Relative thermalization

Lídia del Rio,[1,2,*] Adrian Hutter,[2,3,4] Renato Renner,[2] and Stephanie Wehner[4,5]
[1]*School of Physics, University of Bristol, BS8 1TL Bristol, United Kingdom*
[2]*Institute for Theoretical Physics, ETH Zurich, 8093 Zurich, Switzerland*
[3]*Department of Physics, University of Basel, 4056 Basel, Switzerland*
[4]*Centre for Quantum Technologies, National University of Singapore, 117543 Singapore*
[5]*QuTech, Delft University of Technology, 2628 CJ Delft, The Netherlands*
(Received 14 November 2015; revised manuscript received 4 March 2016; published 3 August 2016)

Locally thermal quantum systems may contradict traditional thermodynamics: heat can flow from a cold body to a hotter one, if the two are highly entangled. We show that to recover thermodynamic laws, we must use a stronger notion of thermalization: a system S is *thermal relative to a reference* R if S is both locally thermal and uncorrelated with R. Considering a general quantum reference is particularly relevant for a thermodynamic treatment of nanoscale quantum systems. We derive a technical condition for relative thermalization in terms of conditional entropies. Established results on local thermalization, which implicitly assume a classical reference, follow as special cases.

DOI: 10.1103/PhysRevE.94.022104

I. MOTIVATION

The *fundamental postulate of statistical physics* is the assumption that systems in contact with an environment equilibrate to a *thermal state* of maximum entropy. Despite its name, this equilibration is not a fundamental law of nature, but rather an emerging characteristic of quantum systems undergoing typical evolutions. More precisely, consider a system S interacting with an environment E. The two systems may be subject to a physical constraint, like energy conservation. Our knowledge of that constraint is expressed by the subspace $\Omega \subseteq S \otimes E$ of allowed states (for instance, Ω could be an energy shell). A naïve reading of the postulate could be that Ω equilibrates to the so-called microcanonical state, $\pi_\Omega := \mathbb{1}_\Omega/|\Omega|$ (here, $|\Omega|$ denotes the dimension of Hilbert space Ω). This is the state that maximizes the entropy of Ω; in other words, a state of maximal ignorance. However, such equilibration is forbidden by the reversibility of time evolution, if $S \otimes E$ is treated as a closed system. For example, an initially pure quantum state cannot evolve unitarily to a more mixed state. Instead, we may look for local subsystem thermalization: S might equilibrate to state $\pi_S = \mathrm{Tr}_E(\pi_\Omega)$, even though the global state of Ω is not equilibrated. In many natural settings, for instance if the environment is large and the Hamiltonian is fully interactive, local thermalization is *typical*, in the sense that S will be approximately in state π_S for most of the time [1–4]. If some extra conditions are satisfied (like weak coupling between S and E), π_S approximates the familiar Gibbs state [5]. Indeed, a small system S appears to thermalize locally because typical evolutions leave it highly correlated (entangled) with the environment, and so the reduced state of S becomes very mixed|a thermal state. For reviews on quantum thermalization, see [6,7].

But simply knowing that a system is thermalized does not allow us to treat it as a heat bath towards any other system, as the curious example of Fig. 1 illustrates. Consider two systems H and C, each in a local thermal state (their reduced states are Gibbs states of different temperatures, π_H and π_C). If we put the two systems in thermal contact, we would expect heat to flow from the hotter bath, H, to the colder one, C. However, if H and C are highly entangled, one can observe an anomalous heat flow from C to H [8–10]. The clue to understand this phenomenon is that H and C are not truly heat baths *relative to each other*, because they are correlated. Nevertheless, H can still act as a normal heat bath towards a different reference system R, provided that they are not initially correlated ($\rho_{HR} = \pi_H \otimes \rho_R$).

Clausius' formulation of the second law of thermodynamics states that heat cannot flow from cold to hot bodies [11]. When this law was originally proposed, there was no microscopic model to suggest the possibility of correlations between such systems. Even today, although we have quantum models for these physical bodies, and quantum correlations have been extensively studied, Clausius' law is implicitly interpreted as "whenever two systems in local thermal states are put in contact, heat cannot flow from the colder system to the hotter one." This reading, however, cannot be correct, given the violation brought about by anomalous heat flows. In order to clarify its meaning, Clausius' law could be reformulated as "whenever two systems which are *thermal relative to each other* are put in contact, heat will not flow from the colder to the hotter body." Let us formalize what we mean by this relative thermalization.

Add to the previous setting a third quantum system, the reference R. The global state is $\rho_{\Omega R}$. We say that S is thermalized relative to the reference R if $\rho_{SR} := \mathrm{Tr}_E(\rho_{\Omega R}) = \pi_S \otimes \rho_R$. Often we can only approximately estimate states, so we may generalize this definition to say that S is δ-thermalized relative to R if ρ_{SR} is δ-close to the above state, according to the trace distance,

$$\tfrac{1}{2}\,\|\rho_{SR} - \pi_S \otimes \rho_R\|_1 \leqslant \delta.$$

Note in particular that the same system S may be thermalized relative to a reference, but not another (see Fig. 2 for two examples). This definition, stronger than the usual ones, forces us to revisit the standard arguments that predict local

*lidia.delrio@bristol.ac.uk

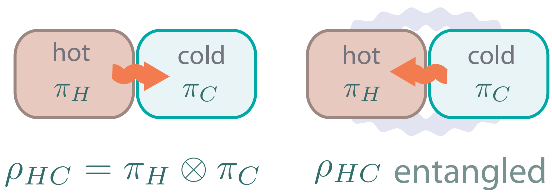

FIG. 1. Anomalous heat flow. If two thermal bodies are put in contact, heat normally flows from the hotter body to the colder one. However, it could be that the two systems are correlated, while still presenting local thermal states. If those correlations are strong enough (for instance if they are highly entangled), heat may flow from the colder to the hotter body. There is no contradiction with the second law, if one formulates it in terms of relative thermalization, because the two bodies are not thermal relative to each other.

thermalization: do they also lead to relative thermalization? In other words, is relative thermalization a typical phenomenon in nature, or an unattainable idealization? Our technical result addresses this question, showing that if a system were to thermalize locally, it is also very likely to thermalize relative to a reference—unless that reference is very entangled with the system and its environment.

II. RESULTS

Formally, we study conditions for relative thermalization of S in the setting of Fig. 3(a). We want to know what kind of initial states $\rho_{\Omega R}$ and physical evolutions in Ω lead to the

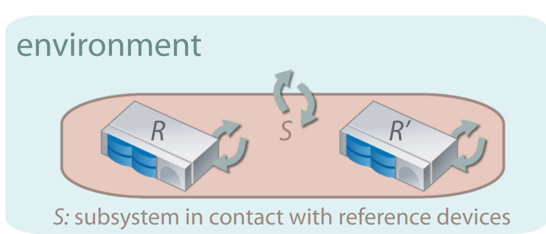

FIG. 2. Local and relative thermalization. The same system S may be thermalized relative to a reference R' but not another, R. For instance, imagine that R and R' are the memories of two observers who measured S. In the case of state $\rho_{SRR'} = (\sum_{i=1}^{|\Omega|} \frac{1}{|\Omega|} \operatorname{Tr}_E |i\rangle\langle i|_\Omega \otimes |i\rangle\langle i|_R) \otimes |0\rangle\langle 0|_{R'}$, S is locally thermalized, with $\rho_S = \pi_S$. Here, the observer with memory R' measured a few macroscopic parameters of $S \otimes E$, enough only to determine the subspace Ω; the reduced state of $S \otimes R'$ is precisely $\pi_S \otimes \rho_{R'}$. A second observer, with more precise measurement instruments, may determine the exact state $|i\rangle$ of $S \otimes E$, and write it down on the memory R. Although S is locally thermalized, S and R are classically correlated. In a more critical example, suppose that the global state is $\rho_{SRR'} = \operatorname{Tr}_E |\Psi\rangle\langle\Psi|_{\Omega R} \otimes \rho_{R'}$, where $|\psi\rangle$ is entangled between Ω and R, $|\Psi\rangle_{\Omega R} = |\Omega|^{-\frac{1}{2}} \sum_{i=1}^{|\Omega|} |i\rangle_\Omega \otimes |i\rangle_R$. Here again S is locally thermalized, and also thermalized towards R', but it is entangled with R. This difference has actual physical consequences: in that limit, a joint evolution of S and R' will likely increase the entropy of the reference R', because π_S is a very mixed state. On the other hand, no global evolution of S and R can increase the entanglement between the two, and therefore the entropy of R—not the typical effect of a thermal bath. In other words, S acts as a source of random noise towards R', but not towards R.

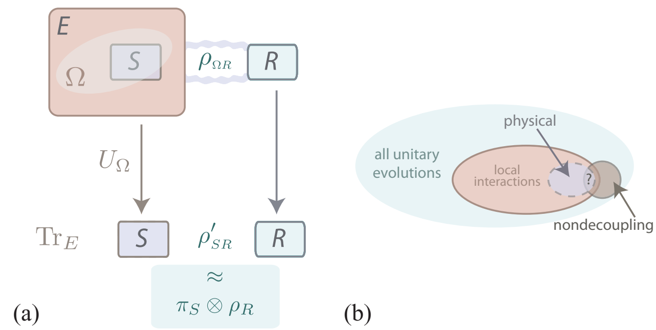

FIG. 3. Setting and nature of our results. A system $S \otimes E$ is subject to a physical constraint Ω. This system may be correlated with a reference R, in an arbitrary global state $\rho_{\Omega R}$. We let $\rho_{\Omega R}$ evolve under a unitary U_Ω acting on Ω, and then ask if S is (approximately) thermalized relative to the reference R (a). We derive an entropic condition on $\rho_{\Omega R}$ and Ω, (1), that guarantees that most unitaries, according to the Haar measure, lead to relative thermalization. This means that, under that condition, if all we know about U_Ω is that it is a unitary in Ω, it is highly likely, from our point of view, that U_Ω will thermalize S relative to R. Usually, though, we know more about U_Ω, for instance, that it is induced by a given local Hamiltonian. As the set of all unitaries in Ω is full of operators that are unrelated to our physical setting (like nonlocal evolutions, ruled out by our knowledge), it is desirable to obtain similar probabilistic statements about smaller sets that still contain U_Ω, like those generated by local interactions (b). This is possible, because the decoupling approach [12–15] used to obtain our results is very general, and can be applied to more physical sets of unitaries, consisting of local two-body interactions [16–18], or time-independent Hamiltonians [3,4,16,18–20]. See the Methods Summary for further discussion.

thermalization of subsystem S relative to the reference R. We show below that, if an entropic condition is satisfied, only an exponentially small fraction of evolutions in Ω do not lead to relative thermalization [Fig. 3(b)]. This formulation includes the traditional case of a classical reference; in that special case, we recover known results on local thermalization [1]. See Fig. 4 for an illustration of our results, applied to systems of weakly interacting spins.

Theorem 1. Let $\rho_{\Omega R}$ be a quantum state in $\Omega \otimes R$, with $\Omega \subseteq S \otimes E$, and let $\pi_\Omega = \frac{\mathbb{1}_\Omega}{|\Omega|}$. Let $\varepsilon,\delta > 0$. If

$$H^{9\varepsilon}(\Omega|R)_\rho \gtrsim H^{1-\varepsilon}(S)_\pi - H^\varepsilon(E)_\pi, \tag{1}$$

then a unitary evolution U_Ω of ρ will typically lead to δ-thermalization of S relative to R.

More precisely, the fraction of unitary evolutions in Ω that do not lead to δ-thermalization, according to the Haar measure, is exponentially small in δ^2 and in the dimension of Ω.

Before we give a physical interpretation to the theorem, let us introduce its protagonist: the *smooth conditional entropy* $H^\varepsilon(\Omega|R)_\rho$, which is a generalization of Boltzmann's entropy for single-shot quantum settings [21]. Smooth entropy measures are widely used in the context of quantum information processing; they measure our ignorance about the exact state of Ω, given access to side information stored in a reference R, which may be correlated with Ω in state $\rho_{\Omega R}$. The smoothing parameter $\varepsilon \in [0,1]$ accounts for our error

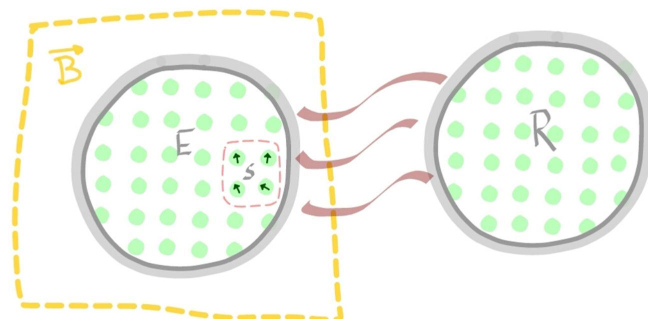

FIG. 4. Application to spin systems. Consider a system of N weakly interacting spins, subject to the Hamiltonian $\hat{H} = \hat{H}_0 + \hat{V}$, where $\hat{H}_0 = J \sum_i |\uparrow\rangle\langle\uparrow|_i$ and $\hat{V}$ is a random nearest-neighbor perturbation that conserves the total spin (with $|\hat{V}| \ll |\hat{H}_0|$); this system is also studied in the preprint version of Ref. [1]. We select $\alpha\, N$ of those spins to be our subsystem S, while the remaining $(1-\alpha)N$ spins are called the environment E. In addition, the spins of $S \otimes E$ may be correlated with a reference spin system R. We want to study thermalization of S relative to R, for an arbitrary initial state ρ_{SER}. Note that the energy subspaces of $S \otimes E$ are invariant under time evolution ruled by $\hat{H}$; therefore we will look at states that lie in one of these invariant subspaces. For mixtures and superpositions over different subspaces, the results follow by linearity. Each energy shell $\{\Omega_k\}$ is generated by states with a fixed number k of spins up, $\Omega_k = \text{span}\{|\Psi\rangle_{SE} : \hat{H}_0 |\Psi\rangle = k\, J\, |\Psi\rangle\}$. The initial state is ρ_{SER}, with $\rho_{SE} \in \text{End}(\Omega_k)$ for some k. We apply the weaker condition for relative thermalization, $H^{\varepsilon}(\Omega|R)_\rho > \log_2|S| - \log_2|\Omega|$. The dimension of S is $2^{\alpha N}$, while $|\Omega_k| = \binom{N}{k}$. For large N, $\log_2\binom{N}{k} \approx N\, h(k/N)$, where $h(p) = -p\, \log_2 p - (1-p)\log_2(1-p)$ is the binary entropy of p. We obtain the following result: if $H^{\varepsilon}(\Omega|R)_\rho > N[2\alpha - h(k/N)]$, then S will be δ-thermalized relative to R after most evolutions. Conversely, if $H^{\varepsilon}(\Omega|R)_\rho < -H^1(E)_\pi$, then no evolution in Ω_k leads to relative thermalization of S (Theorem 2; here $H^1(E)_\pi = \log_2[\text{Supp}\,[\text{HypergeometricDistribution}[\alpha N, k N, N]]]$).

tolerance|that is, our willingness to ignore highly unlikely events, like the spontaneous repair of a shattered glass. In many natural scenarios, we want ε to be small but nonzero (see details in Appendix A). To give an idea of the values that H^{ε} takes, consider the limit $\varepsilon \to 0$. Then, $H^{\varepsilon}(\Omega|R)_\rho$ is at most $\log_2|\Omega|$, which is achieved for the decoupled state $\rho_{\Omega R} = \pi_\Omega \otimes \rho_R$, is zero if ρ_Ω is pure, and becomes negative if Ω and R are entangled, with a minimum at $-\log_2|\Omega|$ for maximally entangled states.

We may now analyze the meaning of Theorem 1. The right-hand side of (1), $H^{1-\varepsilon}(S)_\pi - H^{\varepsilon}(E)_\pi$, depends only on the canonical state π_Ω. That is, these entropies are determined by the structure of the physical constraint Ω, given by factors like the Hamiltonian of $S \otimes E$ and the relative dimensions of those systems. In fact, the whole right-hand side can be approximately replaced by $2\log_2|S| - \log_2|\Omega|$ (see Corollary C.6 in Appendix C). On the left-hand side of (1), we have $H^{9\varepsilon}(\Omega|R)_\rho$, which depends on the global initial state. This term gives us an information-theoretical condition for relative thermalization: if the reference is not highly correlated with $S \otimes E$, then a typical evolution in Ω is likely to "sweep" correlations with S to the environment, leaving S thermalized relative to R.

Sometimes, the reference is so correlated with $S \otimes E$ that no evolution in Ω can decouple S from it (intuitively, the correlations cannot all be "moved" to the environment). Theorem 2 characterizes states that can never achieve relative thermalization.

Theorem 2. Let $\rho_{\Omega R}$ be a quantum state in $\Omega \otimes R$, with $\Omega \subseteq S \otimes E$, and let $\pi_\Omega = \frac{\mathbb{1}_\Omega}{|\Omega|}$. Let $\delta, \varepsilon > 0$. If

$$H^{\varepsilon}(\Omega|R)_\rho \lesssim -H^1(E)_\pi, \tag{2}$$

then no unitary evolution of ρ in Ω can lead to δ-thermalization of S relative to R.

Note that (2) is close to a converse of the direct bound (1), in the thermodynamic limit of a large environment E and small subsystem S, when it is reasonable to neglect a term of the order $\log_2|S|$. In other words, the conditions for relative thermalization are tight in this typical setting.

Technically, our results rely on decoupling [12–15] and smooth entropy measures [21–23], tools that have been recently developed in the field of quantum information theory, and have powerful applications in quantum cryptography, error correction, and thermodynamics [13,15,24]. Smooth conditional entropies, which quantify the size of subsystems likely to decouple from a reference, are defined in Appendix A and further characterized in Appendix D. Decoupling is introduced in Appendix B, and a technical and stronger version of our results and proofs lie in Appendix C. In particular, we show that for most unitaries U_Ω, the bound of Theorem 1 is tight even when S is large compared to the environment.

III. DISCUSSION

Two factors make the study of relative thermalization particularly relevant today. One is that, unlike previous approaches, it does not assume any classicality of the reference. For classical references, we can describe the global state as a quantum-classical density matrix, $\rho_{\Omega R} = \sum_x p_x\, |x\rangle\langle x|_R \otimes \rho_\Omega^x$. Crucially, this means that for each fixed value x in the reference, we assign a "conditional" density matrix ρ_Ω^x to Ω. We could then read the state $|x\rangle$ of the reference and study local thermalization of a subsystem S for the initial state ρ_Ω^x. Clearly, if S thermalizes, it is also uncorrelated with R. In fact, this is implicitly done in the current literature, when the "initial knowledge" about Ω is mentioned [2]. However, we cannot define these "conditional states" when the reference is itself a quantum system, a more general and natural setting than imposing classicality on the reference|R may be simply a system that has interacted with Ω, and became entangled with it. In order to study the evolution of Ω with respect to R in this general framework, we need to consider their joint density matrix.

There is also a question of scale. Traditionally, thermodynamics deals with large objects, and quantum correlations between systems can be neglected. This is because most degrees of freedom are irrelevant for the macroscopic behavior of a system, or the performance of a heat engine: we are only interested in the average energy of a gas or the position of a piston, while correlations are typically encoded in finer details of the particles' wave functions. However, as modern technologies miniaturize to the nanoscale, a comprehensive understanding of the thermodynamics of small quantum systems is essential to identify and harness their power. As the number of degrees of freedom decreases, correlations become more likely to influence the relevant parameters of an experiment, and can no longer be neglected.

For example, correlations between heat baths have been shown to affect the performance of three-qubit heat engines [25]. These engines only behave like traditional Carnot machines if the baths involved are thermalized relative to each other. Relative thermalization was also found to be crucial to prove Landauer's principle, which quantifies the work cost of information-processing tasks in physical systems [24,26,27]. For Landauer's bound to apply, it is necessary that the system of interest be decoupled from a thermal bath|otherwise we could exploit correlations with this "bath" to extract extra work.

Our results can be naturally applied to settings where conservation laws demand that more than one quantity be preserved [28–31], for example total energy and spin. In this case, we simply take Ω to reflect those global constraints: for example, it can be the subspace of fixed total quantities $\{\langle A_i\rangle\}_i$. In the case where the preserved quantities are local (e.g., noninteracting particles with local Hamiltonian and spin), it was shown that the microcanonical state π_Ω locally looks like a generalized Gibbs state, $\pi_S \propto e^{\sum_i \beta_i A_i}$, with a different notion of inverted temperature β_i for each quantity [31]. These states are also shown to be completely passive under the respective conservation laws, and are therefore a generalization of the usual Gibbs state [28–31]. Our results immediately give conditions for typical relative thermalization towards the generalized Gibbs state π_S.

A natural direction of research now is to approach other aspects of relative thermalization [6,7]. An example is the study of time scales for thermalization of systems with fixed Hamiltonians [4,5], which can be generalized to a setting with an explicit quantum reference. Another example is the apparent thermalization of isolated quantum systems under realistic observables [32,33]—there we may ask whether we can distinguish the actual state of a system Ω from π_Ω after a measurement, if in addition we hold a quantum reference correlated with Ω.

Thermalization results can be strengthened by restricting the class of unitary evolutions to smaller sets, thereby excluding obviously unphysical, nonlocal evolutions [Fig. 3(b)]. One option is given by random local circuits. These model systems like a chain of atoms or a particle gas, where neighboring particles undergo successive two-body unitary evolutions. Local quantum circuits typically achieve decoupling after an initial equilibration period [12,15,16,18–20,34]. Since our results are based on decoupling, it is straightforward to apply them to systems described by local interactions [15,19]. More generally, relative thermalization motivates the search for restricted and physical classes of unitaries that achieve decoupling.

We have seen that relative thermalization of physical systems is necessary to apply traditional thermodynamics; indeed it is the very foundation of resource theories for quantum thermodynamics, used to derive concepts like the free energy and the efficiency of heat engines. The next step is to ask whether some thermodynamic statements can be recovered from independence conditions that are strictly weaker than relative thermalization. These conditions may be studied under general frameworks for resource theories [35].

ACKNOWLEDGMENTS

We thank J. Eisert, P. Faist, A. Fomins, C. Gogolin, P. Kammerlander, L. Krämer, and Y.-C. Liang for discussions and comments on the manuscript. We acknowledge support from the Swiss National Science Foundation (L.d.R. and R.R., Grants No. 200020-135048 and No. 200021-153296 and the NCCR QSIT), the Portuguese Fundação para a Ciência e Tecnologia (L.d.R., Grant No. SFRH/BD/43263/2008), the European Research Council (R.R., Grant No. 258932, and L.d.R., AdG NLST), the Engineering and Physical Sciences Research Council (Grant No. DIQIP), and the National Research Foundation and the Ministry of Education, Singapore (A.H. and S.W.) STW Netherlands (S.W.). We further acknowledge the Singapore programme Mathematical Horizons for Quantum Physics 2, and the COST Action MP1209 "Thermodynamics in the quantum regime" for supporting our collaboration.

APPENDIX A: INTRODUCTION TO SMOOTH ENTROPY MEASURES

A word on notation. We use $\mathcal{S}(A)$ to denote the set of density matrices acting on Hilbert space A, i.e.,

$$\mathcal{S}(A) = \{\rho \in \mathrm{End}(A) : \rho \geqslant 0, \mathrm{Tr}\,\rho = 1\},$$

where $\mathrm{End}(A)$ denotes endomorphisms on A. Similarly, the set of subnormalized positive semidefinite operators ($\rho \geqslant 0$, $\mathrm{Tr}\,\rho \leqslant 1$) is denoted by $\mathcal{S}_\leqslant(A)$. For instance, $\rho_{AB} \in \mathcal{S}(A \otimes B)$ is the (possibly mixed) state of a bipartite quantum system, consisting of subsystems A and B.

The identity operator on Hilbert space A is denoted by $\mathbb{1}_A \in \mathrm{End}(A)$, while the identity map acting on operators of A is denoted by $\mathcal{I}_A \in \mathrm{End}\,(\mathrm{End}(A))$.

For simplicity, we use $U_A \cdot \rho_{AB}$ to denote $[U_A \otimes \mathbb{1}_B]\, \rho_{AB}\, [U_A^\dagger \otimes \mathbb{1}_B]$.

Our results rely on decoupling [12–14], which is tightly characterized by smooth entropies, a natural class of entropies quantifying correlations between quantum systems in single-shot settings. From this class, we choose a particular conditional entropy, sometimes called the hypothesis-testing entropy [21], to express our final results, and we use smooth min- and max entropies [22,23,36–38] throughout the proofs. In this section, we define and characterize these entropy measures.

1. Smooth min and max entropies

Most of our technical proofs use conditional smooth min and max entropies [22,23,36–38]. These have convenient properties, used to derive the final form of our results (for example, duality; see (A15)). For a comprehensive discussion of these entropies, their properties, and applications, we refer to [23].

a. Purified distance

The purified distance [39] is used to smooth the min and max entropies, and is defined for subnormalized states $\rho,\sigma \in$

$\mathcal{S}_{\leqslant}(A)$. Let us first recall the definition of fidelity,

$$F(\rho,\sigma) := \|\sqrt{\rho}\,\sqrt{\sigma}\|_1, \tag{A1}$$

where $\|\cdot\|_1$ is the L_1 norm. The generalized fidelity is defined for subnormalized states as

$$\bar{F}(\rho,\sigma) := F(\rho,\sigma) + \sqrt{(1-\operatorname{Tr}\rho)(1-\operatorname{Tr}\sigma)}. \tag{A2}$$

Note that if at least one of the two states is normalized, we recover the usual fidelity. Finally, the purified distance is defined in terms of the generalized fidelity,

$$d(\rho,\sigma) := \sqrt{1-\bar{F}(\rho,\sigma)^2}. \tag{A3}$$

The purified distance is a metric, is invariant under purifications and extensions, and can only decrease under physical operations and projections [39]. It relates to the trace distance as [39]

$$\begin{aligned}\tfrac{1}{2}\|\rho-\sigma\|_1 + \tfrac{1}{2}|\operatorname{Tr}\rho - \operatorname{Tr}\sigma| &\leqslant d(\rho,\sigma)\\ &\leqslant \sqrt{\|\rho-\sigma\|_1 + |\operatorname{Tr}\rho-\operatorname{Tr}\sigma|}.\end{aligned} \tag{A4}$$

The ε ball around a positive operator $\rho \in \mathcal{S}_{\leqslant}(A)$ is defined as usually,

$$\mathcal{B}^{\varepsilon}(\rho) := \{\tilde{\rho} \in \mathcal{S}_{\leqslant}(A) : d(\rho,\sigma) \leqslant \varepsilon\}.$$

b. Smooth min entropy

The conditional smooth min entropy $H^{\varepsilon}_{\min}(A|B)_\rho$ can be used to quantify the size of a subsystem of A that can be decoupled from B [14]. In classical cryptography, it is applied to privacy amplification, giving us the length of a secret key that can be securely extracted from A such that it is inaccessible to an adversary that controls system B. The nonsmooth version of the min entropy is defined as

$$H_{\min}(A|B)_\rho := \sup_{\sigma_B\in\mathcal{S}(B)} \sup_{\lambda\in\mathbb{R}}\{\lambda : 2^{-\lambda}\mathbb{1}_A\otimes\sigma_B \geqslant \rho_{AB}\}. \tag{A5}$$

In the particular case where the two systems are independent, $\rho_{AB} = \rho_A\otimes\rho_B$, the min entropy is simply $-\log_2\|\rho_A\|_\infty$, where $\|\rho_A\|_\infty$ is the maximum eigenvalue of ρ_A.

Smoothing is made by optimizing the min entropy over a small neighborhood of ρ, according to the purified distance,

$$H^{\varepsilon}_{\min}(A|B)_\rho := \sup_{\tilde{\rho}\in\mathcal{B}^{\varepsilon}(\rho)} H_{\min}(A|B)_{\tilde{\rho}}. \tag{A6}$$

The *smoothness parameter* $\varepsilon \geqslant 0$ is usually chosen to be small but nonzero. In most contexts, it corresponds to a small error probability.

c. Smooth max entropy

The smooth conditional max entropy $H^{\varepsilon}_{\max}(A|B)_\rho$ can be used to quantify the number of bits necessary to reconstruct the state of system A, given quantum side information B. In thermodynamics, it characterizes the work cost of erasure of A, given access to B [24]. In classical information theory, the nonconditional max entropy quantifies the compression rate of a random source A. The nonsmooth conditional max entropy can be defined as

$$H_{\max}(A|B)_\rho := \sup_{\sigma_B\in\mathcal{S}(B)} \log_2 F(\rho_{AB},\mathbb{1}_A\otimes\sigma_B)^2, \tag{A7}$$

where F is the fidelity [Eq. (A1)]. We smooth the max entropy as we did with the min entropy,

$$H^{\varepsilon}_{\max}(A|B)_\rho := \inf_{\tilde{\rho}\in\mathcal{B}^{\varepsilon}(\rho)} H_{\max}(A|B)_{\tilde{\rho}}. \tag{A8}$$

2. Generalized smooth entropy

Our final results are expressed in terms of a generalized smooth entropy, introduced in [21]. For $\varepsilon > 0$, it is defined as

$$H^{\varepsilon}(A|B)_\rho := -D^{\varepsilon}_H(\rho_{AB}||\mathbb{1}_A\otimes\rho_B), \tag{A9}$$

where D^{ε}_H is the hypothesis-testing relative entropy, defined as

$$2^{-D^{\varepsilon}_H(\rho||\sigma)} := \frac{1}{\varepsilon}\inf_Q\{\operatorname{Tr}(Q\sigma) : 0\leqslant Q\leqslant\mathbb{1}\wedge\operatorname{Tr}(Q\rho)\geqslant\varepsilon\}. \tag{A10}$$

This corresponds precisely to the setting of hypothesis testing: we are given one of two states ρ and σ at random, and we want to distinguish them with a single measurement, trying to be right on ρ with probability at least ε. We start from the set of all POVMs with two outcomes, $\{Q,\mathbb{1}-Q\}$: our guessing strategy is to say that the state is ρ if we obtain Q and σ if we obtain $\mathbb{1}-Q$. First we restrict the set to those POVMs such that the probability of guessing correctly ρ if the outcome is Q is at least ε. To further optimize our overall guessing probability, we pick the POVM that minimizes the probability of obtaining Q when measuring σ.

Further operational interpretations of the generalized smooth entropy come from its relation to the smooth min and max entropies, given below. In short, for small ε it behaves like the smooth min entropy, and for large ε it approximates the smooth max entropy.

3. Basic properties

(a) Trivial bounds. For any state ρ_{AB}, the three entropy measures are lower bounded by $-\min\{\log_2|A|,\log_2|B|\}$, and upper bounded by $\log_2|A|$.

(b) Examples. For $\varepsilon\to 0$, all three smooth entropies are 0 if ρ_A is pure, $\log_2|A|$ if $\rho_{AB} = \frac{1}{|A|}\mathbb{1}_A\otimes\rho_B$, and $-\log_2|A|$ if ρ_{AB} is maximally entangled.

(c) Pure bipartite states. The nonconditional versions of our entropies only depend on the spectrum of the reduced state, so, if ρ_{AB} is pure, we have $H^{\varepsilon}(A) = H^{\varepsilon}(B)$, $H^{\varepsilon}_{\min}(A) = H^{\varepsilon}_{\min}(B)$ and $H^{\varepsilon}_{\max}(A) = H^{\varepsilon}_{\max}(B)$ (by Schmidt decomposition).

(d) Conditioning on classical information. [23, Prop. 4.6] For quantum-classical states of the form $\rho_{ABC} = \sum_k p_k\,\tau^k_{AB}\otimes|k\rangle\langle k|_C$, the conditional min and max entropies have the form

$$H_{\min}(A|BC)_\rho = -\log_2\left(\sum_k p_k 2^{-H_{\min}(A|B)_{\tau^k}}\right), \tag{A11}$$

$$H_{\max}(A|BC)_\rho = \log_2\left(\sum_k p_k 2^{H_{\max}(A|B)_{\tau^k}}\right). \tag{A12}$$

(e) Product states. The conditional entropy equals the nonconditional entropy for product states,

$$H^{\varepsilon}(A|B)_{\rho_A\otimes\rho_B} = H^{\varepsilon}(A)_{\rho_A}. \tag{A13}$$

Equation (A13) also applies to the smooth min and max entropies.

(f) Data-processing inequality. The entropy of A conditioned on B cannot decrease if information is locally processed at B. Formally,

$$H^{\varepsilon}(A|B)_{\rho} \leqslant H^{\varepsilon}(A|B')_{[\mathcal{I}\otimes\mathcal{E}](\rho)}, \tag{A14}$$

where $[\mathcal{I}\otimes\mathcal{E}](\rho)$ is the state obtained from ρ_{AB} after applying a trace-preserving completely positive map $\mathcal{E}$ on system B. Smooth entropies are invariant under local unitaries $U_A \otimes U_B$. This relation also holds for the smooth min and max entropies.

4. Chain rules

The hypothesis-testing entropy satisfies a chain rule.

Lemma A.1 (Corollary 1 from [21]). *Let* $\rho_{ABC} \in \mathrm{End}(A\otimes B\otimes C)$ be an arbitrary normalized state, and $\epsilon,\epsilon' > 0$. Then,

$$H^{\epsilon+\sqrt{8\epsilon'}}(AB|C)_{\rho} \geqslant H^{\epsilon}(A|BC)_{\rho} + H^{\epsilon'}(B|C)_{\rho} - \log_2 \frac{\epsilon+\sqrt{8\epsilon'}}{\epsilon}.$$

Smooth entropies satisfy several chain rules, for different combinations of min and max entropies [40]. Here we present those needed for our proofs.

Lemma A.2 (Lemma A.7 from [14]). *Let* $\varepsilon > 0$ and $\varepsilon',\varepsilon'' \geqslant 0$. Then

$$H_{\min}^{\varepsilon'}(A|BC)_{\rho} \leqslant H_{\min}^{\varepsilon+2\varepsilon'+\varepsilon''}(AB|C)_{\rho} - H_{\min}^{\varepsilon''}(B|C)_{\rho} + \log_2 \frac{1}{1-\sqrt{1-\varepsilon^2}}.$$

Lemma A.3 (Dual of Theorem 15 from [40]). *Let* $\varepsilon > 0$ and $\varepsilon',\varepsilon'' \geqslant 0$. Then

$$H_{\max}^{2\varepsilon+\varepsilon'+2\varepsilon''}(A|BC)_{\rho} \leqslant H_{\max}^{\varepsilon'}(AB|C)_{\rho} - H_{\min}^{\varepsilon''}(B|C)_{\rho} + 3\log_2 \frac{1}{1-\sqrt{1-\varepsilon^2}}.$$

Lemma A.4 (Theorem 14 from [40]). *Let* $\varepsilon > 0$ and $\varepsilon',\varepsilon'' \geqslant 0$. Then

$$H_{\min}^{\varepsilon+\varepsilon'+\varepsilon''}(A|BC)_{\rho} \geqslant H_{\min}^{\varepsilon'}(AB|C)_{\rho} - H_{\max}^{\varepsilon''}(B|C)_{\rho} - 2\log_2 \frac{1}{1-\sqrt{1-\varepsilon^2}}.$$

Lemma A.5 (Dual of Lemma A.2). *Let* $\varepsilon > 0$ and $\varepsilon',\varepsilon'' \geqslant 0$. Then

$$H_{\max}^{\varepsilon''}(A|BC)_{\rho} \geqslant H_{\max}^{\varepsilon+2\varepsilon'+\varepsilon''}(AB|C)_{\rho} - H_{\max}^{\varepsilon'}(B|C)_{\rho} - \log_2 \frac{1}{1-\sqrt{1-\varepsilon^2}}.$$

5. Relations between the different smooth entropies

a. Duality between smooth min and max entropies

For any tripartite pure state ρ_{ABC}, we have [37,39]

$$H_{\min}^{\varepsilon}(A|C)_{\rho} = -H_{\max}^{\varepsilon}(A|B)_{\rho}. \tag{A15}$$

b. Hypothesis-testing entropy interpolates between smooth min and max entropies

(a) H^{ε} *and* $H_{\min}^{\varepsilon'}$. For small ε, H^{ε} behaves approximately like the smooth min entropy,

$$H^{\varepsilon^2/2}(A|B)_{\rho} \leqslant H_{\min}^{\varepsilon}(A|B)_{\rho} \leqslant H^{11\sqrt{\varepsilon}}(A|B)_{\rho} + \frac{5}{2}\log_2\left(\frac{3}{\varepsilon}\right) + \log_2\left(\frac{2}{1-\varepsilon}\right). \tag{A16}$$

The lower bound comes from [21, Prop. 4.1]. The upper bound is proved in Lemma D.8.

(b) H^{ε} *and* $H_{\max}^{\varepsilon'}$. [21, Prop. 8] For large ε, the the hypothesis-testing entropy behaves approximately like max entropy,

$$H_{\max}(A|B)_{\rho} + \log_2 \frac{1}{\epsilon^2} \geqslant H^{1-\epsilon}(A|B)_{\rho}. \tag{A17}$$

There is also a known bound for the nonconditional smooth max entropy,

$$H^{1-\epsilon}(A)_{\rho} \geqslant H_{\max}^{\sqrt{2\epsilon}}(A)_{\rho} + \log_2 \frac{1}{(1-\epsilon)}. \tag{A18}$$

c. Smooth entropies and von Neumann entropy

For a bipartite quantum state ρ_{AB}, the von Neumann entropy of A conditioned on B is defined as $H(A|B)_{\rho} = H(AB)_{\rho} - H(B)_{\rho}$, where $H(X)_{\sigma} = -\mathrm{Tr}(\sigma_X \log_2 \sigma_X)$ is the usual (nonconditional) von Neumann entropy of σ_X. The conditional von Neumann entropy is always bounded by the smooth min and max entropies in the limit of small ε [41],

$$\lim_{\varepsilon\to 0} H_{\min}^{\varepsilon}(A|B)_{\rho} \leqslant H(A|B)_{\rho} \leqslant \lim_{\varepsilon\to 0} H_{\max}^{\varepsilon}(A|B)_{\rho}. \tag{A19}$$

In particular, if the smooth min and max entropies coincide, they are automatically equal to the von Neumann entropy.

Asymptotic equipartition property. Smooth entropy measures converge to the von Neumann entropy in the limit of many identical and independently distributed systems, when the global state has the form $\rho_{A^{\otimes n}B^{\otimes n}} = {\sigma_{AB}}^{\otimes n}$ [21,41]. Formally, for any $0 < \varepsilon < 1$,

$$\begin{aligned}\lim_{n\to\infty} \frac{1}{n} H^{\varepsilon}(A^{\otimes n}|B^{\otimes n})_{\sigma^{\otimes n}} &= \lim_{n\to\infty} \frac{1}{n} H_{\max}^{\varepsilon}(A^{\otimes n}|B^{\otimes n})_{\sigma^{\otimes n}} \\ &= \lim_{n\to\infty} \frac{1}{n} H_{\min}^{\varepsilon}(A^{\otimes n}|B^{\otimes n})_{\sigma^{\otimes n}} \\ &= H(A|B)_{\sigma}.\end{aligned} \tag{A20}$$

In information theory, this limit is applied to many sequential uses of the same resources, or repetitions of an experiment—which is why the von Neumann entropy is used to characterize the success rate of information-processing tasks. In thermodynamics, we do not always have the luxury of arbitrarily repeating experiments (like letting a cup of coffee thermalize several times), and are usually interested in predictions for a single instance of an event ("what is the probability that this cup of coffee cools down now?"). The same limit emerges, however, in the treatment of large systems made out of many uncorrelated subsystems, like an ideal gas.

APPENDIX B: DECOUPLING THEOREMS

Decoupling theorems [13–15] capture the idea that, given two quantum systems A and R not perfectly correlated, most (random) subsystems of A up to a certain size are decoupled from R. The maximal size of decoupled subsystems depends on correlations between A and R, as measured by conditional entropies. This result has powerful applications in quantum cryptography, error correction, and thermodynamics [13,15,24].

Theorem B.1 (Decoupling [adapted from Theorem 3.1 of [14]]). *Let* $\rho_{AR} \in \mathcal{S}(A \otimes R)$. Let $\mathcal{T}_{A\to B}$ be a trace nonincreasing, completely positive map from End(A) to End(B). Let τ be the Choi-Jamiołkowski representation of $\mathcal{T}$,

$$\tau_{A'B} = [\mathcal{I}_{A'} \otimes \mathcal{T}_{A\to B}]\,(|\Psi\rangle\langle\Psi|_{A'A}),$$

where $|\Psi\rangle_{A'A} = |A|^{-1/2} \sum_i^{|A|} |i\rangle_A |i\rangle_{A'}$ is maximally entangled between A' and a virtual system A. Finally, let $\varepsilon, \Delta, \delta > 0$.

If the entropic relation

$$H^{\varepsilon}_{\min}(A|R)_\rho + H^{\varepsilon}_{\min}(A'|B)_\tau \geqslant 2\log_2 \frac{1}{\Delta - 12\varepsilon}$$

holds, then the fraction (over the set of all unitaries $\{U_A\}$ acting on A, according to the Haar measure) of unitaries such that

$$\|[\mathcal{T} \otimes \mathcal{I}_R](U_A \cdot \rho_{AR}) - \tau_B \otimes \rho_R\|_1 \geqslant \Delta + \delta$$

is at most $2\, e^{-(|A|/16)\delta^2}$.

Note that δ, Δ, and ε do not scale with the size of the systems involved, whereas the entropies do.

The converse theorem gives us tightness of the bound above for trace-preserving maps.

Theorem B.2 (Converse). *Let* $\rho_{AR} \in \mathcal{S}(A \otimes R)$. Let $\mathcal{T}_{A\to B}$ be a trace-preserving completely positive map from End(A) to End(B). Let

$$\tilde{\rho}_{BA'} := [\mathcal{T}_{A\to B} \otimes \mathcal{I}_{A'}](\rho_{AA'}),$$

where $\rho_{AA'}$ is a purification of $\rho_A = \mathrm{Tr}_R(\rho_{AR})$ on a virtual system A'. For any $\varepsilon' > 0$ and any $\varepsilon, \varepsilon'' \geqslant 0$, if

$$H^{2\sqrt{2\varepsilon+6\varepsilon''}+2\sqrt{\varepsilon'}+\varepsilon''}_{\min}(A|R)_\rho + H^{\varepsilon''}_{\max}(A'|B)_{\tilde{\rho}} < -\log_2 \frac{1}{\varepsilon'}, \tag{B1}$$

then

$$\|[\mathcal{T} \otimes \mathcal{I}_R](\rho_{AR}) - \mathcal{T}(\rho_A) \otimes \rho_R\| > \varepsilon.$$

The following corollary is useful to compare the final state with the canonical state.

Corollary B.3. In the setting of Theorem B.2, if condition (B1) holds, then

$$\|[\mathcal{T} \otimes \mathcal{I}_R](\rho_{AR}) - \mathcal{T}(\sigma_A) \otimes \rho_R\| > \frac{\varepsilon}{2},$$

for any normalized density operator σ_A on A.

Proof. First we use the fact that the trace distance cannot decrease under trace-preserving completely positive maps, like the partial trace, to show

$$\begin{aligned}
&\|[\mathcal{T} \otimes \mathcal{I}_R](\rho_{AR}) - \mathcal{T}(\sigma_A) \otimes \rho_R\| \\
&\quad\geqslant \|\mathcal{T}(\rho_A) - \mathcal{T}(\sigma_A)\| \\
&\quad= \|\mathcal{T}(\rho_A) \otimes \rho_R - \mathcal{T}(\sigma_A) \otimes \rho_R\|.
\end{aligned}$$

Using the triangle inequality for the trace distance, we obtain

$$\begin{aligned}
\varepsilon &< \|[\mathcal{T} \otimes \mathcal{I}_R](\rho_{AR}) - \mathcal{T}(\rho_A) \otimes \rho_R\| \\
&< \|[\mathcal{T} \otimes \mathcal{I}_R](\rho_{AR}) - \mathcal{T}(\sigma_A) \otimes \rho_R\| \\
&\quad + \|\mathcal{T}(\sigma_A) \otimes \rho_R - \mathcal{T}(\rho_A) \otimes \rho_R\| \\
&< 2\, \|[\mathcal{T} \otimes \mathcal{I}_R](\rho_{AR}) - \mathcal{T}(\sigma_A) \otimes \rho_R\|.
\end{aligned}$$

■

APPENDIX C: DETAILED RESULTS AND PROOFS

1. Thermalization of typical subsystems

In this section we prove our main result on thermalization after a random evolution (or thermalization of random subsystems), Theorem 1. The first step is to apply the decoupling theorem (Theorem B.1), setting $A = \Omega$, $B = S$, and $\mathcal{T}_{\Omega\to S} = \mathrm{Tr}_E$.

Lemma C.1. Let $\rho_{SER} \in \mathcal{S}(\Omega \otimes R)$, with $\Omega \subseteq S \otimes E$. For any $\tilde{\varepsilon} \geqslant 0$, and any $\Delta > 0$, if

$$H^{\tilde{\varepsilon}}_{\min}(\Omega|R)_\rho + H^{\tilde{\varepsilon}}_{\min}(\Omega'|S)_\tau \geqslant -2\log_2(\Delta - 12\tilde{\varepsilon}) \tag{C1}$$

holds, then, for any $\delta > 0$, the fraction of unitaries $\{U_\Omega\}$ acting on Ω such that

$$\|\,\mathrm{Tr}_E(U_\Omega \cdot \rho_{\Omega R}) - \pi_S \otimes \rho_R\|_1 \geqslant \Delta + \delta$$

is at most $2\, e^{-(|\Omega|/16)\delta^2}$, according to the Haar measure.

In the above, $\tau_{\Omega' S} = \mathrm{Tr}_E(|\Psi\rangle\langle\Psi|_{\Omega'\Omega})$, for the maximally entangled state $|\Psi\rangle_{\Omega'\Omega}$. Note that the reduced state in S is the canonical state, $\tau_S = \mathrm{Tr}_{\Omega'}\,\mathrm{Tr}_E(|\Psi\rangle\langle\Psi|_{\Omega'\Omega}) = \pi_S$.

Now we are ready to state our main theorem in terms of the smooth min and max- entropies. A final reformulation in terms of H^ε follows (Corollary C.3).

Theorem C.2. Let $\rho_{SER} \in \mathcal{S}(\Omega \otimes R)$, with $\Omega \subseteq S \otimes E$. For any $\varepsilon_2, \varepsilon_3 \geqslant 0$, any $\varepsilon_1 > \varepsilon_2 + \varepsilon_3$, and any $\Delta > 0$, if the entropic relation

$$\begin{aligned}
&H^{\varepsilon_1}_{\min}(SE|R)_\rho + H^{\varepsilon_2}_{\min}(E)_\pi - H^{\varepsilon_3}_{\max}(S)_\pi \\
&\quad\geqslant 2\log_2 \frac{1}{\left(1 - \sqrt{1 - (\varepsilon_1 - \varepsilon_2 - \varepsilon_3)^2}\right)(\Delta - 12\varepsilon_1)}
\end{aligned}$$

holds, then, for any $\delta > 0$, the fraction of unitaries $\{U_\Omega\}$ acting on Ω such that

$$\|\,\mathrm{Tr}_E(U_\Omega \cdot \rho_{\Omega R}) - \pi_S \otimes \rho_R\|_1 \geqslant \Delta + \delta$$

is at most $2\, e^{-(|\Omega|/16)\delta^2}$, according to the Haar measure.

Proof. We start from Lemma C.1, and break down the left-hand side of condition (C1). First off, we observe that $H^{\tilde{\varepsilon}}_{\min}(\Omega|R)_\rho = H^{\tilde{\varepsilon}}_{\min}(SE|R)_\rho$. We use the chain rule from Lemma A.4 to bound the other entropy. Setting $\tilde{\varepsilon} = \varepsilon_1 + \varepsilon_2 + \varepsilon_3$, we have

$$\begin{aligned}
H^{\varepsilon_1+\varepsilon_2+\varepsilon_3}_{\min}(\Omega'|S)_\tau &\geqslant H^{\varepsilon_2}_{\min}(\Omega' S)_\tau - H^{\varepsilon_3}_{\max}(S)_\tau \\
&\quad + 2\log_2\left(1 - \sqrt{1 - \varepsilon_1^2}\right).
\end{aligned}$$

Since $|\Psi\rangle_{\Omega' SE}$ is a pure state, we have that $H^{\varepsilon_2}_{\min}(\Omega' S)_\tau = H^{\varepsilon_2}_{\min}(E)_\pi$. Condition (C1) becomes

$$\begin{aligned}
&H^{\varepsilon_1+\varepsilon_2+\varepsilon_3}_{\min}(SE|R)_\rho + H^{\varepsilon_2}_{\min}(E)_\pi - H^{\varepsilon_3}_{\max}(S)_\pi \\
&\quad + 2\log_2\left(1 - \sqrt{1 - \varepsilon_1^2}\right) \geqslant -2\log_2(\Delta - \varepsilon_1 - \varepsilon_2 - \varepsilon_3).
\end{aligned}$$

To clean up, we take $\varepsilon_1 + \varepsilon_2 + \varepsilon_3 \to \varepsilon_1$. ■

We may now write this result in terms of the hypothesis-testing entropy, and simplify the ε terms at the cost of little generality.

Corollary C.3. Let $\rho_{SER} \in \mathcal{S}(\Omega \otimes R)$, with $\Omega \subseteq S \otimes E$. Let $\varepsilon, \Delta > 0$.

If the entropic relation

$$H^{9\varepsilon}(SE|R)_\rho + H^{\varepsilon}(E)_\pi - H^{1-\varepsilon}(S)_\pi \geqslant 2\log_2 \frac{1}{(1-\sqrt{1-2\,\varepsilon})(\Delta - 36\sqrt{2\varepsilon})} - \log_2 \frac{1}{1-\varepsilon} \quad \text{(C2)}$$

holds, then, for any $\delta > 0$, the fraction of unitaries $\{U_\Omega\}$ acting on Ω such that

$$\|\operatorname{Tr}_E(U_\Omega \cdot \rho_{\Omega R}) - \pi_S \otimes \rho_R\|_1 \geqslant \Delta + \delta$$

is at most $2\, e^{-(|\Omega|/16)\delta^2}$, according to the Haar measure.

Proof. Starting from

$$H_{\min}^{\varepsilon_1}(SE|R)_\rho + H_{\min}^{\varepsilon_2}(E)_\pi - H_{\max}^{\varepsilon_3}(S)_\pi \geqslant 2\log_2 \frac{1}{(1-\sqrt{1-(\varepsilon_1-\varepsilon_2-\varepsilon_3)^2})(\Delta-12\varepsilon_1)}, \quad \text{(C3)}$$

we use relations (A16) and (A18) to obtain

$$H_{\min}^{\varepsilon_1}(SE|R)_\rho \geqslant H^{{\varepsilon_1}^2/2}(SE|R)_\rho,$$

$$H_{\min}^{\varepsilon_2}(E)_\pi \geqslant H^{{\varepsilon_2}^2/2}(E)_\pi,$$

$$-H_{\max}^{\varepsilon_3}(S)_\pi \geqslant -H^{1-{\varepsilon_3}^2/2}(S)_\pi + \log_2 \frac{1}{1-{\varepsilon_3}^2/2}.$$

Applying these bounds to (C3), we get

$$H^{{\varepsilon_1}^2/2}(SE|R)_\rho + H^{{\varepsilon_2}^2/2}(E)_\pi - H^{1-{\varepsilon_3}^2/2}(S)_\pi \geqslant 2\log_2 \frac{1}{(1-\sqrt{1-(\varepsilon_1-\varepsilon_2-\varepsilon_3)^2})(\Delta-12\varepsilon_1)} - \log_2 \frac{1}{1-{\varepsilon_3}^2/2}.$$

To simplify, we consider the special case $\tilde\varepsilon = \frac{\varepsilon_1}{3} = \varepsilon_2 = \varepsilon_3$. This gives us

$$H^{9\tilde\varepsilon^2/2}(SE|R)_\rho + H^{\tilde\varepsilon^2/2}(E)_\pi - H^{1-\tilde\varepsilon^2/2}(S)_\pi \geqslant 2\log_2 \frac{1}{(1-\sqrt{1-\tilde\varepsilon^2})(\Delta-36\,\tilde\varepsilon)} - \log_2 \frac{1}{1-\tilde\varepsilon^2/2}.$$

Finally, we take $\varepsilon = \frac{\tilde\varepsilon^2}{2}$ to obtain the statement of the corollary. ■

We may simplify it further by taking $\Delta = \delta$.

Since we are usually interested in the limit of small ε, it might at first appear that the right-hand side of (C2) diverges in that limit. However, the divergence is only logarithmic in ε, and does not depend on the size of the systems involved. The entropic terms, on the other hand, grow with the size of the systems. In the thermodynamic limit of large systems, the logarithmic divergence is negligible.

2. Converse

The converse bound follows. A friendlier, if weaker, bound can be found in Corollary C.5.

Theorem C.4 (Tightness). Let $\rho_{SER} \in \mathcal{S}(\Omega \otimes R)$, with $\Omega \subseteq S \otimes E$. Let $\delta, \varepsilon_1, \varepsilon_2 > 0$ and $\varepsilon_3, \varepsilon_4 \geqslant 0$.

For readability, we set $\tilde\varepsilon = 2\sqrt{\delta + 3(\varepsilon_2+\varepsilon_3+\varepsilon_4)} + \varepsilon_1 + \varepsilon_2 + \varepsilon_3 + \varepsilon_4$.

If

$$H_{\min}^{2\tilde\varepsilon}(\Omega|R)_\rho + \max_{\sigma\in\mathcal{S}(\Omega)} \left[H_{\max}^{2\varepsilon_3}(E)_\sigma - H_{\min}^{\varepsilon_4}(S)_\sigma\right] < -\log_2 \frac{1}{{\varepsilon_1}^2} - 3\log_2 \frac{1}{1-\sqrt{1-{\varepsilon_2}^2}}, \quad \text{(C4)}$$

then

$$\|\operatorname{Tr}_E(U_\Omega \cdot \rho_{AR}) - \pi_S \otimes \rho_R\| > \delta,$$

for any unitary U_Ω acting on Ω.

Proof. We start from Corollary B.3, setting $A = \Omega$, $B = S$, $\mathcal{T}(\cdot) = \operatorname{Tr}_E(U_\Omega \cdot)$, and $\sigma_A = \pi_\Omega$. This gives us the condition

$$H_{\min}^{2\sqrt{2\varepsilon_1+6\varepsilon_2}+2\sqrt{\varepsilon_3}+\varepsilon_2}(\Omega|R)_\rho + H_{\max}^{\varepsilon_2}(\Omega'|S)_{\tilde\rho} < -\log_2 \frac{1}{\varepsilon_3}, \quad \text{(C5)}$$

which implies

$$\|\operatorname{Tr}_E(U_\Omega \cdot \rho_{AR}) - \operatorname{Tr}_E(\pi_\Omega) \otimes \rho_R\| > \frac{\varepsilon_1}{2}.$$

Here, $\tilde\rho = U_\Omega \cdot \rho_{\Omega\Omega'}$, where $\rho_{\Omega\Omega'}$ is a purification of ρ_Ω.

We will look for an upper bound for $H_{\max}^{\varepsilon_2}(\Omega'|S)_{\tilde\rho}$, as we might not know which unitary U_Ω was applied. We will use a little of brute force, maximizing the conditional entropy over all states σ_Ω in $\mathcal{S}(\Omega)$, with purification $\sigma_{\Omega\Omega'}$ (this is stronger than maximizing over all unitaries U_Ω). Also, in order to use a chain rule, let us set $\varepsilon_2 = 2\varepsilon_4 + \varepsilon_5 + 2\varepsilon_6$. We have

$$\begin{aligned}
H_{\max}^{2\varepsilon_4+\varepsilon_5+2\varepsilon_6}(\Omega'|S)_{\tilde\rho} &\leqslant \max_{\sigma\in\mathcal{S}(\Omega)} H_{\max}^{2\varepsilon_4+\varepsilon_5+2\varepsilon_6}(\Omega'|S)_{\sigma_{\Omega\Omega'}} \\
&\leqslant \max_{\sigma\in\mathcal{S}(\Omega)} \left[H_{\max}^{\varepsilon_5}(S\Omega')_\sigma - H_{\min}^{\varepsilon_6}(S)_\sigma\right] + 3\log_2 \frac{1}{1-\sqrt{1-\varepsilon_4^2}} \quad \text{[Lemma A.3]} \\
&= \max_{\sigma\in\mathcal{S}(\Omega)} \left[H_{\max}^{\varepsilon_5}(E)_\sigma - H_{\min}^{\varepsilon_6}(S)_\sigma\right] + 3\log_2 \frac{1}{1-\sqrt{1-\varepsilon_4^2}} \quad [\sigma_{SE\Omega'} \text{ pure}].
\end{aligned}$$

Condition (C5) becomes

$$H_{\min}^{2\sqrt{2\varepsilon_1+6(2\varepsilon_4+\varepsilon_5+2\varepsilon_6)}+2\sqrt{\varepsilon_3}+(2\varepsilon_4+\varepsilon_5+2\varepsilon_6)}(\Omega|R)_\rho + \max_{\sigma\in\mathcal{S}(\Omega)} \left[H_{\max}^{\varepsilon_5}(E)_\sigma - H_{\min}^{\varepsilon_6}(S)_\sigma\right] < -\log_2 \frac{1}{\varepsilon_3} - 3\log_2 \frac{1}{1-\sqrt{1-\varepsilon_4^2}},$$

which we cannot hope to make much more readable without losing generality (we do simplify it in the corollary ahead). For now, let us just relabel

$$\varepsilon_1 \to 2\delta, \quad \varepsilon_3 \to \varepsilon_1^2, \quad \varepsilon_4 \to \varepsilon_2, \quad \varepsilon_5 \to 2\varepsilon_3, \quad \varepsilon_6 \to \varepsilon_4,$$

to obtain the beauty

$$H_{\min}^{2\left(2\sqrt{\delta+3(\varepsilon_2+\varepsilon_3+\varepsilon_4)}+\varepsilon_1+\varepsilon_2+\varepsilon_3+\varepsilon_4\right)}(\Omega|R)_\rho + \max_{\sigma\in\mathcal{S}(\Omega)} \left[H_{\max}^{2\varepsilon_3}(E)_\sigma - H_{\min}^{\varepsilon_4}(S)_\sigma\right] < -\log_2 \frac{1}{{\varepsilon_1}^2} - 3\log_2 \frac{1}{1-\sqrt{1-{\varepsilon_2}^2}}.$$

■

In the following corollary we simplify some of the terms. In particular, we neglect a term with the smooth min entropy of S, for an optimal state. In the typical case where S is much smaller than E, this is only a small loss.

Corollary C.5. Let $\rho_{SER} \in \mathcal{S}(\Omega \otimes R)$, with $\Omega \subseteq S \otimes E$. Let $\delta > 0$ and let $\varepsilon > 4\sqrt{\delta}$. We define $f(\varepsilon,\delta) := \frac{1}{16}(6+\varepsilon - 2\sqrt{9+3\,\varepsilon+4\delta})^2$.

If

$$H^{11\sqrt{\varepsilon}}(\Omega|R)_\rho + H^1(E)_\pi < -\log_2 \frac{1}{f(\varepsilon,\delta)} - 3\log_2 \frac{1}{1-\sqrt{1-f(\varepsilon,\delta)}} - \frac{5}{2}\log_2\left(\frac{3}{\varepsilon}\right) - \log_2\left(\frac{2}{1-\varepsilon}\right), \tag{C6}$$

then

$$\| \mathrm{Tr}_E(U_\Omega \cdot \rho_{AR}) - \pi_S \otimes \rho_R \| > \delta,$$

for any unitary U_Ω acting on Ω.

Proof. We start from the condition of Theorem C.4,

$$H_{\min}^{2(2\sqrt{\delta+3(\varepsilon_2+\varepsilon_3+\varepsilon_4)}+\varepsilon_1+\varepsilon_2+\varepsilon_3+\varepsilon_4)}(\Omega|R)_\rho + \max_{\sigma\in\mathcal{S}(\Omega)} \left[H_{\max}^{2\varepsilon_3}(E)_\sigma - H_{\min}^{\varepsilon_4}(S)_\sigma\right] < -\log_2 \frac{1}{{\varepsilon_1}^2} - 3\log_2 \frac{1}{1-\sqrt{1-{\varepsilon_2}^2}}.$$

We are looking for a simpler, tighter condition, i.e., an upper bound to the left-hand side of the inequality and a lower bound to the right-hand side.[1] First we neglect the term with the nonconditional entropy of S, as

$$\max_{\sigma\in\mathcal{S}(\Omega)} \left[H_{\max}^{\varepsilon_3}(E)_\sigma - H_{\min}^{\varepsilon_4}(S)_\sigma\right] \leqslant \max_{\sigma\in\mathcal{S}(\Omega)}\ H_{\max}^{\varepsilon_3}(E)_\sigma.$$

Now we apply the upper bound for the max entropy given by Lemma D.2,

$$\max_{\sigma\in\mathcal{S}(\Omega)}\ H_{\max}^{\varepsilon_3}(E)_\sigma \leqslant \max_{\sigma\in\mathcal{S}(\Omega)}\ H^1(E)_\sigma = \max_{\sigma\in\mathcal{S}(\Omega)}\ \log_2|\mathrm{supp}\ \sigma_E|. \tag{C7}$$

Finally, we show that for all states $\sigma \in \mathcal{S}(\Omega)$, it stands that $\mathrm{supp}\ \sigma_E \subseteq \mathrm{supp}\ \pi_E$, and therefore (C7) is upper bounded by $H^1(E)_\pi$. For every $\sigma \in \mathcal{S}(\Omega)$, there exists a basis $\{|k\rangle\}_k$ of Ω that diagonalizes it,

$$\sigma = \sum_k^{|\Omega|} p_k |k\rangle\langle k|_\Omega.$$

Since Ω is a subspace of $S\otimes E$, we can expand each element $|k\rangle_\Omega$ in any basis of $S \otimes E$; in particular, we can choose a product basis $\{|i\rangle_S \otimes |j\rangle_E\}_{i,j}$, such that

$$|k\rangle_\Omega = \sum_i^{|S|}\sum_j^{|E|} c_{ij}^k\ |i\rangle_S \otimes |j\rangle_E, \qquad \sum_{i,j} \left|c_{ij}^k\right|^2 = 1,\ \forall\, k.$$

We may now expand σ in this basis,

$$\sigma_\Omega = \sum_k^{|\Omega|} p_k \sum_{i,i'}^{|S|}\sum_{j,j'}^{|E|} c_{ij}^k\ \left(c_{i'j'}^k\right)^*\ |i\rangle\langle i'|_S \otimes |j\rangle\langle j'|_E,$$

$$\sigma_E = \mathrm{Tr}_S\ \bar{\sigma}_\Omega = \sum_k^{|\Omega|} p_k \sum_i^{|S|}\sum_{j,j'}^{|E|} c_{ij}^k\ \left(c_{ij'}^k\right)^*\ |j\rangle\langle j'|_E.$$

Note that the canonical state is given by

$$\pi_\Omega = \sum_k^{|\Omega|} \frac{1}{|\Omega|}|k\rangle\langle k| = \sum_k^{|\Omega|} \frac{1}{|\Omega|} \sum_{i,i'}^{|S|}\sum_{j,j'}^{|E|} c_{ij}^k\ \left(c_{i'j'}^k\right)^*\ |i\rangle\langle i'|_S \otimes |j\rangle\langle j'|_E,$$

$$\pi_E = \sum_k^{|\Omega|} \frac{1}{|\Omega|} \sum_i^{|S|}\sum_{j,j'}^{|E|} c_{ij}^k\ \left(c_{ij'}^k\right)^*\ |j\rangle\langle j'|_E,$$

so clearly $\mathrm{supp}\ \sigma_E \subseteq \mathrm{supp}\ \pi_E$.

Let us see where we stand. We may set $\varepsilon_3 = \varepsilon_4 = 0$, and $\varepsilon_1 = \varepsilon_2 =: \tilde{\varepsilon}$. Our condition becomes

$$H_{\min}^{4(\sqrt{\delta+3\tilde{\varepsilon}}\ +\tilde{\varepsilon})}(\Omega|R)_\rho + H^1(E)_\pi < -\log_2 \frac{1}{\tilde{\varepsilon}^2} - 3\log_2 \frac{1}{1-\sqrt{1-\tilde{\varepsilon}^2}}.$$

We may also bound the term with the smooth min entropy using Eq. (A16). We set $\varepsilon := 4(\sqrt{\delta+3\tilde{\varepsilon}}+\tilde{\varepsilon})$, and have

$$H_{\min}^{\varepsilon}(\Omega|R)_\rho \leqslant H^{11\sqrt{\varepsilon}}(\Omega|R)_\rho - \frac{5}{2}\log_2\left(\frac{\varepsilon}{3}\right) + \log_2\left(\frac{2}{1-\varepsilon}\right).$$

This leaves us with the condition

$$H^{11\sqrt{\varepsilon}}(\Omega|R)_\rho + H^1(E)_\pi < -\log_2 \frac{1}{\tilde{\varepsilon}^2} - 3\log_2 \frac{1}{1-\sqrt{1-\tilde{\varepsilon}^2}} - \frac{5}{2}\log_2\left(\frac{3}{\varepsilon}\right) - \log_2\left(\frac{2}{1-\varepsilon}\right).$$

Now we should make the dependence in δ a little more explicit. In order to keep the above expression only moderately foul, we bound the logarithmic terms on the right-hand side. [We used $\tilde{\varepsilon}^2 = \frac{1}{16}(6+\varepsilon-2\sqrt{9+3\,\varepsilon+4\delta})^2$ and applied the expansion $1-\sqrt{1-x^2} \geqslant \frac{x^2}{2}$ twice.] The new bound is

$$H^{11\sqrt{\varepsilon}}(\Omega|R)_\rho + H^1(E)_\pi < -4\log_2 \frac{6+\varepsilon}{4\left(\frac{\varepsilon^2}{16}-\delta\right)^2} - \frac{5}{2}\log_2\left(\frac{3}{\varepsilon}\right) - \log_2\left(\frac{16}{1-\varepsilon}\right).$$

■

3. Dimension bounds

To give an intuitive idea of the magnitude of the entropic terms in our results, we present a coarser version of our direct bounds.

Corollary C.6. Let $\rho_{SER} \in \mathrm{End}(\Omega \otimes R)$ be a normalized density operator, with $\Omega \subseteq S \otimes E$. For any $\varepsilon \geqslant 0$, and any

[1]In other words, we start from an inequality like $A < B$, and search for suitable $\bar{A}$ and $\bar{B}$ such that $A \leqslant \bar{A}$ and $\bar{B} \leqslant B$. Therefore, $\bar{A} < \bar{B}$ implies the original condition $A < B$.

$\Delta > 0$, if

$$H^\varepsilon(\Omega|R)_\rho + \log_2|\Omega| - 2\log_2|S| \geqslant 2\log_2\left(\frac{1}{\delta - \sqrt{2\,\varepsilon}}\right) \quad \text{(C8)}$$

holds, then, for any $\delta > 0$, the fraction of unitaries $\{U_\Omega\}$ acting on Ω such that

$$\tfrac{1}{2}\,\|\,\mathrm{Tr}_E(U_\Omega \cdot \rho_{\Omega R}) - \pi_S \otimes \rho_R\|_1 \geqslant \delta$$

is at most $2\,e^{-(|\Omega|/16)\delta^2}$, according to the Haar measure.

This corollary follows directly from Lemma C.1, combined with Lemma D.3, and the relation between the smooth min entropy and the hypothesis-testing entropy. We also set $\Delta = \delta$.

APPENDIX D: EXTRA TECHNICAL RESULTS FOR SMOOTH ENTROPIES

In order to present our physical results, we had to prove some of the properties of smooth entropies stated in Appendix A. This appendix is a collection of technical lemmas, mostly adaptations of similar results for other entropy measures. Please have no expectations of elegance or originality as you read through.

The highlights of the appendix are Lemma D.7, where we show that H^ε is continuous on the quantum state (in a way that does not depend on the dimension of the quantum systems involved; in other words, it is "smooth"), and Lemma D.8, where we give a bound for H^ε in terms of the conditional smooth min entropy.

1. A few more definitions

a. Hypothesis-testing relative entropy as a semi-definite program

We can write the hypothesis-testing relative entropy as a semidefinite program (SDP) [21,42,43]. The primal and dual SDPs for $2^{-D^\varepsilon(\rho||\sigma)}$ are

primal		dual	
minimize	$\frac{1}{\varepsilon}\,\mathrm{Tr}(Q\,\sigma)$	maximize	$\mu - \dfrac{\mathrm{Tr}\,X}{\varepsilon}$
subject to	$\mathrm{Tr}(Q\,\rho) \geqslant \varepsilon,$	subject to	$\mu\,\rho \leqslant \sigma + X,$
	$0 \leqslant Q \leqslant \mathbb{1}$		$X,\mu \geqslant 0.$

In the above, it is required that ρ and σ be Hermitian operators. Remember that the generalized conditional smooth entropy is defined as $H^\varepsilon(A|B)_\rho = -D^\varepsilon(\rho_{AB}||\mathbb{1}_A \otimes \rho_B)$.

b. Alternative smooth min entropy

$\hat{H}^\varepsilon_{\min}$ is an alternative entropy measure similar to the smooth min entropy, except that we do not optimize over the choice of the marginal σ_B [22],

$$\hat{H}^\varepsilon_{\min}(A|B)_\rho := \max_{\tilde{\rho}_{AB}\in\mathcal{B}^\varepsilon(\rho)} \sup_{\lambda\in\mathbb{R}}\{\lambda : 2^{-\lambda}\mathbb{1}_A \otimes \tilde{\rho}_B \geqslant \tilde{\rho}_{AB}\}.$$

The optimization is made over the set of subnormalized states that are ε-close to ρ_{AB}, according to the purified distance.

2. A couple of trivial bounds for the smooth entropies

Lemma D.1. Let $\rho \in \mathcal{S}(A)$. Then we have $H^1(A)_\rho = \log_2|\mathrm{supp}\,\rho|$.

Proof. To show that $H^1(A)_\rho \leqslant \log_2|\mathrm{supp}\,\rho|$, we look at the primal program for $H^1(A)_\rho$,

$$\begin{aligned} 2^{H^1(A)_\rho} &= \min \mathrm{Tr}(Q_A\,\mathbb{1}_A),\\ \mathrm{Tr}(Q_A\,\rho_A) \geqslant 1, &\quad 0 \leqslant Q_A \leqslant \mathbb{1}_A. \end{aligned}$$

We take as a candidate the projector onto the support of ρ, $Q = \Pi_\rho$. We have $\mathrm{Tr}(\Pi_\rho\,\rho) = 1$, so Π_ρ is a feasible candidate for the minimization. Therefore we have $2^{H^1(A)_\rho} \leqslant \mathrm{Tr}(\Pi_\rho\,\mathbb{1}_A) = |\mathrm{supp}\,\rho|$.

Now we show that $H^1(A)_\rho \geqslant \log_2|\mathrm{supp}\,\rho|$. The dual program for the generalized smooth entropy $H^1(A)_\rho$ is, in the nonconditional case,

$$\begin{aligned} 2^{H^1(A)_\rho} &= \max \mu - \mathrm{Tr}\,X,\\ \mu\rho \leqslant \mathbb{1} + X, &\quad \mu, X \geqslant 0. \end{aligned}$$

Let us take the candidate $X = \mu\rho - \Pi_\rho$. We have

$$\mathbb{1} + X = \mathbb{1} + \mu\rho - \Pi_\rho \geqslant \mu\rho,$$

so X is a feasible candidate for the dual SDP. This gives us

$$\begin{aligned} 2^{H^1(A)_\rho} \geqslant \mu - \mathrm{Tr}\,X &= \mu - \mathrm{Tr}(\mu\rho - \Pi_\rho)\\ &= \mu - \mu\,\mathrm{Tr}\,\rho + \mathrm{Tr}(\Pi_\rho) = |\mathrm{supp}\,\rho|. \end{aligned}$$

■

Lemma D.2. Let $\rho \in \mathcal{S}(A)$. The nonconditional max entropy is upper bounded as

$$H^\varepsilon_{\max}(A)_\rho \leqslant H^1(A)_\rho.$$

Proof. Let $\rho_A = \sum_k p_k\,|k\rangle\langle k|_A$, for some basis $\{|k\rangle\}_k$ of the support of ρ in A. We note that $H^\varepsilon_{\max}(A)_\rho \leqslant H^0_{\max}(A)_\rho = \log_2 F(\rho_A,\mathbb{1}_A)^2$, and

$$\begin{aligned} F(\rho_A,\mathbb{1}_A)^2 &= \mathrm{Tr}\,(|\sqrt{\rho_A}\,\sqrt{\mathbb{1}_A}|)^2 = \sum_{k,\ell}^{|\mathrm{supp}\,\rho|}\sqrt{p_k}\sqrt{p_\ell}\\ &\leqslant \sum_{k,\ell}^{|\mathrm{supp}\,\rho|}\frac{p_k + p_\ell}{2} \quad \text{[inequality of arithmetic and geometric means]}\\ &= |\mathrm{supp}\rho_A|. \end{aligned}$$

Combining this with Lemma D.1, we obtain $H^\varepsilon_{\max}(A)_\rho \leqslant H^1(A)_\rho$. ■

The following lemma is used to bound our condition for relative thermalization in terms of system dimensions (see Appendix C 3).

Lemma D.3. Let $\rho_{AB} \in \mathcal{S}(A \otimes B)$ be a quantum state with a fully mixed marginal in A, $\rho_A = \frac{\mathbb{1}_A}{|A|}$. Then, for any $\varepsilon \geqslant 0$,

$$H^\varepsilon_{\min}(A|B)_\rho \geqslant \log_2|A| - 2\log_2|B|.$$

Proof. We start by going to the nonsmooth version of the min entropy,

$$\forall \varepsilon \geqslant 0, \qquad H^\varepsilon_{\min}(A|B)_\rho \geqslant H_{\min}(A|B)_\rho.$$

It is convenient to formulate the min entropy as an SDP. The primal SDP for $2^{-H_{\min}(A|B)_\rho}$ is

$$\begin{aligned}
\text{minimize}\quad & \gamma\\
\text{subject to}\quad & \rho_{AB} \leqslant \gamma\, \mathbb{1}_A \otimes \sigma_B,\\
& \sigma_B \in \mathcal{S}(B),\\
& \gamma \geqslant 0.
\end{aligned}$$

We want to show that $\gamma = \frac{|B|^2}{|A|}$ is a feasible candidate for the optimization problem, so that $H_{\min}(A|B)_\rho \geqslant \log_2 \frac{|A|}{|B|^2}$. We apply [23, Lemma A.2], which says that for positive operators $\rho \in \text{End}(A\otimes B)$, it holds that $\rho_{AB} \leqslant |B|\, \rho_A \otimes \mathbb{1}_B$. This gives us

$$\rho_{AB} \leqslant |B|\, \rho_A \otimes \mathbb{1}_B = |B|\, \tfrac{\mathbb{1}_A}{|A|} \otimes \mathbb{1}_B = \tfrac{|B|^2}{|A|}\, \mathbb{1}_A \otimes \underbrace{\frac{\mathbb{1}_B}{|B|}}_{=:\sigma_B}.$$

■

3. Three recycled lemmas

The following lemmas come from [41, Lemma 15]. We need them to prove smoothness of H^ε, so we repeat them here for completeness.

Lemma D.4. Let $\sigma,\Delta \in \mathcal{S}_\leqslant(A)$. The operator

$$G := \sigma^{1/2}(\sigma+\Delta)^{-1/2}$$

is a contraction, i.e., $G \geqslant 0$ and $\|G\|_\infty \leqslant 1$. In particular, conjugating any positive operator X with G can only decrease the trace: $\text{Tr}(G\,X\,G^\dagger) \leqslant \text{Tr}(X)$.

Proof. We conjugate the following with $(\sigma+\Delta)^{-1/2}$:

$$\begin{gathered}
\sigma \leqslant \sigma + \Delta,\\
(\sigma+\Delta)^{-1/2}\sigma(\sigma+\Delta)^{-1/2} \leqslant (\sigma+\Delta)^{-1/2}(\sigma+\Delta)(\sigma+\Delta)^{-1/2},\\
G^\dagger G \leqslant \mathbb{1}.
\end{gathered}$$

Now we use the fact that, for the operator norm, $\Rightarrow \|G\|_\infty{}^2 = \|G^\dagger G\|_\infty \leqslant \|\mathbb{1}\|_\infty = 1$. The second claim comes from $\text{Tr}(GXG^\dagger) = \text{Tr}(X\,G^\dagger\,G) \leqslant \text{Tr}(X\,\mathbb{1})$. ■

Lemma D.5. Let $\rho_{AB} \in \mathcal{S}(A\otimes B)$, and $\sigma_B,\Delta_B \in \mathcal{S}_\leqslant(B)$, such that $\rho_B \leqslant \sigma_B + \Delta_B$. Let $G_B = \sigma^{1/2}(\sigma+\Delta)^{-1/2}$. Then,

$$\|\rho_{AB} - (\mathbb{1}_A\otimes G_B)\rho_{AB}(\mathbb{1}_A \otimes G_B^\dagger)\|_1 \leqslant 2\sqrt{2\,\text{Tr}\ \Delta}.$$

Proof. First we work with the fidelity between the two states, and later we relate it to the trace distance. Using Uhlmann's theorem, we bound the fidelity using a purification of ρ_{AB}. Note that if $|\psi\rangle_{RAB}$ purifies ρ_{AB}, then $(\mathbb{1}_{RA}\otimes G_B)|\psi\rangle$ purifies $(\mathbb{1}_A\otimes G_B)\rho_{AB}(\mathbb{1}\otimes G_B^\dagger)$, and in particular it purifies $G_B\rho_B G_B^\dagger$. We have

$$\begin{aligned}
&F(\rho_{AB},(\mathbb{1}_A\otimes G_B)\rho_{AB}(\mathbb{1}_A\otimes G_B^\dagger))\\
&\quad\geqslant F(|\psi\rangle,(\mathbb{1}_{RA}\otimes G_B)|\psi\rangle)\\
&\quad= |\langle\psi|(\mathbb{1}_{RA}\otimes G_B)|\psi\rangle|\\
&\quad= |\text{Tr}\,((\mathbb{1}_{RA}\otimes G_B)|\psi\rangle\langle\psi|)|\\
&\quad= |\text{Tr}(G_B\ \rho_B)|\ \text{[real and imaginary parts]}\\
&\quad= \sqrt{\mathcal{R}[\text{Tr}(G_B\ \rho_B)]^2 + \mathcal{I}[\text{Tr}(G_B\ \rho_B)]^2}\\
&\quad\geqslant \mathcal{R}[\text{Tr}(G_B\ \rho_B)]\\
&\quad= \text{Tr}\left(\tfrac{1}{2}(G_B + G_B^\dagger)\rho_B\right).
\end{aligned}$$

From Lemma D.4 we know that G is a contraction. Note that $\frac{1}{2}(G+G^\dagger)$ is also a contraction, as $\|\frac{1}{2}(G+G^\dagger)\|_\infty \leqslant \frac{1}{2}\|G\|_\infty + \frac{1}{2}\|G^\dagger\|_\infty \leqslant 1$. We omit the subscript B in most of the following. We have

$$\begin{aligned}
1 - \text{Tr}&\left(\frac{1}{2}(G+G^\dagger)\rho_B\right)\\
&= \text{Tr}\left(\underbrace{\left[\mathbb{1}_B - \frac{1}{2}(G+G^\dagger)\right]}_{\geqslant 0}\rho_B\right)\\
&\leqslant \text{Tr}\left(\left[\mathbb{1}_B - \frac{1}{2}(G+G^\dagger)\right](\sigma+\Delta)\right)\\
&= \text{Tr}(\sigma+\Delta) - \frac{1}{2}\,\text{Tr}[(G+G^\dagger)(\sigma+\Delta)]\\
&= \text{Tr}(\sigma+\Delta) - \frac{1}{2}\,\text{Tr}[\sigma^{1/2}(\sigma+\Delta)^{-1/2}(\sigma+\Delta)]\\
&\quad - \frac{1}{2}\,\text{Tr}[(\sigma+\Delta)^{-1/2}\sigma^{1/2}(\sigma+\Delta)]\\
&= \text{Tr}(\sigma+\Delta) - \text{Tr}[\sigma^{1/2}(\sigma+\Delta)^{1/2}]\\
&\leqslant \text{Tr}(\sigma+\Delta) - \text{Tr}(\sigma)\\
&= \text{Tr}(\Delta),
\end{aligned}$$

so $F(\rho_{AB},(\mathbb{1}_A\otimes G_B)\rho_{AB}(\mathbb{1}_A\otimes G_B^\dagger)) \geqslant 1 - \text{Tr}(\Delta)$. From the relation between trace distance and fidelity, we have

$$\begin{aligned}
&\|\rho_{AB} - (\mathbb{1}_A\otimes G_B)\rho_{AB}(\mathbb{1}_A\otimes G_B^\dagger)\|_1\\
&\quad\leqslant 2\sqrt{1 - F(\rho_{AB},(\mathbb{1}_A\otimes G_B)\rho_{AB}(\mathbb{1}_A\otimes G_B^\dagger))^2}\\
&\quad\leqslant 2\sqrt{1 - [1-\text{Tr}(\Delta)]^2}\\
&\quad= 2\sqrt{1 - 1 + 2\,\text{Tr}(\Delta) - \text{Tr}(\Delta)^2}\\
&\quad\leqslant 2\sqrt{2\ \text{Tr}(\Delta)}.
\end{aligned}$$

■

The following lemma is simply an adaptation of [23, Lemma 5.2] for the alternative smooth min entropy. The proof is identical.

Lemma D.6. Let $\rho \in \mathcal{S}(A\otimes B), \varepsilon \geqslant 0$. Then, there is an embedding from A to $A \oplus \bar{A}$ and a normalized state $\hat{\rho} \in \mathcal{S}[(A\oplus\bar{A})\otimes B]$ such that

$$\hat{H}^\varepsilon_{\min}(A|B)_\rho = \hat{H}_{\min}(A\oplus\bar{A}|B)_{\hat{\rho}},$$

with $\hat{\rho} \in \mathcal{B}^\varepsilon(\rho)$ (according to the purified distance), and $|\bar{A}| = \lceil \varepsilon\, 2^{\hat{H}^\varepsilon_{\min}(A|B)_\rho} \rceil$.

Proof. Let us choose the subnormalized state $\tilde{\rho}_{AB} \in \mathcal{S}_\leqslant(A\otimes B)$ that achieves the maximum in the definition of the entropy, i.e.,

$$\begin{gathered}
\lambda = \hat{H}^\varepsilon_{\min}(A|B)_\rho,\\
\tilde{\rho}_{AB} \leqslant 2^{-\lambda}\ \mathbb{1}_A \otimes \tilde{\rho}_B.
\end{gathered}$$

Now we construct the direct sum space $A\oplus\bar{A}$, where $\bar{A}$ is a Hilbert space of dimension $|\bar{A}| \geqslant \varepsilon\, 2^\lambda$. In that space, we write a normalized extension of $\tilde\rho$,

$$\hat\rho = \tilde\rho_{AB} \oplus \left((1-\operatorname{Tr}\tilde\rho)\frac{\mathbb{1}_{\bar A}}{|\bar A|} \otimes \tilde\rho_B \right) \quad \in \mathcal{S}[(A\oplus\bar A)\otimes B].$$

Note that $\hat\rho_B \propto \tilde\rho_B$. We have

$$\begin{aligned}\hat\rho &= \underbrace{\tilde\rho_{AB}}_{\leqslant 2^{-\lambda}\ \mathbb{1}_A\otimes\tilde\rho_B} \oplus \left(\underbrace{(1-\operatorname{Tr}\tilde\rho)}_{\leqslant\varepsilon}\frac{\mathbb{1}_{\bar A}}{|\bar A|} \otimes \tilde\rho_B \right)\\ &\leqslant (2^{-\lambda}\ \mathbb{1}_A \otimes \tilde\rho_B) \oplus \left(\frac{\varepsilon}{\varepsilon\, 2^{\lambda}}\mathbb{1}_{\bar A}\otimes\tilde\rho_B\right)\\ &= 2^{-\lambda}(\mathbb{1}_A\oplus\mathbb{1}_{\bar A})\otimes\tilde\rho_B.\end{aligned}$$

This tells us that λ is a feasible candidate for the primal SDP of the entropy of $\hat\rho$, and therefore

$$\hat H_{\min}(A\oplus\bar A|B)_{\hat\rho} \geqslant \lambda = \hat H^{\varepsilon}_{\min}(A|B)_\rho.$$

Now we show that $\hat\rho \in \mathcal{B}^\varepsilon(\rho)$, according to the purified distance. It suffices to show that $F(\rho,\hat\rho) = F(\rho,\tilde\rho)$. The fidelity is linear under direct sums, i.e., for two states $\sigma = \sigma_1\oplus\sigma_2$ and $\tau = \tau_1\oplus\tau_2$, we have

$$F(\sigma,\tau) = \|\sqrt{\sigma}\sqrt{\tau}\|_1 = \|\sqrt{\sigma_1}\sqrt{\tau_1}\|_1 + \|\sqrt{\sigma_2}\sqrt{\tau_2}\|_1.$$

In our case, we have $\rho = \rho_{AB}\oplus 0_{\bar AB}$ and $\hat\rho = \tilde\rho_{AB}\oplus((1-\operatorname{Tr}\tilde\rho)\frac{\mathbb{1}_{\bar A}}{|\bar A|}\otimes\tilde\rho_B)$, so

$$\begin{aligned}F(\rho,\hat\rho) &= \|\sqrt{\rho}\sqrt{\hat\rho}\|_1\\ &= \|\sqrt{\rho_{AB}}\sqrt{\tilde\rho_{AB}}\|_1\\ &\quad + \|\sqrt{0_{\bar AB}}\sqrt{\left((1-\operatorname{Tr}\tilde\rho)\frac{\mathbb{1}_{\bar A}}{|\bar A|}\otimes\tilde\rho_B\right)}\|_1\\ &= F(\rho,\tilde\rho)+0,\end{aligned}$$

which implies $\hat\rho\in\mathcal{B}^\varepsilon(\rho)$. Therefore we have, by definition of the smooth entropy,

$$\hat H_{\min}(A\oplus\bar A|B)_{\hat\rho} \leqslant \hat H^\varepsilon_{\min}(A|B)_\rho,$$

and the equality follows. ■

4. Smoothness of the hypothesis-testing entropy and relation to the smooth min entropy

The next lemma proves that the generalized smooth entropy H^ε is actually "smooth," i.e., if two states ρ and σ are close according to the trace distance, then their generalized smooth entropies are also close.

Lemma D.7. Let $\rho_{AB},\sigma_{AB}\in\mathcal{S}(A\otimes B)$ be two positive, normalized density operators, such that $\|\rho_{AB}-\sigma_{AB}\|_1\leqslant\delta$. Then, for any $\varepsilon>0$,

$$H^\varepsilon(A|B)_\rho \leqslant H^{\varepsilon+\delta+2\sqrt{2\delta}}(A|B)_\sigma + \log_2\frac{\varepsilon+\delta+2\sqrt{2\delta}}{\varepsilon}.$$

Proof. This proof is made of two parts. First we will relate $H^\varepsilon(A|B)_\rho$ to $-D^\varepsilon(\rho_{AB}-\Delta'||\mathbb{1}_A\otimes\sigma_B)$, where Δ' is a positive operator with trace at most $2\sqrt{2\delta}$. Later we bound $-D^\varepsilon(\rho_{AB}-\Delta'||\mathbb{1}_A\otimes\sigma_B)$ in terms of $H^{\varepsilon+\delta+2\sqrt{2\delta}}(A|B)_\sigma$.

We have

$$\|\rho_B-\sigma_B\|_1 \leqslant \|\rho_{AB}-\sigma_{AB}\|_1\leqslant\delta,$$

and therefore there exist positive operators Δ^+ and Δ^- such that

$$\begin{gathered}\rho_B-\sigma_B = \Delta^+-\Delta^-,\qquad \Delta^+,\Delta^-\geqslant 0,\\ \operatorname{Tr}(\Delta^+),\quad \operatorname{Tr}(\Delta^-)\leqslant\delta.\\ \rho_B\leqslant\sigma_B+\Delta^+.\end{gathered}$$

Consider the pair (μ,X) that forms the optimal solution of the dual SDP for $H^\varepsilon(A|B)_\rho$,

$$2^{H^\varepsilon(A|B)_\rho} = \mu - \operatorname{Tr}\frac{X}{\varepsilon},\qquad \mu\,\rho_{AB}\leqslant\mathbb{1}_A\otimes\rho_B+X_{AB}.$$

We can define the operator

$$G = \sigma_B^{\frac12}(\sigma_B+\Delta^+)^{-1/2}.$$

We conjugate the feasibility condition for the dual program with $\mathbb{1}_A\otimes G$),

$$\begin{aligned}&\mu\,\rho_{AB}\leqslant\mathbb{1}_A\otimes\rho_B+X_{AB},\quad \mu\,(\mathbb{1}\otimes G)\,\rho_{AB}\,(\mathbb{1}\otimes G^\dagger)\\ &\qquad\leqslant\mathbb{1}_A\otimes G\,\rho_B\,G^\dagger+(\mathbb{1}\otimes G)X_{AB}(\mathbb{1}\otimes G^\dagger).\end{aligned}$$

On the right-hand side, we have

$$\begin{aligned}G\,\rho_B\,G^\dagger &\leqslant G\,(\sigma_B+\Delta^+)G^\dagger\\ &= \sigma_B^{1/2}(\sigma_B+\Delta^+)^{-1/2}(\sigma_B+\Delta^+)(\sigma_B+\Delta^+)^{-1/2}\sigma_B^{1/2}\\ &=\sigma_B.\end{aligned}$$

Note also that G is a contraction (see Lemma D.4), and therefore $\operatorname{Tr}[(\mathbb{1}\otimes G)X_{AB}(\mathbb{1}\otimes G^\dagger)]\leqslant\operatorname{Tr}(X_{AB})$. On the left-hand side, we apply Lemma D.5,

$$\begin{aligned}&\|\rho_{AB}-(\mathbb{1}_A\otimes G_B)\rho_{AB}(\mathbb{1}_A\otimes G_B^\dagger)\|_1\leqslant 2\sqrt{2\operatorname{Tr}\Delta^+}\\ &\quad\Rightarrow\quad \exists\Delta':\quad \rho_{AB}-\Delta'\leqslant(\mathbb{1}_A\otimes G_B)\rho_{AB}(\mathbb{1}_A\otimes G_B^\dagger),\\ &\qquad\quad \Delta'\geqslant 0,\ \operatorname{Tr}(\Delta')\leqslant 2\sqrt{2\operatorname{Tr}\Delta^+}\leqslant 2\sqrt{2\delta}.\end{aligned}$$

This gives us

$$\mu\,(\rho_{AB}-\Delta')\leqslant\mathbb{1}_A\otimes\sigma_B+\underbrace{(\mathbb{1}\otimes G)X_{AB}(\mathbb{1}\otimes G^\dagger)}_{=:X'\geqslant 0}.$$

The above inequality tells us that (μ,X') form a candidate pair for the dual SDP of $D^\varepsilon(\rho_{AB}-\Delta'||\mathbb{1}_A\otimes\sigma_B)$. Note that $\rho_{AB}-\Delta'$ is Hermitian (as both ρ_{AB} and Δ' are positive operators, and therefore Hermitian), so both primal and dual SDPs for $D^\varepsilon(\rho_{AB}-\Delta'||\mathbb{1}_A\otimes\sigma_B)$ are well defined. Since the dual program is a maximization over all feasible pairs, we have

$$\begin{aligned}2^{-D^\varepsilon(\rho_{AB}-\Delta'||\mathbb{1}_A\otimes\sigma_B)} &\geqslant \mu-\frac{\operatorname{Tr}\ X'}{\varepsilon}\\ &=\mu-\frac{\operatorname{Tr}[(\mathbb{1}\otimes G)X_{AB}(\mathbb{1}\otimes G^\dagger)]}{\varepsilon}\\ &\geqslant\mu-\frac{\operatorname{Tr}(X_{AB})}{\varepsilon}\\ &=2^{H^\varepsilon(A|B)_\rho}.\end{aligned}$$

This gives us the bound

$$-D^\varepsilon(\rho_{AB}-\Delta'||\mathbb{1}_A\otimes\sigma_B)\geqslant H^\varepsilon(A|B)_\rho. \tag{D1}$$

Now we just need to relate $D^\varepsilon(\rho_{AB}-\Delta'||\mathbb{1}_A\otimes\sigma_B)$ to the smooth conditional entropy of σ, $H^{\varepsilon'}(A|B)_\sigma$ (which, as we will see, might have a different smoothing parameter, ε'). First we observe that the operator $\rho_{AB} - \Delta'$ is close to σ,

$$\begin{aligned}\|(\rho_{AB} - \Delta') - \sigma_{AB}\|_1 &\leqslant \delta + \operatorname{Tr}\Delta' \\ &\leqslant \underbrace{\delta + 2\sqrt{2\delta}}_{=:\delta'}.\end{aligned}$$

To shorten notation, it is convenient to define $\delta' := \delta + 2\sqrt{2\delta}$. The trace distance gives us an upper bound for the probability of distinguishing two states by applying any POVM $\{Q, \mathbb{1} - Q\}$,

$$\max_{0 \leqslant Q \leqslant \mathbb{1}} |\operatorname{Tr}(Q\,[\rho_{AB} - \Delta']) - \operatorname{Tr}(Q\,\sigma_{AB})| \leqslant \delta'. \tag{D2}$$

We start by writing down the primal SDP for $2^{H^{\varepsilon+\delta'}(A|B)_\sigma}$,

$$\begin{aligned}&\text{minimize} \frac{1}{\varepsilon + \delta'} \operatorname{Tr}(Q\ \mathbb{1}_A \otimes \sigma_B), \quad \text{subject to} \quad \operatorname{Tr}(Q\ \sigma_{AB}) \\ &\geqslant \varepsilon + \delta', 0 \leqslant Q \leqslant \mathbb{1}.\end{aligned}$$

We take the operator Q that achieves the minimum, and show that Q is a feasible candidate for the primal SDP of $2^{-D^\varepsilon(\rho_{AB} - \Delta' || \mathbb{1}_A \otimes \sigma_B)}$. To make that clear, let us first write this SDP,

$$\begin{aligned}\text{minimize} \quad & \frac{1}{\varepsilon} \operatorname{Tr}(P\ \mathbb{1}_A \otimes \sigma_B), \\ \text{subject to} \quad & \operatorname{Tr}(P\ [\rho_{AB} - \Delta']) \geqslant \varepsilon, \\ & 0 \leqslant P \leqslant \mathbb{1}.\end{aligned}$$

We may relate the feasibility conditions of the two SDPs using inequality (D2), which gives us

$$\begin{aligned}\operatorname{Tr}(Q\ [\rho_{AB} - \Delta']) &\geqslant \operatorname{Tr}(Q\ \sigma_{AB}) - \delta' \\ &\geqslant \varepsilon + \delta' - \delta' = \varepsilon.\end{aligned}$$

Therefore we can bound $2^{-D^\varepsilon(\rho_{AB} - \Delta' || \mathbb{1}_A \otimes \sigma_B)}$ as

$$\begin{aligned}2^{-D^\varepsilon(\rho_{AB} - \Delta' || \mathbb{1}_A \otimes \sigma_B)} &\leqslant \frac{1}{\varepsilon} \operatorname{Tr}(Q\ \mathbb{1}_A \otimes \sigma_B) \\ &= \frac{1}{\varepsilon}(\varepsilon + \delta') 2^{H^{\varepsilon+\delta'}(A|B)_\sigma}.\end{aligned}$$

Taking the logarithm and using $\delta' = \delta + 2\sqrt{2\delta}$, we obtain

$$\begin{aligned}H^{\varepsilon+\delta+2\sqrt{2\delta}}(A|B)_\sigma &\geqslant -D_H^\varepsilon(\rho_{AB} - \Delta' || \mathbb{1}_A \otimes \sigma_B) \\ &\quad - \log_2 \frac{\varepsilon + \delta + 2\sqrt{2\delta}}{\varepsilon} \\ &\geqslant H^\varepsilon(A|B)_\rho - \log_2 \frac{\varepsilon + \delta + 2\sqrt{2\delta}}{\varepsilon}.\end{aligned}$$

■

where we used Eq. (D1) in the last step. In the following lemma, we find a lower bound for H^ε in terms of the smooth min entropy. An upper bound is given in [21, Prop. 4.1].

Lemma D.8. Let $\rho \in \mathcal{S}(A \otimes B)$, and let $\varepsilon \in]0, \frac{1}{2}]$. Then,

$$H_{\min}^\varepsilon(A|B)_\rho \leqslant H^{11\sqrt{\varepsilon}}(A|B)_\rho - \frac{5}{2}\log_2\left(\frac{3}{\varepsilon}\right) + \log_2\left(\frac{2}{1-\varepsilon}\right).$$

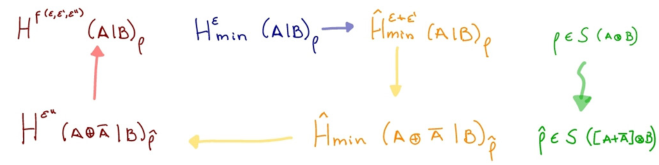

FIG. 5. Proof steps for Lemma D.8. We start from $H_{\min}^\varepsilon(A|B)_\rho$ and bound it successively until we end up with the generalized smooth entropy for the same state. Along the way we need to extend our state ρ to a larger Hilbert space (below).

Proof. See Fig. 5 for a schematic representation of the different steps of this proof.

From [44, Lemma 19] we have that

$$\begin{aligned}&\forall \rho \in \mathcal{S}(A \otimes B), \forall \varepsilon, \varepsilon' \in]0,1], \quad H_{\min}^\varepsilon(A|B)_\rho \\ &\leqslant \hat{H}_{\min}^{\varepsilon+\varepsilon'}(A|B)_\rho + \log_2\left(\frac{2}{\varepsilon'^2} + \frac{1}{1-\varepsilon}\right).\end{aligned}$$

Now we use Lemma D.6 to find a normalized state $\hat\rho \in \mathcal{B}^{\varepsilon+\varepsilon'}(\rho)$ in a larger Hilbert space $(A \oplus \bar{A}) \otimes B$ that attains the optimization. This gives us

$$\begin{aligned}H_{\min}^\varepsilon(A|B)_\rho &\leqslant \hat{H}_{\min}^{\varepsilon+\varepsilon'}(A|B)_\rho + \log_2\left(\frac{2}{\varepsilon'^2} + \frac{1}{1-\varepsilon}\right) \\ &= \hat{H}_{\min}(A \oplus \bar{A}|B)_{\hat\rho} + \log_2\left(\frac{2}{\varepsilon'^2} + \frac{1}{1-\varepsilon}\right).\end{aligned}$$

It follows from the definition of $\hat{H}_{\min}$ that [21, Prop. 4.1]

$$\forall\, \varepsilon'' \in]0,1]: \quad \hat{H}_{\min}(A \oplus \bar{A}|B)_{\hat\rho} \leqslant H^{\varepsilon''}(A \oplus \bar{A}|B)_{\hat\rho},$$

which leaves us with

$$H_{\min}^\varepsilon(A|B)_\rho \leqslant H^{\varepsilon''}(A \oplus \bar{A}|B)_{\hat\rho} + \log_2\left(\frac{2}{\varepsilon'^2} + \frac{1}{1-\varepsilon}\right).$$

Now we only need to relate $H^{\varepsilon''}(A \oplus \bar{A}|B)_{\hat\rho}$ back to the smooth entropy of ρ. Since the two states $\rho, \hat\rho$ are normalized, we have $\|\rho - \hat\rho\|_1 \leqslant 2(\varepsilon + \varepsilon')$. We can use Lemma D.7 to obtain

$$\begin{aligned}H^{\varepsilon''}(A \oplus \bar{A}|B)_{\hat\rho} &\leqslant H^{\varepsilon''+2\varepsilon+2\varepsilon'+4\sqrt{\varepsilon+\varepsilon'}}(A \oplus \bar{A}|B)_\rho \\ &\quad + \log_2 \frac{\varepsilon'' + 2\varepsilon + 2\varepsilon' + 4\sqrt{\varepsilon + \varepsilon'}}{\varepsilon''}.\end{aligned}$$

Now we observe that ρ has no support on $\bar{A}$, therefore, for any smoothing factor $\tilde\varepsilon \in [0,1]$,

$$H^{\tilde\varepsilon}(A \oplus \bar{A}|B)_\rho = H^{\tilde\varepsilon}(A|B)_\rho.$$

All in all, we have

$$\begin{aligned}H_{\min}^\varepsilon(A|B)_\rho &\leqslant H^{\varepsilon''+2\varepsilon+2\varepsilon'+4\sqrt{\varepsilon+\varepsilon'}}(A|B)_\rho \\ &\quad + \log_2\left(\frac{\varepsilon'' + 2\varepsilon + 2\varepsilon' + 4\sqrt{\varepsilon+\varepsilon'}}{\varepsilon''}\right) + \log_2\left(\frac{2}{\varepsilon'^2} + \frac{1}{1-\varepsilon}\right).\end{aligned}$$

To clean up, we consider the special case $\varepsilon = \varepsilon' = \varepsilon''$, which gives us

$$H_{\min}^{\varepsilon}(A|B)_\rho \leqslant H^{5\varepsilon+4\sqrt{2\,\varepsilon}}(A|B)_\rho + \log_2\left(5 + \frac{4\sqrt{2}}{\sqrt{\varepsilon}}\right) + \log_2\left(\frac{2}{\varepsilon^2} + \frac{1}{1-\varepsilon}\right),$$

and finally we upper bound the additive terms and smoothing factors with simpler terms (the factors were found numerically). We obtain

$$H_{\min}^{\varepsilon}(A|B)_\rho \leqslant H^{11\sqrt{\varepsilon}}(A|B)_\rho - \frac{5}{2}\log_2\left(\frac{\varepsilon}{3}\right) + \log_2\left(\frac{2}{1-\varepsilon}\right).$$

■

[1] S. Popescu, A. J. Short, and A. Winter, Entanglement and the foundations of statistical mechanics, Nat. Phys. **2**, 754 (2006).

[2] N. Linden, S. Popescu, A. J. Short, and A. Winter, Quantum mechanical evolution towards thermal equilibrium, Phys. Rev. E **79**, 061103 (2009).

[3] F. G. S. L. Brandão, P. Ćwikliski, M. Horodecki, P. Horodecki, J. K. Korbicz, and M. Mozrzymas, Convergence to equilibrium under a random Hamiltonian, Phys. Rev. E **86**, 031101 (2012).

[4] L. Masanes, A. J. Roncaglia, and A. Acin, Complexity of energy eigenstates as a mechanism for equilibration, Phys. Rev. E **87**, 032137 (2013).

[5] A. Riera, C. Gogolin, and J. Eisert, Thermalization in Nature and on a Quantum Computer, Phys. Rev. Lett. **108**, 080402 (2012).

[6] J. Eisert, M. Friesdorf, and C. Gogolin, Quantum many-body systems out of equilibrium, Nat. Phys. **11**, 124 (2015).

[7] J. Goold, M. Huber, A. Riera, L. del Rio, and P. Skrzypczyk, The role of quantum information in thermodynamicsa topical review, J. Phys. A: Math. Theor. **49**, 143001 (2016).

[8] M. Hossein Partovi, Quantum thermodynamics, Phys. Lett. A **137**, 440 (1989).

[9] D. Jennings and T. Rudolph, Entanglement and the thermodynamic arrow of time, Phys. Rev. E **81**, 061130 (2010).

[10] S. Jevtic, D. Jennings, and T. Rudolph, Maximally and Minimally Correlated States Attainable Within a Closed Evolving System, Phys. Rev. Lett. **108**, 110403 (2012).

[11] R. Clausius, *The Mechanical Theory of Heat: With its Applications to the Steam-engine and to the Physical Properties of Bodies* (J. Van Voorst, London, 1867), p. 376.

[12] P. Hayden and J. Preskill, Black holes as mirrors: Quantum information in random subsystems, J. High Energy Phys. 09 (2007) 120.

[13] F. Dupuis, The decoupling approach to quantum information theory, Ph.D. thesis, Université de Montréal, 2010.

[14] F. Dupuis, M. Berta, J. Wullschleger, and R. Renner, One-Shot Decoupling, Commun. Math. Phys. **328**, 251 (2014).

[15] O. Szehr, F. Dupuis, M. Tomamichel, and R. Renner, Decoupling with unitary almost two-designs, New J. Phys. **15**, 053022 (2013).

[16] F. G. S. L. Brandão, A. W. Harrow, and M. Horodecki, Local random quantum circuits are approximate polynomial-designs, arXiv:1208.0692.

[17] W. Brown and O. Fawzi, Scrambling speed of random quantum circuits, arXiv:1210.6644.

[18] W. Brown and O. Fawzi, Decoupling with random quantum circuits, Commun. Math. Phys. **340**, 867 (2015).

[19] Y. Nakata, C. Hirche, C. Morgan, and A. Winter, Decoupling with random diagonal-unitary matrices, arXiv:1509.05155.

[20] Y. Nakata, C. Hirche, C. Morgan, and A. Winter, Implementing unitary 2-designs using random diagonal-unitary matrices, arXiv:1502.07514.

[21] F. Dupuis, L. Krämer, P. Faist, J. M. Renes, and R. Renner, Generalized entropies, in *Proceedings of the XVIIth International Congress on Mathematical Physics* (Aalborg, Denmark, 2012), Chap. 9, pp. 134–153.

[22] R. Renner, Security of quantum key distribution, Ph.D. thesis, ETH Zurich, 2005.

[23] M. Tomamichel, A framework for non-asymptotic quantum information theory, Ph.D. thesis, ETH Zurich, 2012.

[24] L. del Rio, J. Aberg, R. Renner, O. Dahlsten, and V. Vedral, The thermodynamic meaning of negative entropy, Nature (London) **474**, 61 (2011).

[25] N. Brunner, M. Huber, N. Linden, S. Popescu, R. Silva, and P. Skrzypczyk, Entanglement enhances cooling in microscopic quantum refrigerators, Phys. Rev. E **89**, 032115 (2014).

[26] P. Faist, F. Dupuis, J. Oppenheim, and R. Renner, The minimal work cost of information processing, Nat. Commun. **6**, 7669 (2015).

[27] D. Reeb and M. M. Wolf, An improved Landauer principle with finite-size corrections, New J. Phys. **16**, 103011 (2014).

[28] M. Lostaglio, D. Jennings, and T. Rudolph, Thermodynamic resource theories, non-commutativity and maximum entropy principles, arXiv:1511.04420.

[29] Y. Guryanova, S. Popescu, A. J. Short, R. Silva, and P. Skrzypczyk, Thermodynamics of quantum systems with multiple conserved quantities, arXiv:1512.01190.

[30] M. Perarnau-Llobet, A. Riera, R. Gallego, H. Wilming, and J. Eisert, Work and entropy production in generalised Gibbs ensembles, arXiv:1512.03823.

[31] N. Y. Halpern, P. Faist, J. Oppenheim, and A. Winter, Microcanonical and resource-theoretic derivations of the Non-Abelian Thermal State, arXiv:1512.01189.

[32] C. Ududec, N. Wiebe, and J. Emerson, Information-Theoretic Equilibration: The Appearance of Irreversibility Under Complex Quantum Dynamics, Phys. Rev. Lett. **111**, 080403 (2013).

[33] P. Reimann, Equilibration of isolated macroscopic quantum systems under experimentally realistic conditions, Phys. Scr. **86**, 058512 (2012).

[34] R. A. Low, Pseudo-randomness and learning in quantum computation, Ph.D. thesis, University of Bristol, 2010.

[35] L. del Rio, L. Krämer, and R. Renner, Resource theories of knowledge, arXiv:1511.08818.

[36] J. M. Renes and R. Renner, One-shot classical data compression with quantum side information and the distillation of common randomness or secret keys, IEEE Trans. Inf. Theory **58**, 1985 (2012).
[37] R. König, R. Renner, and C. Schaffner, The operational meaning of min- and max-entropy, IEEE Trans. Inf. Theory **55**, 4337 (2009).
[38] N. Datta and R. Renner, Smooth entropies and the quantum information spectrum, IEEE Trans. Inf. Theory **55**, 2807 (2009).
[39] M. Tomamichel, R. Colbeck, and R. Renner, Duality between smooth min- and max-entropies, IEEE Trans. Inf. Theory **56**, 4674 (2010).
[40] A. Vitanov, F. Dupuis, M. Tomamichel, and R. Renner, Chain Rules for Smooth Min- and Max-Entropies, IEEE Trans. Inf. Theory **59**, 2603 (2013).
[41] M. Tomamichel, R. Colbeck, and R. Renner, A fully quantum asymptotic equipartition property, IEEE Trans. Inf. Theory **55**, 5840 (2009).
[42] S. Boyd and L. Vandenbergh, *Convex Optimization* (Cambridge University Press, Cambridge, England, 2004).
[43] J. Watrous, Semidefinite programs for completely bounded norms, Theory Comput. **5**, 217 (2009).
[44] M. Tomamichel, C. Schaffner, A. Smith, and R. Renner, Leftover hashing against quantum side information, IEEE Trans. Inf. Theory **57**, 5524 (2011).