_{i=1}^n a_i P_{M_i}.
\]

\[ \square \]
Proof. Let $Q := |w_1\rangle\langle w_1| + \cdots + |w_k\rangle\langle w_k|$. Since

$$Qx = \langle w_1, x \rangle w_1 + \cdots + \langle w_k, x \rangle w_k \in \text{span}\{w_1, \ldots, w_k\}$$

for all $x \in \mathbb{C}^n$, we have $\text{ran} \ Q \subseteq \text{span}\{w_1, \ldots, w_k\}$. Since $|w_i\rangle\langle w_i| \leq Q$, we have

$$w_i \in \text{ran} \ |w_i\rangle\langle w_i| \subseteq \text{ran} \ Q, \quad 1 \leq i \leq k.$$ 

Hence we have $\text{span}\{w_1, \ldots, w_k\} \subseteq \text{ran} \ Q$. The proof of the latter assertion is similar. \hfill \square

Thanks to the lemma we can restate Theorem 2.4 as follows:

**Theorem 2.6.** Let $\Phi : M_n \rightarrow M_{n'}$ be a positive linear map. Let $A \in M_n^+$ be given with the spectral decomposition $A = \sum_{k=1}^m a_k P_k$, where $a_1 > a_2 > \cdots > a_m > 0$. Define

$$\mathcal{M}_1 := \text{ran} \ \Phi(P_1),$$

$$\mathcal{M}_k := \text{ran} \ \Phi(P_1 + \cdots + P_k) \ominus \text{ran} \ \Phi(P_1 + \cdots + P_{k-1}), \quad 2 \leq k \leq m.$$ 

Then

$$\lim_{p \rightarrow \infty} \Phi(A^p)^{1/p} = \sum_{k=1}^m a_k P_{\mathcal{M}_k}.$$ 

**Example 2.7.** Consider a linear map $\Phi : M_{2n} \rightarrow M_n$ given by

$$\Phi \left( \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix} \right) := \frac{X_{11} + X_{22}}{2}, \quad X_{ij} \in M_n.$$ 

Clearly, $\Phi$ is completely positive. For any $A, B \in M_n^+$ and $p > 0$ we have

$$\Phi \left( \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}^p \right)^{1/p} = \left( \frac{A^p + B^p}{2} \right)^{1/p}.$$ 

Thus, it is well-known \cite{12} that

$$\lim_{p \rightarrow \infty} \Phi \left( \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}^p \right)^{1/p} = \lim_{p \rightarrow \infty} \left( \frac{A^p + B^p}{2} \right)^{1/p} = A \lor B,$$ 

where $A \lor B$ is the supremum of $A, B$ in the spectral order. Here let us show (2.7) from Theorem 2.6. The spectral decompositions of $A, B$ are given as

$$A = \sum_{i=1}^m a_i P_i, \quad B = \sum_{j=1}^{m'} b_j Q_j,$$

where $a_1 > \cdots > a_m \geq 0$, $b_1 > \cdots > b_{m'} \geq 0$ and $\sum_{i=1}^m P_i = \sum_{j=1}^{m'} Q_j = I$. Then

$$A \oplus B = \sum_{k=1}^l c_k R_k,$$
where \( \{ c_k \}_{k=1}^l = \{ a_i \}_{i=1}^m \cup \{ b_j \}_{j=1}^{m'} \) with \( c_1 > \cdots > c_l \) and

\[
R_k = \begin{cases} 
  P_i \oplus Q_j & \text{if } a_i = b_j = c_k, \\
  P_i \oplus 0 & \text{if } a_i = c_k \text{ and } b_j \neq c_k \text{ for all } j, \\
  0 \oplus Q_j & \text{if } b_j = c_k \text{ and } a_i \neq c_k \text{ for all } i.
\end{cases}
\]

Note that

\[
\Phi(R_1 + \cdots + R_k) = \frac{1}{2} \left( \sum_{i:a_i \geq c_k} P_i + \sum_{j:b_j \geq c_k} Q_j \right)
\]

so that by Lemma 2.5 the support projection \( F_k \) (i.e., the orthogonal projection onto the range) of \( \Phi(R_1 + \cdots + R_k) \) is

\[
F_k = \left( \sum_{i:a_i \geq c_k} P_i \right) \vee \left( \sum_{j:b_j \geq c_k} Q_j \right).
\]

Theorem 2.6 implies that

\[
\lim_{p \to \infty} \Phi \left( \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \right)^{1/p} = C := \sum_{k=1}^l c_k (F_k - F_{k-1}).
\]

For every \( x \in \mathbb{R} \) we denote by \( E_{[x,\infty)}(A) \) the spectral projection of \( A \) corresponding to the interval \([x, \infty)\), i.e.,

\[
E_{[x,\infty)}(A) := \sum_{i:a_i \geq x} P_i,
\]

and similarly for \( E_{[x,\infty)}(B) \) and \( E_{[x,\infty)}(C) \). If \( c_k \geq x > c_{k+1} \) for some \( 1 \leq k < l \), then we have

\[
E_{[x,\infty)}(C) = F_k = E_{[x,\infty)}(A) \vee E_{[x,\infty)}(B).
\]

This holds also when \( x > c_1 \) and \( x \leq c_l \). Indeed, when \( x > c_1 \), \( E_{[x,\infty)}(C) = 0 = E_{[x,\infty)}(A) \vee E_{[x,\infty)}(B) \). When \( x \geq c_1 \), \( E_{[x,\infty)}(C) = I = E_{[x,\infty)}(A) \vee E_{[x,\infty)}(B) \). This description of \( C \) is the same as \( A \vee B \) in [12], so we have \( C = A \vee B \).

**Example 2.8.** The example here is relevant to quantum information. For density matrices \( \rho, \sigma \in \mathbb{M}_n \) (i.e., \( \rho, \sigma \in \mathbb{M}_n^+ \) with \( \operatorname{Tr} \rho = \operatorname{Tr} \sigma = 1 \)) and for a parameter \( \alpha \in (0, \infty) \setminus \{1\} \), the traditional Rényi relative entropy is

\[
D_{\alpha}(\rho \| \sigma) := \begin{cases} 
  \frac{1}{\alpha-1} \log \left[ \operatorname{Tr} \rho^\alpha \sigma^{1-\alpha} \right] & \text{if } \rho^0 \leq \sigma^0 \text{ or } 0 < \alpha < 1, \\
  +\infty & \text{otherwise},
\end{cases}
\]

where \( \rho^0 \) denotes the support projection of \( \rho \). On the other hand, the new concept recently introduced and called the sandwiched Rényi relative entropy [14] [18] is

\[
\tilde{D}_{\alpha}(\rho \| \sigma) := \begin{cases} 
  \frac{1}{\alpha-1} \log \left[ \operatorname{Tr} \left( \sigma^{\frac{1-\alpha}{2\alpha}} \rho \sigma^{\frac{1-\alpha}{2\alpha}} \right)^\alpha \right] & \text{if } \rho^0 \leq \sigma^0 \text{ or } 0 < \alpha < 1, \\
  +\infty & \text{otherwise}.
\end{cases}
\]
By taking the limit we also consider
\[ D_0(\rho\|\sigma) := \lim_{\alpha \searrow 0} D_\alpha(\rho\|\sigma) = -\log \text{Tr} (\rho^0 \sigma), \]
\[ \tilde{D}_0(\rho\|\sigma) := \lim_{\alpha \searrow 0} \tilde{D}_\alpha(\rho\|\sigma) = -\log \left[ \lim_{\alpha \searrow 0} \text{Tr} \left( \sigma^{1-\alpha} \rho \sigma^{1-\alpha} \right)^\alpha \right]. \]

(We remark that the notations \( D_\alpha \) and \( \tilde{D}_\alpha \) are interchanged from those in \[8\].) Here, note that
\[ \lim_{\alpha \searrow 0} \text{Tr} \left( \sigma^{1-\alpha} \rho \sigma^{1-\alpha} \right)^\alpha = \lim_{p \to \infty} \text{Tr} \left( \sigma^{p/2} \rho \sigma^{p/2} \right)^{1/p} = \lim_{p \to \infty} \text{Tr} \left( \rho^0 \sigma^p \rho^0 \right)^{1/p}, \]
where the existence of \( \lim_{p \to \infty} \text{Tr} \left( \rho^0 \sigma^p \rho^0 \right)^{1/p} \) follows from the Araki-Lieb-Thirring inequality \[4\] (also \[1\]), and the latter equality above follows since \( \lambda \rho^0 \leq \rho \leq \mu \rho^0 \) for some \( \lambda, \mu > 0 \) and
\[ \lambda^{1/p} \text{Tr} \left( \sigma^{p/2} \rho^0 \sigma^{p/2} \right)^{1/p} \leq \text{Tr} \left( \sigma^{p/2} \rho \sigma^{p/2} \right)^{1/p} \leq \mu^{1/p} \text{Tr} \left( \sigma^{p/2} \rho^0 \sigma^{p/2} \right)^{1/p}. \]

It was proved in \[8\] that
\[ \tilde{D}_0(\rho\|\sigma) \leq D_0(\rho\|\sigma) \quad (2.8) \]
and equality holds in \[2.8\] if \( \rho^0 = \sigma^0 \). Let us here prove the following:

1. \( \tilde{D}_0(\rho\|\sigma) = -\log \tilde{Q}_0(\rho\|\sigma) \), where
\[ \tilde{Q}_0(\rho\|\sigma) := \max \{ \text{Tr} (P\sigma) : P \text{ an orthogonal projection}, \]
\[ [P, \sigma] = 0, (P \rho^0 P)^0 = P \}. \]

2. \( \tilde{D}_0(\rho\|\sigma) = D_0(\rho\|\sigma) \) holds if and only if \( [\rho^0, \sigma] = 0 \). (Obviously, \( [\rho^0, \sigma] = 0 \) if \( \rho^0 = \sigma^0 \).)

Indeed, to prove (1), first note that \( (P \rho^0 P)^0 = P \) means that the dimension of \( \text{ran} \rho^0 P \) is equal to that of \( P \), that is, \( \rho^0 v_1, \ldots, \rho^0 v_d \) are linearly independent when \( \{v_1, \ldots, v_d\} \) is an orthonormal basis of \( \text{ran} P \). Choose \( 1 \leq l_1 < l_2 < \cdots < l_m \) as in the first paragraph of this section (before Theorem \[2.1\]) for \( A = \sigma \) and \( K = \rho^0 \). Let \( P_0 \) be the orthogonal projection onto \( \text{span}\{v_1, \ldots, v_{l_m}\} \). Then \( [P_0, \sigma] = 0, (P_0 \rho^0 P_0)^0 = P_0 \), and Theorem \[2.1\] gives
\[ \lim_{p \to \infty} \text{Tr} \left( \rho^0 \sigma^p \rho^0 \right)^{1/p} = \sum_{k=1}^m a_{i_k} = \text{Tr} (P_0 \sigma). \]
On the other hand, let \( P \) be an orthogonal projection with \( [P, \sigma] = 0 \) and \( (P \rho^0 P)^0 = P \). From \( [P, \sigma] = 0 \) we may assume that \( P = \sum_{k=1}^d |v_i \rangle \langle v_i| \) for some \( 1 \leq i_1 < \cdots < i_d \leq n \).
(after, if necessary, changing \( v_i \) for degenerate eigenvalues \( a_i \)). Since \((P \rho^0 P)^0 = P\) implies that \(\rho^0 v_1, \ldots, \rho^0 v_d\) are linearly independent, we have \(d \leq m\) and

\[
\text{Tr} (P \sigma) = \sum_{k=1}^{d} a_{ik} \leq \sum_{k=1}^{m} a_{ik} = \text{Tr} (P_0 \sigma).
\]

Next, to prove (2), note that \(\text{Tr} (\rho^0 \sigma^p \rho^0)^{1/p}\) is increasing in \(p > 0\) by the Araki-Lieb-Thirring inequality mentioned above, which shows that

\[
\text{Tr} (\rho^0 \sigma) \leq \lim_{p \to \infty} \text{Tr} (\rho^0 \sigma^p \rho^0)^{1/p}.
\]

This means inequality (2.8), and equality holds in (2.8) if and only if \(\text{Tr} (\rho^0 \sigma^p \rho^0)^{1/p}\) is constant for \(p \geq 1\). By [10, Theorem 2.1] this is equivalent to the commutativity \(\rho^0 \sigma = \sigma \rho^0\).

Finally, we consider the complementary convergence of \(\Phi(A^p)^{1/p}\) as \(p \to -\infty\), or \(\Phi(A^{-p})^{-1/p}\) as \(p \to \infty\). Here, the expression \(\Phi(A^{-p})^{-1/p}\) for \(p > 0\) is defined in such a way that the \((-p)\)-power of \(A\) is restricted to the support of \(A\), i.e., defined in the sense of the generalized inverse, and the \((-1/p)\)-power of \(\Phi(A^{-p})\) is also in this sense.

The next theorem is the complementary counterpart of Theorem 2.6.

**Theorem 2.9.** Let \(\Phi : M_n \to M_{n'}\) be a positive linear map. Let \(A \in M_n^+\) be given with the spectral decomposition \(A = \sum_{k=1}^{m} a_k P_k\), where \(a_1 > a_2 > \cdots > a_m > 0\). Define

\[
\tilde{\mathcal{M}}_k := \text{ran} \Phi(P_k + \cdots + P_m) \ominus \text{ran} \Phi(P_{k+1} + \cdots + P_m), \quad 1 \leq i \leq m - 1,
\]

\[
\tilde{\mathcal{M}}_m := \text{ran} \Phi(P_m).
\]

Then

\[
\lim_{p \to -\infty} \Phi(A^{-p})^{-1/p} = \sum_{k=1}^{m} a_k P_{\tilde{\mathcal{M}}_k}.
\]

**Proof.** The proof is just a simple adaptation of Theorem 2.6. We can write for any \(p > 0\)

\[
\Phi(A^{-p})^{-1/p} = \left\{ \Phi((A^{-1})^p)^{1/p} \right\}^{-1},
\]

where \(A^{-1}\) and \(\{\cdots\}^{-1}\) are defined in the sense of the generalized inverse so that

\[
A^{-1} = \sum_{k=1}^{m} a_k^{-1} P_k = \sum_{k=1}^{m} a_{m+1-k}^{-1} P_{m+1-k}
\]

with \(a_m^{-1} > \cdots > a_1^{-1} > 0\). By Theorem 2.6 we have

\[
\lim_{p \to -\infty} \Phi((A^{-1})^p)^{1/p} = \sum_{k=1}^{m} a_{m+1-k}^{-1} P_{\tilde{\mathcal{M}}_{m+1-k}} = \sum_{k=1}^{m} a_k^{-1} P_{\tilde{\mathcal{M}}_k}.
\]
where

\[ \tilde{M}_m := \text{ran } \Phi(P_{m+1}) = \text{ran } \Phi(P_m), \]
\[ \tilde{M}_{m+1-k} := \text{ran } \Phi(P_{m+1} \oplus \cdots \oplus P_{m+1-k}) \oplus \text{ran } \Phi(P_{m+1} \oplus \cdots \oplus P_{m+1-(k-1)}) \]
\[ = \text{ran } \Phi(P_{m+1-k} \oplus \cdots \oplus P_m) \oplus \text{ran } \Phi(P_{m+2-k} \oplus \cdots \oplus P_m), \quad 2 \leq k \leq m. \]

According to the proofs of Theorems 2.1 and 2.4, we see that the \(i\)th eigenvalue \(\lambda_i(p)\) of \(\Phi((A^{-1})^p)^{1/p}\) converges to a positive real as \(p \to \infty\), or otherwise, \(\lambda_i(p) = 0\) for all \(p > 0\). That is, \(\lambda_i(p) \to 0\) as \(p \to \infty\) occurs only when \(\lambda_i(p) = 0\) for all \(p > 0\). This implies that

\[ \lim_{p \to \infty} \Phi(A^{-p})^{1/p} = \left\{ \lim_{p \to \infty} \Phi((A^{-1})^p)^{1/p} \right\}^{-1} = \sum_{k=1}^{m} a_k P_{\tilde{M}_k}. \]

\[ \square \]

**Remark 2.10.** Assume that \(\Phi : M_n \to M_n\) is a unital positive linear map. Let \(A \in M_n\) be positive definite and \(1 \leq p < q\). Since \(x^{p/q}\) and \(x^{1/p}\) are operator monotone on \([0, \infty)\), we have \(\Phi(A^q)^{p/q} \geq \Phi(A^p)\) and so \(\Phi(A^q)^{1/q} \geq \Phi(A^p)^{1/p}\). Hence \(\Phi(A^p)^{1/p}\) increases as \(1 \leq p \nearrow\). Similarly, \(\Phi(A^{-q})^{p/q} \geq \Phi(A^{-p})\) and so \(\Phi(A^{-q})^{-1/q} \leq \Phi(A^{-p})^{-1/p}\) since \(x^{-1/p}\) is operator monotone decreasing on \((0, \infty)\). Hence \(\Phi(A^{-p})^{-1/p}\) decreases as \(1 \leq p \nearrow\). Moreover, since \(x^{-1}\) is operator convex on \((0, \infty)\), we have \(\Phi(A^{-1})^{-1} \leq \Phi(A)\). (See [5, Theorem 2.1] for more details.) Combining altogether, when \(A\) is positive definite, we have

\[ \Phi(A^{-p})^{-1/p} \leq \Phi(A^q)^{1/q}, \quad p, q \geq 1, \quad (2.9) \]

and in particular,

\[ \lim_{p \to \infty} \Phi(A^{-p})^{-1/p} \leq \lim_{p \to \infty} \Phi(A^p)^{1/p}. \]

However, the latter inequality does not hold unless \(\Phi\) is unital. For example, let

\[ P_1 := \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad P_2 := \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \quad Q_1 := \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}, \quad Q_2 := \begin{bmatrix} 1/2 & -1/2 \\ -1/2 & 1/2 \end{bmatrix}, \]

and consider \(\Phi : M_2 \to M_2\) given by

\[ \Phi \left( \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \right) := a_{11}P_1 + a_{22}Q_1, \]

and \(A := aP_1 + bP_2\) where \(a > b > 0\). Since \(P_{\text{ran } \Phi(P_1+P_2)} = I\), \(P_{\text{ran } \Phi(P_1)} = P_1\) and \(P_{\text{ran } \Phi(P_2)} = Q_1\), Theorems 2.6 and 2.9 give

\[ \lim_{p \to \infty} \Phi(A^p)^{1/p} = aP_1 + b(I - P_1) = aP_1 + bP_2, \]
\[ \lim_{p \to \infty} \Phi(A^{-p})^{-1/p} = a(I - Q_1) + bQ_1 = aQ_2 + bQ_1. \]
We compute
\[(aP_1 + bP_2) - (aQ_2 + bQ_1) = \begin{bmatrix} a-b & a-b \\ \frac{a-b}{2} & \frac{a-b}{2} \end{bmatrix},\]
which is not positive semidefinite.

**Remark 2.11.** We may always assume that \(\Phi : M_n \to M_{n'}\) is strictly positive. Indeed, we may consider \(\Phi\) as \(\Phi : M_n \to Q_0M_{n'}Q_0 \cong M_{n''}\), where \(Q_0\) is the support projection of \(\Phi(I_n)\). Under this convention, another reasonable definition of \(\Phi(A^{-p})^{-1/p}\) for \(p \geq 1\) is
\[
\Phi(A^{-p})^{-1/p} := \lim_{\varepsilon \downarrow 0} \Phi((A + \varepsilon I_n)^{-p})^{-1/p},
\]
which is well defined since \(\Phi((A + \varepsilon I)^{-p})\) is increasing so that \(\Phi((A + \varepsilon I)^{-p})^{-1/p}\) is decreasing as \(\varepsilon \downarrow 0\). But this definition is different from the above definition of \(\Phi(A^{-p})^{-1/p}\). For example, let \(\Phi : M_2 \to M_2\) be given by \(\Phi(A) := KAK^*\) with an invertible \(K = \begin{bmatrix} a & b \\ c & d \end{bmatrix}\), and let \(A = P = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\). Then \(A^{-p} = P\) (in the generalized inverse) so that
\[
KA^{-p}K^* = \begin{bmatrix} |a|^2 & a\overline{c} \\ \overline{ac} & |c|^2 \end{bmatrix}
\]
and so
\[
(KA^{-p}K^*)^{-1/p} = \frac{1}{(|a|^2 + |c|^2)^{1-\frac{1}{p}}} \begin{bmatrix} |a|^2 & a\overline{c} \\ \overline{ac} & |c|^2 \end{bmatrix}. \tag{2.10}
\]
On the other hand,
\[
\lim_{\varepsilon \downarrow 0} (K(A + \varepsilon I)^{-p}K^*)^{-1/p} = \lim_{\varepsilon \downarrow 0} (K^{*^{-1}}K^{-1})^{1/p} = (K^{*^{-1}}A^{-1}P)^{-1/p} = (K^{*^{-1}}PK^{-1})^{1/p}
\]
is equal to
\[
\frac{1}{|ad - bc|^2/p(|b|^2 + |d|^2)^{1-\frac{1}{p}}} \begin{bmatrix} |d|^2 & -b\overline{d} \\ -b\overline{d} & |b|^2 \end{bmatrix}. \tag{2.11}
\]
Hence we find that (2.10) and (2.11) are very different, even after taking the limits as \(p \to \infty\).

Here is a simpler example. Let \(\varphi : M_2 \to \mathbb{C} = M_1\) be a state (hence, unital) with density matrix \(\begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}\), and let \(A = P = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\). For the first definition we have
\[
\lim_{p \to \infty} \varphi(A^{-p})^{-1/p} = \lim_{p \to \infty} 2^{1/p} = 1.
\]
For the second definition,
\[
\lim_{\varepsilon \downarrow 0} \varphi((A + \varepsilon I)^{-p})^{-1/p} = \lim_{\varepsilon \downarrow 0} \left\{ \frac{(1 + \varepsilon)^{-p} + \varepsilon^{-p}}{2} \right\}^{-1/p} = 0
\]
Hence we find that (2.10) and (2.11) are very different, even after taking the limits as \(p \to \infty\).
for all $p > 0$. Moreover, since $\varphi(A^p)^{1/p} = 2^{-1/p}$ for $p > 0$, this example says also that (2.9) does not hold for general positive semidefinite $A$.

**Problem 2.12.** It is also interesting to consider the limit of $(A^pBA^p)^{1/p}$ as $p \to \infty$ for $A, B \in \mathbb{M}_n^+$, a version different from the limit treated in Theorem 2.1. To consider $\lim_{p \to \infty} (A^pBA^p)^{1/p}$, we may assume without loss of generality that $B$ is an orthogonal projection $E$ (see the argument around (3.2) below). Since $(A^pEA^p)^{1/p} = (A^pE^{2p}A^p)^{1/p}$ converges as $p \to \infty$ by [6, Theorem 2.5], the existence of the limit $\lim_{p \to \infty} (A^pBA^p)^{1/p}$ follows. But it seems that the description of the limit is a combinatorial problem much more complicated than that in Theorem 2.1.

3 \quad \lim_{p \to \infty} (A^p\sigma B)^{1/p}$ for operator means $\sigma$

In theory of operator means due to Kubo and Ando [13], a main result says that each operator mean $\sigma$ is associated with a non-negative operator monotone function $f$ on $[0, \infty)$ with $f(1) = 1$ in such a way that

$$A\sigma B = A^{1/2}f(A^{-1/2}BA^{-1/2})A^{1/2}$$

for $A, B \in \mathbb{M}_n^+$ with $A > 0$, which is further extended to general $A, B \in \mathbb{M}_n^+$ as

$$A\sigma B = \lim_{\varepsilon \searrow 0} (A + \varepsilon I)\sigma(B + \varepsilon I).$$

We write $\sigma_f$ for the operator mean associated with $f$ as above. For $0 \leq \alpha \leq 1$, the operator mean corresponding to the function $x^\alpha$ ($x \geq 0$) is the *weighted geometric mean* $\#_\alpha$, i.e.,

$$A\#_\alpha B = A^{1/2}(A^{-1/2}BA^{-1/2})^\alpha A^{1/2}$$

for $A, B \in \mathbb{M}_n^+$ with $A > 0$. In particular, $\# = \#_{1/2}$ is the so-called *geometric mean* first introduced by Pusz and Woronowicz [15].

The transpose of $f$ above is given by

$$\tilde{f}(x) := x f(x^{-1}), \quad x > 0,$$

which is again an operator monotone function on $[0, \infty)$ (after extending to $[0, \infty)$ by continuity) corresponding to the transposed operator mean of $\sigma_f$, i.e., $A\sigma_f B = B\sigma_f A$. We also write

$$\tilde{f}(x) := \begin{cases} 
\tilde{f}(x^{-1}) = f(x)/x & \text{if } x > 0, \\
0 & \text{if } x = 0.
\end{cases} \quad (3.1)$$

In the rest of the section, let $f$ be such an operator monotone function as above and $\sigma_f$ be the corresponding operator mean. We are concerned with the existence and
the description of the limit \( \lim_{p \to \infty} (A^p \sigma f B)^{1/p} \), in particular, \( \lim_{p \to \infty} (A^p \#_\alpha B)^{1/p} \) for \( A, B \in \mathbb{M}_n^+ \). For this, we may assume without loss of generality that \( B \) is an orthogonal projection. Indeed, let \( E \in A, B \) \( \lambda < \mu \) such that \( \lambda E \leq B \leq \mu E \). Thanks to monotonicity and positive homogeneity of \( \sigma_f \), we have

\[
\lambda(A^p \sigma f E) = (\lambda A^p) \sigma f (\lambda E) \leq A^p \sigma f B \leq (\mu A^p) \sigma f (\mu E) = \mu(A^p \sigma f E).
\]

Hence, for every \( p \geq 1 \), since \( x^{1/p} (x \geq 0) \) is operator monotone,

\[
\lambda^{1/p}(A^p \sigma f E)^{1/p} \leq (A^p \sigma f B)^{1/p} \leq \mu^{1/p}(A^p \sigma f B), \tag{3.2}
\]

so that \( \lim_{p \to \infty} (A^p \sigma f B)^{1/p} \) exists if and only if \( \lim_{p \to \infty} (A^p \sigma f E)^{1/p} \) does, and in this case, both limits are equal. In particular, when \( B > 0 \), since \( (A^p \sigma f I)^{1/p} = \tilde{f}(A^p)^{1/p} \), we note that

\[
\lim_{p \to \infty} (A^p \sigma f B)^{1/p} = \tilde{f}^{(\infty)}(A)
\]

whenever \( \tilde{f}^{(\infty)}(x) := \lim_{p \to \infty} \tilde{f}(x^p)^{1/p} \) exists for all \( x \geq 0 \). For instance,

- if \( f(x) = 1 - \alpha + \alpha x \) where \( 0 \leq \alpha < 1 \), then \( \sigma_f = \nabla_\alpha \), the \( \alpha \)-arithmetic mean \( A \nabla_\alpha B := (1 - \alpha) A + \alpha B \), and \( \tilde{f}^{(\infty)}(x) = \max\{x, 1\} \),
- if \( f(x) = x^\alpha \) where \( 0 \leq \alpha \leq 1 \), then \( \sigma_f = \#_\alpha \) and \( \tilde{f}^{(\infty)}(x) = \tilde{f}(x) = x^{1-\alpha} \),
- if \( f(x) = x/((1-\alpha)x + \alpha) \) where \( 0 < \alpha < 1 \), then \( \sigma_f = !_\alpha \), the \( \alpha \)-harmonic mean \( A !_\alpha B := (A^{-1} \nabla_\alpha B)^{-1} \), and \( \tilde{f}^{(\infty)}(x) = \min\{x, 1\} \).

But it is unknown to us that, for any operator monotone function \( f \) on \([0, \infty)\), the limit \( \lim_{p \to \infty} f(x^p)^{1/p} \) exists for all \( x \geq 0 \), while it seems so.

When \( E \) is an orthogonal projection, the next proposition gives a nice expression for \( A \sigma f E \). This was shown in [11, Lemma 4.7], while the proof is given here for the convenience of the reader.

**Lemma 3.1.** Assume that \( f(0) = 0 \). If \( A \in \mathbb{M}_n \) is positive definite and \( E \in \mathbb{M}_n \) is an orthogonal projection, then

\[
A \sigma f E = \hat{f}(EA^{-1}E), \tag{3.3}
\]

where \( \hat{f} \) is given in (3.1).

**Proof.** For every \( m = 1, 2, \ldots \) we have

\[
A^{-1/2}(EA^{-1}E)^m A^{-1/2} = (A^{-1/2}E)A^{-1/2})^{m+1}. \tag{3.4}
\]

Note that the eigenvalues of \( EA^{-1}E \) and those of \( A^{-1/2}E A^{-1/2} \) are the same including multiplicities. Choose a \( \delta > 0 \) such that the positive eigenvalues of \( EA^{-1}E \) and
\(A^{-1/2}EA^{-1/2}\) are included in \([\delta, \delta^{-1}]\). Then, since \(\hat{f}(x)\) is continuous on \([\delta, \delta^{-1}]\), one can choose a sequence of polynomials \(p_k(x)\) with \(p_k(0) = 0\) such that \(p_k(x) \to \hat{f}(x)\) uniformly on \([\delta, \delta^{-1}]\) as \(n \to \infty\). By (3.3) we have
\[
A^{-1/2}p_k(EA^{-1}E)A^{-1/2} = A^{-1/2}EA^{-1/2}p_k(A^{-1/2}EA^{-1/2})
\]
for every \(k\). Since \(\hat{f}(0) = 0\) by definition, we have
\[
p_k(EA^{-1}E) \to \hat{f}(EA^{-1}E)
\]
and
\[
A^{-1/2}EA^{-1/2}p_k(A^{-1/2}EA^{-1/2}) \to A^{-1/2}EA^{-1/2} \hat{f}(A^{-1/2}EA^{-1/2})
\]
as \(k \to \infty\). Since \(f(0) = 0\) by assumption, we have \(f(x) = x\hat{f}(x)\) for all \(x \in [0, \infty)\). This implies that
\[
A^{-1/2}EA^{-1/2} \hat{f}(A^{-1/2}EA^{-1/2}) = f(A^{-1/2}EA^{-1/2}).
\]
Therefore,
\[
A^{-1/2} \hat{f}(EA^{-1}E)A^{-1/2} = f(A^{-1/2}EA^{-1/2})
\]
so that we have \(\hat{f}(EA^{-1}E) = A^{1/2}f(A^{-1/2}EA^{-1/2})A^{1/2} = A\sigma f\), as asserted. \(\square\)

Formula (3.3) can equivalently be written as
\[
A\sigma f = \tilde{f}((EA^{-1}E)^{-1}),
\]
where \((EA^{-1}E)^{-1}\) is the inverse restricted to \(\text{ran } E\) (in the sense of the generalized inverse) and \(\tilde{f}((EA^{-1}E)^{-1})\) is also restricted to \(\text{ran } E\). In particular, if \(f\) is symmetric (i.e., \(f = \tilde{f}\)) with \(f(0) = 0\), then
\[
A\sigma f = f((EA^{-1}E)^{-1}).
\]

**Example 3.2.** Assume that \(0 < \alpha \leq 1\) and \(A, E\) are as in Lemma 3.1.

1. When \(f(x) = x^\alpha\) and \(\sigma_f = \#_\alpha\), \(\hat{f}(x) = x^{\alpha-1}\) for \(x > 0\) so that
\[
A\#_\alpha E = (EA^{-1}E)^{\alpha-1},
\]
where the \((\alpha-1)\)-power in the right-hand side is defined with restriction to \(\text{ran } E\).

2. When \(f(x) = x/(1 - \alpha)x + \alpha\) and \(\sigma_f = \!\alpha\), \(\hat{f}(x) = (1 - \alpha + \alpha x)^{-1}\) for \(x > 0\) so that
\[
A\!\alpha E = \{(1 - \alpha)E + \alpha EA^{-1}E\}^{-1} = \{E((1 - \alpha)I + \alpha A^{-1}E)^{-1}\}^{-1},
\]
where the inverse of \(E((1 - \alpha)I + \alpha A^{-1}E)\) in the right-hand side is restricted to \(\text{ran } E\).
(3) When \( f(x) = (x - 1)/\log x \) and so \( \sigma_f \) is the logarithmic mean, \( \tilde{f}(x) = (1 - x^{-1})/\log x \) for \( x > 0 \) so that

\[
A \sigma_f E = (E - (EA^{-1}E)^{-1})(\log EA^{-1}E)^{-1},
\]

where the right-hand side is defined with restriction to \( \text{ran } E \).

**Theorem 3.3.** Assume that \( f(0) = 0 \) and \( f(x^r) \geq f(x)^r \) for all \( x > 0 \) and all \( r \in (0, 1) \). Let \( A \in \mathbb{M}_n^+ \) and \( E \in \mathbb{M}_n \) be an orthogonal projection. Then

\[
(A^p \sigma_f E)^{1/p} \geq (A^q \sigma_f E)^{1/q} \quad \text{if } 1 \leq p < q. \tag{3.6}
\]

**Proof.** First, note that \( \tilde{f}(x^r) = x^r f(x^{-r}) \geq x^r f(x^{-1})^r = \tilde{f}(x)^r \) for all \( x > 0, r \in (0, 1) \). By replacing \( A \) with \( A + \varepsilon I \) and taking the limit as \( \varepsilon \to 0 \), we may assume that \( A \) is positive definite. Let \( 1 \leq p < q \) and \( r := p/q \in (0, 1) \). By (3.5) we have

\[
(A^q \sigma_f E)^r = \tilde{f}((EA^{-q}E)^{-1}) \leq \tilde{f}((EA^{-q}E)^{-r}). \tag{3.7}
\]

Since \( x^r \) is operator monotone on \([0, \infty)\), we have by Hansen’s inequality \([9]\)

\[
(EA^{-q}E)^r \geq EA^{-qr}E = EA^{-p}E
\]

so that \( (EA^{-q}E)^{-r} \leq (EA^{-p}E)^{-1} \). Since \( \tilde{f}(x) \) is operator monotone on \([0, \infty)\), we have

\[
\tilde{f}((EA^{-q}E)^{-r}) \leq \tilde{f}((EA^{-p}E)^{-1}) = A^p \sigma_f E. \tag{3.8}
\]

Combining (3.7) and (3.8) gives

\[
(A^q \sigma_f E)^r \leq A^p \sigma_f E.
\]

Since \( x^{1/p} \) is operator monotone on \([0, \infty)\), we finally have

\[
(A^q \sigma_f E)^{1/q} \leq (A^p \sigma_f E)^{1/p}.
\]

**Corollary 3.4.** Assume that \( f(0) = 0 \) and \( f(x^r) \geq f(x)^r \) for all \( x > 0, r \in (0, 1) \). Then for every \( A, B \in \mathbb{M}_n^+ \), the limit

\[
\lim_{p \to \infty} (A^p \sigma_f B)^{1/p}
\]

exists.

**Proof.** From the argument around (3.2) we may assume that \( B \) is an orthogonal projection \( E \). Then Theorem 3.3 implies that \( (A^p \sigma_f E)^{1/p} \) converges as \( p \to \infty \).
Remark 3.5. Following [17], an operator monotone function \( f \) on \([0, \infty)\) is said to be power monotone increasing (p.m.i. for short) if \( f(x^r) \geq f(x)^r \) for all \( x > 0, r > 1 \) (equivalently, \( f(x^r) \leq f(x)^r \) for all \( x > 0, r \in (0,1) \)), and power monotone decreasing (p.m.d.) if \( f(x^r) \leq f(x)^r \) for all \( x > 0, r > 1 \). These conditions play a role to characterize the operator means \( \sigma_f \) satisfying Ando-Hiai’s inequality [1], see [17] Lemmas 2.1, 2.2. For instance, the p.m.d. condition is satisfied for \( f \) in (1) and (2) of Example 3.2 while \( f \) in Example 3.2(3) does the p.m.i. condition. Hence, for any \( \alpha \in [0,1] \), \((A^p\#\alpha E)^{1/p}\) and \((A^p\sigma_f E)^{1/p}\) converge decreasingly as \( 1 \leq p \nearrow \infty \). In fact, for the harmonic mean, we have the limit \( A \wedge B := \lim_{p \to \infty} (A^p B^p)^{1/p} \), the decreasing limit as \( 1 \leq p \nearrow \infty \) for any \( A, B \geq 0 \), which is the infimum counterpart of \( A \vee B \) in [12] (see also Example 2.7). The reader might be wondering if the opposite inequality to (3.6) holds (i.e., \((A^p f E)^{1/p}\) is increasing as \( 1 \leq p \nearrow \infty \)) when \( f \) satisfies the p.m.i. condition. Although this is the case when \( \sigma = \nabla_\alpha \) the weighted arithmetic mean, it is not the case in general. In fact, if it were true, \((A^p\#\alpha E)^{1/p}\) must be constant for \( p \geq 1 \) since \( x^\alpha \) satisfies both p.m.i. and p.m.d. conditions, that is impossible.

Finally, for the weighted geometric mean \( \#_\alpha \) we obtain the explicit description of \( \lim_{p \to \infty} (A^p\#\alpha E)^{1/p} \) for any \( A \in \mathbb{M}_n^+ \). For the trivial cases \( \alpha = 0,1 \) note that \((A^p\#0 E)^{1/p} = A\) and \((A^p\#1 E)^{1/p} = E\) for all \( p > 0 \).

Theorem 3.6. Assume that \( 0 < \alpha < 1 \). Let \( A \in \mathbb{M}_n^+ \) be given with the spectral decomposition \( A = \sum_{k=1}^m a_k P_k \) where \( a_1 > \cdots > a_m > 0 \), and \( E \in \mathbb{M}_n \) be an orthogonal projection. Then

\[
\lim_{p \to \infty} (A^p\#_\alpha E)^{1/p} = \sum_{k=1}^m a_k^{1-\alpha} Q_k, \tag{3.9}
\]

where

\[
Q_1 := P_1 \wedge E,
Q_k := (P_1 + \cdots + P_k) \wedge E - (P_1 + \cdots + P_{k-1}) \wedge E, \quad 2 \leq k \leq m.
\]

Proof. First, assume that \( A \) is positive definite so that \( P_1 + \cdots + P_m = I \). When \( f(x) = x^\alpha \) with \( 0 < \alpha < 1 \), formula (3.5) is given as

\[
A\#_\alpha E = (EA^{-1}E)^{-(1-\alpha)}.
\]

Since

\[
\lim_{p \to \infty} (A^p \#_\alpha E)^{1/p} = \lim_{p \to \infty} (EA^{-p}E)^{-(1-\alpha)/p} = \lim_{p \to \infty} (E(A^{1-\alpha}E)^{-p}E)^{-1/p},
\]

it follows from Theorem 2.9 that

\[
\lim_{p \to \infty} (A^p \#_\alpha E)^{1/p} = \sum_{k=1}^m a_k^{1-\alpha} P_k \bar{M}_k.
\]
where
\[ \mathcal{M}_k := \text{ran } E(P_k + \cdots + P_{m})E \ominus \text{ran } E(P_{k+1} + \cdots + P_{m})E, \quad 1 \leq k \leq m - 1, \]
\[ \mathcal{M}_m := \text{ran } EP_mE. \]

From Lemma 3.7 below we have
\[ \mathcal{M}_1 = \text{ran } E \ominus \text{ran } EP_1^\perp E = \text{ran } P_1 \wedge E, \]
and for \( 2 \leq k \leq m, \)
\[ \mathcal{M}_k = \text{ran } (P_1 + \cdots + P_{k-1})^\perp E \ominus \text{ran } (P_1 + \cdots + P_k)^\perp E \]
\[ = [\text{ran } E \ominus \text{ran } (P_1 + \cdots + P_k)^\perp E] \ominus [\text{ran } E \ominus \text{ran } (P_1 + \cdots + P_{k-1})^\perp E] \]
\[ = \text{ran } (P_1 + \cdots + P_k) \wedge E \ominus \text{ran } (P_1 + \cdots + P_{k-1}) \wedge E \]
\[ = \text{ran } [(P_1 + \cdots + P_k) \wedge E - (P_1 + \cdots + P_{k-1}) \wedge E]. \]

Therefore, (3.9) is established when \( A \) is positive definite.

Next, when \( A \) is not positive definite, let \( P_{m+1} := (P_1 + \cdots + P_m)^\perp \). For any \( \varepsilon \in (0, a_m) \) define \( A_\varepsilon := A + \varepsilon P_{m+1} \). Then the above case implies that
\[ \lim_{p \to \infty} (A_\varepsilon A_\varepsilon^* E)^{1/p} = \sum_{k=1}^{m} a_k^{1-\alpha} Q + \varepsilon^{1-\alpha} Q_{m+1}, \]
where
\[ Q_{m+1} := E - (P_1 + \cdots + P_m) \wedge E. \]

Assume that \( 0 < \varepsilon < \varepsilon' < a_m \). For every \( p \geq 1 \), since \( A_\varepsilon^p \leq A_{\varepsilon'}^p \), we have \( A_\varepsilon^p A_\varepsilon^* E \leq A_{\varepsilon'}^p A_{\varepsilon'}^* E \) and hence \( (A_{\varepsilon}^p A_{\varepsilon}^* E)^{1/p} \leq (A_{\varepsilon'}^p A_{\varepsilon'}^* E)^{1/p} \). Furthermore, since \( A_\varepsilon^p A_\varepsilon^* E \to A^p A^* E \) as \( a_m > \varepsilon \wedge 0 \), we have
\[ (A^p A^* E)^{1/p} = \lim_{a_m > \varepsilon \wedge 0} (A_{\varepsilon}^p A_{\varepsilon}^* E)^{1/p} \text{ decreasingly. } (3.10) \]

Now, we can perform a calculation of limits as follows:
\[ \lim_{1 \leq p \to \infty} (A^p A^* E)^{1/p} = \lim_{1 \leq p \to \infty} \lim_{a_m > \varepsilon \wedge 0} (A_{\varepsilon}^p A_{\varepsilon}^* E)^{1/p} \]
\[ = \lim_{a_m > \varepsilon \wedge 0} \lim_{1 \leq p \to \infty} (A_{\varepsilon}^p A_{\varepsilon}^* E)^{1/p} \]
\[ = \lim_{a_m > \varepsilon \wedge 0} \left( \sum_{k=1}^{m} a_k^{1-\alpha} Q + \varepsilon^{1-\alpha} Q_{m+1} \right) \]
\[ = \sum_{k=1}^{m} a_k^{1-\alpha} Q_k. \]

In the above, the second equality (the exchange of two limits) is confirmed as follows. Let \( X_{p,\varepsilon} := (A_{\varepsilon}^p A_{\varepsilon}^* E)^{1/p} \) for \( p \geq 1 \) and \( 0 < \varepsilon < a_m \). Then \( X_{p,\varepsilon} \) is decreasing as
1 \leq p \to \infty \text{ by Theorem 3.3} \text{ and also decreasing as } a_m > \varepsilon \to 0 \text{ as seen in (3.10). Let } X_p := \lim_{\varepsilon} X_{p,\varepsilon} (= (A^p \&_\alpha E)^{1/p}), X_{\varepsilon} := \lim_{p} X_{p,\varepsilon}, \text{ and } X := \lim_{p} X_p. \text{ Since } X_{p,\varepsilon} \geq X_p, \text{ we have } X_{\varepsilon} \geq X \text{ and hence } \lim_{\varepsilon} X_{\varepsilon} \geq X. \text{ On the other hand, since } X_{p,\varepsilon} \geq X_{\varepsilon}, \text{ we have } X_p \geq \lim_{\varepsilon} X_{\varepsilon} \text{ and hence } X \geq \lim_{\varepsilon} X_{\varepsilon}. \text{ Therefore, } X = \lim_{\varepsilon} X_{\varepsilon}, \text{ which gives the assertion.}

In particular, when \( A = P \) is an orthogonal projection, we have \((A^p \&_\alpha E)^{1/p} = P \& E\) for all \( p > 0 \) (see [13, Theorem 3.7]) so that both sides of (3.9) are certainly equal to \( P \& E \).

**Lemma 3.7.** For every orthogonal projections \( E \) and \( P \),

\[
\text{ran } EP^\perp E = \text{ran } (E - P \& E),
\]

or equivalently,

\[
\text{ran } P \& E = \text{ran } E \ominus \text{ran } EP^\perp E.
\]

**Proof.** According to the well-known representation of two projections (see [16 pp. 306–308]), we write

\[
E = I \oplus 0 \oplus \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \oplus 0,
\]

\[
P = I \oplus 0 \oplus I \oplus \begin{bmatrix} C^2 & CS \\ CS & S^2 \end{bmatrix} \oplus 0,
\]

where \( 0 < C, S < I \) with \( C^2 + S^2 = I \). We have

\[
P \& E = I \oplus 0 \oplus 0 \oplus 0 \oplus 0.
\]

Since

\[
P^\perp = 0 \oplus I \oplus 0 \oplus \begin{bmatrix} S^2 & -CS \\ -CS & C^2 \end{bmatrix} \oplus I
\]

we also have

\[
EP^\perp E = 0 \oplus I \oplus 0 \oplus \begin{bmatrix} S^2 & 0 \\ 0 & 0 \end{bmatrix} \oplus 0,
\]

whose range is that of

\[
0 \oplus I \oplus 0 \oplus \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} \oplus 0 = E - P \& E,
\]

which yields the conclusion.

**Acknowledgments**

This work was supported by JSPS KAKENHI Grant Number JP17K05266.
References

[1] T. Ando and F. Hiai, Log majorization and complementary Golden-Thompson type inequalities, *Linear Algebra Appl.* 197/198 (1994), 113–131.

[2] J. Antezana, G. Corach and D. Stojanoff, Spectral shorted matrices, *Linear Algebra Appl.* 381 (2004), 197–217.

[3] J. Antezana, G. Corach and D. Stojanoff, Spectral shorted operators, *Integral Equations Operator Theory* 55 (2006), 169–188.

[4] H. Araki, On an inequality of Lieb and Thirring, *Lett. Math. Phys.* 19 (1990), 167–170.

[5] K. M. R. Audenaert and F. Hiai, On matrix inequalities between the power means: counterexamples, *Linear Algebra Appl.* 439 (2013), 1590–1604.

[6] K. M. R. Audenaert and F. Hiai, Reciprocal Lie-Trotter formula, *Linear and Multilinear Algebra* 64 (2016), 1220–1235.

[7] J.-C. Bourin, Convexity or concavity inequalities for Hermitian operators, *Math. Ineq. Appl.* 7 (2004), 607–620.

[8] N. Datta and F. Leditzky, A limit of the quantum Rényi divergence, *J. Phys. A: Math. Theor.* 47 (2014), 045304.

[9] F. Hansen, An operator inequality, *Math. Ann.* 246 (1980), 249–250.

[10] F. Hiai, Equality cases in matrix norm inequalities of Golden-Thompson type, *Linear and Multilinear Algebra* 36 (1994), 239–249.

[11] F. Hiai, A generalization of Araki’s log-majorization, *Linear Algebra Appl.* 501 (2016), 1–16.

[12] T. Kato, Spectral order and a matrix limit theorem, *Linear and Multilinear Algebra* 8 (1979), 15–19.

[13] F. Kubo and T. Ando, Means of positive linear operators, *Math. Ann.* 246 (1980), 205–224.

[14] M. Müller-Lennert, F. Dupuis, O. Szehr, S. Fehr and M. Tomamichel, On quantum Rényi entropies: A new generalization and some properties, *J. Math. Phys.* 54 (2013), 122203.

[15] W. Pusz and S. L. Woronowicz, Functional calculus for sesquilinear forms and the purification map, *Rep. Math. Phys.* 8 (1975), 159–170.
[16] M. Takesaki, *Theory of Operator Algebras I*, Encyclopaedia of Mathematical Sciences, Vol. 124, Springer-Verlag, Berlin, 2002.

[17] S. Wada, Some ways of constructing Furuta-type inequalities, *Linear Algebra Appl.* 457 (2014), 276–286.

[18] M. M. Wilde, A. Winter and D. Yang, Strong converse for the classical capacity of entanglement-breaking and Hadamard channels via a sandwiched Rényi relative Entropy, *Comm. Math. Phys.* 331 (2014), 593–622.