Journal of Mathematical Physics

RESEARCH ARTICLE | SEPTEMBER 22 2023

Continuity bounds on observational entropy and measured relative entropies

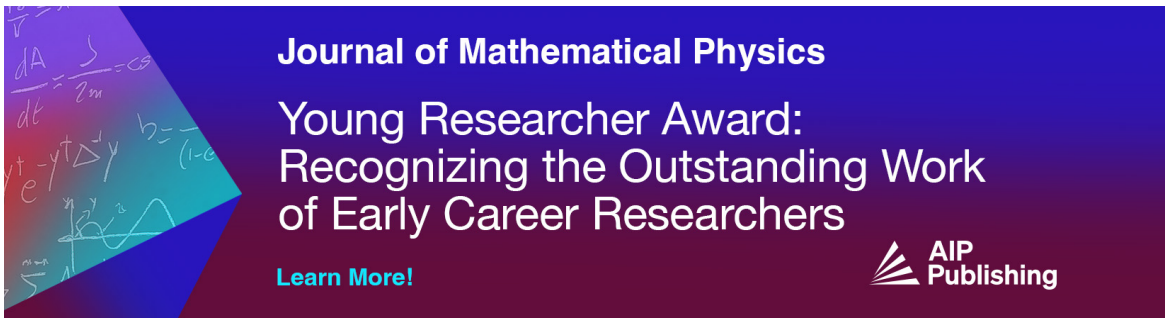
Joseph Schindler ; Andreas Winter

Check for updates

J. Math. Phys. 64, 092201 (2023)
https://doi.org/10.1063/5.0147294

View Online
Export Citation
CrossMark

04 December 2023 10:33:46



AIP Publishing

Journal of Mathematical Physics ARTICLE pubs.aip.org/aip/jmp

Continuity bounds on observational entropy and measured relative entropies

Cite as: J. Math. Phys. 64, 092201 (2023); doi: 10.1063/5.0147294
Submitted: 21 February 2023 • Accepted: 28 July 2023 •
Published Online: 22 September 2023

View Online Export Citation CrossMark

Joseph Schindler[1,a)] and Andreas Winter[1,2,3,a)]

AFFILIATIONS

[1] Física Teòrica: Informació i Fenòmens Quàntics, Departament de Física, Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain
[2] Institució Catalana de Recerca i Estudis Avançats, Pg. Lluís Companys, 23, 08010 Barcelona, Spain
[3] Institute for Advanced Study, Technische Universität München, Lichtenbergstraße 2a, 85748 Garching, Germany

[a)] **Authors to whom correspondence should be addressed:** JosephC.Schindler@uab.cat and andreas.winter@uab.cat

ABSTRACT

We derive a measurement-independent asymptotic continuity bound on the observational entropy for general positive operator valued measures measurements, making essential use of its property of bounded concavity. The same insight is used to obtain continuity bounds for other entropic quantities, including the measured relative entropy distance to a convex set of states under a general set of measurements. As a special case, we define and study conditional observational entropy, which is an observational entropy in one (measured) subsystem conditioned on the quantum state in another (unmeasured) subsystem. We also study continuity of relative entropy with respect to a jointly applied channel, finding that observational entropy is uniformly continuous as a function of the measurement. But we show by means of an example that this continuity under measurements cannot have the form of a concrete asymptotic bound.

© 2023 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/). https://doi.org/10.1063/5.0147294

04 December 2023 10:33:46

I. INTRODUCTION

Observational entropy has recently emerged (in fact, re-emerged, cf. Refs. 1–4) in studies of non-equilibrium statistical mechanics as a useful unifying framework to describe coarse-grained entropy in classical and quantum systems,[5–7] with applications across thermodynamics and quantum information theory.[8–27]

In the quantum case, for any measurement described by a positive operator valued measures (POVM) $M = (M_i)_{i\in\mathcal{I}}$, $(M_i \geq 0, \sum_i M_i = \mathbb{1})$, and quantum state described by a density matrix ρ, observational entropy

$$S_M(\rho) = -\sum_i p_i \log \frac{p_i}{V_i}, \qquad p_i = \mathrm{tr}(M_i\rho), \qquad V_i = \mathrm{tr}(M_i), \tag{1}$$

describes an entropy of the state ρ "coarse-grained" by measurement M. The p_i describe a probability distribution over "macrostates" (i.e. measurement outcomes), and the V_i describe "volumes" of macrostates analogous to the phase space volumes defining Boltzmann entropy.

The formula (1) defining observational entropy can be recast in relative entropy forms more suggestive of its information-theoretic content. In particular, with ρ defined on a Hilbert space $\mathscr{H}$ of finite dimension d,

$$D(\Phi_M(\rho) \,\|\, \Phi_M(\mathbb{1}/d)) = D(p \,\|\, q) = \log d - S_M(\rho), \tag{2}$$

© Author(s) 2023

with D denoting either quantum or classical relative entropy depending on its arguments. Here M is implemented by quantum-classical channel

$$\Phi_M(\sigma) = \sum_i \mathrm{tr}(\sigma M_i)|i\rangle\langle i|, \tag{3}$$

called the measuring channel of M, and $p_i = \mathrm{tr}(M_i\rho)$, $q_i = \mathrm{tr}(M_i \mathbb{1}/d)$ are the observed probability distributions induced by M on ρ and on the maximally mixed state, respectively.

It is evident from the definition that observational entropy is a continuous function of the state ρ. For practical purposes, however, it is most useful to realize this continuity statement in terms of concrete continuity bounds, as has also been done for many other entropic quantities.[28–32]

One naive version of such a bound follows quickly from other known bounds. Expanding the definition (1) one can separate observational entropy into a sum of observed Shannon entropy $H(p) = -\sum_i p_i \log p_i$ over measurement outcomes, and a mean Boltzmann entropy,

$$S_M(\rho) = H(p) + \sum_i p_i \log V_i. \tag{4}$$

Bounding the Shannon and Boltzmann terms separately one can obtain, given two states ρ, σ such that $\frac{1}{2}\|\rho - \sigma\|_1 \le \epsilon$,

$$\begin{aligned} |S_M(\rho) - S_M(\sigma)| &\le \left|H(p^\rho) - H(p^\sigma)\right| + \left|\sum_i (p_i^\rho - p_i^\sigma) \log V_i\right| \\ &\le h(\epsilon) + \epsilon \log|M| + \epsilon \|\log V_i\|_\infty. \end{aligned} \tag{5}$$

Here $|M|$ is the number of outcomes of the POVM M (cardinality of the index set $\mathcal{I}$), the function $h(x) = -x\log x - (1-x)\log(1-x)$ is the binary Shannon entropy, and $\|\log V_i\|_\infty = \max_i |\log V_i|$. The first two terms in the bottom line arise from the classical Zhang–Audenaert continuity bound on Shannon entropy,[32–34] while the latter term provides a loose bound on the Boltzmann term. Both terms are derived using elementary properties relating the quantum trace norm to probabilities.[34]

Though apparently useful insofar as it implies $|\Delta S| \to 0$ as $\epsilon \to 0$ for any particular M, the bound (5) fails to be universal, depending on the measurement M both through the number of outcomes $|M|$ and through the volume terms V_i. Moreover, it is unacceptably loose, as evidenced by the following pathology. Given any POVM M, one can define another one M', with twice as many outcomes, by splitting each POVM element in half and counting it twice ($M_i \to$ two copies $M_i/2$). Letting $M \to M'$ the observational entropies $S_M(\rho)$ and $S_M(\sigma)$ are both invariant, but by repeating this transformation the bound (5) can be made arbitrarily loose (noting that after $V_i \le 1$ each iteration separately loosens both terms). This pathology highlights the need for an improved and universal bound.

One expects the shortcomings above to be common to any bound attempting to treat the Shannon and Boltzmann terms separately. On the other hand, one might hope that the presence of normalizing volume terms in the definition (1) may actually help improve the bound compared to a Shannon entropy rather than [as occurs in (5)] making it looser. We will show this is indeed the case, improving the bound by treating the terms together.

In particular, the purpose of this note is to demonstrate that, for any measurement M and any states ρ, σ such that $\frac{1}{2}\|\rho - \sigma\|_1 \le \epsilon$, observational entropy obeys a continuity bound of the form

$$|S_M(\rho) - S_M(\sigma)| \le g(\epsilon) + \epsilon \log d. \tag{6}$$

This bound is universal in that it is independent of M, and depends only on the Hilbert space dimension d and the trace distance between the states.

The function $g(\epsilon)$ arising in the bound is defined by $g(x) = -x\log x + (1+x)\log(1+x)$ for $x > 0$, and $g(0) \equiv 0$. It derives here from the binary Shannon entropy $h(x) = -x\log x - (1-x)\log(1-x)$, where it arises as

$$g(\epsilon) = (1+\epsilon)\, h\left(\frac{\epsilon}{1+\epsilon}\right). \tag{7}$$

Elsewhere it arises as an entropy in bosonic systems.[35] This $g(x)$ is particularly convenient, as throughout $x \in [0, \infty)$ it is continuous, monotonic increasing, non-negative, and concave.

The main bound (6) appears in Theorem 6. In order to establish it, we first will observe that observational entropy has the same bounded concavity property typical of other entropic quantities, that is, it is concave but not too concave. We then establish a general continuity bound based only on bounded concavity, obtaining the main result as a corollary. The method of proof follows closely the Proof of Lemma 2 of Ref. 32, which obtained a closely related tight bound on quantum conditional entropy, and which in turn is closely related to methods used in Refs. 36 and 37, this method sometimes being known as the Alicki–Fannes–Winter (AFW) trick.[32,36] The general bound Proposition 5 given here is closely related to Lemma 7 of Ref. 32, and provides a slightly different generalization thereof. Continuity bounds of this form are often referred to as asymptotic continuity due to the bound per qudit scaling as $\epsilon \log d$ in the n-copy regime.[38,39]

In addition to the main observational entropy result, we also discuss continuity results for other entropic quantities that may be of general interest. In Proposition 8 we consider relative entropy distances of the form

$$\inf_{\sigma \in \chi} D(\mathcal{M}(\rho) \,\|\, \mathcal{N}(\sigma)) \tag{8}$$

04 December 2023 10:33:46

© Author(s) 2023

to a convex set of states χ under channels $\mathcal{M}, \mathcal{N}$. In Theorem 11 we then consider measured relative entropies under restricted measurements[40]

$$\inf_{\sigma\in\chi} \sup_{M\in\mathscr{M}} D_M(\rho \,\|\, \sigma), \tag{9}$$

where $D_M(\rho\|\sigma)$ is the classical relative entropy observed under measurement M. And in Proposition 12 we apply these results to the case of an observational conditional entropy $S_{M_A}(A|B)_\rho$ motivated by relation to entropic uncertainty principles. All of these bounds are derived by appealing to a general bound in the form

$$|Z(\rho) - Z(\sigma)| \le g(\epsilon) + \epsilon\,\kappa \tag{10}$$

shown in Proposition 5 for $Z(\rho)$ any function obeying bounded concavity/convexity. As a final step, in Corollary 15 we show that $S_M(\rho)$ and $D_M(\rho\|\sigma)$ are continuous as functions of the measurement M (in an appropriate topology on POVMs). This result is a special case of Theorem 14, where we show uniform continuity of

$$D(\Phi(\rho) \,\|\, \Phi(\sigma)) \tag{11}$$

as a function of channel Φ. In Example 13 we show limitations preventing simple asymptotic bounds from being obtained for the measurement continuity.

In the following we continue to consider a finite dimensional Hilbert space $\mathscr{H}$ of dimension d unless otherwise specified. Quantum states ρ are positive semidefinite Hermitian operators normalized to unit trace. Measurements are described by POVMs (positive operator valued measures), defined as collections of positive semidefinite Hermitian operators summing to the identity, assumed here to have a finite number of outcomes. The von Neumann entropy is denoted by $S(\rho) = -\mathrm{tr}(\rho\log\rho)$, and Shannon entropy by $H(p) = -\sum_i p_i \log p_i$. The binary Shannon entropy function is denoted by $h(x)$, and $g(x)$ is as above. Observational entropy $S_M(\rho)$ is defined by (1).

II. BOUNDED CONCAVITY

An elementary property of Shannon entropy is bounded concavity[41] (in the case of classical Shannon entropy this property can be derived directly from a simple chain rule computation, but it can also be inferred as a subcase of the more commonly emphasized quantum generalization), which we state here as a lemma for later use.

Lemma 1. Let λ_k be a probability distribution and p^k a probability distribution for each k. The Shannon entropy H is concave but not too concave,

$$0 \le \left[H\left(\sum_k \lambda_k p^k\right) - \sum_k \lambda_k H(p^k)\right] \le H(\lambda). \tag{12}$$

Note that $(p^k)_i = p_{i|k}$ is a conditional probability distribution and $\left(\sum_k \lambda_k p^k\right)_i = \sum_k \lambda_k p_{i|k}$.

Bounded concavity of observational entropy follows from the bounded concavity of Shannon entropy.

Lemma 2 [observational entropy (OE) Bounded Concavity]. Let λ_k be a probability distribution and ρ_k a state for each k. The observational entropy S_M is concave but not too concave,

$$0 \le \left[S_M\left(\sum_k \lambda_k \rho_k\right) - \sum_k \lambda_k S_M(\rho_k)\right] \le H(\lambda). \tag{13}$$

Proof. Let $\rho = \sum_k \lambda_k \rho_k$ and $p_i = \mathrm{tr}(M_i\rho)$. Then $p_i = \sum_k \lambda_k p_{i|k}$, where $p_{i|k} = \mathrm{tr}(M_i\rho_k)$. Observe that only the difference in Shannon entropy contributes to the difference in OE, since the mean Boltzmann entropy is equal for the two distributions. That is, we have [abusing the notation $H(p_i) \leftrightarrow H(p)$ for simplicity where unambiguous]

$$\begin{aligned}
\Delta S &\equiv S_M(\rho) - \sum_k \lambda_k S_M(\rho_k) && (14)\\
&= H(p_i) + \sum_i p_i \log V_i - \sum_k \lambda_k \left[H(p_{i|k}) + \sum_i p_{i|k} \log V_i\right] && (15)\\
&= H(p_i) - \sum_k \lambda_k H(p_{i|k}) && (16)\\
&\equiv \Delta H. && (17)
\end{aligned}$$

But Lemma 1 precisely says that $0 \le \Delta H \le H(\lambda)$, the desired result. □

In general bounded concavity for real functions on convex sets may be defined as follows.

Definition 3 (General Bounded Concavity/Convexity). A function $Z : \mathcal{S} \to \mathbb{R}$ defined on a convex set $\mathcal{S}$ is bounded concave if, for any finite convex combination $\rho = \sum_k \lambda_k \rho_k$,

$$0 \le \left[Z(\rho) - \sum_k \lambda_k Z(\rho_k)\right] \le H(\lambda). \tag{18}$$

A function Z is bounded convex if $-Z$ is bounded concave.

04 December 2023 10:33:46

© Author(s) 2023

In the next section we derive a continuity bound for functions of quantum states that applies to any function with a bounded concavity or convexity property of this form, with observational entropy continuity as a corollary.

III. CONTINUITY

The continuity bound proved below is derived by appealing to bounded concavity. Similar methods have been used to derive continuity results for other entropic quantities in the literature, such as in Refs. 32, 36–39, and 42–46. The following decomposition is essential to performing the main trick of the proof, sometimes referred to as the Alicki-Fannes-Winter trick (cf. Refs. 32 and 36). Whereas the previous section only involved convex structure, the derivation of the following decomposition also involves topology via the trace norm. The resulting continuity bound depends on both the convex and topological structures of the space of quantum states.

Lemma 4 ($\omega\Delta$ Decomposition). Let ρ, σ be quantum states, and define $\epsilon = \frac{1}{2}\|\rho - \sigma\|_1$. There exists a state ω and states $\Delta_\pm$ such that

$$\begin{aligned} \omega &= \frac{1}{1+\epsilon}\,\rho + \frac{\epsilon}{1+\epsilon}\,\Delta_- \\ &= \frac{1}{1+\epsilon}\,\sigma + \frac{\epsilon}{1+\epsilon}\,\Delta_+. \end{aligned} \tag{19}$$

Proof. Let ρ, σ be states such that $\|\rho - \sigma\|_1 = 2\epsilon$, and let

$$\epsilon\,\Delta_\pm = \pm(\rho - \sigma)_\pm. \tag{20}$$

The positive and negative parts $(\rho - \sigma)_\pm$ are defined in the usual way,[34] by diagonalizing the Hermitian matrix $\rho - \sigma$, then separating into the sum of a non-negative diagonal and a non-positive diagonal matrix. That is, $\rho - \sigma = U^\dagger D U = U^\dagger(D_+ + D_-)U = \epsilon\Delta_+ - \epsilon\Delta_-$. By definition $\Delta_\pm \geq 0$ since they are Hermitian with only non-negative eigenvalues. Since $\rho - \sigma$ is traceless, we have $\mathrm{tr}(\Delta_+) = \mathrm{tr}(\Delta_-) = \|\Delta_\pm\|_1$, with the last equality by positive semidefiniteness. Further, $\|\rho - \sigma\|_1 = \|D_+ + D_-\|_1 = \|D_+\|_1 + \|D_-\|_1 = \|\epsilon\Delta_+\|_1 + \|\epsilon\Delta_-\|_1 = 2\epsilon\|\Delta_\pm\|_1 = 2\epsilon$ using unitary invariance of the trace norm and the fact that $D_\pm$ are diagonal. Thus $\mathrm{tr}(\Delta_\pm) = 1$, and $\Delta_\pm$ are states.

Now define

$$\omega = \frac{1}{1+\epsilon}\,\frac{(\rho+\sigma)}{2} + \frac{\epsilon}{1+\epsilon}\,\frac{(\Delta_+ + \Delta_-)}{2}. \tag{21}$$

This ω is a convex combination of states, so $\mathrm{tr}(\omega) = 1$ and ω is a state. The above relations taken together with ω can be viewed as expressing a linear decomposition

$$\rho - \sigma = \epsilon\,(\Delta_+ - \Delta_-) \tag{22}$$

$$\rho + \sigma = 2(1+\epsilon)\,\omega - \epsilon(\Delta_+ + \Delta_-). \tag{23}$$

From (21) with (22) it is straightforward to verify that

$$\omega = \frac{1}{1+\epsilon}\,\rho + \frac{\epsilon}{1+\epsilon}\,\Delta_- = \frac{1}{1+\epsilon}\,\sigma + \frac{\epsilon}{1+\epsilon}\,\Delta_+, \tag{24}$$

which completes the proof. □

Next appears a general continuity bound for any function $Z(\rho)$ obeying bounded concavity on quantum states. Since (25) is invariant under $Z \to -Z$ this implies the same bound for functions obeying bounded convexity.

Proposition 5. Let $Z(\rho)$ be any function satisfying bounded concavity [cf. (18)] on quantum states. We have the following continuity bound. Let ρ, σ be states such that $\frac{1}{2}\|\rho - \sigma\|_1 \leq \epsilon$. Then

$$|Z(\rho) - Z(\sigma)| \leq g(\epsilon) + \epsilon\,\kappa, \tag{25}$$

where $g(\epsilon)$ is as in (7) and $\kappa = \sup_{(\mu,\nu)}|Z(\mu) - Z(\nu)|$ is the supremum absolute difference in Z between any two states.

Proof. Suppose $\epsilon = \frac{1}{2}\|\rho - \sigma\|_1$, and suppose there exist a state ω and states $\Delta_\pm$ such that

$$\omega = \frac{1}{1+\epsilon}\,\rho + \frac{\epsilon}{1+\epsilon}\,\Delta_- = \frac{1}{1+\epsilon}\,\sigma + \frac{\epsilon}{1+\epsilon}\,\Delta_+. \tag{26}$$

The existence of such states was demonstrated in Lemma 4. From (26) we have two different convex decompositions $\omega = \lambda\,\omega_1 + (1-\lambda)\omega_2$, both with the same coefficients $\lambda = 1/(1+\epsilon)$. By the assumption (18) of bounded concavity,

$$[\lambda\,Z(\omega_1) + (1-\lambda)\,Z(\omega_2)] \leq Z(\omega) \leq h(\lambda) + [\lambda\,Z(\omega_1) + (1-\lambda)\,Z(\omega_2)] \tag{27}$$

04 December 2023 10:33:46

© Author(s) 2023

holds for both of these convex decompositions. The trick is simply to use $\omega_1, \omega_2 = \rho, \Delta_-$ for one side of the inequality and $\omega_1, \omega_2 = \sigma, \Delta_+$ for the other. This amounts to

$$\lambda Z(\rho) + (1-\lambda) Z(\Delta_-) \leq h(\lambda) + \lambda Z(\sigma) + (1-\lambda) Z(\Delta_+). \tag{28}$$

Rearranging yields

$$(Z(\rho) - Z(\sigma)) \leq \frac{h(\lambda) + (1-\lambda)\left(Z(\Delta_+) - Z(\Delta_-)\right)}{\lambda}. \tag{29}$$

Finally we take the magnitude and apply triangle inequality,

$$|Z(\rho) - Z(\sigma)| \leq \frac{h(\lambda)}{\lambda} + \frac{(1-\lambda)}{\lambda} |Z(\Delta_+) - Z(\Delta_-)| \tag{30}$$

$$\leq \frac{h(\lambda)}{\lambda} + \frac{(1-\lambda)}{\lambda} \kappa. \tag{31}$$

The second inequality follows since a magnitude difference in Z between two states is by definition less than or equal to κ. Finally with $\lambda = 1/(1+\epsilon)$ and $h(\lambda) = h(1-\lambda)$ we obtain

$$|Z(\rho) - Z(\sigma)| \leq g(\epsilon) + \epsilon \kappa. \tag{32}$$

One can confirm the bound is monotonic in ϵ. Therefore if it holds for $\epsilon = \frac{1}{2}\|\rho - \sigma\|_1$, as was just shown, then it also holds for $\epsilon' \geq \epsilon$. This completes the proof. □

In the following sections this general bound is evaluated for the observational entropy and extended to a form that includes several related quantities.

IV. OBSERVATIONAL ENTROPY

The general statement Proposition 5 implies a continuity bound on observational entropy.

Theorem 6 (OE Continuity). *Let ρ, σ be quantum states such that $\frac{1}{2}\|\rho - \sigma\|_1 \leq \epsilon$. Then observational entropy obeys the continuity bound*

$$|S_M(\rho) - S_M(\sigma)| \leq g(\epsilon) + \epsilon \log d, \tag{33}$$

where $g(\epsilon)$ is as in (7) and d is the Hilbert space dimension.

Proof. By Lemma 2, S_M obeys bounded concavity, so the bound follows from Proposition 5. An elementary property of observational entropy is that $0 \leq S_M(\rho) \leq \log d$.[10] Thus we have $\kappa = \sup_{(\mu,\nu)} |S_M(\mu) - S_M(\nu)| = \sup_\mu S_M(\mu) = \log d$. □

The main result, a measurement-independent continuity bound on observational entropy, has therefore been established.

As an aside we note that bounds on both the Shannon and von Neumann entropies can be derived from the observational entropy continuity, as shown in the following corollary. The bounds thus obtained are looser than the known optimized forms.[30]

Corollary 7. Let ρ, σ be states such that $\frac{1}{2}\|\rho - \sigma\|_1 \leq \epsilon$. Then von Neumann entropy S obeys

$$|S(\rho) - S(\sigma)| \leq g(\epsilon) + \epsilon \log d. \tag{34}$$

Let p_i, q_i be N-outcome probability distributions such that $\frac{1}{2}\|\vec{p} - \vec{q}\|_1 \leq \epsilon$. Then Shannon entropy H obeys

$$|H(p_i) - H(q_i)| \leq g(\epsilon) + \epsilon \log N. \tag{35}$$

In this form these are both special cases of Theorem 6, however, tighter bounds can be derived by other methods.

Proof. We derive these as corollaries of Theorem 6. For the quantum case suppose (wlg) that $S(\rho) \geq S(\sigma)$. Let M_σ measure in the σ eigenbasis. Then $|S_{M_\sigma}(\rho) - S_{M_\sigma}(\sigma)| = S_{M_\sigma}(\rho) - S(\sigma) \geq S(\rho) - S(\sigma) = |S(\rho) - S(\sigma)|$. For the classical case, let $\rho = \sum_i p_i |i\rangle\langle i|$ and $\sigma = \sum_i q_i |i\rangle\langle i|$ in N-dimensional Hilbert space, and M_0 measure in the $|i\rangle\langle i|$ basis. (At least N dimensions are needed to embed p, q classically.) Since the states are diagonal we have $\|\rho - \sigma\|_1 = \|\vec{p} - \vec{q}\|_1 \leq 2\epsilon$. Then measuring M_0 obtains p_i, q_i and we have $|S_{M_0}(\rho) - S_{M_0}(\sigma)| = |H(p_i) - H(q_i)|$. □

04 December 2023 10:33:46

© Author(s) 2023

V. MEASURED RELATIVE ENTROPY

The general Proposition 5 likewise provides a continuity bound on measured relative entropies associated with a particular measurement, as well as on the quantum relative entropy.

Proposition 8. Let χ be a convex set of states (or more generally, of positive semidefinite Hermitian operators). Let $\mathcal{M}$ and $\mathcal{N}$ be arbitrary quantum channels. Then

$$Z(\rho) = \inf_{\sigma\in\chi} D(\mathcal{M}(\rho) \parallel \mathcal{N}(\sigma)), \tag{36}$$

is bounded convex. Thus $Z(\rho)$ obeys the bound (25) of Proposition 5. This ensures continuity whenever the maximum variation κ is finite.

Proof. Let $\rho = \sum_k \lambda_k \rho_k$. First note that for any fixed $\sigma \in \chi$ we have

$$\sum_k \lambda_k D(\mathcal{M}(\rho_k) \parallel \mathcal{N}(\sigma)) - D(\mathcal{M}(\rho) \parallel \mathcal{N}(\sigma)) = S(\mathcal{M}(\rho)) - \sum_k \lambda_k S(\mathcal{M}(\rho_k)), \tag{37}$$

since $D(\mathcal{M}(\rho) \parallel \mathcal{N}(\sigma)) = -S(\mathcal{M}(\rho)) - \operatorname{tr}\mathcal{M}(\rho)\log\mathcal{N}(\sigma)$ and applying linearity of the channels. Thus

$$0 \le \sum_k \lambda_k D(\mathcal{M}(\rho_k) \parallel \mathcal{N}(\sigma)) - D(\mathcal{M}(\rho) \parallel \mathcal{N}(\sigma)) \le H(\lambda) \tag{38}$$

by bounded concavity of von Neumann entropy. Note that $D(\mathcal{M}(\rho) \parallel \mathcal{N}(\sigma)) = \infty$ if and only if $\sum_k \lambda_k D(\mathcal{M}(\rho_k) \parallel \mathcal{N}(\sigma)) = \infty$ because the von Neumann terms are finite in finite dimension, and that (38) can be rearranged so the infinite case $\infty = \infty = \infty$ is well defined.

Now we introduce the infimum over $\sigma \in \chi$. It is immediate that the upper bound on convexity is retained, since

$$\sum_k \lambda_k \inf_{\sigma_k\in\chi} D(\mathcal{M}(\rho_k) \parallel \mathcal{N}(\sigma_k)) \le \inf_{\sigma\in\chi} \sum_k \lambda_k D(\mathcal{M}(\rho_k) \parallel \mathcal{N}(\sigma)) \le H(\lambda) + Z(\rho), \tag{39}$$

using that the sum of infima is no greater than the infimum of the sum and (38).

The upper convexity bound just derived holds for the infimum of any family of bounded convex functions. In the particular case considered here, the lower convexity bound also holds, using convexity of the set χ and joint convexity of the relative entropy. In particular, since $Z(\rho)$ is an infimum, it follows that for any ρ and any $\delta > 0$, there exists a state $\gamma \in \chi$ such that $D(\mathcal{M}(\rho) \parallel \mathcal{N}(\gamma)) \le Z(\rho) + \delta$. Thus fix $\delta > 0$ and let $D(\mathcal{M}(\rho_k) \parallel \mathcal{N}(\gamma_k)) \le Z(\rho_k) + \delta$, with $\nu = \sum_k \lambda_k \gamma_k$. Then

$$\sum_k \lambda_k Z(\rho_k) + \delta \ge \sum_k \lambda_k D(\mathcal{M}(\rho_k) \parallel \mathcal{N}(\gamma_k)) \ge D(\mathcal{M}(\rho) \parallel \mathcal{N}(\nu)) \ge Z(\rho). \tag{40}$$

Since this holds for all $\delta > 0$, it follows that $\sum_k \lambda_k Z(\rho_k) \ge Z(\rho)$. The first inequality expresses the initial infimum assumption, the second is joint convexity of relative entropy and linearity of the channels, and the last that $\nu \in \chi$ using convexity of the set. □

Remark 9. If $\mathcal{N}(\chi)$ contains an element of full rank the maximum variation κ is necessarily finite. If $\mathcal{M} = \mathcal{N}$ and χ is a set of states containing the maximally mixed state, then $\kappa \le \log d$.

Proof. If $\mathcal{N}(\chi)$ contains an element $\mathcal{N}(\sigma)$ of full rank then all $|Z(\rho)| \le \log d + \|\log\mathcal{N}(\sigma)\|_\infty$ and therefore κ is finite. Meanwhile if $\mathcal{M} = \mathcal{N}$ and $\chi \ni \sigma$ is a set of states, then $D(\rho \parallel \sigma) \ge D(\mathcal{M}(\rho) \parallel \mathcal{M}(\sigma)) \ge 0$, so if also $(\mathbb{1}/d) \in \chi$ then $\kappa \le \sup_\rho S(\rho) \le \log d$. □

In particular, Proposition 8 in the case $\mathcal{M} = \mathcal{N} = \Phi_M$ [cf. (3)] demonstrates a bound on measured relative entropy distance to a convex set of states, given the single measurement M. A natural question is whether this extends to include measured relative entropies associated with an allowed class of measurements $\mathscr{M}$,[40,42]

$$D_{\mathscr{M}}(\rho \parallel \sigma) = \sup_{M\in\mathscr{M}} D(\Phi_M(\rho) \parallel \Phi_M(\sigma)). \tag{41}$$

In full generality we take χ a convex set of states and $\mathscr{M}$ a set of measurements, and consider the quantity

$$Z(\rho) = \inf_{\sigma\in\chi} D_{\mathscr{M}}(\rho \parallel \sigma). \tag{42}$$

Due to the supremum in $D_{\mathscr{M}}$ this $Z(\rho)$ fails to obey bounded convexity, and the bound (25) cannot be applied directly. Nonetheless, we demonstrate below that a closely related continuity result can be obtained. This provides a more general and systematic extension to related results previously appearing in Ref. 42, and also subsumes the entropic bounds already considered above.

04 December 2023 10:33:46

© Author(s) 2023

A crucial step is the following minimax lemma.

Lemma 10. Let χ be a compact convex set of states. Let $\mathscr{M}$ be a set of quantum channels that is closed under flagged finite convex combination, that is, such that if $\mathcal{M}_k \in \mathscr{M}$, and λ_k is a finite probability distribution, then also $\sum_k \lambda_k \mathcal{M}_k \otimes |k\rangle\langle k| \in \mathscr{M}$. Then

$$\inf_{\sigma\in\chi} \sup_{\mathcal{M}\in\mathscr{M}} D(\mathcal{M}(\rho) \parallel \mathcal{M}(\sigma)) = \sup_{\mathcal{M}\in\mathscr{M}} \inf_{\sigma\in\chi} D(\mathcal{M}(\rho) \parallel \mathcal{M}(\sigma)). \tag{43}$$

Operationally motivated classes $\mathscr{M}$ of transformations often have the required property automatically, such as LOCC, SEP, PPT channels in composite systems.

Proof. This lemma has been established previously in Ref. 47 on the grounds of a general minimax theorem of Ref. 48. For completeness we include its demonstration here. We argue based on the minimax theorem in the form of Sion.[49] Namely, consider the function $\sum_i p_i D(\mathcal{M}_i(\rho) \parallel \mathcal{M}_i(\sigma))$ of two arguments, one a state $\sigma \in \chi$, the other a finite-support probability distribution over channels from $\mathscr{M}$. The first set (χ) is convex and compact by assumption, the second (denoted $\mathbb{P}\mathscr{M}$) is convex by construction but otherwise an arbitrary space. For simplicity assume that our objective function $f(\sigma, p)$ is finite; it is then continuous, and it is convex in the first and linear—in particular concave—in the second argument. Hence, by Ref. 49 we get

$$\begin{aligned}\inf_{\sigma\in\chi} \sup_{\mathcal{M}\in\mathscr{M}} D(\mathcal{M}(\rho) \parallel \mathcal{M}(\sigma)) &= \inf_{\sigma\in\chi} \sup_{p\in\mathbb{P}\mathscr{M}} \sum_i p_i D(\mathcal{M}_i(\rho) \parallel \mathcal{M}_i(\sigma)) \\ &= \sup_{p\in\mathbb{P}\mathscr{M}} \inf_{\sigma\in\chi} \sum_i p_i D(\mathcal{M}_i(\rho) \parallel \mathcal{M}_i(\sigma)).\end{aligned}$$

Note that in general we need the sum over i weighted by p_i's, otherwise the final supremum is not large enough. But if the set of channels is closed under taking flagged convex combinations, i.e. with $\mathcal{M}_i$ it also contains $\mathcal{M}' := \sum_i p_i \mathcal{M}_i \otimes |i\rangle\langle i|$, then $\sum_i p_i D(\mathcal{M}_i(\rho) \parallel \mathcal{M}_i(\sigma)) = D(\mathcal{M}'(\rho) \parallel \mathcal{M}'(\sigma))$, hence

$$\sup_{p\in\mathbb{P}\mathscr{M}} \inf_{\sigma\in\chi} \sum_i p_i D(\mathcal{M}_i(\rho) \parallel \mathcal{M}_i(\sigma)) = \sup_{\mathcal{M}\in\mathscr{M}} \inf_{\sigma\in\chi} D(\mathcal{M}(\rho) \parallel \mathcal{M}(\sigma)),$$

which concludes the proof in the case that f is always finite. To cover the case of values $+\infty$, one can modify the argument, removing the infinities from the objective function; or use the minimax theorem from Ref. 48 as done in Ref. 47. □

For classes $\mathscr{M}$ of measurements closed under disjoint convex combination, and compact convex sets χ of states, we have—with the help of the minimax lemma—the following continuity theorem for measured relative entropies $Z(\rho) = \inf_{\sigma\in\chi} D_{\mathscr{M}}(\rho \parallel \sigma)$.

We remark that many commonly invoked measurement classes (such as LOCC, SEP, PPT measurements in composite systems) automatically fulfill the required closure property. Furthermore, given some primitive set of measurements $\mathscr{M}_0$ not already doing so, from an operational perspective it is perfectly feasible to complete this to an $\mathscr{M}$ that is closed in this way, given the capability to choose which measurements to perform based on classical randomness.

Theorem 11. *Let χ be a compact convex set of states. Let $\mathscr{M}$ be a convex set of measurements that is also closed under disjoint convex combination (that is, convex combinations taken on the disjoint union of the outcome sets). Let Φ_M as in (3) and define*

$$Z(\rho) = \inf_{\sigma\in\chi} \sup_{M\in\mathscr{M}} Z_{M,\sigma}(\rho), \qquad Z_{M,\sigma}(\rho) = D(\Phi_M(\rho) \parallel \Phi_M(\sigma)). \tag{44}$$

Then for any two states ρ, ρ' such that $\frac{1}{2}\|\rho - \rho'\|_1 \le \epsilon$,

$$\left|Z(\rho) - Z(\rho')\right| \le g(\epsilon) + \epsilon\,\kappa. \tag{45}$$

Here

$$\kappa = \sup_{M\in\mathscr{M}} \left[\sup_{\mu,\nu} |Z_M(\mu) - Z_M(\nu)|\right], \qquad Z_M(\rho) = \inf_{\sigma\in\chi} Z_{M,\sigma}(\rho), \tag{46}$$

and $g(\epsilon)$ is as in (7).

Proof. Let $\Phi_{\mathscr{M}}$ denote the set of measuring channels associated with class $\mathscr{M}$, defined as in (3). Clearly $Z(\rho) = \inf_{\sigma\in\chi} \sup_{M\in\mathscr{M}} D(\Phi_M(\rho) \parallel \Phi_M(\sigma)) = \inf_{\sigma\in\chi} \sup_{\Phi\in\Phi_{\mathscr{M}}} D(\Phi_M(\rho) \parallel \Phi_M(\sigma))$. Now the measuring channel is convex linear in M on a fixed outcome set, $\Phi_{\sum_k \lambda_k M_k} = \sum_k \lambda_k \Phi_{M_k}$. To perform disjoint convex combination, which we denote $\oplus_k \lambda_k M_k$, one first extends each outcome set to the disjoint union of outcome sets, padding with zeroes, then convexly combines on the fixed set. Suppose $\Phi_{M_k} \in \Phi_{\mathscr{M}}$. It can be seen that $\Phi = \sum_k \lambda_k \Phi_{M_k} \otimes |k\rangle\langle k| = \Phi_{\oplus_k \lambda_k M_k}$, which is also in $\Phi_{\mathscr{M}}$ since $\oplus_k \lambda_k M_k \in \mathscr{M}$. Therefore the assumptions of the Lemma 10 are fulfilled, and it holds

$$\inf_{\sigma\in\chi} \sup_{M\in\mathscr{M}} Z_{M,\sigma}(\rho) = \sup_{M\in\mathscr{M}} \inf_{\sigma\in\chi} Z_{M,\sigma}(\rho). \tag{47}$$

04 December 2023 10:33:46

© Author(s) 2023

This provides the key step to relate $Z(\rho)$ to a form where bounded convexity can be exploited.

Proceeding, first using the minimax equality, and next that a difference of suprema is no less than the supremum difference, one obtains

$$|Z(\rho) - Z(\rho')| = \left| \inf_{\sigma\in\chi} \sup_{M\in\mathscr{M}} Z_{M,\sigma}(\rho) - \inf_{\sigma\in\chi} \sup_{M\in\mathscr{M}} Z_{M,\sigma}(\rho') \right| \tag{48}$$

$$= \left| \sup_{M\in\mathscr{M}} \inf_{\sigma\in\chi} Z_{M,\sigma}(\rho) - \sup_{M\in\mathscr{M}} \inf_{\sigma\in\chi} Z_{M,\sigma}(\rho') \right| \tag{49}$$

$$\leq \sup_{M\in\mathscr{M}} \left| Z_M(\rho) - Z_M(\rho') \right|. \tag{50}$$

But by Proposition 8 this $Z_M(\rho)$ is bounded convex and obeys (25), hence further

$$|Z(\rho) - Z(\rho')| \leq \sup_{M\in\mathscr{M}} [g(\epsilon) + \epsilon\,\kappa_M] \tag{51}$$

$$= g(\epsilon) + \epsilon \sup_{M\in\mathscr{M}} \kappa_M, \tag{52}$$

establishing the result. □

This provides asymptotic continuity of the entropies of restricted measurement first considered by Piani.[40] Similar results were also obtained recently in independent work, in the context of local classes of measurements, and will appear elsewhere.[50]

The main limitation of the Theorem 11 is the requirement that χ be convex. One may also be interested in measured relative entropy distances to non-convex sets of states such as coherent or classically correlated states. In this regard the theorem can be compared to results of Ref. 42, which obtained a looser bound but allows χ to be any set star-shaped around the maximally mixed state. Obtaining a bound for star-shaped χ with an optimal dimensional factor remains a useful future direction.

VI. CONDITIONAL OBSERVATIONAL ENTROPY

The above results are applied in this section to derive continuity bounds on an observational conditional entropy in bipartite systems, defined by

$$S_{M_A}(A|B)_\rho = -D(\,(\Phi_{M_A} \otimes I_B)\,\rho \,\|\, (\Phi_{M_A} \otimes I_B)(\mathbb{1}_A \otimes \rho_B)), \tag{53}$$

a quantity motivated by applications to entropic uncertainty principles with memory.[51] This conditional form describes quantum uncertainty in system B having made a measurement M_A in system A.

We have the following bound. Note that the dimensional factor of $\log d$ arising from maximum variation is reduced by a factor of two from the corresponding bound on quantum conditional entropy $S(A|B)_\rho = -D(\rho \| \mathbb{1}_A \otimes \rho_B)$.[32]

Proposition 12. Let ρ, σ be quantum states such that $\frac{1}{2}\|\rho - \sigma\|_1 \leq \epsilon$. Then

$$|S_{M_A}(A|B)_\rho - S_{M_A}(A|B)_\sigma| \leq g(\epsilon) + \epsilon \log d_A, \tag{54}$$

where $g(\epsilon)$ is as in (7) and d_A is the Hilbert space dimension in system A.

Proof. We derive the bound as a consequence of Theorem 8. To establish the link we note that the observational conditional entropy $S_{M_A}(A|B)_\rho$ defined by (53) admits the variational formula

$$S_{M_A}(A|B)_\rho = \log d_A - \inf_{\omega_B} D\Big((\Phi_{M_A} \otimes I_B)\,\rho \,\|\, (\Phi_{M_A} \otimes I_B)(\frac{\mathbb{1}_A}{d_A} \otimes \omega_B) \Big), \tag{55}$$

where the infimum is over all states ω_B in system B. (We will return to the justification of this variational formula after concluding the rest of the proof.) In light of this equality, the difference $S_{M_A}(A|B)_\rho - S_{M_A}(A|B)_\sigma$ is precisely equal to a difference between infimized relative entropies of the form in Theorem 8, in particular with $\mathcal{M} = \mathcal{N} = \phi_{M_A} \otimes I_B$ and $\chi = \{\frac{\mathbb{1}_A}{d_A} \otimes \omega_B \,|\, \omega_B \text{ a state on } B\}$. Thus the Theorem 8 demonstrates the bound (25) holds, and it only remains to evaluate the maximum variation κ.

04 December 2023 10:33:46

© Author(s) 2023

To determine κ it suffices to obtain the bound $0 \le S_{M_A}(A|B)_\rho \le \log d_A$ for all ρ. From non-negativity of the relative entropy, the upper bound $S_{M_A}(A|B)_\rho \le \log d_A$ follows immediately. To establish the lower bound we will make use of the following observations:

(i) With Φ_M a measurement channel as in (3), $(\Phi_M \otimes I)\rho = \sum_j |j\rangle\langle j| \otimes p_j\rho_j$ yields a classical-quantum state, with $p_j = \mathrm{tr}(M_j\rho_A)$ the induced probability distribution over M outcomes, and $\rho_j = \mathrm{tr}_A((M_j \otimes \mathbb{1})\rho)/p_j$ some conditional states.
(ii) For any two such classical-quantum states, $D(\rho\|\rho') = D(p_j\|p'_j) + \sum_j p_j D(\rho_j\|\rho'_j)$.
(iii) With $p_j = \mathrm{tr}(M_j\rho_A)$ and $q_j = \mathrm{tr}(M_j)/d_A$, one has $S_M(\rho_A) = \log d_A - D(p_j\|q_j)$ the marginal observational entropy in system A.
(iv) With p_j, ρ_j as in (i), one has $\sum_j p_j\rho_j = \rho_B$ since $\sum_j M_j = \mathbb{1}_A$.
(v) $S_{M_A}(A|B)_\rho$ is concave in ρ.We thus have

$$S_{M_A}(A|B)_\rho = \log d_A - D\left((\Phi_{M_A} \otimes I_B)\,\rho \,\middle\|\, (\Phi_{M_A} \otimes I_B)\left(\frac{\mathbb{1}_A}{d_A} \otimes \rho_B\right)\right) \tag{56}$$

$$= \log d_A - D(p_j\|q_j) - \sum_j p_j D(\rho_j\|\rho_B) \tag{57}$$

$$= S_{M_A}(\rho_A) - S(\rho_B) + \sum_j p_j S(\rho_j). \tag{58}$$

The first equality is by definition, the second uses (i) and (ii), and the third uses (iii) and (iv). Finally, we note that for pure states $|\psi_{AB}\rangle$ we have that $S(\rho_A) = S(\rho_B)$ and therefore that $S_{M_A}(\rho_A) \ge S(\rho_B)$. Thus for pure states we have $S_{M_A}(A|B)_{|\psi\rangle\langle\psi|} \ge 0$. But by concavity this non-negativity extends to convex combinations and so in general it holds that $S_{M_A}(A|B)_\rho \ge 0$.

To complete the proof, note that since $0 \le S_{M_A}(A|B)_\rho \le \log d_A$, the same bounds hold for the infimized relative entropy in (55). Thus the maximum variation κ in Theorem 8 is given by $\kappa = \sup_{\mu,\nu}|Z(\mu) - Z(\nu)| \le \sup_\mu Z(\mu) \le \log d_A$.

It remains to justify the variational formula (55). Applying the same methods used to obtain (58) earlier, (55) can be seen to become $S_{M_A}(\rho_A) + \sum_j p_j S(\rho_j) - \inf_{\omega_B} \mathrm{tr}(-\rho_B \log \omega_B)$. But it follows immediately from $D(\rho_B\|\omega_B) \ge 0$ that $\mathrm{tr}(-\rho_B \log \omega_B) \ge \mathrm{tr}(-\rho_B \log \rho_B)$ for any ω_B, establishing that $\omega_B = \rho_B$ saturates the infimum. □

VII. CONTINUITY WITH RESPECT TO MEASUREMENTS

The preceding sections have demonstrated concrete uniform continuity bounds on observational entropy $S_M(\rho)$ (and related quantities) under variations in the state ρ. An equally important question is that of continuity under changes in the measurement M. With a suitable norm in the space of measurements (namely the diamond norm between measuring channels), can similar bounds be obtained on $|S_M(\rho) - S_{M'}(\rho)|$ for a fixed state?

We will show that although the observational entropy $S_M(\rho)$ is indeed a continuous function of M, no simple concrete bounds of the type obtained earlier can hold. A strong limitation is exhibited by the following example.

Example 13. Let $M = (\Pi, \mathbb{1} - \Pi)$ be a binary measurement, and $\tilde{M} = (\mathbb{1} - \Pi, \Pi)$ its permutation. Consider a convex combination of this binary M with its permutation,

$$M_\lambda = (1-\lambda)M + \lambda\tilde{M}. \tag{59}$$

This has POVM elements $M_\lambda = ([(1-2\lambda)\Pi + 2\lambda(\mathbb{1}/2)], [(1-2\lambda)(\mathbb{1} - \Pi) + 2\lambda(\mathbb{1}/2)])$. It follows that also

$$M_\lambda = (1-2\lambda)\,M + 2\lambda\left[\frac{\mathbb{1}}{2}\right], \tag{60}$$

where $\left[\frac{\mathbb{1}}{2}\right]$ denotes the trivial binary measurement with elements $(\mathbb{1}/2, \mathbb{1}/2)$. This shows that mixing M with λ of its permutation is equivalent to mixing M with 2λ of pure noise.

Observe that measuring channels are convex linear in the measurement, namely $\Phi_{(1-\lambda)M+\lambda N} = (1-\lambda)\Phi_M + \lambda\Phi_N$. Thus the diamond norm distance between channels implementing M_λ and M is bounded by $\frac{1}{2}\|\Phi_{M_\lambda} - \Phi_M\|_\diamond = \frac{1}{2}\|\lambda(\Phi_{\tilde{M}} - \Phi_M)\|_\diamond \le \lambda$ due to the universal bounds on trace distances between states.[34] This provides a strong sense in which M_λ and M are close together for continuity purposes, and ensures in particular that the probability distributions (outcome statistics) induced by M_λ, M on any state differ by no more than λ in the total variational distance (half of the ℓ_1 norm).

Continuing, suppose $\rho = |\psi\rangle\langle\psi|$ is pure and $M = (|\psi\rangle\langle\psi|, \mathbb{1} - |\psi\rangle\langle\psi|)$ measures ρ perfectly. The observational entropy of M and its permutation are equal in general, and in the present case one finds $S_M(\rho) = S_{\tilde{M}}(\rho) = 0$. On the other hand, for the convex combination M_λ one has probabilities and volumes

$$\begin{aligned} p_0 &= 1-\lambda, & V_0 &= 1 + \lambda(d-2), \\ p_1 &= \lambda, & V_1 &= (d-1) - \lambda(d-2), \end{aligned} \tag{61}$$

04 December 2023 10:33:46

© Author(s) 2023

leading to an observational entropy

$$S_{M_\lambda}(\rho) = (1-\lambda)\log\left(\frac{1+\lambda(d-2)}{1-\lambda}\right) + \lambda\ \log\left(\frac{(d-1)-\lambda(d-2)}{\lambda}\right). \tag{62}$$

This can be viewed as Boltzmann and Shannon terms by expanding the logs.

Now the entropy $S_{M_\lambda}(\rho)$ may be used to analyze both the bounded concavity, and directly the continuity, of observational entropy viewed as a function of measurement. First consider the bounded concavity. In the present case we have an entropy concavity difference

$$\left|S_{(1-\lambda)M+\lambda\tilde{M}}(\rho) - [(1-\lambda)S_M(\rho) + \lambda S_{\tilde{M}}(\rho)]\right| = S_{M_\lambda}(\rho) \tag{63}$$

since $M_\lambda = (1-\lambda)M + \lambda\tilde{M}$ with the $M, \tilde{M}$ each giving zero entropy. And second consider directly the continuity. In the present case we have measurements such that $\frac{1}{2}\|\Phi_{M_\lambda} - \Phi_M\|_\diamond \le \lambda$ and whose entropy difference is

$$|S_{M_\lambda}(\rho) - S_M(\rho)| = S_{M_\lambda}(\rho) \tag{64}$$

since the latter term gives zero. Therefore general bounds on continuity and concavity can be no better than in this particular case.

We make the following observations {with $S_\lambda \equiv S_{M_\lambda}(\rho)$ and $d \ge 2$ and $\lambda \in [0, 1/2]$}:

- Any $\lambda > 0$ implies $S_\lambda > 0$. For fixed d, as $\lambda \to 0$ also $S_\lambda \to 0$, consistent with continuity.
- $S_{1/2} = \log d$, so (63) shows that the naive concavity bound $H(\lambda)$ that held for states [cf. (18)] can be violated by a maximum amount.
- For fixed $\lambda > 0$, one can achieve $S_\lambda/\log d$ arbitrarily close to 1 by taking d sufficiently large. This shows that there is no continuous $f(\lambda)$ such that $f(0) = 0$ and $S_\lambda \le f(\lambda)\log d$.

The final observation implies a strong limitation on asymptotic continuity bounds for measurements: there cannot be any continuous function $f(\lambda)$ with $f(0) = 0$ such that

$$|S_N(\rho) - S_M(\rho)| \le f(\lambda)\log d \tag{65}$$

for all $\frac{1}{2}\|\Phi_N - \Phi_M\|_\diamond \le \lambda$. This form includes all bounds of the type considered earlier for continuity of various entropies with respect to states.

Moreover, one might have hoped to approach the measurement continuity problem similarly to the method used for states, by attempting some form of an AFW trick (modified for channels) based on bounded concavity. However, the example shows that not only does the naive concavity bound by $H(\lambda)$ fail, but also (likewise with the continuity above) that no concavity bound of the form $f(\lambda)\log d$ is possible.

Due to these limitations, at present we settle for the following more abstract statement of continuity over measurements. We give here a general statement containing observational entropy as a special case.

Theorem 14. *Let ρ, σ be states such that $D(\rho\|\sigma) < \infty$. Then $F(\Phi) = D(\Phi(\rho)\,\|\,\Phi(\sigma))$ is a uniformly continuous function of the quantum channel Φ.*

Proof. We are in finite dimension, so the quantum channels form a finite-dimensional convex set and all norms induce the same topology. For concreteness we use the diamond norm. To show continuity we have to show that for any Φ and any $\delta > 0$, there exists an $\epsilon > 0$ such that $\|\Phi - \Phi'\|_\diamond < \epsilon$ implies $|F(\Phi) - F(\Phi')| < \delta$.

To begin, for any channel Φ define $\Phi_s = (1-s)\Phi + s\tau$, which mixes Φ with noise, letting $\tau(X) = \mathrm{tr}(X)\mathbb{1}/d$ prepare maximally mixed state. Observe that $F(\Phi)$ is convex in Φ and is finite and non-negative, and that $F(\tau) = 0$. It follows that $F(\Phi) \ge (1-s)F(\Phi) \ge F(\Phi_s)$. Therefore for any Φ one obtains $|F(\Phi) - F(\Phi_s)| \le sF(\Phi)$.

Additionally, for any Φ, the states $\Phi_s(\rho)$ and $\Phi_s(\sigma)$ are positive definite. The set $\mathcal{P}$ of positive definite matrices is an open set in the 2-norm matrix topology (the usual Euclidean topology on matrix elements). Over the domain $\mu, \nu \in \mathcal{P}$, the function $D(\mu\|\nu)$ is continuous, by the elementary reasoning that it is a composition of matrix multiplication, linear combination, trace, and log, which are all continuous functions on the positive set $\mathcal{P}$. Thus for any $\delta_s > 0$ there exists an $\epsilon_s > 0$ such that if both $\|\Phi_s(\rho) - \Phi_s'(\rho)\|_2 < \epsilon_s$ and $\|\Phi_s(\sigma) - \Phi_s'(\sigma)\|_2 < \epsilon_s$, then $|F(\Phi_s) - F(\Phi_s')| < \delta_s$. This ϵ_s must be chose sufficiently small so that both ϵ_s-balls lie entirely in $\mathcal{P}$.

Meanwhile, for channels Φ, Φ' it holds $\|\Phi_s - \Phi_s'\|_\diamond = \|(1-s)(\Phi - \Phi')\|_\diamond \le \|\Phi - \Phi'\|_\diamond$. And therefore for any state ν we have $\|\Phi_s(\nu) - \Phi_s'(\nu)\|_2 \le \|\Phi_s(\nu) - \Phi_s'(\nu)\|_1 \le \|\Phi_s - \Phi_s'\|_\diamond \le \|\Phi - \Phi'\|_\diamond$.

Now we combine the ingredients. For any $1 > s > 0$, $\delta_s > 0$ let ϵ_s be as above, and suppose $\|\Phi - \Phi'\|_\diamond < \epsilon_s$. Observe that

$$|F(\Phi) - F(\Phi')| = |F(\Phi) - F(\Phi_s) + F(\Phi_s) - F(\Phi_s') + F(\Phi_s') - F(\Phi')| \tag{66}$$

$$\le |F(\Phi) - F(\Phi_s)| + |F(\Phi_s) - F(\Phi_s')| + |F(\Phi_s') - F(\Phi')| \tag{67}$$

$$\le sF(\Phi) + \delta_s + sF(\Phi') \tag{68}$$

$$\le \delta_s + 2sD(\rho\,\|\,\sigma). \tag{69}$$

04 December 2023 10:33:46

© Author(s) 2023

The first inequality is triangle inequality. The second follows from the observations listed above. The last is relative entropy monotonicity.

Finally, choose a $\delta > 0$. Since we assumed $D(\rho\|\sigma) < \infty$, we can choose $s > 0$ sufficiently small so that $2sD(\rho\|\sigma) < \delta$. Next choose any $\delta_s > 0$ sufficiently small so that also $2sD(\rho\|\sigma) + \delta_s < \delta$. There exists an $\epsilon_s > 0$ as above so that $\left\|\Phi - \Phi'\right\|_\diamond < \epsilon_s$ implies $|F(\Phi_s) - F(\Phi'_s)| < \delta_s$ and therefore also that $|F(\Phi) - F(\Phi')| < \delta$. This completes the proof of continuity. The finite dimensional quantum channels form a compact set, thus continuity of F implies it is also uniformly continuous. One can also argue directly (but with more technical difficulty) for uniform continuity by fixing $s = \delta/4D(\rho\|\sigma)$ and using uniform continuity of log on the domain $[s/d, \infty)$. □

Observational entropy is determined by $D(\Phi(\rho)\|\Phi(\sigma))$ in the case $\sigma = \mathbb{1}/d$ and Φ a measuring channel. Since $D(\rho \,\|\, \mathbb{1}/d) < \infty$ always, this implies that (in the topology induced by the diamond norm on Φ_M) observational entropy $S_M(\rho)$ is a continuous function of the measurement. In this form the continuity is useful only for measurements on the same outcome set.

The continuity under measurements can be extended to measurements on arbitrary outcome sets by defining an appropriate topology. Such a topology can be induced by a *simulation distance* defined (for POVMs M, N) by

$$\gamma(M,N) = \frac{\vec{\gamma}(M,N) + \vec{\gamma}(N,M)}{2}, \qquad \vec{\gamma}(M,N) = \inf_\Lambda \frac{1}{2}\|\Phi_{\Lambda M} - \Phi_N\|_\diamond, \tag{70}$$

where the infimum defining the "one-way" simulation distance $\vec{\gamma}(M,N)$ is over stochastic maps Λ from the M outcome set to the N outcome set, so that ΛM and N are measurements with the same set of outcomes. The function $\gamma(M,N)$ can be shown to obey the axioms of a (pseudo-) metric, and therefore induces a topology on the space of POVMs. This distance captures how well two POVMs can approximate one another using classical postprocessing by stochastic channels—if both can emulate each other perfectly by postprocessing alone, they are effectively the same POVM. In terms of this topology we have continuity of measured relative entropy D_M with respect to M, and therefore also of S_M the observational entropy.

Corollary 15. Let ρ, σ be states such that $D(\rho\|\sigma) < \infty$. Then $F(M) = D_M(\rho\|\sigma)$ is a continuous function of the measurement M in the simulation distance topology. This also implies continuity of $F_0(M) = S_M(\rho)$ as a particular case. The continuity is uniform (at least for $d < \infty$).

Proof. Fix $\delta > 0$ and let $\epsilon = \epsilon(\delta)$ as supplied by Theorem 14. Choose any c with $\epsilon > c > 0$, and let $\epsilon' = (\epsilon - c)/4 > 0$. Suppose $\gamma(M,N) < \epsilon'$. From the definition of γ, there exist Λ, Λ' such that $\|\Phi_{\Lambda M} - \Phi_N\|_\diamond < \epsilon$ and $\|\Phi_{\Lambda' N} - \Phi_M\|_\diamond < \epsilon$. Thus $|D_{\Lambda M} - D_N| < \delta$ and $|D_{\Lambda' N} - D_M| < \delta$. We then have $D_M \geq D_{\Lambda M} > D_N - \delta$ and also $D_M < D_{\Lambda' N} + \delta \leq D_N + \delta$, which follow using relative entropy monotonicity, and thus $|D_M - D_N| < \delta$. This completes the proof.

Put more simply, we stated that for small $\gamma(M,N)$ we have $D_N \geq D_{\Lambda' N} \approx D_M \geq D_{\Lambda M} \approx D_N$, which implies $D_M \approx D_N$, where $\geq$ are by monotonicity and $\approx$ are by the earlier continuity. □

VIII. CONCLUDING REMARKS

The main Theorem 6 described a continuity bound (33) on observational entropy that is universal, in the sense of being independent of the POVM M defining the coarse-graining, and with a dimensional factor $\log d$ of the Hilbert space dimension.

This bound improves upon the naive bound (5) insofar as it scales only with $\log d$ independently of the number of measurement outcomes $|M|$. This reflects the intuitive status of observational entropy as a classical entropy in quantum systems, compared, for example to the Shannon entropy of an observable. The improved universality does come with a small trade-off, as when M is a projective measurement in a complete orthonormal basis, the naive bound (5) is slightly stronger. We note that (33) can only be relevant when ρ, σ have overlapping support. In case the supports are orthogonal, the trace distance $\epsilon = 1$ is maximized, and (33) is weaker than the trivial $|\Delta S_M| \leq \log d$. This demonstrates that (33) cannot be tight in general—for fixed ϵ, there is not generally a choice of ρ, σ, M for which the inequality is saturated.

Meanwhile, in the case of continuity with respect to measurements, it was seen that although observational entropy is continuous, explicit bounds on the convergence are difficult (if not impossible) to obtain. Other notable limitations on the present results were the necessity of a convex reference set for measured relative entropy continuity (as discussed earlier), and the restriction throughout the paper to finite dimensional spaces. Extending our results to the infinite dimensional case encounters difficulties associated with the possibility of infinite Boltzmann and von Neumann entropies, and doing so carefully will be a useful continuation.

One can also consider a classical version of the observational entropy development for density distributions ρ, σ on phase space Γ. The trace is an integral on Γ, the norm the L^1 norm, and the measurement M a partition of unity. The dimensional factor d (better labelled V in the classical case) is the total volume of phase space in an appropriately chosen measure. The resulting bound (33) holds equally well in the classical case so long as one can restrict consideration to a finite region of phase space.

As with the quantum case, the classical (33) can be nontrivial (stronger than $|\Delta S_M| \leq \log V$) only when ρ, σ have support on an overlapping region of phase space. In this sense, the closeness of entropy ensured by the bound arises only from a probability to be in exactly the same state. Further one may observe that the bound will be time independent for isolated systems, whose dynamics are norm preserving. This can be considered in terms of the second law of thermodynamics for coarse-grained entropies. Suppose ρ is concentrated at some particular "second-law-violating" point in phase space, one which evolves to a lower entropy state. Any σ sufficiently close to ρ in the norm distance must also be second-law-violating, which at first glance seems to violate the principle that such states are rare. Fortunately, by the above considerations,

04 December 2023 10:33:46

© Author(s) 2023

σ may be concentrated in any phase space region disjoint from ρ, even arbitrarily nearby, without any restriction on its entropic dynamics. In this way the bound is consistent with the principle that nearby phase space points of a chaotic system may have drastically different dynamics.

ACKNOWLEDGMENTS

The authors thank Ludovico Lami for comments regarding the measured relative entropy continuity and the minimax lemma, and Niklas Galke, Philipp Strasberg, and Giulio Gasbarri for helpful discussions.

The authors acknowledge support by MICIIN with funding from European Union NextGenerationEU (Grant No. PRTR-C17.I1) and by Generalitat de Catalunya. AW is furthermore supported by the European Commission QuantERA grant ExTRaQT (Spanish MICINN Project No. PCI2022-132965), by the Spanish MINECO (Project No. PID2019-107609GB-I00) with the support of FEDER funds, the Generalitat de Catalunya (Project No. 2017-SGR-1127), by the Alexander von Humboldt Foundation, as well as the Institute for Advanced Study of the Technical University Munich.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Joseph Schindler: Conceptualization (equal); Formal analysis (equal); Investigation (equal); Writing – original draft (equal); Writing – review & editing (equal). **Andreas Winter**: Conceptualization (equal); Formal analysis (equal); Investigation (equal); Writing – review & editing (equal).

DATA AVAILABILITY

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

REFERENCES

[1] J. von Neumann, "Proof of the ergodic theorem and the H-theorem in quantum mechanics," Eur. Phys. J. H **35**, 201 (2010), Translation of: Beweis des Ergodensatzes und des H-Theorems in der neuen Mechanik; arXiv:1003.2133 [physics.hist-ph].
[2] N. G. Van Kampen, "Quantum statistics of irreversible processes," Physica **20**, 603 (1954).
[3] J. von Neumann, *Mathematical Foundations of Quantum Mechanics* (Princeton University Press, 1955).
[4] A. Wehrl, "General properties of entropy," Rev. Mod. Phys. **50**, 221 (1978).
[5] D. Šafránek, J. M. Deutsch, and A. Aguirre, "Quantum coarse-grained entropy and thermodynamics," Phys. Rev. A **99**, 010101 (2019); arXiv:1707.09722 [quant-ph].
[6] D. Šafránek, A. Aguirre, J. Schindler, and J. M. Deutsch, "A brief introduction to observational entropy," Found. Phys. **51**, 101 (2021); arXiv:2008.04409 [quant-ph].
[7] P. Strasberg and A. Winter, "First and second law of quantum thermodynamics: A consistent derivation based on a microscopic definition of entropy," PRX Quantum **2**, 030202 (2021); arXiv:2002.08817 [quant-ph].
[8] D. Šafránek, J. M. Deutsch, and A. Aguirre, "Quantum coarse-grained entropy and thermalization in closed systems," Phys. Rev. A **99**, 012103 (2019); arXiv:1803.00665 [quant-ph].
[9] D. Šafránek, A. Aguirre, and J. M. Deutsch, "Classical dynamical coarse-grained entropy and comparison with the quantum version," Phys. Rev. E **102**, 032106 (2020); arXiv:1905.03841 [cond-mat.stat-mech].
[10] D. Šafránek and J. Thingna, "Quantifying information extraction using generalized quantum measurements," Phys. Rev. A **108**, 032413 (2023).
[11] A. Wehrl, "On the relation between classical and quantum-mechanical entropy," Rep. Math. Phys. **16**, 353 (1979).
[12] A. Riera-Campeny, A. Sanpera, and P. Strasberg, "Quantum systems correlated with a finite bath: Nonequilibrium dynamics and thermodynamics," PRX Quantum **2**, 010340 (2021); arXiv:2008.02184 [quant-ph].
[13] P. Strasberg, M. G. Díaz, and A. Riera-Campeny, "Clausius inequality for finite baths reveals universal efficiency improvements," Phys. Rev. E **104**, L022103 (2021); arXiv:2012.03262 [quant-ph].
[14] F. Buscemi, J. Schindler, and D. Šafránek, "Observational entropy, coarse quantum states, and Petz recovery: Information-theoretic properties and bounds," New J. Phys. **25**, 053002 (2022); arXiv:2209.03803 [quant-ph].
[15] L. Amadei, H. Liu, and A. Perez, "Unitarity and information in quantum gravity: A simple example," Front. Astron. Space Sci. **8**, 46 (2021); arXiv:1912.09750 [gr-qc].
[16] X. Zhou and Z.-J. Zheng, "Relations between the quantum correlation entropy and quantum discord for X-states in multipartite systems," Eur. Phys. J. Plus **137**, 625 (2022).
[17] X. Zhou and Z.-J. Zheng, "Relations between the observational entropy and Rényi information measures," Quantum Inf. Process. **21**, 228 (2022).
[18] R. Hamazaki, "Speed limits for macroscopic transitions," PRX Quantum **3**, 020319 (2022); arXiv:2110.09716 [cond-mat.stat-mech].
[19] D. Faiez, D. Šafránek, J. M. Deutsch, and A. Aguirre, "Typical and extreme entropies of long-lived isolated quantum systems," Phys. Rev. A **101**, 052101 (2020); arXiv:1908.07083 [quant-ph].
[20] R. Modak and S. Aravinda, "Observational entropic study of Anderson localization," Phys. Rev. A **106**, 062217 (2022); arXiv:2209.10273 [quant-ph].
[21] A. Stokes, "Nonconjugate quantum subsystems," Phys. Rev. E **106**, 034111 (2022); arXiv:2106.11017 [quant-ph].

04 December 2023 10:33:46

© Author(s) 2023

[22] P. Strasberg, *Quantum Stochastic Thermodynamics: Foundations and Selected Applications* (Oxford University Press, 2022).
[23] P. Strasberg, A. Winter, J. Gemmer, and J. Wang, "Classicality, Markovianity, and local detailed balance from pure-state dynamics," Phys. Rev. A **108**, 012225 (2022); arXiv:2209.07977 [quant-ph].
[24] D. Šafránek, D. Rosa, and F. Binder, "Work extraction from unknown quantum sources," Phys. Rev. Lett. **130**, 210401 (2022); arXiv:2209.11076 [quant-ph].
[25] S. PG, R. Modak, and S. Aravinda, "Witnessing quantum chaos using observational entropy," Phys. Rev. E **107**, 064204 (2022); arXiv:2212.01585 [quant-ph].
[26] X. Zhou, "Dynamical behavior of quantum correlation entropy under the noisy quantum channel for multiqubit systems," Int. J. Theor. Phys. **62**, 23 (2023).
[27] D. Šafránek and D. Rosa, "Expectation values from any quantum measurement," Phys. Rev. A **108**, 022208 (2023); arXiv:2301.10428 [quant-ph].
[28] M. Fannes, "A continuity property of the entropy density for spin lattice systems," Commun. Math. Phys. **31**, 291 (1973).
[29] M. Grabowski and P. Staszewski, "On continuity properties of the entropy of an observable," Rep. Math. Phys. **11**, 233 (1977).
[30] K. M. R. Audenaert, "A sharp continuity estimate for the von Neumann entropy," J. Phys. A: Math. Theor. **40**, 8127 (2007); arXiv:quant-ph/0610146 [quant-ph].
[31] K. M. R. Audenaert and J. Eisert, "Continuity bounds on the quantum relative entropy," J. Math. Phys. **46**, 102104 (2005); arXiv:quant-ph/0503218 [quant-ph].
[32] A. Winter, "Tight uniform continuity bounds for quantum entropies: Conditional entropy, relative entropy distance and energy constraints," Commun. Math. Phys. **347**, 291 (2016); arXiv:1507.07775 [quant-ph].
[33] Z. Zhang, "Estimating mutual information via Kolmogorov distance," IEEE Trans. Inf. Theory **53**, 3280 (2007).
[34] M. M. Wilde, *Quantum Information Theory* (Cambridge University Press, 2013); arXiv:1106.1445v8.
[35] A. S. Holevo and R. F. Werner, "Evaluating capacities of Bosonic Gaussian channels," Phys. Rev. A **63**, 032312 (2001); arXiv:quant-ph/9912067 [quant-ph].
[36] R. Alicki and M. Fannes, "Continuity of quantum conditional information," J. Phys. A: Math. Gen. **37**, L55 (2004); arXiv:quant-ph/0312081 [quant-ph].
[37] M. Mosonyi and F. Hiai, "On the quantum Rényi relative entropies and related capacity formulas," IEEE Trans. Inf. Theory **57**, 2474 (2011); arXiv:0912.1286 [quant-ph].
[38] M. J. Donald and M. Horodecki, "Continuity of relative entropy of entanglement," Phys. Lett. A **264**, 257 (1999); arXiv:quant-ph/9910002 [quant-ph].
[39] B. Synak-Radtke and M. Horodecki, "On asymptotic continuity of functions of quantum states," J. Phys. A: Math. Gen. **39**, L423 (2006); arXiv:quant-ph/0507126 [quant-ph].
[40] M. Piani, "Relative entropy of entanglement and restricted measurements," Phys. Rev. Lett. **103**, 160504 (2009); arXiv:0904.2705 [quant-ph].
[41] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge University Press, 2010).
[42] K. Li and A. Winter, "Relative entropy and squashed entanglement," Commun. Math. Phys. **326**, 63 (2014); arXiv:1210.3181 [quant-ph].
[43] M. E. Shirokov, "Advanced Alicki–Fannes–Winter method for energy-constrained quantum systems and its use," Quantum Inf. Process. **19**, 164 (2020); arXiv:1907.02458 [quant-ph].
[44] A. Bluhm, Á. Capel, P. Gondolf, and A. Pérez-Hernández, "Continuity of quantum entropic quantities via almost convexity," IEEE Trans. Inf. Theory **69**, 5869 (2022); arXiv:2208.00922 [quant-ph].
[45] A. Bluhm, Á. Capel, P. Gondolf, and A. Pérez-Hernández, "General continuity bounds for quantum relative entropies," arXiv:2305.10140 [quant-ph] (2023).
[46] M. E. Shirokov, "Quantifying continuity of characteristics of composite quantum systems," Phys. Scr. **98**, 042002 (2023); arXiv:2201.11477 [quant-ph].
[47] F. G. S. L. Brandão, A. W. Harrow, J. R. Lee, and Y. Peres, "Adversarial hypothesis testing and a quantum Stein's lemma for restricted measurements," IEEE Trans. Inf. Theory **66**, 5037 (2020); arXiv:1308.6702.
[48] B. Farkas and S. G. Révész, "Potential theoretic approach to rendezvous numbers," Monatsh. Math. **148**, 309 (2006); arXiv:math/0503423 [math.CA].
[49] M. Sion, "On general minimax theorems," Pac. J. Math. **8**, 171 (1958).
[50] L. Lami, private communication (2022).
[51] P. J. Coles, M. Berta, M. Tomamichel, and S. Wehner, "Entropic uncertainty relations and their applications," Rev. Mod. Phys. **89**, 015002 (2017); arXiv:1511.04857 [quant-ph].

04 December 2023 10:33:46

© Author(s) 2023