Universal recoverability in quantum information

Conference Paper

Author(s):
Junge, Marius; Renner, Renato; Sutter, David; Wilde, Mark M.; Winter, Andreas

Publication date:
2016

Permanent link:
https://doi.org/10.3929/ethz-b-000120060

Rights / license:
In Copyright - Non-Commercial Use Permitted

Originally published in:
https://doi.org/10.1109/ISIT.2016.7541748

Funding acknowledgement:
258932 - Generalized (quantum) information theory (EC)
Abstract—The quantum relative entropy is well known to obey a monotonicity property (i.e., it does not increase under the action of a quantum channel). Here we present several refinements of this entropy inequality, some of which have a physical interpretation in terms of recovery from the action of the channel. The recovery channel given here is explicit and universal, depending only on the channel and one of the arguments to the relative entropy.

I. INTRODUCTION

The quantum relative entropy is a fundamental measure of distinguishability in quantum information theory [1], finding an operational interpretation in the context of quantum hypothesis testing [2], [3]. Let $\mathcal{L}(\mathcal{H})$ denote the set of linear operators acting on a finite-dimensional Hilbert space $\mathcal{H}$, $\mathcal{L}_+(\mathcal{H})$ the set of positive semi-definite operators, and let $\mathcal{D}(\mathcal{H})$ denote the set of density operators (positive semi-definite with unit trace). The quantum relative entropy is defined as the following function of $\rho \in \mathcal{D}(\mathcal{H})$ and $\sigma \in \mathcal{L}_+(\mathcal{H})$:

$$D(\rho||\sigma) = \begin{cases} \text{Tr}\{\rho[\log\rho - \log\sigma]\} & \text{if supp}(\rho) \subseteq \text{supp}(\sigma) \\ +\infty & \text{else} \end{cases}$$

where $\log$ denotes the natural logarithm here and throughout for convenience. One of the main properties of $D(\rho||\sigma)$ obeys is monotonicity [4], [5]. Let $\mathcal{N} : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H}')$ be a quantum channel (a completely positive, trace-preserving map), where $\mathcal{H}'$ is finite-dimensional. Then the following monotonicity inequality holds [4], [5]

$$D(\rho||\sigma) \geq D(\mathcal{N}(\rho)||\mathcal{N}(\sigma)).$$

The equality conditions for this entropy inequality are that $D(\rho||\sigma) = D(\mathcal{N}(\rho)||\mathcal{N}(\sigma))$ if and only if there exists a recovery channel $\mathcal{R}$, depending only on $\sigma$ and $\mathcal{N}$, such that $\rho = (\mathcal{R} \circ \mathcal{N})(\rho)$ and $\sigma = (\mathcal{R} \circ \mathcal{N})(\sigma)$ [6], [7].

The main purpose of this paper is to present several further refinements of the entropy inequality in (1), recently reported in [8] and [9] (see also the earlier contributions on this topic in [10], [11] and the more recent one in [12]). One of the refinements can be summarized informally as follows: if the decrease in quantum relative entropy between two quantum states after a quantum channel is relatively small, then it is possible to perform a recovery channel, such that we can perfectly recover one state while approximately recovering the other. This can be interpreted as quantifying how well one can reverse the action of a quantum channel. Throughout, we take $\rho$, $\sigma$, and $\mathcal{N}$ as given in the following definition:

Definition 1: Let $\rho \in \mathcal{D}(\mathcal{H})$ and $\sigma \in \mathcal{L}_+(\mathcal{H})$, such that $\text{supp}(\rho) \subseteq \text{supp}(\sigma)$. Let $\mathcal{N} : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H}')$ be a quantum channel.

The formal statement of the main theorem is as follows:

Theorem 2: Given $\rho$, $\sigma$, and $\mathcal{N}$ as in Definition 1, there exists an explicit recovery channel $\mathcal{R} : \mathcal{L}(\mathcal{H}') \to \mathcal{L}(\mathcal{H})$, depending only on $\sigma$ and $\mathcal{N}$, such that

$$D(\rho||\sigma) \geq D(\mathcal{N}(\rho)||\mathcal{N}(\sigma)) - \log F(\rho, (\mathcal{R} \circ \mathcal{N})(\rho)).$$

The fact that $\mathcal{R}$ satisfies (2) for all $\rho \in \mathcal{D}(\mathcal{H})$ implies that

$$\sigma = (\mathcal{R} \circ \mathcal{N})(\sigma).$$

Given that the quantum fidelity $F(\omega, \tau) \equiv \|\sqrt{\omega}\sqrt{\tau}\|_1^2$ [13] takes values between zero and one for density operators $\omega$ and $\tau$, we can immediately conclude that

$$-\log F(\rho, (\mathcal{R} \circ \mathcal{N})(\rho)) \geq 0,$$

so that the above theorem implies (1) as a consequence. Furthermore, the recovery channel satisfying (2) has the property that it perfectly recovers $\sigma$ from $\mathcal{N}(\sigma)$ (satisfying (3)), a fact which we prove later and which makes (2) non-trivial.

The proof for Theorem 2 relies on the method of complex interpolation [14], [15] and the notion of a Rényi generalization of a relative entropy difference [16]. We review this background before going through the proof. We also present another lower bound for the relative entropy difference $D(\rho||\sigma) - D(\mathcal{N}(\rho)||\mathcal{N}(\sigma))$ and give two upper bounds for special cases, which were previously reported in [9], [11]. One of the consequences of Theorem 2 is to provide physically meaningful improvements to many well known quantum entropy inequalities, such as strong subadditivity [17], joint convexity of quantum relative entropy, and concavity of conditional quantum entropy. These are discussed in [8], [11] and improve upon the results in [18], [19].

II. BACKGROUND

The proof of Theorem 2 given here requires some mathematical background. So we first begin by defining the Schatten norms and several of their properties. We then review some essential results from complex analysis, that lead to a complex interpolation theorem known as the Stein–Hirschman interpolation theorem [20], [14].
A. Schatten Norms and Duality

An important technical tool in the proof given here is the Schatten $p$-norm of an operator $A$, defined as

$$\|A\|_p = \left( \text{Tr} \{ |A|^p \} \right)^{1/p},$$

where $A \in \mathcal{L}(\mathcal{H})$, $|A| = \sqrt{A^*A}$, and $p \geq 1$. $\|A\|_p$ is equal to the $p$-norm of the singular values of $A$. That is, if $\sigma_i(A)$ is the vector of singular values of $A$, then

$$\|A\|_p = \left( \sum_i \sigma_i(A)^p \right)^{1/p}.$$  

The convention is for $\|A\|_\infty$ to be defined as the largest singular value of $A$ because $\|A\|_p$ converges to this in the limit as $p \to \infty$. In the proof of Theorem 2, we repeatedly use the fact that $\|A\|_p$ is unitarily invariant. That is, $\|A\|_p$ is invariant with respect to linear isometries, in the sense that $\|A\|_p = \|UV^*\|_p$, where $U, V \in \mathcal{L}(\mathcal{H}, \mathcal{H}')$ are linear isometries satisfying $U^*U = I_\mathcal{H}$ and $V^*V = I_\mathcal{H}$. From these norms, one can define information measures relating quantum states and channels, with the one used here known as a Rényi generalization of a relative entropy difference [21], [16].

Throughout we adopt the common convention and define $f(A)$ for a function $f$ and a positive semi-definite operator $A$ as follows: $f(A) \equiv \sum_{\lambda_i \neq \lambda_j} f(\lambda_i) |i\rangle \langle i|$, where $A = \sum_i \lambda_i |i\rangle \langle i|$ is a spectral decomposition of $A$. We denote the support of $A$ by $\text{supp}(A)$, and we let $\Pi_A$ denote the projection onto the support of $A$.

B. Complex Analysis

We now review a few concepts from complex analysis [22]. We will not prove these results in detail, but the purpose instead is to recall them, and the interested reader can follow references to books on complex analysis for details of proofs.

The culmination of the development is the Stein–Hirschman complex interpolation theorem (Theorem 5).

The well known maximum modulus principle [22] has an extension to an unbounded strip in $\mathbb{C}$, which we call the maximum modulus principle on a strip. Let $S$ denote the standard strip in $\mathbb{C}$, $\overline{S}$ its closure, and $\partial S$ its boundary:

$$S \equiv \{ z \in \mathbb{C} : 0 < \text{Re}(z) < 1 \},$$

$$\overline{S} \equiv \{ z \in \mathbb{C} : 0 \leq \text{Re}(z) \leq 1 \},$$

$$\partial S \equiv \{ z \in \mathbb{C} : \text{Re}(z) = 0 \lor \text{Re}(z) = 1 \}.$$  

Let $f : \overline{S} \to \mathbb{C}$ be bounded on $\overline{S}$, holomorphic on $S$, and continuous on $\partial S$. Then the supremum of $|f|$ is attained on $\partial S$. That is, $\sup_{z \in \overline{S}} |f(z)| = \sup_{z \in \partial S} |f(z)|$.

The maximum modulus principle on a strip implies a result known as the Hadamard three-lines theorem:

**Theorem 3 (Hadamard Three-Lines):** Let $f : \overline{S} \to \mathbb{C}$ be a function that is bounded on $\overline{S}$, holomorphic on $S$, and continuous on the boundary $\partial S$. Let $θ \in (0, 1)$ and $M(θ) \equiv \sup_{t \in \mathbb{R}} |f(θ + it)|$. Then $\log M(θ)$ is a convex function on $[0, 1]$, implying that

$$\log M(θ) \leq (1 - θ) \log M(0) + θ \log M(1).$$

There is a strengthening of the Hadamard three-lines theorem due to Hirschman [20], which in fact implies the Hadamard three-lines theorem:

**Theorem 4 (Hirschman):** Let $f(z) : \overline{S} \to \mathbb{C}$ be a function that is bounded on $\overline{S}$, holomorphic on $S$, and continuous on the boundary $\partial S$. Then for $θ \in (0, 1)$, the following holds

$$\log |f(θ)| \leq \int_{-\infty}^{\infty} dt \left( α_θ(t) \log \left[ |f(it)|^{1-θ} \right] + β_θ(t) \log \left[ |f(1+it)|^θ \right] \right),$$

where

$$α_θ(t) ≡ \frac{\sin(πθ)}{2(1 - θ)[\cosh(πt) - \cos(πt)]},$$

$$β_θ(t) ≡ \frac{\sin(πθ)}{2θ[\cosh(πt) + \cos(πt)]}.$$  

For a fixed $θ \in (0, 1)$, $α_θ(t), β_θ(t) \geq 0$ for all $t \in \mathbb{R}$ and

$$\int_{-\infty}^{\infty} dt \alpha_θ(t) = \int_{-\infty}^{\infty} dt \beta_θ(t) = 1,$$

(see, e.g., [22, Exercise 1.3.8]) so that $α_θ(t)$ and $β_θ(t)$ can be interpreted as probability density functions. Furthermore,

$$\lim_{θ \downarrow 0} β_θ(t) = \frac{π}{2} [\cosh(πt) + 1]^{-1} ≡ β_0(t),$$

where $β_0$ is also a probability density function on $\mathbb{R}$. With these observations, we can see that Hirschman’s theorem implies the Hadamard three-lines theorem, given that an expectation can never exceed a supremum.

C. Complex Interpolation of Schatten Norms

We can extend much of the development above to operator-valued functions, which is needed to prove Theorem 2. Let $G : \mathcal{C} \to \mathcal{L}(\mathcal{H})$ be an operator-valued function. We say that $G(z)$ is holomorphic if every function mapping $z$ to a matrix entry is holomorphic. For our purposes in what follows, we are interested in operator-valued functions of the form $A^z$, where $A$ is a positive semi-definite operator. In this case, we apply the aforementioned convention and take $A^z = \sum_{\lambda_i \neq \lambda_j} A_i |i\rangle \langle i|$, where $A = \sum_i \lambda_i |i\rangle \langle i|$ is an eigendecomposition of $A$ with $\lambda_i \geq 0$ for all $i$. Given that $x^z$, with $x > 0$, is holomorphic, combined with the closure properties of holomorphic functions, we can conclude that $A^z$ is holomorphic if $A$ is positive semi-definite.

We can now state a version of the Hirschman theorem which applies to operator-valued functions and allows for bounding their Schatten norms [14]. The proof is standard and recalled in [8]. This theorem is one of the main technical tools that we need to establish Theorem 2.

**Theorem 5 (Stein–Hirschman):** Let $G : \overline{S} \to \mathcal{L}(\mathcal{H})$ be an operator-valued function that is bounded on $\overline{S}$, holomorphic on $S$, and continuous on the boundary $\partial S$. Let $θ \in (0, 1)$ and define $p_θ$ by

$$\frac{1}{p_θ} = \frac{1 - θ}{p_0} + \frac{θ}{p_1},$$

(16)
where \( p_0, p_1 \in [1, \infty] \). Then the following bound holds

\[
\log \| G(\theta) \|_{p_0} \leq \int_{-\infty}^{\infty} dt \alpha_0(t) \log \left[ \| G(it) \|_{p_0}^{-\theta} \right] + \int_{-\infty}^{\infty} dt \beta_0(t) \log \left[ \| G(1+it) \|_{p_1}^{\theta} \right],
\]

(17)

where \( \alpha_0(t) \) and \( \beta_0(t) \) are defined in (12)–(13).

### III. PETZ RECOVERY MAP

The recovery channel appearing in the lower bound of Theorem 2 has an explicit form and is constructed from a map known as the Petz recovery map [6], [7], which is defined as follows:

**Definition 6:** Let \( \sigma \in \mathcal{L}_+(\mathcal{H}) \), and let \( \mathcal{N} : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H}') \) be a quantum channel. The Petz recovery map \( \mathcal{P}_{\mathcal{P},\mathcal{N}} : \mathcal{L}(\mathcal{H}') \to \mathcal{L}(\mathcal{H}) \) is a completely positive, trace-non-increasing linear map defined as follows for \( Q \in \mathcal{L}(\mathcal{H}') \):

\[
\mathcal{P}_{\mathcal{P},\mathcal{N}}(Q) \equiv \sigma^{1/2} \mathcal{N} \left( \left[ \mathcal{N}(\sigma) \right]^{-1/2} Q \left[ \mathcal{N}(\sigma) \right]^{-1/2} \right) \sigma^{1/2}.
\]

(18)

The Petz recovery map \( \mathcal{P}_{\mathcal{P},\mathcal{N}} \) is linear, and it is completely positive because it is equal to a serial concatenation of three completely positive maps: \( Q \to \left[ \mathcal{N}(\sigma) \right]^{-1/2} Q \left[ \mathcal{N}(\sigma) \right]^{-1/2}, \) \( Q \to \mathcal{N}(Q), \) and \( M \to \sigma^{1/2} M \sigma^{1/2} \) for \( M \in \mathcal{L}(\mathcal{H}) \). It is trace-non-increasing because the following holds for \( Q \in \mathcal{L}_+(\mathcal{H}') \):

\[
\text{Tr} \{ \mathcal{P}_{\mathcal{P},\mathcal{N}}(Q) \} = \text{Tr} \left\{ \sigma \mathcal{N}^\dagger \left[ \left[ \mathcal{N}(\sigma) \right]^{-1/2} Q \left[ \mathcal{N}(\sigma) \right]^{-1/2} \right] \right\} \leq \text{Tr} \{ Q \}.
\]

(19)

The following lemma is one of the main reasons that \( \Delta_\alpha(\rho,\sigma,\mathcal{N}) \) generalizes a relative entropy difference.

**Lemma 9 ([16], [11]):** The following limit holds for \( \rho, \sigma, \) and \( \mathcal{N} \) as given in Definition 1:

\[
\lim_{\alpha \to 1} \Delta_\alpha(\rho,\sigma,\mathcal{N}) = D(\rho||\sigma) - D(\mathcal{N}(\rho)||\mathcal{N}(\sigma)).
\]

(20)

### IV. RÉNYI INFORMATION MEASURE

Given \( \rho, \sigma, \) and \( \mathcal{N} \) as in Definition 1, we define a Rényi information measure known as a Rényi generalization of a relative entropy difference [16]:

\[
\Delta_\alpha(\rho,\sigma,\mathcal{N}) = \frac{2\alpha}{\alpha - 1} \log \left\| \left[ \left( \mathcal{N}(\rho) \right)^{\frac{1-\alpha}{\alpha}} - I_E \right] \right\|_{2\alpha},
\]

where \( \alpha \in (0, 1) \cup (1, \infty) \) and \( U : \mathcal{H} \to \mathcal{H}' \otimes \mathcal{H}_E \) is an isometric extension of the channel \( \mathcal{N} \). That is, \( U \) is a linear isometry satisfying \( \text{Tr}_E(U \cdot U^\dagger) = \mathcal{N}(\cdot) \) and \( U^\dagger U = I_{\mathcal{H}} \).

Recall that all isometric extensions of a channel are related by an isometry acting on the environment system \( E \) [24], so that the definition above is invariant under any such choice.

The adjoint \( \mathcal{N}^\dagger \) of a channel is given in terms of an isometric extension \( U \) as \( \mathcal{N}(\cdot) = U^\dagger (\cdot) \otimes I_E \) [24, Prop. 5.2.1].

### V. PROOF OF THE RECOVERABILITY THEOREM

This section presents the proof of Theorem 2. In fact, we prove a stronger statement, which implies Theorem 2 for a particular recovery channel that we discuss below.

**Theorem 10:** Let \( \rho, \sigma, \) and \( \mathcal{N} \) be as in Definition 1. Then

\[
\Delta(\rho||\sigma) - D(\mathcal{N}(\rho)||\mathcal{N}(\sigma)) \geq \int_{-\infty}^{\infty} dt \beta_0(t) \log \left[ F \left( \rho, \left( \mathcal{R}_{\mathcal{P},\mathcal{N}}^t \circ \mathcal{N} \right)(\rho) \right) \right],
\]

(21)

where \( \beta_0(t) = \frac{\pi}{2} \left[ \cosh(\pi t) + 1 \right]^{-1} \) is a probability density function for \( t \in \mathbb{R} \) and \( \mathcal{R}_{\mathcal{P},\mathcal{N}}^t \) is a rotated Petz recovery map.

**Proof:** We can prove this result by employing Theorem 5. Let \( U : \mathcal{H} \to \mathcal{H}' \otimes \mathcal{H}_E \) be an isometric extension of the channel \( \mathcal{N} \). Pick

\[
G(z) \equiv \left[ \left( \mathcal{N}(\rho) \right)^{z/2} \left[ \mathcal{N}(\sigma) \right]^{-z/2} \otimes I_E \right] U \sigma^{1/2} \rho^{1/2},
\]

(22)

for \( z \in \mathbb{S}_+ \), \( p_0 = 2, \ p_1 = 1 \), and \( \theta \in (0, 1) \), which fixes \( p_0 = \frac{2}{1+\theta} \). The operator valued-function \( G(z) \) satisfies the conditions needed to apply Theorem 5. For the choices above,

\[
\| G(\theta) \|_{\frac{2}{1+\theta}} \leq \left\| \left( \left[ \mathcal{N}(\rho) \right]^{\frac{\theta}{2}} \left[ \mathcal{N}(\sigma) \right]^{-\frac{\theta}{2}} \otimes I_E \right) U \sigma^{1/2} \rho^{1/2} \right\|_{\frac{2}{1+\theta}},
\]

\[
\| G(\theta) \|_2 = \left\| \left( \left[ \mathcal{N}(\rho) \right]^{it/2} \left[ \mathcal{N}(\sigma) \right]^{-it/2} \otimes I_E \right) U \sigma^{it} \rho^{1/2} \right\|_2
\]

\[
\leq \left\| \sigma^{it} \right\|_2 = 1,
\]

(23)

\[
\| G(1+it) \|_1 = \left\| \left[ \mathcal{N}(\rho) \right]^{it/2} \left[ \mathcal{N}(\sigma) \right]^{-it/2} \otimes I_E \right) U \sigma^{it} \rho^{1/2} \right\|_1
\]

\[
= \sqrt{F} \left( \rho, \left( \mathcal{R}_{\mathcal{P},\mathcal{N}}^t \circ \mathcal{N} \right)(\rho) \right)
\]

\[
= \sqrt{F}(\rho, (\mathcal{R}_{\mathcal{P},\mathcal{N}}^t \circ \mathcal{N})(\rho)).
\]

(24)
Then we can apply Theorem 5 to conclude that
\[
\log \left\| \left([N(\rho)]^{\theta/2} [N(\sigma)]^{-\theta/2} \otimes I_E\right) U^{\theta/2} \rho^{1/2} \right\|^2 \leq \int_{-\infty}^{\infty} dt \, \beta_0(t) \log \left[ F \left( \rho, \left( R_{\sigma,N}^{t/2} \circ N \right)(\rho) \right) \right].
\] (31)

Letting \( \theta = (1 - \alpha)/\alpha \), we see that this is the same as
\[
\Delta_\alpha(\rho, \sigma, N) \geq \int_{-\infty}^{\infty} dt \, \beta_0(t) \log \left[ F \left( \rho, \left( R_{\sigma,N}^{t/2} \circ N \right)(\rho) \right) \right] - \log \left[ F \left( \rho, \left( R_{\sigma,N}^{t/2} \circ N \right)(\rho) \right) \right].
\] (32)

Since the inequality in (31) holds for all \( \theta \in (0, 1) \) and thus (32) holds for all \( \alpha \in (1/2, 1) \), we can take the limit as \( \alpha \to 1 \) and apply (26) and the dominated convergence theorem to conclude that (27) holds.

With the theorem above in hand, Theorem 2 follows as a consequence by taking \( R_{\sigma,N} \) to be as follows:
\[
R_{\sigma,N}(Q) = \int_{-\infty}^{\infty} dt \, \beta_0(t) \, \left( R_{\sigma,N}^{t/2} \circ N \right)(Q) + \text{Tr}\{ (I - \Pi_{N(\sigma)})Q \} \omega_t,
\] (33)
where \( Q \in L(H') \) and \( \omega \in D(H) \). This is because
\[
- \int_{-\infty}^{\infty} dt \, \beta_0(t) \log \left[ F \left( \rho, \left( R_{\sigma,N}^{t/2} \circ N \right)(\rho) \right) \right] \geq - \log \left[ F \left( \rho, \left( \int_{-\infty}^{\infty} dt \, \beta_0(t) R_{\sigma,N}^{t/2} \circ N \right)(\rho) \right) \right] = 0.
\] (34)

VII. OTHER BOUNDS

In this section, we present other bounds for the relative entropy difference \( D(\rho||\sigma) - D(N(\rho)||N(\sigma)) \), one of which is a direct consequence of Theorem 5 and [11, Theorem 4] and two others which were previously reported in [9, Remark 12]. The first bound we present is an upper bound, and to have an interpretation in terms of recoverability, we need to take \( \rho, \sigma, N \) as given in the following definition:

**Definition 12:** Let \( \rho_{SE} \) be a positive definite density operator and let \( \sigma_{SE} \) be a positive definite operator, each acting on a finite-dimensional tensor-product Hilbert space \( \mathcal{H}_S \otimes \mathcal{H}_E \). Let \( N \) be a quantum channel given as follows: \( N(\theta_{SE}) = (U_{SE}^\theta \otimes I_E) N(\theta_{SE}) (U_{SE}^{1-\theta} \otimes I_E) \), where \( U_{SE} \) is a unitary operator taking \( \mathcal{H}_S \otimes \mathcal{H}_E \) to an isomorphic finite-dimensional tensor-product Hilbert space \( \mathcal{H}_B \otimes \mathcal{H}_E \), such that \( N(\rho) \) and \( N(\sigma) \) are each positive definite and act on \( \mathcal{H}_B \).

Following the proof of Theorem 10, we pick
\[
G(z) = \left( \int_{N(\rho)}^{z^{1/2}} [N(\sigma)]^{-1/2} \otimes I_E \right) U^{1/2} \rho^{1/2},
\] (40)
for \( \rho_0 = 2, \rho_1 = \infty \), apply Theorem 5, and arrive at the following:

**Theorem 13:** For \( \rho, \sigma, N \) as given in Definition 12, the following inequality holds
\[
D(\rho||\sigma) - D(N(\rho)||N(\sigma)) \leq \int_{-\infty}^{\infty} dt \, \beta_0(t) \, D_{\text{max}} \left( \rho \left|\left| P_{\sigma,N}^{t/2} N(\rho) \right|\right| \right),
\] (41)
where \( D_{\text{max}}(\omega||\tau) = 2 \log \left\|\omega^{1/2} \tau^{-1/2} \right\|_\infty \) [25].

However, we note that there is an improvement of this bound, which is found as follows. There is an alternate Rényi generalization of a relative entropy difference [16], defined as
\[
\Delta_\alpha(\rho, \sigma, N) = \frac{2}{\alpha - 1} \log \left( \left| \int_{N(\rho)}^{1/2} [N(\sigma)]^{-\alpha/2} \otimes I_E \right| U^{1/2} \rho^{1/2} \right)\|_2^2,
\]
where \( \alpha \in (0, 1) \cup (1, \infty) \) and \( U : \mathcal{H} \to \mathcal{H}' \otimes \mathcal{H}_E \) is an isometric extension of the channel \( \mathcal{N} \). It has the property that
\[
\lim_{n \to 1} \Delta_n(\rho, \sigma, N) = D(\rho || \sigma) - D(\rho || \mathcal{N}(\rho) || \mathcal{N}(\sigma)) [16], [9].
\]
Following the proof of Theorem 10, we pick
\[
G(z) \equiv \left( [\mathcal{N}(\rho)]^{z/2} [\mathcal{N}(\sigma)]^{z/2} \otimes I_E \right) U \sigma^{-z/2} \rho^{(1+z)/2},
\]
p\(_0 = 2, p_1 = 2\), apply Theorem 5, and arrive at the following:

**Theorem 14 (Remark 12 of [9]):** For \( \rho, \sigma, \) and \( \mathcal{N} \) as given in Definition 12, the following inequality holds
\[
D(\rho || \sigma) - D(\rho || \mathcal{N}(\rho) || \mathcal{N}(\sigma)) \leq \int_{-\infty}^{\infty} dt \beta_0(t) D_2\left( \rho \bigg|\bigg| \mathcal{P}_{1/2}^{1/2}(\mathcal{N}(\rho)) \right),
\]
where \( D_2(\rho || \sigma) \equiv \log \text{Tr} \{ \rho^2 - \rho \sigma \} - 1 \). This bound is stronger than that in Theorem 13 because
\[
D_2(\omega || \tau) \leq D_{\text{max}}(\omega || \tau).
\]
where the optimization in the second line is over all density operators \( \gamma \) and the second-to-last equality follows from a variational characterization of the infinity norm.

Following the proof of Theorem 10, we pick
\[
G(z) \equiv \left( [\mathcal{N}(\rho)]^{z/2} [\mathcal{N}(\sigma)]^{z/2} \otimes I_E \right) U \sigma^{-z/2} \rho^{(1+z)/2},
\]
p\(_0 = 2, p_1 = 2\), apply Theorem 5, and arrive at the following:

**Theorem 15 (Remark 12 of [9]):** For \( \rho, \sigma, \) and \( \mathcal{N} \) as given in Definition 1, the following inequality holds
\[
D(\rho || \sigma) - D(\rho || \mathcal{N}(\rho) || \mathcal{N}(\sigma)) \geq \int_{-\infty}^{\infty} dt \beta_0(t) D_0\left( \rho \bigg|\bigg| \mathcal{P}_{1/2}^{1/2}(\mathcal{N}(\rho)) \right),
\]
where \( D_0(\omega || \tau) \equiv -\log \text{Tr} \{ \Pi_\omega, \tau \}. \)

**VIII. CONCLUSION**

We have presented several refinements of the monotonicity of quantum relative entropy. In all cases, these have an interpretation in terms of recoverability. An open question is to refine the lower and upper bounds further, perhaps with different measures or other recovery maps. Note that stronger bounds are possible in the classical case [26].

**Acknowledgements**—We thank O. Fawzi, R. Frank, E. Lieb, V. Scholz, M. Tomamichel, and T. Vidick for helpful discussions. MJ was supported by NSF DMS 1501103 and NSF-DMS 1201886. RR and DS acknowledge support by the European Research Council (ERC) via grant No. 258932, the Swiss National Science Foundation (SNSF) via the National Centre of Competence in Research “QSIT”, and by the European Commission via the project “RAQUEL”. MMW acknowledges the NSF under Award No. CCF-1350397 and DARPA Quiness through award W31PQ-12-1-0019. AWs work was supported by the EU (STREP RAQUEL), the ERC (AdG IRQUAT), the Spanish MINECO (grant FIS2013-40627-P) with the support of FEDER funds, as well as by the Generalitat de Catalunya CIRIT, project 2014-SGR-966.

**REFERENCES**


