Approximate reversibility in the context of entropy gain, information gain, and complete positivity

Francesco Buscemi,[1] Siddhartha Das,[2] and Mark M. Wilde[2,3]

[1]*Graduate School of Information Science, Nagoya University, Chikusa-ku, Nagoya, 464-8601, Japan*
[2]*Hearne Institute for Theoretical Physics, Department of Physics and Astronomy, Louisiana State University, Baton Rouge, Louisiana 70803, USA*
[3]*Center for Computation and Technology, Louisiana State University, Baton Rouge, Louisiana 70803, USA*

(Dated: January 7, 2016)

There are several inequalities in physics which limit how well we can process physical systems to achieve some intended goal, including the second law of thermodynamics, entropy bounds in quantum information theory, and the uncertainty principle of quantum mechanics. Recent results provide physically meaningful enhancements of these limiting statements, determining how well one can attempt to reverse an irreversible process. In this paper, we apply and extend these results to give strong enhancements to several entropy inequalities, having to do with entropy gain, information gain, and complete positivity of physical evolutions. Our first result is a remainder term for the entropy gain of a quantum channel. This result implies that a small increase in entropy under the action of a unital channel is a witness to the fact that the channel's adjoint can be used as a recovery channel to undo the action of the original channel. We apply this result to pure-loss, quantum-limited amplifier, and phase-insensitive quantum Gaussian channels, showing how a quantum-limited amplifier can serve as a recovery from a pure-loss channel and vice versa. Our second result regards the information gain of a quantum measurement, both without and with quantum side information. We find here that a small information gain implies that it is possible to undo the action of the original measurement if it is efficient. The result also has operational ramifications for the information-theoretic tasks known as measurement compression without and with quantum side information. We finally establish that the reduced dynamics of a system-environment interaction are approximately CPTP if and only if the data processing inequality holds approximately.

Keywords: approximate reversibility, recoverability, entropy gain, information gain, completely positive, quantum relative entropy

I. INTRODUCTION

The second law of thermodynamics constitutes a fundamental limitation on our ability to extract energy from physical systems [24]. The data processing inequality represents a limitation on our ability to process information, being the basis for most of the important capacity theorems in quantum information theory [34]. The uncertainty principle of quantum mechanics places a limitation on how well we can measure incompatible observables [10, 15]. These seemingly disparate statements have a common mathematical foundation in an entropy inequality known as the monotonicity of quantum relative entropy [19, 29], which states that the quantum relative entropy cannot increase under the action of a quantum channel. More precisely, the quantum relative entropy between two density operators ρ and σ is defined as [30]

$$D(\rho\|\sigma) \equiv \mathrm{Tr}\{\rho\,[\log\rho - \log\sigma]\}, \qquad (1)$$

and the monotonicity of quantum relative entropy states that

$$D(\rho\|\sigma) \geq D(\mathcal{N}(\rho)\|\mathcal{N}(\sigma)), \qquad (2)$$

where $\mathcal{N}$ is a quantum channel.

Recently, researchers have explored refinements of these statements in various contexts, with the common theme being to understand how well one can attempt to reverse an irreversible process. One of the main technical developments which has allowed for these refined statements is a strengthening of the monotonicity of quantum relative entropy of the following form [35]:

$$D(\rho\|\sigma) \geq D(\mathcal{N}(\rho)\|\mathcal{N}(\sigma)) - \log F(\rho, (\mathcal{R}\circ\mathcal{N})(\rho)), \quad (3)$$

where $F(\omega,\tau) \equiv \|\sqrt{\omega}\sqrt{\tau}\|_1^2$ is the quantum fidelity [28] between two density operators ω and τ, and $\mathcal{R}$ is a recovery channel with the property that it perfectly recovers the σ state, in the sense that $\sigma = (\mathcal{R}\circ\mathcal{N})(\sigma)$ (see also [17, 26] for later developments and [11] for an important earlier development with conditional mutual information). Several applications follow as a consequence. Ref. [32] gave an application in thermodynamics, proving that if the free energies of two states are close and if it is possible to transition from one state to another via a thermal operation such that there is an energy gain in the process, then one can approximately reverse this thermodynamic transition without using any energy at all. Ref. [5] showed how to tighten the uncertainty principle in the presence of quantum memory [3] with another term related to how much disturbance a given measurement causes, thus unifying several aspects of quantum physics, including measurement incompatibility, entanglement, and measurement disturbance, in a single entropic uncertainty relation. Finally, Ref. [35] has given an increased understanding of many well known entropy inequalities in quantum information, such as the joint convexity of quantum relative entropy, the non-negativity of quantum discord, the Holevo bound, and multipartite information inequalities.

In this paper, we continue with this theme and derive several new results:

1. First, we give a strong improvement of the well known statement that the quantum entropy cannot decrease under the action of a unital quantum channel (a channel which preserves the identity operator). The bound that we derive has a rather simple proof, following from the operator concavity of the logarithm (related to the method used in [26]). The main physical implication of this result is that if the entropy gain under the action of a unital channel is not too large, then it is possible to reverse the action of this channel by applying its adjoint (which is a quantum channel in this case).

2. Next, we consider the information gain of a quantum measurement, a concept introduced in [14] and subsequently refined in [4, 7, 36, 37]. The information gain of a quantum measurement quantifies how much data we can gather by performing a quantum measurement on a given state. It has an operational interpretation as the rate at which a sender needs to transmit classical information to a receiver in order for them to simulate a quantum measurement on a given state [37]. Here, we prove that if the information gain is not too large, then it is possible to reverse the action of the measurement and, in the operational context, one can also simulate the measurement well on average without sending any classical data at all. The result also applies if the measurement is performed on one share of a bipartite state.

3. Finally, we give a refinement of the recent link between the data processing inequality and complete positivity of a linear map [6]. In [6], it was shown that the data processing inequality holds if and only if the reduced dynamics of an evolution are completely positive. Here, we show how this result holds approximately, which should allow for experimental tests if desired. That is, we show that the data processing inequality holds approximately if and only if the reduced dynamics of an evolution are approximately completely positive (see Section V for precise statements).

The rest of the paper is devoted to giving more details and explanations of these results. We begin in the next section by setting notation, definitions, and reviewing the prior literature in more detail. We then follow with each of the aforementioned results and conclude in Section VI with a summary.

II. PRELIMINARIES

This section reviews background material on quantum information, all of which is available in [34]. Let $\mathcal{L}(\mathcal{H})$ denote the algebra of bounded linear operators acting on a Hilbert space $\mathcal{H}$. Let $\mathcal{L}_+(\mathcal{H})$ denote the subset of positive semi-definite operators. We also write $X \geq 0$ if $X \in \mathcal{L}_+(\mathcal{H})$. An operator ρ is in the set $\mathcal{D}(\mathcal{H})$ of density operators (or states) if $\rho \in \mathcal{L}_+(\mathcal{H})$ and $\text{Tr}\{\rho\} = 1$. The tensor product of two Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$ is denoted by $\mathcal{H}_A \otimes \mathcal{H}_B$ or $\mathcal{H}_{AB}$. Given a multipartite density operator $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$, we unambiguously write $\rho_A = \text{Tr}_B\{\rho_{AB}\}$ for the reduced density operator on system A. We use ρ_{AB}, σ_{AB}, τ_{AB}, ω_{AB}, etc. to denote general density operators in $\mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$, while ψ_{AB}, φ_{AB}, ϕ_{AB}, etc. denote rank-one density operators (pure states) in $\mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ (with it implicit, clear from the context, and the above convention implying that ψ_A, φ_A, ϕ_A may be mixed if ψ_{AB}, φ_{AB}, ϕ_{AB} are pure). A purification $|\phi^\rho\rangle_{RA} \in \mathcal{H}_R \otimes \mathcal{H}_A$ of a state $\rho_A \in \mathcal{D}(\mathcal{H}_A)$ is such that $\rho_A = \text{Tr}_R\{|\phi^\rho\rangle\langle\phi^\rho|_{RA}\}$. An isometry $U : \mathcal{H} \to \mathcal{H}'$ is a linear map such that $U^\dagger U = I_{\mathcal{H}}$. Often, an identity operator is implicit if we do not write it explicitly (and should be clear from the context).

Throughout this paper, we take the usual convention that $f(A) = \sum_i f(a_i)|i\rangle\langle i|$ when given a function f and a Hermitian operator A with spectral decomposition $A = \sum_i a_i|i\rangle\langle i|$. So this means that A^{-1} is interpreted as a generalized inverse, so that $A^{-1} = \sum_{i:a_i\neq 0} a_i^{-1}|i\rangle\langle i|$, $\log(A) = \sum_{i:a_i>0} \log(a_i)|i\rangle\langle i|$, $\exp(A) = \sum_{i:a_i\neq 0} \exp(a_i)|i\rangle\langle i|$, etc. Throughout the paper, we interpret log as the binary logarithm. We employ the shorthand $\text{supp}(A)$ and $\ker(A)$ to refer to the support and kernel of an operator A, respectively.

A linear map $\mathcal{N}_{A\to B} : \mathcal{L}(\mathcal{H}_A) \to \mathcal{L}(\mathcal{H}_B)$ is positive if $\mathcal{N}_{A\to B}(\sigma_A) \in \mathcal{L}(\mathcal{H}_B)_+$ whenever $\sigma_A \in \mathcal{L}(\mathcal{H}_A)_+$. Let id_A denote the identity map acting on a system A. A linear map $\mathcal{N}_{A\to B}$ is completely positive if the map $\text{id}_R \otimes \mathcal{N}_{A\to B}$ is positive for a reference system R of arbitrary size. A linear map $\mathcal{N}_{A\to B}$ is trace-preserving if $\text{Tr}\{\mathcal{N}_{A\to B}(\tau_A)\} = \text{Tr}\{\tau_A\}$ for all input operators $\tau_A \in \mathcal{L}(\mathcal{H}_A)$. It is trace non-increasing if $\text{Tr}\{\mathcal{N}_{A\to B}(\tau_A)\} \leq \text{Tr}\{\tau_A\}$ for all $\tau_A \in \mathcal{L}_+(\mathcal{H}_A)$. A quantum channel is a linear map which is completely positive and trace-preserving (CPTP). A positive operator-valued measure (POVM) is a set $\{\Lambda^m\}$ of positive semi-definite operators such that $\sum_m \Lambda^m = I$. For $X, Y \in \mathcal{L}(\mathcal{H})$, let $\langle X, Y\rangle \equiv \text{Tr}\{X^\dagger Y\}$ denote the Hilbert–Schmidt inner product. The adjoint $(\mathcal{M}_{A\to B})^\dagger$ of a linear map $\mathcal{M}_{A\to B}$ is the unique linear map satisfying

$$\langle Y_B, \mathcal{M}_{A\to B}(X_A)\rangle = \langle (\mathcal{M}_{A\to B})^\dagger (Y_B), X_A\rangle, \quad (4)$$

for all $X_A \in \mathcal{L}(\mathcal{H}_A)$ and $Y_B \in \mathcal{L}(\mathcal{H}_B)$. A linear map $\mathcal{M}_{A\to B}$ is unital if it preserves the identity, i.e., $\mathcal{M}_{A\to B}(I_A) = I_B$. It then follows that a linear map is unital if and only if its adjoint is trace preserving. A linear map $\mathcal{M}_{A\to B}$ is subunital if $\mathcal{M}_{A\to B}(I_A) \leq I_B$, and this is equivalent to the adjoint of $\mathcal{M}_{A\to B}$ being trace non-increasing.

A quantum instrument is a quantum channel that accepts a quantum system as input and outputs two systems: a classical one and a quantum one. More formally, a quantum instrument is a collection $\{\mathcal{N}^x\}$ of completely positive trace non-increasing maps, such that the sum map $\sum_x \mathcal{N}^x$ is a quantum channel. We can write the action of a quantum instrument on an input operator C as the following quantum channel:

$$C \to \sum_x \mathcal{N}^x(C) \otimes |x\rangle\langle x|, \qquad (5)$$

where $\{|x\rangle\}$ is an orthonormal basis labeling the classical output of the instrument.

The trace distance between two quantum states $\rho, \sigma \in \mathcal{D}(\mathcal{H})$ is equal to $\|\rho - \sigma\|_1$. It has a direct operational interpretation in terms of the distinguishability of these states. That is, if ρ or σ are prepared with equal probability and the task is to distinguish them via some quantum measurement, then the optimal success probability in doing so is equal to $(1 + \|\rho - \sigma\|_1 /2)/2$. The fidelity is defined as $F(\rho, \sigma) \equiv \left\|\sqrt{\rho}\sqrt{\sigma}\right\|_1^2$ [28], and more generally we can use the same formula to define $F(P, Q)$ if $P, Q \in \mathcal{L}_+(\mathcal{H})$. Uhlmann's theorem states that [28]

$$F(\rho_A, \sigma_A) = \max_U |\langle \phi^\sigma|_{RA} U_R \otimes I_A |\phi^\rho\rangle_{RA}|^2, \qquad (6)$$

where $|\phi^\rho\rangle_{RA}$ and $|\phi^\sigma\rangle_{RA}$ are purifications of ρ_A and σ_A, respectively, and the optimization is with respect to all isometries U_R. The same statement holds more generally for $P, Q \in \mathcal{L}_+(\mathcal{H})$. The direct-sum property of the fidelity is that

$$\sqrt{F}(\omega_{XS}, \tau_{XS}) = \sum_x \sqrt{p_X(x) q_X(x)} \sqrt{F}(\omega_S^x, \tau_S^x), \qquad (7)$$

for classical–quantum states

$$\omega_{XS} \equiv \sum_x p_X(x) |x\rangle\langle x|_X \otimes \omega_S^x, \qquad (8)$$

$$\tau_{XS} \equiv \sum_x q_X(x) |x\rangle\langle x|_X \otimes \tau_S^x. \qquad (9)$$

The relative entropy $D(P\|Q)$ between $P, Q \in \mathcal{L}_+(\mathcal{H})$, with $P \neq 0$, is defined as

$$D(P\|Q) = \text{Tr}\{P[\log P - \log Q]\} \qquad (10)$$

if $\text{supp}(P) \subseteq \text{supp}(Q)$ and as $+\infty$ otherwise. The relative entropy $D(P\|Q)$ is non-negative if $\text{Tr}\{P\} \geq \text{Tr}\{Q\}$, a result known as Klein's inequality [18]. Thus, for density operators ρ and σ, the relative entropy is non-negative, and furthermore, it is equal to zero if and only if $\rho = \sigma$. The quantum relative entropy obeys the following property:

$$D(P\|Q) \geq D(P\|Q'), \qquad (11)$$

for $P, Q, Q' \in \mathcal{L}_+(\mathcal{H})$ such that $Q \leq Q'$. The following relationship between fidelity and quantum relative entropy is well known (see, e.g., [20]):

$$D(P\|Q) \geq -\log F(P, Q). \qquad (12)$$

The quantum entropy $H(\rho)$ of a density operator ρ is $H(\rho) = -\text{Tr}\{\rho \log \rho\}$. We often write this as $H(A)_\rho$ if ρ_A is the density operator for system A. The conditional entropy of a bipartite density operator ρ_{AB} is equal to $H(A|B)_\rho \equiv H(AB)_\rho - H(B)_\rho$. The mutual information is equal to $I(A;B)_\rho = H(A)_\rho - H(A|B)_\rho$. The conditional mutual information of a tripartite state ρ_{ABC} is equal to $I(A;B|C)_\rho = H(B|C)_\rho - H(B|AC)_\rho$. The following identities are well known (see, e.g., [34]):

$$H(A)_\rho = -D(\rho_A\|I_A), \qquad (13)$$

$$H(A|B)_\rho = -D(\rho_{AB}\|I_A \otimes \rho_B), \qquad (14)$$

$$I(A;B)_\rho = D(\rho_{AB}\|\rho_A \otimes \rho_B). \qquad (15)$$

The following "recoverability theorem" is an enhancement of the monotonicity of quantum relative entropy (mentioned in (3)) and was proved recently in [17], by an extension of the methods from [35]:

$$D(\rho\|\sigma) \geq D(\mathcal{N}(\rho)\|\mathcal{N}(\sigma)) - \log F(\rho, (\mathcal{R} \circ \mathcal{N})(\rho)), \quad (16)$$

where $\rho \in \mathcal{D}(\mathcal{H})$, $\sigma \in \mathcal{L}_+(\mathcal{H})$, $\mathcal{N} : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H}')$ is a quantum channel, and $\mathcal{R}$ is a recovery quantum channel of the following form:

$$\mathcal{R}(Q) \equiv \text{Tr}\{(I - \Pi_{\mathcal{N}(\sigma)})Q\}\tau + \int_{-\infty}^{\infty} dt\ p(t)\ \mathcal{R}^{t/2}_{\sigma,\mathcal{N}}(Q), \quad (17)$$

where $\Pi_{\mathcal{N}(\sigma)}$ is the projection onto the support of $\mathcal{N}(\sigma)$,

$$\tau \in \mathcal{D}(\mathcal{H}), \qquad p(t) \equiv \frac{\pi}{2}\left[\cosh(t) + 1\right]^{-1} \qquad (18)$$

is a probability distribution on $t \in \mathcal{R}$, $\mathcal{U}_{\omega,t}(X) \equiv \omega^{it} X \omega^{-it}$ for ω positive semi-definite,

$$\mathcal{P}_{\sigma,\mathcal{N}}(Q) \equiv \sigma^{1/2}\mathcal{N}^\dagger\left(\mathcal{N}(\sigma)^{-1/2} Q \mathcal{N}(\sigma)^{-1/2}\right)\sigma^{1/2} \quad (19)$$

is a completely positive, trace non-increasing map known as the Petz recovery map [22, 23], and $\mathcal{R}^t_{\sigma,\mathcal{N}}$ is a rotated or "swiveled" Petz recovery map, defined as

$$\mathcal{R}^t_{\sigma,\mathcal{N}} \equiv \mathcal{U}_{\sigma,-t} \circ \mathcal{P}_{\sigma,\mathcal{N}} \circ \mathcal{U}_{\mathcal{N}(\sigma),t}. \qquad (20)$$

In fact, the following stronger statement holds [17]

$$D(\rho\|\sigma) \geq D(\mathcal{N}(\rho)\|\mathcal{N}(\sigma)) - \int_{-\infty}^{\infty} dt\ p(t) \log F(\rho, (\mathcal{R}^{t/2}_{\sigma,\mathcal{N}} \circ \mathcal{N})(\rho)), \quad (21)$$

which will be useful for our purposes here. The inequality in (16) implies the following one:

$$I(A;B|C)_\rho \geq -\log F(\rho_{ABC}, \mathcal{R}_{C\to AC}(\rho_{BC})), \qquad (22)$$

where $\mathcal{R}_{C\to AC}$ is defined from (17), taking $\sigma = \rho_{AC}$ and $\mathcal{N} = \mathrm{Tr}_A$. This follows from the definition we gave for $I(A;B|C)_\rho$, the equality in (14), and the inequality in (16). Similarly, the following holds as well:

$$I(A;B|C)_\rho \geq -\int_{-\infty}^{\infty} dt\ p(t) \log F(\rho_{ABC}, \mathcal{R}^{t/2}_{\sigma,\mathcal{N}}(\rho_{BC})), \tag{23}$$

taking $\sigma = \rho_{AC}$ and $\mathcal{N} = \mathrm{Tr}_A$.

III. ENTROPY GAIN

It is well known that the quantum entropy cannot decrease under the action of a unital quantum channel [33]:

$$H(\mathcal{N}(\rho)) \geq H(\rho), \tag{24}$$

where $\rho \in \mathcal{D}(\mathcal{H})$ and $\mathcal{N} : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H}')$ is a unital quantum channel. This entropy inequality follows as a simple consequence of the monotonicity of quantum relative entropy, established by picking $\sigma = I$ in (2) and applying that $\mathcal{N}$ is unital, whereby

$$-H(\rho) = D(\rho\|I) \geq D(\mathcal{N}(\rho)\|\mathcal{N}(I)) \tag{25}$$
$$= D(\mathcal{N}(\rho)\|I) \tag{26}$$
$$= -H(\mathcal{N}(\rho)). \tag{27}$$

This entropy inequality has a number of applications in quantum information and other contexts.

The following theorem leads to an enhancement of (24):

Theorem 1 *Let $\rho \in \mathcal{D}(\mathcal{H})$ and let $\mathcal{N} : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H}')$ be a positive and trace-preserving map. Then*

$$H(\mathcal{N}(\rho)) - H(\rho) \geq D(\rho\|(\mathcal{N}^\dagger \circ \mathcal{N})(\rho)). \tag{28}$$

Proof. This follows because

$$H(\mathcal{N}(\rho)) - H(\rho)$$
$$= \mathrm{Tr}\{\rho \log \rho\} - \mathrm{Tr}\{\mathcal{N}(\rho) \log \mathcal{N}(\rho)\} \tag{29}$$
$$= \mathrm{Tr}\{\rho \log \rho\} - \mathrm{Tr}\{\rho \mathcal{N}^\dagger(\log \mathcal{N}(\rho))\} \tag{30}$$
$$\geq \mathrm{Tr}\{\rho \log \rho\} - \mathrm{Tr}\{\rho \log (\mathcal{N}^\dagger \circ \mathcal{N})(\rho)\} \tag{31}$$
$$= D(\rho\|(\mathcal{N}^\dagger \circ \mathcal{N})(\rho)). \tag{32}$$

The second equality is from the definition of the adjoint. The inequality follows from operator concavity of the logarithm and the operator Jensen inequality for positive unital maps [9] (see also [26, Lemma 3.10]). ■

If $\mathcal{N}$ is additionally subunital, then Theorem 1 implies that $\mathcal{N}^\dagger$ is trace non-increasing, which in turn implies that $D(\rho\|(\mathcal{N}^\dagger \circ \mathcal{N})(\rho)) \geq 0$ by Klein's inequality. Thus, in this case, we obtain a significant strengthening of the well known fact that the entropy increases under the action of a unital quantum channel. The resulting entropy inequality also leads to an interpretation in terms of recoverability, in the sense discussed in [35]. That is, we can take the recovery channel to be

$$\mathcal{R}(Y) \equiv \mathcal{N}^\dagger(Y) + \mathrm{Tr}\{(\mathrm{id} - \mathcal{N}^\dagger)(Y)\}\tau, \tag{33}$$

where $\tau \in \mathcal{D}(\mathcal{H})$, and we get that

$$H(\mathcal{N}(\rho)) - H(\rho) \geq D(\rho\|(\mathcal{R} \circ \mathcal{N})(\rho)) \tag{34}$$

by applying (11), because $(\mathcal{R} \circ \mathcal{N})(\rho) \geq (\mathcal{N}^\dagger \circ \mathcal{N})(\rho)$. Note that $\mathcal{R}$ is a positive map if $\mathcal{N}$ is.

Thus, what we find is an improvement over what we would get by applying (16) or the main result of [26]. First, there is a mathematical advantage in the sense that $\mathcal{N}$ is not required to be a channel, but it suffices for it to be a positive map. This addresses the main open question of [26] for a very special case. Some might also consider this to be a physical advantage as well, given the controversy over positive versus completely positive maps for the description of quantum dynamical evolutions (see, e.g., [6] and references therein). Second, the remainder term in (34) features the quantum relative entropy and thus is stronger than the $-\log F$ bound in (16) (cf. (12)) and the "measured relative entropy" term from [26].

A. Application to bosonic channels

Theorem 1 finds application for practical bosonic channels that have a long history in quantum information theory, in particular, the pure-loss and quantum-limited amplifier channels, and even all phase insensitive Gaussian channels [31]. A pure-loss channel is defined from the following input-output Heisenberg-picture relation:

$$\hat{b} = \sqrt{\eta}\hat{a} + \sqrt{1-\eta}\hat{e}, \tag{35}$$

where $\hat{a}$, $\hat{b}$, and $\hat{e}$ are the field-mode annihilation operators representing the sender's input, the receiver's output, and the environmental input of the channel. The parameter $\eta \in [0,1]$ represents the average fraction of photons that make from the sender to receiver. For the pure-loss channel, the environment is prepared in the vacuum state. Let $\mathcal{B}_\eta$ denote the CPTP map corresponding to this channel. A quantum-limited amplifier channel is defined from the following input-output Heisenberg-picture relation:

$$\hat{b} = \sqrt{G}\hat{a} + \sqrt{G-1}\hat{e}^\dagger, \tag{36}$$

where $\hat{a}$, $\hat{b}$, and $\hat{e}$ have the same physical meaning as given for the pure-loss channel. The parameter $G \in [1,\infty)$ represents the gain or amplification factor of the channel. For the quantum-limited amplifier channel, the environment is prepared in the vacuum state. Let $\mathcal{A}_G$ denote the CPTP map corresponding to this channel.

One of the critical insights of [16] is that these channels are "almost unital," in the sense that

$$\mathcal{B}_\eta(I) = \eta^{-2} I, \qquad \mathcal{A}_G(I) = G^{-2} I, \tag{37}$$

and that their adjoints are strongly related, in the sense that

$$\mathcal{B}_\eta^\dagger = \eta^{-2}\mathcal{A}_{1/\eta}, \tag{38}$$

$$\mathcal{A}_G^\dagger = G^{-2}\mathcal{B}_{1/G}. \tag{39}$$

Observe that the pure-loss channel is superunital and the amplifier channel is subunital. These facts allow us to apply Theorem 1 and the fact that $D(\rho\|c\sigma) = D(\rho\|\sigma) - \log c$ for $c > 0$ to find that

$$H(\mathcal{B}_\eta(\rho)) - H(\rho) \geq D(\rho\|(\mathcal{A}_{\frac{1}{\eta}} \circ \mathcal{B}_\eta)(\rho)) + 2\log\eta, \tag{40}$$

$$H(\mathcal{A}_G(\rho)) - H(\rho) \geq D(\rho\|(\mathcal{B}_{\frac{1}{G}} \circ \mathcal{A}_G)(\rho)) + 2\log G. \tag{41}$$

These bounds demonstrate that a quantum-limited amplifier suffices as a reversal channel for a pure-loss channel and vice versa. Note that the reversal is only good if the entropy does not change too much (i.e., if $\eta \approx 1$ or $G \approx 1$). We can also conclude that

$$H((\mathcal{A}_G \circ \mathcal{B}_\eta)(\rho)) - H(\rho) \geq D(\rho\|(\mathcal{A}_{1/\eta} \circ \mathcal{B}_{1/G} \circ \mathcal{A}_G \circ \mathcal{B}_\eta)(\rho)) + 2\log[\eta G], \tag{42}$$

because

$$(\mathcal{A}_G \circ \mathcal{B}_\eta)^\dagger = [\eta G]^{-2}\mathcal{A}_{1/\eta} \circ \mathcal{B}_{1/G}. \tag{43}$$

The above bound applies to any phase insensitive quantum Gaussian channel, given that any such channel can be written as a serial concatenation of a pure-loss channel and a quantum-limited amplifier channel [8, 13].

B. Optimized entropy gain

In [1], the minimal entropy gain of a quantum channel was defined as

$$G(\mathcal{N}) \equiv \inf_\rho [H(\mathcal{N}(\rho)) - H(\rho)], \tag{44}$$

and the following bounds were established for a channel with the same input and output Hilbert space $\mathcal{H}$:

$$-\log\dim(\mathcal{H}) \leq G(\mathcal{N}) \leq 0. \tag{45}$$

Applying Theorem 1 gives the following alternate lower bound for the entropy gain of a quantum channel:

$$G(\mathcal{N}) \geq \inf_\rho D(\rho\|(\mathcal{N}^\dagger \circ \mathcal{N})(\rho)). \tag{46}$$

C. Entropy gain in the presence of quantum side information

A generalization of the entropy inequality in (34) holds for the case of the conditional quantum entropy, found by applying the same method:

Corollary 2 *Let $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ and $\mathcal{N} : \mathcal{L}(\mathcal{H}_A) \to \mathcal{L}(\mathcal{H}_{A'})$ be a positive and trace-preserving map. Then*

$$H(A'|B)_\sigma - H(A|B)_\rho \geq D(\rho_{AB}\|((\mathcal{N}_{A\to A'})^\dagger \circ \mathcal{N}_{A\to A'})(\rho_{AB})), \tag{47}$$

where $\sigma_{A'B} \equiv (\mathcal{N}_{A\to A'} \otimes \mathrm{id}_B)(\rho_{AB})$.

Proof. This follows by applying Theorem 1 and definitions. From Theorem 1, we can conclude that

$$H(A'B)_\sigma - H(AB)_\rho \geq D(\rho_{AB}\|((\mathcal{N}_{A\to A'})^\dagger \circ \mathcal{N}_{A\to A'})(\rho_{AB})). \tag{48}$$

Consider also that

$$H(A'B)_\sigma - H(AB)_\rho = H(A'B)_\sigma - H(B)_\sigma - [H(AB)_\rho - H(B)_\rho] \tag{49}$$

$$= H(A'|B)_\sigma - H(A|B)_\rho, \tag{50}$$

where we have used that $H(B)_\rho = H(B)_\sigma$. Combining these gives (47). ■

IV. INFORMATION GAIN

Let $\{\mathcal{N}^x\}$ constitute a quantum instrument, where each $\mathcal{N}^x : \mathcal{L}(\mathcal{H}_A) \to \mathcal{L}(\mathcal{H}_{A'})$ is a completely positive, trace-non-increasing map. Groenewold originally defined the information gain of a quantum instrument $\{\mathcal{N}^x\}$, when performed on a quantum state ρ_A, as follows [14]:

$$H(\rho_A) - \sum_x p_X(x)H(\rho_{A'}^x), \tag{51}$$

where

$$\rho_{A'}^x \equiv \frac{\mathcal{N}_{A\to A'}^x(\rho_A)}{p_X(x)}, \qquad p_X(x) \equiv \mathrm{Tr}\{\mathcal{N}_{A\to A'}^x(\rho_A)\}. \tag{52}$$

This definition was based on the physical intuition that information gain should be identified with the entropy reduction of the measurement. However, it was later realized that the entropy reduction can be negative, and that this happens if and only if the instrument is not an efficient measurement [21] (such that each $\mathcal{N}^x$ consists of a single Kraus operator [12]).

Apparently without realizing the connection to Groenewold's information gain of a measurement, Winter considered the operational, information-theoretic task [37] of determining the rate at which classical information would need to be communicated from a sender to a receiver in order to simulate the action of the measurement on a given state (if shared randomness is allowed for free between sender and receiver). He called this task "measurement compression," given that the goal is to send the classical output of the measurement at the smallest rate possible, in such a way that a third party would not be able to distinguish the output of the protocol performed on many copies of ϕ^{ρ}_{RA} from the same number of copies of the following state:

$$\sigma_{RX} \equiv \sum_x \mathrm{Tr}_{A'}\{(\mathrm{id}_R \otimes \mathcal{N}^x_{A\to A'})(\phi^{\rho}_{RA})\} \otimes |x\rangle\langle x|_X, \quad (53)$$

where ϕ^{ρ}_{RA} is a purification of ρ and $\{|x\rangle\}$ is an orthonormal basis for the classical output X of the measurement. He found that the optimal rate of measurement compression is equal to the mutual information of the measurement $I(R;X)_\sigma$.

After Winter's development, Ref. [7] suggested that the information gain of the measurement should be defined as its mutual information. The advantage of such an approach is that the mutual information $I(R;X)_\sigma$ is non-negative and has a clear operational interpretation. Furthermore, it is equal to the entropy reduction in (51) for efficient measurements [7] and thus connects with Groenewold's original intuition.

Winter's result was later extended in two different directions. First, Ref. [36] allowed for a correlated initial state ρ_{AB}, shared between the sender and receiver before communication begins. In this case, the optimal rate at which the sender needs to transmit classical information in order to simulate the measurement is equal to the conditional mutual information $I(R;X|B)_\omega$, where the conditional mutual information is with respect to the following state:

$$\omega_{RBX} \equiv \sum_x \mathrm{Tr}_{A'}\{(\mathrm{id}_R \otimes \mathcal{N}^x_{A\to A'})(\phi^{\rho}_{RAB})\} \otimes |x\rangle\langle x|_X, \quad (54)$$

and ϕ^{ρ}_{RAB} is a purification of ρ_{AB}. We can thus call $I(R;X|B)_\omega$ the information gain in the presence of quantum side information (IG-QSI), and the information-processing task is known as measurement compression with quantum side information [36]. In general, the IG-QSI is smaller than $I(RB;X)_\omega$, which is the rate at which classical communication would need to be transmitted if the receiver does not make use of the B system. The other extension of Winter's result was to determine the rate required to simulate the instrument on an arbitrary input state, and the optimal rate was proved to be equal to the optimized information gain

$$\max_\rho I(R;X)_\sigma, \quad (55)$$

where the optimization is with respect to all input states ρ_A leading to a purification ϕ^{ρ}_{RA} [4].

A. Information gain without quantum side information

In what follows, we demonstrate how the refined entropy inequality in (16) has implications for the information gain of a quantum measurement, both without and with quantum side information. We begin with the following result, which applies to the setting without quantum side information:

Theorem 3 *Let $\rho \in \mathcal{D}(\mathcal{H}_A)$ and $\{\mathcal{N}^x\}$ be a quantum instrument, where each $\mathcal{N}^x : \mathcal{L}(\mathcal{H}_A) \to \mathcal{L}(\mathcal{H}_{A'})$. Then the following inequality holds*

$$I(R;X)_\sigma \geq -\log F(\sigma_{RX}, \sigma_R \otimes \sigma_X). \quad (56)$$

If the quantum instrument is efficient, then the above inequality implies that

$$I(R;X)_\sigma \geq -2\log\left[\sum_x p_X(x)\sqrt{F}(\mathcal{U}^x_{A'\to A}(\phi^{\rho_x}_{RA'}), \phi^{\rho}_{RA})\right], \quad (57)$$

for some collection $\{\mathcal{U}^x_{A'\to A}\}$, where each $\mathcal{U}^x_{A'\to A}$ is an isometric quantum channel, $\phi^{\rho_x}_{RA'}$ is a purification of $\rho^x_{A'}$ defined in (52), and $p_X(x)$ is defined in (52).

Proof. The inequality in (56) is a simple consequence of (15) and (12). The inequality in (57) follows because

$$\sqrt{F}(\sigma_{RX}, \sigma_R \otimes \sigma_X) = \sum_x p_X(x)\sqrt{F}(\phi^{\rho_x}_R, \phi^{\rho}_R). \quad (58)$$

Applying Uhlmann's theorem (see (6)), we can conclude that there exist isometric channels $\mathcal{U}^x_{A'\to A}$ such that $F(\phi^{\rho_x}_R, \phi^{\rho}_R) = F(\mathcal{U}^x_{A'\to A}(\phi^{\rho_x}_{RA'}), \phi^{\rho}_{RA})$ for all x. ■

The implication of the inequality in (57) is that if the information gain of the measurement is small, so that

$$I(R;X)_\sigma \approx 0, \quad (59)$$

then it possible to reverse the action of the measurement approximately, in such a way as to restore the post-measurement state to the original state with a fidelity

$$\sum_x p_X(x)\sqrt{F}(\mathcal{U}^x_{A'\to A}(\phi^{\rho_x}_{RA'}), \phi^{\rho}_{RA}) \approx 1. \quad (60)$$

We can thus view this result as a one-sided information-disturbance trade-off. [7, Theorem 1] contains an observation related to this, but the result here is stronger because it makes a statement about average fidelity and restoration of the full state on both systems RA. The observation above is also related to the general one from [25] (upon which [7, Theorem 1] relies), but the result given here is again stronger: an inability to find correction isometries, which leads to a small fidelity, is a witness to having a large information gain $I(R;X)_\sigma$, due to the presence of the negative logarithm in (57).

The inequality in (57) also has an operational implication for Winter's measurement compression task. If the information gain is small, so that (59) holds, then the sender and receiver can simulate the measurement with a high fidelity per copy of the source state, in such a way that the sender does not need to transmit any classical information at all. The receiver can just prepare many copies of ρ_A locally, perform the measurements, and deliver the outputs of the measurements as the classical data. This situation occurs because the reference system R is approximately decoupled from the classical output, in the sense that $F(\sigma_{RX}, \sigma_R \otimes \sigma_X) \approx 1$ if $I(R;X)_\sigma \approx 0$.

B. Information gain with quantum side information

We can readily extend the above results to the case of quantum side information, by employing the inequality in (22). This leads to the following theorem:

Theorem 4 *Let $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ and $\{\mathcal{N}^x\}$ be a quantum instrument, where each $\mathcal{N}^x : \mathcal{L}(\mathcal{H}_A) \to \mathcal{L}(\mathcal{H}_{A'})$. Then the following inequality holds*

$$I(R;X|B)_\omega \geq -2\int_{-\infty}^{\infty} dt\; p(t) \log\left[\sum_x p_X(x)\sqrt{F}(\omega^x_{RB}, \mathcal{R}^{x,t/2}_B(\omega_{RB}))\right], \tag{61}$$

where

$$\omega_{RBX} \equiv \sum_x \operatorname{Tr}_{A'}\{\mathcal{N}^x_{A\to A'}(\phi^\rho_{RAB})\} \otimes |x\rangle\langle x|_X, \tag{62}$$

ϕ^ρ_{RAB} is a purification of ρ_{AB},

$$\omega^x_{RBA'} \equiv \frac{\mathcal{N}^x_{A\to A'}(\phi^\rho_{RAB})}{p_X(x)}, \tag{63}$$

$$p_X(x) \equiv \operatorname{Tr}\{\mathcal{N}^x_{A\to A'}(\phi^\rho_{RAB})\}, \tag{64}$$

$\{p_X(x)\mathcal{R}^{x,t/2}_B\}$ is a quantum instrument defined by

$$\mathcal{R}^{x,t/2}_B(\omega_{RB}) \equiv (\omega^x_B)^{\frac{1-it}{2}}\, \omega_B^{\frac{1+it}{2}}(\omega_{RB})\omega_B^{\frac{1-it}{2}}\, (\omega^x_B)^{\frac{1+it}{2}}, \tag{65}$$

and $p(t)$ is defined in (18). If the instrument $\{\mathcal{N}^x\}$ is efficient, then the following inequality holds as well:

$$I(R;X|B)_\omega \geq -2\int_{-\infty}^{\infty} dt\; p(t)\log\left[\sum_x p_X(x)\sqrt{F_{x,t}}\right], \tag{66}$$

for some collection $\{\mathcal{U}^{x,t}_{A\to A'}\}$, where

$$F_{x,t} \equiv F(\omega^x_{RBA'}, (\mathcal{R}^{x,t/2}_B \otimes \mathcal{U}^{x,t}_{A\to A'})(\phi^\rho_{RBA})) \tag{67}$$

and each $\mathcal{U}^{x,t}_{A\to A'}$ is an isometric quantum channel.

Proof. We begin by proving the inequality in (61). Consider that

$$I(R;X|B)_\omega \geq -\int_{-\infty}^{\infty} dt\; p(t)\log F(\omega_{RBX}, \mathcal{R}^{t/2}_{B\to BX}(\omega_{RB})), \tag{68}$$

which is a direct consequence of (23). By a direct calculation, we find that

$$\mathcal{R}^{t/2}_{B\to BX}(\omega_{RB}) = \sum_x p_X(x)|x\rangle\langle x|_X \otimes \mathcal{R}^{x,t/2}_B(\omega_{RB}). \tag{69}$$

This then leads to the inequality in (61), by applying the direct sum property of fidelity. The inequality in (66) is an application of Uhlmann's theorem, after observing that the rank-one operator $\mathcal{R}^{x,t/2}_B(\phi^\rho_{RBA})$ purifies $\mathcal{R}^{x,t/2}_B(\omega_{RB})$ and the rank-one operator $\omega^x_{RBA'}$ purifies ω^x_{RB}. The aforementioned operators are rank-one if the measurement is efficient (which is what we assumed in the statement of the theorem). ■

The implications of Theorem 4 are similar to those of Theorem 3, except they apply to a setting in which quantum side information is available. If the information gain of the measurement is small, so that

$$I(R;X|B)_\omega \approx 0, \tag{70}$$

then it possible to reverse the action of the measurement approximately, in such a way as to restore the post-measurement state of systems RA' to the original state on systems RA with a fidelity larger than

$$\int_{-\infty}^{\infty} dt\; p(t)\sum_x p_X(x)\sqrt{F_{x,t}} \approx 1. \tag{71}$$

This follows from the concavity of the fidelity. The reversal operation consists of two steps. First, Bob performs the instrument $\{p_X(x)\mathcal{R}^{x,t/2}_B\}$. He then forwards the outcomes to Alice, who performs a channel corresponding to the inverse of the isometric quantum channel $\mathcal{U}^{x,t}_{A\to A'}$. Then, the average fidelity is high if the information gain is small. We can view this result as a one-sided information-disturbance trade-off which extends the aforementioned one without quantum side information.

The inequality in (66) also has an operational implication for measurement compression with quantum side information [36]. If the IG-QSI is small, so that (70) holds, then the sender and receiver can simulate the measurement with a high fidelity per copy of the source state, in such a way that the sender does not need to transmit any classical information at all. The receiver can just perform the instrument $\{p_X(x)\mathcal{R}^{x,t/2}_B\}$ with probability $p(t)$ on the individual B systems of many copies of ρ_{AB} and deliver the classical outputs of the measurements as the classical data. This situation occurs because the X system of ω_{RBX} is approximately recoverable from B alone, in the sense that $\int_{-\infty}^{\infty} dt\; p(t)\sum_x p_X(x)\sqrt{F_{x,t}} \approx 1$ if $I(R;X|B)_\omega \approx 0$. This latter result might have implications for quantum communication complexity (cf. [27]).

V. COMPLETE POSITIVITY

This section demonstrates how the inequality in (22) and the Alicki–Fannes–Winter inequality [2, 38] lead to a robust version of the main conclusion of [6], which links the data processing inequality to complete positivity. Let us recall the operational framework considered in [6].

1. At some time $t = \tau$, we fix a tripartite configuration, i.e., an arbitrary tripartite density operator ρ_{RQE}, describing the initial correlations between

the system Q, its environment E, and a reference R. It can be helpful to imagine that the dynamics are "frozen" at the intermediate time $t = \tau$, when the correlations between R, Q, and E are arbitrary.

2. We move to the next instant in time, i.e., $t = \tau + \Delta$. We assume that the system-environment pair evolves from τ to $\tau + \Delta$ according to some unitary operator V, while the reference R remains unchanged.

3. Denoting by Q' and E' the system and the environment after the evolution described by V has taken place, the tripartite configuration ρ_{RQE} has evolved to the tripartite configuration $\sigma_{RQ'E'} = (I_R \otimes V_{QE})\rho_{RQE}(I_R \otimes V_{QE}^\dagger)$.

4. We then look at the reduced reference-system dynamics (i.e., the transformation mapping ρ_{RQ} to $\sigma_{RQ'}$) and check whether they are compatible with the application of a completely positive trace-preserving linear map on the system Q *alone*. More explicitly, we check whether there exists a completely positive trace-preserving linear map $\mathcal{E}$, mapping Q to Q', such that $\sigma_{RQ'} = (\mathrm{id}_R \otimes \mathcal{E}_Q)(\rho_{RQ})$.

Theorem 5 *Fix a tripartite configuration ρ_{RQE}. Suppose that the data processing inequality holds approximately for all joint system-environment evolutions $V_{QE\to Q'E'}$, i.e.,*

$$I(R;Q')_\sigma \le I(R;Q)_\rho + \varepsilon, \tag{72}$$

where $\varepsilon > 0$ and

$$\sigma_{RQ'E'} = V_{QE\to Q'E'}\rho_{RQE}V_{QE\to Q'E'}^\dagger. \tag{73}$$

Then the conditional mutual information is nearly equal to zero:

$$I(R;E|Q)_\rho \le \varepsilon, \tag{74}$$

and the reduced dynamics are approximately CPTP, i.e., to every unitary interaction $V_{QE\to Q'E'}$, there exists a CPTP map $\mathcal{E}_{Q\to Q'}$ such that

$$-\log F\left(\sigma_{RQ'}, \mathcal{E}_{Q\to Q'}\left(\rho_{RQ}\right)\right) \le \varepsilon. \tag{75}$$

Proof. We begin by proving (74) with the same approach used in [6]. Consider the particular evolution in which $Q' = QE$ and system E' is trivial. The assumption that data processing holds approximately gives that

$$I(R;Q)_\rho + \varepsilon \ge I(R;Q')_\sigma = I(R;QE)_\rho. \tag{76}$$

We can rewrite this inequality using the chain rule for conditional mutual information as

$$\varepsilon \ge I(R;QE)_\rho - I(R;Q)_\rho = I(R;E|Q)_\rho, \tag{77}$$

which proves (74). Now, from the inequality in (22), we know that there exists a recovery map $\mathcal{R}_{Q\to QE}$ such that

$$I(R;E|Q)_\rho \ge -\log F\left(\rho_{RQE}, \mathcal{R}_{Q\to QE}\left(\rho_{RQ}\right)\right) \tag{78}$$

Since the fidelity is invariant with respect to unitaries, we find (abbreviating $V_{QE\to Q'E'}$ as V) that

$$\begin{aligned} &F\left(\rho_{RQE}, \mathcal{E}_{Q\to QE}\left(\rho_{RQ}\right)\right) \\ &= F\left(V\rho_{RQE}V^\dagger, V\mathcal{R}_{Q\to QE}\left(\rho_{RQ}\right)V^\dagger\right) \quad (79)\\ &= F\left(\sigma_{RQ'E'}, V\mathcal{R}_{Q\to QE}\left(\rho_{RQ}\right)V^\dagger\right) \quad (80)\\ &\le F\left(\sigma_{RQ'}, \mathrm{Tr}_{E'}\left\{V\mathcal{R}_{Q\to QE}\left(\rho_{RQ}\right)V^\dagger\right\}\right), \quad (81) \end{aligned}$$

where the inequality follows from monotonicity of fidelity under the discarding of subsystems. By defining the map

$$\mathcal{E}_{Q\to Q'}\left(\cdot\right) \equiv \mathrm{Tr}_{E'}\left\{V_{QE\to Q'E'}\mathcal{R}_{Q\to QE}\left(\cdot\right)V_{QE\to Q'E'}^\dagger\right\}, \tag{82}$$

we find that

$$\varepsilon \ge -\log F\left(\sigma_{RQ'}, \mathcal{E}_{Q\to Q'}\left(\rho_{RQ}\right)\right), \tag{83}$$

establishing (75). ■

Theorem 6 *Suppose that the reduced dynamics are approximately CPTP, i.e., that to every unitary interaction $V_{QE\to Q'E'}$ leading to*

$$\sigma_{RQ'} = \mathrm{Tr}_{E'}\left\{V_{QE\to Q'E'}\rho_{RQE}V_{QE\to Q'E'}^\dagger\right\}, \tag{84}$$

there exists a CPTP map $\mathcal{E}_{Q\to Q'}$ such that

$$\frac{1}{2}\left\|\sigma_{RQ'} - \mathcal{E}_{Q\to Q'}\left(\rho_{RQ}\right)\right\|_1 \le \varepsilon, \tag{85}$$

where $\varepsilon \in [0,1]$. Then the quantum data processing inequality is satisfied approximately, in the sense that

$$\begin{aligned} I(R;Q')_\sigma &\le I(R;Q)_\rho \\ &\quad + 2\varepsilon\log|R| + (1+\varepsilon)\,h_2(\varepsilon/(1+\varepsilon)), \quad (86) \end{aligned}$$

and the conditional mutual information is nearly equal to zero as well:

$$I(R;E|Q)_\rho \le 2\varepsilon\log|R| + (1+\varepsilon)\,h_2(\varepsilon/(1+\varepsilon)). \tag{87}$$

Proof. This follows directly from the assumption in (85), the Alicki–Fannes–Winter inequality, and the quantum data processing inequality:

$$\begin{aligned} &I(R;Q')_\sigma \\ &= H(R)_\sigma - H(R|Q')_\sigma \\ &= H(R)_{\mathcal{E}(\rho)} - H(R|Q')_\sigma \\ &\le H(R)_{\mathcal{E}(\rho)} - H(R|Q')_{\mathcal{E}(\rho)} \\ &\qquad + 2\varepsilon\log|R| + (1+\varepsilon)\,h_2(\varepsilon/(1+\varepsilon)) \\ &= I(R;Q')_{\mathcal{E}(\rho)} + 2\varepsilon\log|R| + (1+\varepsilon)\,h_2(\varepsilon/(1+\varepsilon)) \\ &\le I(R;Q)_\rho + 2\varepsilon\log|R| + (1+\varepsilon)\,h_2(\varepsilon/(1+\varepsilon)). \quad (88) \end{aligned}$$

The inequality for conditional mutual information follows the same reasoning we used to arrive at (77). ■

VI. CONCLUSION

We have shown how recent results regarding recoverability give enhancements to several entropy inequalities, having to do with entropy gain, information gain, and complete positivity. Our first result is a remainder term for the entropy gain of a quantum channel, which for unital channels is stronger than that which is obtained by directly applying the results of [17, 26]. This result implies that a small increase in entropy under a unital channel is a witness to the fact that the channel's adjoint can be used as a recovery channel to undo the action of the original channel. Our second result regards the information gain of a quantum measurement, both without and with quantum side information. We find here that a small information gain implies that it is possible to undo the action of the original measurement (if it is efficient). The result also has operational ramifications for the information-theoretic tasks known as measurement compression without and with quantum side information. We finally provide a robust extension of the main result of [6], establishing that the reduced dynamics of a system-environment interaction are approximately CPTP if and only if the data processing inequality holds approximately.

Acknowledgments

We acknowledge helpful discussions with Mario Berta, Ryan Glasser, Eneet Kaur, Prabha Mandayam, David Sutter, Marco Tomamichel, Dave Touchette, Andreas Winter. SD and MMW acknowledge the NSF under Award No. CCF-1350397 and startup funds from the Department of Physics and Astronomy and the Center for Computation and Technology at Louisiana State University.

---

[1] Robert Alicki. Isotropic quantum spin channels and additivity questions. February 2004. arXiv:quant-ph/0402080.

[2] Robert Alicki and Mark Fannes. Continuity of quantum conditional information. *Journal of Physics A: Mathematical and General*, 37(5):L55–L57, February 2004. arXiv:quant-ph/0312081.

[3] Mario Berta, Matthias Christandl, Roger Colbeck, Joseph M. Renes, and Renato Renner. The uncertainty principle in the presence of quantum memory. *Nature Physics*, 6:659–662, 2010. arXiv:0909.0950.

[4] Mario Berta, Joseph M. Renes, and Mark M. Wilde. Identifying the information gain of a quantum measurement. *IEEE Transactions on Information Theory*, 60(12):7987–8006, December 2014. arXiv:1301.1594.

[5] Mario Berta, Stephanie Wehner, and Mark M. Wilde. Entropic uncertainty and measurement reversibility. November 2015. arXiv:1511.00267.

[6] Francesco Buscemi. Complete positivity, Markovianity, and the quantum data-processing inequality, in the presence of initial system-environment correlations. *Physical Review Letters*, 113(14):140502, October 2014. arXiv:1307.0363.

[7] Francesco Buscemi, Masahito Hayashi, and Michał Horodecki. Global information balance in quantum measurements. *Physical Review Letters*, 100(21):210504, May 2008. arXiv:quant-ph/0702166.

[8] Filippo Caruso, Vittorio Giovannetti, and Alexander S. Holevo. One-mode bosonic Gaussian channels: A full weak-degradability classification. *New Journal of Physics*, 8(12):310, 2006. arXiv:quant-ph/0609013.

[9] Man-Duen Choi. A Schwarz inequality for positive linear maps on C^*-algebras. *Illinois Journal of Mathematics*, 18(4):565–574, December 1974.

[10] Patrick Coles, Mario Berta, Marco Tomamichel, and Stephanie Wehner. Entropic uncertainty relations and their applications. November 2015. arXiv:1511.04857.

[11] Omar Fawzi and Renato Renner. Quantum conditional mutual information and approximate Markov chains. *Communications in Mathematical Physics*, 340(2):575–611, December 2015. arXiv:1410.0664.

[12] Christopher A. Fuchs and Kurt Jacobs. Information-tradeoff relations for finite-strength quantum measurements. *Physical Review A*, 63(6):062305, May 2001. arXiv:quant-ph/0009101.

[13] Raul Garcia-Patron, Carlos Navarrete-Benlloch, Seth Lloyd, Jeffrey H. Shapiro, and Nicolas J. Cerf. Majorization theory approach to the Gaussian channel minimum entropy conjecture. *Physical Review Letters*, 108:110505, March 2012. arXiv:1111.1986.

[14] Hilbrand J. Groenewold. A problem of information gain by quantal measurements. *International Journal of Theoretical Physics*, 4(5):327–338, 1971.

[15] Werner Heisenberg. Über quantentheoretische umdeutung kinematischer und mechanischer beziehungen. *Zeitschrift für Physik*, 33:879–893, 1925.

[16] J. Solomon Ivan, Krishna K. Sabapathy, and Rajiah Simon. Operator-sum representation for bosonic Gaussian channels. *Physical Review A*, 84(4):042311, 2011. arXiv:1012.4266.

[17] Marius Junge, Renato Renner, David Sutter, Mark M. Wilde, and Andreas Winter. Universal recovery from a decrease of quantum relative entropy. September 2015. arXiv:1509.07127.

[18] Oscar E. Lanford and Derek W. Robinson. Mean entropy of states in quantum-statistical mechanics. *Journal of Mathematical Physics*, 9(7):1120–1125, 1968.

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

[20] 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, December 2013. arXiv:1306.3142.

[21] Masanao Ozawa. On information gain by quantum mea-

surements of continuous observables. *Journal of Mathematical Physics*, 27(3):759–763, 1986.

[22] Denes Petz. Sufficient subalgebras and the relative entropy of states of a von Neumann algebra. *Communications in Mathematical Physics*, 105(1):123–131, 1986.

[23] Denes Petz. Sufficiency of channels over von Neumann algebras. *The Quarterly Journal of Mathematics*, 39(1):97–108, 1988.

[24] Takahiro Sagawa. *Lectures on Quantum Computing, Thermodynamics and Statistical Physics*, chapter Second Law-Like Inequalities with Quantum Relative Entropy: An Introduction, page 127. World Scientific, 2012. arXiv:1202.0983.

[25] Benjamin Schumacher and Michael D. Westmoreland. Approximate quantum error correction. *Quantum Information Processing*, 1(1/2):5–12, April 2002. arXiv:quant-ph/0112106.

[26] David Sutter, Marco Tomamichel, and Aram W. Harrow. Strengthened monotonicity of relative entropy via pinched Petz recovery map. July 2015. arXiv:1507.00303.

[27] Dave Touchette. Quantum information complexity and amortized communication. *Proceedings of the Forty-Seventh Annual ACM on Symposium on Theory of Computing*, pages 317–326, June 2015. arXiv:1404.3733.

[28] Armin Uhlmann. The "transition probability" in the state space of a *-algebra. *Reports on Mathematical Physics*, 9(2):273–279, 1976.

[29] Arminn Uhlmann. Relative entropy and the Wigner-Yanase-Dyson-Lieb concavity in an interpolation theory. *Communications in Mathematical Physics*, 54(1):21–32, 1977.

[30] Hisaharu Umegaki. Conditional expectations in an operator algebra IV (entropy and information). *Kodai Mathematical Seminar Reports*, 14(2):59–85, 1962.

[31] Christian Weedbrook, Stefano Pirandola, Raúl García-Patrón, Nicolas J. Cerf, Timothy C. Ralph, Jeffrey H. Shapiro, and Seth Lloyd. Gaussian quantum information. *Reviews of Modern Physics*, 84:621–669, May 2012. arXiv:1110.3234.

[32] Stephanie Wehner, Mark M. Wilde, and Mischa P. Woods. Work and reversibility in quantum thermodynamics. 2015. arXiv:1506.08145.

[33] Alfred Wehrl. General properties of entropy. *Reviews of Modern Physics*, 50(2):221–260, April 1978.

[34] Mark M. Wilde. *From Classical to Quantum Shannon Theory*. 2015. arXiv:1106.1445v6.

[35] Mark M. Wilde. Recoverability in quantum information theory. *Proceedings of the Royal Society A*, 471(2182):20150338, October 2015. arXiv:1505.04661.

[36] Mark M. Wilde, Patrick Hayden, Francesco Buscemi, and Min-Hsiu Hsieh. The information-theoretic costs of simulating quantum measurements. *Journal of Physics A: Mathematical and Theoretical*, 45(45):453001, November 2012. arXiv:1206.4121.

[37] Andreas Winter. "Extrinsic" and "intrinsic" data in quantum measurements: Asymptotic convex decomposition of positive operator valued measures. *Communications in Mathematical Physics*, 244(1):157–185, January 2004. arXiv:quant-ph/0109050.

[38] Andreas Winter. Tight uniform continuity bounds for quantum entropies: conditional entropy, relative entropy distance and energy constraints. July 2015. arXiv:1507.07775.