Communications in
**Mathematical
Physics**

CrossMark

Tight Uniform Continuity Bounds for Quantum Entropies: Conditional Entropy, Relative Entropy Distance and Energy Constraints

Andreas Winter

ICREA & Física Teòrica: Informació i Fenòmens Quàntics, Universitat Autònoma de Barcelona, 08193 Bellaterra (Barcelona), Spain. E-mail: andreas.winter@uab.cat

Abstract: We present a bouquet of continuity bounds for quantum entropies, falling broadly into two classes: first, a tight analysis of the Alicki–Fannes continuity bounds for the conditional von Neumann entropy, reaching almost the best possible form that depends only on the system dimension and the trace distance of the states. Almost the same proof can be used to derive similar continuity bounds for the relative entropy distance from a convex set of states or positive operators. As applications, we give new proofs, with tighter bounds, of the asymptotic continuity of the relative entropy of entanglement, E_R, and its regularization E_R^∞, as well as of the entanglement of formation, E_F. Using a novel "quantum coupling" of density operators, which may be of independent interest, we extend the latter to an asymptotic continuity bound for the regularized entanglement of formation, aka entanglement cost, $E_C = E_F^\infty$. Second, we derive analogous continuity bounds for the von Neumann entropy and conditional entropy in infinite dimensional systems under an energy constraint, most importantly systems of multiple quantum harmonic oscillators. While without an energy bound the entropy is discontinuous, it is well-known to be continuous on states of bounded energy. However, a quantitative statement to that effect seems not to have been known. Here, under some regularity assumptions on the Hamiltonian, we find that, quite intuitively, the Gibbs entropy at the given energy roughly takes the role of the Hilbert space dimension in the finite-dimensional Fannes inequality.

1. Introduction

On finite dimensional systems, the von Neumann entropy $S(\rho) = -\operatorname{Tr} \rho \log \rho$ is continuous, but this becomes useful only once one has explicit continuity bounds, most significantly the one due to Fannes [13], the sharpest form of which is the following:

Lemma 1 (Audenaert [3], Petz [32]). *For states ρ and σ on a Hilbert space A of dimension $d = |A| < \infty$, if $\frac{1}{2}\|\rho - \sigma\|_1 \leq \epsilon \leq 1$, then*

$$|S(\rho) - S(\sigma)| \leq \begin{cases} \epsilon \log(d-1) + h(\epsilon) & \text{if } \epsilon \leq 1 - \frac{1}{d}, \\ \log d & \text{if } \epsilon > 1 - \frac{1}{d}, \end{cases}$$

with $h(x) = H(x, 1-x) = -x \log x - (1-x) \log(1-x)$ the binary entropy. A simplified, but universal bound reads

$$|S(\rho) - S(\sigma)| \leq \epsilon \log d + h(\epsilon).$$

We include a short proof for self-containedness, and also because it deserves to be known better. It seems that it was first found by Petz [32, Thm. 3.8], who credits Csiszár for the classical case; the latter seems to have appeared first in Zhang's paper [56] (see also [34]).

Proof. We only have to treat the case $\epsilon \leq 1 - \frac{1}{d}$. We begin with the classical case of two probability distributions p and q on the same ground set of d elements. It is well known, and in fact elementary to confirm, that one can find two jointly distributed random variables, $X \sim p$ and $Y \sim q$ (meaning X is distributed according to the probability law p, and Y according to q), with $\Pr\{X \neq Y\} = \frac{1}{2}\|p - q\|_1 \leq \epsilon$. The crucial idea is to let $\Pr\{X = Y = x\} = \min(p_x, q_x)$ and to distribute the remaining probability weight suitably off the diagonal. (This is also the minimum probability over all such coupled random variables [56]. For the reader with a taste for the sophisticated, this is the Kantorovich–Rubinstein dual formula for the Wasserstein distance in the case of the trivial metric $d(x, y) = 1$ for all $x \neq y$ and $d(x, x) = 0$, cf. the broad survey [5].) Then, by the monotonicity of the Shannon entropy under taking marginals and Fano's inequality (see [9]),

$$H(X) - H(Y) \leq H(XY) - H(Y)$$
$$= H(X|Y) \leq \epsilon \log(d-1) + h(\epsilon),$$

and likewise for $H(Y) - H(X)$. [For the simplified bound, we use $H(X|Y) \leq \epsilon \log d + h(\epsilon)$.]

Next, we reduce the quantum case to the classical one: W.l.o.g. $S(\rho) \leq S(\sigma)$, and consider the dephasing operation E in the eigenbasis of ρ, which maps ρ to itself, a diagonal matrix with a probability distribution p along the diagonal, and σ to $E(\sigma)$, a diagonal matrix with a probability distribution q along the diagonal. Hence

$$H(p) = S(\rho) \leq S(\sigma) \leq S\big(E(\sigma)\big) = H(q).$$

At the same time, $\|p - q\|_1 = \|E(\rho) - E(\sigma)\|_1 \leq \|\rho - \sigma\|_1$, and so, using the classical case,

$$|S(\rho) - S(\sigma)| \leq H(q) - H(p) \leq \epsilon \log(d-1) + h(\epsilon).$$

Note that the inequality is tight for all ϵ and d, e.g. by $\sigma = |0\rangle\langle 0|$ and $\rho = (1 - \epsilon)|0\rangle\langle 0| + \frac{\epsilon}{d-1}(\mathbb{1} - |0\rangle\langle 0|)$. $\quad\square$

We are interested in bounds of the above form, i.e., only referring to the trace distance of the states and some general global parameter specifying the system, for a number of entropic quantities, starting with the conditional von Neumann entropy, relative entropy distances from certain sets, etc, which have numerous applications in quantum information theory and quantum statistical physics. Furthermore, and perhaps even more urgently, in situations of infinite dimensional Hilbert spaces, where the above form of the Fannes inequality becomes trivial.

The rest of the paper is structured as follows: in Sect. 2 we present and prove an almost tight version of Lemma 1 for the conditional entropy (originally due to Alicki and Fannes [1]), then in Sect. 3 we generalize the principle behind our proof to a family of relative entropy distance measures from a convex set; in these two sections we also present some illustrative applications of the conditional entropy bounds to two entanglement measures, E_R and E_F, as well as their regularizations. In Sect. 4 we expand the methodology of the first part of the paper to infinite dimensional systems, where Fannes-type continuity bounds are obtained under an energy constraint for a broad class of Hamiltonians, and specifically for quantum harmonic oscillators. All entropy continuity bounds are stated as *Lemmas*, while the applications appear as *Corollaries*, and two auxiliary results (on "quantum coupling" of density matrices) as *Propositions*. The absence of *Theorems* is meant to encourage readers to apply the results presented here.

2. Conditional Entropy

Alicki and Fannes [1] proved an extension of the Fannes inequality for the conditional entropy

$$S(A|B)_\rho = S(\rho^{AB}) - S(\rho^B),$$

defined for states ρ on a bipartite (tensor product) Hilbert space $A \otimes B$. While a double application of Lemma 1 would yield such a bound involving both the dimensions of A and B, Alicki and Fannes show that if $\|\rho - \sigma\|_1 \le \epsilon \le 1$, then

$$\left| S(A|B)_\rho - S(A|B)_\sigma \right| \le 4\epsilon \log |A| + 2h(\epsilon).$$

In particular, this form is independent of the dimension of B, which might even be infinite. Note that for classical, Shannon, conditional entropy, an inequality like the above can be obtained from Lemma 1 by convex combination, resulting in a bound like that of Lemma 1 (see below).

The Alicki–Fannes inequality has several applications in quantum information theory, from the proof of asymptotic continuity of entanglement measures—most notably squashed entanglement [7] and conditional entanglement of mutual information (CEMI) [55]—to the continuity of quantum channel capacities [26], and on to the recent discussion of approximately degradable channels [44].

We present a simple proof of the Alicki–Fannes inequality that yields the stronger form of Lemma 2. One of the themes of the present paper, to which we draw attention here, is the use of entropy inequalities in the proofs. In particular, we make use of the concavity of the conditional entropy (which is equivalent to strong subadditivity of the von Neumann entropy) [27]. In the following proof we will specifically rely on two inequalities expressing the concavity of the entropy and the fact that it is not "too concave" [23]:

$$\sum_i p_i S(\rho_i) \le S\left(\sum_i p_i \rho_i\right) \le \sum_i p_i S(\rho_i) + H(p). \tag{1}$$

By introducing a bipartite state $\rho = \sum_i p_i \rho_i^A \otimes |i\rangle\langle i|^I$, this is seen to be equivalent to

$$S(A|I) \le S(A) \le S(AI) = S(A|I) + S(I),$$

which consists of two applications of strong subadditivity.

Lemma 2. *For states ρ and σ on a Hilbert space $A \otimes B$, if $\frac{1}{2}\|\rho - \sigma\|_1 \leq \epsilon \leq 1$, then*

$$\left| S(A|B)_\rho - S(A|B)_\sigma \right| \leq 2\epsilon \log |A| + (1 + \epsilon)\, h\!\left(\frac{\epsilon}{1 + \epsilon}\right).$$

If B is classical in the sense that both ρ and σ are so-called qc-states, i.e. with an orthonormal basis $\{|x\rangle\}$,

$$\rho = \sum_x p_x \rho_x^A \otimes |x\rangle\langle x|^B, \quad \sigma = \sum_x q_x \sigma_x^A \otimes |x\rangle\langle x|^B,$$

and analogously if both are cq-states, then this can be tightened to

$$\left| S(A|B)_\rho - S(A|B)_\sigma \right| \leq \epsilon \log |A| + (1 + \epsilon)\, h\!\left(\frac{\epsilon}{1 + \epsilon}\right).$$

Proof. The right hand side is monotonic in ϵ, hence we may assume $\frac{1}{2}\|\rho - \sigma\|_1 = \epsilon$. Let $\epsilon\Delta = (\rho - \sigma)_+$ be the positive part of $\rho - \sigma$. Note that because this difference is traceless and its trace norm equals 2ϵ, Δ is a bona fide state. Furthermore,

$$\begin{aligned}
\rho &= \sigma + (\rho - \sigma) \\
&\leq \sigma + \epsilon\Delta \\
&= (1 + \epsilon)\left(\frac{1}{1 + \epsilon}\sigma + \frac{\epsilon}{1 + \epsilon}\Delta\right) \\
&=: (1 + \epsilon)\omega.
\end{aligned}$$

By letting $\epsilon\Delta' := (1 + \epsilon)\omega - \rho$, we obtain another state Δ', such that

$$\omega = \frac{1}{1 + \epsilon}\sigma + \frac{\epsilon}{1 + \epsilon}\Delta = \frac{1}{1 + \epsilon}\rho + \frac{\epsilon}{1 + \epsilon}\Delta'. \tag{2}$$

This is a slightly optimized version of the trick in the proof of Alicki and Fannes [1]; cf. [28].

Now, we use the following well-known variational characterization of the conditional entropy:

$$-S(A|B)_\omega = \min_\xi D\!\left(\omega^{AB} \| \mathbb{1}^A \otimes \xi^B\right),$$

where $D(\rho\|\sigma) = \operatorname{Tr}\rho(\log\rho - \log\sigma)$ is the quantum relative entropy [30,47]. Choosing an optimal state ξ for ω (which is $\xi = \omega^B$), we have, from Eq. (2),

$$\begin{aligned}
S(A|B)_\omega &= -D\!\left(\omega^{AB} \| \mathbb{1}^A \otimes \xi^B\right) \\
&= S(\omega) + \operatorname{Tr}\omega\log\xi^B \\
&\leq h\!\left(\frac{\epsilon}{1 + \epsilon}\right) + \frac{1}{1 + \epsilon}S(\rho) + \frac{\epsilon}{1 + \epsilon}S(\Delta') \\
&\quad + \frac{1}{1 + \epsilon}\operatorname{Tr}\rho\log\xi^B + \frac{\epsilon}{1 + \epsilon}\operatorname{Tr}\Delta'\log\xi^B \\
&= h\!\left(\frac{\epsilon}{1 + \epsilon}\right) - \frac{1}{1+\epsilon}D(\rho\|\mathbb{1}\otimes\xi) - \frac{\epsilon}{1+\epsilon}D(\Delta'\|\mathbb{1}\otimes\xi) \\
&\leq h\!\left(\frac{\epsilon}{1 + \epsilon}\right) + \frac{1}{1 + \epsilon}S(A|B)_\rho + \frac{\epsilon}{1 + \epsilon}S(A|B)_{\Delta'},
\end{aligned}$$

where in the third line we have used the concavity upper bound from Eq. (1). Using the other decomposition in Eq. (2), the concavity of the conditional entropy, i.e. the lower bound in Eq. (1), gives

$$S(A|B)_\omega \geq \frac{1}{1+\epsilon} S(A|B)_\sigma + \frac{\epsilon}{1+\epsilon} S(A|B)_\Delta.$$

Putting these two bounds together and multiplying by $1 + \epsilon$, we arrive at

$$S(A|B)_\sigma - S(A|B)_\rho \leq \epsilon\big(S(A|B)_{\Delta'} - S(A|B)_\Delta\big) + (1+\epsilon)\, h\!\left(\frac{\epsilon}{1+\epsilon}\right).$$

The proof of the general bound is concluded observing that the conditional entropy of any state is bounded between $-\log|A|$ and $+\log|A|$.

For the case of two qc-states or two cq-states as above, note that the states Δ and Δ' are of the same, qc-form (cq-form, resp.), and so their conditional entropies are between 0 and $\log|A|$.   □

Remark 3. Lemma 2 is almost best possible, as we can see by considering the example of $\sigma^{AB} = \Phi_d$, the maximally entangled state on $A = B = \mathbb{C}^d$, and $\rho^{AB} = (1-\epsilon)\Phi_d + \frac{\epsilon}{d^2-1}(\mathbb{1} - \Phi_d)$. Clearly, $\frac{1}{2}\|\rho - \sigma\|_1 = \epsilon$, while

$$S(A|B)_\rho - S(A|B)_\sigma = \big(\epsilon \log(d^2 - 1) + h(\epsilon) - \log d\big) - (-\log d)$$
$$= 2\epsilon \log d + h(\epsilon) - O\left(\frac{\epsilon}{d^2}\right).$$

This asymptotically matches Lemma 2 for large d and small ϵ.

As an application of Lemma 2, we can obtain tighter continuity bounds on various quantum channel capacities, simply substituting our tighter bound rather than the original formulation of Alicki and Fannes in the proofs of Leung and Smith [26].

As a token, we demonstrate a tight version of the asymptotic continuity of the entanglement of formation [4],

$$E_F(\rho) = \inf \sum_x p_x S(\mathrm{Tr}_B \rho_x) \text{ s.t. } \rho = \sum_x p_x \rho_x$$

for a state ρ^{AB} on the bipartite system $A \otimes B$, originally due to Nielsen [29]. We then go on to prove asymptotic continuity for its regularization, the entanglement cost [16],

$$E_C(\rho) = E_F^\infty(\rho) = \lim_{n\to\infty} \frac{1}{n} E_F(\rho^{\otimes n}),$$

which, albeit following the general "telescoping" strategy of [26], requires a new idea, and seems not to have been known before [8]. Note that E_C is different from E_F [15].

Corollary 4. *Let ρ and σ be states on the system $A \otimes B$, denoting the smaller of the two dimensions by d. Then, $\frac{1}{2}\|\rho - \sigma\|_1 \leq \epsilon$ implies, with $\delta = \sqrt{\epsilon(2-\epsilon)}$,*

$$|E_F(\rho) - E_F(\sigma)| \leq \delta \log d + (1+\delta)\, h\!\left(\frac{\delta}{1+\delta}\right),$$
$$|E_C(\rho) - E_C(\sigma)| \leq 2\delta \log d + (1+\delta)\, h\!\left(\frac{\delta}{1+\delta}\right).$$

Note that these bounds only depend on the smaller of the two dimensions, in contrast to [29]; in particular, they apply even in the case that one of the two Hilbert spaces is infinite dimensional.

Proof. We may assume w.l.o.g. that $E_F(\rho) \geq E_F(\sigma)$ and $|B| \geq |A| = d$. Choose a purifying system $R \simeq AB$, and pure states φ^{ABR} and ψ^{ABR} with $\varphi^{AB} = \rho$ and $\psi^{AB} = \sigma = \psi^R$ such that

$$|\langle \varphi | \psi \rangle| = F(\rho, \sigma) \geq 1 - \epsilon,$$

thus $\frac{1}{2}\|\varphi - \psi\|_1 \leq \delta = \sqrt{1 - (1 - \epsilon)^2}$. Here, $F(\rho, \sigma) = \|\sqrt{\rho}\sqrt{\sigma}\|_1$ is the fidelity between two quantum states, and we have used that it is related to the trace distance by these well-known inequalities [14]:

$$1 - F(\rho, \sigma) \leq \frac{1}{2}\|\rho - \sigma\|_1 \leq \sqrt{1 - F(\rho, \sigma)^2}. \tag{3}$$

By an observation of Schrödinger (which he called "steering") in the context of his investigation of quantum entanglement [35,36], cf. [22], for any convex decomposition $\sigma = \sum_x p_x \sigma_x$, there exists a measurement POVM (M_x) on R such that $p_x \sigma_x = \text{Tr}_R \psi(\mathbb{1}^{AB} \otimes M_x^R)$. Introducing the qc-channel $\mathcal{M}(\xi) = \sum_x \text{Tr}\, \xi M_x |x\rangle\langle x|$ from R to a suitable space X, we then have

$$\widetilde{\sigma} := (\text{id}_{AB} \otimes \mathcal{M})\psi = \sum_x p_x \sigma_x^{AB} \otimes |x\rangle\langle x|^X,$$

$$\text{and} \quad S(A|X)_{\widetilde{\sigma}} = \sum_x p_x S(\text{Tr}_B \sigma_x). \tag{4}$$

Let us choose an optimal decomposition for the purpose of entanglement of formation, and the corresponding POVM and quantum channel, i.e. $E_F(\sigma) = S(A|X)_{\widetilde{\sigma}}$. Applying the same to φ^{ABR}, we obtain

$$\widetilde{\rho} := (\text{id}_{AB} \otimes \mathcal{M})\varphi = \sum_x q_x \rho_x^{AB} \otimes |x\rangle\langle x|^X,$$

with $q_x = \text{Tr}\, \varphi^R M_x$. Hence,

$$E_F(\rho) \leq \sum_x p_x S(\text{Tr}_B \rho_x) = S(A|X)_{\widetilde{\rho}}.$$

Observe that by the contractivity of the trace norm under cptp maps,

$$\delta \geq \|\psi - \varphi\|_1 \geq \|\widetilde{\sigma} - \widetilde{\rho}\|_1.$$

Now we can invoke the classical part of Lemma 2,

$$E_F(\rho) - E_F(\sigma) \leq S(A|X)_{\widetilde{\rho}} - S(A|X)_{\widetilde{\sigma}}$$

$$\leq \delta \log d + (1 + \delta)\, h\left(\frac{\delta}{1 + \delta}\right),$$

and we are done.

For the regularization, consider any integer n and

$$
\left| E_F\left(\rho^{\otimes n}\right) - E_F\left(\sigma^{\otimes n}\right) \right|
$$

$$
= \left| \sum_{t=1}^{n} E_F\left(\rho^{\otimes t} \otimes \sigma^{\otimes n-t}\right) - E_F\left(\rho^{\otimes t-1} \otimes \sigma^{\otimes n-t+1}\right) \right|
$$

$$
\leq \sum_{t=1}^{n} \left| E_F(\rho \otimes \Omega_t) - E_F(\sigma \otimes \Omega_t) \right|, \tag{5}
$$

with $\Omega_t = \rho^{\otimes t-1} \otimes \sigma^{\otimes n-t}$. The proof will be concluded by showing that for any $\Omega^{A'B'}$,

$$
\left| E_F(\rho \otimes \Omega) - E_F(\sigma \otimes \Omega) \right| \leq 2\delta \log d + (1 + \delta) h\left(\frac{\delta}{1 + \delta}\right),
$$

as this will imply from Eq. (5) that

$$
\frac{1}{n}\left| E_F\left(\rho^{\otimes n}\right) - E_F\left(\sigma^{\otimes n}\right) \right| \leq 2\delta \log d + (1 + \delta) h\left(\frac{\delta}{1 + \delta}\right).
$$

To see this, assume again w.l.o.g. that $E_F(\rho \otimes \Omega) \geq E_F(\sigma \otimes \Omega)$, and choose a purification υ of Ω on $A'B'R'$, with $R' \simeq A'B'$. Besides the purification ψ^{ABR} of σ, we now need a state (not generally pure) Θ^{ABR} with $\Theta^{AB} = \rho$ and $\Theta^R = \psi^R$. Proposition 5 below guarantees the existence of such a state with $F(\psi, \Theta) \geq 1 - \epsilon$, hence $\frac{1}{2}\|\psi - \Theta\|_1 \leq \delta$, once more invoking Eq. (3). As before we choose an optimal decomposition of $\sigma^{AB} \otimes \Omega^{A'B'}$ into states on $AA' : BB'$, which we can represent by a POVM and associated cptp map $\mathcal{M} : RR' \longrightarrow X$:

$$
\widetilde{\sigma} := (\mathrm{id}_{AA'BB'} \otimes \mathcal{M})(\psi \otimes \upsilon)
$$

$$
= \sum_x p_x \sigma_x^{AA'BB'} \otimes |x\rangle\langle x|^X,
$$

$$
E_F(\sigma \otimes \Omega) = S(AA'|X)_{\widetilde{\sigma}} = \sum_x p_x S(\mathrm{Tr}_{BB'} \sigma_x).
$$

Applying the same map to $\omega \otimes \upsilon$, we get

$$
\widetilde{\rho} := (\mathrm{id}_{AA'BB'} \otimes \mathcal{M})(\Theta \otimes \upsilon)
$$

$$
= \sum_x p_x \rho_x^{AA'BB'} \otimes |x\rangle\langle x|^X,
$$

$$
E_F(\sigma \otimes \Omega) \leq S(AA'|X)_{\widetilde{\rho}} = \sum_x p_x S(\mathrm{Tr}_{BB'} \rho_x),
$$

where we observe that, crucially, the same p_x appear in the expressions for $\widetilde{\rho}$ and $\widetilde{\sigma}$. Using $\Theta^R = \sigma^T = \psi^R$, we even have

$$
\widetilde{\rho}^{A'B'X} = (\mathrm{id}_{A'B'} \otimes \mathcal{M})(\sigma^T \otimes \upsilon) = \widetilde{\sigma}^{A'B'X}.
$$

Thus with Lemma 2, as desired,

$$
\begin{aligned}
E_F(\rho \otimes \Omega) - E_F(\sigma \otimes \Omega) &\leq S(AA'|X)_{\widetilde{\rho}} - S(AA'|X)_{\widetilde{\sigma}} \\
&= S(A|A'X)_{\widetilde{\rho}} - S(A|A'X)_{\widetilde{\sigma}} \\
&\leq 2\delta \log d + (1+\delta) h\left(\frac{\delta}{1+\delta}\right),
\end{aligned}
$$

where in the second line we have used the chain rule $S(AA'|X) = S(A'|X) + S(A|A'X)$, as well as $S(A'|X)_{\widetilde{\rho}} = S(A'|X)_{\widetilde{\sigma}}$. $\quad\square$

Proposition 5 (Quantum coupling). *Given states ρ and σ on a Hilbert space A, with $\frac{1}{2}\|\rho - \sigma\|_1 \leq \epsilon$, there exist purifications $|\varphi\rangle$ of ρ and $|\psi\rangle$ of σ, and a (sub-normalized) vector $|\vartheta\rangle$, all three in the tensor square Hilbert space $A \otimes A =: A_1 A_2$, such that*

$$
\rho^T = \varphi^{A_2}, \quad \rho = \varphi^{A_1} \geq \mathrm{Tr}_{A_2} |\vartheta\rangle\langle\vartheta|,
$$

$$
\sigma = \psi^{A_1}, \quad \sigma^T = \psi^{A_2} \geq \mathrm{Tr}_{A_1} |\vartheta\rangle\langle\vartheta|,
$$

and

$$
|\langle\psi|\vartheta\rangle|, \; |\langle\varphi|\vartheta\rangle| \geq 1 - \epsilon.
$$

Here, $\cdot^T$ denotes the transpose of a matrix with respect to a chosen basis.

Consequently, there exists a state $\Theta^{A_1 A_2}$ with the properties $\Theta^{A_1} = \rho$ and $\Theta^{A_2} = \psi^{A_2} = \sigma^T$, and such that $F(\psi, \Theta)$, $F(\varphi, \Theta) \geq 1 - \epsilon$.

This proposition can be viewed as a quantum analogue of the coupling of random variables $X \sim p$ and $Y \sim q$ such that $\Pr\{X \neq Y\} = \frac{1}{2}\|p - q\|_1$, on which the proof of Lemma 1 relied.

Proof. Fixing an orthonormal basis $\{|i\rangle\}$ of A, and introducing the unnormalized maximally entangled vector

$$
|\Phi\rangle = \sum_i |i\rangle^{A_1} |i\rangle^{A_2},
$$

we have the following two "pretty good purifications" [54] of ρ and σ:

$$
|\varphi\rangle := (\sqrt{\rho} \otimes \mathbb{1})|\Phi\rangle = \left(\mathbb{1} \otimes \sqrt{\rho}^T\right)|\Phi\rangle,
$$

$$
|\psi\rangle := (\sqrt{\sigma} \otimes \mathbb{1})|\Phi\rangle = \left(\mathbb{1} \otimes \sqrt{\sigma}^T\right)|\Phi\rangle,
$$

the claimed properties of which can be readily checked.

To obtain $|\vartheta\rangle$, we use once more Eq. (2) from the proof of Lemma 2:

$$
\omega = \frac{1}{1+\epsilon}\sigma + \frac{\epsilon}{1+\epsilon}\Delta = \frac{1}{1+\epsilon}\rho + \frac{\epsilon}{1+\epsilon}\Delta',
$$

with states Δ and Δ'. Then define

$$
\begin{aligned}
|\vartheta\rangle &:= \left(\frac{1}{\sqrt{1+\epsilon}}\rho^{1/2}\omega^{-1/2}\sigma^{1/2} \otimes \mathbb{1}\right)|\Phi\rangle \\
&= \left(X \otimes \sqrt{\sigma}^T\right)|\Phi\rangle = (X \otimes \mathbb{1})|\psi\rangle \\
&= \left(\sqrt{\rho} \otimes Y\right)|\Phi\rangle = (\mathbb{1} \otimes Y)|\varphi\rangle,
\end{aligned}
$$

using $(Z \otimes \mathbb{1})|\Phi\rangle = (\mathbb{1} \otimes Z^T)|\Phi\rangle$, with

$$X = \frac{1}{\sqrt{1+\epsilon}} \rho^{1/2} \omega^{-1/2},$$

$$Y = \frac{1}{\sqrt{1+\epsilon}} (\sigma^T)^{1/2} (\omega^T)^{-1/2}.$$

We claim that $\|X\|,\ \|Y\| \leq 1$. Indeed, $\omega \geq \frac{1}{1+\epsilon}\rho$, so

$$XX^{\dagger} = \frac{1}{1+\epsilon} \sqrt{\rho}\,\omega^{-1}\sqrt{\rho}$$

$$\leq \frac{1}{1+\epsilon} \sqrt{\rho}\left[(1+\epsilon)\rho^{-1}\right]\sqrt{\rho} = \mathbb{1},$$

and similarly $YY^{\dagger} \leq \mathbb{1}$. From this it follows that

$$\vartheta^{A_2} = \mathrm{Tr}_{A_1}(X^{\dagger}X \otimes \mathbb{1})\psi \leq \psi^{A_2} = \sigma^T, \text{ and}$$

$$\vartheta^{A_1} = \mathrm{Tr}_{A_2}(\mathbb{1} \otimes Y^{\dagger}Y)\varphi \leq \varphi^{A_1} = \rho.$$

It remains to bound the inner product $|\langle\psi|\vartheta\rangle|$ (the other one, $|\langle\varphi|\vartheta\rangle|$, is completely analogous):

$$
\begin{aligned}
|\langle\psi|\vartheta\rangle| &= \frac{1}{\sqrt{1+\epsilon}} \left|\langle\Phi|\left(\rho^{1/2}\omega^{-1/2}\sigma^{1/2} \otimes \sqrt{\sigma}^T\right)|\Phi\rangle\right| \\
&= \frac{1}{\sqrt{1+\epsilon}} \left|\mathrm{Tr}\,\sqrt{\rho}\,\omega^{-1/2}\sigma\right| \\
&= \frac{1}{\sqrt{1+\epsilon}} \left|\mathrm{Tr}\,\sqrt{\rho}\,\omega^{-1/2}[(1+\epsilon)\omega - \epsilon\Delta]\right| \\
&= \frac{1}{\sqrt{1+\epsilon}} \left|(1+\epsilon)\,\mathrm{Tr}\,\sqrt{\rho}\sqrt{\omega} - \epsilon\,\mathrm{Tr}\,\sqrt{\rho}\,\omega^{-1/2}\Delta\right| \\
&\geq \mathrm{Tr}\,\sqrt{\rho}\sqrt{(1+\epsilon)\omega} - \epsilon\left|\mathrm{Tr}\,X\Delta\right| \\
&\geq \mathrm{Tr}\,\sqrt{\rho}\sqrt{\rho} - \epsilon\|X\|\,\|\Delta\|_1 \geq 1 - \epsilon,
\end{aligned}
$$

where we have first used the definitions of $|\psi\rangle$, $|\vartheta\rangle$ and $|\Phi\rangle$, and then the identity between ω and σ; the fifth line is by triangle inequality, in the sixth we used $(1+\epsilon)\omega \geq \rho$ once more, the operator monotonicity of the square root, and the Hölder inequality $|\mathrm{Tr}\,X\Delta| \leq \|X\|\,\|\Delta\|_1$; in the last step we use the fact that both ρ and Δ are states and $\|X\| \leq 1$.

Finally, to obtain Θ, we write

$$\rho = |\vartheta\rangle\langle\vartheta|^{A_1} + (1 - \langle\vartheta|\vartheta\rangle)\Delta_1,$$

$$\sigma^T = |\vartheta\rangle\langle\vartheta|^{A_2} + (1 - \langle\vartheta|\vartheta\rangle)\Delta_2,$$

with bona fide states Δ_1 and Δ_2. It is straightforward to check that the definition

$$\Theta := |\vartheta\rangle\langle\vartheta| + (1 - \langle\vartheta|\vartheta\rangle)\Delta_1 \otimes \Delta_2$$

satisfies all requirements on Θ. $\square$

Remark 6. Although the above proof refers to the unnormalized vector $|\Phi\rangle$, and thus taken literally only makes sense for finite dimensional Hilbert spaces, the proposition remains true also in the infinite dimensional (separable) case. This can be seen either by finite dimensional approximation, or by considering $|\Phi\rangle$ as a formal device to mediate between normalized entangled vectors ($|\varphi\rangle$, $|\psi\rangle$, $|\vartheta\rangle$, etc) and Hilbert-Schmidt class operators ($\sqrt{\rho}$, $\sqrt{\sigma}$, $\rho^{1/2}\omega^{-1/2}\sigma^{1/2}$, etc).

3. Relative Entropy Distances

The same method employed in Lemma 2 can be used to derive asymptotic continuity bounds for the relative entropy distance with respect to any closed convex set C of states, or more generally positive semidefinite operators, on a Hilbert space A, cf. [46]),

$$D_C(\rho) = \min_{\gamma \in C} D(\rho\|\gamma). \tag{6}$$

Unlike [46], C has to contain only at least one full-rank state, so that D_C is guaranteed to be finite; in addition, C should be bounded, so that D_C is bounded from below. We recover the conditional entropy $S(A|B)_\rho$ for a bipartite state ρ on $A \otimes B$, as $D_C(\rho)$ with

$$C = \left\{ \mathbb{1}^A \otimes \sigma^B : \sigma \text{ a state on } B \right\}.$$

Lemma 7. *For a closed, convex and bounded set C of positive semidefinite operators, containing at least one of full rank, let*

$$\kappa := \sup_{\tau, \tau'} D_C(\tau) - D_C(\tau')$$

be the largest variation of D_C. Then, for any two states ρ and σ with $\frac{1}{2}\|\rho - \sigma\|_1 \leq \epsilon$,

$$|D_C(\rho) - D_C(\sigma)| \leq \epsilon \kappa + (1 + \epsilon) h\left(\frac{\epsilon}{1 + \epsilon}\right). \tag{7}$$

Proof. The only modification with respect to the proof of Lemma 2 is that we replace the invocation of concavity of the conditional entropy with the joint convexity of the relative entropy, which makes D_C a convex functional.

Namely, with ω as in Eq. (2), we have on the one hand,

$$D_C(\omega) \leq \frac{1}{1 + \epsilon} D_C(\sigma) + \frac{\epsilon}{1 + \epsilon} D_C(\Delta).$$

On the other hand, with an optimal $\gamma \in C$,

$$\begin{aligned}
D_C(\omega) &= D\left(\omega\|\gamma\right) \\
&= -S(\omega) - \operatorname{Tr} \omega \log \gamma \\
&\geq -h\left(\frac{\epsilon}{1 + \epsilon}\right) - \frac{1}{1 + \epsilon} S(\rho) - \frac{\epsilon}{1 + \epsilon} S(\Delta') \\
&\quad - \frac{1}{1 + \epsilon} \operatorname{Tr} \rho \log \gamma - \frac{\epsilon}{1 + \epsilon} \operatorname{Tr} \Delta' \log \gamma \\
&= -h\left(\frac{\epsilon}{1 + \epsilon}\right) + \frac{1}{1 + \epsilon} D(\rho\|\gamma) + \frac{\epsilon}{1 + \epsilon} D(\Delta'\|\gamma) \\
&\geq -h\left(\frac{\epsilon}{1 + \epsilon}\right) + \frac{1}{1 + \epsilon} D_C(\rho) + \frac{\epsilon}{1 + \epsilon} D_C(\Delta').
\end{aligned}$$

Putting these two inequalities together yields the claim of the lemma. □

In particular, in the case that

$$C = \text{SEP}(A : B)$$
$$:= \text{conv}\{\alpha^A \otimes \beta^B : \alpha, \ \beta \text{ states on } A, \ B, \ \text{resp.}\}$$

is the set of separable states, we obtain the relative entropy of entanglement of a state ρ on bipartite system $A \otimes B$, $E_R(\rho) = D_{\text{SEP}(A:B)}(\rho)$ [48]. Furthermore, we consider its regularization

$$E_R^\infty(\rho) = \lim_{n \to \infty} \frac{1}{n} E_R(\rho^{\otimes n}),$$

which is known to be different from $E_R(\rho)$ in general [49].

Corollary 8 (Cf. Donald/Horodecki [11] and Christandl [6]). *For any two states ρ and σ on the composite system $A \otimes B$, denoting the smaller of the dimensions $|A|$, $|B|$ by d, $\frac{1}{2}\|\rho - \sigma\|_1 \leq \epsilon$ implies*

$$|E_R(\rho) - E_R(\sigma)| \leq \epsilon \log d + (1 + \epsilon) h\left(\frac{\epsilon}{1 + \epsilon}\right),$$
$$|E_R^\infty(\rho) - E_R^\infty(\sigma)| \leq \epsilon \log d + (1 + \epsilon) h\left(\frac{\epsilon}{1 + \epsilon}\right).$$

Note that this bound only depends on the smaller of the two dimensions, in contrast to [11]; in particular, it applies even in the case that one of the two Hilbert spaces is infinite dimensional.

Proof. The first bound, on the single-letter E_R is a direct application of Lemma 7 to the case where C is the set of all separable states on $A \otimes B$.

For the regularization, consider any integer n and

$$\left| E_R\left(\rho^{\otimes n}\right) - E_R\left(\sigma^{\otimes n}\right) \right|$$
$$= \left| \sum_{t=1}^{n} E_R\left(\rho^{\otimes t} \otimes \sigma^{\otimes n-t}\right) - E_R\left(\rho^{\otimes t-1} \otimes \sigma^{\otimes n-t+1}\right) \right|$$
$$\leq \sum_{t=1}^{n} |E_R(\rho \otimes \Omega_t) - E_R(\sigma \otimes \Omega_t)|,$$

with $\Omega_t = \rho^{\otimes t-1} \otimes \sigma^{\otimes n-t}$. Now for each t, Lemma 7 gives

$$|E_R(\rho \otimes \Omega_t) - E_R(\sigma \otimes \Omega_t)| \leq \epsilon \kappa_t + (1 + \epsilon) h\left(\frac{\epsilon}{1 + \epsilon}\right),$$

with $\kappa_t = \sup_{\tau, \tau'}\left(E_R(\tau \otimes \Omega_t) - E_R(\tau' \otimes \Omega_t)\right)$. To see this, we have to look into the proof of the lemma, and observe that for states $\rho \otimes \Omega_t$ and $\sigma \otimes \Omega_t$, also the auxiliary operators Δ and Δ' are of the form $\tau \otimes \Omega_t$ and $\tau' \otimes \Omega_t$. However, by LOCC monotonicity,

$$E_R(\tau' \otimes \Omega_t) \geq E_R(\Omega_t),$$

and similarly

$$E_R(\tau \otimes \Omega_t) \leq E_R(\Phi_d \otimes \Omega_t) \leq \log d + E_R(\Omega_t),$$

so that $\kappa_t \leq \log d$. Although we do not need it, the right hand inequality is in fact an equality, $E_R(\Phi_d \otimes \Omega_t) = \log d + E_R(\Omega_t)$ [33]. Thus, we obtain for all n,

$$\left| \frac{1}{n} E_R(\rho^{\otimes n}) - \frac{1}{n} E_R(\sigma^{\otimes n}) \right| \leq \epsilon \log d + (1 + \epsilon) h\left(\frac{\epsilon}{1 + \epsilon}\right),$$

and taking the limit $n \to \infty$ concludes the proof. □

Again, in Lemma 7 and Corollary 8, the constant in the linear term (proportional to ϵ) is essentially best possible, as we see by taking two states maximizing the difference $D_C(\rho) - D_C(\sigma)$, i.e. attaining κ, since $\frac{1}{2}\|\rho - \sigma\|_1 \leq 1 =: \epsilon$.

Remark 9. Lemma 7 improves upon similar-looking general bounds by Synak-Radtke and Horodecki [46], which were subsequently optimized by Mosonyi and Hiai [28, Prop. VI.1]. The latter paper also explains lucidly (in Sec. VI) that the coefficient $\frac{1}{1+\epsilon}$ in the convex decomposition of ω in two ways, into ρ and Δ' and into σ and Δ, is optimal, and gives a nice geometric interpretation of ω as a max-relative entropy center of ρ and σ (cf. [24]). Thus, at least following the same strategy one cannot improve the bound any more.

That the regularized relative entropy measure E_R^∞ is asymptotically continuous followed previously from its non-lockability [21], which it inherits from E_R. This has been worked out in [8, Prop. 13], following [6, Prop. 3.23], with a different linear term.

Remark 10. It would be interesting to lift the restriction that C has to be a convex set: Natural examples are the case that C is the set of all product states in a bipartite (multipartite) system, in which case D_C becomes the quantum mutual information (multiinformation); or the case that C is the closure of the set of all Gibbs states for a suitable Hamiltonian operator H,

$$C = \overline{\left\{ \frac{1}{\text{Tr } e^{-\beta H}} e^{-\beta H} : \beta > 0 \right\}}.$$

Both examples have in common that C is an exponential family (or the closure of one); it is known that at least in some cases D_C is continuous, but counterexamples of discontinuous behaviour are known [52].

4. Bounded Energy

If the Hilbert space in the Fannes inequality (Lemma 1) has infinite dimension, or likewise A in the Alicki–Fannes inequality (Lemma 2), then the bound becomes trivial: the right hand side is infinite. This is completely natural, since the entropy is not even continuous, and these Fannes-type bounds imply a sort of uniform continuity. Continuity is restored, however, when restricting to states of finite energy, for instance of a quantum harmonic oscillator [51], see also [12,39] for more recent results and excellent surveys on the status of continuity of the entropy. Shirokov [40] has developed an approach to prove (local) continuity of entropic quantities, based on certain finite entropy assumptions, in which he uses Alicki–Fannes inequalities on finite approximations.

Uniform bounds are still out of the question, but what we shall show here is that the Fannes and Alicki–Fannes inequalities discussed above have satisfying analogues, with a dependence on the energy of the states rather than the Hilbert space dimension.

Abstractly, our setting is this: Consider a Hamiltonian H on a infinite dimensional separable Hilbert space A. If there is another system B and we consider bipartite states and conditional entropy, we implicitly assume trivial Hamiltonian on B, i.e., global Hamiltonian $H = H^A \otimes \mathbb{1}^B$. We shall need a number of assumptions on H, to start with that it has discrete spectrum and that it is bounded from below; for normalization purposes we fix the ground state energy of H to be 0. The mathematically precise assumption is the following.

Gibbs Hypothesis. For every $\beta > 0$, let the partition function $Z(\beta) := \operatorname{Tr} e^{-\beta H} < \infty$ be finite, so that $\frac{1}{Z(\beta)} e^{-\beta H}$ is a bona fide state with finite entropy. In this case, for every energy E in the spectrum of H, the (unique) maximizer of the entropy $S(\rho)$ subject to $\operatorname{Tr} \rho H \leq E$ is of the Gibbs form:

$$\gamma(E) = \frac{1}{Z(\beta(E))} e^{-\beta(E)H},$$

where $\beta = \beta(E)$ is decreasing with E and is the solution to the equation

$$\operatorname{Tr} e^{-\beta H}(H - E) = 0.$$

The entropy in this case is given by

$$S(\gamma(E)) = \log Z + \beta(E)(\log e)E.$$

This implies that the spectrum is unbounded above, and that the energy levels cannot become "too dense" with growing energy value.

Let us immediately draw some conclusions from these assumptions; the following is a simply consequence of Shirokov's [38, Prop. 1], for which we present an elementary proof.

Proposition 11. *For a Hamiltonian H satisfying the Gibbs Hypothesis, $S(\gamma(E))$ is a strictly increasing, strictly concave function of the energy E.*

Proof. It is clear from the maximum entropy characterization of $\gamma(E)$ that the entropy as a function of E must be non-decreasing; it is unbounded by looking at the formula for the entropy in terms of $\log Z$.

Furthermore, for energies E_1 and E_2, and $0 \leq p \leq 1$,

$$\operatorname{Tr}(p\gamma(E_1) + (1 - p)\gamma(E_2))H \leq pE_1 + (1 - p)E_2 =: E,$$

and so concavity follows:

$$\begin{aligned} S(\gamma(E)) &\geq S(p\gamma(E_1) + (1 - p)\gamma(E_2)) \\ &\geq pS(\gamma(E_1)) + (1 - p)S(\gamma(E_2)). \end{aligned} \tag{8}$$

From this it follows that $S(\gamma(E))$ is strictly increasing, because otherwise $S(\gamma(E_1)) = S(\gamma(E_2))$ for some $E_1 < E_2$, but then $S(\gamma(E_2)) < S(\gamma(E_3))$ for some $E_2 < E_3$, since the entropy grows to infinity as $E \to \infty$, contradicting concavity.

But this means that for $E_1 \neq E_2$, necessarily $\gamma(E_1) \neq \gamma(E_2)$, and so by the strict concavity of the von Neumann entropy, we have strict inequality in the second line of Eq. (8) for $0 < p < 1$. □

Corollary 12. *If H satisfies the Gibbs Hypothesis, then for any* $\delta > 0$,

$$\sup_{0 < \lambda \leq \delta} \lambda \, S\big(\gamma(E/\lambda)\big) = \delta \, S\big(\gamma(E/\delta)\big).$$

Proof. The right hand side is clearly attained by letting $\lambda = \delta$. To prove "$\leq$" for any admissible λ, observe that by concavity (Proposition 11),

$$S\big(\gamma(tF)\big) \geq t \, S\big(\gamma(F)\big) + (1-t) \, S\big(\gamma(0)\big) \geq t \, S\big(\gamma(F)\big).$$

Letting $t = \frac{\lambda}{\delta} \leq 1$ and $F = \frac{E}{\lambda}$ concludes the proof.   □

Remark 13. Another useful fact proved by Shirokov [38, Prop. 1(ii)], which we shall invoke later, is that under our assumptions, $S\big(\gamma(E)\big) = o(E)$, which can be recast as saying that $\delta \, S\big(\gamma(E/\delta)\big) \to 0$ for every finite E and $\delta \to 0$.   □

We start with an easy-to-prove continuity bound for the entropy, inspired by the proof of Lemma 1, though for the conditional entropy we shall have to resort to a different argument. It uses a quantum coupling as in Proposition 5 (which implies a weaker bound in the following, with the square of the expression on the right hand side).

Proposition 14. *Let ρ and σ be states on the same Hilbert space A, and consider the tensor square $A \otimes A =: A_1 A_2$ of the quantum system. Then, there exists a state ω with $\omega^{A_1} = \rho$, $\omega^{A_2} = \sigma$ and such that*

$$\|\omega\|_\infty \geq 1 - \frac{1}{2}\|\rho - \sigma\|_1.$$

Proof. Choose spectral decompositions

$$\rho = \sum_i r_i |e_i\rangle\langle e_i|,$$

$$\sigma = \sum_i s_i |f_i\rangle\langle f_i|,$$

of the two states, with $r_1 \geq r_2 \geq \cdots$ and $s_1 \geq s_2 \geq \cdots$; then, the ℓ^1-distance between the probability vectors (r_i) and (s_i) is not larger than the trace distance between ρ and σ:

$$\|\rho - \sigma\|_1 \geq \|(r_i) - (s_i)\|_1 =: 2\epsilon.$$

(This is known as Mirksy's inequality [20, Cor. 7.4.9.3].)

Defining a vector

$$|\phi\rangle := \sum_i \sqrt{\min\{r_i, s_i\}} |e_i\rangle^{A_1} |f_i\rangle^{A_2}$$

in $A_1 A_2$, we clearly have $\mathrm{Tr}\, |\phi\rangle\langle\phi| = 1 - \epsilon$, and $\phi^{A_1} \leq \rho$, $\phi^{A_2} \leq \sigma$, thus we can write

$$\rho = |\phi\rangle\langle\phi|^{A_1} + \epsilon\Delta_1, \quad \sigma = |\phi\rangle\langle\phi|^{A_2} + \epsilon\Delta_2,$$

with bona fide states Δ_1 and Δ_2.

It is straightforward to check that the definition $\omega := |\phi\rangle\langle\phi| + \epsilon\Delta_1 \otimes \Delta_2$ satisfies all requirements on ω.   □

Lemma 15. *Let the Hamiltonian H on A be satisfying the Gibbs Hypothesis. Then for any two states ρ and σ on A with $\operatorname{Tr}\rho H$, $\operatorname{Tr}\sigma H \leq E$ and $\frac{1}{2}\|\rho - \sigma\|_1 \leq \epsilon \leq 1$,*

$$\left|S(\rho) - S(\sigma)\right| \leq 2\epsilon S\big(\gamma(E/\epsilon)\big) + h(\epsilon).$$

Proof. Pick a state ω on $A_1 A_2$, according to Proposition 14: $\omega^{A_1} = \rho$, $\omega^{A_2} = \sigma$, and with largest eigenvalue $\geq 1 - \epsilon$, meaning that we can write

$$\omega = (1 - \epsilon)|\psi\rangle\langle\psi| + \epsilon\omega',$$

with a pure state $|\psi\rangle$ (the normalized vector $|\phi\rangle$ from the proof of Proposition 14) and some other state ω'. Hence,

$$\begin{aligned}
\left|S(\rho) - S(\sigma)\right| &= \left|S(\omega^{A_1}) - S(\omega^{A_2})\right| \\
&\leq S(\omega^{A_1 A_2}) \\
&\leq \epsilon S(\omega') + h(\epsilon) \\
&\leq 2\epsilon S\big(\gamma(E/\epsilon)\big) + h(\epsilon).
\end{aligned}$$

Here, we have first used the marginals of ω, then in the second line the Araki–Lieb "triangle" inequality [2], in the third line strong subadditivity, and in the last step the maximum entropy principle, noting that with respect to the Hamiltonian $H^{A_1} \otimes \mathbb{1}^{A_2} + \mathbb{1}^{A_1} \otimes H^{A_2}$, ω has energy $\leq 2E$, and so the energy of ω' is bounded by $2E/\epsilon$. For the last line, observe that the Gibbs state at energy $2E/\epsilon$ of the composite system is $\gamma(E/\epsilon)^{\otimes 2}$. □

The following two general bounds lack perhaps the simple elegance of Lemma 15, but they turn out to be more flexible, and stronger in certain regimes.

Meta-Lemma 16 (Entropy). *For a Hamiltonian H on A satisfying the Gibbs Hypothesis and any two states ρ and σ with $\operatorname{Tr}\rho H$, $\operatorname{Tr}\sigma H \leq E$, $\frac{1}{2}\|\rho - \sigma\|_1 \leq \epsilon < \epsilon' \leq 1$, and $\delta = \frac{\epsilon' - \epsilon}{1 + \epsilon'}$,*

$$\left|S(\rho) - S(\sigma)\right| \leq (\epsilon' + 2\delta)S\big(\gamma(E/\delta)\big) + h(\epsilon') + h(\delta).$$

Meta-Lemma 17 (Conditional entropy). *For states ρ and σ on the bipartite system $A \otimes B$ and otherwise the same assumption as before,*

$$\left|S(A|B)_\rho - S(A|B)_\sigma\right| \leq (2\epsilon' + 4\delta)S\big(\gamma(E/\delta)\big) + (1 + \epsilon')h\left(\frac{\epsilon'}{1 + \epsilon'}\right) + 2h(\delta).$$

To interpret these bounds, we remark that in a certain sense they show that the Gibbs entropy at the cutoff energy E/ϵ (E/δ) takes on the role of the logarithm of the dimension in the finite dimensional case. Before we launch into their proof, let us introduce some notation: define the energy cutoff projectors

$$P_\leq := \sum_{0 \leq E_n \leq E/\delta} |n\rangle\langle n|, \quad P_> := \mathbb{1} - P_\leq,$$

where $|n\rangle$ is the eigenvector of eigenvalue E_n of the Hamiltonian H. We shall also consider the pinching map

$$\mathcal{T}(\xi) = P_\le \xi P_\le + P_> \xi P_>,$$

which is a unital channel, as well as its action on the original ρ and σ:

$$\mathcal{T}(\rho) =: (1 - \lambda)\rho_\le + \lambda\rho_>,$$
$$\mathcal{T}(\sigma) =: (1 - \mu)\sigma_\le + \mu\sigma_>.$$

Note that because H commutes with the action of $\mathcal{T}$, we have $\mathrm{Tr}\,\xi H = \mathrm{Tr}\,\mathcal{T}(\xi)H$, and so the energy bound E applies also to $\mathcal{T}(\rho)$ and $\mathcal{T}(\sigma)$. Hence,

$$\lambda \le \delta, \ \lambda\,\mathrm{Tr}\,\rho_> H \le E, \quad \mu \le \delta, \ \mu\,\mathrm{Tr}\,\sigma_> H \le E. \tag{9}$$

Our strategy will be to relate $S(\rho)$ to $S(\rho_\le)$ (and the same for σ and $\sigma_\le$) via entropy inequalities, including concavity, similar to the first part of the paper, and then apply the usual Fannes (Alicki–Fannes) inequalities to $\rho_\le$ and $\sigma_\le$.

Proof of Lemma 16. First of all, by concavity of the entropy (monotonicity under unital cptp maps),

$$S(\rho) \le S(\mathcal{T}(\rho)) = h(\lambda) + (1 - \lambda)S(\rho_\le) + \lambda S(\rho_>). \tag{10}$$

Now, by Eq. (9), the maximum entropy principle and Corollary 12,

$$\lambda S(\rho_>) \le \lambda S\big(\gamma(E/\lambda)\big) \le \delta S\big(\gamma(E/\delta)\big).$$

Thus, from Eq. (10), observing $\delta \le \frac{1}{2}$, we get

$$S(\rho) \le S(\rho_\le) + h(\delta) + \delta S\big(\gamma(E/\delta)\big), \tag{11}$$

and likewise for σ.

Second, we have

$$S(\sigma) \ge (1 - \mu)S(\sigma_\le) + \mu S(\sigma_>). \tag{12}$$

To see this, we think of the action of $\mathcal{T}$ as a binary measurement on the system A, which we can implement coherently with two ancilla qubits X and X',

$$|\varphi\rangle \longmapsto (P_\le|\varphi\rangle)^A|00\rangle^{XX'} + (P_>|\varphi\rangle)^A|11\rangle^{XX'}.$$

Applying this to σ, we have by unitary invariance and the Araki-Lieb "triangle" inequality,

$$\begin{aligned}
S(\sigma) = S(AXX') &\ge S(AX) - S(X') \\
&= S(AX) - S(X) \\
&= S(A|X) = (1 - \mu)S(\sigma_\le) + \mu S(\sigma_>).
\end{aligned}$$

Thus, using that the energy of $\sigma_\le$ is at most E/δ by construction, and so $S(\sigma_\le) \le S\big(\gamma(E/\delta)\big)$,

$$S(\sigma) \ge (1 - \mu)S(\sigma_\le) \ge S(\sigma_\le) - \delta S\big(\gamma(E/\delta)\big). \tag{13}$$

Third, by definitions, contractivity of the trace norm and triangle inequality,

$$\begin{aligned}
2\epsilon &\geq \|\rho - \sigma\|_1 \\
&\geq \left\| P_\leq \rho P_\leq - P_\leq \sigma P_\leq \right\|_1 \\
&= \left\| (1-\lambda)\rho_\leq - (1-\mu)\sigma_\leq \right\|_1 \\
&= \left\| (1-\delta)(\rho_\leq - \sigma_\leq) + (\delta - \lambda)\rho_\leq + (\mu - \delta)\sigma_\leq \right\|_1 \\
&\geq (1-\delta)\|\rho_\leq - \sigma_\leq\|_1 - 2\delta,
\end{aligned}$$

and so

$$\frac{1}{2}\|\rho_\leq - \sigma_\leq\|_1 \leq \frac{\epsilon + \delta}{1-\delta} = \epsilon'. \tag{14}$$

Hence by the Fannes inequality in the form of Lemma 1,

$$\begin{aligned}
|S(\rho_\leq) - S(\sigma_\leq)| &\leq \epsilon' \log \mathrm{Tr}\, P_\leq + h(\epsilon') \\
&\leq \epsilon' S\big(\gamma(E/\delta)\big) + h(\epsilon').
\end{aligned} \tag{15}$$

The latter inequality holds because the state $\frac{1}{\mathrm{Tr}\, P_\leq} P_\leq$ clearly has energy bounded by E/δ, and so cannot have entropy larger than the Gibbs state.

With these three elements we can conclude the proof: W.l.o.g. $S(\rho) \geq S(\sigma)$, and so from Eqs. (11), (13) and (15),

$$\begin{aligned}
S(\rho) - S(\sigma) &\leq S(\rho_\leq) - S(\sigma_\leq) + h(\delta) + 2\delta S\big(\gamma(E/\delta)\big) \\
&\leq (\epsilon' + 2\delta)S\big(\gamma(E/\delta)\big) + h(\epsilon') + h(\delta),
\end{aligned}$$

as advertised. □

Proof of Lemma 17. It is very similar to the previous one, only that we have to be a bit more careful in some details, as the conditional entropy can be negative.

The first step goes through almost unchanged, with the map $\mathcal{T} \otimes \mathrm{id}_B$, since the conditional entropy is concave as well (equivalent to strong subadditivity) [27]:

$$\begin{aligned}
S(A|B)_\rho &\leq S(A|B)_{\mathcal{T}(\rho)} \\
&= h(\lambda) + (1-\lambda)S(A|B)_{\rho_\leq} + \lambda S(A|B)_{\rho_>}.
\end{aligned}$$

The remainder term $\lambda S(A|B)_{\rho_>}$ is upper bounded by $\lambda S(\rho_>^A)$ (again by strong subadditivity), hence the upper bound $\lambda S\big(\gamma(E/\lambda)\big)$ still applies. The only change is due to the fact that the conditional entropy can be negative. However, for any bipartite state ξ^{AB},

$$-S(\xi^A) \leq S(A|B)_\xi \leq S(\xi^A).$$

Here, the right hand inequality is strong subadditivity that we have used before; introducing a purification $|\varphi\rangle^{ABC}$ of the state, we have $-S(A|B)_\varphi = S(A|C)_\varphi \leq S(\xi^A)$, which is the left hand inequality. Thus,

$$(1-\lambda)S(A|B)_{\rho_\leq} \leq S(A|B)_{\rho_\leq} + \delta S\big(\gamma(E/\delta)\big).$$

Altogether,

$$S(A|B)_\rho \leq S(A|B)_{\rho_\leq} + 2\delta S\big(\gamma(E/\delta)\big) + h(\delta). \tag{16}$$

Also the second step requires only minor modifications: With the notation of the previous proof, and using the Araki-Lieb "triangle" inequality once again,

$$
\begin{aligned}
S(AXX'|B) &= S(AXX'B) - S(B) \\
&\geq S(ABX) - S(X') - S(B) \\
&= S(ABX) - S(BX) - S(X) - S(B) + S(XB) \\
&= S(A|BX) - I(X:B) \\
&\geq S(A|BX) - h(\delta).
\end{aligned}
$$

Again, since conditional entropies can be negative, we have to be more careful with remainder terms and get

$$
S(A|B)_\sigma \geq S(A|B)_{\sigma_\leq} - 2\delta S\big(\gamma(E/\delta)\big) - h(\delta). \tag{17}
$$

In the third step, the trace norm estimate (14) goes through unchanged, and then we apply the Alicki–Fannes inequality in the form of Lemma 2:

$$
\begin{aligned}
\Big| S(A|B)_{\rho_\leq} &- S(A|B)_{\sigma_\leq} \Big| \\
&\leq 2\epsilon' \log \operatorname{Tr} P_\leq + (1+\epsilon') h\left(\frac{\epsilon'}{1+\epsilon'}\right) \\
&\leq 2\epsilon' S\big(\gamma(E/\delta)\big) + (1+\epsilon') h\left(\frac{\epsilon'}{1+\epsilon'}\right).
\end{aligned}
$$

Putting this together with Eqs. (16) and (17), assuming w.l.o.g. that $S(A|B)_\rho \geq S(A|B)_\sigma$, we obtain

$$
\begin{aligned}
S(A|B)_\rho - S(A|B)_\sigma &\leq S(A|B)_{\rho_\leq} - S(A|B)_{\sigma_\leq} \\
&\quad + 2h(\delta) + 4\delta S\big(\gamma(E/\delta)\big) \\
&\leq (2\epsilon' + 4\delta) S\big(\gamma(E/\delta)\big) \\
&\quad + (1+\epsilon') h\left(\frac{\epsilon'}{1+\epsilon'}\right) + 2h(\delta),
\end{aligned}
$$

and we are done. □

The bounds of Lemmas 15, 16 and 17 are very general, and it may not be immediately apparent how useful they are. However, thanks to [38, Prop. 1(ii)], restated in Remark 13, $\delta S\big(\gamma(E/\delta)\big) \to 0$ for every finite E, as $\delta \to 0$ (cf. [41, Cor. 4]). Thus, choosing $\epsilon' = \sqrt{\epsilon}$, the lemmas do prove continuity of the entropy and conditional entropy in general, and uniformly for each fixed energy.

We now specialize our bounds to the important case of a collection of ℓ quantum harmonic oscillators, where we shall see that the bounds are asymptotically tight. The Hamiltonian is

$$
H = \sum_{i=1}^{\ell} \hbar\omega_i\, a_i^\dagger a_i, \tag{18}
$$

where ω_i is the native frequency of the i-th oscillator and a_i is its annihilation (aka lowering) operator (see e.g. [25] or [50]). Note that we chose the slightly unusual energy convention such that the ground state has energy 0, rather than $\sum_i \frac{1}{2}\hbar\omega_i$, to be able to apply directly our above results. In the case of a single mode, and choosing units such

that $\hbar\omega_1 = 1$, the Hamiltonian simply becomes the number operator N. In that case, it is well-known that

$$S(\gamma(N)) = g(N) := (N+1)\log(N+1) - N\log N$$
$$\leq \log(N+1) + \log e.$$

Crucially, and in accordance with Proposition 11, g is a concave, monotone increasing function of N.

In the general case of Eq. (18), $\gamma(E) = \bigotimes_{i=1}^{\ell} \gamma_i(E_i)$, with $E = \sum_i E_i$ and where $\gamma_i(E_i)$ is the Gibbs state of the i-th mode with energy E_i. Maximizing the entropy,

$$S\left(\bigotimes_{i=1}^{\ell} \gamma_i(E_i)\right) = \sum_{i=1}^{\ell} g\left(\frac{E_i}{\hbar\omega_i}\right),$$

over all allocations of the total energy over the ℓ modes leads to a transcendental equation, but we do not need to solve it as we only want an upper bound, via $g(N) \leq \log(N+1) + \log e$. By a straightforward Lagrange multiplier calculation we see that the optimum is to divide the energy equally among the modes:

$$S(\gamma(E)) \leq \max \sum_{i=1}^{\ell} \left[\log\left(\frac{E_i}{\hbar\omega_i} + 1\right) + \log e\right]$$
$$= (\log e)\ell + \sum_{i=1}^{\ell} \log\left(\frac{\overline{E}}{\hbar\omega_i} + 1\right), \tag{19}$$

with $E =: \ell\overline{E}$.

By using this upper bound in Lemmas 16 and 17, for $\delta = \alpha\epsilon(1-\epsilon)$, with a parameter α between 0 and $\frac{1}{2}$, and introducing

$$\widetilde{h}(x) := \begin{cases} h(x) & \text{for } x \leq \frac{1}{2}, \\ 1 & \text{for } x \geq \frac{1}{2}, \end{cases}$$

we obtain directly the following:

Lemma 18. *Consider two states ρ and σ of the ℓ-oscillator system* (18)*, whose energies are bounded* $\text{Tr}\,\rho H$, $\text{Tr}\,\sigma H \leq E = \ell\overline{E}$*. Then,* $\frac{1}{2}\|\rho - \sigma\|_1 \leq \epsilon < 1$ *implies*

$$\left|S(\rho) - S(\sigma)\right| \leq \epsilon\left(\frac{1+\alpha}{1-\alpha} + 2\alpha\right)\left[\sum_{i=1}^{\ell} \log\left(\frac{\overline{E}}{\hbar\omega_i} + 1\right) + \ell\log\frac{e}{\alpha(1-\epsilon)}\right]$$
$$+ (\ell+2)\left(\frac{1+\alpha}{1-\alpha} + 2\alpha\right)\widetilde{h}\left(\frac{1+\alpha}{1-\alpha}\epsilon\right).$$

If the states live on a system composed of the ℓ oscillators (A) and another system B, then

$$\left| S(A|B)_\rho - S(A|B)_\sigma \right| \le 2\epsilon \left(\frac{1+\alpha}{1-\alpha} + 2\alpha \right) \left[\sum_{i=1}^{\ell} \log \left(\frac{\overline{E}}{\hbar\omega_i} + 1 \right) + \ell \log \frac{e}{\alpha(1-\epsilon)} \right]$$
$$+ (2\ell + 4) \left(\frac{1+\alpha}{1-\alpha} + 2\alpha \right) \widetilde{h} \left(\frac{1+\alpha}{1-\alpha} \epsilon \right).$$

<div align="right">□</div>

Remark 19. For each fixed $\epsilon \le 1$, we can make α arbitrarily small, and then for large energy $E \gg \sum_i \hbar\omega_i$, the bounds of Lemma 18 are asymptotically tight, in the sense that apart from the additive offset terms, the factor multiplying ϵ (2ϵ, resp.) cannot be smaller than

$$S(\gamma(E)) \approx \sum_{i=1}^{\ell} \log \left(\frac{\overline{E}}{\hbar\omega_i} + 1 \right).$$

This can be seen in the entropy case by comparing the vacuum state $\rho = |0\rangle\langle 0|^{\otimes\ell}$ of all ℓ modes with the state $\sigma = (1-\epsilon)|0\rangle\langle 0|^{\otimes\ell} + \epsilon\gamma(E)$; in the conditional entropy case, take ρ to be a purification of the Gibbs state $\gamma(E)$ on $A \otimes B$, and $\sigma = (1-\epsilon)\rho + \epsilon\gamma(E)^A \otimes \tau^B$ with an arbitrary state τ on B.

5. Conclusions

Using entropy inequalities, specifically concavity, we improved the appearance of the Alicki–Fannes inequality for the conditional von Neumann entropy to an almost tight form. It would be nice to know the ultimately best form among all formulas that depend only on the dimension of the Hilbert space and the trace distance, but we have to leave this as an open problem.

In particular, it would be curious to find the optimal form of the fidelity in Proposition 5,

$$\widetilde{F} := \max F(\psi, \Theta) \text{ s.t. } \Theta^{A_1} = \rho, \ \Theta^{A_2} = \psi^{A_2},$$
$$\ge 1 - \frac{1}{2}\|\rho - \sigma\|_1,$$

with a fixed purification ψ of σ, and of Proposition 14,

$$1 - \frac{1}{2}\|\rho - \sigma\|_1 \le \max \|\omega\|_\infty \text{ s.t. } \omega^{A_1} = \rho, \ \omega^{A_2} = \sigma,$$

which may be regarded as quantum state analogues of the coupling random variables,

$$\frac{1}{2}\|p - q\|_1 = \min \Pr\{X \ne Y\} \text{ s.t. } X \sim p, \ Y \sim q.$$

Furthermore, are there versions of these statements that would allow for alternative proofs or tighter versions of Lemmas 2 and 17 for the conditional entropy?

The same principle lead to the apparently first uniform continuity bounds of the entropy and conditional on infinite dimensional Hilbert spaces under a bound on the expected energy (or, for that matter, bounded expectation of any sufficiently well-behaved Hermitian operator). In the case of a system of harmonic oscillators, we have seen that

the bound is, in a certain sense, asymptotically tight, even though here we are much further away from a universally optimal form.

The Fannes and Alicki–Fannes inequalities already are known to have many applications in quantum information theory. These include the continuity of certain entanglement measures such as entanglement of formation [29], relative entropy of entanglement [11], squashed entanglement [7] and conditional entanglement of mutual information [55], and of various quantum channel capacities [26]. In fact, we always get explicit continuity bounds in terms of the trace distance of the states or diamond norm distance of the channels, respectively. While in many applications it is of minor interest to have the optimal form of the bound (for example when ϵ goes to 0), it pays off to have a tighter bound than [1] in the setting of approximately degradable channels [44]. Indeed, this results even in new, tighter upper bounds on the quantum capacity of very quiet depolarizing channels [44], by way of an extension of the methodology of [42].

The infinite dimensional versions of these entropy bounds under an energy constraint are awaiting applications, though it seems clear that explicit bounds on the continuity and asymptotic continuity of entanglement measures [12], (e.g. for squashed entanglement since the first posting of the present manuscript [41]) and channel capacities [17–19,37] in infinite dimension should be among the first, as well as the extension of approximate degradability [44] to Bosonic channels [45].

References

1. Alicki, R., Fannes, M.: Continuity of quantum conditional information. J. Phys. A Math. Gen. **37**, L55–L57 (2004)
2. Araki, H., Lieb, E.H.: Entropy inequalities. Commun. Math. Phys. **18**, 160–170 (1970)
3. Audenaert, K.M.R.: A sharp continuity estimate for the von Neumann entropy. J. Math. Phys. A Math. Theor. **40**(28), 8127–8136 (2007)
4. Bennett, C.H., DiVincenzo, D.P., Smolin, J.A., Wootters, W.K.: Mixed-state entanglement and quantum error correction. Phys. Rev. A **54**(5), 3824–3851 (1996)
5. Bogachev, V.I., Kolesnikov, A.V.: The Monge–Kantorovich problem: achievements, connections, and perspectives. Russian Math. Surv. **67**(5), 785–890 (2012)
6. Christandl, M.: The structure of bipartite quantum states — insights from group theory and cryptography. PhD Thesis, Department of Applied Mathematics and Theoretical Physics, Cambridge University (2006). arXiv:quant-ph/0604183
7. Christandl, M., Winter, A.: Squashed entanglement—an additive entanglement measure. J. Math. Phys. **45**(3), 829–840 (2003)
8. Christandl, M., Schuch, N., Winter, A.: Entanglement of the antisymmetric state. Commun. Math. Phys. **311**, 397–422 (2012)
9. Cover, T.M., Thomas, J.A.: Elements of Information Theory, 2nd ed. Wiley, New York (2006)
10. Devetak, I., Shor, P.W.: The capacity of a quantum channel for simultaneous transmission of classical and quantum information. Commun. Math. Phys. **256**, 287–303 (2005)
11. Donald, M.J., Horodecki, M.: Continuity of relative entropy of entanglement. Phys. Lett. A **264**, 257–260 (1999)
12. Eisert, J., Simon, C., Plenio, M.B.: On the quantification of entanglement in infinite-dimensional quantum systems. J. Phys. A Math. Gen. **35**(17), 3911–3923 (2002)

13. Fannes, M.: A continuity property of the entropy density for spin lattice systems. Commun. Math. Phys. **31**, 291–294 (1973)
14. Fuchs, C.A., van de Graaf, J.: Cryptographic distinguishability measures for quantum-mechanical states. IEEE Trans. Inf. Theory **45**(4), 1216–1227 (1999)
15. Hastings, M.B.: Superadditivity of communication capacity using entangled inputs. Nature Phys. **5**, 255–257 (2009)
16. Hayden, P.M., Horodecki, M., Terhal, B.M.: The asymptotic entanglement cost of preparing a quantum state. J. Phys. A Math. Gen. **34**(35), 6891–6898 (2001)
17. Holevo A.S.: On Quantum Communication Channels with Constrained Inputs. arXiv:quant-ph/9705054 (1997)
18. Holevo, A.S.: Entanglement-assisted capacities of constrained quantum channels. Theory Probab. Appl. **48**(2), 243–255 (2006)
19. Holevo, A.S., Shirokov, M.E.: Continuous ensembles and the capacity of infinite-dimensional quantum channels. Theory Probab. Appl. **50**(1), 86–98 (2006)
20. Horn, R., Johnson, Ch.: Matrix Analysis, 2nd ed. Cambridge University Press, Cambridge (2013)
21. Horodecki, K., Horodecki, M., Horodecki, P., Oppenheim, J.: Locking entanglement with a single qubit. Phys. Rev. Lett. **94**, 200501 (2005)
22. Hughston, L.P., Jozsa, R., Wootters, W.K.: A complete classification of quantum ensembles having a given density matrix. Phys. Lett. A **183**, 14–18 (1993)
23. Kim, I., Ruskai, M.-B.: Bounds on the concavity of entropy. J. Math. Phys. **55**, 092201 (2014)
24. Kimura, G., Miyadera, T., Imai, H.: Optimal state discrimination in general probabilistic theories. Phys. Rev. A **79**, 062306 (2009)
25. Kok, P., Lovett, B.W.: Introduction to Optical Quantum Information Processing. Cambridge University Press, Cambridge (2010)
26. Leung, D., Smith, G.: Continuity of quantum channel capacities. Commun. Math. Phys. **292**(1), 201–215 (2009)
27. Lieb, E.H., Ruskai, M.-B.: Proof of the strong subadditivity of quantum-mechanical entropy. J. Math. Phys. **14**(12), 1938–1941 (1973)
28. Mosonyi, M., Hiai, F.: On the quantum Rényi relative entropies and related capacity formulas. IEEE Trans. Inf. Theory **57**(4), 2474–2487 (2011)
29. Nielsen, M.A.: Continuity bounds for entanglement. Phys. Rev. A **61**, 064301 (2000)
30. Ohya, M., Petz, D.: Quantum Entropy and Its Use, 2nd ed. Springer, Berlin (2004)
31. Ouyang, Y.: Channel covariance, twirling, contraction, and some upper bounds on the quantum capacity. Quantum Inf. Comput. **14**(11, 12):917–936 (2014). arXiv:1106.2337v6 [quant-ph]
32. Petz, D.: Quantum Information Theory and Quantum Statistics. Springer, Berlin (2008)
33. Piani, M., Winter, A.: Extra-weak additivity of the relative entropy of entanglement **(in eternal preparation)**
34. Sason, I.: Entropy bounds for discrete random variables via maximal coupling. IEEE Trans. Inf. Theory **59**(11), 7118–7131 (2013)
35. Schrödinger, E.: Discussion of probability relations between separated systems. Math. Proc. Cambridge Phil. Soc. **31**(4), 555–563 (1935)
36. Schrödinger, E.: Probability relations between separated systems. Math. Proc. Cambridge Phil. Soc. **32**(3), 446–452 (1936)
37. Shirokov, M.E., Holevo, A.S.: On approximation of infinite-dimensional quantum channels. Probl. Inf. Transm. **44**(2), 73–90 (2008)
38. Shirokov, M.E.: Entropy characteristics of subsets of states. I. Izvestiya Math. **70**(6), 1265–1292 (2006)
39. Shirkov, M.E.: Continuity of the von Neumann entropy. Commun. Math. Phys. **296**(3), 625–654 (2010)
40. Shirokov, M.E.: Measures of quantum correlations in infinite-dimensional systems (2015). arXiv:1506.06377 [quant-ph]
41. Shirokov, M.E.: Squashed entanglement in infinite dimensions (2015). arXiv:1507.08964v3 [quant-ph]
42. Smith, G., Smolin, J.A., Winter, A.: The quantum capacity with symmetric side channels. IEEE Trans. Inf. Theory **54**(9), 4208–4217 (2008)
43. Smith, G.: Private classical capacity with a symmetric side channel and its application to quantum cryptography. Phys. Rev. A **78**, 022306 (2008)
44. Sutter, D., Scholz, V.B., Renner, R., Winter, A.: Approximate Degradable Quantum Channels (2015). arXiv:1412.0980v2 [quant-ph]
45. Sutter, D., Scholz, V.B., Renner, R., Winter, A. **(in preparation)**
46. Synak-Radtke, B., Horodecki, M.: On asymptotic continuity of functions of quantum states. J. Phys. A Math. Gen. **39**(2006), L 423–L 437 (2005). arXiv:quant-ph/0507126
47. Umegaki, H.: Conditional expectations in an operator alegebra IV (entropy and information). Kodai Math. Sem. Rep. **14**, 59–85 (1962)

48. Vedral, V., Plenio, M.B., Rippin, M.A., Knight, P.L.: Quantifying entanglement. Phys. Rev. Lett. **78**(12), 2275–2279 (1997)
49. Vollbrecht, K.G.H., Werner, R.F.: Entanglement measures under symmetry. Phys. Rev. A **64**, 062307 (2001)
50. Weedbrook, C., Pirandola, S., García-Patrón, R., Cerf, N.J., Ralph, T.C., Shapiro, J.H., Lloyd, S.: Gaussian quantum information. Rev. Mod. Phys **84**(2), 621–669 (2012)
51. Wehrl, A.: General properties of entropy. Rev. Mod. Phys. **50**(2), 221–260 (1978)
52. Weis, S., Knauf, A.: Entropy distance: new quantum phenomena. J. Math. Phys. **53**, 102206 (2012)
53. Wilde, M.M.: Quantum Information Theory. Cambridge University Press, Cambridge (2013)
54. Winter, A.: 'Extrinsic' and 'intrinsic' data in quantum measurements: asymptotic convex decomposition of positive operator valued measures. Commun. Math. Phys. **244**, 157–185 (2004)
55. Yang, D., Horodecki, M., Wang, Z.D.: An additive and operational entanglement measure: conditional entanglement of mutual information. Phys. Rev. Lett. **101**, 140501 (2008). arXiv:quant-ph/0701149
56. Zhang, Z.: Estimating mutual information via Kolmogorov distance. IEEE Trans. Inf. Theory **53**(9), 3280–3282 (2007)

Communicated by M. M. Wolf