IEEE Transactions on Information Theory

Continuity of entropies via integral representations

SCHOLARONE™
Manuscripts

Continuity of entropies via integral representations

Mario Berta,[1] Ludovico Lami,[2,3,4] and Marco Tomamichel[5,6]

[1]*Institute for Quantum Information, RWTH Aachen University, Germany*

[2]*QuSoft, Science Park 123, 1098 XG Amsterdam, The Netherlands*

[3]*Korteweg–de Vries Institute for Mathematics, University of Amsterdam,
Science Park 105-107, 1098 XG Amsterdam, The Netherlands*

[4]*Institute for Theoretical Physics, University of Amsterdam,
Science Park 904, 1098 XH Amsterdam, The Netherlands*

[5]*Center for Quantum Technologies, National University of Singapore*

[6]*Department of Electrical and Computer Engineering,
College of Design and Engineering, National University of Singapore*

We show that Frenkel's integral representation of the quantum relative entropy provides a natural framework to derive continuity bounds for quantum information measures. Our main general result is a dimension-independent semi-continuity relation for the quantum relative entropy with respect to the first argument. Using it, we obtain a number of results: (1) a tight continuity relation for the conditional entropy in the case where the two states have equal marginals on the conditioning system, resolving a conjecture by Wilde in this special case; (2) a stronger version of the Fannes–Audenaert inequality on quantum entropy; (3) better estimates on the quantum capacity of approximately degradable channels; (4) an improved continuity relation for the entanglement cost; (5) general upper bounds on asymptotic transformation rates in infinite-dimensional entanglement theory; and (6) a proof of a conjecture due to Christandl, Ferrara, and Lancien on the continuity of 'filtered' relative entropy distances.

CONTENTS

2

I. INTRODUCTION

In a series of recent works, a new integral formula for Umegaki's quantum relative entropy was discovered. Originally found by Frenkel [1] and later refined by Jenčova [2] and Hirche and Tomamichel [3], it can be written as

$$D(\rho\|\sigma) = (\log e) \int_1^\infty \mathrm{d}\gamma \left(\frac{1}{\gamma} E_\gamma(\rho\|\sigma) + \frac{1}{\gamma^2} E_\gamma(\sigma\|\rho) \right). \tag{1}$$

Here, Umegaki's quantum *relative entropy* [4] of a quantum state ρ with regards to another quantum state σ is given as

$$D(\rho\|\sigma) := \mathrm{Tr}\, \rho \left(\log \rho - \log \sigma \right), \tag{2}$$

assuming that the support ρ is contained in the support σ as otherwise the quantity is not finite. Moreover, for any given $\gamma \geq 1$, the corresponding *hockey-stick divergence* is defined by

$$E_\gamma(\rho\|\sigma) := \mathrm{Tr}\left(\rho - \gamma\sigma \right)_+, \tag{3}$$

where $X_+ := \sum_i \max\{x_i, 0\} |i\rangle\langle i|$ is the positive part of a Hermitian operator X with spectral decomposition $X = \sum_i x_i |i\rangle\langle i|$.

The quantum relative entropy can be seen as a parent quantity for quantum entropies, as other information measures such as the von Neumann entropy, the conditional entropy, and the mutual information can all be expressed in terms of it. The operational importance of the relative entropy, in turn, stems from the quantum Stein's lemma due to Hiai and Petz [5], which endows it with an operational interpretation in the context of quantum hypothesis testing. The above expression of the relative entropy in terms of hockey-stick divergences is particularly fruitful because the latter quantities, thanks to their operational significance in binary hypothesis testing, exhibit many useful mathematical properties and are thus comparatively easy to manipulate.

Integral representations of quantum relative entropies have already been employed to derive continuity bounds (see [6] for an early example and [3, Section 5.2] for the most recent instance). One of our main goals is to derive fully quantum extensions of the continuity inequality for the Shannon entropy. Its tight version, due to Csiszár, states that for two probability distributions P_X, Q_X on a finite alphabet $\mathcal{X}$ with $\frac{1}{2}\|P_X - Q_X\|_1 \leq \varepsilon \leq 1 - \frac{1}{|\mathcal{X}|}$ in variational distance, one has that

$$\left| H(X)_P - H(X)_Q \right| \leq \varepsilon \log(|\mathcal{X}| - 1) + h_2(\varepsilon), \tag{4}$$

where $H(X)_P := -\sum_{x\in\mathcal{X}} P(x) \log P(x)$ denotes the Shannon entropy, and

$$h_2(p) := p \log \frac{1}{p} + (1 - p) \log \frac{1}{1 - p} \tag{5}$$

is the *binary entropy function*. The above inequality admits a quantum generalisation due to Fannes [7], Audenaert [8], and Petz [9, Theorem 3.8]:

$$\left| S(\rho) - S(\sigma) \right| \leq \varepsilon \log(d - 1) + h_2(\varepsilon), \tag{6}$$

where $S(\rho) := -\mathrm{Tr}\, \rho \log \rho$, called the *von Neumann entropy*, is the extension of the Shannon entropy to density operators ρ on Hilbert spaces of finite dimension d, and ε, which is assumed to satisfy $\frac{1}{2}\|\rho - \sigma\|_1 \leq \varepsilon \leq 1 - \frac{1}{d}$, is a bound on the *trace distance* between the density operators ρ and σ, where $\|\rho - \sigma\|_1 := \mathrm{Tr}\,|\rho - \sigma|$. Yet another generalisation of (4) is to conditional entropies: for any

two probability distributions P_{XY}, Q_{XY} on a product alphabet $\mathcal{X} \times \mathcal{Y}$, if $\frac{1}{2}\|P_{XY} - Q_{XY}\|_1 \leq \varepsilon \leq 1 - \frac{1}{|\mathcal{X}|}$ then Alhejji and Smith proved that [10]

$$\left| H(X|Y)_P - H(X|Y)_Q \right| \leq \varepsilon \log \left(|\mathcal{X}| - 1 \right) + h_2(\varepsilon) , \tag{7}$$

where $H(X|Y) \coloneqq H(XY) - H(Y)$ is the conditional entropy. The above inequality was then extended by Wilde to the quantum-classical case where X, the conditioned system, is promoted to a quantum system [11]. Wilde also conjectured that a similar continuity relation should hold also for the fully quantum case [11, Eq. (58)], namely that

$$\left| H(A|B)_\rho - H(A|B)_\sigma \right| \overset{?}{\leq} \varepsilon \log \left(|A|^2 - 1 \right) + h_2(\varepsilon) , \tag{8}$$

where $H(A|B)_\rho \coloneqq S(\rho_{AB}) - S(\rho_B)$ is the conditional entropy, $|A|$ is the dimension of the quantum system A, and $\frac{1}{2}\left\|\rho_{AB} - \sigma_{AB}\right\|_1 \leq \varepsilon \leq 1 - \frac{1}{|A|^2}$ is an estimate on the trace distance between the two states ρ_{AB} and σ_{AB}.

Our intuition to unify all these different bounds and move towards a proof of (8) is to lift the problem from quantum entropies to the corresponding parent quantity, the relative entropy. Our main general result is Theorem 1 below, which presents a semi-continuity bound of the form

$$D(\rho\|\omega) - D(\sigma\|\omega) \leq \varepsilon \log(M - 1) + h_2(\varepsilon) , \tag{9}$$

where $\frac{1}{2}\|\rho - \sigma\|_1 \leq \varepsilon \leq 1 - \frac{1}{M}$, and M is any number such that the operator inequality $\rho \leq M\omega$, meaning that $M\omega - \rho$ is positive semi-definite, holds. Note that the left-hand side of (9) does not contain an absolute value. This feature, which makes ours a *semi*-continuity bound rather than a plain continuity bound like (4), is rather fundamental, as the expression on the left-hand side can happen to diverge to $-\infty$. Eq. (9) is in fact tight, in the sense that for all $M \geq 1$ and $\varepsilon \in \left[0, 1 - \frac{1}{M}\right]$ one can find a triple of states ρ, σ, ω obeying the above conditions and saturating (9). Our proof of (9) uses the integral representation in (1) together with the properties of hockey-stick divergences. The conceptual novelty of our approach, therefore, is that it works directly at the level of quantum entropies, without first operating a reduction to the classical case, like all previous approaches to prove the Fannes–Audenaert inequality [7–9].

The formal resemblance between (4) and (6), on one side, and (9), on the other, is clear. Indeed, as we will show, the latter inequality constitutes a strict generalisation of the former two. In fact, in Section III A we show how to use our (9) to obtain an improved version of the Fannes–Audenaert inequality (6) (Corollary 3).

As an immediate consequence of (9), in Section III B we also prove Wilde's conjecture (8) in the special case where ρ_{AB} and σ_{AB} have equal marginals on the conditioning system, in formula $\rho_B = \sigma_B$. Our inequality takes the form

$$\left| H(A|B)_\rho - H(A|B)_\sigma \right| \leq \varepsilon \log \left(|A| \min\{|A|, |B|\} - 1 \right) + h_2(\varepsilon) \qquad (\rho_B = \sigma_B) , \tag{10}$$

where $\frac{1}{2}\left\|\rho_{AB} - \sigma_{AB}\right\|_1 \leq \varepsilon \leq 1 - (|A| \min\{|A|, |B|\})^{-1}$; it generalises the Fannes–Audenaert inequality (6), and it directly implies (8) in the case of equal marginals. Note that the case where $|B| \leq |A|$ is less interesting, as it follows immediately from (6) applied to the global system AB. The case where the dimension of B is larger than that of A, and perhaps even infinite, is, on the contrary, novel.

Remarkably, in proving (10) we cannot afford to restrict ourselves to the classical case, as one typically does to prove (6), because the conditional entropy, unlike the von Neumann entropy,

is not invariant under (global) unitaries. It is here that our fully quantum methods grant us a decisive advantage.

While our solution of Wilde's conjecture covers only the case of equal marginals, it is precisely this case that turns out to have a plethora of fruitful applications throughout quantum information theory. Using (10), we are able to derive continuity bounds on several fundamental quantities that are tighter than anything found in the prior literature. First, in Section III C we tackle the entanglement cost, a key entanglement measure that quantifies how much entanglement is needed to prepare a given quantum state in the asymptotic limit. Then, in Section III D, we employ once again (10) to derive the tightest available bounds on the magnitude of additivity violations in the quantum capacity of approximately degradable channels [12]. We continue by applying our general result (9) in a somewhat indirect way to resolve a conjecture by Christandl, Ferrara, and Lancien [13, Conjecture 7] on a special form of continuity of certain relative-entropy-based resource quantifiers. Finally, in Section III F we show how to employ the dimension-independent inequality (9) to constrain transformation rates in infinite-dimensional entanglement theory, by-passing known technical hurdles related to the failure of asymptotic continuity in infinite-dimensional systems, a phenomenon referred to as the 'asymptotic continuity catastrophe' in [14].

II. A SEMI-CONTINUITY RELATION FOR THE RELATIVE ENTROPY

We now present our main technical result, which employs the integral formula in (1) to derive a new semi-continuity relation for the relative entropy with respect to the first argument, when the second one is fixed.

Theorem 1. *Let ρ, σ, ω be quantum states on the same separable Hilbert space such that*

$$\frac{1}{2} \left\| \rho - \sigma \right\|_1 \leq \varepsilon, \qquad \exp\left[D_{\max}(\rho \| \omega) \right] := \inf\{m : \rho \leq m\omega\} \leq M \tag{11}$$

for some $\varepsilon \in [0, 1]$ and $M \geq 1$. Then, we have that

$$D(\rho\|\omega) - D(\sigma\|\omega) \leq \begin{cases} \varepsilon \log(M-1) + h_2(\varepsilon) & \text{if } \varepsilon < 1 - \frac{1}{M}, \\ \log M & \text{if } \varepsilon \geq 1 - \frac{1}{M}. \end{cases} \tag{12}$$

Moreover, for every such ε and M, there exists a triple of states ρ, σ and ω obeying (11) *for which the inequality* (12) *is tight.*

Remark 2. It is possible to give a slightly simplified bound that is looser than the right-hand side of (12) but does not involve a function defined piece-wise. Namely, it holds that

$$D(\rho\|\omega) - D(\sigma\|\omega) \leq \varepsilon \log M + h_2(\varepsilon), \tag{13}$$

simply because the right-hand side of the above equation can be shown to be never smaller than that of (12).

Proof of Theorem 1. If $\varepsilon \geq 1 - 1/M$ there is nothing to prove, as it is clear that

$$D(\rho\|\omega) - D(\sigma\|\omega) \leq D(\rho\|\omega) \leq D_{\max}(\rho\|\omega) \leq \log M; \tag{14}$$

therefore, let us assume that $\varepsilon < 1 - 1/M$ and thus also $M > 1$. We start by writing

$$\frac{1}{\log e} \left(D(\rho\|\omega) - D(\sigma\|\omega) \right) = \int_1^\infty \frac{\mathrm{d}\gamma}{\gamma} \left(E_\gamma(\rho\|\omega) - E_\gamma(\sigma\|\omega) \right) + \int_1^\infty \frac{\mathrm{d}\gamma}{\gamma^2} \left(E_\gamma(\omega\|\rho) - E_\gamma(\omega\|\sigma) \right).$$
(15)

We will now bound the two above integrals separately. First, it holds that

$$\begin{aligned}
&\int_1^\infty \frac{\mathrm{d}\gamma}{\gamma} \left(E_\gamma(\rho\|\omega) - E_\gamma(\sigma\|\omega) \right) \\
&\stackrel{(i)}{=} \int_1^{(1-\varepsilon)M} \frac{\mathrm{d}\gamma}{\gamma} \left(E_\gamma(\rho\|\omega) - E_\gamma(\sigma\|\omega) \right) + \int_{(1-\varepsilon)M}^M \frac{\mathrm{d}\gamma}{\gamma} \left(E_\gamma(\rho\|\omega) - E_\gamma(\sigma\|\omega) \right) \\
&\quad + \int_M^\infty \frac{\mathrm{d}\gamma}{\gamma} \left(E_\gamma(\rho\|\omega) - E_\gamma(\sigma\|\omega) \right) \\
&\stackrel{(ii)}{\leq} \int_1^{(1-\varepsilon)M} \frac{\mathrm{d}\gamma}{\gamma}\, \varepsilon + \int_{(1-\varepsilon)M}^M \frac{\mathrm{d}\gamma}{\gamma} \left(1 - \frac{\gamma}{M} \right) \\
&= \varepsilon \ln \left((1-\varepsilon)M \right) + \ln \frac{1}{1-\varepsilon} - \varepsilon \\
&= -(1-\varepsilon)\ln(1-\varepsilon) + \varepsilon \ln M - \varepsilon .
\end{aligned}$$
(16)

Here, in (i) we split the integral into three parts, noting that $(1-\varepsilon)M \geq 1$. The justification of (ii) is a bit more involved, as we used three different upper bounds for the integrand in each of the three regions: namely, for $1 \leq \gamma \leq (1-\varepsilon)M$ we have that

$$E_\gamma(\rho\|\omega) = \operatorname{Tr}(\rho - \gamma\omega)_+ \leq \operatorname{Tr}(\rho - \sigma)_+ + \operatorname{Tr}(\sigma - \gamma\omega)_+ \leq \varepsilon + E_\gamma(\sigma\|\omega); \tag{17}$$

in the third region, i.e. for $\gamma \geq M$, it holds that

$$E_\gamma(\rho\|\omega) - E_\gamma(\sigma\|\omega) \leq E_\gamma(\rho\|\omega) \leq E_M(\rho\|\omega) = \operatorname{Tr}(\rho - M\omega)_+ = 0, \tag{18}$$

where the last equality follows because $\rho \leq M\omega$ by definition; finally, in the second region, i.e. for $(1-\varepsilon)M \leq \gamma \leq M$, we write

$$\gamma = pM + (1-p)(1-\varepsilon)M, \qquad p := \frac{\gamma - (1-\varepsilon)M}{\varepsilon M}, \tag{19}$$

and then

$$\begin{aligned}
E_\gamma(\rho\|\omega) &\leq p\, E_M(\rho\|\omega) + (1-p)\, E_{(1-\varepsilon)M}(\rho\|\omega) \\
&= (1-p)\, E_{(1-\varepsilon)M}(\rho\|\omega) \\
&= (1-p)\operatorname{Tr}(\rho - (1-\varepsilon)M\omega)_+ \\
&\leq (1-p)\operatorname{Tr}(\rho - (1-\varepsilon)\rho)_+ \\
&= (1-p)\,\varepsilon \\
&= 1 - \frac{\gamma}{M},
\end{aligned}$$
(20)

where the first inequality is by convexity of $\gamma \mapsto E_\gamma(\rho\|\omega)$, in the second line we observed again that $E_M(\rho\|\omega) = 0$, and the inequality in the fourth line holds due to the operator inequality $M\omega \geq \rho$, together with the fact that $\operatorname{Tr} X_+$ is monotonically non-decreasing in X with respect to the positive semi-definite (Löwner) ordering. This completes the justification of (16).

We now move on to the second integral in (15). It holds that

$$
\begin{aligned}
\int_1^\infty & \frac{\mathrm{d}\gamma}{\gamma^2} \left(E_\gamma(\omega\|\rho) - E_\gamma(\omega\|\sigma) \right) \\
&\overset{\text{(iii)}}{=} \int_1^{\frac{M-1}{M\varepsilon}} \frac{\mathrm{d}\gamma}{\gamma^2} \left(E_\gamma(\omega\|\rho) - E_\gamma(\omega\|\sigma) \right) + \int_{\frac{M-1}{M\varepsilon}}^\infty \frac{\mathrm{d}\gamma}{\gamma^2} \left(E_\gamma(\omega\|\rho) - E_\gamma(\omega\|\sigma) \right) \\
&\overset{\text{(iv)}}{\leq} \int_1^{\frac{M-1}{M\varepsilon}} \frac{\mathrm{d}\gamma}{\gamma^2}\, \gamma\varepsilon + \int_{\frac{M-1}{M\varepsilon}}^\infty \frac{\mathrm{d}\gamma}{\gamma^2} \left(1 - \frac{1}{M} \right) \\
&= \varepsilon \ln \frac{M-1}{M\varepsilon} + \left(1 - \frac{1}{M} \right) \frac{M\varepsilon}{M-1} \\
&= \varepsilon \ln(M-1) - \varepsilon \ln M - \varepsilon \ln \varepsilon + \varepsilon \,.
\end{aligned}
\tag{21}
$$

Here: in (iii) we split the integral, noting that $\frac{M-1}{M\varepsilon} \geq 1$ by assumption; in (iv) we observed that

$$
E_\gamma(\omega\|\rho) = \mathrm{Tr}\left(\omega - \gamma\rho \right)_+ \leq \mathrm{Tr}\left(\omega - \gamma\sigma \right)_+ + \gamma\,\mathrm{Tr}(\sigma - \rho)_+ \leq E_\gamma(\omega\|\sigma) + \gamma\varepsilon \,,
\tag{22}
$$

which gives the upper bound on the first term, and moreover

$$
E_\gamma(\omega\|\rho) - E_\gamma(\omega\|\sigma) \leq E_\gamma(\omega\|\rho) \leq E_{1/M}(\omega\|\rho) = \mathrm{Tr}\left(\omega - \frac{1}{M}\rho \right)_+ = 1 - \frac{1}{M} \,,
\tag{23}
$$

because rather obviously

$$
\gamma \geq \frac{M-1}{M\varepsilon} \geq 1 \geq \frac{1}{M} \,,
\tag{24}
$$

which gives the upper bound on the second term. This completes the justification of (21). Putting (16) and (21) together finally proves the claim.

To show optimality, consider (commuting) qubit states

$$
\rho = |0\rangle\langle0| \,, \quad \sigma = (1-\varepsilon)\,|0\rangle\langle0| + \varepsilon\,|1\rangle\langle1| \,, \quad \text{and} \quad \omega = \frac{1}{M}\,|0\rangle\langle0| + \left(1 - \frac{1}{M} \right)|1\rangle\langle1| \,,
\tag{25}
$$

which satisfy $\frac{1}{2}\|\rho - \sigma\|_1 = \varepsilon$ and $D_{\max}(\rho\|\omega) = \log M$, as well as

$$
D(\rho\|\omega) - D(\sigma\|\omega) = \varepsilon \log(M-1) + h_2(\varepsilon) =: f(M, \varepsilon)
\tag{26}
$$

We note that f is monotone in M but achieves a maximum of $\log M$ at $\varepsilon = 1 - \frac{1}{M}$ as a function of ε. If $\varepsilon \leq 1 - \frac{1}{M}$, the inequality in (12) is clearly saturated. If $\varepsilon > 1 - \frac{1}{M}$, instead, we can consider the above three states for $\varepsilon \mapsto 1 - \frac{1}{M}$. Since we only require that $\frac{1}{2}\|\rho - \sigma\|_1 \leq \varepsilon$, this once again proves the claim. $\qquad\square$

III. APPLICATIONS

A. Improved Fannes–Audenaert inequality

Theorem 1 evaluated for $\omega = \pi := \mathbb{1}/d$ immediately implies a slightly strengthened version of the Fannes–Audenaert inequality in (6). We recall that this inequality states that given two

states ρ, σ at trace distance $\varepsilon := \frac{1}{2} \|\rho - \sigma\|_1$, their entropies obey the inequality $|S(\rho) - S(\sigma)| \leq \varepsilon \log(d - 1) + h_2(\varepsilon)$, where d is the dimension of the underlying Hilbert space. The original, looser form of this continuity relation was found by Fannes [7], while the tight version was established by Audenaert [8] and Petz [9, Theorem 3.8]. Our bound yields an improvement for mixed states.

Corollary 3. *Let ρ, σ be any two states at trace distance $\frac{1}{2}\|\rho - \sigma\|_1 \leq \varepsilon \leq 1 - \frac{1}{d\,\lambda_{\max}(\sigma)}$ on a d-dimensional quantum system. Here, $\lambda_{\max}(\sigma)$ denotes the maximal eigenvalue of σ. Then, we have that*

$$S(\rho) - S(\sigma) \leq \varepsilon \log\left(d\,\lambda_{\max}(\sigma) - 1\right) + h_2(\varepsilon). \tag{27}$$

Remark 4. The restriction on the range of ε in the statement of Corollary 3 is motivated by considerations that are similar to those found at the end of the proof of Theorem 1. Namely, for $\varepsilon = 1 - \frac{1}{d\,\lambda_{\max}(\sigma)}$ the right-hand side of (27) achieves its maximum value, equal to $\log\left(d\lambda_{\max}(\sigma)\right)$. That the left-hand side of (27) must be upper bounded by this number is however easy to show, so for larger values of ε the statement of Corollary 3 would be somewhat empty.

Another approach to avoid restricting the range of ε is to define ε as the exact trace distance $\frac{1}{2}\|\rho - \sigma\|_1$, instead of taking it to be a generic upper bound on it. While this is the approach followed by some authors, it is not what is most useful in applications, where one often has an estimate of the trace distance but not its exact value. Similar considerations will apply to the restrictions on the range of ε in, e.g., the forthcoming Theorem 5 and Proposition 8.

We stress that our proof, in contrast to previous arguments, does not go via a reduction to the commutative case (e.g. via dephasing) but instead the integral representation method applies directly in the quantum setting.

B. Improved continuity of conditional entropy

Our Theorem 1 can be applied to the question about the tight continuity of the quantum conditional entropy independent of the dimension of the conditioning system [15–17]. Our tightest bound features the mixed state Schmidt number [18]

$$\mathrm{SN}(\rho_{AB}) := \inf_{\rho = \sum_i p_i \psi_i} \max_i \mathrm{SN}\left(|\psi_i\rangle_{AB}\right), \tag{28}$$

where the minimisation is over pure state convex decompositions $\rho_{AB} = \sum_i p_i |\psi_i\rangle\langle\psi_i|_{AB}$, and $\mathrm{SN}\left(|\psi_i\rangle_{AB}\right)$ denotes the regular pure state Schmidt number, i.e. the rank of the reduced state $\mathrm{Tr}_B |\psi_i\rangle\langle\psi_i|_{AB}$.

Theorem 5. *Let AB be a bipartite quantum system where A has finite dimension $|A|$. Let $\rho = \rho_{AB}$ and $\sigma = \sigma_{AB}$ be two states at trace distance $\frac{1}{2}\|\rho - \sigma\|_1 \leq \varepsilon$ and having the same marginal on B, i.e. such that*

$$\rho_B = \sigma_B. \tag{29}$$

Then, assuming $\varepsilon \leq 1 - \frac{1}{|A|\,\mathrm{SN}(\rho_{AB})}$, it holds that

$$H(A|B)_\rho - H(A|B)_\sigma \leq \varepsilon \log\left(|A| \cdot \mathrm{SN}(\rho_{AB}) - 1\right) + h_2(\varepsilon). \tag{30}$$

It then immediately follows that

$$\left|H(A|B)_\rho - H(A|B)_\sigma\right| \leq \varepsilon \log\left(|A| \min\{|A|, |B|\} - 1\right) + h_2(\varepsilon) \tag{31}$$

for quantum states ρ_{AB}, σ_{AB} with $\rho_B = \sigma_B$ and $\varepsilon \leq 1 - \frac{1}{|A|\min\{|A|,|B|\}}$, whereas for separable states the bound is improved to

$$\left|H(A|B)_\rho - H(A|B)_\sigma\right| \leq \varepsilon \log\left(|A| - 1\right) + h_2(\varepsilon) \tag{32}$$

for $\varepsilon \leq 1 - \frac{1}{|A|}$. This pleasantly reproduces the tight classical (7) from [10] and the corresponding quantum-classical bound [11] (that both equally hold even when $\rho_B \neq \sigma_B$).

Proof of Theorem 5. Set

$$\omega = \omega_{AB} = \pi_A \otimes \rho_B = \pi_A \otimes \sigma_B. \tag{33}$$

For any positive integer $k \in \mathbb{N}^+$, the map $X \mapsto k\mathbb{1}\operatorname{Tr} X - X$ is known to be k-positive (see e.g. [19, Eq. (3.11) and discussion thereafter]). By taking $k = \operatorname{SN}(\rho_{AB})$ and applying the map to ρ_{AB}, we deduce the operator inequality

$$\rho_{AB} \leq \operatorname{SN}(\rho_{AB})\,\mathbb{1}_A \otimes \rho_B = |A| \cdot \operatorname{SN}(\rho_{AB})\,\pi_A \otimes \rho_B = |A| \cdot \operatorname{SN}(\rho_{AB})\,\omega_{AB}, \tag{34}$$

which immediately implies that

$$D_{\max}\left(\rho_{AB}\,\middle\|\,\omega_{AB}\right) \leq \log\left(|A| \cdot \operatorname{SN}(\rho_{AB})\right). \tag{35}$$

We thus obtain that

$$\begin{aligned}
H(A|B)_\sigma - H(A|B)_\rho &= D\left(\rho_{AB}\,\middle\|\,\pi_A \otimes \rho_B\right) - D\left(\sigma_{AB}\,\middle\|\,\pi_A \otimes \sigma_B\right) \\
&= D(\rho_{AB}\|\omega_{AB}) - D(\sigma_{AB}\|\omega_{AB}) \\
&\leq \varepsilon \log\left(|A| \cdot \operatorname{SN}(\rho_{AB}) - 1\right) + h_2(\varepsilon),
\end{aligned} \tag{36}$$

where in the last step we applied Theorem 1. □

As mentioned in the introduction, it was then also conjectured by Wilde [11] that (31) holds generally for finite-dimensinal states:

Conjecture 6 [11]. *For two finite-dimensional quantum states ρ_{AB}, σ_{AB} with trace distance $\frac{1}{2}\|\rho_{AB} - \sigma_{AB}\|_1 \leq \varepsilon \leq 1 - \frac{1}{|A|^2}$, do we have that*

$$\left|H(A|B)_\rho - H(A|B)_\sigma\right| \stackrel{?}{\leq} \varepsilon \log\left(|A|^2 - 1\right) + h_2(\varepsilon). \tag{37}$$

This would indeed be stronger than the state-of-the-art Alicki–Fannes–Winter bound [16, 20–22], which states for $\frac{1}{2}\|\rho_{AB} - \sigma_{AB}\|_1 \leq \varepsilon \leq 1$ that

$$\left|H(A|B)_\rho - H(A|B)_\sigma\right| \leq \varepsilon \log|A|^2 + (1 + \varepsilon) \cdot h_2\left(\frac{\varepsilon}{1 + \varepsilon}\right). \tag{38}$$

While we fail to prove Conjecture 6 in full generality, a multitude of comments around our Theorem 5 are noteworthy:

- Eq. (31) is exactly tight for $|A| = |B|$ in all dimensions, as choosing $\rho_{AB} = \Phi_{AB}$ maximally entangled and $\sigma_{AB} = (1 - \varepsilon)\Phi_{AB} + \frac{\varepsilon}{|A|^2-1}\left(\mathbb{1}_{AB} - \Phi_{AB}\right)$ gives equality. This is in contrast to the not exactly tight (38).

- For the general case, our Theorem 1 implies that

$$H(A|B)_\sigma - H(A|B)_\rho \le \varepsilon \log\left(|A| \min\{|A|, |B|\} - 1\right) + h_2(\varepsilon) + D(\sigma_B \| \rho_B). \tag{39}$$

- Employing the Fannes–Audenaert inequality on AB for the special case $\rho_B = \sigma_B$ yields

$$\left|H(A|B)_\rho - H(A|B)_\sigma\right| \le \varepsilon \log(|A||B| - 1) + h_2(\varepsilon). \tag{40}$$

- Following [23, Proposition 3.5.3], one resolved quantum case is when $H(A|B)_\rho \ge H(A|B)_\sigma$ and σ_{AB} is diagonal in a maximally entangled basis (which also implies that $|A| = |B|$). This, however, then also directly follows from Fannes–Audenaert inequality — as remarked in [23, Section 3.5].

- In the classical case, one does have the tight bound (7), which is not directly implied by Theorem 1. The classical proof from [10] employs sophisticated pre-processing of the input distributions P_{XY} and Q_{XY}, making use of the conditional majorisation monotonicity of the conditional entropy. This then reduces everything to the product case with a uniform Y-marginal. The proof method also extends to A quantum, by means of conditional dephasing [11].

- For the general quantum case, one might then try similar pre-processing ideas as well, making use of the monotonicity of the quantum conditional entropy under A-unital operations [24], also known as quantum conditional majorisation [25, 26] — which is also how the above example of σ_{AB} diagonal in a maximally entangled basis works.

- Another approach is to employ divergence centers for the choice of ω in Theorem 1. For the quantum relative entropy it is defined as [27]

$$\omega^*(\rho, \sigma) := \operatorname{argmin} \max\left\{D(\rho\|\omega), D(\sigma\|\omega)\right\}, \tag{41}$$

for which $D(\sigma_B \| \omega_B^*) - D(\rho_B \| \omega_B^*) = 0$. However, it is not a priori clear how to tightly upper bound $D_{\max}(\rho_{AB} \| \pi_A \otimes \omega_B^*)$ for that choice. It is known that [27, 28]

$$\omega^*(\rho, \sigma) = \lambda \cdot \rho + (1 - \lambda)\sigma \text{ for } \lambda \in [0, 1] \tag{42}$$

and one might choose some general λ with $\omega_B^\lambda(\rho_B, \sigma_B)$. Bounding

$$D_{\max}\left(\rho_{AB} \,\|\, \pi_A \otimes \omega_B^\lambda\right) \le \log\left(\lambda^{-1} d^2\right) \tag{43}$$

and then further optimising the choice of $\lambda \in [0, 1]$ against the terms

$$D\left(\sigma_B \,\|\, \omega_B^\lambda\right) - D\left(\rho_B \,\|\, \omega_B^\lambda\right) \tag{44}$$

also leads to some bounds. In fact, this approach leads to the conjectured bound as long as $H_{\min}(A|B)_\rho \ge -\log\left(\lambda|A|\right)$, solidifying the intuition that the most difficult case to handle is when ρ_{AB} is highly entangled.

- The original proofs [16, 17] on the continuity of the quantum conditional entropy instead work with the max-relative entropy divergence center [21]

$$\omega_{\max}^*(\rho, \sigma) := \operatorname{argmin} \max\left\{D_{\max}(\rho\|\omega), D_{\max}(\sigma\|\omega)\right\}, \tag{45}$$

where the ***max-relative entropy*** between two states τ and ξ on the same separable Hilbert space is defined by [29]

$$D_{\max}(\tau\|\xi) := \inf\left\{\lambda \,:\, \tau \le 2^\lambda \xi\right\}. \tag{46}$$

The above state $\omega^*_{\max}(\rho, \sigma)$ takes the explicit form

$$\omega^*_{\max}(\rho, \sigma) = \frac{\rho + 2[\rho - \sigma]_+}{1 + \frac{1}{2}\|\rho - \sigma\|_1} = \frac{\sigma + 2[\rho - \sigma]_-}{1 + \frac{1}{2}\|\rho - \sigma\|_1}. \tag{47}$$

This choice also leads to bounds in our formalism, which, however, do not seem to be exactly tight either.

We remark that Theorem 5 can be generalised with a suitable application of the chain rule, showing that only the marginals that differ contribute to the continuity bound. Namely, for ABC a tripartite quantum system with A of finite dimension $|A|$ and $\rho = \rho_{ABC}$ and $\sigma = \sigma_{ABC}$ two states at trace distance $\frac{1}{2}\|\rho - \sigma\|_1 \le \varepsilon \le 1 - \frac{1}{|A|^2}$ and satisfying $\rho_{BC} = \sigma_{BC}$, we have that

$$\left|H(AB|C)_\rho - H(AB|C)_\sigma\right| \le \varepsilon \log\left(|A|^2 - 1\right) + h_2(\varepsilon). \tag{48}$$

This follows from $H(AB|C) = H(A|BC) + H(B|C)$ together with the fact that the latter entropy has the same value on ρ_{BC} and σ_{BC}.

Further, this then also implies bounds for entropic quantum channel expressions (cf. [30, 31]). Those often involve so-called regularisations [32], which make the derivation of the corresponding continuity statements slightly more involved. As an example, we state an improved continuity bound for the output entropy of n-fold tensor product channels in terms of the diamond norm, which, for quantum channels $\mathcal{M}_{A\to B}$ and $\mathcal{N}_{A\to B}$, is defined as

$$\left\|\mathcal{M}_{A\to B} - \mathcal{N}_{A\to B}\right\|_\diamond := \sup_{|\Psi\rangle_{AA'}} \left\|\left(\left(\mathcal{M}_{A\to B} - \mathcal{N}_{A\to B}\right) \otimes I_{A'}\right)\left(|\Psi\rangle\!\langle\Psi|_{AA'}\right)\right\|_1, \tag{49}$$

where the supremum is over all ancillary systems A' and all pure states $|\Psi\rangle_{AA'}$ on A'.

Lemma 7. *Let $\mathcal{M}_{A\to B}$ and $\mathcal{N}_{A\to B}$ be quantum channels, with B finite dimensional and $\frac{1}{2}\|\mathcal{M}_{A\to B} - \mathcal{N}_{A\to B}\|_\diamond \le \varepsilon \le 1 - \frac{1}{|B|^2}$. Then, for all reference systems A' and states $\rho_{A^n A'}$ on $A^{\otimes n}A'$ we have that*

$$\left|S\left(\left(\mathcal{M}^{\otimes n}_{A\to B} \otimes I_{A'}\right)(\rho_{A^n A'})\right) - S\left(\left(\mathcal{N}^{\otimes n}_{A\to B} \otimes I_{A'}\right)(\rho_{A^n A'})\right)\right| \le n \cdot \left(\varepsilon \log\left(|B|^2 - 1\right) + h_2(\varepsilon)\right). \tag{50}$$

This improves on the previous bound [30, Theorem 11]

$$\left|S\left(\left(\mathcal{M}^{\otimes n}_{A\to B} \otimes I_{A'}\right)(\rho_{A^n A'})\right) - S\left(\left(\mathcal{N}^{\otimes n}_{A\to B} \otimes I_{A'}\right)(\rho_{A^n A'})\right)\right| \le n \cdot \left(4\varepsilon \log|B| + 2h_2(\varepsilon)\right). \tag{51}$$

Proof of Lemma 7. Following the telescoping argument from [30, Theorem 11] and employing our improved continuity of the quantum conditional entropy for states with equal marginals from Theorem 5, we immediately deduce the claim. $\square$

C. Improved continuity of entanglement cost

The *entanglement cost* [33] is the fundamental entanglement measure that quantifies the amount of pure entanglement that is needed to prepare a given mixed state in the asymptotic limit [34–37] using only local operations and classical communication (LOCC) [38]. It represents the maximal asymptotic entanglement measure [39–41], and it can be computed for any state $\rho = \rho_{AB}$ of a bipartite quantum system AB as

$$E_c(\rho) = \lim_{n \to \infty} \frac{1}{n} E_f\left(\rho^{\otimes n}\right), \tag{52}$$

$$E_f(\rho) := \inf_{\{p_x, \psi_x\}_x : \sum_x p_x \psi_x^{AB} = \rho_{AB}} \sum_x p_x S\left(\psi_x^A\right). \tag{53}$$

Throughout this section, we are going to assume that at least one between A and B is finite dimensional. The measure in (53) is called the *entanglement of formation*. The infimum in its definition runs over all decompositions of ρ as a convex combination of pure states $\psi_x^{AB} = |\psi_x\rangle\langle\psi_x|_{AB}$. For a precise operational definition of the entanglement cost, we refer the reader to [34] or to the explanation below (96). However, for the sake of this discussion we could also take directly (52) as a definition.

A continuity relation for the entanglement of formation in (53) was proved already in 2000 by Nielsen [42] and then later refined in [11] based on the tight continuity bounds for the classical conditional entropy from [10]. The analysis of the continuity properties of the entanglement cost (52) proved significantly more complicated, mainly due to the presence of the regularisation, i.e. the limit $n \to \infty$ over the number of copies, in its definition. The first continuity relation was established in 2015 by Winter [17, Corollary 4], and proving it required considerable technical progress [17, Proposition 5], which requires conditioning on quantum systems. Here, we show how to use our new Theorem 1 to provide an improvement over Winter's result.

Proposition 8. *Let AB be a bipartite quantum system, with either A or B finite dimensional. Let ρ_{AB} and σ_{AB} be two states on AB whose trace distance satisfies*

$$\frac{1}{2}\left\|\rho_{AB} - \sigma_{AB}\right\|_1 \leq \varepsilon \leq 1 - \frac{\sqrt{2d^2 - 1}}{d^2}. \tag{54}$$

Then, setting $\delta := \sqrt{\varepsilon(2 - \varepsilon)}$ and $d := \min\{|A|, |B|\} < \infty$, we have that

$$\left|E_c(\rho_{AB}) - E_c(\sigma_{AB})\right| \leq \delta \log\left(d^2 - 1\right) + h_2(\delta). \tag{55}$$

This improves on the previous bound

$$\left|E_c(\rho_{AB}) - E_c(\sigma_{AB})\right| \leq \delta \log\left(d^2\right) + (1 + \delta) \cdot h_2\left(\frac{\delta}{1 + \delta}\right) \tag{56}$$

from [17, Corollary 4].

Proof of Proposition 8. The key observation is that Theorem 5 suffices to improve the continuity estimate in the original proof of the (asymptotic) continuity of the entanglement cost, given by Winter [17, Corollary 4]. With Winter's notation, we can explain the idea very concisely as follows. The entire proof of [17, Corollary 4] goes through, but in the last line of the last equation we apply our Theorem 5 instead of [17, Lemma 2]. This is possible because $\widetilde{\rho}$ and $\widetilde{\sigma}$ have the same marginals on $A'X$ (in fact, even on $A'B'X$), as Winter observes in the equation before. Also, note that because of the restriction on the range of ε we have that $\delta \leq 1 - \frac{1}{d^2}$. $\qquad\square$

Remark 9. The technical reason why our fully quantum continuity relation is needed in the proof of continuity of the entanglement cost, but not in that of the entanglement of formation, is that in the aforementioned last equation of the proof of [17, Corollary 4] the system A' is un-measured and thus quantum. This does not happen in the proof of continuity for the entanglement of formation, as in that case the conditioning system is fully measured (cf. the first part of the proof of [17, Corollary 4]).

D.   Approximate degradability

Let $\mathcal{N}_{A \to B}$ be a quantum channel with input system A and output system B. Let $V_{A \to BE}$ be one of its Stinespring dilations, so that $\mathcal{N}_{A \to B}(\rho_A) = \mathrm{Tr}_E \, V_{A \to BE} \rho_A V_{A \to BE}^\dagger$ for all states ρ_A on A. We denote with $\mathcal{N}_{A \to E}^c$ the associated complementary channel, that acts as $\mathcal{N}_{A \to E}^c(\rho_A) = \mathrm{Tr}_B \, V_{A \to BE} \rho_A V_{A \to BE}^\dagger$. For some $\varepsilon \in [0, 1]$, we say that a channel $\Theta_{B \to \tilde{E}}$, where $\tilde{E} \simeq E$ is a copy of the quantum system E, is an ε-**degrading channel** for $\mathcal{N}_{A \to B}$ if it holds that [12]

$$\frac{1}{2} \left\| \mathcal{N}_{A \to E}^c - \iota_{\tilde{E} \to E} \circ \Theta_{B \to \tilde{E}} \circ \mathcal{N}_{A \to B} \right\|_\diamond \le \varepsilon , \tag{57}$$

where $\iota_{\tilde{E} \to E}$ is the embedding of $\tilde{E} \simeq E$ into E, and the diamond norm is given by (49).

The notion of approximate degradability was introduced to study quantum capacities [12, 30, 43]. To compute these latter quantities, a useful notion is a suitably modified version of the coherent information. We remind the reader that the coherent information of a channel $\mathcal{N} = \mathcal{N}_{A \to B}$ with complementary channel $\mathcal{N}^c = \mathcal{N}_{A \to E}^c$ is given by

$$I_c(\mathcal{N}) := \sup_{\rho = \rho_A} \left\{ S(\mathcal{N}(\rho)) - S(\mathcal{N}^c(\rho)) \right\} , \tag{58}$$

and due to the Lloyd–Shor–Devetak theorem [44–47], the quantum capacity can be written as

$$Q(\mathcal{N}) = \lim_{n \to \infty} \frac{1}{n} I_c(\mathcal{N}^{\otimes n}) = \sup_{n \in \mathbb{N}^+} \frac{1}{n} I_c(\mathcal{N}^{\otimes n}) . \tag{59}$$

For degradable (i.e. '0-approximate degradable') channels, the above regularisation is unnecessary, because the coherent information turns out to be additive; one thus obtains the single-letter formula $Q(\mathcal{N}_{A \to B}) = I_c(\mathcal{N}_{A \to B})$. In the case of approximately degradable channels, the coherent information in general ceases to be additive and can be strictly super-additive, entailing that the regularisation in (59) cannot be removed.

To obtain an additive quantity, one needs to consider a suitably modified version of the coherent information. Let $\Theta_{B \to \tilde{E}}$ be an ε-approximate degrading map for $\mathcal{N}_{A \to B}$. Then we can imagine to modify (58) by replacing the complementary channel $\mathcal{N}_{A \to E}^c$ on the right-hand side by its approximate version $\Theta_{B \to \tilde{E}} \circ \mathcal{N}_{A \to B}$ (we can safely ignore the isometric embedding $\iota_{\tilde{E} \to E}$ if we only compute entropies). By doing so, one obtains the quantity [12, Eq. (5)]

$$\begin{aligned} U_\Theta(\mathcal{N}) &:= \sup_{\rho = \rho_A} \left\{ S(\mathcal{N}(\rho)) - S((\Theta \circ \mathcal{N})(\rho)) \right\} \\ &= \sup_{\rho_A} H(F | \tilde{E}) , \end{aligned} \tag{60}$$

where F is an environment for a Stinespring dilation $W_{B \to \tilde{E} F}$ of $\Theta_{B \to F}$. We can compare directly (60) and (58) by re-writing the latter as

$$I_c(\mathcal{N}) = \sup_{\rho_A} \left\{ H(\tilde{E} F) - H(E) \right\} . \tag{61}$$

Remarkably, the quantity in (60) is additive, meaning that [12, Lemma 3.3]

$$U_{\Theta_1 \otimes \Theta_2}(\mathcal{N}_1 \otimes \mathcal{N}_2) = U_{\Theta_1}(\mathcal{N}_1) + U_{\Theta_2}(\mathcal{N}_2) \tag{62}$$

holds for all choices of channels $\mathcal{N}_1, \mathcal{N}_2$ and Θ_1, Θ_2.

We are now ready to state the main result of this section.

Proposition 10. *Let $\mathcal{N}$ be a quantum channel, and let Θ be an ε-degrading channel for $\mathcal{N}$. Provided that $\varepsilon \leq 1 - \frac{1}{|E|^2}$, where $|E|$ is the dimension of the environment of a Stinespring dilation of $\mathcal{N}$, the quantum capacity $Q(\mathcal{N})$ of $\mathcal{N}$ can be bounded as*

$$I_c(\mathcal{N}) \leq Q(\mathcal{N}) \leq U_{\Theta}(\mathcal{N}) + \varepsilon \log\left(|E|^2 - 1\right) + h_2(\varepsilon), \tag{63}$$

where $U_{\Theta}(\mathcal{N})$ is defined by (60). Furthermore, if $\varepsilon \leq 1 - \frac{1}{|E|}$, then it also holds that

$$I_c(\mathcal{N}) \leq Q(\mathcal{N}) \leq I_c(\mathcal{N}) + \varepsilon \log\left[(|E| - 1)^2(|E| + 1)\right] + 2h_2(\varepsilon), \tag{64}$$

This improves on the previous bound in [12, Theorem 3.4], which featured worse fudge terms.

Proof of Proposition 10. We follow the main ideas employed in the proof of [12, Theorem 3.4, (i) and (ii)], starting with (63). The first inequality is trivial, and follows by direct inspection of (59). As for the second one, pick a positive integer $n \in \mathbb{N}^+$, and write as in [12, proof of Theorem 3.4(ii)]

$$
\begin{aligned}
H\left(\tilde{E}^n F^n\right) - H(E^n) &= H\left(F^n | \tilde{E}^n\right) + H\left(\tilde{E}^n\right) - H(E^n) \\
&\overset{(i)}{=} H\left(F^n | \tilde{E}^n\right) + \sum_{t=1}^{n}\left(H\left(\tilde{E}_{\leq t} E_{>t}\right) - H\left(\tilde{E}_{<t} E_{\geq t}\right)\right) \\
&\overset{(ii)}{=} H\left(F^n | \tilde{E}^n\right) + \sum_{t=1}^{n}\left(H\left(\tilde{E}_t | \tilde{E}_{<t} E_{>t}\right) - H\left(E_t | \tilde{E}_{<t} E_{>t}\right)\right).
\end{aligned}
\tag{65}
$$

Here, we used the shorthand notation $H\left(\tilde{E}_{\leq t} E_{>t}\right) := H\left(\tilde{E}_1 \ldots \tilde{E}_t E_{t+1} \ldots E_n\right)$, and similarly $H\left(\tilde{E}_{<t} E_{\geq t}\right) := H\left(\tilde{E}_1 \ldots \tilde{E}_{t-1} E_t \ldots E_n\right)$. We used a telescopic identity in (i) and the chain rule in (ii). To continue, we observe that the state of the system $\tilde{E}_{\leq t} E_{>t}$ is of the form

$$\sigma^{(n,t)} = \sigma^{(n,t)}_{\tilde{E}_{\leq t} E_{>t}} := \left((\Theta \circ \mathcal{N})^{\otimes t} \otimes (\mathcal{N}^c)^{\otimes(n-t)}\right)(\omega^n), \tag{66}$$

where $\omega^n = \omega^n_{A^n}$ is the initial state. The state of the system $\tilde{E}_{<t} E_{\geq t}$ can instead be written as

$$\rho^{(n,t)} = \rho^{(n,t)}_{\tilde{E}_{<t} E_{\geq t}} := \left((\Theta \circ \mathcal{N})^{\otimes(t-1)} \otimes (\mathcal{N}^c)^{\otimes(n-t+1)}\right)(\omega^n). \tag{67}$$

Identifying $\tilde{E} \simeq E$, we have that

$$\frac{1}{2}\left\|\rho^{(n,t)} - \sigma^{(n,t)}\right\|_1 \leq \frac{1}{2}\left\|\Theta \circ \mathcal{N} - \mathcal{N}^c\right\|_{\diamond} \leq \varepsilon \leq 1 - \frac{1}{|E|^2}, \tag{68}$$

where $\tilde{E}_{<t} E_{>t}$ plays the role of the ancillary system A' in (49). Note that

$$\rho^{(n,t)}_{\tilde{E}_{<t} E_{>t}} = \mathrm{Tr}_{E_t}\, \rho^{(n,t)}_{\tilde{E}_{<t} E_{\geq t}} = \left((\Theta \circ \mathcal{N})^{\otimes(t-1)} \otimes (\mathcal{N}^c)^{\otimes(n-t)}\right)(\omega^n) = \mathrm{Tr}_{\tilde{E}_t}\, \sigma^{(n,t)}_{\tilde{E}_{\leq t} E_{>t}} = \sigma^{(n,t)}_{\tilde{E}_{<t} E_{>t}}, \tag{69}$$

i.e. the marginals of $\rho^{(n,t)}$ and $\sigma^{(n,t)}$ on $\tilde{E}_{<t}E_{>t}$ coincide. This implies, via Theorem 5, that

$$
\begin{aligned}
H\left(\tilde{E}_t \middle| \tilde{E}_{<t}E_{>t}\right) - H\left(E_t \middle| \tilde{E}_{<t}E_{>t}\right) &= H\left(\tilde{E}_t \middle| \tilde{E}_{<t}E_{>t}\right)_{\sigma^{(n,t)}} - H\left(E_t \middle| \tilde{E}_{<t}E_{>t}\right)_{\rho^{(n,t)}} \\
&\leq \varepsilon \log\left(|E|^2 - 1\right) + h_2(\varepsilon).
\end{aligned}
\tag{70}
$$

Now, plugging (70) into (65), optimising over the initial state ω^n, and dividing by n yields the inequality

$$
\begin{aligned}
\frac{1}{n} I_c\left(\mathcal{N}^{\otimes n}\right) &\leq \frac{1}{n} U_{\Theta^{\otimes n}}\left(\mathcal{N}^{\otimes n}\right) + \varepsilon \log\left(|E|^2 - 1\right) + h_2(\varepsilon) \\
&\overset{\text{(iii)}}{=} U_\Theta(\mathcal{N}) + \varepsilon \log\left(|E|^2 - 1\right) + h_2(\varepsilon),
\end{aligned}
\tag{71}
$$

where (iii) follows from the additivity of U_Θ in (62). Taking the limit $n \to \infty$ in (71) yields (63), according to (59).

As for (64), it is obtained precisely as in [12, Theorem 3.4(i)], by combining the above statement with the continuity relation that links $I_c(\mathcal{N})$ and $U_\Theta(\mathcal{N})$. This takes the form [12, Proposition 3.2]

$$
|I_c(\mathcal{N}) - U_\Theta(\mathcal{N})| \leq \varepsilon \log(|E| - 1) + h_2(\varepsilon),
\tag{72}
$$

provided that $\varepsilon \leq 1 - \frac{1}{|E|}$. $\qquad\square$

Remark 11. It is possible to obtain tighter versions of the relations (63) and (64) by using the full power of our main inequality (12). To this end, one needs to introduce two channel divergences, the ***un-stabilised*** and the ***stabilised max-relative entropy***, respectively given by

$$
\overline{D}_{\max}(\Lambda_1 \| \Lambda_2) := \sup_\rho D_{\max}\left(\Lambda_1(\rho) \middle\| \Lambda_2(\rho)\right),
\tag{73}
$$

$$
D_{\max}(\Lambda_1 \| \Lambda_2) := \sup_\rho D_{\max}\left((\Lambda_1 \otimes I)(\rho) \middle\| (\Lambda_2 \otimes I)(\rho)\right)
\tag{74}
$$

for any two quantum channels Λ_1, Λ_2 with the same input and output systems. Here, the max-relative entropy is defined by (46), and the optimisation in (73) is over all states on the input system, while in (74) we included also an ancillary system (a procedure known as 'stabilisation'). Importantly, due to [48, Lemma 12] the stabilised max-relative entropy can be computed via the semi-definite program

$$
D_{\max}(\Lambda_1 \| \Lambda_2) = \min\{\gamma : \ 2^\gamma \Lambda_2 - \Lambda_1 \in \mathrm{CP}\},
\tag{75}
$$

where CP denotes the set of completely positive maps.

With this notation, we can state the following refinements of (63) and (64): *provided that $\varepsilon \leq 1 - 2^{-D_{\max}(\Theta \circ \mathcal{N} \| \pi \operatorname{Tr})}$, where $(\pi \operatorname{Tr})(X) := \pi_E \operatorname{Tr} X$ is the completely depolarising channel, we have that*

$$
I_c(\mathcal{N}) \leq Q(\mathcal{N}) \leq U_\Theta(\mathcal{N}) + \varepsilon \log\left(2^{D_{\max}(\Theta \circ \mathcal{N} \| \pi \operatorname{Tr})} - 1\right) + h_2(\varepsilon).
\tag{76}
$$

Moreover, *if also the condition $\varepsilon \leq 1 - 1/M$ is satisfied for some $M \geq 2^{\overline{D}_{\max}(\Theta \circ \mathcal{N} \| \pi \operatorname{Tr})}$ (for example, one could set $M = |E|$), then*

$$
I_c(\mathcal{N}) \leq Q(\mathcal{N}) \leq I_c(\mathcal{N}) + \varepsilon \log\left[(M-1)\left(2^{D_{\max}(\Theta \circ \mathcal{N} \| \pi \operatorname{Tr})} - 1\right)\right] + 2h_2(\varepsilon).
\tag{77}
$$

All of these approximate degradability bounds can be evaluated both numerically and analytically on concrete quantum channels (cf. [12, 43]), and we note that our methods can be combined with other ways of extracting bounds (see, e.g., the latest [49]).

E. Improved continuity of filtered entropies

Our goal in this section is to use the above Theorem 1 to prove the following conjecture of Christandl, Ferrara, and Lancien [13, Conjecture 7], formulated below in a general fashion that the interested reader can find discussed in [13] below their statement of the problem.

Conjecture 12 [13]. *Let $\mathcal{F} \subseteq \mathcal{D}(\mathcal{H})$ be a set of states star-shaped around the maximally mixed state in dimension $D = \dim(\mathcal{H}) < \infty$, and let $\mathcal{L} \subseteq \mathrm{CPTP}(\mathcal{H} \to \mathcal{H}')$ be any set of channels. Then, does there exist a universal constant κ and a function $g : [0,1] \to \mathbb{R}_+$ with $\lim_{\varepsilon \to 0^+} g(\varepsilon) = 0$ such that for every $\rho, \sigma \in \mathcal{D}(\mathcal{H})$ satisfying*

$$\frac{1}{2}\|\rho - \sigma\|_{\mathcal{L}} \le \varepsilon \tag{78}$$

we have that

$$\left| D^{\mathcal{L}}(\rho\|\mathcal{F}) - D^{\mathcal{L}}(\sigma\|\mathcal{F}) \right| \overset{?}{\le} \kappa \varepsilon \log D + g(\varepsilon) . \tag{79}$$

Here, we defined

$$\|X\|_{\mathcal{L}} := \sup_{\Lambda \in \mathcal{L}} \|\Lambda(X)\|_1 , \tag{80}$$

$$D^{\mathcal{L}}(\rho\|\omega) := \sup_{\Lambda \in \mathcal{L}} D\big(\Lambda(\rho) \,\big\| \, \Lambda(\omega)\big), \tag{81}$$

$$D^{\mathcal{L}}(\rho\|\mathcal{F}) := \inf_{\omega \in \mathcal{F}} D^{\mathcal{L}}(\rho\|\omega) . \tag{82}$$

In essence, (79) is a sort of asymptotic continuity statement, but stronger, because it is given with respect to the '$\mathcal{L}$-filtered' norm $\|\cdot\|_{\mathcal{L}}$ instead of the full trace norm $\|\cdot\|_1$. If the full trace norm is considered, then asymptotic continuity is known, e.g. from [50, Theorem 11].

The above conjecture is known to be true when $\mathcal{L}$ is a set of measurements [51, Proposition 3], or a specific subset of channels [13, Proposition 6]. We will see that we can tackle the case of general sets of channels. Before we state the general result that will allow us to solve the conjecture, let us fix some terminology. An alternative measure of $\mathcal{L}$-filtered distance from a set of states $\mathcal{F}$ can be constructed by using the max-relative entropy (46) instead of the Umegaki relative entropy in (82). Namely,

$$D^{\mathcal{L}}_{\max}(\rho\|\omega) := \sup_{\Lambda \in \mathcal{L}} D_{\max}\big(\Lambda(\rho) \,\big\| \, \Lambda(\omega)\big), \tag{83}$$

$$D^{\mathcal{L}}_{\max}(\rho\|\mathcal{F}) := \inf_{\omega \in \mathcal{F}} D^{\mathcal{L}}_{\max}(\rho\|\omega) . \tag{84}$$

Before proceeding, we also need to construct the real function

$$g(x) := (1 + x)\log(1 + x) - x \log x , \tag{85}$$

defined for all $x \ge 0$, with the usual convention that $0 \log 0 = 0$. We now state the following general result.

Proposition 13. *Let $\mathcal{H}$ be an arbitrary (possibly infinite-dimensional) separable Hilbert space. Let $\mathcal{F} \subseteq \mathcal{D}(\mathcal{H})$ be a set of states that is star-shaped around some $\tau \in \mathcal{F}$, and let $\mathcal{L} \subseteq \mathrm{CPTP}(\mathcal{H} \to \mathcal{H}')$ be any set of channels. Then for all $\rho, \sigma \in \mathcal{D}(\mathcal{H})$ satisfying $D_{\max}(\rho\|\tau) < \infty$ and*

$$\frac{1}{2}\|\rho - \sigma\|_{\mathcal{L}} \le \varepsilon \tag{86}$$

it holds that

$$D^{\mathcal{L}}(\rho\|\mathcal{F}) - D^{\mathcal{L}}(\sigma\|\mathcal{F}) \leq \varepsilon D^{\mathcal{L}}_{\max}(\rho\|\tau) + g(\varepsilon) + h_2(\varepsilon)$$
$$\leq \varepsilon D_{\max}(\rho\|\tau) + g(\varepsilon) + h_2(\varepsilon)\,, \tag{87}$$

where the function g is defined by (85).

In particular, *if $\mathcal{F}$ is convex, then minimising over $\tau \in \mathcal{F}$ yields*

$$D^{\mathcal{L}}(\rho\|\mathcal{F}) - D^{\mathcal{L}}(\sigma\|\mathcal{F}) \leq \varepsilon D^{\mathcal{L}}_{\max}(\rho\|\mathcal{F}) + g(\varepsilon) + h_2(\varepsilon)\,. \tag{88}$$

As a simple corollary, we obtain a solution to Conjecture 12 by Christandl, Ferrara, and Lancien.

Corollary 14. *Let $\mathcal{F} \subseteq \mathcal{D}(\mathcal{H})$ be a set of states star-shaped around the maximally mixed state in dimension $D = \dim(\mathcal{H}) < \infty$, and let $\mathcal{L} \subseteq \mathrm{CPTP}(\mathcal{H} \to \mathcal{H}')$ be any set of channels. Then for all $\rho, \sigma \in \mathcal{D}(\mathcal{H})$ satisfying*

$$\frac{1}{2}\|\rho - \sigma\|_{\mathcal{L}} \leq \varepsilon\,, \tag{89}$$

we have that

$$\left| D^{\mathcal{L}}(\rho\|\mathcal{F}) - D^{\mathcal{L}}(\sigma\|\mathcal{F}) \right| \leq \varepsilon \log D + h_2(\varepsilon) + g(\varepsilon)\,, \tag{90}$$

where the function g is defined by (85).

The above result is remarkable because:

- it solves Conjecture 12;

- it generalises [50, Theorem 11] in three ways, namely (a) lifting the asymptotic continuity from the trace distance to $\mathcal{L}$-filtered distance (see also [51, Proposition 3]), (b) removing the assumption on the convexity of $\mathcal{F}$, and (c) removing any assumption on the set $\mathcal{L}$, which doesn't even need to be made of measurements;

- it generalises and significantly tightens [51, Proposition 3] and [13, Proposition 6], removing any assumption on the set of channels, which in particular no longer need to be (partial) measurements; finally,

- the estimate on the right-hand side seems reasonably tight, as its form is not too different from that found in [17, Lemma 7] (note that $g(x) = (1 + x)\, h_2\!\left(\frac{x}{1+x}\right)$).

We will now prove Proposition 13 and then deduce from it Corollary 14. In order to prove Proposition 13 with Theorem 1, we need to show that we can restrict the optimisation over free states in (82) to states with respect to which ρ has a controlled max-relative entropy (see (46) for a definition). This is needed because of the $D_{\max}$ term on the right-hand side of (12). We fix this with a little lemma that is already essentially contained in the proof of [51, Proposition 3] (and that we re-discovered independently).

Lemma 15. *For an arbitrary (possibly infinite-dimensional) separable Hilbert space $\mathcal{H}$, let $\mathcal{F} \subseteq \mathcal{D}(\mathcal{H})$ be a set of states star-shaped around some $\tau \in \mathcal{F}$,[1] and let $\mathcal{L} \subseteq \mathrm{CPTP}(\mathcal{H} \to \mathcal{H}')$ be any set of channels. Then for all $\rho \in \mathcal{D}(\mathcal{H})$ and all $q \in (0, 1)$, it holds that*

$$\inf_{\omega \in \mathcal{F},\ \omega \geq q\tau} D^{\mathcal{L}}(\rho\|\omega) + \log(1 - q) \leq D^{\mathcal{L}}(\rho\|\mathcal{F}) \leq \inf_{\omega \in \mathcal{F},\ \omega \geq q\tau} D^{\mathcal{L}}(\rho\|\omega)\,, \tag{91}$$

where $D^{\mathcal{L}}(\rho\|\mathcal{F})$ is defined by (82).

---

[1] This means that for all $\sigma \in \mathcal{F}$ and all $\lambda \in [0, 1]$, it holds that $\lambda\sigma + (1 - \lambda)\tau \in \mathcal{F}$.

Proof. The upper bound is trivial. As for the lower bound, for every $\omega \in \mathcal{F}$ one can construct $\omega' = (1-q)\omega + q\tau$, which is again a state in $\mathcal{F}$ due to the fact that this set is star-shaped around τ. Note also that $\omega \leq \frac{1}{1-q}\omega'$ and moreover $\omega' \geq q\tau$, so that by operator monotonicity of the logarithm (or, in other words, by applying [27, Proposition 5.1(i)]) we obtain

$$D^{\mathcal{L}}(\rho\|\omega) \geq D^{\mathcal{L}}\left(\rho \,\middle\|\, \frac{\omega'}{1-q}\right) = D^{\mathcal{L}}(\rho\|\omega') + \log(1-q) \geq \inf_{\omega'' \in \mathcal{F},\ \omega'' \geq q\tau} D^{\mathcal{L}}(\rho\|\omega') + \log(1-q). \quad (92)$$

Taking the infimum over $\omega \in \mathcal{F}$ on the left-hand side completes the proof. $\qquad\square$

We are now ready to prove Proposition 13 and Corollary 14.

Proof of Proposition 13. Let us fix an arbitrary $q \in (0,1)$. Then

$$
\begin{aligned}
D^{\mathcal{L}}(\rho\|\mathcal{F}) &\overset{(i)}{\leq} \inf_{\omega \in \mathcal{F},\ \omega \geq q\tau} D^{\mathcal{L}}(\rho\|\omega) \\
&= \inf_{\omega \in \mathcal{F},\ \omega \geq q\tau} \sup_{\Lambda \in \mathcal{L}} D\left(\Lambda(\rho) \,\middle\|\, \Lambda(\omega)\right) \\
&\overset{(ii)}{\leq} \inf_{\omega \in \mathcal{F},\ \omega \geq q\tau} \sup_{\Lambda \in \mathcal{L}} \left\{ D\left(\Lambda(\rho) \,\middle\|\, \Lambda(\omega)\right) + \varepsilon D^{\mathcal{L}}_{\max}(\rho\|\tau) + \varepsilon \log \frac{1}{q} + h_2(\varepsilon) \right\} \\
&\overset{(iii)}{\leq} D^{\mathcal{L}}(\sigma\|\mathcal{F}) + \log \frac{1}{1-q} + \varepsilon D^{\mathcal{L}}_{\max}(\rho\|\tau) + \varepsilon \log \frac{1}{q} + h_2(\varepsilon).
\end{aligned}
$$

Here: in (i) we restricted the infimum (the easy direction in Lemma 15); in (ii) we applied Theorem 1 in the simplified form (13), noting that $\frac{1}{2}\left\|\Lambda(\rho) - \Lambda(\sigma)\right\|_1 \leq \frac{1}{2}\|\rho - \sigma\|_{\mathcal{L}} \leq \varepsilon$ and observing that

$$D_{\max}\left(\Lambda(\rho) \,\middle\|\, \Lambda(\omega)\right) \leq D_{\max}\left(\Lambda(\rho) \,\middle\|\, \Lambda(\tau)\right) + \log \frac{1}{q} \leq D^{\mathcal{L}}_{\max}(\rho\|\tau) + \log \frac{1}{q} \quad (93)$$

via the triangle inequality for the max-relative entropy, applicable because $\omega \geq q\tau$; finally, in (iii) we used again Lemma 15 (this time the less trivial part of it).

It remains only to optimise over $q \in (0,1)$. Fortunately this can be done analytically, thanks to the simple formula

$$g(\varepsilon) = \inf_{0 < q < 1} \left\{ \varepsilon \log \frac{1}{q} + \log \frac{1}{1-q} \right\}. \quad (94)$$

Using it, we obtain that

$$
\begin{aligned}
D^{\mathcal{L}}(\rho\|\mathcal{F}) &\leq D^{\mathcal{L}}(\sigma\|\mathcal{F}) + \varepsilon D^{\mathcal{L}}_{\max}(\rho\|\tau) + g(\varepsilon) + h_2(\varepsilon) \\
&\overset{(iv)}{\leq} D^{\mathcal{L}}(\sigma\|\mathcal{F}) + \varepsilon D^{\mathcal{L}}_{\max}(\rho\|\tau) + g(\varepsilon) + h_2(\varepsilon),
\end{aligned}
\quad (95)
$$

where (iv) is simply by data processing. Finally, if $\mathcal{F}$ is convex, and hence star-shaped around all of its states, then we can minimise over $\tau \in \mathcal{F}$, thus obtaining also (88) and completing the proof. $\qquad\square$

Proof of Corollary 14. It suffices to apply Proposition 13 with $\tau = \mathbb{1}/D$ equal to the maximally mixed state. Indeed, noting that $\rho \leq \mathbb{1} = D\tau$, we have that $D_{\max}(\rho\|\mathcal{F}) \leq D_{\max}(\rho\|\tau) \leq \log D$, and similarly for σ. Applying (87) twice, the second time with ρ and σ exchanged, we obtain precisely (90). $\qquad\square$

F. Upper bound on transformation rates in infinite dimensions

Let $\rho = \rho_{AB}$ and $\sigma = \sigma_{A'B'}$ be two bipartite states of possibly infinite-dimensional systems AB and $A'B'$. A key quantity of interest in entanglement theory is the ***asymptotic transformation rate*** $\rho \to \sigma$, i.e. the maximum rate of production of copies of σ that can be achieved by consuming copies of ρ and using only local operations and classical communication (LOCC) [38]. This is formally defined by

$$R_{\mathrm{LOCC}}(\rho \to \sigma) := \sup \left\{ R : \lim_{n \to \infty} \inf_{\Lambda \in \mathrm{LOCC}_n} \frac{1}{2} \left\| \Lambda_n \left(\rho^{\otimes n} \right) - \sigma^{\otimes \lceil Rn \rceil} \right\|_1 = 0 \right\}, \tag{96}$$

where LOCC_n denotes the set of channels with input system $A^n : B^n$ and output system $A'^m : B'^m$, and $m := \lceil Rn \rceil$. The operationally most important entanglement measures, the distillable entanglement and the entanglement cost (see also Section III C), can be seen as special case of the above notion, according to the formulae $E_{d,\mathrm{LOCC}}(\rho) := R_{\mathrm{LOCC}}(\rho \to \Phi_2)$ and $E_{c,\mathrm{LOCC}}(\rho) := R_{\mathrm{LOCC}}(\Phi_2 \to \rho)^{-1}$.

In spite of their operational importance, asymptotic transformation rates are extremely difficult to evaluate and especially estimate from above. For this reason, one is often interested in tight upper bounds on them. A general strategy to construct upper bounds is to use the ***regularised relative entropy of entanglement***, defined by [52, 53]

$$E_R^\infty(\rho) := \lim_{n \to \infty} \frac{1}{n} E_R \left(\rho^{\otimes n} \right), \tag{97}$$

$$E_R(\rho) := D \left(\rho \,\|\, \mathcal{S}(A:B) \right) = \min_{\omega \in \mathcal{S}(A:B)} D(\rho \| \omega). \tag{98}$$

Here, $\mathcal{S}(A:B)$ denotes the set of ***separable states*** on the bipartite system AB, defined by [54]

$$\mathcal{S}(A:B) := \mathrm{cl\,conv} \left\{ |\alpha\rangle\langle\alpha|_A \otimes |\beta\rangle\langle\beta|_B : |\alpha\rangle_A \in \mathcal{H}_A, \ |\beta\rangle_B \in \mathcal{H}_B, \ \langle\alpha|\alpha\rangle = 1 = \langle\beta|\beta\rangle \right\}, \tag{99}$$

where cl conv represents the closed convex hull (with respect to the trace norm topology). The limit in (97) exists due to Fekete's lemma [55], because of the sub-additivity property $E_R \left(\rho^{\otimes n} \right) \leq n E_R(\rho)$. Furthermore, the infimum in the definition of $D \left(\rho \,\|\, \mathcal{S}(A:B) \right)$ in (98) is always achieved, also for infinite-dimensional systems [56, Theorem 5]. The importance of the regularised relative entropy of entanglement stems from its operational interpretation in the context of entanglement testing, guaranteed by the recently established generalised quantum Stein's lemma [57–60].

If ρ and σ are finite-dimensional states, with $\sigma \notin \mathcal{S}(A:B)$ entangled, i.e. not separable, it is known that [33, Eq. (146)]

$$R_{\mathrm{LOCC}}(\rho \to \sigma) \leq \frac{E_R^\infty(\rho)}{E_R^\infty(\sigma)}. \tag{100}$$

Note that $E_R^\infty(\sigma) > 0$, since E_R^∞ is a faithful measure of entanglement [57, 61]. The problem with (100) is that its standard proof [33, XV.E.2] makes use of *asymptotic continuity*, a strong form of continuity that does not apply to infinite-dimensional systems. Our main contribution in this section is to generalise (100) to the infinite-dimensional setting, thus avoiding the 'asymptotic continuity catastrophe' described in [14]. The price we will have to pay is a stronger regularity assumption on the target state σ; however, crucially, no regularity assumption is made on the LOCC protocols that can be employed in (96) to effect the transformation. Our result is as follows.

Proposition 16. *Let AB, $A'B'$ be arbitrary bipartite quantum systems, finite or infinite dimensional. Let $\rho = \rho_{AB}$ and $\sigma = \sigma_{A'B'}$ be two states with $\sigma \notin \mathcal{S}(A':B')$ and $E_{R,\max}(\sigma) < \infty$. Then*

$$R_{\mathrm{LOCC}}(\rho \to \sigma) \le R_{\mathrm{NE}}(\rho \to \sigma) \le \frac{E_R^\infty(\rho)}{E_R^\infty(\sigma)}. \tag{101}$$

In the above statement, $E_{R,\max}$ denotes the max-relative entropy version of (98), i.e.

$$E_{R,\max}(\rho) := \min_{\omega \in \mathcal{S}(A:B)} D_{\max}(\rho\|\omega), \tag{102}$$

also called the **logarithmic robustness of entanglement** [62–64]. As for the standard relative entropy of entanglement, also in this case we have the sub-additivity property

$$E_{R,\max}(\rho^{\otimes n}) \le n\, E_{R,\max}(\rho). \tag{103}$$

Furthermore, in (101) we denoted with $R_{\mathrm{NE}}(\rho \to \sigma)$ the asymptotic transformation rate under **non-entangling (NE) operations** [58, 65–67], obtained from (96) by replacing LOCC with the set of all channels that map separable states to separable states.

Proof of Proposition 16. Since the set of LOCC operations is a subset of that of NE operations, it is immediate to see that

$$R_{\mathrm{LOCC}}(\rho \to \sigma) \le R_{\mathrm{NE}}(\rho \to \sigma). \tag{104}$$

Thus, it suffices to prove the second inequality in (101). To this end, let R be an achievable rate, i.e. let it belong to the set in (96), with LOCC replaced by NE. Then there exists a sequence of NE operations Λ_n such that the states $\omega_n := \Lambda_n(\rho^{\otimes n})$ satisfy $\varepsilon_n := \frac{1}{2}\left\|\omega_n - \sigma^{\otimes \lceil Rn \rceil}\right\|_1 \xrightarrow[n\to\infty]{} 0$. We can now write that

$$\begin{aligned}
E_R\left(\sigma^{\otimes \lceil Rn \rceil}\right) - E_R\left(\rho^{\otimes n}\right) &\overset{(i)}{\le} E_R\left(\sigma^{\otimes \lceil Rn \rceil}\right) - E_R(\omega_n) \\
&\overset{(ii)}{\le} \varepsilon_n\, E_{R,\max}\left(\sigma^{\otimes \lceil Rn \rceil}\right) + g(\varepsilon_n) + h_2(\varepsilon_n) \\
&\overset{(iii)}{\le} \lceil Rn \rceil\, \varepsilon_n\, E_{R,\max}(\sigma) + g(\varepsilon_n) + h_2(\varepsilon_n).
\end{aligned} \tag{105}$$

The justification of the above inequalities is as follows. In (i) we simply observed that due to monotonicity under NE operations [66, Lemma IV.1] it holds that $E_R(\rho^{\otimes n}) \ge E_R(\Lambda_n(\rho^{\otimes n})) = E_R(\omega_n)$. In (ii) we applied Proposition 13, and in particular (88), with $\mathcal{F} = \mathcal{S}(A'^{\lceil Rn \rceil} : B'^{\lceil Rn \rceil})$ and $\mathcal{L} = \mathrm{CPTP}(\mathcal{H}_{AB} \to \mathcal{H}_{AB})$, so that $D^{\mathcal{L}}(\cdot\|\mathcal{F}) = E_R(\cdot)$ and $D_{\max}^{\mathcal{L}}(\cdot\|\mathcal{F}) = E_{R,\max}(\cdot)$. Finally, in (iii) we employed (103).

Diving both sides of (105) by n and taking the limit $n \to \infty$ yields immediately $RE_R^\infty(\sigma) - E_R^\infty(\rho) \le 0$. The claim follows by taking the supremum over R and dividing by $E_R^\infty(\sigma)$, which is strictly positive due to the faithfulness of E_R^∞, whose proof in [61] works equally well for finite- and infinite-dimensional systems. $\qquad\square$

IV. CONCLUSION

Further applications in quantum information are waiting to be explored, but the main open question remains Conjecture 6 about the conditional entropy for general bipartite quantum states.

Whereas many of applications do seem to have fixed marginals, one important question we have to leave open concerns tight continuity bounds for quantum conditional mutual information, which would improve on [16, 20–22]. A strictly related question is that of establishing tight continuity bounds for the squashed entanglement [15, 68], improving on the original one by Alicki–Fannes [16].

Another special case that is particularly relevant for improving the known continuity bounds on capacities of quantum channels (see [30, 31]) would be a version of Theorem 5 in terms of quantum mutual information, where one marginal of the bipartite states is equal (but not the other one). This is not covered by Theorem 1 and our proof techniques, but one might nevertheless conjecture even more generally that for states ρ_{AB}, σ_{AB} and with trace distance $\frac{1}{2}\|\rho_{AB} - \sigma_{AB}\|_1 \leq \varepsilon$ small enough, one has something like

$$\left| I(A:B)_\rho - I(A:B)_\sigma \right| \overset{?}{\leq} \varepsilon \log \left(\min \left\{ |A|^2, |B|^2 \right\} - 1 \right) + h_2(\varepsilon) \tag{106}$$

with the quantum mutual information $I(A:B)_\rho := H(A)_\rho + H(B)_\rho - H(AB)_\rho$. In fact, to the best of our knowledge, this question seems to be open even in the classical case.[2]

Note. Our result on the uniform continuity bound for the conditional entropy in the case of equal marginals (Theorem 5) was obtained with different techniques by Audenaert, Bergh, Datta, Jabbour, Capel, and Gondolf.

---

[1] P. E. Frenkel. Integral formula for quantum relative entropy implies data processing inequality. *Quantum*, 7:1102, 2023. 2

[2] A. Jenčová. Recoverability of quantum channels via hypothesis testing. *Lett. Math. Phys.*, 114(1):31, 2024. 2

[3] C. Hirche and M. Tomamichel. Quantum Rényi and f-divergences from integral representations. *Commun. Math. Phys.*, 405(9):208, 2024. 2

[4] H. Umegaki. Conditional expectation in an operator algebra. IV. Entropy and information. *Kodai Math. Sem. Rep.*, 14(2):59–85, 1962. 2

[5] F. Hiai and D. Petz. The proper formula for relative entropy and its asymptotics in quantum probability. *Comm. Math. Phys.*, 143(1):99–114, 1991. 2

[6] K. M. R. Audenaert and J. Eisert. Continuity bounds on the quantum relative entropy. *J. Math. Phys.*, 46(10):102104, 2005. 2

[7] M. Fannes. A continuity property of the entropy density for spin lattice systems. *Commun. Math. Phys.*, 31(4):291–294, 1973. 2, 3, 7

[8] K. M. R. Audenaert. A sharp continuity estimate for the von Neumann entropy. *J. Phys. A*, 40(28):8127, 2007. 2, 7

[2] We have not intensively tested the above mutual information inequality numerically.

[9] D. Petz. *Quantum Information Theory and Quantum Statistics*. Theoretical and Mathematical Physics. Springer-Verlag, Berlin, 2008. 2, 3, 7

[10] M. A. Alhejji and G. Smith. A tight uniform continuity bound for equivocation. In *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, pages 2270–2274, 2020. 3, 8, 9, 11

[11] M. M. Wilde. Optimal uniform continuity bound for conditional entropy of classical–quantum states. *Quantum Inf. Process.*, 19(2):61, 2020. 3, 8, 9, 11, 20

[12] D. Sutter, V. B. Scholz, A. Winter, and R. Renner. Approximate degradable quantum channels. *IEEE Trans. Inf. Theory*, 63(12):7832–7844, 2017. 4, 12, 13, 14

[13] M. Christandl, R. Ferrara, and C. Lancien. Random private quantum states. *IEEE Trans. Inf. Theory*, 66(7):4621–4640, 2020. 4, 15, 16

[14] G. Ferrari, L. Lami, T. Theurer, and M. B. Plenio. Asymptotic state transformations of continuous variable resources. *Commun. Math. Phys.*, 398(1):291–351, 2023. 4, 18

[15] M. Christandl and A. Winter. Squashed entanglement: An additive entanglement measure. *J. Math. Phys.*, 45(3):829–840, 2004. 7, 20

[16] R. Alicki and M. Fannes. Continuity of quantum conditional information. *J. Phys. A*, 37(5):L55, 2004. 8, 9, 20

[17] A. Winter. Tight uniform continuity bounds for quantum entropies: Conditional entropy, relative entropy distance and energy constraints. *Commun. Math. Phys.*, 347(1):291–313, 2016. 7, 9, 11, 12, 16

[18] B. M. Terhal and P. Horodecki. Schmidt number for density matrices. *Phys. Rev. A*, 61:040301, Mar 2000. 7

[19] M. M. Wolf. Quantum Channels & Operations: Guided Tour. Lecture notes, 2012. 8

[20] B. Synak-Radtke and M. Horodecki. On asymptotic continuity of functions of quantum states. *J. Phys. A*, 39(26):L423–L437, 2006. 8, 20

[21] M. Mosonyi and F. Hiai. On the quantum Rényi relative entropies and related capacity formulas. *IEEE Trans. Inf. Theory*, 57(4):2474–2487, 2011. 9

[22] A. Winter and D. Yang. Operational resource theory of coherence. *Phys. Rev. Lett.*, 116:120404, 2016. 8, 20

[23] M. A. Alhejji. *Some Problems Concerning Quantum Channels and Entropies*. PhD thesis, University of Colorado, 2023. 9

[24] M. Vempati, S. Shah, N. Ganguly, and I. Chakrabarty. A-unital operations and quantum conditional entropy. *Quantum*, 6:641, 2022. 9

[25] S. Brandsen, I. J. Geng, M. M. Wilde, and G. Gour. Quantum conditional entropy from information-theoretic principles. *Preprint arXiv:2110.15330*, 2021. 9

[26] G. Gour, A. Grudka, M. Horodecki, W. Kłobus, J. Łodyga, and V. Narasimhachar. Conditional uncertainty principle. *Phys. Rev. A*, 97(4):042130, 2018. 9

[27] M. Ohya and D. Petz. *Quantum Entropy and Its Use*. Theoretical and Mathematical Physics. Springer Berlin Heidelberg, 2004. 9, 17

[28] M. Tomamichel and V. Y. F. Tan. Second-order asymptotics for the classical capacity of image-additive quantum channels. *Commun. Math. Phys.*, 338(1):103–137, 2015. 9

[29] N. Datta. Min- and max-relative entropies and a new entanglement monotone. *IEEE Trans. Inf. Theory*, 55(6):2816–2826, 2009. 10

[30] D. Leung and G. Smith. Continuity of quantum channel capacities. *Commun. Math. Phys.*, 292(1):201–215, 2009. 10, 12, 20

[31] M. E. Shirokov. Tight uniform continuity bounds for the quantum conditional mutual information, for the Holevo quantity, and for capacities of quantum channels. *Journal of Mathematical Physics*, 58(10):102202, 10 2017. 10, 20

[32] M. M. Wilde. *Quantum Information Theory*. Cambridge University Press, 2nd edition, 2017. 10

[33] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki. Quantum entanglement. *Rev. Mod. Phys.*, 81:865–942, 2009. 11, 18

[34] P. M. Hayden, M. Horodecki, and B. M. Terhal. The asymptotic entanglement cost of preparing a quantum state. *J. Phys. A*, 34(35):6891–6898, 2001. 11

[35] G. Vidal and J. I. Cirac. Irreversibility in asymptotic manipulations of entanglement. *Phys. Rev. Lett.*, 86:5803–5806, 2001.

[36] D. Yang, M. Horodecki, R. Horodecki, and B. Synak-Radtke. Irreversibility for all bound entangled states. *Phys. Rev. Lett.*, 95:190501, 2005.

[37] H. Yamasaki, K. Kuroiwa, P. Hayden, and L. Lami. Entanglement cost for infinite-dimensional physical systems. *Preprint arXiv:2401.09554*, 2024. 11

[38] E. Chitambar, D. Leung, L. Mančinska, M. Ozols, and A. Winter. Everything you always wanted to know about LOCC (but were afraid to ask). *Commun. Math. Phys.*, 328(1):303–326, 2014. 11, 18

[39] M. Horodecki, P. Horodecki, and R. Horodecki. Limits for entanglement measures. *Phys. Rev. Lett.*, 84:2014–2017, 2000. 11

[40] M. J. Donald, M. Horodecki, and O. Rudolph. The uniqueness theorem for entanglement measures. *J. Math. Phys.*, 43(9):4252–4272, 2002.

[41] M. Horodecki. Entanglement measures. *Quantum Inf. Comput.*, 1(1):3–26, 2001. 11

[42] M. A. Nielsen. Continuity bounds for entanglement. *Phys. Rev. A*, 61:064301, 2000. 11

[43] F. Leditzky, D. Leung, and G. Smith. Quantum and private capacities of low-noise channels. *Phys. Rev. Lett.*, 120:160503, Apr 2018. 12, 14

[44] S. Lloyd. Capacity of the noisy quantum channel. *Phys. Rev. A*, 55:1613–1622, 1997. 12

[45] P. Shor. Lecture notes. MSRI Workshop on Quantum Computation, 2002.

[46] I. Devetak. The private classical capacity and quantum capacity of a quantum channel. *IEEE Trans. Inf. Theory*, 51(1):44–55, 2005.

[47] D. Kretschmann and R. F. Werner. Tema con variazioni: quantum channel capacity. *New J. Phys.*, 6:26–26, 2004. 12

[48] M. M. Wilde, M. Berta, C. Hirche, and E. Kaur. Amortized channel divergence for asymptotic quantum channel discrimination. *Lett. Math. Phys.*, 2020. 14

[49] M. G. Jabbour and N. Datta. Tightening continuity bounds on entropies and bounds on quantum capacities. *Preprint arXiv:2310.17329*, 2023. 14

[50] J. Schindler and A. Winter. Continuity bounds on observational entropy and measured relative entropies. *J. Math. Phys.*, 64(9):092201, 2023. 15, 16

[51] K. Li and A. Winter. Relative entropy and squashed entanglement. *Commun. Math. Phys.*, 326(1):63–80, 2014. 15, 16

[52] V. Vedral, M. B. Plenio, M. A. Rippin, and P. L. Knight. Quantifying entanglement. *Phys. Rev. Lett.*, 78:2275–2279, 1997. 18

[53] V. Vedral and M. B. Plenio. Entanglement measures and purification procedures. *Phys. Rev. A*, 57:1619–1633, 1998. 18

[54] R. F. Werner. Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model. *Phys. Rev. A*, 40:4277–4281, 1989. 18

[55] M. Fekete. Über die Verteilung der Wurzeln bei gewissen algebraischen Gleichungen mit ganzzahligen Koeffizienten. *Math. Z.*, 17(1):228–249, 1923. 18

[56] L. Lami and M. E. Shirokov. Attainability and lower semi-continuity of the relative entropy of entanglement and variations on the theme. *Ann. Henri Poincaré*, 24(12):4069–4137, 2023. 18

[57] F. G. S. L. Brandão and M. B. Plenio. A generalization of quantum Stein's lemma. *Commun. Math. Phys.*, 295(3):791–828, 2010. 18

[58] M. Berta, F. G. S. L. Brandão, G. Gour, L. Lami, M. B. Plenio, B. Regula, and M. Tomamichel. On a gap in the proof of the generalised quantum Stein's lemma and its consequences for the reversibility of quantum resources. *Quantum*, 7:1103, 2023. 19

[59] M. Hayashi and H. Yamasaki. Generalized quantum Stein's lemma and second law of quantum resource theories. *Preprint arXiv:2408.02722*, 2024.

[60] L. Lami. A solution of the generalised quantum stein's lemma. *Preprint arXiv:2408.06410*, 2024. 18

[61] M. Piani. Relative entropy of entanglement and restricted measurements. *Phys. Rev. Lett.*, 103:160504, 2009. 18, 19

[62] G. Vidal and R. Tarrach. Robustness of entanglement. *Phys. Rev. A*, 59:141–155, 1999. 19

[63] M. Steiner. Generalized robustness of entanglement. *Phys. Rev. A*, 67:054305, 2003.

[64] N. Datta. Max-relative entropy of entanglement, alias log robustness. *Int. J. Quantum Inf.*, 07(02):475–491, 2009. 19

[65] F. G. S. L. Brandão and M. B. Plenio. Entanglement theory and the second law of thermodynamics. *Nat. Phys.*, 4:873–877, 2008. 19

[66] F. G. S. L. Brandão and M. B. Plenio. A reversible theory of entanglement and its relation to the second law. *Commun. Math. Phys.*, 295(3):829–851, 2010. 19

[67] L. Lami and B. Regula. No second law of entanglement manipulation after all. *Nat. Phys.*, 19(2):184–189, 2023. 19

[68] R. R. Tucci. Quantum entanglement and conditional information transmission. *Preprint arXiv:quant-ph/9909041*, 1999. 20