Quantum Markov Chains and Logarithmic Trace Inequalities

David Sutter*, Mario Berta†, and Marco Tomamichel‡§
*Institute for Theoretical Physics, ETH Zurich, Switzerland
†Institute for Quantum Information and Matter, Caltech, USA
‡School of Physics, The University of Sydney, Australia
§Centre for Quantum Software and Information, University of Technology Sydney, Australia

***Abstract*—A Markov chain is a tripartite quantum state ρ_{ABC} where there exists a recovery map $\mathcal{R}_{B\to BC}$ such that $\rho_{ABC} = \mathcal{R}_{B\to BC}(\rho_{AB})$. More generally, an approximate Markov chain ρ_{ABC} is a state whose distance to the closest recovered state $\mathcal{R}_{B\to BC}(\rho_{AB})$ is small. Recently it has been shown that this distance can be bounded from above by the conditional mutual information $I(A:C|B)_\rho$ of the state. We improve on this connection by deriving the first bound that is tight in the commutative case and features an explicit recovery map that only depends on the reduced state ρ_{BC}. The key tool in our proof is a multivariate extension of the Golden-Thompson inequality, which allows us to extend logarithmic trace inequalities from two to arbitrarily many matrices.**

I. Introduction

A state ρ_{ABC} on a tripartite quantum system $A \otimes B \otimes C$ forms a *(quantum) Markov chain* if it can be recovered from its reduced state ρ_{AB} on $A \otimes B$ by a quantum operation $\mathcal{R}_{B\to BC}$ from B to $B \otimes C$, i.e.,

$$\rho_{ABC} = \mathcal{R}_{B\to BC}(\rho_{AB}) \,. \tag{1}$$

An equivalent characterization of ρ_{ABC} being a Markov chain is that the *conditional mutual information*

$$I(A:C|B)_\rho := H(AB)_\rho + H(BC)_\rho - H(B)_\rho - H(ABC)_\rho \,,$$

is equal to zero [1]. Here $H(A)_\rho := -\operatorname{tr}\rho_A \log \rho_A$ denotes the von Neumann entropy. In addition, for Markov chains ρ_{ABC} the recovery quantum operation can without loss of generality be chosen as

$$\mathcal{T}_{B\to BC}(\cdot) := \rho_{BC}^{\frac{1}{2}}\big(\rho_B^{-\frac{1}{2}}(\cdot)\rho_B^{-\frac{1}{2}} \otimes \mathrm{id}_C\big)\rho_{BC}^{\frac{1}{2}} \,, \tag{2}$$

the *Petz recovery* or *transpose map*. The structure of Markov chains has been studied in various works. In particular, it has been shown that A and C can be viewed as independent conditioned on B, for an operationally useful notion of conditioning [2].

The idea of *approximate quantum Markov chains* became a powerful concept with the recent breakthrough work in [3]. It was shown that for any state ρ_{ABC} there exists a quantum operation $\mathcal{R}_{B\to BC}$ (the recovery map) in the sense that

$$I(A:C|B)_\rho \geq -\log \sup_{\mathcal{R}_{B\to BC}} F\big(\rho_{ABC}, \mathcal{R}_{B\to BC}(\rho_{AB})\big) \,, \tag{3}$$

where the supremum is over trace-preserving completely positive maps from B to $B \otimes C$. Here $F(\omega,\xi) := \big\|\sqrt{\omega}\sqrt{\xi}\big\|_1^2$ denotes the fidelity and the right-hand side of (3) is also known as minus the logarithm of the *fidelity of recovery* [4].

Inequality (3) is of interest for various reasons. First, it strengthens the celebrated *strong subadditivity of quantum entropy* (SSA) which states that $I(A:C|B)_\rho \geq 0$ [5]. (We note that $F(\omega,\xi) \in [0,1]$ for quantum states ω and ξ.) Second, (3) shows that states with small conditional mutual information are approximately recoverable in the sense that there exists a recovery map such that (1) holds approximately. Such states are therefore called approximate quantum Markov chains.

In the case that ρ_{ABC} is a classical state, meaning that ρ_{ABC} and all its reduced states are diagonal in the same basis and therefore commute, the lower bound for the conditional mutual information given in (3) is not tight. To see this, we rewrite the conditional mutual information as

$$I(A:C|B)_\rho = D\big(\rho_{ABC}\,\|\, \exp(\log\rho_{AB} + \log\rho_{BC} - \log\rho_B)\big)$$

in terms of the relative entropy defined as

$$D(\omega\|\xi) := \operatorname{tr}\omega(\log\omega - \log\xi) \text{ if } \omega \ll \xi \,,$$

and $+\infty$ otherwise (where $\omega \ll \xi$ denotes that the support of ω is contained in the support of ξ). This rewriting holds for all states, but now only when ρ_{ABC} is classical we have $\exp(\log\rho_{AB} + \log\rho_{BC} - \log\rho_B) = \mathcal{T}_{B\to BC}(\rho_{AB})$ for the Petz recovery map $\mathcal{T}_{B\to BC}$ from (2) leading to [6]:

$$I(A:C|B)_\rho = D\big(\rho_{ABC}\,\|\,\mathcal{T}_{B\to BC}(\rho_{AB})\big) \,. \tag{4}$$

It is well-known that $D(\omega\|\xi) \geq -\log F(\omega,\xi)$ with strict inequality in the generic case (see, e.g., [7]) and hence we can conclude that (3) cannot be tight in the commutative case. From this line of arguments we can also see the reason why lower bounds on the conditional mutual information as in (3) are difficult to prove: in general matrices do not commute and the matrix exponential cannot just be taken apart.

The inequality (3) has been strengthened in different ways and several alternative proofs have been developed [3], [8], [9], [10], [11], [12], [13]. Here we improve on all these results by showing that

$$I(A:C|B)_\rho \geq D_{\mathbb{M}}\big(\rho_{ABC}\,\|\,\mathcal{R}_{B\to BC}(\rho_{AB})\big) \,, \tag{5}$$

for the rotated Petz recovery map

$$\mathcal{R}_{B\to BC}(\cdot) := \int_{-\infty}^{\infty} \mathrm{d}t\, \beta_0(t)\, \rho_{BC}^{\frac{1+\mathrm{i}t}{2}} \left(\rho_B^{-\frac{1+\mathrm{i}t}{2}}(\cdot)\rho_B^{-\frac{1-\mathrm{i}t}{2}} \otimes \mathrm{id}_C\right) \rho_{BC}^{\frac{1-\mathrm{i}t}{2}} ,$$

with β_0 some fixed probability distribution on $\mathbb{R}$ (independent of the other parameters) defined in (12), and $D_{\mathbb{M}}$ the measured relative entropy defined in (7). We refer to Prop. V.1 for a precise statement and discussion.

Statements like (5) correspond to matrix trace inequalities, or more specifically logarithmic trace inequalities. For our proofs we use a recent multivariate extension of the Golden-Thompson matrix trace inequality [14], [15] to prove (5) as well as other related logarithmic trace inequalities. Our paper is structured as follows. We discuss multivariate trace inequalities in Sec. III and IV, and in Sec. V we show how these can be used to prove entropic inequalities such as (5).

II. Preliminaries

A *quantum state* is a positive semi-definite matrix with trace equal to one.

A. Schatten Norms

Let us define the *Schatten p-norm* of any matrix L as

$$\|L\|_p := \left(\mathrm{tr}|L|^p\right)^{\frac{1}{p}} \quad \text{for} \quad p \geq 1 ,$$

where $|L| := \sqrt{L^\dagger L}$. We will also employ this expression for $0 < p < 1$ although it is no longer a norm then. In the limit $p \to \infty$ we recover the *operator norm* and for $p = 1$ we obtain the *trace norm*. Schatten norms are functions of the singular values and thus unitarily invariant. They satisfy $\|L\|_p = \|L^\dagger\|_p$ and $\|L\|_{2p}^2 = \|LL^\dagger\|_p = \|L^\dagger L\|_p$.

B. Variational Formulas for Relative Entropies

The relative entropy can be expressed as the solution of a convex optimization problem [16].

Lemma II.1. *Let ρ and σ be positive definite matrices such that $\mathrm{tr}\rho = 1$. Then, we have*

$$D(\rho\|\sigma) = \sup_{\omega > 0} \mathrm{tr}\rho \log \omega + 1 - \mathrm{tr}\exp(\log\sigma + \log\omega) . \quad (6)$$

This variational formula is often useful because the dependence on ρ and σ is split up in two separate terms. (To see why (6) is a convex optimization note that for A positive definite $\log A$ becomes Hermitian, and that the set of Hermitian matrices is convex. Furthermore the function $H_1 \mapsto \mathrm{tr}\exp(H_1 + H_2)$ is convex on the set of Hermitian matrices [17].)

The *measured relative entropy* is defined as [18], [19]

$$D_{\mathbb{M}}(\rho\|\sigma) := \sup_{(\mathcal{X},M)} D\left(P_{\rho,M} \middle\| P_{\sigma,M}\right) , \quad (7)$$

where the optimization is over positive operator valued measures (POVMs) M on the power-set of a finite set $\mathcal{X}$, the probability mass functions are given by $P_{\rho,M}(x) = \mathrm{tr}\rho M(x)$, and $D(P\|Q)$ is the *Kullback-Leibler divergence* [20]. The data processing inequality implies that $D_{\mathbb{M}}(\rho\|\sigma) \leq D(\rho\|\sigma)$ and equality holds if and only if ρ and σ commute [21]. It is further known that (see, e.g., [22])

$$D_{\mathbb{M}}(\rho\|\sigma) \geq -\log F(\rho,\sigma) . \quad (8)$$

The measured relative entropy also features a variational formula [21]

Lemma II.2. *Let ρ and σ be positive definite matrices such that $\mathrm{tr}\rho = 1$. Then, we have*

$$D_{\mathbb{M}}(\rho\|\sigma) = \sup_{\omega > 0} \mathrm{tr}\rho \log \omega + 1 - \mathrm{tr}\sigma\omega . \quad (9)$$

III. Multivariate Trace Inequalities

Trace inequalities are mathematical relations between different multivariate trace functionals. Often these relations are straightforward equalities if the involved matrices commute — and are highly non-trivial to prove for the non-commuting case. Arguably one of the best-known trace inequalities is the *Golden-Thompson (GT) inequality* [23], [24]. It states that for any two Hermitian matrices H_1 and H_2 we have

$$\mathrm{tr}\exp(H_1 + H_2) \leq \mathrm{tr}\exp(H_1)\exp(H_2) . \quad (10)$$

We note that in case H_1 and H_2 commute, (10) holds with equality (and only then). We also note that straightforward extensions of this inequality to three matrices are incorrect (see [14] for more explanations). However, recently the GT inequality has been extended to arbitrarily many matrices [14], [15].

Proposition III.1. *Let $p > 0$, $n \in \mathbb{N}$, and consider a collection $\{H_k\}_{k=1}^n$ of Hermitian matrices. Then, we have*

$$\log \left\| \exp\left(\sum_{k=1}^n H_k\right) \right\|_p \leq \int_{-\infty}^{\infty} \mathrm{d}t\, \beta_0(t) \, \log \left\| \prod_{k=1}^n \exp\big((1+\mathrm{i}t)H_k\big) \right\|_p \quad (11)$$

with the probability density

$$\beta_0(t) := \frac{\pi}{2}\big(\cosh(\pi t) + 1\big)^{-1} \text{ on } \mathbb{R} . \quad (12)$$

Note that the expression $\exp((1+\mathrm{i}t)H_k)$ decomposes as $\exp(H_k)\exp(\mathrm{i}tH_k)$, where the latter is a unitary rotation. Since the Schatten p-norm is unitarily invariant, it follows that the integrand in (11) is independent of t for $n = 2$. Inequality (11) thus constitutes an n-matrix extension of the GT inequality and further simplifies to (10) for $n = 2$ and $p = 2$.

IV. Logarithmic Trace Inequalities

The straightforward logarithmic analog of the GT inequality is a relation between $\mathrm{tr}\log A_1 A_2$ and $\mathrm{tr}\log A_1 + \mathrm{tr}\log A_2$ for A_1, A_2 positive definite matrices. As the determinant is multiplicative and since $\mathrm{tr}\log A_1 = \log\det A_1$ we find that

$$\mathrm{tr}\log A_1 + \mathrm{tr}\log A_2 = \mathrm{tr}\log A_2^{\frac{1}{2}} A_1 A_2^{\frac{1}{2}} .$$

This trivially extends to n matrices. Another logarithmic trace inequality states that for $p > 0$,

$$\frac{1}{p}\mathrm{tr}A_1 \log A_1^{\frac{p}{2}} A_2^p A_1^{\frac{p}{2}} \geq \mathrm{tr}\, A_1(\log A_1 + \log A_2) \quad (13)$$

$$\geq \frac{1}{p}\mathrm{tr}A_1 \log A_2^{\frac{p}{2}} A_1^p A_2^{\frac{p}{2}}, \quad (14)$$

with equalities in the limit $p \to 0$ [25], [26]. Prop. III.1 implies the following n-matrix extension of the lower bound in (14).

Proposition IV.1. *Let $q > 0$, β_0 as defined in (12), $n \in \mathbb{N}$, and consider a collection $\{A_k\}_{k=1}^n$ of positive definite matrices. Then, we have*

$$\sum_{k=1}^n \mathrm{tr}\, A_1 \log A_k \geq \int_{-\infty}^{\infty} \mathrm{d}t\, \beta_0(t) \cdot$$
$$\frac{1}{q}\mathrm{tr}\, A_1 \log A_n^{\frac{q(1+\mathrm{i}t)}{2}} \cdots A_3^{\frac{q(1+\mathrm{i}t)}{2}} A_2^{\frac{q}{2}} A_1^q A_2^{\frac{q}{2}} A_3^{\frac{q(1-\mathrm{i}t)}{2}} \cdots A_n^{\frac{q(1-\mathrm{i}t)}{2}},$$

with equality in the limit $q \to 0$.

Proof. First, note that the statement that we aim to show is invariant under multiplication of the matrices $A_1, A_2, \ldots, A_n$ with positive scalars $a_1, a_2, \ldots, a_n > 0$, and hence additional constraints on the norms of the matrices can be introduced without loss of generality.

Let us first show the inequality for $q > 0$, where we suppose that $\mathrm{tr}A_1 = 1$. By definition of the relative entropy we have

$$\sum_{k=1}^n \mathrm{tr}A_1 \log A_k$$
$$= D\Big(A_1 \Big\| \exp\Big(\sum_{k=2}^n \log A_k^{-1}\Big)\Big)$$
$$= \sup_{\omega>0} \mathrm{tr}A_1 \log\omega + 1 - \mathrm{tr}\exp\Big(\log\omega - \sum_{k=2}^n \log A_k\Big), \quad (15)$$

where we used the variational formula for the relative entropy given in Lem. II.1. Now note that the n-matrix extension of the GT inequality (Prop. III.1) can for $pH_k = \log B_k$ and $p = \frac{1}{q}$ be relaxed to

$$\mathrm{tr}\exp\left(\sum_{k=1}^n \log B_k\right) \leq \int_{-\infty}^{\infty} \mathrm{d}t\, \beta_0(t)$$
$$\times\, \mathrm{tr}\left(B_n^{\frac{q}{2}} \cdots B_3^{\frac{q(1+\mathrm{i}t)}{2}} B_2^{\frac{q(1+\mathrm{i}t)}{2}} B_1^q B_2^{\frac{q(1-\mathrm{i}t)}{2}} B_3^{\frac{q(1-\mathrm{i}t)}{2}} \cdots B_n^{\frac{q}{2}}\right)^{\frac{1}{q}}$$

using the concavity of the logarithm and Jensen's inequality. Applying this to (15) we find

$$\sum_{k=1}^n \mathrm{tr}A_1 \log A_k$$
$$\geq \sup_{\omega>0}\left\{\int_{-\infty}^{\infty}\mathrm{d}t\beta_0(t)\mathrm{tr}A_1\log\omega + 1 - \mathrm{tr}\Big(A_2^{-\frac{q}{2}} A_3^{-\frac{q(1+\mathrm{i}t)}{2}} \cdots\right.$$
$$\left. A_n^{-\frac{q(1+\mathrm{i}t)}{2}} \omega^q A_n^{-\frac{q(1-\mathrm{i}t)}{2}} \cdots A_3^{-\frac{q(1-\mathrm{i}t)}{2}} A_2^{-\frac{q}{2}}\Big)^{\frac{1}{q}}\right\}. \quad (16)$$

Now since

$$\omega := \left(A_n^{\frac{q(1+\mathrm{i}t)}{2}} \cdots A_3^{\frac{q(1+\mathrm{i}t)}{2}} A_2^{\frac{q}{2}} A_1^q A_2^{\frac{q}{2}} A_3^{\frac{q(1-\mathrm{i}t)}{2}} \cdots A_n^{\frac{q(1-\mathrm{i}t)}{2}}\right)^{\frac{1}{q}}$$

is a positive definite matrix, we can insert this into (16), which then proves the assertion for $q > 0$.

Next, we show that in the limit $q \to 0$ the inequality in Prop. IV.1 also holds in the opposite direction. For the following we suppose that $A_i \geq 1$ for all $i \in \{1, 2, \ldots, n\}$. We use that $\log X \geq 1 - X^{-1}$ for $X > 0$ and hence

$$\mathrm{tr}\, A_1 \log A_n^{\frac{q(1+\mathrm{i}t)}{2}} \cdots A_2^{\frac{q}{2}} A_1^q A_2^{\frac{q}{2}} \cdots A_n^{\frac{q(1-\mathrm{i}t)}{2}}$$
$$\geq \underbrace{\mathrm{tr}\, A_1\Big(1 - A_n^{\frac{-q(1-\mathrm{i}t)}{2}} \cdots A_2^{-\frac{q}{2}} A_1^{-q} A_2^{-\frac{q}{2}} \cdots A_n^{-\frac{q(1+\mathrm{i}t)}{2}}\Big)}_{=:Z_q(t)}.$$

By assumption on our matrices we have that $A_i^{-1} \leq 1$ for all $i \in \{1, 2, \ldots, n\}$ and thus $Z_q(t) \geq 0$ for all $t \in \mathbb{R}$. By Fatou's lemma, we further find

$$\liminf_{q\to 0} \int_{-\infty}^{\infty} \mathrm{d}t\, \beta_0(t)\, \frac{Z_q(t)}{q} \geq \int_{-\infty}^{\infty} \mathrm{d}t\, \beta_0(t) \liminf_{q\to 0} \frac{Z_q(t)}{q}.$$

Moreover, since $Z_0(t) \equiv 0$ and

$$\frac{\mathrm{d}}{\mathrm{d}q} Z_q(t)\bigg|_{q=0} = \sum_{k=1}^n \mathrm{tr}\, A_1 \log A_k\,,$$

for all $t \in \mathbb{R}$, an application of l'Hôpital's rule yields

$$\liminf_{q\to 0} \frac{Z_q(t)}{q} = \sum_{k=1}^n \mathrm{tr}\, A_1 \log A_k\,.$$

Since $\beta_0(t)$ is normalized this proves the assertion. ☐

We note that there exist other trace inequalities that have been extended to arbitrarily many matrices, such as the Araki-Lieb-Thirring inequality. We refer the reader to [14], [15].

V. Connection to Quantum Markov Chains

As motivated in the introduction, we would like to obtain a lower bound on the conditional mutual information — strengthening (3). Logarithmic trace inequalities such as the one given by Prop. IV.1 promise to give good bounds. (Note that the left-hand side of the inequality in Prop. IV.1 for $n = 4$ constitutes a conditional mutual information.) However, a direct application of Prop. IV.1 is not strong enough and we need to proceed slightly differently.

We first show an inequality that is more general than (3) and covers the general problem of *recoverability of quantum information* (see [11] for an extended discussion). The proof is given in Section VI. We then explain how our bound can be used to characterize (the special case of) approximate quantum Markov chains.

Proposition V.1. *Let ρ, σ be positive semi-definite matrices such that $\rho \ll \sigma$, $\mathrm{tr}\,\rho = 1$, and $\mathcal{N}$ be a trace-preserving completely positive map acting on these matrices. Then, we have*

$$D(\rho\|\sigma) - D\left(\mathcal{N}(\rho)\|\mathcal{N}(\sigma)\right) \geq D_{\mathbb{M}}\left(\rho\|\mathcal{R}_{\sigma,\mathcal{N}} \circ \mathcal{N}(\rho)\right), \quad (17)$$

with the rotated Petz recovery map given by

$$\mathcal{R}_{\sigma,\mathcal{N}}(\cdot) := \int_{-\infty}^{\infty} \mathrm{d}t\, \beta_0(t)\, \mathcal{R}^{[t]}_{\sigma,\mathcal{N}}(\cdot) \quad \text{with}$$
$$\mathcal{R}^{[t]}_{\sigma,\mathcal{N}}(\cdot) := \sigma^{\frac{1+it}{2}} \mathcal{N}^{\dagger}\left(\mathcal{N}(\sigma)^{-\frac{1+it}{2}}(\cdot)\mathcal{N}(\sigma)^{-\frac{1-it}{2}}\right)\sigma^{\frac{1-it}{2}}.$$

We note that the recovery map in $\mathcal{R}_{\sigma,\mathcal{N}}$ is *explicit* and *universal*, i.e., independent of ρ. In addition, it perfectly recovers σ from $\mathcal{N}(\sigma)$, that is

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

Let us explain how this can be used to understand the entropic structure of approximate quantum Markov chains. If we choose $\rho := \rho_{ABC}$, $\sigma := \mathrm{id}_A \otimes \rho_{BC}$, and $\mathcal{N}(\cdot) = \mathrm{tr}_C(\cdot)$ we immediately find a characterization of approximate quantum Markov chains in terms of the conditional mutual information.

Corollary V.2. *Let ρ_{ABC} be a quantum state on $A \otimes B \otimes C$. Then, we have*

$$I(A:C|B)_\rho \geq D_{\mathbb{M}}\big(\rho_{ABC}\big\|\mathcal{R}_{B\to BC}(\rho_{AB})\big), \qquad (18)$$

with the rotated Petz recovery map given by

$$\mathcal{R}_{B\to BC}(\cdot) := \int_{-\infty}^{\infty} \mathrm{d}t\, \beta_0(t)\, \mathcal{R}^{[t]}_{B\to BC}(\cdot) \quad \textit{with}$$
$$\mathcal{R}^{[t]}_{B\to BC}(\cdot) := \rho_{BC}^{\frac{1+it}{2}}\left(\rho_B^{-\frac{1+it}{2}}(\cdot)\rho_B^{-\frac{1-it}{2}} \otimes \mathrm{id}_C\right)\rho_{BC}^{\frac{1-it}{2}}.$$

The recovery map $\mathcal{R}_{B\to BC}$ is explicit, universal (i.e., it only depends on the reduced state ρ_{BC}), and satisfies $\mathcal{R}_{B\to BC}(\rho_B) = \rho_{BC}$. We note that (18) together with (8) imply (3) which we set out to improve. In addition we know from (4) that (18) is tight in the commutative case, since the measured relative entropy then coincides with the relative entropy and the recovery map $\mathcal{R}_{B\to BC}$ simplifies to the Petz recovery map $\mathcal{T}_{B\to BC}$ from (2).

Cor. V.2 is the first lower bound on the conditional mutual information in terms of recoverability that is tight in the commutative case and has a recovery map that is explicit and universal. We note that Prop. V.1 and Cor. V.2 are no longer valid if we replace the measured relative entropy in (17) with the relative entropy. This leads us to believe that (17) and (18) cannot be further improved.

VI. Proof of Prop. V.1

We first prove Prop. V.1 for the special case where $\mathcal{N}$ is a partial trace (see Lem. VI.1). This can then be lifted to the full general statement.

Lemma VI.1. *Let ρ_{AB} and σ_{AB} be positive semi-definite matrices on $A \otimes B$ such that $\rho_{AB} \ll \sigma_{AB}$ and $\mathrm{tr}\,\rho_{AB} = 1$. Then, we have*

$$D(\rho_{AB}\|\sigma_{AB}) - D(\rho_A\|\sigma_A) \geq D_{\mathbb{M}}\big(\rho_{AB}\|\mathcal{R}_{\sigma_{AB},\mathrm{tr}_B}(\rho_A)\big)$$

with the rotated Petz recovery map given by

$$\mathcal{R}_{\sigma_{AB},\mathrm{tr}_B}(\cdot) := \int_{-\infty}^{\infty} \mathrm{d}t\, \beta_0(t)\, \mathcal{R}^{[t]}_{\sigma_{AB},\mathrm{tr}_B}(\cdot) \quad \text{and}$$
$$\mathcal{R}^{[t]}_{\sigma_{AB},\mathrm{tr}_B}(\cdot) := \sigma_{AB}^{\frac{1+it}{2}}\left(\sigma_A^{-\frac{1+it}{2}}(\cdot)\sigma_A^{-\frac{1-it}{2}} \otimes \mathrm{id}_B\right)\sigma_{AB}^{\frac{1-it}{2}}.$$

Proof. Let us recall Prop. III.1 applied for $n = 4$ and $p = 2$. Using the concavity of the logarithm and Jensen's inequality, this yields

$$\begin{aligned}&\mathrm{tr}\exp(H_1 + H_2 + H_3 + H_4)\\ &\leq \int_{-\infty}^{\infty} \mathrm{d}t\beta_0(t)\,\mathrm{tr}\exp(H_1)\exp\left(\frac{1+it}{2}H_2\right)\exp\left(\frac{1+it}{2}H_3\right)\\ &\qquad \times \exp(H_4)\exp\left(\frac{1-it}{2}H_3\right)\exp\left(\frac{1-it}{2}H_2\right) \quad (19)\end{aligned}$$

for Hermitian matrices $\{H_k\}_{k=1}^4$. Moreover, by definition of the relative entropy for positive definite operators ρ_{AB} and σ_{AB}, we have

$$\begin{aligned}&D(\rho_{AB}\|\sigma_{AB}) - D(\rho_A\|\sigma_A)\\ &= D\big(\rho_{AB}\|\exp(\log\sigma_{AB} + \log\rho_A \otimes \mathrm{id}_B - \log\sigma_A \otimes \mathrm{id}_B)\big)\,. \quad (20)\end{aligned}$$

For positive semi-definite operators ρ_{AB} and σ_{AB}, the Hermitian operators $\log\sigma_{AB}$, $\log\rho_A$ and $\log\sigma_A$ are well-defined under the convention $\log 0 = 0$. Under this convention, the above (20) also holds for positive semi-definite operators as long as $\rho_{AB} \ll \sigma_{AB}$, which is required by the proposition. By the variational formula for the relative entropy (Lem. II.1) we thus find

$$\begin{aligned}&D(\rho_{AB}\|\sigma_{AB}) - D(\rho_A\|\sigma_A)\\ &= \sup_{\omega_{AB}>0}\ \mathrm{tr}\,\rho_{AB}\log\omega_{AB} + 1\\ &\quad - \mathrm{tr}\exp(\log\sigma_{AB} + \log\rho_A \otimes \mathrm{id}_B - \log\sigma_A \otimes \mathrm{id}_B + \log\omega_{AB})\\ &\geq \sup_{\omega_{AB}>0}\ \mathrm{tr}\,\rho_{AB}\log\omega_{AB} + 1\\ &\quad - \int_{-\infty}^{\infty}\mathrm{d}t\,\beta_0(t)\,\mathrm{tr}\,\sigma_{AB}^{\frac{1+it}{2}}\left(\sigma_A^{-\frac{1+it}{2}}\rho_A\sigma_A^{-\frac{1-it}{2}} \otimes \mathrm{id}_B\right)\sigma_{AB}^{\frac{1-it}{2}}\,\omega_{AB}\\ &= D_{\mathbb{M}}\left(\rho_{AB}\middle\|\int_{-\infty}^{\infty}\mathrm{d}t\beta_0(t)\,\sigma_{AB}^{\frac{1+it}{2}}\left(\sigma_A^{-\frac{1+it}{2}}\rho_A\sigma_A^{-\frac{1-it}{2}} \otimes \mathrm{id}_B\right)\sigma_{AB}^{\frac{1-it}{2}}\right)\end{aligned}$$

where the single inequality follows by the four matrix extension of the GT inequality from (19). The final step uses the variational formula from (9) for the measured relative entropy. □

Lem. VI.1 now readily implies our result about the connection of quantum Markov chains to the conditional mutual information (Prop. V.1) by means of Stinespring dilation.

Proof of Prop. V.1. Let us introduce the Stinespring dilation of $\mathcal{N}$, denoted U, and the states $\rho_{AB} = U\rho U^{\dagger}$, $\sigma_{AB} = U\sigma U^{\dagger}$ such that $\mathcal{N}(\rho) = \rho_A$ and $\mathcal{N}(\sigma) = \sigma_B$. Then, using the fact that the relative entropy is invariant under isometries, we have

$$\begin{aligned}D(\rho\|\sigma) - D\left(\mathcal{N}(\rho)\|\mathcal{N}(\sigma)\right) &= D(\rho_{AB}\|\sigma_{AB}) - D(\rho_A\|\sigma_A)\\ &\geq D_{\mathbb{M}}\big(\rho_{AB}\|\mathcal{R}_{\sigma_{AB},\mathrm{tr}_B}(\rho_A)\big)\\ &= D_{\mathbb{M}}\left(\rho\|\mathcal{R}_{\sigma,\mathcal{N}} \circ \mathcal{N}(\rho)\right),\end{aligned}$$

where the inequality is due to Lem. VI.1, and the last equality uses again invariance under isometries and the fact that

$$U^\dagger \mathcal{R}^{[t]}_{\sigma_{AB},\mathrm{tr}_B}(\cdot)U$$
$$= U^\dagger U \sigma^{\frac{1+it}{2}} U^\dagger \Big(\mathcal{N}(\sigma)^{-\frac{1+it}{2}}(\cdot)\mathcal{N}(\sigma)^{-\frac{1-it}{2}} \otimes \mathrm{id}_B\Big) U \sigma^{\frac{1-it}{2}} U^\dagger U$$
$$= \sigma^{\frac{1+it}{2}} \mathcal{N}^\dagger \left(\mathcal{N}(\sigma)^{-\frac{1+it}{2}}(\cdot)\mathcal{N}(\sigma)^{-\frac{1-it}{2}}\right) \sigma^{\frac{1-it}{2}} = \mathcal{R}^{[t]}_{\sigma,\mathcal{N}}(\cdot)\,.$$

□

VII. Conclusion

As shown in the introduction, if a state ρ_{ABC} is classical it is straightforward to see that $I(A:C|B)_\rho = D\big(\rho_{ABC}\|\mathcal{T}_{B\to BC}(\rho_{AB})\big)$ with the Petz recovery map $\mathcal{T}_{B\to BC}$ from (2). Our corresponding inequality for the general, non-commutative case is

$$I(A:C|B)_\rho \geq D_{\mathbb{M}}\big(\rho_{ABC}\|\mathcal{R}_{B\to BC}(\rho_{AB})\big) \tag{21}$$

with the rotated Petz recovery map $\mathcal{R}_{B\to BC}$ as given in Cor. V.2. This result is of particular interest as it strengthens the celebrated SSA of quantum entropy which has been vastly useful in quantum information theory (see, e.g., [27]).

Moreover, it is natural to ask if there is also a relative entropy distance type upper bound for the conditional mutual information in terms of Markovianity for the non-commutative case. To put it differently, (21) states that whenever the conditional mutual information is small there exists a quantum operation acting on the B-system only that recovers ρ_{ABC} well out of ρ_{AB}. The desired opposite statement would be that whenever the conditional mutual information is large there is no recovery map acting on B that recovers ρ_{ABC} well from ρ_{AB}. (See [28], [14], [12] for partial results.) A very related question that remains open is if the logarithmic trace inequality upper bound from (13) can be extended from two to arbitrarily many matrices.

Acknowledgements: DS acknowledges support by the Swiss National Science Foundation (SNSF) via the National Centre of Competence in Research "QSIT" and by the European Commission via the project "RAQUEL". MB acknowledges funding by the SNSF through a fellowship, funding by the Institute for Quantum Information and Matter (IQIM), an NSF Physics Frontiers Center (NFS Grant PHY-1125565) with support of the Gordon and Betty Moore Foundation (GBMF-12500028), and funding support form the ARO grant for Research on Quantum Algorithms at the IQIM (W911NF-12-1-0521). MT is funded by an ARC Discovery Early Career Researcher Award (DECRA) fellowship and acknowledges support from the ARC Centre of Excellence for Engineered Quantum Systems (EQUS).

References

[1] D. Petz, "Sufficient subalgebras and the relative entropy of states of a von Neumann algebra," *Communications in Mathematical Physics*, vol. 105, no. 1, pp. 123–131, 1986.

[2] P. Hayden, R. Jozsa, D. Petz, and A. Winter, "Structure of states which satisfy strong subadditivity of quantum entropy with equality," *Communications in Mathematical Physics*, vol. 246, no. 2, pp. 359–374, 2004.

[3] O. Fawzi and R. Renner, "Quantum conditional mutual information and approximate Markov chains," *Communications in Mathematical Physics*, vol. 340, no. 2, pp. 575–611, 2015.

[4] K. P. Seshadreesan and M. M. Wilde, "Fidelity of recovery, squashed entanglement, and measurement recoverability," *Physical Review A*, vol. 92, p. 042321, 2015.

[5] E. H. Lieb and M. B. Ruskai, "Proof of the strong subadditivity of quantum-mechanical entropy," *Journal of Mathematical Physics*, vol. 14, no. 12, pp. 1938–1941, 1973.

[6] B. Ibinson, N. Linden, and A. Winter, "Robustness of quantum markov chains," *Communications in Mathematical Physics*, vol. 277, no. 2, pp. 289–304, 2008.

[7] M. Müller-Lennert, F. Dupuis, O. Szehr, S. Fehr, and M. Tomamichel, "On quantum Rényi entropies: A new generalization and some properties," *Journal of Mathematical Physics*, vol. 54, no. 12, 2013.

[8] F. G. S. L. Brandão, A. W. Harrow, J. Oppenheim, and S. Strelchuk, "Quantum conditional mutual information, reconstructed states, and state redistribution," *Physical Review Letters*, vol. 115, no. 5, p. 050501, 2015.

[9] M. Berta and M. Tomamichel, "The fidelity of recovery is multiplicative," *IEEE Transactions on Information Theory*, vol. 62, no. 4, pp. 1758–1763, 2016.

[10] D. Sutter, O. Fawzi, and R. Renner, "Universal recovery map for approximate Markov chains," *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 472, no. 2186, 2016.

[11] M. M. Wilde, "Recoverability in quantum information theory," *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 471, no. 2182, p. 20150338, 2015.

[12] D. Sutter, M. Tomamichel, and A. W. Harrow, "Strengthened monotonicity of relative entropy via pinched Petz recovery map," *IEEE Transactions on Information Theory*, vol. 62, no. 5, pp. 2907–2913, 2016.

[13] M. Junge, R. Renner, D. Sutter, M. M. Wilde, and A. Winter, "Universal recovery from a decrease of quantum relative entropy," 2015, arXiv:1509.07127.

[14] D. Sutter, M. Berta, and M. Tomamichel, "Multivariate trace inequalities," *Communications in Mathematical Physics*, vol. 352, no. 1, pp. 37–58, 2017. [Online]. Available: http://dx.doi.org/10.1007/s00220-016-2778-5

[15] F. Hiai, R. König, and M. Tomamichel, "Generalized log-majorization and multivariate trace inequalities," *Annales Henri Poincaré*, pp. 1–23, 2017. [Online]. Available: http://dx.doi.org/10.1007/s00023-017-0569-y

[16] D. Petz, "A variational expression for the relative entropy," *Communications in Mathematical Physics*, vol. 114, no. 2, pp. 345–349, 1988.

[17] E. Carlen, *Trace Inequalities and Quantum Entropy: An Introductory Course*. Contemporary Mathematics, 2009.

[18] M. J. Donald, "On the relative entropy," *Communications in Mathematical Physics*, vol. 105, no. 1, pp. 13–34, 1986.

[19] F. Hiai and D. Petz, "The proper formula for relative entropy and its asymptotics in quantum probability," *Communications in Mathematical Physics*, vol. 143, no. 1, pp. 99–114, 1991.

[20] S. Kullback and R. A. Leibler, "On information and sufficiency," *Annals of Mathematical Statistics*, vol. 22, no. 1, pp. 79–86, 1951.

[21] M. Berta, O. Fawzi, and M. Tomamichel, "On variational expressions for quantum relative entropies," 2015, arXiv:1512.02615.

[22] C. A. Fuchs, "Distinguishability and accessible information in quantum theory," *PhD Thesis, University of New Mexico*, 1996, arXiv:quant-ph/9601020.

[23] S. Golden, "Lower bounds for the Helmholtz function," *Physical Review*, vol. 137, pp. B1127–B1128, 1965.

[24] C. J. Thompson, "Inequality with applications in statistical mechanics," *Journal of Mathematical Physics*, vol. 6, no. 11, pp. 1812–1813, 1965.

[25] F. Hiai and D. Petz, "The Golden-Thompson trace inequality is complemented," *Linear Algebra and its Applications*, vol. 181, pp. 153 – 185, 1993.

[26] T. Ando and F. Hiai, "Log majorization and complementary Golden-Thompson type inequalities," *Linear Algebra and its Applications*, vol. 197, pp. 113 – 131, 1994.

[27] M. M. Wilde, *Quantum Information Theory*. Cambridge University Press, June 2013.

[28] F. Dupuis and M. M. Wilde, "Swiveled Rényi entropies," *Quantum Information Processing*, vol. 15, no. 3, pp. 1309–1345, 2016.