Exponential Decay of Correlations Implies Area Law

Exponential Decay of Correlations Implies Area Law

Fernando G. S. L. Brandão[1], **Michał Horodecki**[2]

[1] Institute for Theoretical Physics, ETH Zürich, Zürich 8093, Switzerland.
   E-mail: fgslbrandao@gmail.com
[2] Institute for Theoretical Physics and Astrophysics, University of Gdańsk, Gdańsk 80-952, Poland.
   E-mail: fizmh@ug.edu.pl

Abstract: We prove that a finite correlation length, i.e., exponential decay of correlations, implies an area law for the entanglement entropy of quantum states defined on a line. The entropy bound is exponential in the correlation length of the state, thus reproducing as a particular case Hastings's proof of an area law for groundstates of 1D gapped Hamiltonians.

As a consequence, we show that 1D quantum states with exponential decay of correlations have an efficient classical approximate description as a matrix product state of polynomial bond dimension, thus giving an equivalence between injective matrix product states and states with a finite correlation length.

The result can be seen as a rigorous justification, in one dimension, of the intuition that states with exponential decay of correlations, usually associated with non-critical phases of matter, are simple to describe. It also has implications for quantum computing: it shows that unless a pure state quantum computation involves states with long-range correlations, decaying at most algebraically with the distance, it can be efficiently simulated classically.

The proof relies on several previous tools from quantum information theory—including entanglement distillation protocols achieving the hashing bound, properties of single-shot smooth entropies, and the quantum substate theorem—and also on some newly developed ones. In particular we derive a new bound on correlations established by local random measurements, and we give a generalization to the max-entropy of a result of Hastings concerning the saturation of mutual information in multiparticle systems. The proof can also be interpreted as providing a limitation on the phenomenon of data hiding in quantum states.

1. Introduction

Quantum states of many particles are fundamental to our understanding of many-body physics. Yet they are extremely daunting objects, requiring in the worst case an expo-
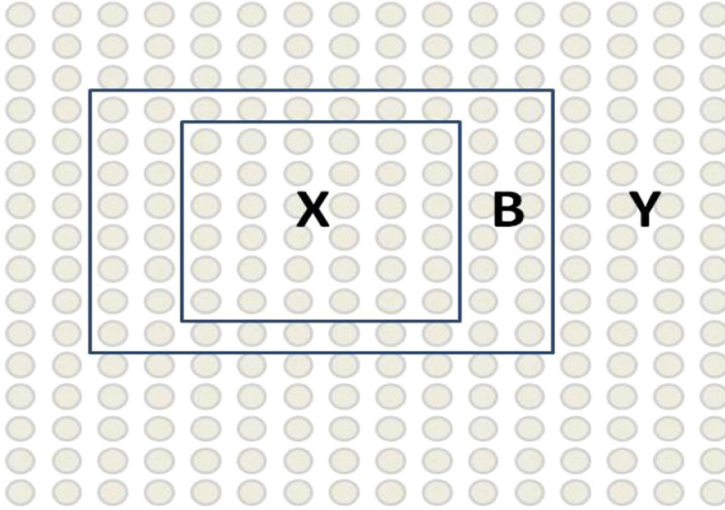
Fig. 1. Example of a partition of the lattice into regions X, B and Y. The separation between X and Y in the example consists of one site

nential number of parameters in the number of subsystems to be even approximately described. How then can multiparticle quantum states be useful for giving predictions to physical observables? The intuitive explanation, based on several decades of developments in condensed matter physics and more recently also on complementary input from quantum information theory, is that physically relevant quantum states, defined as the ones appearing in nature, are usually much simpler than generic quantum states. In this paper we prove a new theorem about quantum states that gives further justification to this intuition.

One physically meaningful way of limiting the set of quantum states is to put restrictions on their *correlations*. Given a bipartite quantum state ρ_{XY}, we can quantify the correlations between X and Y by

$$\mathrm{Cor}(X : Y) := \max_{\|M\| \leq 1, \|N\| \leq 1} |\mathrm{tr}((M \otimes N)(\rho_{XY} - \rho_X \otimes \rho_Y))|, \tag{1}$$

where $\|M\|$ is the operator norm of M. Such a correlation function generalizes the more well-known two-point correlation function, widely studied in condensed matter physics, in which both X and Y are composed of a single site.

We say a quantum state $\rho_{1,\ldots,n}$ composed of n qubits defined on a finite dimensional lattice has (ξ, l_0)-exponential decay of correlations if for every $l \geq l_0$ and every two regions X and Y separated by more than l sites (see Fig. 1),

$$\mathrm{Cor}(X : Y) \leq 2^{-l/\xi}. \tag{2}$$

Here ξ is the correlation length of the state and l_0 the minimum length for which correlations start decreasing. Such a form of exponential decay of correlations is sometimes also termed the exponential clustering property (see e.g. [1–3]).

An important class of quantum states with exponential decay of correlations are groundstates of non-critical, gapped, local quantum Hamiltonians. In a seminal work,

Hastings proved that in any fixed dimension the groundstate of a gapped local Hamiltonian exhibits exponential decay of correlations, with l_0 of unit order and the correlation length ξ of order of the inverse spectral gap of the model [4] (see also [3,5–7] for subsequent developments and [1,2] for a previous analogous result in the context of relativistic quantum systems). Recently a simpler proof of exponential decay of correlations, of a combinatorial flavour, was given in Ref. [8] for the class of frustration-free models with a unique groundstate.

Another physically motivated way of limiting the set of quantum states is to put restrictions on their *entanglement* [9]. Given a quantum state $|\psi\rangle$ we say it satisfies an area law for entanglement entropy if for every region X, $S(\rho_X) \leq O(\partial X)$, where ρ_X is the reduced density matrix of $|\psi\rangle$ in the region X, $S(X) := -\text{tr}(\rho_X \log \rho_X)$ the von Neumann entropy, and ∂X the perimeter of the region X [10]. The term entanglement entropy is due to the interpretation of $S(X)$ as the amount of entanglement between X and $\overline{X}$ (the complementary region) in the quantum pure state $|\psi\rangle_{X\overline{X}}$ [11]. Generic states typically have an extensive behaviour of entanglement entropy. Thus states satisfying an area law have a strong restriction on their amount of entanglement.

What kind of quantum states are expected to obey an area law? Starting with the seminal work of Bekenstein in the context of black hole entropy [12], and more recently in the context of quantum spin systems [13,14] and quantum harmonic systems [15–17] (see [10] for more references), an increasing body of evidence appeared suggesting that states corresponding to the ground or to low-lying energy eigenstates of local models satisfy an area law.[1],[2] It turns out however that this is not always the case: One can construct local Hamiltonians, even in one dimension [18] and with translational symmetry [19,20], whose groundstate entanglement entropy follows a volume law. Yet, these are critical models, with the spectral gap shrinking to zero with the number of sites. Gapped models, on the other hand, are expected to satisfy an area law, although a general proof in arbitrary dimensions is still lacking.

In a ground-breaking work, Hastings proved that this is indeed the case for one-dimensional systems, i.e., 1D gapped Hamiltonians with a unique groundstate always obey an area law [21]. Hastings' proof is based on locality estimates in quantum spin Hamiltonians known as Lieb-Robinson bounds [22] and gives an exponential dependency between the entanglement entropy and the spectral gap of the model. Recently, Arad et al. [23,24] found a combinatorial proof of the area law for groundstates of 1D gapped models, with a bound on entanglement entropy only polynomially large in the spectral gap, matching (up to polynomial factors) explicit examples [19,25].

Both proofs of the area law explore extensively the fact that the quantum state is a groundstate of a local Hamiltonian with a constant spectral gap, which gives much more structure than merely the fact that the state has exponential decay of correlations. However, as pointed out already in Ref. [26], exponential decay of correlations by itself already suggests the entanglement of the state should satisfy an area law. Indeed, consider a quantum state $|\psi\rangle_{XBY}$ as in Fig. 1, with B the boundary region between X and Y. If $|\psi\rangle$ has exponential decay of correlations and the separation between X and Y is of order of the correlation length of the state, X will have almost no correlations with Y,

---

[1] With a logarithmic correction in the case of critical systems (see e.g. Ref. [10]).

[2] For thermal states of local Hamiltonians the situation is simpler: Jaynes' principle of maximum entropy implies a general area law for mutual information valid in arbitrary dimensions, with the mutual information of a region with the complementary region being proportional to the perimeter of the region and the inverse temperature [33].
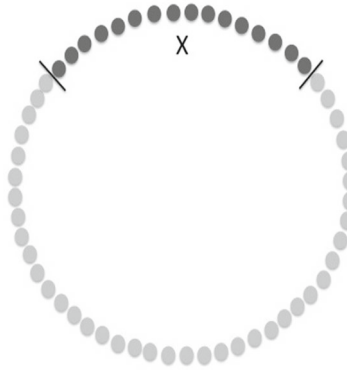
Fig. 2. Arrangement of qubits on a ring considered in Theorem 1

and one would expect the entanglement of X with BY to be only due to correlations with the region B, thus obeying an area law.

Perhaps surprisingly, and as also pointed out in Ref. [26], the argument presented above is flawed. This is because there are quantum states for which X and Y have almost no correlations, yet X has very large entropy. This is not only a pathological case among quantum states, but it is actually the general rule: the overwhelming majority of quantum states will have such peculiar type of correlations, whenever the regions X and Y have approximately equal sizes [27]. Quantum states with this property are termed *quantum data hiding* states, due to their use in hiding correlations (and information) from local measurements [28]. Developing a classification of entanglement in such states is one of the outstanding challenges in our understanding of quantum correlations.[3]

We therefore see that it is not possible to obtain an area law from exponential decay of correlations by considering only one fixed partition of the system (into XBY regions). However, this leaves open the possibility that an area law could be established by exploring exponential decay of correlations simultaneously in several *different* partitions of the system.

2. Results

Area Law for Pure States: Quantum data hiding states, and the related quantum expander states [31,32], have been largely recognized as an obstruction for obtaining an area law for entanglement entropy from exponential decay of correlations (see e.g. [8,21,24,26,31–35]). In this paper we show that such an implication, at least for states defined on a 1-dimensional lattice, is in fact *false*. Our main result is the following (see Fig. 2):

Theorem 1. *Let* $|\psi\rangle_{1,...,n}$ *be a state defined on a ring with* (ξ, l_0)*-exponential decay of correlations and* $n \geq Cl_0/\xi$*. Then for any connected region* $X \subset [n]$ *and every* $l \geq 8\xi$*,*

$$H_{\max}^{2^{-\frac{l}{8\xi}}}(X) \leq c'l_0 \exp(c \log(\xi)\xi) + l, \tag{3}$$

with $C, c, c' > 0$ *universal constants.*

---

[3] As a concrete example, in the problem of determining if a quantum bipartite state ρ_{XY} is entangled or not, quantum data hiding states appear to be the hardest instances from a computational point of view [29,30].

An announcement of the proof can be found in [36].

We leave it as an open question whether a similar statement holds true in dimensions larger than one.

The quantity $H_{\max}^\varepsilon(X)$ is the ε-smooth max-entropy of $\rho_X := \mathrm{tr}_{\backslash X}(|\psi\rangle\langle\psi|)$, where $\mathrm{tr}_{\backslash X}(|\psi\rangle\langle\psi|)$ denotes the partial trace of $|\psi\rangle\langle\psi|$ with respect to all sites expect the ones in region X. The ε-smooth max-entropy is defined as [37,38]

$$H_{\max}^\varepsilon(X) = H_{\max}^\varepsilon(\rho_X) := \min_{\tilde\rho_X \in \mathcal{B}_\varepsilon(\rho_X)} \log\left(\mathrm{rank}(\tilde\rho_X)\right), \tag{4}$$

with $B_\varepsilon(\rho_X) := \{\tilde\rho_X : D(\rho_X, \tilde\rho_X) \le \varepsilon\}$ the ε-ball of quantum states around ρ_X and $D(\sigma_1, \sigma_2)$ the purified distance between states σ_1 and σ_2.[4] In words, $H_{\max}^\varepsilon(X)$ gives the logarithm of the approximate support of ρ_X, i.e., the number of qubits needed to store an ε-approximation of the state ρ_X.

We now present two simple corollaries of Theorem 1. The first is an area law for the von Neumann entropy:

Corollary 2. *Let $|\psi\rangle_{1,\ldots,n}$ be a state defined on a ring with (ξ, l_0)-exponential decay of correlations and $n \ge Cl_0/\xi$. Then for any connected region $X \subset [n]$,*

$$S(X) \le c'l_0 \exp\left(c\log(\xi)\xi\right), \tag{5}$$

with $C, c, c' > 0$ universal constants.

Note that by combining Corollary 2 with the fact that groundstates of gapped local Hamiltonians have exponential decay of correlations [4], we recover Hastings' area law for groundstates of gapped models, with the same exponential dependence of the entropy on the inverse spectral gap of the model [21].

The second corollary gives an approximation of a state $|\psi\rangle_{1,\ldots,n}$ in terms of a matrix product state of small bond dimension. A matrix product representation of the state $|\psi\rangle_{1,\ldots,n}$ is given by

$$|\psi\rangle_{1,\ldots,n} = \sum_{i_1=1}^{d} \cdots \sum_{i_n=1}^{d} \mathrm{tr}(A_{i_1}^{[1]} \ldots A_{i_n}^{[n]})|i_1, \ldots, i_n\rangle, \tag{6}$$

with $D \times D$ matrices $A^{[j]}$, $j \in [n]$. The parameter D is termed the bond dimension and measures the complexity of the matrix product representation. When $D = \mathrm{poly}(n)$ the quantum state $|\psi\rangle_{1,\ldots,n}$ admits an efficient classical description in terms of its matrix product representation, with only polynomially many parameters and in which expectation values of local observables can be calculated efficiently. We call such states themselves matrix product states (MPS) [39–42].

It is known that every so-called injective matrix product state (which is the generic case) has a finite correlation length. Combining Theorem 1 and the results of Ref. [41] (see also [26]) we find that the converse statement also holds true:

Corollary 3. *Let $|\psi\rangle_{1,\ldots,n}$ be a state defined on a ring with (ξ, l_0)-exponential decay of correlations. For every $\delta > 0$, there is a matrix product state $|\phi_k\rangle$ of bond dimension $k = \mathrm{poly}(n^{\xi\log(\xi)}, \delta)$ such that $|\langle\psi|\phi_k\rangle| \ge 1 - \delta$.*

---

[4] See "Appendix A" for the definition of the purified distance.

Thus we see that one-dimensional pure quantum states with exponential decay of correlations have a very simple structure, admitting an efficient classical parametrization. In fact, such a representation is of more than theoretical interest, as matrix product states are the variational ansatz for the most successful known method for simulating one-dimensional quantum systems, the Density Matrix Renormalization Group (DMRG) [40,43].

Area Law for Mixed States: So far we have focused on the case of pure states satisfying exponential decay of correlations. How about mixed states? It is clear that Theorem 1 cannot be true in this case. As an example, consider the maximally mixed state: It has no correlations, yet the local entropies are all maximum. Although we do not present a full answer to the mixed state case in this paper, we can prove the following extension of Theorem 1:

Theorem 4. *Let $\rho_{1,...,n}$ be a state defined on ring with (ξ, l_0)-exponential decay of correlations and $n \geq C l_0 / \xi$. Then for any connected region $X \subset [n]$ and every $l \geq 8\xi$,*

$$H_{\max}^{2^{-\frac{l}{8\xi}}}(X) \leq c' l_0 \exp\left(c \log(\xi)\xi\right)(1 + H_{\max}(\rho)) + l, \tag{7}$$

with $C, c, c' > 0$ universal constants.

Long-range Correlations in Quantum Computation: A perennial question in quantum information science is to understand what is responsible for the apparent superiority of quantum computation over classical computation. A fruitful approach in this direction is to find conditions under which quantum circuits have an efficient classical simulation (see e.g. [41,46–52]). In this way one can at least say what properties a quantum circuit *should* have if it is supposed to give a superpolynomial speed-up over classical computing.

In [41] Vidal gave an interesting result in this context: Unless a quantum computation in the circuit model involves states that violate an area law, with the entropy of a certain subregion being bigger than the logarithm of the number of boundary qubits, it can be simulated classically in polynomial time. A direct corollary of this result and Theorem 1 is the following:

Corollary 5. *Consider a family of quantum circuits $V = V_k, \ldots, V_2 V_1$ acting on n qubits arranged in a ring and composed of two qubit gates V_k. Let $|\psi_t\rangle := V_t, \ldots, V_2 V_1 |0^n\rangle$ be the state after the t-th gate has been applied. Then if there are constants ξ, l_0 independent of n such that, for all n and $t \in [n]$, $|\psi_t\rangle$ has (ξ, l_0)-exponential decay of correlations, one can classically simulate the evolution of the quantum circuit in* poly(n, k) *time.*

The corollary says that one must have at least algebraically decaying correlations in a quantum circuit if it is supposed to solve a classically hard problem. Interestingly such long range correlations are usually associated to critical phases of matter. From a quantum information perspective, the result gives a limitation to the possibility of hiding information in 1D quantum circuits.

Random States and Quantum Expanders: It is interesting to analyse how Theorem 1 fits together with the entanglement properties of random states and quantum expander states, since these have been thought of as giving an obstruction to a statement of a similar flavour [8,21,24,26,31–35].
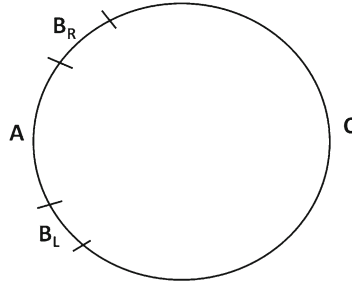
Fig. 3. Partition of the one-dimensinal lattice into regions A, B_L, B_R, and C

Random States: The situation for random quantum states is fairly simple: Even though they have exponentially small two-point correlations, they do *not* exhibit general exponential decay of correlations, and thus nothing prevents the extensive behaviour of entanglement entropy found in such states. Indeed we show in "Appendix D" that for a quantum state $|\psi\rangle_{1,\ldots,n}$ drawn from the Haar measure, with overwhelming probability, for every region X with $|X| \leq n/4$,

$$D(\rho_X, \tau_X) \leq 2^{-\Omega(n)}, \tag{8}$$

where τ_X is the maximally mixed state on X. Now let us consider the sites $\{1, \ldots, n\}$ arranged in a ring and divide the region X into two subregions A and B of equal size $n/8$, and call C the region composed by the remaining $3n/4$ sites. The region B is further divided into two subregions B_L and B_R of equal sizes, one to the left and the other to the right of A (see Fig. 3). In quantum information terminology, Eq. (8) says, that A is approximately *decoupled* from B, since their joint state is close to a product state [44]. From Uhlmann's theorem (Lemma 11 in "Appendix A") it thus follows that there is an isometry $V : C \to C_1 C_2$ that can be applied to C such that

$$D((\text{id}_{AB} \otimes V)|\psi\rangle_{ABC}\langle\psi|(\text{id}_{AB} \otimes V)^\dagger, |\Phi\rangle_{AC_1}\langle\Phi| \otimes |\Phi\rangle_{BC_2}\langle\Phi|) \leq 2^{-\Omega(n)}, \tag{9}$$

with $|\Phi\rangle_{AC_1} = \dim(A)^{-1/2} \sum_{k=1}^{\dim(A)} |k, k\rangle$ a maximally entangled state between A and C_1. Thus it is clear that the regions A and C, separated by $n/16$ sites, are correlated, since A and C share a maximally entangled state (and are separated by a distance given by the size of B_L and B_R). Indeed, choosing $M = \sum_{i=1}^{\dim(A)/2} |k\rangle\langle k|$ and $N = V(M \otimes \text{id}_{C_2}) V^\dagger$, we find from Eq. (9) that

$$\text{Cor}(A : C) \geq \text{tr}\Big((M_A \otimes N_C)(\rho_{AC} - \rho_A \otimes \rho_C)\Big) = 1/4 - 2^{-\Omega(n)}. \tag{10}$$

Expander States: Quantum expander states were introduced in Refs. [31,32] in order to find a more explicit example of states in which exponential decay of correlations does not directly imply an area law. Let $\Lambda : \mathcal{D}(\mathbb{C}^D) \to \mathcal{D}(\mathbb{C}^D)$, with $\mathcal{D}(\mathbb{C}^D)$ the set of density matrices acting on $\mathbb{C}^D$, be a quantum channel with Kraus decomposition $\Lambda(\rho) = \sum_{i=1}^d A_i \rho A_i^\dagger$ for $\rho \in \mathcal{D}(\mathbb{C}^D)$. We say it is an (η, d)-quantum expander if its maximum eigenvalue (seen Λ as a linear map) is one and its second largest eigenvalue is smaller than $\eta < 1$ [31,32,45] (we mean here eigenvalues of channel treated as a linear operator). Given a quantum expander, its associated quantum expander state is defined

as the translational-invariant matrix product state whose correlation matrices are given by the $\{A_i\}_{i=1}^d$:

$$|\psi\rangle_{1,\ldots,n} := \sum_{i_1=1}^d \cdots \sum_{i_n=1}^d \operatorname{tr}(A_{i_1} \ldots A_{i_n})|i_1, \ldots, i_n\rangle. \tag{11}$$

Following the original approach of Fannes, Nachtergaele and Werner [39] (see also [33]) we show in Appendix E that for regions A and C separated by l sites:

$$\operatorname{Cor}(A:C) \leq D\eta^l, \tag{12}$$

Thus the state $|\psi\rangle_{1,\ldots,n}$ has at least $\Big(O(\log^{-1}(1/\eta)), O(\log(D)\log^{-1}(1/\eta)) \Big)$-exponential decay of correlations.

For some choices of the matrices $\{A_i\}_{i=1}^d$, one can show that the previous bound on the correlation length is tight. For that we use the approach of Ref. [31], where it was shown that choosing $A_i = U_i/\sqrt{D}$, with U_i drawn independently from the Haar measure, one obtains with high probability a $(d^{-1/2}, d)$-quantum expander. By a similar approach we show in "Appendix F" that, with high probability, the corresponding quantum expander state is such that there *are* strong correlations between regions separated by less than $c\log(D)/\log(d)$ sites, for a constant c.

Therefore we have the following situation: There are correlations between regions separated by less than $c\log(D)/\log(d)$ sites, for a constant c, while there is exponential decay of correlations for regions separated by more than $C\log(D)/\log(d)$ sites, for another constant $C > c$. This is consistent with Theorem 1 and in fact shows that the linear dependence of l_0 in the entropy bound is optimal (since the entropy of quantum expander states saturates at approximately $\log(D)$).

We note that one can also derive a bound on correlations in a MPS independent of the bond dimension, but only for regions which are sufficiently far away from the boundary of the chain. See [53] for a derivation of the bound and a discussion of its relation with the results of this paper.

3. Proof of Theorem 1

Before we turn to the proof in earnest, we first provide a sketch of it and an outline of the main techniques used, in order to give a more global view of its structure. In a nutshell, we explore the idea that if an area law is violated, then a *random measurement* in one region, together with another carefully chosen measurement in another region far apart, give rise to stronger correlations than what would be allowed by the finite correlation length of the state.

Proof Sketch: The key idea of the proof comes from the analysis provided before for random states. There we could find a partition of the system into three regions ABC such that A was approximately decoupled from B, and the state on A was close to the maximally mixed state. Then by Uhlmann's theorem we could show that A was strongly correlated with C. Therefore the fact that the entropy on AB was close to maximum implied a lower bound on the correlations between A and C. In the general case AB might not have entropy close to maximum, and so A might not be decoupled from B. However one can still try to follow the same reasoning by applying a measurement on A that *decouples* it from B.

The above problem—of decoupling A from B while making the state on A maximally mixed—has been studied before and is known as entanglement distillation protocol [54–56]. Given the state $|\psi\rangle_{ABC}^{\otimes n} \in (\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C)^{\otimes n}$, with n sufficiently large, consider a random measurement on $(\mathcal{H}_A)^{\otimes n}$, consisting of $N \approx 2^{nI(A:B)}$ Haar distributed projectors $\{P_k\}_{k=1}^N$ of equal dimension summing up to the identity. Here $I(A : B) := H(A) + H(B) - H(AB)$ is the mutual information of A and B. Then it was shown in Refs. [54–56] that with high probability the post-selected state $|\phi\rangle_{A'B^nC^n} := (P_k \otimes \mathrm{id}_{B^nC^n})|\psi\rangle_{ABC}^{\otimes n}/\|(P_k \otimes \mathrm{id}_{B^nC^n})|\psi\rangle_{ABC}^{\otimes n}\|$ is such that, if $H(B) \leq H(C)$,

$$D\left(\rho_{A'B^n}, \tau_{A'} \otimes \rho_{B^n}\right) \approx 0 \quad \text{and} \quad |A'| \approx 2^{-nH(A|C)}, \tag{13}$$

with $\tau_{A'}$ the maximally mixed state on A' and $\rho_{A'B^n}$ the $A'B^n$ reduced density matrix of $|\phi\rangle_{A'B^nC^n}$. We say

$$-H(A|C) := H(C) - H(AC) = H(C) - H(B) \tag{14}$$

is the entanglement distillation rate of the protocol, as it gives the number of EPR pairs shared by A and C after the random measurement on A. Here $H(A|C)$ is the conditional entropy of A given the side information C. Using that the state $|\psi\rangle_{ABC}$ is pure the entanglement rate could alternatively be written as $H(A|B) = H(AB) - H(B)$.[5] This bound on the entanglement distillation rate is known as hashing bound.

Thus considering many copies of the state and making an appropriate measurement on A we end up again in the situation where A' is close to maximally mixed and decoupled from B, implying that if $H(A|C) < 0$, A' is maximally entangled with (part of) C. The argument thus suggests that in order not to have long-range correlations between A and C one must have $H(C) \leq H(B)$, which gives an area law for region C if B has constant size.

There are two challenges for making this idea work. The first concerns the fact that the entanglement distillation protocol of [54–56] is devised only in the limit of infinitely many copies of the state, but in our problem we have only a single copy of it. The second is the fact that we only get a particular outcome k with probability $\approx 2^{-nI(A:B)}$, which could be catastrophic if we want to show that the correlations between A and C are large. Let us ignore the first challenge in this sketch and focus on the second. Thus we are going to assume that the entanglement distillation protocol works in the same way for a single copy of the state as it does asymptotically; later we will see how we can make this precise by using tools from single-shot quantum information theory [38,64].

Considering the simplifying assumption that the distillation protocol works for a single copy of the state, the upshot of the result of [54–56] is that $H(C) \geq H(B)$ implies $\mathrm{Cor}(A : C) \geq 2^{-I(A:B)}$. Indeed, one could first make a random measurement on A, obtaining one of the possible outcomes with probability $2^{-I(A:B)}$ and distilling a maximally entangled state between A and C, and then measure the correlations in the maximally entangled state. We can also write the previous relation as

$$\mathrm{Cor}(A : C) \leq 2^{-I(A:B)} \quad \text{implies} \quad H(C) \leq H(B). \tag{15}$$

From exponential decay of correlations we have that $\mathrm{Cor}(A : C) \leq 2^{-l/\xi}$, with $l = \mathrm{size}(B)$. Thus if $I(A : B) \leq l/\xi$, we must have $H(C) \leq H(B)$, which constitutes

---

[5] A variant of this latter expression will be the one used later in the proof once we consider the single-shot entanglement distillation protocol.
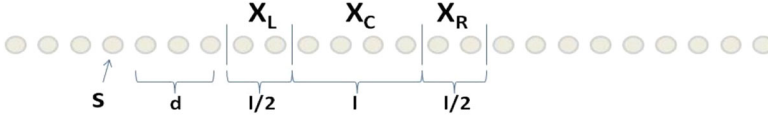
Fig. 4. Structure of the regions for Lemma 16 and Lemma 7. In Lemma 16, $d, l \leq l_0 \exp(O(1/\varepsilon))$, while in Lemma 7, $d, l \leq l_0 \exp(O(\log(1/\varepsilon)/\varepsilon))$

an area law for region C if B has fixed size. Unfortunately we do not have any guarantee that $I(A : B) \leq l/\xi$.

The second idea of the proof is to show that at a distance at most $l_0 \exp(O(\xi))$ sites from C, we can find a region B of size smaller than $l_0 \exp(O(\xi))$ such that $H(B) \leq l/(2\xi)$, which implies $I(A : B) \leq 2H(B) \leq l/\xi$. Then by the argument above we can get a bound of $l_0 \exp(O(\xi))$ on the entropy of a region which differs from C by less than $l_0 \exp(O(\xi))$ sites, and so by subadditivity of entropy and the Araki and Lieb inequality [57] we have a bound of $l_0 \exp(O(\xi))$ on the entropy of C as well.

In order to prove the existence of a region B with $H(B) \leq l/(2\xi)$ we apply a variant of a result due to Hastings [21], used in his proof of an area law for groundstates of 1D gapped models, concerning the saturation of mutual information in a multiparticle state. It appears as Lemma 16 in "Appendix B" and states: For all $\varepsilon > 0$ and a particular site s there exist neighbouring regions $X_L X_C X_R$ at a distance at most $l_0 \exp(O(1/\varepsilon))$ sites from s, with X_L and X_R each of size l and X_C of size $2l$, such that $I(X_C : X_L X_R) \leq \varepsilon l$ and $l \leq l_0 \exp(O(1/\varepsilon))$ (see Fig. 4). Let us consider the partition $X_L X_C X_R R$, with R the region composed of all the remaining sites not in $X_L X_C X_R$. Then we choose $\varepsilon = 1/(2\xi)$ and use once more the fact that the state has exponential decay of correlations to find that $\mathrm{Cor}(X_C : R) \leq 2^{-l/\xi} \leq 2^{I(X_C : X_L X_R)}$. Then using Eq. (15) we get $H(R) \leq H(X_L X_R)$. But since $H(X_C) + H(X_L X_R) - H(R) = I(X_C : X_L X_R) \leq l/(2\xi)$, we find $H(X_C) \leq 1/(2\xi)$, which gives the desired relation setting $B = X_C$.

To summarize, we employ the entanglement distillation protocol in the form of Eq. (15) and the assumption of exponential decay of correlations *twice*. One in conjunction with the result about saturation of mutual information in order to get a region of constant size and not so large entropy, and the second to boost this into an area law for regions of arbitrary size. This finishes the sketch of the proof.   □

In order to turn the sketch above into a proof, we will need two lemmas, which we believe might be of independent interest. Both provide single-shot analogues of the entropic results used in the sketch. To this goal we use the recent framework of single-shot quantum information theory [38,64], whose objective is to analyse the protocols of quantum information theory *without* the assumption of having an arbitrarily large number of copies of the state (or uses of the channel in dynamical problems).

The first lemma gives three relations between correlations and smooth entropies in a general tripartite quantum pure state. To state it we will need the following single-shot analogue of the conditional entropy, which has an important role in the proof: Given a bipartite state $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$, the min-entropy of A conditioned on B is defined as

$$H_{\min}(A|B)_\rho := \max_{\sigma \in \mathcal{D}(\mathcal{H}_B)} \sup \left\{ \lambda \in \mathbb{R} : 2^{-\lambda} \mathrm{id}_A \otimes \sigma_B \geq \rho_{AB} \right\}. \qquad (16)$$

The ε-smooth min-entropy of A conditioned on B of ρ_{AB} is given by

$$H_{\min}^{\varepsilon}(A|B)_\rho := \max_{\tilde{\rho}_{AB} \in \mathcal{B}_\varepsilon(\rho_{AB})} H_{\min}(A|B)_{\tilde{\rho}}. \qquad (17)$$

An operational interpretation for the conditional min-entropy, which we explore in the proof of the next lemma, was given in Ref. [58] (see also Lemma 28 in "Appendix G" for a formal statement). Let $|\psi\rangle_{ABC}$ be an arbitrary purification of ρ_{AC}.

$H_{\min}^{\varepsilon}(A|B)$ is the entanglement distillation rate of the single-shot protocol, where one has a *single* copy of the state and allows an ε-error in the output state. Moreover, the classical communication cost of the protocol is given by the max-mutual information of A with the purification B of AC [59] defined as[6]

$$I_{\max}^{\varepsilon}(A:B) := H_{\max}^{\varepsilon}(A) - H_{\min}^{\varepsilon}(A|B). \tag{18}$$

Lemma 6 (Correlations versus entropies). *Given a tripartite state $|\psi\rangle_{ABC}$ and real numbers $\delta, \nu > 0$, the following holds*

1. $H_{\min}^{\delta}(A|B) \geq -2\log(\delta) + 5$ implies that

$$\mathrm{Cor}\,(A:C) \geq \left(\frac{1}{256} - 26\sqrt{\delta}\right) 2^{-\left(I_{\max}^{\delta}(A:B) - 2\log(\delta) + 5\right)}. \tag{19}$$

2. $H_{\max}^{\nu}(C) \geq 3H_{\max}^{\delta}(B)$ and $\frac{\nu}{2} - 2\delta \geq \sqrt{\frac{4\delta}{1-2\nu-2\delta}}$ imply that

$$\mathrm{Cor}\,(A:C) \geq \left(\frac{\nu}{2} - 2\delta - \sqrt{\frac{4\delta}{1-2\nu-2\delta}}\right)^2 2^{-3\left(H_{\max}^{\delta}(B) + \log(\delta^{-2} + \nu^{-2}) + 5\right)}. \tag{20}$$

3. Setting $\gamma := \mathrm{Cor}(A:C)^{1/2}(1/2 - \delta)^{-1/2}$,

$$H_{\max}^{2\gamma}(A) \leq H_{\max}^{\delta}(A) + 2\log|B| + \log\left(\frac{2}{\gamma^2}\right). \tag{21}$$

Part 1 of Lemma 6 is a direct consequence of the entanglement distillation protocol of Ref. [58] (see "Appendix G"). Part 3 also follows directly from the definition of correlation function together with basic properties of the max-entropy and the purified distance.

Part 2 of the lemma, in turn, is a novel relation: Since one might have $H_{\max}^{\nu}(C) \geq 3H_{\max}^{\delta}(B)$ but $H_{\min}^{\delta}(A|B) \leq 0$, it shows that even in situations where one cannot establish EPR pairs between two parties thus decoupling the third purifying party (which can only happen when $H_{\min}^{\delta}(A|B) > 0$ [58]), one can still have large correlations between the two parties. Moreover, as shown in the proof such correlations are manifested by choosing a random measurement on A and optimizing over a measurement on C. So we are operating in a regime where a random measurement on one party does not decouple it from either of the other two parties, but it does generate correlations with one of them.

The next lemma is an analogue for the max-mutual information of Lemma 16 in "Appendix B", which itself is a slight modification of a result originally shown by Hastings in his proof of an area law for groundstates of 1D gapped models [21]

Lemma 7 (Saturation of max-mutual information). *Let $\rho_{1,\ldots,n} \in \mathcal{D}\left((\mathbb{C}^2)^{\otimes n}\right)$ be a state defined on a line with (ξ, l_0)-exponential decay of correlations, and let s be a particular site. Then for all $\delta > 0$, $\min\{\delta/2, 1/\xi\} \geq \varepsilon > 0$, and $\bar{l}_0 \geq l_0$, there is an l satisfying $O(\xi^2 \log(2/\varepsilon)\,\varepsilon^{-1}) \leq l/\bar{l}_0 \leq \exp(O(\log(1/\varepsilon)/\varepsilon))$ and a connected region $X_{2l} :=$*

---

[6] We note that there are other possible definitions of single-shot mutual information such as the one given in Ref. [60].

$X_{L,l/2} X_{C,l} X_{R,l/2}$ of $2l$ sites (*the borders $X_{L,l/2}$ and $X_{R,l/2}$ with $l/2$ sites each, and the central region $X_{C,l}$ with l sites*) *centred at most $\bar{l}_0 \exp(O(\log(1/\varepsilon)/\varepsilon))$ sites away from s* (*see Fig.* 4) *such that*

$$I_{\max}^{\delta}(X_{C,l} : X_{L,l/2} X_{R,l/2}) \leq \frac{8\varepsilon l}{\delta} + \delta. \tag{22}$$

Although the proof of Lemma 16 in "Appendix B" is straightforward, only involving a chain of applications of subadditivity of entropy, the proof of the single-shot counterpart given by Lemma 7 is considerably more involved. The main difficulty is that one might have big gaps between min- and max-entropies, and thus it is not clear how one can concatenate several uses of the subadditivity inequality.

This difficulty is handled by exploring two ideas: The first is to apply the quantum substate theorem of Jain, Radhakrishnan, and Sen [61] (see also [62] and Lemma 15 in "Appendix B") in order to turn the max-mutual information into a hybrid mutual information, involving both von Neumann and smooth max entropies. The second idea, necessary to handle the remaining gaps that might exist between von Neumann and max entropies, is to use the hypothesis of exponential decay of correlations, together with the quantum equipartition property [63,64], to upper bound the max-entropy by a sum of von Neumann entropies.

We are now ready to prove Theorem 1. We will make use of the three parts of Lemma 6, each in conjunction with exponential decay of correlations, as well as Lemma 7, which itself is based on exponential decay of correlations, which in total amounts to *four* applications of the assumption that the state has exponential decay of correlations.

Theorem 1 (restatement). *Let $|\psi\rangle_{1,\ldots,n}$ be a state defined on a ring with (ξ, l_0)-exponential decay of correlations and $n \geq C l_0/\xi$. Then for any connected region $X \subset [n]$ and every $l \geq 8\xi$,*

$$H_{\max}^{2^{-\frac{l}{8\xi}}}(X) \leq c' l_0 \exp(c \log(\xi)\xi) + l, \tag{23}$$

with $C, c, c' > 0$ universal constants.

Proof. We start by applying Lemma 7 twice to the two boundaries of the region X, with $\delta := 10^{-8}$, $\varepsilon := \min\left(1/(10^5 \xi), \delta/8\right)$ and $\bar{l}_0 := \max\left(l_0, cl_0\varepsilon/(\xi \log(2/\varepsilon))\right)$, for a constant $c > 0$. Then we find that there are regions Y and $\tilde{Y}$ (see Fig. 5a), each of size $2l$ with $O(\xi^2 \log(2/\varepsilon)) \leq l/\bar{l}_0 \leq \exp(O(\log(1/\varepsilon)/\varepsilon))$, and at a distance at most $\bar{l}_0 \exp(O(\log(1/\varepsilon)/\varepsilon))$ sites away from the two boundaries of X, respectively, such that

$$I_{\max}^{\delta/4}(Y_{C,l} : Y_{L,l/2} Y_{R,l/2}) = H_{\max}^{\delta/4}(Y_{C,l}) - H_{\min}^{\delta/4}(Y_{C,l}|Y_{L,l/2} Y_{R,l/2}) \leq \frac{l}{8\xi}, \tag{24}$$

and likewise for $\tilde{Y}$. We note that from the choice of parameters it follows by assuming that $l \geq 300\xi$.

We now argue that exponential decay of correlations implies that the entropy of $\rho_{Y_{C,l}}$ must be small. Let R be the complementary region to Y (see Fig. 5b). Then from exponential decay of correlations we have $\text{Cor}(Y_{C,l} : R) \leq 2^{-l/(2\xi)}$. Applying part (1) of Lemma 6 we find that

$$-H_{\min}^{\delta/4}(Y_{C,l}|Y_{L,l} Y_{R,l}) < -2 \log(\delta) + 9, \tag{25}$$
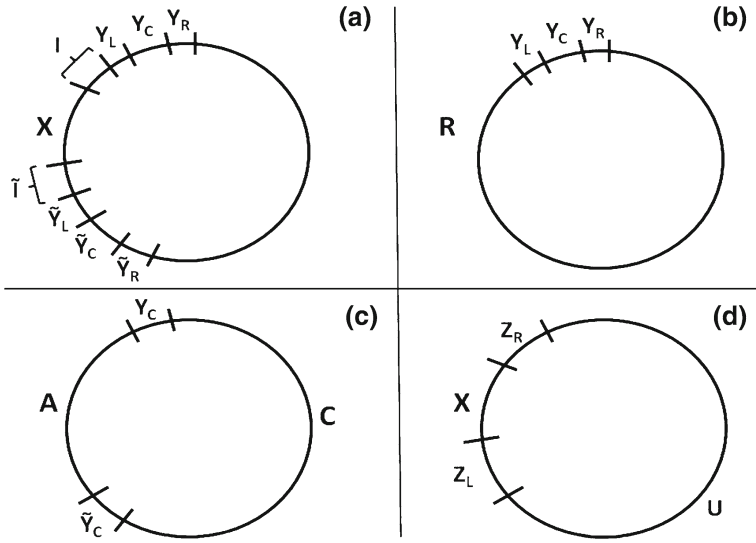
Fig. 5. a Region X separated from $Y := Y_L Y_C Y_R$ by $I \leq l_0 \exp(O(\log(1/\varepsilon)/\varepsilon))$ sites and from $\tilde{Y} := \tilde{Y}_L \tilde{Y}_C \tilde{Y}_R$ by $\tilde{I} \leq l_0 \exp(O(\log(1/\varepsilon)/\varepsilon))$ sites; **b** partition into regions $Y := Y_L Y_C Y_R$ and R; **c** partition into regions A, $B := Y_C \tilde{Y}_C$, and C; **d** partition into regions X, $Z := Z_L Z_R$, and U

since otherwise

$$\mathrm{Cor}(Y_{C,l} : R) \geq \left(\frac{1}{256} - 26\sqrt{\delta}\right) 2^{-I_{\max}^{\delta/4}(Y_{C,l}:Y_{L,l/2}Y_{R,l/2})-2\log(\delta)+9} \geq 2^{-\left(\frac{l}{8\xi}-34\right)}, \quad (26)$$

in contradiction with the correlation length being ξ (since $l \geq 300\xi$). Then from Eqs. (25) and (24) we have

$$H_{\max}^{\delta/4}(Y_{C,l}) \leq \frac{l}{8\xi} - 2\log(\delta) + 9. \quad (27)$$

Applying the same reasoning to $\tilde{Y}$ we find that also $H_{\max}^{\delta/4}(\tilde{Y}_{C,l}) \leq \frac{l}{8\xi} - 2\log(\delta) + 9$.

Define A as the region between $Y_{C,l}$ and $\tilde{Y}_{C,l}$, $B := Y_{C,l}\tilde{Y}_{C,l}$, and C the complementary region to AB (see Fig. 5c). Note that A differs from X by at most $2\bar{l}_0 \exp(O(\log(1/\varepsilon)/\varepsilon))$ sites and so by Lemma 18, for every $\nu > 0$,

$$H_{\max}^{2\nu}(X) \leq H_{\max}^{\nu}(A) + 2\bar{l}_0 \exp(O(\log(1/\varepsilon)/\varepsilon)) + \log \frac{2}{\nu^2}. \quad (28)$$

We now prove an upper bound on $H_{\max}^{\nu}(A)$, with $\nu = 0.01$. Since A and C are separated by l sites, we have from exponential decay of correlations that

$$\mathrm{Cor}(A : C) \leq 2^{-l/\xi}. \quad (29)$$

From part (2) of Lemma 6 and Eq. (29) we then find that

$$H_{\max}^{\nu}(C) < 3H_{\max}^{\delta}(B), \quad (30)$$

since otherwise, using that $l \geq 300\xi$,

$$\mathrm{Cor}\,(A:C) \geq \left(\frac{\nu}{2} - 2\delta - \sqrt{\frac{4\delta}{1-2\nu-2\delta}}\right)^2 2^{-3\left(H_{\max}^\delta(B)+16\log\left(\delta^{-2}+\nu^{-2}\right)+5\right)} \geq 2^{-3l/(4\xi)},$$

(31)

contradicting Eq. (29). In the last inequality of the equation above we used that

$$\begin{aligned}
H_{\max}^\delta(B) &= H_{\max}^\delta(Y_{C,l}\tilde{Y}_{C,l}) \\
&\leq H_{\max}^{\delta/4}(Y_{C,l}) + H_{\max}^{\delta/4}(\tilde{Y}_{C,l}) + \log\frac{8}{\delta^2} \\
&\leq \frac{l}{4\xi} - 4\log(\delta) + 18 + \log\frac{8}{\delta^2},
\end{aligned}$$

(32)

where the first inequality follows from subadditivity of the max-entropy (Lemma 18) and the second from Eq. (27).

Using subadditivity of the max-entropy (Lemma 18) again, we get

$$\begin{aligned}
H_{\max}^{3\nu+\delta}(A) = H_{\max}^{3\nu+\delta}(BC) &\leq H_{\max}^\delta(B) + H_{\max}^\nu(C) + \log\frac{2}{\nu^2} \\
&\leq 4H_{\max}^\delta(B) + \log\frac{2}{\nu^2} \\
&\leq 2\bar{l}_0 \exp\left(O(\log(1/\varepsilon)/\varepsilon)\right),
\end{aligned}$$

(33)

where we used Eq. (30) and the bound $H_{\max}(B) \leq 2l \leq \bar{l}_0 \exp(O(\log(1/\varepsilon)/\varepsilon))$. Therefore from Eq. (28),

$$H_{\max}^{6\nu+2\delta}(X) \leq 2\bar{l}_0 \exp\left(O(\log(1/\varepsilon)/\varepsilon)\right) = O(l_0) \exp\left(O(\log(1/\varepsilon)/\varepsilon)\right). \quad (34)$$

This is already an area law for X, although with a fixed error (recall that we fixed $\delta = 10^{-8}$ and $\nu = 0.01$). To finish the proof we show how applying exponential decay of correlations once more we can reduce the error. Let Z be a region of size $2l$ separating X from the remaining sites by a distance l, and denote by U the complementary region to XZ (see Fig. 5d).[7] By exponential decay of correlations we have $\mathrm{Cor}(X:U) \leq 2^{-l/\xi}$. Using part (3) of Lemma 6 we then get

$$H_{\max}^{2^{-l/(4\xi)}}(X) \leq H_{\max}^{2^{-l/(2\xi)+1}}(X) \leq H_{\max}^{6\nu+2\delta}(X) + 2l \leq O(l_0) \exp\left(O(\log(1/\varepsilon)/\varepsilon)\right) + 2l,$$

(35)

where the first inequality follows assuming $l \geq 4\xi$. $\square$

4. Correlations Versus Entropies

In this section we present the proof of Lemma 6, which we restate for convenience of the reader.

Lemma 6 (restatement). *Given a tripartite state $|\psi\rangle_{ABC}$ and real numbers $\delta, \nu > 0$, the following holds*

---

[7] Note that the distance l has no relation with the parameter l used in the first part of the proof and in particular it does not have to be bigger than 300ξ.

1. $H_{\min}^{\delta}(A|B) \geq -2\log(\delta) + 5$ *implies that*

$$\text{Cor}(A:C) \geq \left(\frac{1}{256} - 26\sqrt{\delta}\right) 2^{-(I_{\max}^{\delta}(A:B) - 2\log(\delta) + 5)}. \tag{36}$$

2. $H_{\max}^{\nu}(C) \geq 3H_{\max}^{\delta}(B)$ *and* $\frac{\nu}{2} - 2\delta \geq \sqrt{\frac{4\delta}{1 - 2\nu - 2\delta}}$ *imply that*

$$\text{Cor}(A:C) \geq \left(\frac{\nu}{2} - 2\delta - \sqrt{\frac{4\delta}{1 - 2\nu - 2\delta}}\right)^2 2^{-3\left(H_{\max}^{\delta}(B) + \log\left(\delta^{-2} + \nu^{-2}\right) + 5\right)}. \tag{37}$$

3. *Setting* $\gamma := \text{Cor}(A:C)^{1/2}(1/2 - \delta)^{-1/2}$,

$$H_{\max}^{2\gamma}(A) \leq H_{\max}^{\delta}(A) + 2\log|B| + \log\left(\frac{2}{\gamma^2}\right). \tag{38}$$

Proof.

(Part 1)

From Lemma 28 in "Appendix G" it follows there is a POVM element M and an isometry V such that $\text{tr}(M\rho_A) \geq 1/(2N)$ and

$$D\left(\frac{(M^{1/2} \otimes V)|\psi\rangle\langle\psi|_{ACB}(M^{1/2} \otimes V)^{\dagger}}{\text{tr}(M\rho_A)}, |\Phi\rangle\langle\Phi|_{A_1 C_1} \otimes |\psi\rangle\langle\psi|_{C'CB}\right) \leq 26\sqrt{\delta}, \tag{39}$$

with

$$\log(\dim(A_1)) = \lfloor H_{\min}^{\delta}(A|B) - 2\log(\delta) - 4\log(13)\rfloor, \tag{40}$$

and

$$\log(N) = \lfloor I_{\max}^{\delta}(A:B) - 2\log(\delta) - 4\log(13)\rfloor. \tag{41}$$

Define

$$P_{A_1} := \sum_{k=1}^{\dim(A_1)/2} |k\rangle\langle k|, \quad P_{C_1} := \sum_{k=1}^{\dim(C_1)/2} |k\rangle\langle k|, \tag{42}$$

and

$$\tilde{\rho}_{AC} = \frac{(M^{1/2} \otimes V)\rho_{AC}(M^{1/2} \otimes V)^{\dagger}}{\text{tr}(M\rho_A)}. \tag{43}$$

Let $\mu(dU)$ be the Haar measure on unitaries over a vector space isomorphic to A_1 and C_1. Then we have

$$\text{Cor}(A:C) \geq \int \mu(dU)\text{tr}\left([(M^{1/2}U P_{A_1} U^{\dagger}M^{1/2}) \otimes V^{\dagger}(U^* P_{C_1} U^T \otimes \text{id}_{C'C})V]\right.$$
$$\left. \times (\rho_{AC} - \rho_A \otimes \rho_C)\right)$$
$$= \text{tr}(M\rho_A)\int \mu(dU)\text{tr}\left((U P_{A_1} U^{\dagger} \otimes U^* P_{C_1} U^T)(\tilde{\rho}_{A_1 C_1} - \sigma_{A_1} \otimes \tilde{\sigma}_{C_1})\right)$$
$$=: M, \tag{44}$$

with σ_{A_1} and σ_{C_1} states on A_1 and C_1, respectively. We have

$$M = \text{tr}(M\rho_A)\text{tr}\left((P_{A_1} \otimes P_{C_1})(\tilde{\Delta}(\rho_{A_1 C_1}) - \Delta(\sigma_{A_1} \otimes \tilde{\sigma}_{C_1}))\right), \tag{45}$$

where the cptp map Δ is given by

$$\Delta(X) := \int \mu(\text{d}U)(U \otimes U^*)X(U \otimes U^*)^\dagger. \tag{46}$$

By monotonicity of trace-norm under cptp maps and Eq. (39),

$$D(\Delta(\rho_{A_1 C_1}), |\Phi\rangle\langle\Phi|_{A_1 C_1}) \leq 26\sqrt{\delta}. \tag{47}$$

Moreover since $\sigma_{A_1} \otimes \tilde{\sigma}_{C_1}$ is a separable state (even product), we have

$$\Delta(\sigma_{A_1} \otimes \tilde{\sigma}_{C_1})) = p|\Phi\rangle\langle\Phi|_{A_1 C_1} + (1-p)\tau_{A_1 C_1}, \tag{48}$$

with $p \leq 2/|A_1|$. Therefore

$$M \geq \text{tr}(M\rho_A)\left(1/256 - 26\sqrt{\delta}\right). \tag{49}$$

(Part 2)

Part 1 of lemma 14 gives projectors P_A, P_B such that $\text{tr}(P_A\rho_A) \geq 1 - 2\nu$, $|P_A| = 2^{H^\nu_{\max}(A)}$ and $\text{tr}(P_B\rho_B) \geq 1 - 2\delta$, $|P_B| = 2^{H^\delta_{\max}(B)}$. Then part 3 of Lemma 14 allows us to write ρ_{AC} as the following mixture

$$\rho_{AC} = (1 - 2\delta)\pi_{AC} + 2\delta\sigma_{AC}, \tag{50}$$

where π_{AC} is the reduced density matrix of the pure state $|\pi\rangle_{ABC}\langle\pi| := P_B|\psi\rangle\langle\psi|P_B/\text{tr}(P_B\rho_B)$, and $H_{\max}(\pi_{AC}) = H_{\max}(\pi_B) \leq H^\delta_{\max}(\rho_B)$. Let Q be a projector satisfying $Q \leq P_A$ and chosen from the Haar measure, i.e., $Q = UQ_0U^\dagger$ for a fixed arbitrary projector $Q_0 \leq P_A$ and U is a unitary drawn from the Haar measure in the space in which P_A projects on. Define $\tilde{\rho}_C := \text{tr}_A(Q \otimes I_C\rho_{AC}Q \otimes I_C)/\text{tr}(\rho_A Q)$, $\tilde{\pi}_C := \text{tr}_A(Q \otimes I_C\pi_{AC}Q \otimes I_C)/\text{tr}(\pi_A Q)$, and $\tilde{\sigma}_C := \text{tr}_A(Q \otimes I_C\sigma_{AC}Q \otimes I_C)/\text{tr}(\sigma_A Q)$. Let $|\tilde{\pi}\rangle\langle\tilde{\pi}|_{ABC} := Q|\pi\rangle\langle\pi|Q/\text{tr}(Q\pi_A)$.

Let us first prove the claim of the lemma under two assumptions and then show how we can relax them. They read

$$\delta|Q| \geq 4\sqrt{|P_A|}, \tag{51}$$

and

$$H^\delta_{\max}(\rho_B) \geq \log \frac{8}{\nu^2} \tag{52}$$

The first step in the proof is to show that the entropy of $\tilde{\rho}_C$ is bounded from above as follows:

$$H^\eta_{\max}(\tilde{\rho}_C) \leq \log|Q| + H^\delta_{\max}(\rho_B) \tag{53}$$

where $\eta := \sqrt{\frac{4\delta}{1 - 2\nu - 2\delta}}$. To prove it first note that by subadditivity (Lemma 18),

$$H_{\max}(\tilde{\pi}_C) \leq H_{\max}(\tilde{\pi}_A) + H_{\max}(\tilde{\pi}_B). \tag{54}$$

However $H_{\max}(\tilde{\pi}_A) \leq \log |Q|$ and $H_{\max}(\tilde{\pi}_B) \leq H_{\max}(\pi_B) \leq H_{\max}^\delta(\rho_B)$ (which follows from the fact that the application of a projection to one system of a bipartite state cannot increase the rank of the other system). Hence we obtain

$$H_{\max}(\tilde{\pi}_C) \leq \log |Q| + H_{\max}^\delta(\rho_B) \tag{55}$$

Applying Lemma 8 it follows that with non-zero probability

$$D(\tilde{\pi}_C, \tilde{\rho}_C) \leq \sqrt{\frac{4\delta}{1 - 2\nu - 2\delta}} = \eta. \tag{56}$$

We thus obtain

$$H_{\max}^\eta(\tilde{\rho}_C) \leq H_{\max}(\tilde{\pi}_C), \tag{57}$$

which together with Eq. (55) gives the required upper bound on the entropy of $\tilde{\rho}_C$ given by Eq. (53).

Define

$$\mu := 2 \operatorname{Cor}(A : C) \frac{|P_A|}{|Q|}. \tag{58}$$

Then from Lemma 9, $D(\tilde{\rho}_C, \rho_C) \leq \sqrt{\mu}$, and hence

$$H_{\max}^{\sqrt{\mu}+\eta}(\rho_C) \leq H_{\max}^\eta(\tilde{\rho}_C). \tag{59}$$

We now set

$$|Q| = \left\lfloor \frac{|P_A|}{2^{3H_{\max}^\delta(B)}} \right\rfloor. \tag{60}$$

Recalling that $|P_A| = 2^{H_{\max}^\nu(\rho_A)}$, and using Eqs. (59) and (53) we obtain

$$H_{\max}^{\sqrt{\mu}+\eta}(\rho_C) \leq H_{\max}^\delta(\rho_B) + H_{\max}^\nu(\rho_A) - 3H_{\max}^\delta(\rho_B). \tag{61}$$

By subadditivity of max-entropy (Lemma 18) for system BC with $\epsilon = \frac{\nu}{2}$, $\epsilon' = \frac{\nu}{2} - 2\delta$, and $\epsilon'' = \delta$,

$$H_{\max}^\nu(\rho_A) = H_{\max}^\nu(\rho_{BC}) \leq H_{\max}^\delta(\rho_B) + H_{\max}^{\frac{\nu}{2}-2\delta}(\rho_C) + \log \frac{8}{\nu^2}, \tag{62}$$

and so

$$H_{\max}^{\sqrt{\mu}+\eta}(\rho_C) \leq H_{\max}^{\frac{\nu}{2}-2\delta}(\rho_C) - H_{\max}^\delta(\rho_B) + \log \frac{8}{\nu^2}. \tag{63}$$

From Eq. (52) we obtain $H_{\max}^{\sqrt{\mu}+\eta}(\rho_C) \leq H_{\max}^{\frac{\nu}{2}-2\delta}(\rho_C)$ so that

$$\sqrt{\mu} + \eta \geq \frac{\nu}{2} - 2\delta. \tag{64}$$

Eqs. (60) and (58) give $\operatorname{Cor}(A : C) \geq \mu 2^{-3H_{\max}^\delta(\rho_B)}$, which together with Eq. (64) implies

$$\operatorname{Cor}(A : C) \geq \left(\frac{\nu}{2} - 2\delta - \sqrt{\frac{4\delta}{1 - 2\nu - 2\delta}} \right)^2 2^{-3H_{\max}^\delta(B)}. \tag{65}$$

We just derived Eq. (65) under the two assumptions given by Eqs. (51) and (52). Let us finally show how we can extend it to a slightly weaker relation without the need of any assumption. Consider the state $|\phi\rangle_{AA':BB':C} := |\psi\rangle_{ABC} \otimes |\Phi\rangle_{A'B'}$, where $|\Phi\rangle_{A'B'}$ is a maximally entangled state of dimension $|A'| = |B'| = \lfloor 16(\delta^{-2} + \nu^{-2}) \rfloor$. Then the assumptions of Eqs. (51) and (52) are satisfied. Using Eq. (65) applied to $|\phi\rangle_{AA':BB':C}$ we get

$$\text{Cor}(A : C) = \text{Cor}(AA' : C) \geq \left(\frac{\nu}{2} - 2\delta - \sqrt{\frac{4\delta}{1 - 2\nu - 2\delta}} \right)^2 2^{-3H_{\max}^\delta(BB')}, \quad (66)$$

where the first equality follows from the fact that A' is decoupled from AC. Finally we note that $H_{\max}^\delta(BB') \leq H_{\max}^\delta(B) + \log |B'| \leq H_{\max}^\delta(B) + \log(\delta^{-2} + \nu^{-2}) + 5$ and we are done.

(Part 3)

By Lemma 14 there is a projector P such that $\text{tr}(P\rho_A) \geq 1 - 2\delta$ and $|P| = 2^{H_{\max}^\delta(A)}$, with ρ_A the reduced state of $|\psi\rangle_{ABC}$. Then applying Lemma 9 with M equals P we get

$$\text{Cor}(A : C) \geq \left(\frac{1}{2} - \delta \right) D(\tilde{\rho}_C, \rho_C)^2, \quad (67)$$

where $\tilde{\rho}_C = \text{tr}_{AB} \left(|\tilde{\psi}\rangle\langle\tilde{\psi}|_{ABC} \right)$, with $|\tilde{\psi}\rangle_{ABC} = (P_A \otimes \text{id}_{BC}) |\psi\rangle_{ABC} / \sqrt{\text{tr}(P\rho_A)}$ the postselected total state. Thus setting $\gamma := \text{Cor}(A : C)^{1/2} (1/2 - \delta)^{-1/2}$,

$$H_{\max}^\gamma(\rho_C) \leq H_{\max}(\tilde{\rho}_C) \quad (68)$$

By subadditivity of max-entropy (Lemma 18) we have

$$H_{\max}^{2\gamma}(\rho_A) = H_{\max}^{2\gamma}(\rho_{BC}) \leq H_{\max}^\gamma(\rho_C) + \log |B| + \log \left(\frac{2}{\gamma^2} \right), \quad (69)$$

where the entropies are computed with respect to the state $|\psi\rangle_{ABC}$. Again by subadditivity of max-entropy, this time applied to the postselected state $|\tilde{\psi}\rangle_{ABC}$,

$$H_{\max}(\tilde{\rho}_C) \leq H_{\max}(\tilde{\rho}_A) + \log |B| \leq H_{\max}^\delta(\rho_A) + \log |B|, \quad (70)$$

where we used that $\text{rank}(\tilde{\rho}_A) \leq 2^{H_{\max}^\delta(A)}$, due to the dimension of the projector P. The last two inequalities together with Eq. (68) imply Eq. (38). □

The next lemma is used in the proof of Theorem 6.

Lemma 8. *Let $\rho_{AC} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ be such that*

$$\rho_{AC} = (1 - \delta)\pi_{AC} + \delta\sigma_{AC}, \quad (71)$$

with $0 \leq \delta \leq 1$ and $\pi, \sigma \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$. Let P be a projector on system A such that $\text{tr}(P\rho_A) \geq 1 - 2\nu$, with $0 \leq \nu \leq 1$. Let Q be a random projector given by $U Q_0 U^\dagger$, where U is a Haar distributed unitary on the support of P and $Q_0 \leq P$ is an arbitrary projector. Define $\tilde{\rho}_C := tr_A(Q \otimes I_C \rho_{AC} Q \otimes I_C)/tr(\rho_A Q)$, $\tilde{\pi}_C :=$

$tr_A(Q \otimes I_C \pi_{AC} Q \otimes I_C)/tr(\pi_A Q)$, and $\tilde{\sigma}_C := tr_A(Q \otimes I_C \sigma_{AC} Q \otimes I_C)/tr(\sigma_A Q)$. *Then with non-zero probability over the choice of U,*

$$D(\tilde{\rho}_C, \tilde{\pi}_C) \leq \sqrt{\alpha}, \tag{72}$$

with

$$\alpha := \frac{\delta \frac{|Q_0|}{|P|} + \frac{8}{\sqrt{|P|}}}{(1 - 2\nu)\frac{|Q_0|}{|P|} - \frac{8}{\sqrt{|P|}}} \tag{73}$$

Proof. From Eq. (71) and the definitions of $\tilde{\rho}_C$, $\tilde{\pi}_C$, and $\tilde{\sigma}_C$:

$$\tilde{\rho}_C = (1 - \gamma)\tilde{\pi}_C + \gamma\tilde{\sigma}_C, \tag{74}$$

with

$$\gamma := \delta \frac{tr(\sigma_A Q)}{tr(\rho_A Q)}. \tag{75}$$

We then have[8]

$$D(\tilde{\pi}_C, \tilde{\rho}_C) \leq \sqrt{\gamma}. \tag{76}$$

We now show that, with non-zero probability, $\gamma \leq \alpha$ with α given by Eq. (73). The function $f(U) := tr(\rho_A U Q_0 U^\dagger)$ is 1-Lipschitz continuous on the set of unitaries acting on a space of dimension $|P|$. Hence from Levy's lemma [65] we have

$$\Pr(|tr(\sigma_A Q) - \mathbb{E}[tr(\sigma_A Q)]| \geq \epsilon) \leq 1 - 4e^{-\frac{(2|P|+1)\epsilon^2}{16}}, \tag{77}$$

and similarly

$$\Pr(|tr(\rho_A Q) - \mathbb{E}[tr(\rho_A Q)]| \geq \epsilon) \leq 1 - 4e^{-\frac{(2|P|+1)\epsilon^2}{16}}. \tag{78}$$

We now take $\epsilon = 8d^{-\frac{1}{2}}$ so that both probabilities above are smaller than $1/2$. Then, since

$$\mathbb{E}[tr(\sigma_A Q)] = tr(\sigma_A P)\frac{|Q|}{|P|}, \quad \mathbb{E}[tr(\rho_A Q)] = tr(\rho_A P)\frac{|Q|}{|P|}, \tag{79}$$

by the union bound we have that with non-zero probability:

$$tr(\sigma_A Q) \leq tr(\sigma_A P)\frac{|Q|}{|P|} + 8|P|^{-\frac{1}{2}} \leq \frac{|Q|}{|P|} + 8|P|^{-\frac{1}{2}}$$
$$tr(\rho_A Q) \geq tr(\rho_A P)\frac{|Q|}{|P|} + 8|P|^{-\frac{1}{2}} \geq (1 - 2\nu)\frac{|Q|}{|P|} + 8|P|^{-\frac{1}{2}}, \tag{80}$$

and we are done. □

---

[8] Indeed, for normalized pure states ρ, σ, $D(\rho, \sigma) = \sqrt{1 - F^2(\rho, \sigma)}$. We also have $F(\rho, \sigma) = \max |\langle\psi|\phi\rangle|$, with the maximum taken over all $|\psi\rangle, |\phi\rangle$ which are purifications of ρ and σ. Let $|\psi\rangle, |\phi\rangle$ be purifications of $\tilde{\pi}_C$ and $\tilde{\sigma}_C$, respectively. Then $|\eta\rangle := \sqrt{1 - \gamma}|\psi\rangle \otimes |0\rangle + \sqrt{\gamma}|\phi\rangle \otimes |1\rangle$ is a purification of $\tilde{\rho}_C$. Therefore $D(\tilde{\pi}_C, \tilde{\rho}_C) \leq \sqrt{1 - |\langle\eta|\psi, 0\rangle|^2}$, giving Eq. (76).

In the proof of parts 1 and 3 of Lemma 6 we make use of the following simple lemma.

Lemma 9. *Given $\rho_{AC} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ and an operator M on system A satisfying $0 \leq M \leq \mathrm{id}$,*

$$Cor(A : C) \geq \frac{p}{2} D(\tilde{\rho}_C, \rho_C)^2, \tag{81}$$

where $p := tr(M\rho_A)$ and $\tilde{\rho}_C := tr_A\left[\left(\sqrt{M} \otimes \mathrm{id}_C\right)\rho_{AC}\left(\sqrt{M} \otimes \mathrm{id}_C\right)\right]/p$ (i.e., $\tilde{\rho}_C$ is postselected state of system C after the measurement on system A).

Proof. We have

$$tr((M \otimes N)(\rho_{AC} - \rho_A \otimes \rho_C)) = p\,tr[N(\tilde{\rho}_C - \rho_C)]. \tag{82}$$

Maximizing over $0 \leq N \leq 1$ we get

$$Cor(A : C) \geq p D_1(\tilde{\rho}_C, \rho_C) \geq \frac{p}{2} D(\tilde{\rho}_C, \rho_C)^2, \tag{83}$$

where we used Eq. (A5) and Lemma 12 in "Appendix A". □

5. Saturation of Max-Mutual Information

We now turn to the proof of Lemma 7:

Lemma 7 (restatement). *Let $\rho_{1,\ldots,n} \in \mathcal{D}\left((\mathbb{C}^2)^{\otimes n}\right)$ be a state defined on a line with (ξ, l_0)-exponential decay of correlations, and let s be a particular site. Then for all $\delta > 0$, $\min\{\delta/2, 1/\xi\} \geq \varepsilon > 0$, and $\bar{l}_0 \geq l_0$, there is an l satisfying $O\left(\xi^2 \log(2/\varepsilon)\varepsilon^{-1}\right) \leq l/\bar{l}_0 \leq \exp\left(O\left(\log(1/\varepsilon)/\varepsilon\right)\right)$ and a connected region $X_{2l} := X_{L,l/2}X_{C,l}X_{R,l/2}$ of $2l$ sites(the borders $X_{L,l/2}$ and $X_{R,l/2}$ with $l/2$ sites each, and the central region $X_{C,l}$ with l sites) centred at most $\bar{l}_0 \exp(O(\log(1/\varepsilon)/\varepsilon))$ sites away from s (see Fig. 4) such that*

$$I_{\max}^\delta(X_{C,l} : X_{L,l/2}X_{R,l/2}) \leq \frac{8\varepsilon l}{\delta} + \delta. \tag{84}$$
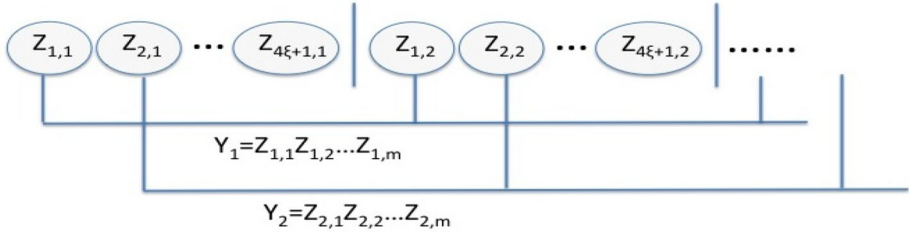
Proof. The first step of the proof is to relate the max-mutual information appearing on the L.H.S. of Eq. (84) to an hybrid mutual information, in which $H_{\min}^\delta(X_{C,l}|X_{L,l/2}X_{R,l/2})$ is replaced by its von Neumann counterpart. We have

$$\begin{aligned}
I_{\max}^\delta(X_{C,l} : X_{L,l/2}X_{R,l/2}) &= H_{\max}^\delta(X_{C,l}) - H_{\min}^\delta(X_{C,l}|X_{L,l/2}X_{R,l/2}) \\
&= H_{\max}^\delta(X_{C,l}) + \min_\sigma S_{\max}^\delta(\rho_{X_{2l}}||\mathrm{id} \otimes \sigma) \\
&\leq H_{\max}^\varepsilon(X_{C,l}) + \min_\sigma S_{\max}^\delta(\rho_{X_{2l}}||\mathrm{id} \otimes \sigma),
\end{aligned} \tag{85}$$

where the last inequality follows from the assumption that $\varepsilon \leq \delta/2$. Here $S_{\max}^\delta$ is the smooth max-relative entropy defined in Eqs. (B3) and (B4) of "Appendix B".

Using Lemma 13 of "Apendix B", we can find a $\tilde{\rho}_{X_{2l}} \in \mathcal{B}_\varepsilon(\rho_{X_{2l}})$ such that $\mathrm{rank}(\tilde{\rho}_{X_{C,l}}) = 2^{H_{\max}^\varepsilon(X_{C,l})}$. Let P be a projector onto its support. Then using $\varepsilon \leq \delta/2$,

$$\begin{aligned}
I_{\max}^\delta(X_{C,l} : X_{L,l/2}X_{R,l/2}) &\leq H_{\max}^\varepsilon(X_{C,l}) + \min_\sigma S_{\max}^{\delta/2}(\tilde{\rho}_{X_{2l}}||P \otimes \sigma) \\
&= \min_\sigma S_{\max}^{\delta/2}(\tilde{\rho}_{X_{2l}}||(2^{-H_{\max}^\varepsilon(X_{C,l})}P) \otimes \sigma)
\end{aligned} \tag{86}$$

Fig. 6. Definition of regions Y's and Z's

Using the quantum substate theorem (Lemma 15), a version of Fannes inequality (Lemma 17), and the identity $\min_\sigma S(\rho_{AB}||\text{id} \otimes \sigma) = -H(A|B)$.[9] we find

$$I^\delta_{\max}(X_{C,l} : X_{L,l/2}X_{R,l/2}) \leq \frac{2}{\delta}\left(\min_\sigma S(\tilde{\rho}_{X_{2l}}||(2^{-H^\varepsilon_{\max}(X_{C,l})}P) \otimes \sigma)\right) + \delta$$

$$= \frac{2}{\delta}\left(H^\varepsilon_{\max}(X_{C,l}) + H(\tilde{\rho}_{X_{L,l/2}X_{R,l/2}}) - H(\tilde{\rho}_{X_{2l}})\right) + \delta$$

$$\leq \frac{2}{\delta}\left(H^\varepsilon_{\max}(X_{C,l}) + H(X_{L,l/2}X_{R,l/2})\right.$$
$$\left. - H(X_{2l}) + 2\varepsilon l + 2h(\epsilon)\right) + \delta. \tag{87}$$

Thus in order to prove the theorem it suffices to show the existence of an l satisfying $2\bar{l}_0(4\xi + 1)^2 \log\left(\frac{2}{\varepsilon}\right)\varepsilon^{-1} \leq l \leq \bar{l}_0 \exp\left(O(\log(1/\varepsilon)/\varepsilon)\right)$ such that

$$H^\varepsilon_{\max}(X_{C,l}) + H(X_{L,l/2}X_{R,l/2}) - H(X_{2l}) \leq \varepsilon l. \tag{88}$$

Following an idea of Ref. [21], we prove Eq. (88) by contradiction. Suppose that for all $2\bar{l}_0(4\xi + 1)^2 \log\left(\frac{2}{\varepsilon}\right)\varepsilon^{-1} \leq l \leq \bar{l}_0 \exp(O(\log(1/\varepsilon)/\varepsilon))$,

$$H(X_{2l}) \leq H^\varepsilon_{\max}(X_{C,l}) + H(X_{L,l/2}X_{R,l/2}) - \varepsilon l. \tag{89}$$

Then we will show that this leads to the entropy $H(X_{2l})$ being negative for $l = \bar{l}_0 \exp(O(\log(1/\varepsilon)/\varepsilon))$.

Lemma 16 gives an argument for the von Neumann mutual information. Thus the challenge is how to adapt the proof of Lemma 16 to Eq. (B7), in which one of the von Neumann entropies is replaced by the max-entropy. In the sequel we show how to do it exploiting the fact that the state $\rho_{1,...,n}$ has exponential decay of correlations.[10] The idea we explore is to relate $H^\varepsilon_{\max}(X_l)$ to the von Neumann entropy. For general states this is not possible, but by assuming exponential decay of correlations we will achieve it.

Let us break the region $X_{C,l}$ into a partition of $(4\xi + 1)$ equally sized subregions, labelled by $\{Y_k\}_{k=1}^{(4\xi+1)}$, each containing $l/(4\xi + 1)$ sites (with ξ the correlation length of the state). Each region Y_k is given by $m := c/(4\xi + 1)$ blocks of l/c sites (for a constant c to be determined later), themselves labelled by $Z_{k,j}$. The $Z_{k,j}$ are separated by a distance of $(4\xi l)/c$ sites to each other, i.e., $Z_{k,j}$ is $(4\xi l)/c$ sites away from $Z_{k,j'}$ for every $j' \neq j$. The structure of the regions is depicted in Fig. 6.

---

[9] It is clear that it suffices that id is the identity on the support of ρ_A.

[10] We leave as an open question whether the result holds true for general states without having to assume exponential decay of correlations.

By iteratively applying subadditivity of $H_{\max}^\varepsilon$ (Lemma 18),

$$H_{\max}^\varepsilon(X_{C,l}) \le \sum_{k=1}^{(4\xi+1)} H_{\max}^{4^{-(4\xi+1)}\varepsilon}(Y_k) + (4\xi+1)^2 \log \frac{32}{\varepsilon}. \tag{90}$$

We claim that by exponential decay of correlations, for every k, the state ρ_{Y_k} is close in trace norm to a m-fold tensor product of states, each with l/c sites. Indeed using the identity

$$\rho_{Z_{k,1},\ldots,Z_{k,m}} - \rho_{Z_{k,1}} \otimes \cdots \otimes \rho_{Z_{k,m}} = \rho_{Z_{k,1},\ldots,Z_{k,m}} - \rho_{Z_{k,1}} \otimes \rho_{Z_{k,2},\ldots,Z_{k,m}}$$
$$+\rho_{Z_{k,1}} \otimes \rho_{Z_{k,2},\ldots,Z_{k,m}} - \rho_{Z_{k,1}} \otimes \rho_{Z_{k,2}} \otimes \rho_{Z_{k,3},\ldots,Z_{k,m}}$$
$$+\rho_{Z_{k,1}} \otimes \rho_{Z_{k,2}} \otimes \rho_{Z_{k,3},\ldots,Z_{k,m}} - \cdots\cdots$$
$$+\rho_{Z_{k,1}} \otimes \cdots \otimes \rho_{Z_{k,m-1},Z_{k,m}} - \rho_{Z_{k,1}} \otimes \cdots \otimes \rho_{Z_{k,m}}, \tag{91}$$

the triangle inequality, and Lemma 20:

$$\|\rho_{Z_{k,1},\ldots,Z_{k,m}} - \rho_{Z_{k,1}} \otimes \cdots \otimes \rho_{Z_{k,m}}\|_1 \le \sum_{j=1}^m \|\rho_{Z_{k,j},\ldots,Z_{k,m}} - \rho_{Z_{k,j}} \otimes \rho_{Z_{k,j+1},\ldots,Z_{k,m}}\|_1$$

$$\le 2^{\frac{3l}{c}} \sum_{j=1}^m \mathrm{Cor}\left(Z_{k,j}, (Z_{k,j+1}, \ldots, Z_{k,m})\right)$$

$$\le m 2^{\frac{3l}{c}} 2^{-\frac{4\xi l}{c\xi}} = m 2^{-\frac{l}{c}}, \tag{92}$$

where the last inequality follows from $\mathrm{Cor}\left(Z_{k,j}, (Z_{k,j+1}, \ldots, Z_{k,m})\right) \le 2^{-\frac{4\xi l}{c\xi}}$, which is a consequence of exponential decay of correlations and the fact that $Z_{k,j}$ is $4\xi l/c$ sites away from $Z_{k,j+1}, \ldots, Z_{k,m}$. By Lemma 12,

$$D\left(\rho_{Z_{k,1},\ldots,Z_{k,m}}, \rho_{Z_{k,1}} \otimes \cdots \otimes \rho_{Z_{k,m}}\right) \le \sqrt{2m} 2^{-\frac{l}{2c}}. \tag{93}$$

From Eqs. (90) and (92),

$$H_{\max}^\varepsilon(X_{C,l}) \le \sum_{k=1}^{(4\xi+1)} H_{\max}^\delta(\rho_{Z_{k,1}} \otimes \cdots \otimes \rho_{Z_{k,m}}) + (4\xi+1)^2 \log \frac{2}{\varepsilon}, \tag{94}$$

with $\delta := 4^{-(4\xi+1)}\varepsilon - \sqrt{2m} 2^{-l/(2c)}$. By the quantum equipartition property (Lemma 19),

$$H_{\max}^\delta(\rho_{Z_{k,1}} \otimes \cdots \otimes \rho_{Z_{k,m}}) \le H(\rho_{Z_{k,1}} \otimes \cdots \otimes \rho_{Z_{k,m}}) + 10\sqrt{\log(1/\delta)}\sqrt{m}\frac{l}{c}. \tag{95}$$

Setting $c := \frac{10^3}{\xi^3 \varepsilon^2 \log(1/\varepsilon)}$, we find

$$H_{\max}^\delta(\rho_{Z_{k,1}} \otimes \cdots \otimes \rho_{Z_{k,m}}) \le \sum_{j=1}^m H(\rho_{Z_{k,j}}) + \frac{\varepsilon l}{4(4\xi+1)}. \tag{96}$$

and thus

$$H_{\max}^{\varepsilon}(X_{C,l}) \leq \sum_{k=1}^{(4\xi+1)} \sum_{j=1}^{m} H(Z_{k,j}) - \varepsilon l/2, \qquad (97)$$

where we used that $l \geq \bar{l}_0(4\xi + 1)^2 \log\left(\frac{2}{\varepsilon}\right)\varepsilon^{-1}$.

We can now apply Eqs. (97) to (89), and apply subadditivity of von Neumann entropy repeatedly to the region $X_{L,l/2}X_{R,l/2}$ until we reach regions of size l/c to obtain:

$$H(X_{2l}) \leq 2cH(X_{l/c}) - \frac{\varepsilon}{2}l. \qquad (98)$$

Appplying the equation above recursively $\log_c(l/\bar{l}_0)$ times we get

$$H(X_{2l}) \leq \frac{2l}{\bar{l}_0}H(X_{\bar{l}_0}) - \frac{\varepsilon}{2}l \log_c(l/\bar{l}_0), \qquad (99)$$

which leads to a contradiction for $l = \bar{l}_0 c^{2/\varepsilon} = \bar{l}_0 \exp\left(O\left(\log(1/\epsilon)/\epsilon\right)\right)$. $\quad\square$

6. Proof of Theorem 4

In this section we prove Theorem 4. The main idea is to consider a purification of the state and follow the proof of the pure state case, adding the purifying system to the part attributed to the reference system, or environment system, in each of the two applications of the entanglement distillation protocol. Apart from this, the proof will follow closely the argument in the pure state case.

Theorem 4 (restatement). *Let $\rho_{1,\ldots,n}$ be a state defined on ring with (ξ, l_0)-exponential decay of correlations and $n \geq Cl_0/\xi$. Then for any connected region $X \subset [n]$ and every $l \geq 8\xi$,*

$$H_{\max}^{2^{-\frac{l}{8\xi}}}(X) \leq c'l_0 \exp\left(c \log(\xi)\xi\right)(1 + H_{\max}(\rho)) + l, \qquad (100)$$

with $C, c, c' > 0$ universal constants.

Proof. Let $|\psi\rangle_{SE}$ be a purification of the state ρ, i.e., $\mathrm{tr}_R(|\psi\rangle_{SE}\langle\psi|) = \rho$. We start by applying Lemma 7 twice to the two boundaries of the region X, with $\delta := 10^{-8}$, $\varepsilon := \min\left(1/(10^5\xi), \delta/8\right)$, and $\bar{l}_0 := \max\left(4(1 + H_{\max}(\rho))l_0, 4(1 + H_{\max}(\rho))cl_0\varepsilon/(\xi\log(2/\varepsilon))\right)$, for a constant $c > 0$. Then we find there are regions Y and $\tilde{Y}$ (see Fig. 5a), each of size $2l$ with $O(\xi^2 \log\left(2/\varepsilon\right)) \leq l/\bar{l}_0 \leq \exp(O(\log(1/\varepsilon)/\varepsilon))$, and at a distance at most $\bar{l}_0 \exp(O(\log(1/\varepsilon)/\varepsilon))$ sites away from the two boundaries of X, respectively, such that

$$I_{\max}^{\delta/4}(Y_{C,l} : Y_{L,l/2}Y_{R,l/2}) = H_{\max}^{\delta/4}(Y_{C,l}) - H_{\min}^{\delta/4}(Y_{C,l}|Y_{L,l/2}Y_{R,l/2}) \leq \frac{l}{8\xi}, \qquad (101)$$

and likewise for $\tilde{Y}$. We note that from the choice of parameters if follows that $l \geq 300\xi$. By sub-additivity we then find

$$I_{\max}^{\delta}(Y_{C,l} : Y_{L,l/2}Y_{R,l/2}E) \leq \frac{l}{8\xi} + H_{\max}(E). \qquad (102)$$

We now argue that exponential decay of correlations implies that the entropy of $\rho_{Y_{C,l}}$ must be small. Let R be the complementary region to Y (see Fig. 5b). Then from exponential decay of correlations we have $\text{Cor}(Y_{C,l} : R) \leq 2^{-l/(2\xi)}$. Applying part (1) of Lemma 6 we find that

$$H_{\min}^{\delta/4}(Y_{C,l}|Y_{L,l}Y_{R,l}) < -2\log(\delta) + 9, \tag{103}$$

since otherwise

$$\text{Cor}(Y_{C,l} : R) \geq \left(\frac{1}{256} - 26\sqrt{\delta}\right) 2^{-I_{\max}^{\delta/4}(Y_{C,l}:Y_{L,l/2}Y_{R,l/2}E) - 2\log(\delta) + 9} \geq 2^{-\left(\frac{l}{8\xi} - 34\right)}, \tag{104}$$

in contradiction with the correlation length being ξ, since $l \geq 300\xi$.

From Eqs. (103) and (102) we thus have

$$H_{\max}^{\delta/4}(Y_{C,l}) \leq \frac{l}{8\xi} - 2\log(\delta) + 9 + H_{\max}(\rho), \tag{105}$$

where we used $H(E) = H(\rho)$, since E is purification of the state.

Applying the same reasoning to $\tilde{Y}$ we find that also $H_{\max}^{\delta/4}(\tilde{Y}_{C,l}) \leq \frac{l}{8\xi} - 2\log(\delta) + 9 + H_{\max}(\rho)$.

Define A as the region between $Y_{C,l}$ and $\tilde{Y}_{C,l}$, $B := Y_{C,l}\tilde{Y}_{C,l}$, and C the complementary region to AB (see Fig. 5c). Note that A differs from X by at most $2\bar{l}_0 \exp(O(\log(1/\varepsilon)/\varepsilon))$ sites and so by Lemma 18, for every $\nu > 0$,

$$H_{\max}^{2\nu}(X) \leq H_{\max}^{\nu}(A) + 2\bar{l}_0 \exp(O(\log(1/\varepsilon)/\varepsilon)) + \log\frac{2}{\nu^2}. \tag{106}$$

We now prove an upper bound on $H_{\max}^{\nu}(A)$, with $\nu = 0.01$. Since A and C are separated by l sites, we have from exponential decay of correlations that

$$\text{Cor}(A : C) \leq 2^{-l/\xi}. \tag{107}$$

From part (2) of Lemma 6 and Eq. (107) we then find that

$$H_{\max}^{\nu}(C) < 3H_{\max}^{\delta}(BE), \tag{108}$$

since otherwise, using that $l \geq 300\xi$,

$$\text{Cor}(A : C) \geq \left(\frac{\nu}{2} - 2\delta - \sqrt{\frac{4\delta}{1 - 2\nu - 2\delta}}\right)^2 2^{-3\left(H_{\max}^{\delta}(BE) + 16\log(\delta^{-2} + \nu^{-2}) + 5\right)} \geq 2^{-3l/(4\xi)}, \tag{109}$$

contradicting Eq. (107). In the last inequality of the equation above we used that

$$\begin{aligned}
H_{\max}^{\delta}(BE) &= H_{\max}^{\delta}(Y_{C,l}\tilde{Y}_{C,l}E) \\
&\leq H_{\max}(\rho) + H_{\max}^{\delta/4}(Y_{C,l}) + H_{\max}^{\delta/4}(\tilde{Y}_{C,l}) + \log\frac{8}{\delta^2} \\
&\leq 3H_{\max}(\rho) + \frac{l}{4\xi} - 4\log(\delta) + 18 + \log\frac{8}{\delta^2}, \tag{110}
\end{aligned}$$

where the first inequality follows from subadditivity of the max-entropy (Lemma 18) and the second from Eq. (105).

Using subadditivity of the max-entropy (Lemma 18) again, we get

$$H_{\max}^{3\nu+\delta}(A) = H_{\max}^{3\nu+\delta}(BCE) \leq H_{\max}(\rho) + H_{\max}^{\delta}(B) + H_{\max}^{\nu}(C) + \log \frac{2}{\nu^2}$$

$$\leq 4H_{\max}^{\delta}(B) + \log \frac{2}{\nu^2}$$

$$\leq 2\bar{l}_0 \exp(O(\log(1/\varepsilon)/\varepsilon)), \tag{111}$$

where we used Eq. (108) and that $H_{\max}(B) \leq 2l \leq \bar{l}_0 \exp\left(O(\log(1/\varepsilon)/\varepsilon)\right)$. Therefore by Eq. (106),

$$H_{\max}^{6\nu+2\delta}(X) \leq 2\bar{l}_0 \exp\left(O(\log(1/\varepsilon)/\varepsilon)\right) = O(l_0) \exp\left(O(\log(1/\varepsilon)/\varepsilon)\right). \tag{112}$$

This is already an area law for X, although with a fixed error (recall that we fixed $\delta = 10^{-8}$ and $\nu = 0.01$). To finish the proof we show how applying exponential decay of correlations once more we can reduce the error. Let Z be a region of size $2l$ separating X from the remaining sites by a distance l, and denote by U the complementary region to XZ (see Fig. 5 d). By exponential decay of correlations we have $\text{Cor}(X : U) \leq 2^{-l/\xi}$. Using part (3) of Lemma 6 we then get

$$H_{\max}^{2^{-l/(4\xi)}}(X) \leq H_{\max}^{2^{-l/(2\xi)+1}}(X) \leq H_{\max}^{6\nu+2\delta}(X) + 2l \leq O(l_0) \exp(O(\log(1/\varepsilon)/\varepsilon)) + 2l, \tag{113}$$

where the first inequality follows assuming $l \geq 4\xi$. $\quad\square$

7. Proofs of Corollaries 2 and 3

Proof Corollary 2: The proof follows from Lemma 2 of [21]. Let $\{\lambda_i\}$ be the eigenvalues of $\rho_X := \text{tr}_{\backslash X}(|\psi\rangle\langle\psi|)$ in decreasing order. Theorem 1 gives that for every l,

$$\sum_{i:i \geq 2^{c'l_0 \exp(c\log(\xi)\xi)+l}} \lambda_i \leq 2^{-l/8\xi}. \tag{114}$$

Eq. (5) follows by maximizing $-\sum_i \lambda_i \log(\lambda_i)$ subject to the constraints of Eq. (114) (see Lemma 2 of [21]).

Proof Corollary 3: The corollary follows from the argument of [21] (section "Matrix Product States").

8. Conclusions and Open Questions

In this work we proved that for one-dimensional quantum states an area law for their entanglement entropy follows merely from the fact that the state has exponential decay of correlations. While intuitively very natural, the relation of exponential decay of correlations and area law was put into question by the peculiar kind of correlations embodied in the so-called quantum data hiding states. The results of the paper thus shows that, despite the difficulties caused by these types of correlations, the physically motivated intuition is nonetheless correct.

In a sense the obstruction provided by ideas from quantum information theory, namely the concept of data hiding states, can be overcome by considering the problem *also* from

the perspective of quantum information theory. In particular we employed the central idea in quantum Shannon theory of *decoupling* two quantum systems by performing a random measurement in one of them—an idea that has been used to derive the best known protocols for a variety of information theoretic tasks [44,55,56,58,60,70,71])— as well as recent developments in the framework of single-shot quantum information theory [38,64]. In this respect the results of this paper represent an interesting application of this framework to a problem *outside* of information theory, giving further evidence of its utility (see e.g. Refs. [72,73] for other examples).

We now list a few open problems:

1. Can we improve the dependency of correlation length in the entanglement entropy bound? The exponential dependency found here seems hardly optimal, and for groundstates of 1D gapped local Hamiltonians—an important class of states with exponential decay of correlations—the recent result of Arad, Kitaev, Landau and Vazirani [23] shows that such an improvement is indeed feasible. A possible direction to get a sharper bound would be to improve the result about saturation of mutual information, getting a better bound on the size of the region one must vary in order to get small mutual information, perhaps exploiting the fact that the state under consideration has exponential decay of correlations.

2. How small can we choose the constant of the smoothing term in Eq. (100) of Theorem 1? We believe this is an interesting question because an improvement from $l/8\xi$ to any number strictly smaller than one would imply a matrix product representation of *sublinear* bond dimension, which by the methods of Refs. [74,75] would lead to a subexponential-time algorithm for obtaining the ground state of gapped 1D Hamiltonians.

3. Can we prove an extension of Theorem 7 for general states, without the assumption of exponential decay of correlations? Hastings' result about the saturation of mutual information [21] (see Lemma 16) and its single-shot counterpart given by Lemma 7 appear to be powerful results about the distribution of correlations on different length-scales. Are there more applications of them?

4. Can we extend the result to mixed states with exponential decay of correlations? Theorem 4 gives a first step in this direction, however the result is not completely satisfactory as the statement is only meaningful for states of low entropy. In turn, it would be very interesting to explore whether for general mixed states exponential decay of correlations implies a bound on the *mutual information* of an arbitrary region with its complementary region (of the same flavour as the area law for thermal states proved in [33]). We believe the techniques developed in this paper might be useful in addressing this problem, although a new idea appears to be needed.

5. Can we prove area law from exponential decay of correlations for states of an infinite number of particles? The challenge here is how to generalize the information-theoretic tools employed in the proof, in particular the entanglement distillation protocol, to the setting of von Neumann algebras. In this respect the recent results of Ref. [69] might be useful.

6. Finally, can the approach of this paper be used to prove an area law from exponential decay of correlations in higher dimensions, or even *just* an area law for groundstates of gapped local Hamiltonians in 2D? The latter is one of the most important open problems in the field of quantum Hamiltonian complexity [35] and we hope this work will drawn the attention of quantum information theorists to it. Again a new idea appears to be needed, the main difficulty being that in higher dimensions the separation distance between two regions is not of the same order as the num-

ber of sites in the separating region, a feature that was crucially explored in this work.

Appendix A: Purified Distance

Let $\mathcal{H}$ be a finite dimensional vector space and $\mathcal{B}(\mathcal{H})$ the set of linear operators on $\mathcal{H}$. We define set of subnormalized states $\mathcal{D}_{\leq}(\mathcal{H}) = \{\rho \in \mathcal{B}(\mathcal{H}) : \mathrm{tr}(\rho) \leq 1\}$ and the set of normalized states $\mathcal{D}(\mathcal{H}) = \{\rho \in \mathcal{B}(\mathcal{H}) : \mathrm{tr}(\rho) = 1\}$.

Let $F(\rho, \sigma) = \mathrm{tr}(\sqrt{\sigma^{1/2} \rho \sigma^{1/2}})$ be the fidelity of ρ and σ. We quantify the distance of quantum states by the purified distance [63]:

Definition 10. Let $\rho, \sigma \in \mathcal{D}_{\leq}(\mathcal{H})$. The purified distance between ρ and σ is defined as

$$D(\rho, \sigma) = \sqrt{1 - \overline{F}(\rho, \sigma)^2}, \tag{A1}$$

where $\overline{F}(\rho, \sigma) = F(\rho, \sigma) + \sqrt{(1 - \mathrm{tr}(\rho))(1 - \mathrm{tr}(\sigma))}$ denotes the generalized fidelity.

We define the ε-ball around ρ as

$$\mathcal{B}^{\varepsilon}(\rho) = \{\rho' \in \mathcal{D}_{\leq}(\mathcal{H}) : D(\rho, \rho') \leq \varepsilon\}. \tag{A2}$$

The next lemma is a slight variant of Uhlmann's theorem for the fidelity.

Lemma 11 (Uhlmann's theorem for purified distance [63]). *Let* $\rho, \sigma \in \mathcal{D}_{\leq}(\mathcal{H})$. *Then*

$$D(\rho, \sigma) = \min_{|\psi\rangle, |\phi\rangle} D(|\psi\rangle\langle\psi|, |\phi\rangle\langle\phi|), \tag{A3}$$

where the minimum is taken over purifications $|\psi\rangle, |\phi\rangle$ *of* ρ *and* σ, *respectively.*

For $\rho, \sigma \in \mathcal{D}_{\leq}(\mathcal{H})$, define

$$D_1(\rho, \sigma) := \frac{1}{2}\|\rho - \sigma\|_1 + \frac{1}{2}|\mathrm{tr}(\rho) - \mathrm{tr}(\sigma)|. \tag{A4}$$

In Ref. [63] it is shown that

$$D_1(\rho, \sigma) = \max_{0 \leq M \leq \mathrm{id}} |\mathrm{tr}(M(\rho - \sigma))|. \tag{A5}$$

Lemma 12 (Purified distance versus trace norm [63]). *For* $\rho, \sigma \in \mathcal{D}_{\leq}(\mathcal{H})$,

$$D_1(\rho, \sigma) \leq D(\rho, \sigma) \leq \sqrt{2D_1(\rho, \sigma)}. \tag{A6}$$

One of the main reasons why it is useful to consider the purified distance is the following:

Lemma 13 (Purified distance of extensions [63]). *Let* $\rho, \sigma \in \mathcal{D}_{\leq}(\mathcal{H})$ *and* $\tilde{\rho} \in \mathcal{D}_{\leq}(\mathcal{H} \otimes \mathcal{H}')$ *be an extension of* ρ. *Then there exists an extension* $\tilde{\sigma} \in \mathcal{D}_{\leq}(\mathcal{H} \otimes \mathcal{H}')$ *of sigma satisfying* $D(\rho, \sigma) = D(\tilde{\rho}, \tilde{\sigma})$.

Appendix B: Properties of Entropies

Lemma 14. *Let $\rho \in \mathcal{D}_{\leq}(\mathcal{H})$. Then,*

1. *There exists a projector P such that $\log |P| = H^{\delta}_{\max}(\rho)$ and $\mathrm{tr}(P\rho) \geq 1 - 2\delta$.*
2. *If a projector P satisfies $\mathrm{tr}(P\rho) \geq 1 - \delta$, then $H^{\sqrt{\delta}}_{\max}(\rho) \leq \log |P|$.*
3. *Let $\{\lambda_k\}$ be the eigenvectors of ρ in decreasing order. Then $\sum_{k=1}^{2^{H^{\delta}_{\max}(\rho)}} \lambda_k \geq 1 - 2\delta$*

Proof. Let ρ_{δ} be such that $D(\rho, \sigma) \leq \delta$ and $H^{\delta}_{\max}(\rho) = H_{\max}(\rho_{\delta})$. Then by Lemma 12, $D_1(\rho, \sigma) \leq \delta$ and since $\mathrm{tr}(\rho) = 1$, $\mathrm{tr}(\rho_{\delta}) \geq 1 - \delta$. Now, let P be the projector onto the support of ρ_{δ}, which satisfies $\log |P| = H^{\delta}_{\max}(\rho)$. We then obtain

$$|\mathrm{tr}(P\rho) - \mathrm{tr}(\sigma)| = |\mathrm{tr}(P(\rho - \sigma))| \leq D_1(\rho, \sigma) \leq \delta. \tag{B1}$$

Therefore $\mathrm{tr}(P\rho) \geq 1 - 2\delta$, proving part 1 of the lemma.

To prove part 2, it is enough to take $\rho_{\delta} := P\rho P$ and note that $F(\rho, \rho_{\delta}) = \sqrt{\mathrm{tr}(P\rho)}$ for normalized ρ, and so $D(\rho, P\rho P) \leq \sqrt{1 - \mathrm{tr}(P\rho)} \leq \sqrt{\delta}$.

Finally, the third part is a consequence of the statement of the first part and the relation

$$\max_Q \left\{ \mathrm{tr}(Q\rho) : Q^2 = Q, \mathrm{rank}(Q) = q \right\} = \sum_{k=1}^{q} \lambda_k. \tag{B2}$$

$\square$

Following Ref. [68] we define the max-relative entropy as

$$S_{\max}(\rho||\sigma) := \min\{\lambda : \rho \leq 2^{\lambda}\sigma\} \tag{B3}$$

and the ε-smooth relative entropy as

$$S^{\varepsilon}_{\max}(\rho||\sigma) := \min_{\tilde{\rho} \in \mathcal{B}_{\varepsilon}(\rho)} S_{\max}(\tilde{\rho}||\sigma). \tag{B4}$$

The next lemma was first proven in Ref. [61] (see also Ref. [62]) and is a fundamental piece in the proof of Theorem 7.

Lemma 15 (Quantum substate theorem [62]). *Let $\rho, \sigma \in \mathcal{D}(\mathbb{C}^d)$ be such that $supp(\rho) \subseteq supp(\sigma)$. For any $\varepsilon \in (0, 1)$,*

$$S^{\varepsilon}_{\max}(\rho||\sigma) \leq \frac{S(\rho||\sigma)}{\varepsilon} + \log \frac{1}{1 - \varepsilon}. \tag{B5}$$

A version of next lemma first appeared in [21] (see also [8]) where it was used to prove an area law for groundstates of 1D gapped Hamiltonians.

Lemma 16 (Saturation of mutual information). *Let $\rho_{1,...,n} \in \mathcal{D}((\mathbb{C}^2)^{\otimes n})$ and $s \in [n]$ be a particular site. Then for all $\varepsilon > 0$ and $l_0 > 0$ there is a l satisfying $1 \leq l/l_0 \leq \exp(O(1/\varepsilon))$ and a connected region $X_{2l} := X_{L,l/2}X_{C,l}X_{R,l/2}$ of $2l$ sites (the borders $X_{L,l/2}$ and $X_{R,l/2}$ with $l/2$ sites each, and the central region $X_{C,l}$ with l sites) centred at most $l_0 \exp(O(1/\varepsilon))$ sites away from s (see Fig. 4) such that*

$$I(X_{C,l} : X_{L,l/2}X_{R,l/2}) \leq \varepsilon l. \tag{B6}$$

Proof. We prove Eq. (B6) by contradiction. Suppose that for all $l_0 \leq l \leq l_0 \exp\left(O(\log(1/\varepsilon))\right)$

$$H(X_{2l}) \leq H(X_{C,l}) + H(X_{L,l/2} X_{R,l/2}) - \varepsilon l. \tag{B7}$$

Then we show that this leads to the entropy $H(X_{2l})$ being negative for $l = l_0 \exp(O(1/\varepsilon))$. Using subadditivity of entropy

$$H(X_{2l}) \leq H(X_{C,L,l/2}) + H(X_{C,R,l/2}) + H(X_{L,l/2}) + H(X_{R,l/2}) - \varepsilon l, \tag{B8}$$

where $X_{C,L,l/2}$ and $X_{C,R,l/2}$ are two regions such that $X_{C,l} = X_{C,L,l/2} X_{C,R,l/2}$. Thus we have

$$H(X_{2l}) \leq 4H(X_{l/2}) - \epsilon l \tag{B9}$$

We can now apply Eq. (B7) to $H(X_{l/2})$. Doing so recursively, until we reach regions of size l_0, we get

$$H(X_{2l}) \leq l - \log_4 (l/l_0) \varepsilon l, \tag{B10}$$

which leads to a contradiction choosing $l = l_0 \exp(O(1/\varepsilon))$. □

The difference of Lemma 16 with the result of [8,21] is that the latter is concerned with the mutual information between two equally sized neighbouring regions $X_{L,l} X_{R,l}$. Note however that the proofs are very similar in both cases.

The next lemma is a version of Fannes inequality [66] for the purified distance, which follows directly from the version of the inequality proved in [67] and Lemma 12.

Lemma 17 (Fannes-type inequality [67]). *Let $\rho, \sigma \in \mathcal{D}(\mathbb{C}^d)$. Then*

$$|H(\rho) - H(\sigma)| \leq \log(d-1)D(\rho,\sigma) + h(D(\rho,\sigma)) \tag{B11}$$

with $h(\varepsilon) := -\varepsilon \log(\varepsilon) - (1-\varepsilon)\log(1-\varepsilon)$ the binary entropy.

In proof of the main theorems we also make repeatedly use of the following version of the subadditivity inequality for max-entropy, which follows directly from a chain rule relation of Ref. [64].

Lemma 18 (Subadditivity smooth max-entropy). *Given a state ρ_{AB} and $\varepsilon, \varepsilon', \varepsilon'' > 0$,*

$$H_{\max}^{\varepsilon+\varepsilon'+2\varepsilon''}(AB) \leq H_{\max}^{\varepsilon'}(A) + H_{\max}^{\varepsilon''}(B) + \log \frac{2}{\varepsilon^2(1-\varepsilon-\varepsilon'-2\varepsilon'')}. \tag{B12}$$

Moreover we have

$$H_{\max}(AB) \leq H_{\max}(A) + H_{\max}(B) \tag{B13}$$

Proof. We first prove the formula (B12). Eq. (5.12) of Ref. [64] with a trivial C system gives

$$H_{1/2}^{\varepsilon+\varepsilon'+2\varepsilon''}(AB) \leq H_{1/2}^{\varepsilon'}(A|B) + H_{1/2}^{\varepsilon''}(B) + \log\frac{2}{\varepsilon^2}. \tag{B14}$$

Using the inequality

$$H_{1/2}^{\varepsilon'}(A|B) \leq H_{1/2}^{\varepsilon'}(A). \tag{B15}$$

we obtain

$$H_{1/2}^{\varepsilon+\varepsilon'+2\varepsilon''}(AB) \leq H_{1/2}^{\varepsilon'}(A) + H_{1/2}^{\varepsilon''}(B) + \log\frac{2}{\varepsilon^2}. \tag{B16}$$

In the case of zero smoothing $\varepsilon' = 0$, this inequality follows directly from the relation $H_{1/2}(A|B) = \max_\sigma \log F(\rho_{AB}, \mathrm{id}_A \otimes \sigma_B)^2$ with id_A the identity onto the support of ρ_A. For $\varepsilon' > 0$, let $\tilde{\rho}_{AB} \in \mathcal{B}_{\varepsilon'}(\rho_{AB})$ be such that $\mathrm{rank}(\tilde{\rho}_A) = 2^{H_{1/2}^{\varepsilon'}(A)}$. Then

$$H_{1/2}^{\varepsilon'}(A|B)_\rho = H_{1/2}(A|B)_{\tilde{\rho}} \leq H_{1/2}(\tilde{\rho}_A) = H_{1/2}^{\varepsilon'}(A). \tag{B17}$$

Eq. (B12) then follows from the relation

$$H_{\max}^{\varepsilon}(X) \leq H_{1/2}^{\varepsilon}(X) - \log(1-\varepsilon). \tag{B18}$$

from Lemma 4.1 of [76].

The formula (B12) just says that the rank of a state is no greater than the product of ranks of its subsystems. □

The next lemma is from Refs. [63,64] is used in the proof of Lemma 7 to relate the max-entropy and the von Neumann entropy. We actually need a slight generalization of the original formulation in which the systems are not necessarily identical. The original proof however carries through to this case without any modification.

Lemma 19 (Quantum equipartition property [63,64]). *Let $\pi_{1,\ldots,n} = \pi_1 \otimes \cdots \otimes \pi_n$, with $\pi_k \in \mathcal{D}(\mathbb{C}^d)$ for all $k \in [n]$. Then*

$$\frac{1}{n} H_{\max}^{\varepsilon}(\pi_{1,\ldots,n}) \leq \frac{1}{n} H(\pi_{1,\ldots,n}) + 4d\sqrt{\frac{\log(2/\varepsilon^2)}{n}}. \tag{B19}$$

Appendix C: Bound on Data Hiding

The next lemma gives a limit to quantum data hiding in terms of the minimal local dimension of a bipartite operator.

Lemma 20. *For every $L \in \mathbb{B}(\mathbb{C}^d \otimes \mathbb{C}^D)$, with $d \leq D$,*

$$\|L\|_1 \leq d^2 \max_{\|X\|\leq 1, \|Y\|\leq 1} |tr((X \otimes Y)L)|. \tag{C1}$$

Proof. From the variational characterization of the trace norm there is a M with $\|M\| \leq 1$ such that $\|L\|_1 = \mathrm{tr}(ML)$. As will be shown below, we can write

$$M = \sum_{k=1}^{d^2} X_k \otimes Y_k, \tag{C2}$$

with $\|X_k\| \leq 1$, $\|Y_k\| \leq 1$. Then we have

$$\|L\|_1 = \mathrm{tr}(ML) = \sum_{k=1}^{d^2} \mathrm{tr}((X_k \otimes Y_k)L) \leq d^2 \max_{\|X\| \leq 1, \|Y\| \leq 1} |\mathrm{tr}((X \otimes Y)L)|. \tag{C3}$$

We now show (C2). To this end we write M in block form:

$$M = \sum_{a=1}^{d} \sum_{b=1}^{d} |a\rangle\langle b| \otimes M_{a,b} \tag{C4}$$

where $M_{a,b}$ are $D \times D$ matrices. It suffices to prove that $||M_{a,b}|| \leq 1$. For this we use that for every operator X,

$$||X|| = \max_{|\psi\rangle, |\phi\rangle} \langle \psi|X|\phi\rangle. \tag{C5}$$

Then

$$1 \geq ||M|| = \max_{|\psi\rangle, |\phi\rangle} \langle \psi|M|\phi\rangle \geq \max_{|a,\psi'\rangle, |b,\phi'\rangle} \langle a, \psi'|M|b, \phi'\rangle$$
$$= \max_{a,b} \max_{|\psi'\rangle, |\phi'\rangle} \langle \psi'|M_{a,b}|\phi'\rangle = \max_{a,b} ||M_{a,b}||. \tag{C6}$$

$\square$

Appendix D: Decoupling in Haar Random States

The objective of this section is to state and prove the following well-known result (see e.g. [27]):

Lemma 21. *Let $|\psi\rangle_{AB}$ be a Haar random state in $\mathcal{H}_A \otimes \mathcal{H}_B$ with $|A| \geq |B|$. Then*

$$\mathbb{E}\left(D(\rho_B, \tau_B)\right) \leq \left(2\frac{|B|}{|A|}\right)^{1/4}, \tag{D1}$$

where ρ_B is the B reduced density matrix of $|\psi\rangle_{AB}$ and τ_B is the maximally mixed state on $\mathcal{H}_B$.

Proof. Lemma V.3 of Ref. [77] gives

$$\mathbb{E}\left(\|\rho_B - \tau_B\|_2\right) \leq \frac{1}{|A|^{1/2}}, \tag{D2}$$

and since for $d \times d$ matrix X, $\|X\|_1 \leq \sqrt{d}\|X\|_2$,

$$\mathbb{E}\left(\|\rho_B - \tau_B\|_1\right) \leq \sqrt{\frac{|B|}{|A|}}. \tag{D3}$$

The statement then follows from Lemma 12. $\square$

Appendix E: Decay of Correlations in MPS

Given a quantum channel $\Lambda(\rho) = \sum_k A_k \rho A_k^\dagger$ we define the following associated operator

$$\pi_\Lambda := \sum_k A_k \otimes \overline{A}_k. \tag{E1}$$

The eigenvalues and eigenvectors of Λ are given by the pairs (λ_X, X) such that $\Lambda(X) = \lambda_X X$. The operator π_Λ is useful because it has the same eigenvalues as Λ and its eigenvectors are in one-to-one correspondence with the eigenvectors of Λ: If X is an eigenvector of Λ, then $(\mathrm{id} \otimes X)|\Phi\rangle$ is an eigenvector of π_Λ, with $|\Phi\rangle$ the maximally entangled state.

The following lemma was first given in Ref. [39] and we present here a proof in order to flesh out the dependency on the bond dimension of the state.

Lemma 22. *Let* $\Lambda : \mathcal{D}(\mathbb{C}^D) \to \mathcal{D}(\mathbb{C}^D)$ *be a unital channel with Kraus decomposition* $\Lambda(\rho) = \sum_k^d A_k \rho A_k^\dagger$ *and second largest eigenvalue* η. *Consider the matrix product state:*

$$|\psi\rangle_{1,\dots,n} := \sum_{i_1=1}^d \dots \sum_{i_n=1}^d tr(A_{i_1} \dots A_{i_n})|i_1, \dots, i_n\rangle. \tag{E2}$$

Then if $A = [1, r]$, $B = [r+1, r+l+1]$, *and* $C = [r+l+2, n]$,

$$Cor(A : C) \leq \|\rho_{AC} - \rho_A \otimes \rho_C\|_1 \leq D\eta^l. \tag{E3}$$

Proof. Let $V : \mathcal{H}_1 \to \mathcal{H}_1 \otimes \mathcal{H}_2$ be an isometric extension of Λ, with $\mathcal{H}_1 \cong \mathbb{C}^D$ and $\mathcal{H}_2 \cong \mathbb{C}^d$, i.e., $\Lambda(\rho) = \mathrm{tr}_{\mathcal{H}_2}\left(V\rho V^\dagger\right)$. The reduced density matrix ρ_{AC} of $|\psi\rangle_{ABC}$ is given by

$$\rho_{AC} = \mathrm{tr}_{\mathcal{H}_1}\left(V^{n_A}\left(\Lambda^l\left(V^{n_C}(\tau)\right)\right)\right), \tag{E4}$$

with n_A and n_C the number of sites in A and C, respectively, and τ the maximally mixed state in $\mathcal{D}(\mathbb{C}^D)$. We used the notation

$$V^k(\tau) := (V \circ \dots \circ V)\, \rho\left(V^\dagger \circ \dots \circ V^\dagger\right), \tag{E5}$$

with the k fold composition of V, and likewise for Λ^l.

Let $\Phi : \mathcal{D}(\mathbb{C}^D) \to \mathcal{D}(\mathbb{C}^D)$ be the completely depolarized channel: $\Phi(\rho) = \tau$. Then

$$\rho_A \otimes \rho_C = \mathrm{tr}_{\mathcal{H}_1}\left(V^{n_A}\left(\Phi^l\left(V^{n_C}(\tau)\right)\right)\right). \tag{E6}$$

Therefore

$$\|\rho_{AC} - \rho_A \otimes \rho_C\|_1 \leq \|\Lambda^l\left(V^{n_C}(\tau)\right) - \Phi^l\left(V^{n_C}(\tau)\right)\|_1 \leq \|\Lambda^l - \Phi\|_\diamond, \tag{E7}$$

where the first inequality follows from the monotonicity of trace norm under trace preserving CP maps and the second inequality from the definition of the diamond norm. Using the relation

$$\|\Lambda_1 - \Lambda_2\|_\diamond \leq D\|\Lambda_1 - \Lambda_2\|_{2\to2} = D\|\pi_{\Lambda_1} - \pi_{\Lambda_2}\|_\infty, \tag{E8}$$

valid for every two channels $\Lambda_1, \Lambda_2 : \mathcal{D}(\mathbb{C}^D) \to \mathcal{D}(\mathbb{C}^D)$, we get

$$\|\Lambda^l - \Phi\|_\diamond \leq D\|\pi_{\Lambda^l} - \pi_\Phi\|_\infty = D\left\|(\pi_\Lambda)^l - \pi_\Phi\right\|_\infty = D\eta^l, \tag{E9}$$

where we used that π_Φ is the projector onto the maximum eigenvector of π_Λ (of value one and corresponding eigenvector τ). $\quad\square$

Appendix F: Correlations in Quantum Expander States

In this section we show how the results of Ref. [31] imply there are correlations between regions separated by of order $\log(D)/\log(d)$ sites in quantum expander states with matrices given by independent Haar unitaries. We start with the following lemma,

Lemma 23. *Let* $\Lambda : \mathcal{D}(\mathbb{C}^D) \to \mathcal{D}(\mathbb{C}^D)$ *be a unital channel with Kraus decomposition* $\Lambda(\rho) = \sum_{k=1}^{d} A_k \rho A_k^{\dagger}$. *Consider the matrix product state*

$$|\psi\rangle_{1,\dots,n} := \sum_{i_1=1}^{d} \dots \sum_{i_n=1}^{d} tr(A_{i_1} \dots A_{i_n})|i_1, \dots, i_n\rangle. \tag{F1}$$

Then the reduced density matrix ρ_l *of l sites is such that*

$$tr\left(\rho_l^2\right) = \frac{1}{D^2} \sum_{i,j} tr\left(\Lambda^l(|i\rangle\langle j|)\Lambda^l(|j\rangle\langle i|)\right). \tag{F2}$$

Proof. We have

$$\rho_l = \sum_{\substack{i_1 \dots i_l=1 \\ j_1 \dots j_l=1}}^{d} tr\left(A_{i_1} \dots A_{i_l} \frac{\mathrm{id}}{D} A_{j_1}^{\dagger} \dots A_{j_l}^{\dagger}\right) |i_1\rangle\langle j_1| \otimes \dots \otimes |i_l\rangle\langle j_l|. \tag{F3}$$

Note that this state is reduced density matrix ρ_E of a tripartite state $|\psi\rangle_{ABE}$, with $|A| = |B| = D$, $|E| = d^l$, defined as

$$|\psi\rangle_{ABE} := \frac{1}{d^{l/2}} \sum_{i_1,\dots,i_l=1}^{d} (\mathrm{id}_A \otimes A_{i_1} \dots A_{i_l})|\Phi\rangle_{AB}|i_1, \dots, i_l\rangle_E, \tag{F4}$$

where the product of operators A_{i_j} acts on system B and

$$|\Phi\rangle_{AB} = \frac{1}{\sqrt{D}} \sum_{k=1}^{D} |k\rangle_A |k\rangle_B \tag{F5}$$

is the maximally entangled state on AB. One then finds that the AB reduced density matrix of this state reads

$$\rho_{AB} = (\mathrm{id}_A \otimes \Lambda_B^l)(|\Phi\rangle\langle\Phi|). \tag{F6}$$

Eq. (F2) then follows from $tr\left(\rho_l^2\right) = tr\left(\rho_E^2\right) = tr\left(\rho_{AB}^2\right)$ and Eq. (F6). □

The next lemma follows from the previous lemma and the results of [31,78].

Lemma 24. *Let* $A_i := \frac{1}{\sqrt{d}} U_i$, *with* U_i *chosen independently at random according to the Haar measure. Then there exists constants* $\gamma > 0, k > 1$ *such that the state given by Eq. (F1) satisfies*

$$\langle tr(\rho_l^2)\rangle \leq \frac{1}{D^2} + \frac{kl}{d^l}. \tag{F7}$$

for any l such that $l \leq D^{\gamma}$, *where* $\langle\cdots\rangle$ *denotes the average over* U_i's.

Proof. For $A_i = \frac{1}{\sqrt{d}} U_i$ Hastings proved an upper bound for the average of the RHS of Eq. (F2) [31,78]. The bound says there are constants $\gamma > 0, k > 1$ such that

$$\sum_{i,j} \text{tr}\left(\Lambda^l(|i\rangle\langle j|)\Lambda^l(|j\rangle\langle i|)\right) \leq 1 + kl\left(\frac{1}{\sqrt{d}}\right)^{2l} D^2 \tag{F8}$$

for any $l \leq D^\gamma$. Using Lemma 23 we get $\text{tr}\left(\rho_l^2\right) \leq \frac{1}{D^2} + \frac{kl}{d^l}$. $\square$

We are now ready to prove the main result of this section. The idea is to reduce the problem to one similar to the case of random states, where one could argue there were correlations by finding two neighbouring regions AB which were decoupled. Here there are no neighbouring regions which are decoupled a priori. However using the result of Lemma 24 that regions of size $O\left(\log(D)/\log(d)\right)$ have very large entropy, we will be able to show that one can find a projector that has large probability of being measured in A, and that decouples it from its neighbouring region B. This idea, of showing the existence of correlations by decoupling one subsystem from its neighbouring one by measurements, will be the central idea in the proof of Theorem 1. Here we consider the particular case in which the entropy is almost the maximum value possible, what makes the argument much simpler.

Proposition 25. *Let $A_i := \frac{1}{\sqrt{d}} U_i$, with U_i chosen independently at random according to the Haar measure. Then the state given by Eq. (F1) has $Cor(A : C) \geq \Omega(1/l)$ between regions A and C separated by l sites for any $l \leq O\left(\log(D)/\log(d)\right)$.*

Proof. Consider a region $\{1, \ldots, l\}$ with $l = O\left(\log(D)/\log(d)\right)$ and divide it into subregions A and B of equal sizes, as in Fig. 3. Let $\varepsilon := 0.01$. Applying Lemma 26 with $N = kl/\varepsilon$, and using Lemma 12, it follows there is a projector Q acting on subsystem A, of dimension bigger than two, such that $\langle\psi|(Q \otimes \text{id}_{BC})|\psi\rangle \geq O(l/\varepsilon)$ and

$$D\left(\rho_{A'B}, \tau_{A'} \otimes \rho_B\right) \leq \sqrt{2\delta}, \tag{F9}$$

with $\rho_{A'B}$ the reduced density matrix of the postselected state $|\phi\rangle_{A'BC} := (Q \otimes \text{id}_{BC})|\psi\rangle_{ABC}/\sqrt{\langle\psi|(Q \otimes \text{id}_{BC})|\psi\rangle}$, and

$$\delta := 2\sqrt{\frac{d^l \varepsilon}{D^2 kl} + \varepsilon} + \frac{2\varepsilon}{kl}. \tag{F10}$$

Choosing $l = \log(D)/\log(d)$ we get $\delta \leq 3\sqrt{2\varepsilon}$.

From Uhlmann's theorem (Lemma 11 in "Appendix A") we find there is an isometry $V : C \to C_1 C_2$ that can be applied to C such that

$$D((\text{id}_{AB} \otimes V)|\psi\rangle_{ABC}\langle\psi|(\text{id}_{AB} \otimes V)^\dagger, |\Phi\rangle_{AC_1}\langle\Phi| \otimes |\Psi\rangle_{BC_2}\langle\Phi|) \leq \sqrt{2\varepsilon}, \tag{F11}$$

with $|\Phi\rangle_{AC_1} = \dim(A)^{-1/2} \sum_{k=1}^{\dim(A)} |k, k\rangle$ a maximally entangled state between A and C_1 and $|\Psi\rangle_{BC_2}$ a purification of ρ_B. Thus defining $M = \sum_{i=1}^{\dim(Q)/2} |k\rangle\langle k|$ and $N = V\left(M \otimes \text{id}_{C_2}\right) V^\dagger$, we find

$$Cor(A : C) \geq \text{tr}\left((QM_A Q \otimes N_C)(\rho_{AC} - \rho_A \otimes \rho_C)\right) \geq \frac{\varepsilon}{8l}. \tag{F12}$$

Therefore there are strong correlations between the regions A and C separated by $l/2 = \log(D)/(2\log(d))$ sites. $\quad\square$

Lemma 26 (Proposition 4 of [56]). *Let $|\psi\rangle_{ABC}$ be a pure state. Then there exists a POVM $\{P_k\}_{k=1}^{N+1}$ consisting of $N = \left\lfloor \frac{|A|}{L} \right\rfloor$ projectors of rank L and one of rank $|A| - NL \le L$ such that*

$$\sum_k p_k \left\| \rho_{A'B}^k - \tau_{A'} \otimes \rho_B \right\|_1 \le 2\sqrt{L|B|tr\left(\rho_{AB}^2\right)} + 2\frac{L}{|A|}, \tag{F13}$$

with $p_k := \langle\psi|P_k \otimes id_{BC}|\psi\rangle$ and $\rho_{A'B}^k$ the reduced density matrix of the postselected state $|\phi\rangle_{A'BC} := (P_k \otimes id_{BC})|\psi\rangle_{ABC}/\sqrt{p_k}$.

Appendix G: Single-shot State Entanglement Distillation

We start this appendix with a formal definition entanglement distillation. In fact we will consider the more demanding task of single-shot state merging, which not only distills EPR pairs but also transfers the state of one of the parties to the other one. Although we do not make use of this further property in the proof, we use it here in order to state the results of Ref. [58] that we make use in the form they originally appeared.

Definition 27. Consider a tripartite state $|\psi\rangle_{ABC}$. Let A_1 and C_1 be registers each of dimension L. A measurement defined by the POVM elements $\{M_k\}_{k=1}^N$, with $M_k : A \to A_1$, and a set of isometries $\{V_k\}_{k=1}^N$, with $V_k : C \to C_1C'C$, define a (L, N, ε)-quantum state merging protocol for $|\psi\rangle_{ABC}$ with error ε, entanglement distillation rate $\log(L)$, and classical communication cost $\log(N)$, if

$$D\left(\sum_{k=1}^N (\sqrt{M_k} \otimes id_B \otimes V_k)|\psi\rangle_{ABC}\langle\psi|(\sqrt{M_k} \otimes id_B \otimes V_k)^\dagger, |\Phi\rangle_{A_1C_1}\langle\Phi| \otimes |\psi\rangle_{BC'C}\langle\psi|\right) \le \varepsilon \tag{G1}$$

We note that in its most general form, the state merging protocol also makes sense when one must use pre-shared entanglement in order to transfer the A part of the state to C. The application of the protocol in this work, however, does not concern this regime. Note also that as we already mentioned, for our application we do not need the state merging part of the protocol per se, but only the entanglement distillation part. However, it is very important that we distill entanglement with the specified classical communication cost achieved in the single-shot version [58] of the state merging protocol of [55,56], i.e.,

Lemma 28 (Single-shot state merging [58]). *Given a tripartite state $|\psi\rangle_{ABC}$, there is protocol for quantum state merging with*

$$\log(N) \le H_{\max}^\varepsilon(A) - H_{\min}^\varepsilon(A|B) - 4\log\varepsilon + 2\log 13, \tag{G2}$$

$$\log(L) \le H_{\min}^\varepsilon(A|B) - 4\log\varepsilon + 2\log 13. \tag{G3}$$

and error $\delta = 13\sqrt{\varepsilon}$.

References

1. Araki, H., Hepp, K., Ruelle, D.: Asymptotic behaviour of Wightman functions. Helv. Phys. Acta **35**, 164 (1962)
2. Fredenhagen, K.: A remark on the cluster theorem. Commun. Math. Phys. **97**, 461 (1985)
3. Nachtergaele, B., Sims, R.: Lieb-Robinson bounds and the exponential clustering theorem. Commun. Math. Phys. **265**, 119 (2006)
4. Hastings, M.B.: Lieb-Schultz-Mattis in higher dimensions. Phys. Rev. B **69**, 104431 (2004)
5. Hastings, M.B.: Locality in quantum and markov dynamics on lattices and networks. Phys. Rev. Lett. **93**, 140402 (2004)
6. M.B. Hastings: Decay of correlations in fermi systems at non-zero temperature. Phys. Rev. Lett. **93**, 126402 (2004)
7. Hastings, M.B., Koma, T.: Spectral gap and exponential decay of correlations. Commun. Math. Phys. **265**, 781 (2006)
8. Aharonov, D., Arad, I., Landau, Z., Vazirani, U.: The detectability lemma and its applications to quantum Hamiltonian complexity. New J. Phys. **13**, 113043 (2011)
9. Horodecki, R., Horodecki, P., Horodecki, M., Horodecki, K.: Quantum entanglement. Rev. Mod. Phys. **81**, 865 (2009)
10. Eisert, J., Cramer, M., Plenio, M.B.: Colloquium: area laws for the entanglement entropy. Rev. Mod. Phys. **82**, 277 (2010)
11. Bennett, C.H., Bernstein, H.J., Popescu, S., Schumacher, B.: Concentrating partial entanglement by local operations. Phys. Rev. A **53**, 2046 (1996)
12. Bekenstein, J.D.: Black holes and entropy. Phys. Rev. D **7**, 233 (1973)
13. Vidal, G., Latorre, J.I., Rico, E., Kitaev, A.: Entanglement in quantum critical phenomena. Phys. Rev. Lett. **90**, 227902 (2003)
14. Calabrese, P., Cardy, J.: Entanglement entropy and quantum field theory, J. Stat. Mech. P06002 (2004)
15. Audenaert, K., Eisert, J., Plenio, M.B., Werner, R.F.: Entanglement properties of the harmonic chain. Phys. Rev. A **66**, 042327 (2002)
16. Plenio, M.B., Eisert, J., Dreissig, J., Cramer, M.: Entropy, entanglement, and area: analytical results for harmonic lattice systems. Phys. Rev. Lett. **94**, 060503 (2005)
17. Wolf, M.M.: Violation of the entropic area law for Fermions. Phys. Rev. Lett. **96**, 010404 (2006)
18. Aharonov, D., Gottesman, D., Irani, S., Kempe, J.: The power of quantum systems on a line. Commun. Math. Phys. **287**, 41 (2009)
19. Irani, S.: Ground state entanglement in one dimensional translationally invariant quantum systems. J. Math. Phys. **51**, 022101 (2010)
20. Gottesman, D., Irani, S.: The quantum and classical complexity of translationally invariant tiling and hamiltonian problems. FOCS ' 09
21. Hastings, M.: An area law for one dimensional quantum systems. JSTAT, P08024 (2007)
22. Lieb, E.H., Robinson, D.W.: The nite group velocity of quantum spin systems. Commun. Math. Phys. **28**, 251 (1972)
23. Arad, I., Kitaev, A., Landau, Z., Vazirani, U.: In preparation, (2012)
24. Arad, I., Landau, Z., Vazirani, U.: An improved 1D area law for frustration-free systems. Phys. Rev. B **85**, 195145 (2012)
25. Gottesman, D., Hastings, M.B.: Entanglement vs. gap for one-dimensional spin systems. New J. Phys. **12**, 025002 (2010)
26. Verstraete, F., Cirac, J.I.: Matrix product states represent ground states faithfully. Phys. Rev. B **73**, 094423 (2006)
27. Hayden, P., Leung, D.W., Winter, A.: Aspects of generic entanglement. Commun. Math. Phys. **265**, 95 (2006)
28. DiVincenzo, D.P., Leung, D.W., Terhal, B.M.: Quantum data hiding. IEEE Trans. Inf. Theory **48**, 580 (2002)
29. Brandão, F.G.S.L., Christandl, M., Yard, J.: A quasipolynomial-time algorithm for the quantum separability problem. In: Proceedings of ACM symposium on theory of computation (STOC'11)
30. Brandão, F.G.S.L., Christandl, M., Yard, J.: Faithful squashed entanglement. Commun. Math. Phys. **306**, 805 (2011)
31. Hastings, M.B.: Random unitaries give quantum expanders. Phys. Rev. A **76**, 032315 (2007)
32. Hastings, M.B.: Entropy and entanglement in quantum ground states. Phys. Rev. B **76**, 035114 (2007)
33. Wolf, M.M., Verstraete, F., Hastings, M.B., Cirac, J.I.: Area laws in quantum systems: mutual information and correlations. Phys. Rev. Lett. **100**, 070502 (2008)
34. Masanes, L.l.: An area law for the entropy of low-energy states. Phys. Rev. A **80**, 052104 (2009)
35. Osborne, T.: Hamiltonian complexity. Rep. Prog. Phys. **75**, 022001 (2012)

36. Brandão, F.G.S.L., Horodecki, M: Entanglement area law from exponential decay of correlations. Nat. Phys. **9**, 721–726 (2013)
37. Tomamichel, M., Colbeck, R., Renner, R.: Duality between smooth min- and max-entropies. IEEE Trans. Inf. Theory **56**, 4674 (2010)
38. Renner, R.: Ph.D. thesis ETH Zurich (2005)
39. Fannes, M., Nachtergaele, B., Werner, R.F.: Finitely correlated states on quantum spin chains. Commun. Math. Phys. **144**, 443 (1992)
40. Ostlund, Y.S., Rommer, S.: Phys. Rev. Lett. **75**, 3537 (1995)
41. Vidal, G.: Efficient classical simulation of slightly entangled quantum computations. Phys. Rev. Lett. **91**, 147902 (2003)
42. Perez-Garcia, D., Verstraete, F., Wolf, M.M., Cirac, J.I.: Matrix product state representations. Q. Inf. Comp. **7**, 401 (2007)
43. White, S.R.: Density matrix formulation for quantum renormalization groups. Phys. Rev. Lett. **69**, 2863 (1992)
44. Abeyesinghe, A., Devetak, I., Hayden, P., Winter, A.: The mother of all protocols: restructuring quantum information's family tree. J. Proc. R. Soc. A **465**, 2537 (2009)
45. Ben-Aroya, A., Ta-Shma, A.: Quantum expanders: motivation and construction. Theory Comput. **6**, 47 (2010)
46. Gottesman, D.: The Heisenberg representation of quantum computers. arXiv:quant-ph/9807006 (1998)
47. Jozsa, R., Linden, N.: On the role of entanglement in quantum computational speed-up. Proc. R. Soc. Lond. A **459**(2036), 2011–2032 (2003)
48. Valiant, L.G.: Quantum circuits that can be simulated classically in polynomial time. SIAM J. Comput. **31**(4), 1229 (2002)
49. DiVincenzo, D., Terhal, B.: Classical simulation of noninteracting-fermion quantum circuits. Phys. Rev. A **65**, 032325 (2002)
50. Markov, I., Shi, Y.: Simulating quantum computation by contracting tensor networks. SIAM J. Comp. **38**, 963 (2008)
51. Vanden Nest, M.: Simulating quantum computers with probabilistic methods. Quant. Inf. Comp. **11**, 784 (2011)
52. Arad, I., Landau, Z.: Quantum computation and the evaluation of tensor networks. Siam J. Comput. **39**(7), 3089–3121 (2010)
53. Hastings, M.B.: Notes on some questions in mathematical physics and quantum information. arXiv:1404.4327
54. Devetak, I.A. Winter: Distillation of secret key and entanglement from quantum states. Proc. R. Soc. Lond. A **461**, 207 (2005)
55. Horodecki, M., Oppenheim, J., Winter, A.: Partial quantum information. Nature **436**, 673 (2005)
56. Horodecki, M., Oppenheim, J., Winter, A.: Quantum state merging and negative information. Commun. Math. Phys. **269**, 107 (2007)
57. Araki H., Lieb E.H. (1970) Entropy inequalities. Commun. Math. Phys. 18:160
58. Dupuis, F., Berta, M., Wullschleger, J., Renner, R.: One-shot decoupling. Commun. Math. Phys. **328**, 251–284 (2014)
59. Datta, N., Hsieh, M.-H.: The apex of the family tree of protocols: optimal rates and resource inequalities. New J. Phys. **13**, 093042 (2011)
60. Berta, M., Christandl, M., Renner, R.: The quantum reverse shannon theorem based on one-shot information theory. Commun. Math. Phys. **306**, 579 (2011)
61. Jain, R., Radhakrishnan, J., Sen, P.: A theorem about relative entropy of quantum states with an application to privacy in quantum communication. arXiv:0705.2437
62. Jain, R., Nayak A.: Short proofs of the quantum substate theorem. IEEE Trans. Inf. Theory 58, no.6 (2012)
63. Tomamichel, M., Colbeck, R., Renner, R.: A fully quantum asymptotic equipartition property. IEEE Trans. Inf. Theory **55**, 5840 (2009)
64. Tomamichel, M.: A framework for non-asymptotic quantum information theory. PhD Thesis, ETH Zürich 2011. arXiv:1203.2142
65. Ledoux, M.: The concentration of measure phenomenon, volume 89 of Mathematical Surveys and Monographs. American Mathematical Society (2001)
66. Fannes, M.: Commun. Math. Phys. 31, 291 (1973)
67. Audenaert, K.M.R.: A sharp fannes-type inequality for the von neumann entropy. J. Phys. A **40**, 8127 (2007)
68. Datta, N.: Min- and max- relative entropies and a new entanglement monotone. IEEE Trans. Inf. Theory **55**, 2816 (2009)
69. Berta, M., Furrer, F., Scholz, V.B.: The smooth entropy formalism on von neumann algebras. arXiv:1107.5460

70. Bennett, C.H., Devetak, I, Harrow, A.W., Shor, P.W., Winter, A.: Quantum reverse shannon theorem. IEEE Trans. Inf. Theory **60**(5), 2926–2959 (2014)
71. Hayden, P., Horodecki, M., Yard, J., Winter, A.: A decoupling approach to the quantum capacity. Open Syst. Inf. Dyn. **15**, 7 (2008)
72. del Rio, L., Aberg, J., Renner, R., Dahlsten, O., Vedral, V.: The thermodynamic meaning of negative entropy. Nature **474**, 61 (2011)
73. Horodecki, M., Oppenheim, J.: Fundamental limitations for quantum and nano thermodynamics. Nat. Commun. **4**, 2059 (2013)
74. Aharonov, D., Arad, I., Irani, S.: An efficient algorithm for approximating 1d ground states. Phys. Rev. A **82**, 012315 (2010)
75. Schuch, N., Cirac, J.I.: Matrix product state and mean field solutions for one-dimensional systems can be found efficiently. Phys. Rev. A **82**, 012314 (2010)
76. Renner, R., Wolf, S.: Smooth renyi entropy and applications. In: Proceedings international symposium on information theory, ISIT 2004 (2004)
77. Brandão, F.G.S.L., Horodecki, M.: On Hastings' counterexamples to the minimum output entropy additivity conjecture. Open Syst. Inf. Dyn. **17**, 31 (2010)
78. Hastings, M.B.: Private communication (2012)

Communicated by A. Winter