Saturating the Data Processing Inequality for $\alpha - z$ Rényi Relative Entropy

Sarah Chehade$^1$ and Anna Vershynina$^2$

$^{1,2}$ Department of Mathematics, University of Houston

Abstract

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 $1 < z \leq \alpha \leq 2z$. Moreover, these conditions coincide whenever $\alpha = z$.

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 [21] about non-commutative conditional expectations. When two states pass through a 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 [19, 20] proved this inequality in the context of von Neumann Algebras. More information about this relative entropy is found in section 2.2.1. 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 [1, 3, 4, 6–10, 13–16, 19–21, 23]. One quantum generalization of the Quantum relative entropy is called the $\alpha$–Rényi relative entropy, and it was proven by Petz in [19, 20] and also by Beny, Mosonyi, et.al in [41] that this entropy measure indeed satisfies the Data Processing Inequality. More information about this relative entropy is found in section 2.2.2. Later, a different generalization was introduced by Müller-Lennert, Dupuis, Szehr, Fehr, and Tomamichel, 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 [1], introduced a two parameter family of relative entropies that generalizes all entropy functions we’ve stated thus far. This family of entropies, 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 [7] gives a nice intuitive summary of the contributions to the Data Processing Inequality, and [23] finishes it up with the final characterizations of the parameters.

It is known that Lindbald and Uhlmann 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 Data Processing Inequality. i.e., when is the relative entropy preserved when states pass through a 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. Our work contributes to the question of recoverability in terms of the \(\alpha-z\) Rényi relative entropy.

Our main result says:

**Corollary** (4.0.1). Let \(\rho \in \mathcal{D}(\mathcal{H})\), \(\sigma \in \mathcal{P}(\mathcal{H})\), and \(\Lambda : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{K})\) be a quantum channel. For any \(1 < z \leq \alpha \leq 2\), whenever saturation of the Data Processing inequality holds, i.e.,

\[
D_{\alpha,z}(\rho||\sigma) = D_{\alpha,z}(\Lambda(\rho)||\Lambda(\sigma)),
\]

our states have the following structure:

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

The paper is arranged as follows: in section 2, we discuss the definitions and notations used throughout this paper. In addition to this, we mention some known results and properties about the different quantum relative entropies of interest. In section 3, we mention the technical results using tools in complex analysis or results about convex/concave trace functionals. Section 4 is dedicated to our 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, we will only consider finite-dimensional Hilbert spaces. (Although some of what is written may be generalized to infinite-dimensional Hilbert spaces.) When we write \(\mathcal{H}_{AB}\), we 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 (\(\rho \geq 0\)), then it is automatically self adjoint, i.e., \(\rho^* = \rho\). This is typically proven using the polarization identity. For a pure state \(|\psi\rangle \in \mathcal{H}\), we denote its corresponding rank one projection in \(\mathcal{B}(\mathcal{H})\), which we also call a pure state, by \(\psi = |\psi\rangle \langle \psi|\). Given any linear operator \(\mathcal{L} : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{K})\), between two Hilbert spaces \(\mathcal{H}\) and \(\mathcal{K}\), the adjoint operator \(\mathcal{L}^* : \mathcal{B}(\mathcal{K}) \to \mathcal{B}(\mathcal{H})\) 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 $p$. We say linear operators $\mathcal{L} : B(\mathcal{H}) \rightarrow B(\mathcal{K})$, between two Hilbert spaces $\mathcal{H}$ and $\mathcal{K}$ are $n$−positive if
\[
1_n \otimes \mathcal{L} : B(\mathbb{C}^n) \otimes B(\mathcal{H}) \rightarrow B(\mathbb{C}^n) \otimes B(\mathcal{K})
\]
is a positive operator, where $1_n$ is the identity operator on $B(\mathbb{C}^n)$.

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

2.0.2 Definition. 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 Other Known Results

2.1.3 Theorem. [2, Theorem 2 Matrix Hölder Inequality] Consider $2m \times m$ matrices $A$ and $B$ as well as their absolute values. Then
\[
|\text{Tr}A^*B| \leq (\text{Tr}|A|^p)^{\frac{1}{p}} (\text{Tr}|B|^q)^{\frac{1}{q}},
\]
where $1 \leq p, q \leq \infty$ such that $\frac{1}{p} + \frac{1}{q} = 1$. Note that equality holds here if and only if $|A|^p = |B|^q$.

2.1.4 Theorem (Young’s Inequality for Numbers). If $\alpha + \beta = 1$, then $a^\alpha b^\beta \leq \alpha a + \beta b$, where $a$ and $b$ are non-negative real numbers.

2.2 Quantum Entropies

2.2.1 Umegaki Relative Entropy

In quantum information theory, the information shared between states is regularly studied through the understanding of quantum entropies. Umegaki relative entropy in [21] is defined as
\[
D(\rho || \sigma) := \text{Tr}(\rho \log \rho - \rho \log \sigma),
\]
where $\rho \in D(\mathcal{H})$ and $\sigma \in D(\mathcal{H})$, provided that $\text{supp}(\rho) \subseteq \text{supp}(\sigma)$. Otherwise, the relative entropy between $\rho$ and $\sigma$ is said to be $\infty$. Reference [7] provides an extensive review of the formulation of this relative entropy, applications, and some of its properties. In this setting, [17] introduces and explain the data processing inequality. 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 quantum noisy channel. So for any relative entropy to have operational meaning the Data Processing Inequality must be satisfied.
2.2.5 Definition. If there is a quantum channel $\Psi$ such that $\Psi$ recovers states $\rho$ and $\sigma$, i.e.,

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

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 Data Processing Inequality was originally proven by Petz in the context of von Neumann algebras in [10,20]. The result states that 2 states $\rho$ and $\sigma$ saturate the Data Processing Inequality 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

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

We index $\Psi$ 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

2.2.6 Theorem. [9, Corollary 1.7] For $\rho$ and $\sigma$ density operators and $N$ a partial trace, the following inequality holds

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

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.

2.2.2 $\alpha$–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

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

provided $\text{supp}(\rho) \subseteq \text{supp}(\sigma)$. For $\alpha \in [0,1) \cup (1,2]$, Theorem 5.1 of [4] proves equality of the DPI if and only if there exists a recovery map that recovers both states $\rho$ and $\sigma$ perfectly well. Furthermore, [4] 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 $z \in \mathbb{C}$. 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

2.2.7 Theorem. [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.

\footnote{It is understood that when $\alpha = 1$ or the limit as $\alpha$ approaches 1, this is the Umegaki Relative Entropy case.}
2.2.3 $\alpha -$Sandwiched Rényi Divergence ($\alpha -$SRD)

This section describes another way to generalize the Umegaki relative entropy and we state some known results as well. For $\alpha \in (-\infty, 1) \cup (1, \infty)$, [15] introduced a new family of Rényi relative entropies called Sandwiched Rényi Relative Entropies. They are defined as

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

provided that supp($\rho$) $\subseteq$ supp($\sigma$). The data processing inequality for all $\alpha \in \left( \frac{1}{2}, 1 \right) \cup (1, \infty)$ was proven in [3]. That is

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

In [16], it was shown that the equality of the data processing inequality is satisfied for all $\alpha > \frac{1}{2}$, if and only if states $\rho$ and $\sigma$ have the following algebraic form:

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

For $\alpha > 1$, the equivalence between saturating the Data Processing Inequality and the sufficiency property is proven in [13]. The techniques use noncommutative interpolated $L_p$ spaces for von Neumann Algebras. The same author proved the equivalence between saturating the Data Processing Inequality and the sufficiency for $\alpha \in \left( \frac{1}{2}, 1 \right)$ in [14] using different norms.

2.2.4 $\alpha - z$ Rényi Relative Entropy ($\alpha - z$ RRE)

We focus on 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 [1] and is defined as

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

provided the supp($\rho$) $\subseteq$ 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 $\alpha \to 1$ or when $z = \alpha$ and $\alpha \to 1$, 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 [1, 7]. We only list the properties that are explicitly used in this paper.

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

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

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 $\rho, \sigma, \tau \in \mathcal{D}(\mathcal{H})$, we have

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

This is due to the fact that the trace of a tensor product between two states is the product of trace of states.
2.2.8 Remark. We will always assume that our 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 data processing inequality holds is outlined in [7] and finally concluded in [23]. This is summarized in the next theorem.

2.2.9 Theorem. [23, Theorem 1.2] The \( \alpha - z \) Rényi relative entropy is monotone under completely positive trace preserving maps (quantum channels) on \( D(\mathcal{H}) \) 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 data processing inequality and joint convexity/concavity of the trace functional defined by the map

\[
(A, B) \mapsto \text{Tr}(B^\frac{\alpha}{z} K^* A^p K B^\frac{1 - \alpha}{z})^s,
\]

where \( A \) and \( B \) are positive operators on \( \mathcal{H} \), \( K \) is any operator in \( B(\mathcal{H}) \), \( p, q > 0 \), 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 we will define the trace functional as

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

The next theorem describes the relationship between the data processing inequality and joint convexity and joint concavity of the trace functional.

2.2.10 Theorem. [7, 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 2.2.9 and Theorem 2.2.10 give the complete picture for data processing of the \( \alpha - z \) Rényi Relative Entropy.

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

3 Preliminaries

This section is dedicated to the technical results needed to prove the main result.
3.1 Tools from Complex Analysis

Let 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 we may use the analytic functional calculus and 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 \(X\), we define
\[
\text{Re}(X) := \frac{X + X^*}{2}
\]
and
\[
\text{Im}(X) := \frac{X + X^*}{2i}.
\]

Then let
\[
I_n^+ := \{ X \in M_n(\mathbb{C}) \text{ s.t. } \text{Im}(X) > 0 \},
\]
where \(n < \infty\) denotes the dimension.

Let us recall a few known facts or results.

3.1.1 Lemma. \([11, \text{Lemma 1.1}]\) If \(X \in I_n^+\), then \(X\) is invertible and moreover we see \(\text{Spec}(X) \subset \mathbb{C}^+\).

3.1.2 Lemma. \([11, \text{Lemma 1.2}]\) Let \(0 < p \leq 1\). If \(X \in I_n^+\), then so is \(X^p\).

3.1.3 Lemma. Any pair of operators \(A\) and \(B\) have the following properties:
\[
\text{Re}(ABA^*) = A(\text{Re}B)A^* \text{ and } \text{Im}(ABA^*) = A(\text{Im}B)A^*.
\]

Proof. Observe that
\[
\text{Im}(ABA^*) = \frac{(ABA^*) - (ABA^*)^*}{2i} = \frac{(ABA^*) - (AB^*A^*)}{2i} = \frac{A(B - B^*)A^*}{2i} = A(\text{Im}B)A^*,
\]
and the real version is similar. \(\square\)

3.1.4 Lemma. If \(B \in I_n^+\), then so is \(ABA^*\).

Proof. By Lemma 3.1.3, we see that \(\text{Im}(ABA^*) = A(\text{Im}B)A^*\). Thus for any \(y\), we have
\[
\langle \text{Im}(ABA^*)y, y \rangle = \langle A(\text{Im}B)A^*y, y \rangle = \langle (\text{Im}B)A^*y, A^*y \rangle = \left\langle \left[ (\text{Im}B)^{\frac{1}{2}} \right] A^*y, \left[ (\text{Im}B)^{\frac{1}{2}} \right] A^*y \right\rangle > 0,
\]
where the last equality is true because \(\text{Im}(B)\) is positive providing a unique square root that is also positive. \(\square\)
3.1.5 Proposition. For $0 < p < 1$, let $\Phi$ be a strictly positive, linear, and self-adjoint map. Then

$$X \mapsto \text{Tr} \left[ \left\{ \Phi \left( X^p \right) A \Phi \left( X^p \right) \right\}^{1/p} \right]$$

is concave, where $X$ and $A$ are both positive.

Proof. To prove the conjecture, it suffices to show that for $X$ and $A$ positive operators and for any $H$ hermitian,

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

for any small $x > 0$. This is because if (3.1) holds, then

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

is concave, and hence we can take a limit as $x \to 0$. Observe that for any $z \in \mathbb{C}^+$, we see that $zX + H \in \mathbb{C}^+$. So $(zX + H)^{\frac{p}{2}}$ is well defined, by the analytic functional calculus. Moreover, by Lemma 3.1.2 we also see that $(zX + H)^{\frac{p}{2}} \in I^+_n$. By the linearity and positivity of $\Phi$, it follows that

$$\text{Im}(\Phi(X)) = \frac{\Phi(X) - (\Phi(X))^*}{2i} = \Phi(\text{Im}(X)) > 0.$$ 

Hence $\Phi(X) \in I^+_n$. Then again by Lemma 3.1.2 we have $\Phi\{\langle X \rangle^{\frac{p}{2}}\} \in I^+_n$. The operator $A \in I^+_n$ because $A$ is assumed to be a positive operator. Define

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

This is analytic in $\mathbb{C}^+$ because a product of analytic functions is analytic. Since $A \in I^+_n$, by Lemma 3.1.4 and by the fact that $\Phi$ is self-adjoint, we see that $F(z) \in I^+_n$. Thus by Lemma 3.1.1 Spec($F(z)$) $\in \mathbb{C}^+$, which is an open subset of $\mathbb{C}$. Hence $(F(z))^{\frac{1}{p}}$ is well defined by the analytic functional calculus. So $\text{Tr}((F(z))^{\frac{1}{p}}) \in \mathbb{C}^+$ holds, by Lemma 3.1.1.

Next, we want to continuously extend function $F(z)$ above onto $(R, \infty)$ for some $R > 0$. To do this, for every $z \in \mathbb{C}^+$ such that $|z| > R$, we can define an analytic function

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

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

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

Similarly, we may write

$$(F(z))^{\frac{1}{p}} = z \left\{ \Phi \left( (X + z^{-1}H)^{\frac{p}{2}} \right) A \Phi \left( (X + z^{-1}H)^{\frac{p}{2}} \right) \right\}^{\frac{1}{p}}.$$
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. Indeed we obtain a Pick function \( \phi \) on \( \mathbb{C} \setminus (-\infty, R) \) such that \( \phi(x) = \text{Tr} \left[ (F(x))^{\frac{1}{p}} \right] \) for all \( x \in (R, \infty) \). Then we can rewrite

\[
x\phi(x^{-1}) = \text{Tr} \left\{ \Phi \left( (X + xH)^{\frac{1}{p}} \right) A \Phi \left( (X + xH)^{\frac{1}{p}} \right) \right\}^{\frac{1}{p}},
\]

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

\[
\phi(z) = a + bz + \int_{-\infty}^{\infty} \frac{1 + t z}{t - z} d\nu(t),
\]

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 \((R, \infty)\). Thus for all \( x \in (0, R^{-1}) \),

\[
x\phi(x^{-1}) = x \left( a + \frac{b}{x} + \int_{-\infty}^{\infty} \frac{1 + tx^{-1}}{t - x^{-1}} d\nu(t) \right)
= ax + b + \int_{-\infty}^{\infty} \frac{x(x + t)}{tx - 1} d\nu(t),
\]

with

\[
\frac{d^2}{dx^2} \left( ax + b + \int_{-\infty}^{\infty} \frac{x(x + t)}{tx - 1} d\nu(t) \right) = \int_{-\infty}^{\infty} \frac{d^2}{dx^2} \left( \frac{x(x + t)}{tx - 1} \right) d\nu(t)
= \int_{-\infty}^{\infty} \frac{2(t^2 + 1)}{(xt - 1)^3} d\nu(t) < 0,
\]

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 \( f_{a,z} (H, \rho, \sigma) \) from equation (3.3) from below, we need the next few results. Recall the equation from (2.1)

\[
(A, B) \mapsto \text{Tr} (B^{\frac{q}{p}} K^* A^p K B^{\frac{q}{p}})^s.
\]

### 3.2.6 Theorem. [23, 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 (3.2) is jointly concave.
2. If \(-1 \leq q \leq p \leq 0 \) and \( s > 0 \), then the map from (3.2) 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 (3.2) is jointly convex.

### 3.2.7 Theorem. [23, Theorem 3.3] For \( r_i > 0, i \in \{0, 1, 2\} \) such that \( \frac{1}{r_0} = \frac{1}{r_1} + \frac{1}{r_2} \), we have that for any invertible \( X, Y \in \mathcal{B}(\mathcal{H}) \) that

\[
\text{Tr} |XY|^r = \max \left\{ \frac{r_1}{r_0} \text{Tr} |XZ|^r_0 - \frac{r_1}{r_2} \text{Tr} |Y^{-1} Z|^r_2 : Z \in \mathcal{B}(\mathcal{H}) \text{ and invertible} \right\}.
\]
3.2.8 Proposition. [7, Proposition 5] For a fixed operator $B$, the map on positive operators $A \mapsto \text{Tr} \left[ (B^* A^p B)^{\frac{1}{p}} \right]$

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

3.2.9 Proposition. 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 $(A, B)$.

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

$$g \left( \sum_i \lambda_i A_i, H \right) \leq \sum_i \lambda_i g(A_i, H)$$

and

$$h \left( \sum_i \lambda_i B_i, H \right) \leq \sum_i \lambda_i h(B_i, H).$$

Thus we have

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

as desired. \hfill \Box

3.2.10 Proposition. If $f(A, B, H)$ is jointly convex in $A$ and $B$, then so is $\sup_{H > 0} \{ f(A, B, H) \}$.

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

3.2.11 Lemma. Let $\rho$ and $\sigma \in \mathcal{D}(\mathcal{H})$, and assume that $\alpha > 1$ and $z > 1$. For any positive operator $H$, define

$$f_{\alpha, z}(H, \rho, \sigma) := z \text{Tr}(\sigma^{\frac{1-\alpha}{2z}} \rho^{\frac{\alpha}{2z}} \sigma^{\frac{1-\alpha}{2z}} H) - (z - 1) \text{Tr} \left[ \left( \sigma^{\frac{z-1}{2z}} H \sigma^{\frac{1-\alpha}{2z}} \right)^{\frac{z-1}{z}} \right]. \quad (3.3)$$

Then $\Psi_{\alpha, z}(\rho||\sigma) = \sup_{H > 0} f_{\alpha, z}(H, \rho, \sigma)$, where the supremum is achieved whenever $H = \sigma^{\frac{1}{2z}} (\sigma^{\frac{1-\alpha}{2z}} \rho^{\frac{\alpha}{2z}} \sigma^{\frac{1}{2z}})^{\frac{z}{z-1}} \sigma^{\frac{1}{2z}}$. 

10
Proof. For $X$ and $Y \in \mathcal{P}(\mathcal{H})$, and for a choice of $1 \leq p, q \leq \infty$, such that $1 = \frac{1}{p} + \frac{1}{q}$, we have that

\[
\text{Tr}(XY) = \text{Tr}(X^*Y) \leq |\text{Tr}(X^*Y)| \leq (\text{Tr}|X|^p)^{\frac{1}{p}} (\text{Tr}|Y|^q)^{\frac{1}{q}} = (\text{Tr}X^p)^{\frac{1}{p}} (\text{Tr}Y^q)^{\frac{1}{q}} \leq \frac{1}{p} \text{Tr}X^p + \frac{1}{q} \text{Tr}Y^q,
\]

where (3.4) follows from the assumption that $X$ is positive, (3.5) follows from Theorem 2 in [2], (3.6) is because $X$ and $Y$ are self adjoint, and (3.7) is from Young’s inequality stated as Theorem 2.1.4. Thus,

\[
p\text{Tr}(XY) - \frac{p}{q} \text{Tr}(Y)^q \leq \text{Tr}X^p,
\]

with equality if and only if $X^p = Y^q$. Take positive operators

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

where $H$ is some positive operator in $\mathcal{B}(\mathcal{H})$. Then it follows that (3.8) becomes

\[
p\text{Tr}(\sigma^{\frac{1-a}{2\alpha}} \rho^{\frac{a}{\alpha}} \sigma^{\frac{1-a}{2\alpha}} H \sigma^{\frac{1-a}{2\alpha}}) - \frac{p}{q} \text{Tr} [(\sigma^{\frac{1-a}{2\alpha}} H \sigma^{\frac{1-a}{2\alpha}})^q] \leq \text{Tr} [(\sigma^{\frac{1-a}{2\alpha}} \rho^{\frac{a}{\alpha}} \sigma^{\frac{1-a}{2\alpha}})^p].
\]

Equality happens

if and only if $X^p = Y^q$

if and only if $\left(\sigma^{\frac{1-a}{2\alpha}} \rho^{\frac{a}{\alpha}} \sigma^{\frac{1-a}{2\alpha}}\right)^p = \left(\sigma^{\frac{1-a}{2\alpha}} H \sigma^{\frac{1-a}{2\alpha}}\right)^q$

if and only if $\sigma^{\frac{1-a}{2\alpha}} \rho^{\frac{a}{\alpha}} \sigma^{\frac{1-a}{2\alpha}} = H$.

By construction, $H$ is unique, hence the left hand side of (3.8) becomes

\[
p\text{Tr} \left(\sigma^{\frac{1-a}{2\alpha}} \rho^{\frac{a}{\alpha}} \sigma^{\frac{1-a}{2\alpha}} \sigma^{\frac{1-a}{2\alpha}} \rho^{\frac{a}{\alpha}} \sigma^{\frac{1-a}{2\alpha}} \left(\sigma^{\frac{1-a}{2\alpha}} \rho^{\frac{a}{\alpha}} \sigma^{\frac{1-a}{2\alpha}}\right)^{\frac{q}{p}} \sigma^{\frac{1-a}{2\alpha}} \rho^{\frac{a}{\alpha}} \sigma^{\frac{1-a}{2\alpha}}\right) - \frac{p}{q} \text{Tr} \left(\sigma^{\frac{1-a}{2\alpha}} \rho^{\frac{a}{\alpha}} \sigma^{\frac{1-a}{2\alpha}} \left(\sigma^{\frac{1-a}{2\alpha}} \rho^{\frac{a}{\alpha}} \sigma^{\frac{1-a}{2\alpha}}\right)^{\frac{q}{p}} \sigma^{\frac{1-a}{2\alpha}} \rho^{\frac{a}{\alpha}} \sigma^{\frac{1-a}{2\alpha}}\right)^q
\]

\[
=p \text{Tr} \left(\sigma^{\frac{1-a}{2\alpha}} \rho^{\frac{a}{\alpha}} \sigma^{\frac{1-a}{2\alpha}} \left(\sigma^{\frac{1-a}{2\alpha}} \rho^{\frac{a}{\alpha}} \sigma^{\frac{1-a}{2\alpha}}\right)^{\frac{q}{p}} \sigma^{\frac{1-a}{2\alpha}} \rho^{\frac{a}{\alpha}} \sigma^{\frac{1-a}{2\alpha}}\right) - \frac{p}{q} \text{Tr} \left(\left(\sigma^{\frac{1-a}{2\alpha}} \rho^{\frac{a}{\alpha}} \sigma^{\frac{1-a}{2\alpha}}\right)^{\frac{q}{p}} \right)^q
\]

\[
=p \text{Tr} \left(\left(\sigma^{\frac{1-a}{2\alpha}} \rho^{\frac{a}{\alpha}} \sigma^{\frac{1-a}{2\alpha}}\right)^{\frac{q}{p}} \right) - \frac{p}{q} \text{Tr} \left(\left(\sigma^{\frac{1-a}{2\alpha}} \rho^{\frac{a}{\alpha}} \sigma^{\frac{1-a}{2\alpha}}\right)^{\frac{q}{p}} \right)^q
\]

\[
=p \text{Tr} \left(\left(\sigma^{\frac{1-a}{2\alpha}} \rho^{\frac{a}{\alpha}} \sigma^{\frac{1-a}{2\alpha}}\right)^{\frac{q}{p}} \right) - \frac{p}{q} \text{Tr} \left(\left(\sigma^{\frac{1-a}{2\alpha}} \rho^{\frac{a}{\alpha}} \sigma^{\frac{1-a}{2\alpha}}\right)^{\frac{q}{p}} \right)^q
\]

\[
= \text{Tr} \left(\left(\sigma^{\frac{1-a}{2\alpha}} \rho^{\frac{a}{\alpha}} \sigma^{\frac{1-a}{2\alpha}}\right)^{\frac{q}{p}} \right) - \frac{p}{q} \text{Tr} \left(\left(\sigma^{\frac{1-a}{2\alpha}} \rho^{\frac{a}{\alpha}} \sigma^{\frac{1-a}{2\alpha}}\right)^{\frac{q}{p}} \right)^q
\]

\[
= \text{Tr} \left(\left(\sigma^{\frac{1-a}{2\alpha}} \rho^{\frac{a}{\alpha}} \sigma^{\frac{1-a}{2\alpha}}\right)^{\frac{q}{p}} \right) - \frac{p}{q} \text{Tr} \left(\left(\sigma^{\frac{1-a}{2\alpha}} \rho^{\frac{a}{\alpha}} \sigma^{\frac{1-a}{2\alpha}}\right)^{\frac{q}{p}} \right)^q
\]

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 $q = \frac{1-z}{z-1}$, where $z > 1$ we define the left hand side of (3.9) as

\[
f_{\alpha,z}(H, \rho, \sigma) := z \text{Tr}(\sigma^{\frac{1-a}{2\alpha}} \rho^{\frac{a}{\alpha}} \sigma^{\frac{1-a}{2\alpha}} H) - (z-1) \text{Tr} \left(\left(\sigma^{\frac{1-a}{2\alpha}} \rho^{\frac{a}{\alpha}} \sigma^{\frac{1-a}{2\alpha}}\right)^{\frac{q}{p}} \right)^q
\]

Then, we get the desired result.

\[
\square
\]
3.2.12 Proposition. Whenever $1 < z \leq \alpha \leq 2z$ and for a fixed $H > 0$, we have that $f_{\alpha,z}(H,\rho,\sigma)$ from equation (3.3) is jointly convex in $\rho$ and $\sigma$.

Proof. Fix $H > 0$. Let $p = \frac{\alpha}{z}$, $q = \frac{z-\alpha}{z}$, $X = \rho^\frac{z}{2}$, and $Y = \sigma^\frac{z}{2}H^{\frac{1}{z}}$. Then $X$ is invertible because $\rho$ is, and similarly, $Y$ is invertible because $H$ and $\sigma$ are. Note that

$$
\text{Tr} \left( \sigma^\frac{z}{2z} \rho^\frac{z}{2z} \sigma^\frac{z}{2z} H \right) = \text{Tr} \left( H^\frac{z}{2} \sigma^\frac{z}{2z} \rho^\frac{z}{2z} \sigma^\frac{z}{2z} H^\frac{1}{2} \right) = \text{Tr} \left( H^\frac{z}{2} \sigma^\frac{z}{2z} \rho^\frac{z}{2z} \sigma^\frac{z}{2z} H^\frac{1}{2} \right) = \text{Tr} \left( Y^* X^* X Y \right) = \text{Tr} |X^2|.
$$

Since $1 < z \leq \alpha \leq 2z$, we see that $1 \leq p = \frac{\alpha}{z} \leq 2$ and $0 \leq -q = -\frac{z-\alpha}{z} \leq 1$. 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 3.2.7 we have that

$$
\text{Tr} \left( \sigma^\frac{z}{2z} \rho^\frac{z}{2z} \sigma^\frac{z}{2z} H \right) = \max \left\{ \rho \text{Tr} |\rho^\frac{z}{2} Z|\frac{z}{2} + q \text{Tr} |H^{-\frac{1}{2}} \sigma^{-\frac{z}{2}} Z|\frac{z}{2} \right\},
$$

where $Z$ is invertible. Since $1 \leq p \leq 2$, it follows from Proposition 3.2.10 part 2 that

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

is convex in $\rho$. For the second term in (3.10), choose $\Phi_Z(X) := Z^* X Z$, which is linear, positive, and self-adjoint. Let $\sigma = X$, $-q = p$, and $A = Z^{-1} H^{-1}(Z^{-1})^*$. Then $A$ is positive. Moreover, for any $y$, since

$$
\langle Z^*(Z^{-1})^* y, y \rangle = \langle (Z^{-1})^* y, Z y \rangle = \langle y, Z^{-1} Z y \rangle = \langle y, y \rangle,
$$

it follows that $Z^*(Z^{-1})^* = 1$, and similarly, $(Z^{-1})^* Z^* = 1$ so then $(Z^*)^{-1} = (Z^{-1})^*$. Thus we may rewrite $A = Z^{-1} H^{-1}(Z^*)^{-1}$. Then by Proposition 3.1.3 we have that

$$
\text{Tr} |H^{-\frac{1}{2}} \sigma^{-\frac{z}{2}} Z|\frac{z}{2} = \text{Tr} \left[ \left( H^{-\frac{1}{2}} \sigma^{-\frac{z}{2}} Z \right)^* \left( H^{-\frac{1}{2}} \sigma^{-\frac{z}{2}} Z \right) \right]^{-\frac{1}{2}} = \text{Tr} \left[ \left( Z^* \sigma^{-\frac{z}{2}} H^{-1} \sigma^{-\frac{z}{2}} Z \right)^{-\frac{1}{2}} \right] = \text{Tr} \left[ \left( Z^* \sigma^{-\frac{z}{2}} Z Z^{-1} H^{-1}(Z^*)^{-1} Z^* \sigma^{-\frac{z}{2}} Z \right)^{-\frac{1}{2}} \right] = \text{Tr} \left[ \left\{ \Phi_Z \left( X^\frac{z}{2} \right) A \Phi_Z \left( X^\frac{z}{2} \right) \right\}^{\frac{1}{2}} \right]
$$

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

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

is convex in $\sigma$. By Propositions 3.2.9 and 3.2.10 we conclude that $\text{Tr} \left( \sigma^\frac{z}{2z} \rho^\frac{z}{2z} \sigma^\frac{z}{2z} H \right)$ is the maximum of a sum of two convex functionals and thus is itself convex in $\rho$ and $\sigma$. 

12
On the other hand, since $\sigma^{\frac{z}{2z}}H\sigma^{\frac{z}{2z}}$ and $H^{\frac{1}{2}}\sigma^{\frac{z}{2z}}H^{\frac{1}{2}}$ have the same nonzero eigenvalues, 

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

So choosing $B = H^{\frac{1}{2}}$, $A = \sigma$, and $\rho = \sigma^{\frac{z}{2z}}$, by Proposition 3.2.8 part 1, we see that 

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

is concave in $\sigma$ whenever $0 \leq \frac{z}{2} \leq 1$. Hence 

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

is convex in $\sigma$. As a consequence, we have that by Proposition 3.2.9, we have that 

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

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

## 4 Main Result

The techniques we use here are inspired by [10] and [23]. Our main result of this section is a consequence of this theorem.

### 4.0.1 Theorem (Necessary Partial Trace Case).

Let $\rho \in \mathcal{D}(\mathcal{H}_{AB})$ and $\sigma \in \mathcal{P}(\mathcal{H}_{AB})$. For any $1 < z \leq \alpha \leq 2z$, whenever saturation of the Data Processing inequality holds, i.e., 

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

the states have the following structure:

$$\sigma^{\frac{1-z}{2z}}_{AB} \left( \sigma^{\frac{1-z}{2z}}_{AB} \rho^{\alpha}_{AB} \sigma^{\frac{1-z}{2z}}_{AB} \right)^{z-1} \text{Tr} \left( \sigma^{\frac{1-z}{2z}}_{AB} \rho^{\alpha}_{AB} \sigma^{\frac{1-z}{2z}}_{AB} \right)^{z-1} \sigma^{\frac{1-z}{2z}}_A.$$

**Proof.** Assume $1 < z \leq \alpha \leq 2z$, 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 matrices.\(^2\) Similarly, define 

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

where $\sigma_{AB} \in \mathcal{D}(\mathcal{H}_{AB})$. For all $i = 1, \ldots, d^2$, define $\lambda_i = \frac{1}{\sigma_i}$, and let 

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

As mentioned in [10]\(^3\), we see that 

$$\tilde{\rho} = \sum_{i=1}^{d^2} \lambda_i \rho_i = \rho_A \otimes \pi_B$$

\(^2\)For more details on the generalized Pauli matrices, see [22] Chapter 3.7 for more details.

\(^3\)See [22] 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.
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 on \( H_B \). i.e., \( \pi_B = \frac{1}{d} I_B \). Define
\[ \tilde{H} := \text{argmax}_{H > 0} f_{\alpha,z}(H, \tilde{\rho}, \tilde{\sigma}) \]
and \( H_i := \text{argmax}_{H > 0} f_{\alpha,z}(H, \rho_i, \sigma_i) \). (4.1)
where
\[ f_{\alpha,z}(H, \rho, \sigma) := z \text{Tr}(\sigma \frac{i^{2\alpha}}{2\pi} \rho \sigma \frac{i^{2\alpha}}{2\pi} H) - (z - 1) \text{Tr} \left[ \left( \sigma \frac{i^{2\alpha}}{2\pi} H \sigma \frac{i^{2\alpha}}{2\pi} \right)^{z-1} \right], \]
as defined and discussed in Lemma 3.2.11. Note that the trace functional \( \Psi_{\alpha,z}(\rho||\sigma) \), mentioned after theorem 2.2.9 is proven jointly convex in \[23\] in a more general setting with general parameters. Then the following chain of inequalities says:
\[ \Psi_{\alpha,z}(\tilde{\rho}||\tilde{\sigma}) = f_{\alpha,z}(\tilde{H}, \tilde{\rho}, \tilde{\sigma}) \]
\[ \leq \sum_{i=1}^{d^2} \lambda_i f_{\alpha,z}(\tilde{H}, \rho_i, \sigma_i) \]
\[ \leq \sum_{i=1}^{d^2} \lambda_i f_{\alpha,z}(H_i, \rho_i, \sigma_i) \]
\[ \leq \sum_{i=1}^{d^2} \lambda_i \Psi_{\alpha,z}(\rho_i||\sigma_i), \]
where lines (4.2) and (4.5) follow from Lemma 3.2.11 line (4.3) follows from the joint convexity of \( f \), which is proven in Proposition 3.2.12, and line (4.4) follows from the fact that \( H_i \) is the maximizer for \( f_{\alpha,z}(H_i, \rho_i, \sigma_i) \) from (4.1).

If we assume saturation of the Data Processing inequality, this is equivalent to
\[ \Psi_{\alpha,z}(\tilde{\rho}||\tilde{\sigma}) = \sum_{i=1}^{d^2} \lambda_i \Psi_{\alpha,z}(\rho_i||\sigma_i). \]

Then the chain of inequalities above is now a chain of equalities and thus by the definition of \( H_i \), the function
\[ f_{\alpha,z}(\tilde{H}, \rho_i, \sigma_i) = f_{\alpha,z}(H_i, \rho_i, \sigma_i) \]
for all \( i = 1, \ldots, d^2 \).

By the uniqueness of the maximizer \( \tilde{H} \), which is proven above in Lemma 3.2.11 we see that \( \tilde{H} = H_i \) for all \( i = 1, \ldots, d^2 \).

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 3.2.11 we see an explicit form of the maximizer:
\[ H_i = \left( \begin{array}{ccc} \frac{i^{2\alpha}}{2\pi} \rho_i & \frac{i^{2\alpha}}{2\pi} \sigma_i \end{array} \right)^{z-1} \left( \begin{array}{ccc} \frac{i^{2\alpha}}{2\pi} \sigma_i & \frac{i^{2\alpha}}{2\pi} \rho_i \end{array} \right)^{-1} \]
\[ = \left( \begin{array}{ccc} (1 \otimes v_i) \sigma_{AB}(1 \otimes v_i^*) \end{array} \right)^{1/z} \left( \frac{i^{2\alpha}}{2\pi} \right)^{1/z} \left( \begin{array}{ccc} (1 \otimes v_i) \sigma_{AB}(1 \otimes v_i^*) \end{array} \right)^{1/z} \left( \frac{i^{2\alpha}}{2\pi} \right)^{1/z}. \]
\[
[(1 \otimes v_i)\sigma_{AB}(1 \otimes v_i^*)]^{1/z} = (1 \otimes v_i)\sigma_{AB}^{1/z} \sigma_{AB}^{-\frac{1}{z}} \rho_{AB}\sigma_{AB}^{\frac{1}{z}} \sigma_{AB}^{-\frac{1}{z}} \sigma_{AB}^{-1} \sigma_{AB}^{1/z}.
\]

This is true for all \(v_i\) due to the fact that \(v_i^*v_i = I\). Therefore for some \(i \in \{1, \ldots, d^2\}\), we see that
\[
H_i = \sigma_{AB}^{1/z} (\sigma_{AB}^{-\frac{1}{z}} \rho_{AB}\sigma_{AB}^{\frac{1}{z}}) \sigma_{AB}^{-1} \sigma_{AB}^{1/z}.
\]

Also by similar calculations,
\[
\hat{H} = (\sigma_A \otimes \pi_B)^{1/z} \left( (\sigma_A \otimes \pi_B)^{-\frac{1}{z}} (\rho_A \otimes \pi_B) \right)^{\frac{1}{z}} (\sigma_A \otimes \pi_B)^{-\frac{1}{z}}
\]
\[
= \sigma_A^{1/z} (\sigma_A^{-\frac{1}{z}} \rho_A\sigma_A^{\frac{1}{z}})^{z-1} \sigma_A^{1/z} \otimes \pi_B \left( (\sigma_A^{1/z} + (\frac{1-z}{z}) \sigma_A^{-\frac{1}{z}}) (z-1) + \frac{1-z}{z} \right)
\]
\[
= \sigma_A^{1/z} (\sigma_A^{-\frac{1}{z}} \rho_A\sigma_A^{\frac{1}{z}})^{z-1} \sigma_A^{1/z} \otimes \mathbb{1}_B,
\]

(4.6)

where (4.6) is true because \(\frac{1-z}{z} + (\frac{1-z}{z}) (z-1) + \frac{1-z}{z} = 0\). Thus
\[
\sigma_{AB}^{1/z} \sigma_{AB}^{-\frac{1}{z}} \rho_A\sigma_{AB}^{\frac{1}{z}} \sigma_{AB}^{-\frac{1}{z}} \sigma_{AB}^{-1} \sigma_{AB}^{1/z} = \sigma_A^{1/z} (\sigma_A^{1/z} \rho_A\sigma_A^{\frac{1}{z}}) \sigma_A^{1/z} \sigma_A^{-\frac{1}{z}} \sigma_A^{1/z},
\]
as desired.

We can generalize the partial case trace using a standard Stine Spring Dilation Theorem argument.

**4.0.2 Corollary.** Let \(\rho \in \mathcal{D}(\mathcal{H})\), \(\sigma \in \mathcal{P}(\mathcal{H})\), and \(\Lambda : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{K})\) be a quantum channel. For any \(1 < z \leq \alpha \leq 2z\), whenever saturation of the Data Processing inequality holds, i.e., \(D_{\alpha,z}(\rho||\sigma) = D_{\alpha,z}(\Lambda(\rho)||\Lambda(\sigma))\), the states have the following structure:
\[
\sigma^{1/z} \sigma^{-\frac{1}{z}} \rho_A\sigma_{AB}^{\frac{1}{z}} \sigma_{AB}^{-\frac{1}{z}} \sigma_{AB}^{1/z} = \Lambda^* \left( \Lambda(\sigma)^{1/z} \Lambda(\rho)^{-\frac{1}{z}} \Lambda(\sigma)^{1/z} \sigma_{AB}^{1/z} \right)^{z-1} \Lambda(\sigma)^{1/z}.
\]

**Proof of 4.0.2** Following [16], for any quantum channel \(\Lambda : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{K})\), by the Stinespring Dilation Theorem, there exists a Hilbert space \(\mathcal{H}'\), a pure state \(\tau \in \mathcal{H}' \otimes \mathcal{K}\), and a unitary operator \(U : \mathcal{H} \otimes \mathcal{H}' \otimes \mathcal{K} \to \mathcal{H} \otimes \mathcal{H}' \otimes \mathcal{K}\) such that for every \(\rho \in \mathcal{B}(\mathcal{H})\), it follows that
\[
\Lambda(\rho) = \text{Tr}_{12}(U(\rho \otimes \tau)U^*),
\]
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., \(\text{Tr}_{12} : \mathcal{H} \otimes \mathcal{H}' \otimes \mathcal{K} \to \mathcal{K}\). So in general, we always have that
\[
D_{\alpha,z}(\rho||\sigma) = D_{\alpha,z}(U(\rho \otimes \tau)U^*||U(\sigma \otimes \tau)U^*)
\]
\[
\geq D_{\alpha,z}([\text{Tr}_{12}(U(\rho \otimes \tau)U^*)||\text{Tr}_{12}(U(\sigma \otimes \tau)U^*)])
\]
\[
= D_{\alpha,z}(\Lambda(\rho)||\Lambda(\sigma)),
\]

where (1.7) is due to properties mentioned in [22.4] and (4.3) is the Data Processing Inequality for partial traces. So if we assume equality, then by Theorem 4.0.1 we see that
\[
\mathbb{1}_{\mathcal{H} \otimes \mathcal{H}'} \otimes \Lambda(\sigma)^{1/z} \Lambda(\rho)^{-\frac{1}{z}} \Lambda(\sigma)^{1/z} \mathbb{1}_{\mathcal{H} \otimes \mathcal{H}'}^{1/z} - \Lambda(\sigma)^{1/z} \Lambda(\rho)^{-\frac{1}{z}} \Lambda(\sigma)^{1/z} \mathbb{1}_{\mathcal{H} \otimes \mathcal{H}'}^{1/z}
\]

\[\text{Stinespring Dilation Theorem can be found in [22].}
\]
\[\text{for } \alpha \text{ and } z \text{ where the Data Processing Inequality makes sense. see Theorem 22.3 for such parameters}
\]
\[= [U(\sigma \otimes \tau)U^*]^{\frac{1}{2\alpha}} \left( [U(\sigma \otimes \tau)U^*]^{\frac{1}{2\beta}} [U(\rho \otimes \tau)U^*]^{\frac{1}{2\alpha}} [U(\sigma \otimes \tau)U^*]^{\frac{1}{2\alpha}} \right)^{\frac{1}{2\beta}} [U(\sigma \otimes \tau)U^*]^{\frac{1}{2\beta}} \]
\[= U(\sigma \frac{1}{2\beta} \rho \frac{1}{2\alpha} \sigma \frac{1}{2\beta})^{\frac{1}{2\alpha}} \sigma \frac{1}{2\beta} \otimes \tau)U^*, \quad (4.9)\]

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
\[\Lambda^*(X) := (\mathbb{1}_\mathcal{H} \otimes \langle \tau |_{\mathcal{H} \otimes \mathcal{K}})U^*(X)U(\mathbb{1}_\mathcal{H} \otimes |\tau\rangle_{\mathcal{H} \otimes \mathcal{K}}). \quad (4.10)\]

So applying (4.10) to (4.9), we get
\[\Lambda^* \left[ \Lambda(\sigma)^{\frac{1}{2\alpha}} \left( \Lambda(\sigma)^{\frac{1}{2\alpha}} \Lambda(\rho)^{\frac{1}{2\alpha}} \Lambda(\sigma)^{\frac{1}{2\alpha}} \right)^{\frac{1}{2\alpha}} \right] =
\[\Lambda^* \left[ U \left( \sigma^{\frac{1}{2\beta}} \left( \sigma^{\frac{1}{2\beta}} \rho \frac{1}{2\alpha} \sigma^{\frac{1}{2\beta}} \right)^{\frac{1}{2\alpha}} \right) U^* \right] =
\[\left( \mathbb{1}_\mathcal{H} \otimes \langle \tau \rangle \right) \left( \sigma^{\frac{1}{2\beta}} \left( \sigma^{\frac{1}{2\beta}} \rho \frac{1}{2\alpha} \sigma^{\frac{1}{2\beta}} \right)^{\frac{1}{2\alpha}} \right) \left( \mathbb{1}_\mathcal{H} \otimes |\tau\rangle \right) = \sigma^{\frac{1}{2\beta}} \left( \sigma^{\frac{1}{2\beta}} \rho \frac{1}{2\alpha} \sigma^{\frac{1}{2\beta}} \right)^{\frac{1}{2\alpha}} \sigma^{\frac{1}{2\beta}} \]

as desired. \(\square\)

**4.0.3 Proposition.** If \(\rho_{AB} \in \mathcal{D}(\mathcal{H}_{AB})\) is a product state and \(\sigma_{AB} \in \mathcal{P}(\mathcal{H}_{AB})\) is any density operator such that they have the following structure,

\[\frac{1}{2\alpha} \sigma_{AB}^{\frac{1}{2\alpha}} \left( \frac{1}{2\beta} \rho_{AB}^{\frac{1}{2\beta}} \sigma_{AB}^{\frac{1}{2\beta}} \right)^{\frac{1}{2\alpha}} \frac{1}{2\beta} = \sigma_{AB} \left( \frac{1}{2\alpha} \rho_{A}^{\frac{1}{2\alpha}} \sigma_{A}^{\frac{1}{2\alpha}} \right)^{\frac{1}{2\alpha}} \frac{1}{2\beta} \otimes \mathbb{1}_B,\]

then the Data processing inequality under partial traces is saturated plus some constant. i.e.,

\[D_{\alpha,z}(\rho_{AB}||\sigma_{AB}) = D_{\alpha,z}(\rho_{A}||\sigma_{A}) + k;\]

where \(k = \frac{1}{a} \log \left[ \text{Tr} \left( \rho_{B}^{\frac{1}{2\alpha}} \right) \right].\)

**Proof.** Since \(\rho_{AB}\) is a product state on \(\mathcal{H}_{AB}\), it follows that \(\rho_{AB} = \rho_{A} \otimes \rho_{B}\), where \(\rho_{A}\) is a density operator on \(\mathcal{H}_{A}\) and \(\rho_{B}\) is a density operator on \(\mathcal{H}_{B}\). Note that \(\rho_{B}\) need not be a pure state. Multiplying the assumed expression by \(\rho_{AB}\) on the left and taking a trace gives \(^6\)

\[\text{Tr} \left( \rho_{AB}^{\frac{1}{2\alpha}} \rho_{AB}^{\frac{1}{2\beta}} \rho_{AB}^{\frac{1}{2\alpha}} \right)^{\frac{1}{2\alpha}} \sigma_{AB}^{\frac{1}{2\beta}} \text{Tr} \left( \sigma_{AB}^{\frac{1}{2\alpha}} \rho_{AB}^{\frac{1}{2\beta}} \sigma_{AB}^{\frac{1}{2\beta}} \right)^{\frac{1}{2\alpha}} \text{Tr} \left( \sigma_{AB}^{\frac{1}{2\alpha}} \rho_{AB}^{\frac{1}{2\beta}} \sigma_{AB}^{\frac{1}{2\beta}} \right)^{\frac{1}{2\alpha}} \sigma_{A}^{\frac{1}{2\beta}} \otimes \mathbb{1}_B.\]

This implies

\[\text{Tr} \left[ \left( \sigma_{AB}^{\frac{1}{2\alpha}} \rho_{AB}^{\frac{1}{2\beta}} \sigma_{AB}^{\frac{1}{2\beta}} \right)^{\frac{1}{2\alpha}} \right] = \text{Tr} \left( \rho_{A}^{\frac{1}{2\beta}} \rho_{A}^{\frac{1}{2\alpha}} \right) \left( \sigma_{A}^{\frac{1}{2\beta}} \rho_{A}^{\frac{1}{2\alpha}} \sigma_{A}^{\frac{1}{2\beta}} \right)^{\frac{1}{2\alpha}} \left( \sigma_{A}^{\frac{1}{2\beta}} \rho_{A}^{\frac{1}{2\alpha}} \sigma_{A}^{\frac{1}{2\beta}} \right)^{\frac{1}{2\alpha}} \sigma_{A}^{\frac{1}{2\beta}} \otimes \mathbb{1}_B,\]

which gives

\[\text{Tr} \left[ \left( \sigma_{AB}^{\frac{1}{2\alpha}} \rho_{AB}^{\frac{1}{2\beta}} \sigma_{AB}^{\frac{1}{2\beta}} \right)^{\frac{1}{2\alpha}} \right] = \text{Tr} \left[ \left( \rho_{A}^{\frac{1}{2\beta}} \right) \left( \sigma_{A}^{\frac{1}{2\beta}} \rho_{A}^{\frac{1}{2\alpha}} \sigma_{A}^{\frac{1}{2\beta}} \right)^{\frac{1}{2\alpha}} \sigma_{A}^{\frac{1}{2\beta}} \right] \otimes \rho_{B}^{\frac{1}{2\alpha}}.\]

\(^6\)Note that we are using the fact that \(\text{Tr}_{AB}(X) = \text{Tr}_{A}(\text{Tr}_{B}(X))\)
Taking the log of both sides and multiplying by $\frac{1}{\alpha - 1}$ gives
\[
\frac{1}{\alpha - 1} \log \left( \text{Tr} \left[ \left( \sigma_{AB}^{1/z} \rho_{AB}^{1/z} \sigma_{AB}^{1/z} \right)^z \right] \right) = \frac{1}{\alpha - 1} \log \left( \text{Tr} \left[ \left( \sigma_{AB}^{1/\alpha} \rho_{AB}^{1/\alpha} \sigma_{AB}^{1/\alpha} \right)^z \right] \text{Tr} \left( \rho_B^z \right) \right),
\]
which is the same as
\[
D_{\alpha,z}(\rho_{AB}||\sigma_{AB}) = D_{\alpha,z}(\rho_A||\sigma_A) + k,
\]
where $k = \frac{1}{\alpha - 1} \log \left[ \text{Tr} \left( \rho_B^z \right) \right]$ as desired.

4.0.4 Remark. It is interesting to see that Theorem 4.0.1 and Proposition 4.0.3 would hold simultaneously and without any condition on $\rho$ if and only if $z = \alpha$. This in turn will result in $\alpha$-SRD, which aligns with the work done in [16].

4.0.5 Remark. If $\rho_B$ in Proposition 4.0.3 is a pure state, then $k = 0$. This immediately leads to another result:

4.0.6 Proposition. If $\rho_{AB} \in \mathcal{D}(\mathcal{H}_{AB})$ is a separable state and $\sigma_{AB} \in \mathcal{P}(\mathcal{H}_{AB})$ is such that they have the following structure,
\[
\sigma_{AB}^{1/\alpha} \left( \sigma_{AB}^{1/\alpha} \rho_{AB}^{1/\alpha} \sigma_{AB}^{1/\alpha} \right)^{z-1} \sigma_{AB}^{1/\alpha} = \sigma_A^{1/\alpha} \left( \sigma_A^{1/\alpha} \rho_A^{1/\alpha} \sigma_A^{1/\alpha} \right)^{z-1} \sigma_A^{1/\alpha},
\]
then the Data processing inequality under partial traces is saturated. i.e.,
\[
D_{\alpha,z}(\rho_{AB}||\sigma_{AB}) = D_{\alpha,z}(\rho_A||\sigma_A).
\]

Proof. Since $\rho_{AB}$ is a separable state on $\mathcal{H}_{AB}$, we may 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, as well as $0 \leq \lambda_i \leq 1$, 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
\[
\text{Tr} \left[ \left( \sigma_{AB}^{1/\alpha} \rho_{AB}^{1/\alpha} \sigma_{AB}^{1/\alpha} \right)^z \right] = \text{Tr} \left[ \left( \sum_{i \in I} \lambda_i |\psi_i\rangle_A \langle \psi_i| \otimes |\phi_i\rangle_B \langle \phi_i| \right)^z \sigma_A^{1/\alpha} \left( \sigma_A^{1/\alpha} \rho_A^{1/\alpha} \sigma_A^{1/\alpha} \right)^{z-1} \sigma_A^{1/\alpha} \right].
\]
This implies
\[
\text{Tr} \left[ \left( \sigma_{AB}^{1/\alpha} \rho_{AB}^{1/\alpha} \sigma_{AB}^{1/\alpha} \right)^z \right] = \text{Tr} \left[ \left( \sum_{i \in I} \lambda_i^{1-\alpha} |\psi_i\rangle_A \langle \psi_i| \sigma_A^{1-\alpha} \left( \sigma_A^{1-\alpha} \rho_A^{1-\alpha} \sigma_A^{1-\alpha} \right)^{z-1} \sigma_A^{1-\alpha} \right) \otimes |\phi_i\rangle_B \langle \phi_i| \right],
\]
which gives
\[
\text{Tr} \left[ \left( \sigma_{AB}^{1-\alpha} \rho_{AB}^{1-\alpha} \sigma_{AB}^{1-\alpha} \right)^z \right] = \text{Tr} \left[ \left( \rho_A^{1-\alpha} \sigma_A^{1-\alpha} \left( \sigma_A^{1-\alpha} \rho_A^{1-\alpha} \sigma_A^{1-\alpha} \right)^{z-1} \sigma_A^{1-\alpha} \right) \right].
\]
Taking the log of both sides and multiplying by $\frac{1}{\alpha - 1}$ gives
\[
D_{\alpha,z}(\rho_{AB}||\sigma_{AB}) = D_{\alpha,z}(\rho_A||\sigma_A)
\]
as desired.

\footnote{Note that we are using the fact that $\text{Tr}_{AB}(X) = \text{Tr}_A(\text{Tr}_B X)$}
5 Closing Remarks

5.1 Future Work

We have shown algebraic conditions equivalent to saturating the data processing inequality for \( 1 < z \leq \alpha \leq 2z \), which generalizes the Sandwiched Rényi relative saturation condition from [10]. Our techniques 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, we define a quantum channel \( \Lambda \) 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 data processing inequality is equivalent to sufficiency of the quantum channel \( \Lambda \). This is proven for the Umegaki relative entropy, \( \alpha \)-RRE, and \( \alpha \)-SRD for their respective parameters mentioned in section 2. In general, it is not known whether sufficiency of the quantum channel is equivalent to saturation of the \( \alpha - z \) data processing inequality, where the data processing inequality makes sense. However in [12], 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. Another way of thinking about this is through non-commutative interpolated \( L_p \) spaces as Jencova did in [15] and [14]. In other words, can we represent the family of \( \alpha - z \) entropies as a family of non-commutative interpolated \( L_p \) norms?

5.2 Acknowledgments

S.C. and A.V. are supported by NSF grant DMS-1812734

References

[1] Koenraad Audenaert and Nilanjana Datta. \( \alpha - z \) rényi relative entropies. Journal of Mathematical Physics, 56(022202), 2015.

[2] Bernhard Baumgartner. An inequality for the trace of matrix products, using absolute values. arXiv preprint arXiv:1106.6189, 2011.

[3] Salman Beigi. Sandwiched rényi divergence satisfies data processing inequality. Journal of Mathematical Physics, 54(12):122202, Dec 2013.

[4] Cédric Bény, Fumio Hiai, Milán Mosonyi, and Dénes Petz. Quantum f-divergences and error correction. Reviews in Mathematical Physics, 23(07):691–747, 2011.

[5] Rajendra Bhatia. Matrix analysis, volume 169. Springer Science & Business Media, 2013.

[6] Eric Carlen, Rupert Frank, and Elliott Lieb. Some operator and trace function convexity theorems. Linear Algebra and its Applications, 490:174–185, 2016.

[7] Eric Carlen, Rupert Frank, and Elliott Lieb. Inequalities for quantum divergences and the audenaert–datta conjecture. Journal of Physics A: Mathematical and Theoretical, 51(48):483001, 2018.
[8] Eric Carlen and Anna Vershynina. Recovery and the data processing inequality for quasi-entropies. *IEEE Transactions on Information Theory*, 64(10):6929–6938, Oct 2018.

[9] Eric Carlen and Anna Vershynina. Recovery map stability for the data processing inequality. *Journal of Physics A: Mathematical and Theoretical*, 53(3), January 2020.

[10] Fumio Hiai. Concavity of certain matrix trace and norm functions. ii. *Linear Algebra and its Applications*, 496:193–220, 2016.

[11] Fumio Hiai et al. Concavity of certain matrix trace functions. *Taiwanese Journal of Mathematics*, 5(3):535–554, 2001.

[12] Fumio Hiai and Milán Mosonyi. Different quantum f-divergences and the reversibility of quantum operations. *Reviews in Mathematical Physics*, 29(07):1750023, 2017.

[13] Anna Jenčová. Preservation of a quantum rényi relative entropy implies existence of a recovery map. *Journal of Physics A: Mathematical and Theoretical*, 50(8):085303, 2017.

[14] Anna Jenčová. Rényi relative entropies and noncommutative $l_p$-spaces ii. *arxiv160908462*, 2017.

[15] Anna Jenčová. Rényi relative entropies and noncommutative $l_p-$ spaces. In *Annales Henri Poincaré*, volume 19, pages 2513–2542. Springer, 2018.

[16] Felix Leditzky, Cambyse Rouzé, and Nilanjana Datta. Data processing for the sandwiched rényi divergence: a condition for equality. *Letters in Mathematical Physics*, 107(1), 2017.

[17] Göran Lindblad. Completely positive maps and entropy inequalities. *Communications in Mathematical Physics*, 40(2):147–151, 1975.

[18] Martin Müller-Lennert, Frédéric Dupuis, Oleg Szehr, Serge Fehr, and Marco Tomamichel. On quantum rényi entropies: A new generalization and some properties. *Journal of Mathematical Physics*, 54(12):122203, Dec 2013.

[19] Dénes Petz. Sufficient subalgebras and the relative entropy of states of a von neumann algebra. *Communications in Mathematical Physics*, 105:123–131, 1986.

[20] Dénes Petz. Sufficiency of channels over von neumann algebras. *The Quarterly Journal of Mathematics*, 39:97–108, 1988.

[21] Hisaharu Umegaki. Conditional expectation in an operator algebra, iv (entropy and information). In *Kodai Mathematical Seminar Reports*, volume 14, pages 59–85. Department of Mathematics, Tokyo Institute of Technology, 1962.

[22] Mark Wilde. *Quantum Information Theory*, volume 1. Cambridge University Press, 2013.

[23] Haonan Zhang. From wigner-yanase-dyson conjecture to carlen-frank-lieb conjecture. *Advances in Mathematics*, 365:107053, 2020.