

# SATURATING THE DATA PROCESSING INEQUALITY FOR $\alpha - z$ RÉNYI RELATIVE ENTROPY

Sarah Chehade  
University of Houston  
sschehad@cougarnet.uh.edu \*

(Received 2020)

It has been shown that the  $\alpha - z$  Rényi relative entropy satisfies the Data Processing Inequality (DPI) for a certain range of  $\alpha$ 's and  $z$ 's. Moreover, the range is completely characterized by Zhang in '20. We prove necessary and algebraically sufficient conditions to saturate the DPI for the  $\alpha - z$  Rényi relative entropy whenever  $1 < \alpha \leq 2$  and  $\frac{\alpha}{2} \leq z \leq \alpha$ . Moreover, these conditions coincide whenever  $\alpha = z$ .

**Keywords:** concave/convex functionals, Pick functions, data processing inequality, quantum relative entropy.

## 1. Introduction

Statistical distinguishability between two states is a central concept in Quantum Information Theory. One basic distinguishability measure, the quantum relative entropy, was introduced by Umegaki in 1962 in his paper [38] about non-commutative conditional expectations. When two states pass through a noisy quantum channel, it is indeed more challenging to measure this distinguishability. This phenomena is described as the data processing inequality. In order for a distinguishability measure to have any operational meaning, it must satisfy the data processing inequality. Petz [31, 30] proved this inequality in the context of von Neumann Algebras. More information about this relative entropy is found in section 2.1.1.. Since then, once generalizations of the quantum relative entropies were defined, the question of whether the data processing inequality holds or not (and for which parameters) generated several publications such as [31, 30, 6, 2, 8, 9, 4, 20, 21, 18, 19, 38, 43, 15, 7, 10]. One quantum generalization of the quantum relative entropy is called the  $\alpha$ -Rényi relative entropy, and it was proven by Petz in [31, 30] and also by Bény, Mosonyi, et.al in [10], that this entropy measure indeed satisfies the data processing inequality. More information about this relative entropy is found in section 2.1.2.. Later, a different generalization was introduced by Müller-Lennert, Dupuis, Szehr, Fehr, and Tomamichel in [28], and independently by Wilde, Winter, and Yang in [41]. This family is called the  $\alpha$ -Sandwiched Rényi relative entropy. Under certain parameters of  $\alpha$ , this relative entropy also satisfies the data processing inequality. Lastly, Audenaert and Datta, in their paper [2], introduced a two parameter family of relative entropies that generalizes all entropy functions stated thus far. This family of entropies,

---

National Science Foundation (NSF) grant DMS-1812734.

the  $\alpha - z$  Rényi relative entropy, was completely characterized in terms of its two parameters,  $\alpha$  and  $z$ , as to when it satisfies the data processing inequality and when it does not. Reference [6] gives a nice intuitive summary of the contributions to the data processing inequality, and [43] finishes it up with the final characterizations of the parameters.

It is known that Lindblad and Uhlmann in [24] and [37] respectively proved that satisfying the data processing inequality is equivalent to proving convexity or concavity of certain trace functionals within the definitions of the relative entropy functions. This is a crucial ingredient in most of results on data processing. In fact, working with these trace functionals is just as important when answering the questions of whether it is possible to saturate these inequalities.

Our interest lies in the question of saturating the DPI. i.e., when is the relative entropy preserved when states pass through a noisy quantum channel. For some of the relative entropies above, the answer to this question is in terms of recoverability of states. Recoverability exists for the quantum relative entropy, the Rényi relative entropy, and the  $\alpha$ -Sandwiched Rényi relative entropy. This work contributes to the question of recoverability in terms of the  $\alpha - z$  Rényi relative entropy.

After this paper first appeared on the arxiv, another paper [42] appeared asking the same question: when is the DPI saturated for the  $\alpha - z$  Rényi relative entropy? In [42], the authors answer this question in a different format than it is presented here. More work is needed to compare the results.

The main result of this paper says:

COROLLARY 1.

Let  $\rho \in \mathcal{D}(\mathcal{H})$ ,  $\sigma \in \mathcal{Q}(\mathcal{H})$ , and  $\Lambda : \mathcal{B}(\mathcal{H}) \rightarrow \mathcal{B}(\mathcal{K})$  be a quantum channel. For any  $1 < \alpha \leq 2$  and  $\frac{\alpha}{2} \leq z \leq \alpha$ , whenever saturation of the DPI holds, i.e.,  $D_{\alpha,z}(\rho||\sigma) = D_{\alpha,z}(\Lambda(\rho)||\Lambda(\sigma))$ , then the states satisfy

$$\sigma^{\frac{1-z}{2z}} (\sigma^{\frac{1-\alpha}{2z}} \rho^{\frac{\alpha}{z}} \sigma^{\frac{1-\alpha}{2z}})^{z-1} \sigma^{\frac{1-z}{2z}} = \Lambda^* \left( \Lambda(\sigma)^{\frac{1-z}{2z}} (\Lambda(\sigma)^{\frac{1-\alpha}{2z}} \rho^{\frac{\alpha}{z}} \Lambda(\sigma)^{\frac{1-\alpha}{2z}})^{z-1} \Lambda(\sigma)^{\frac{1-z}{2z}} \right). \quad (1)$$

The paper is arranged as follows: section 2, discusses the definitions and notations used throughout this paper. In addition to this, some known results and properties about the different quantum relative entropies of interest are mentioned. In section 3, the technical results using tools in complex analysis or results about convex/concave trace functionals are discussed. Section 4 is dedicated to the main result in the context of partial traces followed by the more general consequences. Finally, section 5 concludes with a brief discussion on closing remarks.

## 2. Notations and Definitions

Throughout this paper, only finite-dimensional Hilbert spaces are considered. When  $\mathcal{H}_{AB}$  is written, it is understood to mean a tensor product of Hilbert spaces  $\mathcal{H}_A$  and  $\mathcal{H}_B$ . For a Hilbert space  $\mathcal{H}$ , let  $\mathcal{B}(\mathcal{H})$  denote the set of bounded linear operators on  $\mathcal{H}$ . The set of all positive operators is denoted by

$$\mathcal{P}(\mathcal{H}) := \{A \in \mathcal{B}(\mathcal{H}) : A > 0\},$$

and the space of all *density operators* is defined to be

$$\mathcal{D}(\mathcal{H}) := \{\rho \in \mathcal{P}(\mathcal{H}) : \text{Tr}(\rho) = 1\}.$$

Recall that if an operator  $\rho$  is positive ( $\rho > 0$ ), then it is automatically self adjoint, i.e.,  $\rho^* = \rho$ . Given any linear operator  $\mathcal{L} : \mathcal{B}(\mathcal{H}) \rightarrow \mathcal{B}(\mathcal{K})$ , where  $\mathcal{H}$  and  $\mathcal{K}$  are Hilbert spaces, the adjoint

operator  $\mathcal{L}^* : \mathcal{B}(\mathcal{K}) \rightarrow \mathcal{B}(\mathcal{H})$  is the unique operator satisfying

$$\langle \mathcal{L}(X), Y \rangle_{\mathcal{B}(\mathcal{K})} = \langle X, \mathcal{L}^*(Y) \rangle_{\mathcal{B}(\mathcal{H})},$$

and the inner product here is the Hilbert-Schmidt inner product defined as

$$\langle X, Y \rangle_{HS} = \text{Tr}(X^*Y).$$

In general, the Schatten  $p$ - norm is defined as

$$\|X\|_p := \left( \text{Tr} \left[ (X^*X)^{\frac{p}{2}} \right] \right)^{\frac{1}{p}},$$

for  $p \in [1, \infty)$ . Note that this norm satisfies the Hölder inequality, sub-multiplicativity, and monotonicity in  $p$ . A linear operator  $\mathcal{L} : \mathcal{B}(\mathcal{H}) \rightarrow \mathcal{B}(\mathcal{K})$  is said to be  $n$ -positive if

$$1_n \otimes \mathcal{L} : \mathcal{B}(\mathbb{C}^n) \otimes \mathcal{B}(\mathcal{H}) \rightarrow \mathcal{B}(\mathbb{C}^n) \otimes \mathcal{B}(\mathcal{K})$$

is a positive operator, where  $1_n$  is the identity operator on  $\mathcal{B}(\mathbb{C}^n)$ .

DEFINITION 1. If  $1_n \otimes \mathcal{L}$  is positive for all  $n \in \mathbb{N}$ , then  $\mathcal{L}$  is called a *completely positive map*.

DEFINITION 2. Completely positive maps that also preserve the trace of operators are called *quantum channels*. i.e.,  $\text{Tr}(\mathcal{L}(\rho)) = \text{Tr}(\rho)$ .

## 2.1. Quantum Relative Entropies

### 2.1.1. Umegaki Relative Entropy

In quantum information theory, the information shared between states is regularly studied through the understanding of quantum entropies. Umegaki relative entropy, also known as the quantum relative entropy, in [38] is defined as

$$D(\rho||\sigma) := \text{Tr}(\rho \log \rho - \rho \log \sigma),$$

where  $\rho \in \mathcal{D}(\mathcal{H})$  and  $\sigma \in \mathcal{D}(\mathcal{H})$ , provided that  $\text{supp}(\rho) \subseteq \text{supp}(\sigma)$ . Otherwise, the relative entropy between  $\rho$  and  $\sigma$  is said to be  $\infty$ . Reference [6] provides an extensive review of the formulation of this relative entropy, applications, and some of its properties. In this setting, [38] introduces and explains the DPI. That is, for any quantum channel  $\Lambda$ , the following inequality holds

$$D(\rho||\sigma) \geq D(\Lambda(\rho)||\Lambda(\sigma)).$$

This inequality is interpreted as an increased difficulty in distinguishing states from one another after the states pass through a noisy quantum channel.

DEFINITION 3. If there is a quantum channel  $\Psi$  such that  $\Psi$  recovers states  $\rho$  and  $\sigma$ , i.e.,

$$(\Psi \circ \Lambda)\rho = \rho \text{ and } (\Psi \circ \Lambda)\sigma = \sigma, \tag{2}$$

then we say that  $\Lambda$  is *sufficient* for states  $\rho$  and  $\sigma$ . When this happens, the quantum channel  $\Psi$  is called a *recovery map*.

Saturation of the DPI was originally proven by Petz in the context of von Neumann algebras in [31, 30]. The result states that 2 states  $\rho$  and  $\sigma$  saturate the DPI for a quantum channel  $\Lambda$  if and only if the quantum channel is sufficient for these states. The recovery map  $\Psi_{\sigma, \Lambda}$ , known as the Petz recovery map has an explicit form of

$$\Psi_{\sigma, \Lambda}(\cdot) = \sigma^{\frac{1}{2}} \Lambda^* \left( \Lambda(\sigma)^{-\frac{1}{2}} \cdot \Lambda(\sigma)^{-\frac{1}{2}} \right) \sigma^{\frac{1}{2}}.$$

The map  $\Psi$  is indexed by  $\sigma$  and  $\Lambda$  to indicate that the recovery map depends on this state and this quantum channel. In a different context, you can saturate the DPI with the use of an error term, which was done in

**COROLLARY 2.** [9][Corollary 1.7] For  $\rho$  and  $\sigma$  density operators and  $\mathcal{N}$  a partial trace, the following inequality holds

$$D(\rho||\sigma) - D(\mathcal{N}(\rho)||\mathcal{N}(\sigma)) \geq \left(\frac{\pi}{8}\right)^4 \|\rho^{-1}\|^{-2} \|\mathcal{R}_\rho(\mathcal{N}(\sigma)) - \sigma\|_1^4,$$

where  $\|\cdot\|_1$  is the Schatten 1-norm,  $\|X\|_1 := \text{Tr}|X| = \text{Tr}(X^*X)^{\frac{1}{2}}$ , for an operator  $X$  and  $\mathcal{R}_\rho$  is a Petz recovery map.

### 2.1.2. $\alpha$ -Rényi Relative Entropy ( $\alpha$ -RRE)

One of the first generalizations of the Umegaki relative entropy is defined for  $\alpha \in (-\infty, 1) \cup (1, \infty)$ . The  $\alpha$ -Rényi Relative Entropy is expressed as

$$D_\alpha(\rho||\sigma) := \frac{1}{\alpha - 1} \log \text{Tr}(\rho^\alpha \sigma^{1-\alpha}),$$

provided  $\text{supp}(\rho) \subseteq \text{supp}(\sigma)$ . For  $\alpha \in [0, 1) \cup (1, 2]$ , Theorem 5.1 of [10] proves equality of the DPI if and only if there exists a recovery map that recovers both states  $\rho$  and  $\sigma$  perfectly well. Furthermore, [10] also provides the algebraic necessary and sufficient conditions for the  $\alpha$ -Rényi relative entropy as well. That is for all  $\alpha \in [0, 1) \cup (1, 2]$ , saturation of the DPI is satisfied if and only if

$$\Lambda^*(\Lambda(\sigma)^{-z} \Lambda(\rho)^{-z}) = \sigma^{-z} \rho^z,$$

for all  $z \in \mathbb{C}$ .<sup>1</sup> These proofs are in a more general class of quantum functionals called quantum f-divergence, which is actually a class of quantum quasi-entropies. In fact in [8], similar claims are made however in a different context and with the use of an error bound. Once the bound is proven, necessity and sufficiency follow very easily. Their result is

**THEOREM 1.** [8][Theorem 6.1] For any  $\alpha \in (0, 1)$ , under explicit assumptions defined in the paper for  $\rho$  and  $\sigma$ , the following inequality holds

$$D_\alpha(\rho||\sigma) - D_\alpha(\Lambda(\rho)||\Lambda(\sigma)) \geq \frac{1}{1 - \alpha} \log \left( 1 + K \|\Lambda(\sigma)^{\frac{1}{2}} \Lambda(\rho)^{-\frac{1}{2}} \rho^{\frac{1}{2}} - \sigma^{\frac{1}{2}}\|_2^{6-2\alpha} \right),$$

where  $K$  is a constant calculated in their paper.

---

<sup>1</sup>It is understood that when  $\alpha = 1$  or the limit as  $\alpha$  approaches 1, this is the Umegaki relative entropy case.

### 2.1.3. $\alpha$ – Sandwiched Rényi Divergence ( $\alpha$ –SSRD)

This section describes another way to generalize the Umegaki relative entropy and mentions some known results as well. For  $\alpha \in (-\infty, 1) \cup (1, \infty)$ , [28] and [41] introduce a new family of Rényi relative entropies called *Sandwiched Rényi Relative Entropies* or *Sandwiched Rényi Divergence*. They are defined as

$$\tilde{D}_\alpha(\rho||\sigma) := \frac{1}{\alpha - 1} \log \left( \text{Tr} \left[ \sigma^{\frac{1-\alpha}{2\alpha}} \rho \sigma^{\frac{1-\alpha}{2\alpha}} \right]^\alpha \right),$$

provided that  $\text{supp}(\rho) \subseteq \text{supp}(\sigma)$ . The DPI for all  $\alpha \in [\frac{1}{2}, 1) \cup (1, \infty)$  was proven in [4]. That is

$$\tilde{D}_\alpha(\rho||\sigma) \geq \tilde{D}_\alpha(\Lambda(\rho)||\Lambda(\sigma)).$$

In [21], it was shown that the equality of the DPI is satisfied for all  $\alpha > \frac{1}{2}$ , if and only if states  $\rho$  and  $\sigma$  have the following algebraic form:

$$\sigma^{\frac{1-\alpha}{2\alpha}} (\sigma^{\frac{1-\alpha}{2\alpha}} \rho \sigma^{\frac{1-\alpha}{2\alpha}})^{\alpha-1} \sigma^{\frac{1-\alpha}{2\alpha}} = \Lambda^* \left( \Lambda(\sigma)^{\frac{1-\alpha}{2\alpha}} (\Lambda(\sigma)^{\frac{1-\alpha}{2\alpha}} \rho \Lambda(\sigma)^{\frac{1-\alpha}{2\alpha}})^{\alpha-1} \Lambda(\sigma)^{\frac{1-\alpha}{2\alpha}} \right).$$

For  $\alpha > 1$ , the equivalence between saturating the DPI and the sufficiency property is proven in [20]. The techniques use non-commutative interpolated  $L_p$  spaces for von Neumann Algebras. The same author proved the equivalence between saturating the DPI and the sufficiency for  $\alpha \in (\frac{1}{2}, 1)$  in [?] using different norms.

### 2.1.4. $\alpha - z$ Rényi Relative Entropy ( $\alpha - z$ RRE)

Here is another generalization of relative entropy that combines both  $\alpha$ -RRE and  $\alpha$ -SRD. Let  $\rho, \sigma \in \mathcal{D}(\mathcal{H})$  with  $\alpha \in \mathbb{R} \setminus \{1\}$  and  $z > 0$ . The  $\alpha - z$  Rényi relative entropy was introduced in [2] and is defined as

$$D_{\alpha,z}(\rho||\sigma) := \frac{1}{\alpha - 1} \log \left( \text{Tr} \left[ \left( \sigma^{\frac{1-\alpha}{2z}} \rho^{\frac{\alpha}{z}} \sigma^{\frac{1-\alpha}{2z}} \right)^z \right] \right), \quad (3)$$

provided the  $\text{supp}(\rho) \subseteq \text{supp}(\sigma)$ . Otherwise, the  $\alpha - z$  Rényi relative entropy is said to be  $+\infty$ . When  $z = 1$ , the  $\alpha - z$  RRE reduces to the  $\alpha$ -RRE. When  $z = \alpha$ , the  $\alpha - z$  RRE reduces to the  $\alpha$ -SRD. When  $z = 1$  and  $\alpha \rightarrow 1$  or when  $z = \alpha$  and  $\alpha \rightarrow 1$ , the  $\alpha - z$  RRE reduces to the Umegaki relative entropy.

Many of the interesting properties of the  $\alpha - z$  entropies are explained and introduced in previous works on this family of entropies such as [2, 6]. Only the properties that are explicitly used in this paper are listed here.

1. *Invariance:* The  $\alpha - z$  Rényi entropies are invariant under unitaries. That is for any unitary  $U$ ,

$$D_{\alpha,z}(U\rho U^*||U\sigma U^*) = D_{\alpha,z}(\rho||\sigma).$$

This is because for any unitary  $U$  and for any operator  $A$ , the eigenvalues of  $A$  and  $UAU^*$  are the same.

2. *Tensor Property:* For any  $\rho, \sigma, \tau \in \mathcal{D}(\mathcal{H})$ ,

$$D_{\alpha,z}(\rho \otimes \tau||\sigma \otimes \tau) = D_{\alpha,z}(\rho||\sigma).$$

This is due to the fact that the trace of a tensor product between two states is the product of trace of states.

REMARK 1. In this paper, it is always assumed that the operators,  $\rho$  and  $\sigma$ , are invertible and that  $\text{supp}(\rho) \subseteq \text{supp}(\sigma)$ .

The conjecture for which parameters of  $\alpha$  and  $z$  the DPI holds is outlined in [6] and finally concluded in [43]. This is summarized in the next theorem.

THEOREM 2. [43][Theorem 1.2] *The  $\alpha - z$  Rényi relative entropy is monotone under completely positive trace preserving maps (quantum channels) on  $\mathcal{D}(\mathcal{H})$  for all  $\mathcal{H}$  if and only if one of the following holds*

1.  $0 < \alpha < 1$  and  $z \geq \max\{\alpha, 1 - \alpha\}$ ;
2.  $1 < \alpha \leq 2$  and  $\frac{\alpha}{2} \leq z \leq \alpha$ ;
3.  $2 \leq \alpha < \infty$  and  $\alpha - 1 \leq z \leq \alpha$ .

One way to prove this is through the relationship between the DPI and joint convexity/concavity of the trace functional defined by the map

$$(A, B) \mapsto \text{Tr}(B^{\frac{q}{2}} K^* A^p K B^{\frac{q}{2}})^s, \quad (4)$$

where  $A$  and  $B$  are positive operators on  $\mathcal{H}$ ,  $K$  is any operator in  $\mathcal{B}(\mathcal{H})$ ,  $p, q > 0$ , and  $s \geq \frac{1}{p+q}$ . Here, the case of interest is whenever  $K$  is the identity operator,  $p = \frac{\alpha}{z}$ ,  $q = \frac{1-\alpha}{z}$ ,  $s = \frac{1}{p+q}$ ,  $A = \rho$ , and  $B = \sigma$ . Then the trace functional is defined as

$$\Psi_{\alpha,z}(\rho||\sigma) := \text{Tr} \left[ \left( \sigma^{\frac{1-\alpha}{2z}} \rho^{\frac{\alpha}{z}} \sigma^{\frac{1-\alpha}{2z}} \right)^z \right].$$

The next theorem describes the relationship between the DPI and joint convexity and joint concavity of the trace functional.

THEOREM 3. [6][Proposition 7] *Let  $\alpha, z > 0$  with  $\alpha \neq 1$ . Then  $D_{\alpha,z}$  is monotone under quantum channels on  $\mathcal{P}(\mathcal{H})$  (for any finite dimensional  $\mathcal{H}$ ) if and only if one of the following holds:*

1.  $\alpha < 1$  and  $\Psi_{\alpha,z}(\rho||\sigma)$  is jointly concave.
2.  $\alpha > 1$  and  $\Psi_{\alpha,z}(\rho||\sigma)$  is jointly convex.

Together, Theorem 2 and Theorem 3 give the complete picture for data processing inequality of the  $\alpha - z$  Rényi Relative Entropy.

The next two sections are the technical components used to prove the main result:

### 3. Preliminaries

#### 3.1. Tools from Complex Analysis

Let  $\text{Spec}(X)$  denote the set of eigenvalues for operator  $X$ . If  $\Omega$  is an open subset of  $\mathbb{C}$  such that  $\text{Spec}(X) \subseteq \Omega \subseteq \mathbb{C}$ , then the analytic functional calculus is used to conclude that  $F(X)$  is well defined, for any analytic function  $F$ . Define

$$\mathbb{C}^+ := \{z \in \mathbb{C} \text{ s.t. } \text{Im}(z) > 0\}.$$

Note that  $\mathbb{C}^+$  is an open subset of  $\mathbb{C}$ . For any operator  $X$ , define

$$\operatorname{Re}(X) := \frac{X + X^*}{2} \text{ and } \operatorname{Im}(X) := \frac{X - X^*}{2i}.$$

Then let

$$I_n^+ := \{X \in M_n(\mathbb{C}) \text{ s.t. } \operatorname{Im}(X) > 0\}$$

and

$$I_n^- := \{X \in M_n(\mathbb{C}) \text{ s.t. } \operatorname{Im}(X) < 0\},$$

where  $n < \infty$  denotes the dimension. For  $0 < p \leq 1$ , denote

$$\Gamma_{p\pi} := \{re^{i\theta} : r > 0 \text{ and } 0 < \theta < p\pi\}$$

and

$$\Gamma_{-p\pi} := \{re^{i\theta} : r > 0 \text{ and } 0 > \theta > -p\pi\}.$$

Note that when  $p = 1$ ,  $\Gamma_{p\pi} = \mathbb{C}^+$ . Let us recall a few known facts or results.

LEMMA 1. [17][Lemma 1.1] *If  $X \in I_n^+$ , then  $X$  is invertible and  $\operatorname{Spec}(X) \subset \mathbb{C}^+$ . For  $X$ , an invertible  $n \times n$  matrix,  $X \in I_n^+$  if and only if  $X^{-1} \in I_n^-$ .*

LEMMA 2. [17][Lemma 1.2] *Let  $0 < p \leq 1$ . If  $X \in I_n^+$ , then so is  $X^p$  and  $e^{-ip\pi}X^p \in I_n^-$ . If  $X \in I_n^-$ , then  $X^p \in I_n^-$  and  $e^{ip\pi}X^p \in I_n^+$ .*

LEMMA 3. *Any pair of operators  $A$  and  $B$  have the following properties:*

$$\operatorname{Re}(ABA^*) = A(\operatorname{Re}B)A^* \text{ and } \operatorname{Im}(ABA^*) = A(\operatorname{Im}B)A^*.$$

**Proof:** Observe that

$$\operatorname{Im}(ABA^*) = \frac{(ABA^*) - (ABA^*)^*}{2i} \tag{5}$$

$$= \frac{(ABA^*) - (AB^*A^*)}{2i} \tag{6}$$

$$= \frac{A(B - B^*)A^*}{2i} \tag{7}$$

$$= A(\operatorname{Im}B)A^*, \tag{8}$$

and the real version is similar.  $\square$

LEMMA 4. *If  $B \in I_n^+$ , then so is  $ABA^*$ , for any non-zero operator  $A$ .*

**Proof:** Lemma 3, implies that  $\operatorname{Im}(ABA^*) = A(\operatorname{Im}B)A^*$ . Thus for any  $y$ ,

$$\langle \operatorname{Im}(ABA^*)^*y, y \rangle = \langle A(\operatorname{Im}B)A^*y, y \rangle \tag{9}$$

$$= \langle (\operatorname{Im}B)A^*y, A^*y \rangle \tag{10}$$

$$= \left\langle \left[ (\operatorname{Im}B)^{\frac{1}{2}} \right] A^*y, \left[ (\operatorname{Im}B)^{\frac{1}{2}} \right] A^*y \right\rangle > 0, \tag{11}$$

where the last equality holds because  $\operatorname{Im}(B)$  is positive providing a unique square root that is also positive.  $\square$

REMARK 2. Define

$$\mathcal{I}(\mathcal{H}) := \left\{ X \in \mathcal{P}(\mathcal{H}) : (zX + H)^{\frac{p}{2}} B (zX + H)^{\frac{p}{2}} \in I_n^+ \right\}, \quad (12)$$

for all  $H$  hermitian,  $z \in \mathbb{C}^+$ ,  $0 < p \leq 1$ , and  $B \in \mathcal{P}(\mathcal{H})$ .

PROPOSITION 1. Define  $\Phi_Z(X) := Z^* X Z$ , where  $Z$  is an invertible operator. For any  $0 < p < 1$ , a positive operator  $X$  belonging to any convex subset of  $\mathcal{I}(\mathcal{H})$  is such that the map

$$X \mapsto \text{Tr} \left[ \left\{ \Phi_Z \left( X^{\frac{p}{2}} \right) A \Phi_Z \left( X^{\frac{p}{2}} \right) \right\}^{\frac{1}{p}} \right] \quad (13)$$

is concave for any  $A \in \mathcal{P}(\mathcal{H})$ .

**Proof:** It suffices to show that for any  $H$  hermitian,

$$\frac{d^2}{dx^2} \left( \text{Tr} \left[ \left\{ \Phi_Z \left( (X + xH)^{\frac{p}{2}} \right) A \Phi_Z \left( (X + xH)^{\frac{p}{2}} \right) \right\}^{\frac{1}{p}} \right] \right) \leq 0, \quad (14)$$

for any small  $x > 0$ . This is because if (14) holds, then

$$(X + xH) \mapsto \text{Tr} \left[ \left\{ \Phi_Z \left( (X + xH)^{\frac{p}{2}} \right) A \Phi_Z \left( (X + xH)^{\frac{p}{2}} \right) \right\}^{\frac{1}{p}} \right]$$

is concave, and hence the proposition is proved by taking  $x \rightarrow 0$ .

Observe that for any  $z \in \mathbb{C}^+$ , it follows that  $zX + H \in I_n^+$ . So  $(zX + H)^{\frac{p}{2}}$  is well defined, by the analytic functional calculus. Moreover, by Lemma 2, observe that  $(zX + H)^{\frac{p}{2}} \in I_n^+$ . By linearity of  $\Phi_Z$  and by Lemma 4,  $\Phi_Z \left( (zX + H)^{\frac{p}{2}} \right) \in I_n^+$ . Define

$$F(z) := \Phi_Z \left( (zX + H)^{\frac{p}{2}} \right) A \Phi_Z \left( (zX + H)^{\frac{p}{2}} \right). \quad (15)$$

This function is analytic in  $\mathbb{C}^+$  because a product of analytic functions is analytic. To prove equation (14), we prove the following steps:

1. Show  $\text{Spec}(F(z))$  is contained in some open subset of  $\mathbb{C}$  so that  $(F(z))^{\frac{1}{p}}$  is well defined.
2. Extend  $F(z)$  onto the real line  $(R, \infty)$ .
3. Express  $\text{Tr} \left[ (F(z))^{\frac{1}{p}} \right]$  as a Pick function to admit an integral representation.
4. Show concavity of the above expression through its integral representation.

Step 1: We show  $\text{Spec}(F(z)) \subset \Gamma_{p\pi}$ . As in [17], it suffices to prove the following 3 properties:

1.  $\text{Spec}(F(z)) \subset \Gamma_{p\pi}$ , when  $z = re^{i\theta}$  with fixed  $0 < \theta < \pi$  and sufficiently large  $r > 0$ .
2.  $\text{Spec}(F(z)) \cap [0, \infty) = \emptyset$  for all  $z \in \mathbb{C}^+$ .
3.  $\text{Spec}(F(z)) \cap \{re^{ip\pi}\} = \emptyset$  for all  $z \in \mathbb{C}^+$ .



These statements are sufficient because if, for the sake of contradiction,  $\text{Spec}(F(z)) \not\subset \Gamma_{p\pi}$  for some  $z_o = r_o e^{i\theta_o} \in \mathbb{C}^+$  and the 3 properties are satisfied, then by continuity of eigenvalues of  $F(z)$  and by property (1),  $\text{Spec}(F(z)) \cup \partial\Gamma_{p\pi} \neq \emptyset$ , for some  $z \in \{r e^{i\theta_o} : r > r_o\}$ , which implies either property (2) or property (3) is violated.

To prove property (1), by linearity of  $z$  in  $F(z)$ ,

$$F(z) = z^p \left[ \Phi_Z \left( (X + z^{-1}H)^{\frac{p}{2}} \right) A \Phi_Z \left( (X + z^{-1}H)^{\frac{p}{2}} \right) \right].$$

Whenever  $z = r e^{i\theta_o}$  with a fixed  $0 < \theta_o < \pi$ , note that

$$\text{Spec} \left[ \Phi_Z \left( (X + z^{-1}H)^{\frac{p}{2}} \right) A \Phi_Z \left( (X + z^{-1}H)^{\frac{p}{2}} \right) \right]$$

converges to

$$\text{Spec} \left[ \Phi_Z \left( (X)^{\frac{p}{2}} \right) A \Phi_Z \left( (X)^{\frac{p}{2}} \right) \right] \subset (0, \infty), \text{ as } r \rightarrow \infty.$$

To prove property (2), for any  $0 \leq r < \infty$ ,

$$F(z) - r1_n = \Phi_Z \left( (zX + H)^{\frac{p}{2}} \right) (A - r\Phi_Z((zX + H)^{p/2})^{-2}) \Phi_Z \left( (zX + H)^{\frac{p}{2}} \right). \quad (16)$$

By the assumption of  $X$  in any convex subset of  $\mathcal{I}(\mathcal{H})$  and by Lemma 1,

$$(zX + H)^{-\frac{p}{2}} (Z^{-1})^* (Z^{-1}) (zX + H)^{-\frac{p}{2}} \in I_n^-,$$

and by Lemma 3,

$$\Phi_{(Z^{-1})^*} \left[ (zX + H)^{-\frac{p}{2}} (Z^{-1})^* (Z^{-1}) (zX + H)^{-\frac{p}{2}} \right] \in I_n^-.$$

Note that  $[\Phi_Z((zX + H)^{p/2})]^{-2} = \Phi_{(Z^{-1})^*} [(zX + H)^{-\frac{p}{2}} (Z^{-1})^* (Z^{-1}) (zX + H)^{-\frac{p}{2}}] \in I_n^-$ . This implies that  $A - r [\Phi_Z((zX + H)^{p/2})]^{-2} \in I_n^+$ , and thus  $F(z) - r1_n$  is invertible by Lemma 1. If there exists  $l \in \text{Spec}(F(z)) \cap [0, \infty)$ , then  $F(z) - l1_n = 0$ , which would contradict invertability for all  $r \in [0, \infty)$ . Hence property (2) is satisfied. Proving property (3) is very similar to proving the second property using the assumption from  $X$  in any convex subset of  $\mathcal{I}(\mathcal{H})$  as well as Lemma 2. With this and using the analytic functional calculus,  $(F(z))^{\frac{1}{p}}$  is well defined.

Step 2: For every  $z \in \mathbb{C}^+$  such that  $|z| > R$ , define an analytic function

$$\hat{F}(z) := z^p \Phi_Z \left( (X + z^{-1}H)^{\frac{p}{2}} \right) A \Phi_Z \left( (X + z^{-1}H)^{\frac{p}{2}} \right). \quad (17)$$

Then continuously extending this function to the real line, for every  $x \in (R, \infty)$  gives  $\hat{F}(x) = F(x)$ . Hence for every  $z \in \mathbb{C}^+$  such that  $|z| > R$ , write

$$F(z) = z^p \Phi \left( (X + z^{-1}H)^{\frac{p}{2}} \right) A \Phi \left( (X + z^{-1}H)^{\frac{p}{2}} \right),$$

and

$$(F(z))^{\frac{1}{p}} = z \left\{ \Phi_Z \left( (X + z^{-1}H)^{\frac{p}{2}} \right) A \Phi_Z \left( (X + z^{-1}H)^{\frac{p}{2}} \right) \right\}^{\frac{1}{p}}.$$

Step 3: Given that  $\text{Tr} \left[ (F(z))^{\frac{1}{p}} \right] \in \mathbb{C}^+$  for every  $z \in \mathbb{C}^+$ , and  $\text{Tr} \left[ (F(x))^{\frac{1}{p}} \right] \in \mathbb{R}$  for every  $x \in (R, \infty)$ , by the Schwarz Reflection Principle,  $(F(z))^{\frac{1}{p}}$  can be extended to the lower half plane, that is the set of complex number with negative imaginary parts. And thus,  $\phi$  is a Pick function on  $\mathbb{C} \setminus (-\infty, R)$ , where  $\phi(x) = \text{Tr} \left[ (F(x))^{\frac{1}{p}} \right]$  for all  $x \in (R, \infty)$ . Then

$$x\phi(x^{-1}) = \text{Tr} \left\{ \Phi_Z \left( (X + xH)^{\frac{p}{2}} \right) A \Phi_Z \left( (X + xH)^{\frac{p}{2}} \right) \right\}^{\frac{1}{p}}, \quad (18)$$

for every  $x \in (0, R^{-1})$ . By theory of Pick functions, see [5], every Pick function  $\phi$  admits an integral representation.

$$\phi(z) = a + bz + \int_{-\infty}^{\infty} \frac{1 + tz}{t - z} d\nu(t), \quad (19)$$

where  $a \in \mathbb{R}$ ,  $b \geq 0$ , and  $\nu$  is a finite measure on  $\mathbb{R}$ . The measure  $\nu$  is supported in  $(-\infty, R]$  because  $\phi$  is analytically continued across  $(R, \infty)$ .

Step 4: For all  $x \in (0, R^{-1})$ ,

$$x\phi(x^{-1}) = x \left( a + \frac{b}{x} + \int_{-\infty}^{\infty} \frac{1 + tx^{-1}}{t - x^{-1}} d\nu(t) \right) \quad (20)$$

$$= ax + b + \int_{-\infty}^R \frac{x(x+t)}{tx-1} d\nu(t), \quad (21)$$

with

$$\frac{d^2}{dx^2} \left( ax + b + \int_{-\infty}^R \frac{x(x+t)}{tx-1} d\nu(t) \right) = \int_{-\infty}^R \frac{d^2}{dx^2} \left( \frac{x(x+t)}{tx-1} \right) d\nu(t) \quad (22)$$

$$= \int_{-\infty}^R \left( \frac{2(t^2+1)}{(xt-1)^3} \right) d\nu(t) < 0, \quad (23)$$

for all  $x \in (0, R^{-1})$  and all  $t \in (-\infty, R)$ . □

### 3.2. Convex and Concave Trace Functionals

To prove joint convexity of  $f_{\alpha,z}(H, \rho, \sigma)$  from equation (30) below, the next few results are needed. Recall the equation from (4)

$$(A, B) \mapsto \text{Tr}(B^{\frac{q}{2}} K^* A^p K B^{\frac{q}{2}})^s. \quad (24)$$

**THEOREM 4.** [43][Theorem 1.1] Fix any invertible matrix  $K$ . Suppose that  $p \geq q$  and  $s > 0$ .

1. If  $0 \leq q \leq p \leq 1$  and  $0 < s \leq \frac{1}{p+q}$ , then the map from (24) is jointly concave.
2. If  $-1 \leq q \leq p \leq 0$  and  $s > 0$ , then the map from (24) is jointly convex.
3. If  $-1 \leq q \leq 0$ ,  $1 \leq p \leq 2$ ,  $(p, q) \neq (1, -1)$  and  $s \geq \frac{1}{p+q}$ , then the map from (24) is jointly convex.

**THEOREM 5.** [43][Theorem 3.3] For  $r_i > 0$ ,  $i \in \{0, 1, 2\}$  such that  $\frac{1}{r_0} = \frac{1}{r_1} + \frac{1}{r_2}$ , one has that for any invertible  $X, Y \in \mathcal{B}(\mathcal{H})$  that

$$\mathrm{Tr}|XY|^{r_1} = \max \left\{ \frac{r_1}{r_0} \mathrm{Tr}|XZ|^{r_0} - \frac{r_1}{r_2} \mathrm{Tr}|Y^{-1}Z|^{r_2} : Z \in \mathcal{B}(\mathcal{H}) \text{ and invertible} \right\}. \quad (25)$$

**PROPOSITION 2.** [6][Proposition 5] For a fixed operator  $B$ , the map on positive operators

$$A \mapsto \mathrm{Tr} \left[ (B^* A^p B)^{\frac{1}{p}} \right]$$

1. is concave for  $0 \leq p \leq 1$ , with  $p \neq 0$ .
2. is convex for  $1 \leq p \leq 2$ , with  $p \neq 0$ .

**PROPOSITION 3.** If  $f : \mathcal{D}(\mathcal{H}) \times \mathcal{D}(\mathcal{H}) \times \mathcal{P}(\mathcal{H}) \mapsto [0, \infty)$  is defined as

$$f(A, B, H) := g(A, H) + h(B, H),$$

where  $g$  and  $h$  are continuous, the functional  $g$  is convex in  $A$  and the functional  $h$  is convex in  $B$ , then  $f$  is jointly convex in  $(A, B)$ . Moreover,  $\sup_{H>0} \{f(A, B, H)\}$  is jointly convex in  $(A, B)$  whenever  $f$  is.

**Proof:** For all  $i$  such that  $0 \leq \lambda_i \leq 1$ , with  $\sum_i \lambda_i = 1$ , by convexity of  $g$  and  $h$ ,

$$g \left( \sum_i \lambda_i A_i, H \right) \leq \sum_i \lambda_i g(A_i, H) \text{ and } h \left( \sum_i \lambda_i B_i, H \right) \leq \sum_i \lambda_i h(B_i, H).$$

Thus

$$f \left( \sum_i \lambda_i A_i, \sum_i \lambda_i B_i, H \right) = g \left( \sum_i \lambda_i A_i, H \right) + h \left( \sum_i \lambda_i B_i, H \right) \quad (26)$$

$$\leq \sum_i \lambda_i g(A_i, H) + \sum_i \lambda_i h(B_i, H) \quad (27)$$

$$= \sum_i \lambda_i (g(A_i, H) + h(B_i, H)) \quad (28)$$

$$= \sum_i \lambda_i f(A_i, B_i, H), \quad (29)$$

as desired. □

Recall that

$$\Psi_{\alpha, z}(\rho || \sigma) := \mathrm{Tr} \left[ \left( \sigma^{\frac{1-\alpha}{2z}} \rho^{\frac{\alpha}{z}} \sigma^{\frac{1-\alpha}{2z}} \right)^z \right].$$

**LEMMA 5.** Let  $\rho$  and  $\sigma \in \mathcal{P}(\mathcal{H})$ , and assume that  $\alpha > 1$  and  $z > 1$ . For any positive operator  $H$ , define

$$f_{\alpha, z}(H, \rho, \sigma) := z \mathrm{Tr} \left( \sigma^{\frac{z-\alpha}{2z}} \rho^{\frac{\alpha}{z}} \sigma^{\frac{z-\alpha}{2z}} H \right) - (z-1) \mathrm{Tr} \left[ \left( \sigma^{\frac{z-1}{2z}} H \sigma^{\frac{z-1}{2z}} \right)^{\frac{z}{z-1}} \right]. \quad (30)$$

Then

$$\Psi_{\alpha, z}(\rho || \sigma) = \sup_{H>0} f_{\alpha, z, H}(\rho, \sigma), \quad (31)$$

where the supremum is achieved whenever  $H = \sigma^{\frac{1-z}{2z}} (\sigma^{\frac{1-\alpha}{2z}} \rho^{\frac{\alpha}{z}} \sigma^{\frac{1-\alpha}{2z}})^{z-1} \sigma^{\frac{1-z}{2z}}$ .

**Proof:** For  $X$  and  $Y \in \mathcal{P}(\mathcal{H})$ , it follows that  $\text{Tr}(XY) = \text{Tr}\left(X^{\frac{1}{2}}YX^{\frac{1}{2}}\right)$ , which is positive because  $X^{\frac{1}{2}}YX^{\frac{1}{2}}$  is a positive operator. For a choice of  $1 \leq p, q \leq \infty$ , such that  $1 = \frac{1}{p} + \frac{1}{q}$ ,

$$0 \leq \text{Tr}(XY) = |\text{Tr}(XY)| \quad (32)$$

$$\leq \text{Tr}|XY| \quad (33)$$

$$\leq (\text{Tr}(X^p))^{\frac{1}{p}} (\text{Tr}(Y^q))^{\frac{1}{q}} \quad (34)$$

$$\leq \frac{1}{p} \text{Tr}(X^p) + \frac{1}{q} \text{Tr}(Y^q) \quad (35)$$

where (33) is standard for operators, (34) follows from Theorem 1.1 in [25] for positive operators  $X$  and  $Y$ , and (35) follows from the standard Young's inequality. Note from Theorem 1.1 in [25] that line (34) is saturated if and only if  $X^p = Y^q$  which also implies equality of line (35) as well. Moreover,  $X = Y^{\frac{q}{p}}$  also implies that

$$\text{Tr}|XY| = \text{Tr}|Y^{\frac{q}{p}}Y| \quad (36)$$

$$= \text{Tr}|Y^{\frac{q+p}{p}}| \quad (37)$$

$$= \text{Tr}(Y^q) \quad (38)$$

$$= \text{Tr}(Y^{\frac{q}{p}}Y) \quad (39)$$

$$= \text{Tr}(XY), \quad (40)$$

so that (32) is also saturated whenever  $X^p = Y^q$ . Thus,

$$p \text{Tr}(XY) - \frac{p}{q} \text{Tr}(Y)^q \leq \text{Tr}X^p. \quad (41)$$

Take positive operators

$$X = (\sigma^{\frac{1-\alpha}{2z}} \rho^{\frac{\alpha}{z}} \sigma^{\frac{1-\alpha}{2z}}) \text{ and } Y = (\sigma^{\frac{z-1}{2z}} H \sigma^{\frac{z-1}{2z}}),$$

where  $H$  is some positive operator in  $\mathcal{B}(\mathcal{H})$ . Then it follows that (41) becomes

$$p \text{Tr}(\sigma^{\frac{1-\alpha}{2z}} \rho^{\frac{\alpha}{z}} \sigma^{\frac{1-\alpha}{2z}} \sigma^{\frac{z-1}{2z}} H \sigma^{\frac{z-1}{2z}}) - \frac{p}{q} \text{Tr}[(\sigma^{\frac{z-1}{2z}} H \sigma^{\frac{z-1}{2z}})^q] \leq \text{Tr}[(\sigma^{\frac{1-\alpha}{2z}} \rho^{\frac{\alpha}{z}} \sigma^{\frac{1-\alpha}{2z}})^p], \quad (42)$$

where equality happens if and only if

$$\sigma^{\frac{1-z}{2z}} \left( \sigma^{\frac{1-\alpha}{2z}} \rho^{\frac{\alpha}{z}} \sigma^{\frac{1-\alpha}{2z}} \right)^{\frac{p}{q}} \sigma^{\frac{1-z}{2z}} = H.$$

By construction,  $H$  is unique, hence the left hand side of (41) becomes

$$p\mathrm{Tr}\left(\sigma^{\frac{1-\alpha}{2z}}\rho^{\frac{\alpha}{z}}\sigma^{\frac{1-\alpha}{2z}}\sigma^{\frac{z-1}{2z}}\sigma^{\frac{1-z}{2z}}\left(\sigma^{\frac{1-\alpha}{2z}}\rho^{\frac{\alpha}{z}}\sigma^{\frac{1-\alpha}{2z}}\right)^{\frac{p}{q}}\sigma^{\frac{1-z}{2z}}\sigma^{\frac{z-1}{2z}}\right) - \quad (43)$$

$$\frac{p}{q}\mathrm{Tr}\left[\left(\sigma^{\frac{z-1}{2z}}\sigma^{\frac{1-z}{2z}}\left(\sigma^{\frac{1-\alpha}{2z}}\rho^{\frac{\alpha}{z}}\sigma^{\frac{1-\alpha}{2z}}\right)^{\frac{p}{q}}\sigma^{\frac{1-z}{2z}}\sigma^{\frac{z-1}{2z}}\right)^q\right] \quad (44)$$

$$=p\mathrm{Tr}\left(\sigma^{\frac{1-\alpha}{2z}}\rho^{\frac{\alpha}{z}}\sigma^{\frac{1-\alpha}{2z}}\left(\sigma^{\frac{1-\alpha}{2z}}\rho^{\frac{\alpha}{z}}\sigma^{\frac{1-\alpha}{2z}}\right)^{\frac{p}{q}}\right) - \frac{p}{q}\mathrm{Tr}\left[\left(\left(\sigma^{\frac{1-\alpha}{2z}}\rho^{\frac{\alpha}{z}}\sigma^{\frac{1-\alpha}{2z}}\right)^{\frac{p}{q}}\right)^q\right] \quad (45)$$

$$=p\mathrm{Tr}\left(\left(\sigma^{\frac{1-\alpha}{2z}}\rho^{\frac{\alpha}{z}}\sigma^{\frac{1-\alpha}{2z}}\right)^{\frac{p}{q}+1}\right) - \frac{p}{q}\mathrm{Tr}\left[\left(\sigma^{\frac{1-\alpha}{2z}}\rho^{\frac{\alpha}{z}}\sigma^{\frac{1-\alpha}{2z}}\right)^p\right] \quad (46)$$

$$= \left(p - \frac{p}{q}\right)\mathrm{Tr}\left[\left(\sigma^{\frac{1-\alpha}{2z}}\rho^{\frac{\alpha}{z}}\sigma^{\frac{1-\alpha}{2z}}\right)^p\right] \quad (47)$$

$$=\mathrm{Tr}\left[\left(\sigma^{\frac{1-\alpha}{2z}}\rho^{\frac{\alpha}{z}}\sigma^{\frac{1-\alpha}{2z}}\right)^p\right], \quad (48)$$

where  $1 = \frac{1}{p} + \frac{1}{q}$  implies that  $p + q = pq$  which implies  $\frac{p}{q} + 1 = p$  and  $p - \frac{p}{q} = 1$ . For a choice of  $p = z$  and  $q = \frac{z}{z-1}$ , where  $z > 1$  define the left hand side of (42) as

$$f_{\alpha,z}(H, \rho, \sigma) := z\mathrm{Tr}(\sigma^{\frac{z-\alpha}{2z}}\rho^{\frac{\alpha}{z}}\sigma^{\frac{z-\alpha}{2z}}H) - (z-1)\mathrm{Tr}\left[(\sigma^{\frac{z-1}{2z}}H\sigma^{\frac{z-1}{2z}})^{\frac{z}{z-1}}\right]. \quad (49)$$

Then, the desired result is achieved.  $\square$

**PROPOSITION 4.** Let  $1 < \alpha \leq 2$  and  $\frac{\alpha}{2} \leq z \leq \alpha$ . For a fixed  $H > 0$ , one has that  $f_{\alpha,z}(H, \rho, \sigma)$  from equation (30) is jointly convex in  $\rho$  and  $\sigma$ , where  $\sigma$  is in any convex subset of  $\mathcal{I}(\mathcal{H})$  from Remark 2.

**Proof:** Fix  $H > 0$ . Let  $p = \frac{\alpha}{z}$ ,  $q = \frac{\alpha-z}{z}$ ,  $X = \rho^{\frac{p}{2}}$ , and  $Y = \sigma^{-\frac{q}{2}}H^{\frac{1}{2}}$ . Note that

$$\mathrm{Tr}\left(\sigma^{\frac{z-\alpha}{2z}}\rho^{\frac{\alpha}{z}}\sigma^{\frac{z-\alpha}{2z}}H\right) = \mathrm{Tr}\left(H^{\frac{1}{2}}\sigma^{\frac{z-\alpha}{2z}}\rho^{\frac{\alpha}{z}}\sigma^{\frac{z-\alpha}{2z}}H^{\frac{1}{2}}\right) \quad (50)$$

$$= \mathrm{Tr}\left(H^{\frac{1}{2}}\sigma^{\frac{z-\alpha}{2z}}\rho^{\frac{\alpha}{2z}}\rho^{\frac{\alpha}{2z}}\sigma^{\frac{z-\alpha}{2z}}H^{\frac{1}{2}}\right) \quad (51)$$

$$= \mathrm{Tr}\left(H^{\frac{1}{2}}\sigma^{-\frac{q}{2}}\rho^{\frac{p}{2}}\rho^{\frac{p}{2}}\sigma^{-\frac{q}{2}}H^{\frac{1}{2}}\right) \quad (52)$$

$$= \mathrm{Tr}(Y^*X^*XY) \quad (53)$$

$$= \mathrm{Tr}|XY|^2. \quad (54)$$

Since  $1 < z \leq \alpha \leq 2z$ , observe that  $1 \leq p = \frac{\alpha}{z} \leq 2$  and  $0 < q = \frac{\alpha-z}{z} \leq 1$ . Both  $p$  and  $q$  are both positive numbers, so set  $(r_0, r_1, r_2) = (\frac{2}{p}, 2, \frac{2}{q})$ .

Then by Theorem 5,

$$\mathrm{Tr}\left(\sigma^{\frac{z-\alpha}{2z}}\rho^{\frac{\alpha}{z}}\sigma^{\frac{z-\alpha}{2z}}H\right) = \max\left\{p\mathrm{Tr}|\rho^{\frac{p}{2}}Z|^{\frac{2}{p}} - q\mathrm{Tr}|H^{-\frac{1}{2}}\sigma^{\frac{q}{2}}Z|^{\frac{2}{q}}\right\}, \quad (55)$$

where  $Z$  is invertible. Since  $1 \leq p \leq 2$ , it follows from Proposition 2 part 2 that

$$\rho \mapsto \mathrm{Tr}|\rho^{\frac{p}{2}}Z|^{\frac{2}{p}} = \mathrm{Tr}\left[(Z^*\rho^pZ)^{\frac{1}{p}}\right]$$

is convex in  $\rho$ . For the second term in (55), choose  $\Phi_Z(X) := Z^* X Z$ , which is linear, positive, and self adjoint. Let  $A = Z^{-1} H^{-1} (Z^*)^{-1}$ , then by Proposition 1,

$$\mathrm{Tr} |H^{-\frac{1}{2}} \sigma^{\frac{q}{2}} Z|^{\frac{2}{q}} = \mathrm{Tr} \left( \left[ (H^{-\frac{1}{2}} \sigma^{\frac{q}{2}} Z)^* (H^{-\frac{1}{2}} \sigma^{\frac{q}{2}} Z) \right]^{\frac{1}{q}} \right) \quad (56)$$

$$= \mathrm{Tr} \left( \left[ Z^* \sigma^{\frac{q}{2}} H^{-1} \sigma^{\frac{q}{2}} Z \right]^{\frac{1}{q}} \right) \quad (57)$$

$$= \mathrm{Tr} \left( \left[ Z^* \sigma^{\frac{q}{2}} Z Z^{-1} H^{-1} (Z^*)^{-1} Z^* \sigma^{\frac{q}{2}} Z \right]^{\frac{1}{q}} \right) \quad (58)$$

$$= \mathrm{Tr} \left[ \left\{ \Phi_Z \left( \sigma^{\frac{q}{2}} \right) A \Phi_Z \left( \sigma^{\frac{q}{2}} \right) \right\}^{\frac{1}{q}} \right] \quad (59)$$

is concave in  $\sigma$ . Since  $0 < q$ ,

$$\sigma \mapsto -q \mathrm{Tr} |H^{-\frac{1}{2}} \sigma^{-\frac{q}{2}} Z|^{\frac{2}{q}}$$

is convex in  $\sigma$ . By Proposition 3, one concludes that  $\mathrm{Tr} \left( \sigma^{\frac{z-\alpha}{2z}} \rho^{\frac{\alpha}{z}} \sigma^{\frac{z-\alpha}{2z}} H \right)$  is the maximum of a sum of two convex functionals and thus is itself convex in  $\rho$  and  $\sigma$ .

On the other hand, since  $\sigma^{\frac{z-1}{2z}} H \sigma^{\frac{z-1}{2z}}$  and  $H^{\frac{1}{2}} \sigma^{\frac{z-1}{z}} H^{\frac{1}{2}}$  have the same nonzero eigenvalues,

$$\mathrm{Tr} \left( \left[ \sigma^{\frac{z-1}{2z}} H \sigma^{\frac{z-1}{2z}} \right]^{\frac{z}{z-1}} \right) = \mathrm{Tr} \left( \left[ H^{\frac{1}{2}} \sigma^{\frac{z-1}{z}} H^{\frac{1}{2}} \right]^{\frac{z}{z-1}} \right).$$

By Proposition 2 part 1,  $\mathrm{Tr} \left( \left[ \sigma^{\frac{z-1}{2z}} H \sigma^{\frac{z-1}{2z}} \right]^{\frac{z}{z-1}} \right)$  is concave in  $\sigma$ , for all  $z > 1$ . Hence

$$-(z-1) \mathrm{Tr} \left( \left[ \sigma^{\frac{z-1}{2z}} H \sigma^{\frac{z-1}{2z}} \right]^{\frac{z}{z-1}} \right)$$

is convex in  $\sigma$ . As a consequence, by Proposition 3,

$$f_{\alpha,z}(H, \rho, \sigma) = z \mathrm{Tr} \left( \sigma^{\frac{z-\alpha}{2z}} \rho^{\frac{\alpha}{z}} \sigma^{\frac{z-\alpha}{2z}} H \right) - (z-1) \mathrm{Tr} \left( \left[ \sigma^{\frac{z-1}{2z}} H \sigma^{\frac{z-1}{2z}} \right]^{\frac{z}{z-1}} \right)$$

is a sum of two convex functionals in  $\rho$  and  $\sigma$  as desired.  $\square$

#### 4. Main Result

REMARK 3. Let us denote  $\mathcal{Q}(\mathcal{H}_{AB})$  as

$$\mathcal{Q}(\mathcal{H}_{AB}) := \{X_{AB} \in \mathcal{I}(\mathcal{H}_{AB}) : X_A \otimes \pi_B \in \mathcal{I}(\mathcal{H}_{AB})\}, \quad (60)$$

where  $\pi_B$  is the maximally mixed state on Hilbert space  $\mathcal{H}_B$ .

The techniques used here are inspired by [21] and [43]. The main result of this section is a consequence of the following theorem.

THEOREM 6. [Necessary Partial Trace Case] Let  $\rho_{AB} \in \mathcal{D}(\mathcal{H}_{AB})$  and  $\sigma_{AB} \in \mathcal{Q}(\mathcal{H}_{AB})$ . For any  $1 < \alpha \leq 2$  and  $\frac{\alpha}{2} \leq z \leq \alpha$ , whenever saturation of the DPI holds. i.e.,  $D_{\alpha,z}(\rho_{AB} || \sigma_{AB}) = D_{\alpha,z}(\rho_A || \sigma_A)$ , then the states satisfy

$$\sigma_{AB}^{\frac{1-z}{2z}} \left( \sigma_{AB}^{\frac{1-\alpha}{2z}} \rho_{AB}^{\frac{\alpha}{z}} \sigma_{AB}^{\frac{1-\alpha}{2z}} \right)^{z-1} \sigma_{AB}^{\frac{1-z}{2z}} = \sigma_A^{\frac{1-z}{2z}} \left( \sigma_A^{\frac{1-\alpha}{2z}} \rho_A^{\frac{\alpha}{z}} \sigma_A^{\frac{1-\alpha}{2z}} \right)^{z-1} \sigma_A^{\frac{1-z}{2z}}. \quad (61)$$

**Proof:** Assume  $1 < \alpha \leq 2$  and  $\frac{\alpha}{2} \leq z \leq \alpha$ , denote  $d = \dim(\mathcal{H}_B)$ , and define

$$\rho_i := (1 \otimes v_i) \rho_{AB} (1 \otimes v_i^*),$$

where  $\rho_{AB} \in \mathcal{D}(\mathcal{H}_{AB})$  and  $\{v_i\}_{i=1}^{d^2}$  are the generalized Pauli<sup>1</sup> matrices. Similarly, define

$$\sigma_i := (1 \otimes v_i) \sigma_{AB} (1 \otimes v_i^*),$$

where  $\sigma_{AB} \in \mathcal{I}(\mathcal{H}_{AB})$ . Then  $\sigma_i \in \mathcal{I}(\mathcal{H}_{AB})$ , for all  $i$ . Define  $\lambda_i = \frac{1}{d^2}$ , for all  $i = 1, \dots, d^2$ , and let

$$\tilde{\rho} = \sum_{i=1}^{d^2} \lambda_i \rho_i \text{ and } \tilde{\sigma} = \sum_{i=1}^{d^2} \lambda_i \sigma_i.$$

As mentioned in [21],

$$\tilde{\rho} = \sum_{i=1}^{d^2} \lambda_i \rho_i = \rho_A \otimes \pi_B \text{ and } \tilde{\sigma} = \sum_{i=1}^{d^2} \lambda_i \sigma_i = \sigma_A \otimes \pi_B,$$

where  $\pi_B$  is the completely mixed state<sup>2</sup> on  $\mathcal{H}_B$ . i.e.,  $\pi_B = \frac{1}{d}$ . Note that

$$\tilde{\sigma} = \sigma_A \otimes \pi_B \in \mathcal{Q}(\mathcal{H}_{AB}) \subseteq \mathcal{I}(\mathcal{H}_{AB}).$$

Define

$$\bar{H} := \operatorname{argmax}_{H>0} f_{\alpha,z}(H, \tilde{\rho}, \tilde{\sigma}) \text{ and } H_i := \operatorname{argmax}_{H>0} f_{\alpha,z}(H, \tilde{\rho}_i, \tilde{\sigma}_i), \text{ where} \quad (62)$$

$$f_{\alpha,z}(H, \rho, \sigma) := z \operatorname{Tr}(\sigma^{\frac{z-\alpha}{2z}} \rho^{\frac{\alpha}{z}} \sigma^{\frac{z-\alpha}{2z}} H) - (z-1) \operatorname{Tr} \left[ (\sigma^{\frac{z-1}{2z}} H \sigma^{\frac{z-1}{2z}})^{\frac{z}{z-1}} \right], \quad (63)$$

from Lemma 5. Note that the trace functional  $\Psi_{\alpha,z}(\rho||\sigma)$ , mentioned after theorem 2 is proven jointly convex in [43] in a more general setting. The following chain of inequalities says:

$$\Psi_{\alpha,z}(\tilde{\rho}||\tilde{\sigma}) = f_{\alpha,z}(\bar{H}, \tilde{\rho}, \tilde{\sigma}) \quad (64)$$

$$\leq \sum_{i=1}^{d^2} \lambda_i f_{\alpha,z}(\bar{H}, \rho_i, \sigma_i) \quad (65)$$

$$\leq \sum_{i=1}^{d^2} \lambda_i f_{\alpha,z}(H_i, \rho_i, \sigma_i) \quad (66)$$

$$= \sum_{i=1}^{d^2} \lambda_i \Psi_{\alpha,z}(\rho_i||\sigma_i), \quad (67)$$

where lines (64) and (67) are from Lemma 5, line (65) is from the joint convexity of  $f$ , which is proven in Proposition 4, and line (66) is from the fact that  $H_i$  is the maximizer for  $f_{\alpha,z}(H_i, \rho_i, \sigma_i)$  from (62). Assuming saturation of the DPI is equivalent to

$$\Psi_{\alpha,z}(\tilde{\rho}||\tilde{\sigma}) = \sum_{i=1}^{d^2} \lambda_i \Psi_{\alpha,z}(\rho_i||\sigma_i).$$

<sup>1</sup>For more details on the generalized Pauli matrices, see [40] Chapter 3.7 for more details.

<sup>2</sup>See [40] exercise 4.7.6 for an explanation of how the Generalized Pauli operators randomly applied to any density operator with uniform probability give us a maximally mixed state

Then the chain of inequalities above is now a chain of equalities and thus by the definition of  $H_i$ ,

$$f_{\alpha,z}(\bar{H}, \rho_i, \sigma_i) = f_{\alpha,z}(H_i, \rho_i, \sigma_i) \text{ for all } i = 1, \dots, d^2.$$

By the uniqueness of the maximizer  $\bar{H}$ , which is proven above in Lemma 5,  $\bar{H} = H_i$  for all  $i = 1, \dots, d^2$ .

Recall that because an operator  $X$  and  $UXU^*$  have the same eigenvalues, where  $U$  is any unitary, then for any function  $f$  it follows that  $f(UXU^*) = Uf(X)U^*$ . Therefore from Lemma 5, one has an explicit form of the maximizer:  $H_i =$

$$\sigma_i^{\frac{1-z}{2z}} \left( \sigma_i^{\frac{1-\alpha}{2z}} \rho_i^{\frac{\alpha}{z}} \sigma_i^{\frac{1-\alpha}{2z}} \right)^{z-1} \sigma_i^{\frac{1-z}{2z}} \quad (68)$$

$$= [(1 \otimes v_i) \sigma_{AB} (1 \otimes v_i^*)]^{\frac{1-z}{2z}} \quad (69)$$

$$\left( [(1 \otimes v_i) \sigma_{AB} (1 \otimes v_i^*)]^{\frac{1-\alpha}{2z}} [(1 \otimes v_i) \rho_{AB} (1 \otimes v_i^*)]^{\frac{\alpha}{z}} [(1 \otimes v_i) \sigma_{AB} (1 \otimes v_i^*)]^{\frac{1-\alpha}{2z}} \right)^{z-1} \quad (70)$$

$$[(1 \otimes v_i) \sigma_{AB} (1 \otimes v_i^*)]^{\frac{1-z}{2z}} \quad (71)$$

$$= (1 \otimes v_i) \sigma_{AB}^{\frac{1-z}{2z}} \left( \sigma_{AB}^{\frac{1-\alpha}{2z}} \rho_{AB}^{\frac{\alpha}{z}} \sigma_{AB}^{\frac{1-\alpha}{2z}} \right)^{z-1} \sigma_{AB}^{\frac{1-z}{2z}} (1 \otimes v_i^*). \quad (72)$$

This holds for all  $v_i$  due to the fact that  $v_i^* v_i = I$ . Therefore for some  $i \in \{1, \dots, d^2\}$ ,

$$H_i = \sigma_{AB}^{\frac{1-z}{2z}} \left( \sigma_{AB}^{\frac{1-\alpha}{2z}} \rho_{AB}^{\frac{\alpha}{z}} \sigma_{AB}^{\frac{1-\alpha}{2z}} \right)^{z-1} \sigma_{AB}^{\frac{1-z}{2z}}.$$

Also by similar calculations,  $\bar{H} =$

$$(\sigma_A \otimes \pi_B)^{\frac{1-z}{2z}} \left( (\sigma_A \otimes \pi_B)^{\frac{1-\alpha}{2z}} (\rho_A \otimes \pi_B)^{\frac{\alpha}{z}} (\sigma_A \otimes \pi_B)^{\frac{1-\alpha}{2z}} \right)^{z-1} (\sigma_A \otimes \pi_B)^{\frac{1-z}{2z}} \quad (73)$$

$$= \sigma_A^{\frac{1-z}{2z}} \left( \sigma_A^{\frac{1-\alpha}{2z}} \rho_A^{\frac{\alpha}{z}} \sigma_A^{\frac{1-\alpha}{2z}} \right)^{z-1} \sigma_A^{\frac{1-z}{2z}} \otimes \pi_B^{\left( \frac{1-z}{2z} + \left( \frac{1-\alpha}{2z} + \frac{\alpha}{z} + \frac{1-\alpha}{2z} \right) (z-1) + \frac{1-z}{2z} \right)} \quad (74)$$

$$= \sigma_A^{\frac{1-z}{2z}} \left( \sigma_A^{\frac{1-\alpha}{2z}} \rho_A^{\frac{\alpha}{z}} \sigma_A^{\frac{1-\alpha}{2z}} \right)^{z-1} \sigma_A^{\frac{1-z}{2z}} \otimes 1_B, \quad (75)$$

where (75) holds because  $\frac{1-z}{2z} + \left( \frac{1-\alpha}{2z} + \frac{\alpha}{z} + \frac{1-\alpha}{2z} \right) (z-1) + \frac{1-z}{2z} = 0$ . Thus

$$\sigma_{AB}^{\frac{1-z}{2z}} \left( \sigma_{AB}^{\frac{1-\alpha}{2z}} \rho_{AB}^{\frac{\alpha}{z}} \sigma_{AB}^{\frac{1-\alpha}{2z}} \right)^{z-1} \sigma_{AB}^{\frac{1-z}{2z}} = \sigma_A^{\frac{1-z}{2z}} \left( \sigma_A^{\frac{1-\alpha}{2z}} \rho_A^{\frac{\alpha}{z}} \sigma_A^{\frac{1-\alpha}{2z}} \right)^{z-1} \sigma_A^{\frac{1-z}{2z}}. \quad (76)$$

□

Next is the proof of the main result stated as Corollary 1. This is the generalization of the partial case trace using a standard Stinespring Dilation argument.

**Proof:** Following [21], for any quantum channel  $\Lambda : \mathcal{B}(\mathcal{H}) \rightarrow \mathcal{B}(\mathcal{K})$ , by the Stinespring Dilation Theorem <sup>3</sup>, there exists a Hilbert space  $\mathcal{H}'$ , a pure state  $|\tau\rangle \in \mathcal{H}' \otimes \mathcal{K}$ , and a unitary operator  $U : \mathcal{H} \otimes \mathcal{H}' \otimes \mathcal{K} \rightarrow \mathcal{H} \otimes \mathcal{H}' \otimes \mathcal{K}$  such that for every  $\rho \in \mathcal{B}(\mathcal{H})$ , one has

$$\Lambda(\rho) = \text{Tr}_{12}(U(\rho \otimes \tau)U^*),$$

---

<sup>3</sup>Stinespring Dilation Theorem can be found in [40].



where  $\tau = |\tau\rangle\langle\tau|$  and  $\text{Tr}_{12}$  denotes the partial trace over the first two systems  $\mathcal{H} \otimes \mathcal{H}'$  i.e.,  $\text{Tr}_{12} : \mathcal{H} \otimes \mathcal{H}' \otimes \mathcal{K} \rightarrow \mathcal{K}$ . Then, for the parameters <sup>4</sup> where DPI is satisfied for the  $\alpha - z$  RRE,

$$D_{\alpha,z}(\rho||\sigma) = D_{\alpha,z}(U(\rho \otimes \tau)U^*||U(\sigma \otimes \tau)U^*) \quad (77)$$

$$\geq D_{\alpha,z}(\text{Tr}_{12}(U(\rho \otimes \tau)U^*)||\text{Tr}_{12}(U(\sigma \otimes \tau)U^*)) \quad (78)$$

$$= D_{\alpha,z}(\Lambda(\rho)||\Lambda(\sigma)), \quad (79)$$

where (77) is due to properties mentioned in 2.1.4. and (78) is the DPI for partial traces. By assuming equality and by Theorem 6 one sees that

$$1_{\mathcal{H} \otimes \mathcal{H}'} \otimes \Lambda(\sigma)^{\frac{1-z}{2z}} \left( \Lambda(\sigma)^{\frac{1-\alpha}{2z}} \Lambda(\rho)^{\frac{\alpha}{z}} \Lambda(\sigma)^{\frac{1-\alpha}{2z}} \right)^{z-1} \Lambda(\sigma)^{\frac{1-z}{2z}}. \quad (80)$$

$$= [U(\sigma \otimes \tau)U^*]^{\frac{1-z}{2z}} \quad (81)$$

$$\left( [U(\sigma \otimes \tau)U^*]^{\frac{1-\alpha}{2z}} [U(\rho \otimes \tau)U^*]^{\frac{\alpha}{z}} [U(\sigma \otimes \tau)U^*]^{\frac{1-\alpha}{2z}} \right)^{z-1} [U(\sigma \otimes \tau)U^*]^{\frac{1-z}{2z}} \quad (82)$$

$$= U(\sigma^{\frac{1-z}{2z}} \left( \sigma^{\frac{1-\alpha}{2z}} \rho^{\frac{\alpha}{z}} \sigma^{\frac{1-\alpha}{2z}} \right)^{z-1} \sigma^{\frac{1-z}{2z}} \otimes \tau) U^*, \quad (83)$$

where the last line is due to the fact that  $f(UXU^*) = Uf(X)U^*$ , for every function  $f$  and for any unitary  $U$ . The quantum channel  $\Lambda : \mathcal{B}(\mathcal{H}) \rightarrow \mathcal{B}(\mathcal{K})$ , has a unique adjoint,  $\Lambda^* : \mathcal{B}(\mathcal{H}) \rightarrow \mathcal{B}(\mathcal{K})$ , and it is given by

$$\Lambda^*(X) := (1_{\mathcal{H}} \otimes \langle\tau|_{\mathcal{H}' \otimes \mathcal{K}}) U^*(X) U(1_{\mathcal{H}} \otimes |\tau\rangle_{\mathcal{H}' \otimes \mathcal{K}}). \quad (84)$$

So applying (84) to (83) gives

$$\Lambda^* \left[ \Lambda(\sigma)^{\frac{1-z}{2z}} \left( \Lambda(\sigma)^{\frac{1-\alpha}{2z}} \Lambda(\rho)^{\frac{\alpha}{z}} \Lambda(\sigma)^{\frac{1-\alpha}{2z}} \right)^{z-1} \Lambda(\sigma)^{\frac{1-z}{2z}} \right] = \quad (85)$$

$$\Lambda^* \left[ U \left( \sigma^{\frac{1-z}{2z}} \left( \sigma^{\frac{1-\alpha}{2z}} \rho^{\frac{\alpha}{z}} \sigma^{\frac{1-\alpha}{2z}} \right)^{z-1} \sigma^{\frac{1-z}{2z}} \otimes \tau \right) U^* \right] = \quad (86)$$

$$(1_{\mathcal{H}} \otimes \langle\tau|) \left( \sigma^{\frac{1-z}{2z}} \left( \sigma^{\frac{1-\alpha}{2z}} \rho^{\frac{\alpha}{z}} \sigma^{\frac{1-\alpha}{2z}} \right)^{z-1} \sigma^{\frac{1-z}{2z}} \otimes \tau \right) (1_{\mathcal{H}} \otimes |\tau\rangle) = \quad (87)$$

$$\sigma^{\frac{1-z}{2z}} \left( \sigma^{\frac{1-\alpha}{2z}} \rho^{\frac{\alpha}{z}} \sigma^{\frac{1-\alpha}{2z}} \right)^{z-1} \sigma^{\frac{1-z}{2z}}. \quad (88)$$

□

Here, some algebraic sufficient conditions for saturating the DPI of  $\alpha - z$  RRE are explained.

PROPOSITION 5. If  $\sigma_{AB} \in \mathcal{P}(\mathcal{H}_{AB})$  and  $\rho_{AB} \in \mathcal{D}(\mathcal{H}_{AB})$ , where  $\rho_{AB}$  is a product state such that

$$\sigma_{AB}^{\frac{1-\alpha}{2z}} \left( \sigma_{AB}^{\frac{1-\alpha}{2z}} \rho_{AB}^{\frac{\alpha}{z}} \sigma_{AB}^{\frac{1-\alpha}{2z}} \right)^{z-1} \sigma_{AB}^{\frac{1-\alpha}{2z}} = \sigma_A^{\frac{1-\alpha}{2z}} \left( \sigma_A^{\frac{1-\alpha}{2z}} \rho_A^{\frac{\alpha}{z}} \sigma_A^{\frac{1-\alpha}{2z}} \right)^{z-1} \sigma_A^{\frac{1-\alpha}{2z}} \otimes 1_B,$$

then

$$D_{\alpha,z}(\rho_{AB}||\sigma_{AB}) = D_{\alpha,z}(\rho_A||\sigma_A) + k_{\rho},$$

where  $k_{\rho} = \frac{1}{\alpha-1} \log \left[ \text{Tr} \left( \rho_B^{\frac{\alpha}{z}} \right) \right]$ .

---

<sup>4</sup>for  $\alpha$  and  $z$  where the DPI makes sense. see Theorem 2 for such parameters

**Proof:** Since  $\rho_{AB}$  is a product state on  $\mathcal{H}_{AB}$ , it follows that  $\rho_{AB} = \rho_A \otimes \rho_B$ , where  $\rho_A \in \mathcal{D}(\mathcal{H}_A)$  and  $\rho_B \in \mathcal{D}(\mathcal{H}_B)$ . Multiplying the assumed expression by  $\rho_{AB}$  on the left and taking a trace gives

$$\text{Tr} \left( \rho_{AB}^{\frac{\alpha}{z}} \sigma_{AB}^{\frac{1-\alpha}{2z}} \left( \sigma_{AB}^{\frac{1-\alpha}{2z}} \rho_{AB}^{\frac{\alpha}{z}} \sigma_{AB}^{\frac{1-\alpha}{2z}} \right)^{z-1} \sigma_{AB}^{\frac{1-\alpha}{2z}} \right) = \text{Tr} \left( \rho_{AB}^{\frac{\alpha}{z}} \left( \sigma_A^{\frac{1-\alpha}{2z}} \left( \sigma_A^{\frac{1-\alpha}{2z}} \rho_A^{\frac{\alpha}{z}} \sigma_A^{\frac{1-\alpha}{2z}} \right)^{z-1} \sigma_A^{\frac{1-\alpha}{2z}} \otimes 1_B \right) \right).$$

This implies

$$\text{Tr} \left[ \left( \sigma_{AB}^{\frac{1-\alpha}{2z}} \rho_{AB}^{\frac{\alpha}{z}} \sigma_{AB}^{\frac{1-\alpha}{2z}} \right)^z \right] = \text{Tr} \left( \left( \rho_A^{\frac{\alpha}{z}} \otimes \rho_B^{\frac{\alpha}{z}} \right) \sigma_A^{\frac{1-\alpha}{2z}} \left( \sigma_A^{\frac{1-\alpha}{2z}} \rho_A^{\frac{\alpha}{z}} \sigma_A^{\frac{1-\alpha}{2z}} \right)^{z-1} \sigma_A^{\frac{1-\alpha}{2z}} \right),$$

which gives

$$\text{Tr} \left[ \left( \sigma_{AB}^{\frac{1-\alpha}{2z}} \rho_{AB}^{\frac{\alpha}{z}} \sigma_{AB}^{\frac{1-\alpha}{2z}} \right)^z \right] = \text{Tr} \left[ \left( \rho_A^{\frac{\alpha}{z}} \sigma_A^{\frac{1-\alpha}{2z}} \left( \sigma_A^{\frac{1-\alpha}{2z}} \rho_A^{\frac{\alpha}{z}} \sigma_A^{\frac{1-\alpha}{2z}} \right)^{z-1} \sigma_A^{\frac{1-\alpha}{2z}} \right) \otimes \rho_B^{\frac{\alpha}{z}} \right].$$

Taking the log of both sides and multiplying by  $\frac{1}{\alpha-1}$  gives

$$\frac{1}{\alpha-1} \log \left( \text{Tr} \left[ \left( \sigma_{AB}^{\frac{1-\alpha}{2z}} \rho_{AB}^{\frac{\alpha}{z}} \sigma_{AB}^{\frac{1-\alpha}{2z}} \right)^z \right] \right) = \frac{1}{\alpha-1} \log \left( \text{Tr} \left[ \left( \sigma_A^{\frac{1-\alpha}{2z}} \rho_A^{\frac{\alpha}{z}} \sigma_A^{\frac{1-\alpha}{2z}} \right)^z \right] \text{Tr} \left( \rho_B^{\frac{\alpha}{z}} \right) \right),$$

which is the same as

$$D_{\alpha,z}(\rho_{AB} || \sigma_{AB}) = D_{\alpha,z}(\rho_A || \sigma_A) + k_\rho,$$

where  $k_\rho = \frac{1}{\alpha-1} \log \left[ \text{Tr} \left( \rho_B^{\frac{\alpha}{z}} \right) \right]$  as desired. □

REMARK 4. It is interesting to see that Theorem 6 and Proposition 5 would hold simultaneously if and only if  $z = \alpha$ . This in turn will result in  $\alpha$ -SRD, which aligns with the work done in [21].

REMARK 5. If  $\rho_B$  in Proposition 5 is a pure state, then  $k = 0$ . This immediately leads to another result:

PROPOSITION 6. If  $\sigma_{AB} \in \mathcal{P}(\mathcal{H}_{AB})$  and  $\rho_{AB} \in \mathcal{D}(\mathcal{H}_{AB})$ , where  $\rho_{AB}$  is a separable state, such that

$$\sigma_{AB}^{\frac{1-\alpha}{2z}} \left( \sigma_{AB}^{\frac{1-\alpha}{2z}} \rho_{AB}^{\frac{\alpha}{z}} \sigma_{AB}^{\frac{1-\alpha}{2z}} \right)^{z-1} \sigma_{AB}^{\frac{1-\alpha}{2z}} = \sigma_A^{\frac{1-\alpha}{2z}} \left( \sigma_A^{\frac{1-\alpha}{2z}} \rho_A^{\frac{\alpha}{z}} \sigma_A^{\frac{1-\alpha}{2z}} \right)^{z-1} \sigma_A^{\frac{1-\alpha}{2z}},$$

then the DPI under partial traces is saturated. i.e.,

$$D_{\alpha,z}(\rho_{AB} || \sigma_{AB}) = D_{\alpha,z}(\rho_A || \sigma_A).$$

---

<sup>5</sup>Note that we are using the fact that  $\text{Tr}_{AB}(X) = \text{Tr}_A(\text{Tr}_B X)$

**Proof:** Since  $\rho_{AB}$  is a separable state on  $\mathcal{H}_{AB}$ , write  $\rho_{AB}$  as a convex combination of a tensor product of pure states. i.e.,  $\rho_{AB} = \sum_{i \in I} \lambda_i |\psi_i\rangle_A \langle \psi_i| \otimes |\phi_i\rangle_B \langle \phi_i|$ , where  $\{|\psi_i\rangle\}$  and  $\{|\phi_i\rangle\}$  are sets of pure states on  $\mathcal{H}_A$  and  $\mathcal{H}_B$  respectively. Furthermore,  $0 \leq \lambda_i \leq 1$ , for all  $i \in I$  so that  $\sum_{i \in I} \lambda_i = 1$ . Again, multiplying by  $\rho_{AB}$  on the left and taking a trace of the assumed expression yields

$$\text{Tr} \left[ \left( \sigma_{AB}^{\frac{1-\alpha}{2z}} \rho_{AB}^{\frac{\alpha}{z}} \sigma_{AB}^{\frac{1-\alpha}{2z}} \right)^z \right] = \text{Tr} \left( \left( \sum_{i \in I} \lambda_i |\psi_i\rangle_A \langle \psi_i| \otimes |\phi_i\rangle_B \langle \phi_i| \right)^{\frac{\alpha}{z}} \sigma_A^{\frac{1-\alpha}{2z}} \left( \sigma_A^{\frac{1-\alpha}{2z}} \rho_A^{\frac{\alpha}{z}} \sigma_A^{\frac{1-\alpha}{2z}} \right)^{z-1} \sigma_A^{\frac{1-\alpha}{2z}} \right).$$

This implies

$$\begin{aligned} & \text{Tr} \left[ \left( \sigma_{AB}^{\frac{1-\alpha}{2z}} \rho_{AB}^{\frac{\alpha}{z}} \sigma_{AB}^{\frac{1-\alpha}{2z}} \right)^z \right] = \\ & \text{Tr} \left[ \left( \left( \sum_{i \in I} \lambda_i^{\frac{\alpha}{z}} |\psi_i\rangle_A \langle \psi_i| \right) \sigma_A^{\frac{1-\alpha}{2z}} \left( \sigma_A^{\frac{1-\alpha}{2z}} \rho_A^{\frac{\alpha}{z}} \sigma_A^{\frac{1-\alpha}{2z}} \right)^{z-1} \sigma_A^{\frac{1-\alpha}{2z}} \right) \otimes |\phi_i\rangle_B \langle \phi_i| \right], \end{aligned}$$

which gives

$$\text{Tr} \left[ \left( \sigma_{AB}^{\frac{1-\alpha}{2z}} \rho_{AB}^{\frac{\alpha}{z}} \sigma_{AB}^{\frac{1-\alpha}{2z}} \right)^z \right] = \text{Tr} \left[ \left( \rho_A^{\frac{\alpha}{z}} \sigma_A^{\frac{1-\alpha}{2z}} \left( \sigma_A^{\frac{1-\alpha}{2z}} \rho_A^{\frac{\alpha}{z}} \sigma_A^{\frac{1-\alpha}{2z}} \right)^{z-1} \sigma_A^{\frac{1-\alpha}{2z}} \right) \right].$$

Taking the log of both sides and multiplying by  $\frac{1}{\alpha-1}$  gives

$$D_{\alpha,z}(\rho_{AB} || \sigma_{AB}) = D_{\alpha,z}(\rho_A || \sigma_A).$$

□

## 5. Closing Remarks

We have shown algebraic conditions equivalent to saturating the data processing inequality for  $1 < z \leq \alpha \leq 2z$ , which generalizes the  $\alpha$ -SRD saturation condition from [21]. The techniques in this paper fail for  $\alpha < 1$  because the Hölder inequality requires positive powers, so it would be interesting to find a similar result for this case. As mentioned in section 2., a quantum channel  $\Lambda$  is said to be *sufficient* with respect to  $\rho$  and  $\sigma$  if there exists a quantum channel  $\mathcal{R}$  such that  $(\mathcal{R} \circ \Lambda)(\rho) = \rho$  and  $(\mathcal{R} \circ \Lambda)(\sigma) = \sigma$ . For Umegaki relative entropy,  $\alpha$ -RRE, and  $\alpha$ -SRD, saturation of the DPI is equivalent to sufficiency of the quantum channel  $\Lambda$ . In general, it is not known whether sufficiency of the quantum channel is equivalent to saturation of the  $\alpha - z$  RRE DPI. However in [16], Hiai and Mosonyi do prove such results for a set of density operators fixed under the quantum channel. (i.e.,  $\Lambda(\rho) = \rho$  and  $\Lambda(\sigma) = \sigma$ ). It would be interesting to find a larger class of channels where sufficiency holds.

### 5.1. Acknowledgments

The author is appreciative to Haonan Zhang for carefully reading this manuscript and asking questions that lead to the modification of Proposition 1. The author is also grateful to Anna Vershynina, who is her advisor, for her comments and suggestions throughout the entirety of this paper. This work is supported by the National Science Foundation (NSF) grant DMS-1812734.

## REFERENCES

- [1] T. Ando: Concavity of certain maps on positive definite matrices and applications to Hadamard products *Linear Algebra and its Applications* **26**, 203–241 (1979).
- [2] K. Audenaert and N. Datta:  $\alpha - z$  Rényi relative entropies *Journal of Mathematical Physics* **56**, (2015).
- [3] B. Baumgartner: An inequality for the trace of matrix products, using absolute values *arXiv preprint 1106.6189*, (2011).
- [4] S. Beigi: Sandwiched Rényi divergence satisfies data processing inequality *Journal of Mathematical Physics* **54**, (2013).
- [5] R. Bhatia: *Matrix Analysis*, Springer Science & Business Media, 2013.
- [6] E. Carlen, R. Frank, and E. Lieb: Inequalities for quantum divergences and the Audenaert–Datta conjecture *Journal of Physics A: Mathematical and Theoretical* **51** (2018).
- [7] E. Carlen, R. Frank, and E. Lieb: Some operator and trace function convexity theorems *Linear Algebra and its Applications* **460**, 174–185 (2016).
- [8] E. Carlen and A. Vershynina: Recovery and the Data Processing Inequality for Quasi-Entropies *IEEE Transactions on Information Theory* **64**, 6929–6938 (2018).
- [9] E. Carlen and A. Vershynina: Recovery map stability for the data processing inequality *Journal of Physics A: Mathematical and Theoretical* **53** (2020).
- [10] C. Bény, F. Hiai, M. Mosonyi, and D. Petz: Quantum f-divergences and error correction *Reviews in Mathematical Physics* **23**, 691–747 (2011).
- [11] I. Csiszár: Generalized cutoff rates and Rényi’s information measures *IEEE Transactions on information theory* **41**, 26–34 (1995).
- [12] G. B. Folland: *Real Analysis: Modern Techniques and their Applications*, John Wiley & Sons 1999.
- [13] R. Frank and E. Lieb: Monotonicity of a relative Rényi entropy *Journal of Mathematical Physics* **54** (2013).
- [14] B. Hall: *Quantum Theory for Mathematicians*, Springer-Verlag New York 2013.
- [15] F. Hiai: Concavity of certain matrix trace and norm functions. II *Linear Algebra and its Applications* **496**, 193–220 (2016).
- [16] F. Hiai and M. Mosonyi: Different quantum f-divergences and the reversibility of quantum operations *Reviews in Mathematical Physics* **29** (2017).
- [17] F. Hiai and others: Concavity of certain matrix trace functions *Taiwanese Journal of Mathematics* **5**, 535–554 (2001).
- [18] A. Jenčová: Rényi Relative Entropies and Noncommutative  $L_p$ – Spaces *Annales Henri Poincaré* **19**, 2513–2542 (2018).
- [19] A. Jenčová: Rényi relative entropies and noncommutative  $L_p$ – spaces II *arxiv preprint: 160908462* (2017).
- [20] A. Jenčová: Preservation of a quantum Rényi relative entropy implies existence of a recovery map *Journal of Physics A: Mathematical and Theoretical* **50** (2017).
- [21] F. Leditzky, C. Rouz, and N Datta: Data processing for the sandwiched Rényi divergence: A condition for equality *Letters in Mathematical Physics* **107** (2017).
- [22] E. Lieb: Convex trace functions and the Wigner-Yanase-Dyson conjecture *Les rencontres physiciens-mathématiciens de Strasbourg-RCP25* **19**, 0–35 (1973).

- [23] G. Lindblad: Completely positive maps and entropy inequalities *Communications in Mathematical Physics* **40**, 147–151 (1975).
- [24] G. Linblad: Expectations and entropy inequalities for finite quantum systems *Communications in Mathematical Physics* **39**, 111–119 (1974).
- [25] S. M. Manjegani: Hölder and Young inequalities for the trace of operators *Positivity* **11** (2007).
- [26] M. Mosonyi and F. Hiai: On the quantum Rényi relative entropies and related capacity formulas *IEEE Transactions on Information Theory* **57**, 2474–2487 (2011).
- [27] M. Mosonyi and T. Ogawa: Quantum hypothesis testing and the operational interpretation of the quantum Rényi relative entropies *Communications in Mathematical Physics* **334**, 1617–1648 (2015).
- [28] M. Muller-Lennert and others: On quantum Rényi entropies: A new generalization and some properties *Journal of Mathematical Physics* **54** (2013).
- [29] T. Ogawa and H. Nagaoka: Strong converse and Stein’s lemma in quantum hypothesis testing *Asymptotic Theory Of Quantum Statistical Inference: Selected Papers* 28–42 (2005).
- [30] D. Petz: Sufficiency of Channels over von Neumann Algebras *The Quarterly Journal of Mathematics* **39**, 97–108 (1988).
- [31] D. Petz: Sufficient subalgebras and the relative entropy of states of a von Neumann algebra *Communications in Mathematical Physics* **105**, 123–131 (1986).
- [32] W. Rehder: On the Product of Self-Adjoint Operators *International Journal of Mathematics and Mathematical Sciences* **5** (1982).
- [33] J. R. Retherford and others: Hilbert space: compact operators and the trace theorem *Cambridge University Press* **27** (1993).
- [34] S. Roman, S. Axler, and FW. Gehring: *Advanced Linear Algebra* 2nd ed., Springer 2005.
- [35] M. B. Ruskai: Inequalities for quantum entropy: A review with conditions for equality *Journal of Mathematical Physics* **43**, 4358–4375 (2002).
- [36] W F. Stinespring: Positive functions on C\*-algebras *Proceedings of the American Mathematical Society* **6**, 211–216 (1955).
- [37] A. Uhlmann: Endlich Dimensionale Dichtmatrizen, II *Karl-Marx-Univeristy Leipzig* **22**, 139–177 (1973).
- [38] H. Umegaki: Conditional expectation in an operator algebra, IV (entropy and information) *Kodai Mathematical Seminar Reports* **14**, 59–85 (1962).
- [39] J. von Neumann: *Mathematische Grundlagen der Quantenmechanik*, Springer-Verlag 2013.
- [40] M. Wilde: *Quantum Information Theory*, Cambridge University Press 2013.
- [41] 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 *Communications in Mathematical Physics* **331**, 593–622 (2014).
- [42] H. Zhang: Equality conditions of Data Processing Inequality for  $\alpha - z$  Rényi relative entropies *arXiv preprint arXiv:2007.06644* (2020).
- [43] H. Zhang: From Wigner-Yanase-Dyson conjecture to Carlen-Frank-Lieb conjecture *Advances in Mathematics* **365** (2020).