

Locally accurate MPS approximations for ground states of one-dimensional gapped local Hamiltonians

Alexander M. Dalzell¹ and Fernando G.S.L. Brandão^{1,2}

¹*Institute for Quantum Information and Matter,
California Institute of Technology, Pasadena, CA 91125, USA*
²*Google LLC, Venice, CA 90291, USA*

A key feature of ground states of gapped local 1D Hamiltonians is their relatively low entanglement — they are well approximated by matrix product states (MPS) with bond dimension scaling polynomially in the length N of the chain, while general states require a bond dimension scaling exponentially. We show that the bond dimension of these MPS approximations can be improved to a constant, independent of the chain length, if we relax our notion of approximation to be more local: for all length- k segments of the chain, the reduced density matrices of our approximations are ϵ -close to those of the exact state. If the state is a ground state of a gapped local Hamiltonian, the bond dimension of the approximation scales like $(k/\epsilon)^{1+o(1)}$, and at the expense of worse but still poly($k, 1/\epsilon$) scaling of the bond dimension, we give an alternate construction with the additional features that it can be generated by a constant-depth quantum circuit with nearest-neighbor gates, and that it applies generally for any state with exponentially decaying correlations. For a completely general state, we give an approximation with bond dimension $\exp(O(k/\epsilon))$, which is exponentially worse, but still independent of N . Then, we consider the prospect of designing an algorithm to find a local approximation for ground states of gapped local 1D Hamiltonians. When the Hamiltonian is translationally invariant, we show that the ability to find $O(1)$ -accurate local approximations to the ground state in $T(N)$ time implies the ability to estimate the ground state energy to $O(1)$ precision in $O(T(N)\log(N))$ time.

I. INTRODUCTION

In nature, interactions between particles act locally, motivating the study of many-body Hamiltonians consisting only of terms involving particles spatially near each other. An important method that has emerged from this course of study is the Density Matrix Renormalization Group (DMRG) algorithm [1, 2], which aims to find a description of the ground state of local Hamiltonians on a one-dimensional chain of sites. DMRG has been an indispensable tool for research in many-body physics, but the rationale for its empirical success did not become fully apparent until long after it was being widely used. Its eventual justification required two ingredients: first, that the method can be recast [3, 4] as a variational algorithm that minimizes the energy over the set of matrix product states (MPS), a tensor network ansatz for 1D systems; and second, that the MPS ansatz set actually contains a good approximation to the ground state, at least whenever the Hamiltonian has a nonzero spectral gap. More specifically, Hastings [5] showed that ground states of gapped local Hamiltonians on chains with N sites can be approximated to within trace distance ϵ by an MPS with bond dimension (a measure of the complexity of the MPS) only poly($N, 1/\epsilon$), exponentially smaller than what is needed to describe an arbitrary state on the chain. Even taking into account these observations, DMRG is a heuristic algorithm and is not guaranteed to converge to the global energy minimum as opposed to a local minimum; however, a recently developed alternative algorithm [6–8], sometimes referred to as the “Rigorous RG” (RRG) algorithm, avoids this issue and provides a

way one can guarantee finding an ϵ -approximation to the ground state in poly($N, 1/\epsilon$) time.

These are extremely powerful results, but their value breaks down when the chain becomes very long. The bond dimension required to describe the ground state grows relatively slowly, but it still diverges with N . Meanwhile, if we run the RRG algorithm on longer and longer chains, we will eventually encounter an N too large to handle given finite computational resources. Indeed, often we wish to learn something about the ground state in the thermodynamic limit ($N \rightarrow \infty$) but in this case these results no longer apply. Analogues of DMRG for the thermodynamic limit [9–14] — methods that, for example, optimize over the set of constant bond dimension “uniform MPS” consisting of the same tensor repeated across the entire infinite chain — have been implemented with successful results, but these methods lack the second ingredient that justified DMRG: it is not clear how large we must make the bond dimension to guarantee that the set over which we are optimizing contains a good approximation to the ground state.

Progress toward this ingredient can be found in work by Huang [15] (and later by Schuch and Verstraete [16]), who showed that the ground state of a gapped local 1D Hamiltonian can be approximated *locally* by a matrix product operator (MPO) — a 1D tensor network object that corresponds to a (possibly mixed) density operator as opposed to a quantum state vector — with bond dimension independent of N . Here their approximation sacrifices global fidelity with the ground state, which decays exponentially with the chain length, in exchange for *constant* bond dimension, while retaining high fidelity with the ground state reduced density matrices on all

segments of the chain with constant length. In other words, the statistics for measurements of local operators are faithfully captured by the MPO approximation, a notion of approximation that is often sufficient in practice since many relevant observables, such as the individual terms in the Hamiltonian, are local.

However, the result does not provide the necessary ingredient to justify infinite analogues of DMRG because MPO do not make a good ansatz class for variational optimization algorithms. One can specify the matrix elements for an MPO, but the resulting operator will only correspond to a valid quantum state if it is positive semi-definite, and verifying that this is the case is difficult: it is NP-hard for finite chains, and in the limit $N \rightarrow \infty$ it becomes undecidable [17]. Thus, if we attempt to perform variational optimization over the set of constant bond dimension MPO, we can be certain that our search space contains a good local approximation to the ground state, but we have no way of restricting our search only to the set of valid quantum states; ultimately the minimal energy MPO we find may not correspond to any quantum state at all.

In this work, we fix this problem by showing an analogous result for MPS instead of MPO. We show that for any gapped nearest-neighbor Hamiltonian on a 1D chain with N sites, and for any parameters k and ϵ , there is an MPS representation of a state $|\tilde{\psi}\rangle$ with bond dimension $\text{poly}(k, 1/\epsilon)$ such that the reduced density matrix of $|\tilde{\psi}\rangle\langle\tilde{\psi}|$ on any contiguous region of length k is ϵ -close in trace distance to that of the true ground state. Importantly, the bond dimension is independent of N . For general states (including ground states of non-gapped local Hamiltonians), we give a construction with bond dimension that is also independent of N but exponential in k/ϵ . Thus, we provide the missing ingredient for variational algorithms in the thermodynamic limit as we show that a variational set over constant bond dimension MPS contains a state that simultaneously captures all the local properties of the ground state.

We present two proofs of our claim about ground states of gapped Hamiltonians. The first yields superior scaling of the bond dimension, which grows asymptotically slower than $(k/\epsilon)^{1+\delta}$ for any $\delta > 0$; however, it constructs an MPS approximation $|\tilde{\psi}\rangle$ that is long-range correlated and non-injective. In contrast, the second proof constructs an approximation that is injective and can be generated by a constant-depth quantum circuit with nearest-neighbor gates, while retaining $\text{poly}(k, 1/\epsilon)$ bond dimension. The latter construction also follows merely from the assumption that the state has exponential decay of correlations. The proof idea originates with a strategy first presented in [18] and constructs $|\tilde{\psi}\rangle$ by beginning with the true ground state $|\psi\rangle$ and applying three rounds of operations: first, a set of unitaries that, intuitively speaking, removes the short-range entanglement from the chain; second, a sequence of rank 1 projectors that zeroes-out the long-range entanglement; and third, the set of inverse unitaries from step 1 to add back the short-range entan-

glement. Intuitively, the method works because ground states of gapped Hamiltonians have a small amount of long-range entanglement. The non-trivial part is arguing that the local properties are preserved even as the small errors induced in step 2 accumulate to bring the global fidelity with the ground state to zero. The fact that $|\tilde{\psi}\rangle$ can be produced by a constant-depth quantum circuit acting on an initial product state suggests the possibility of an alternative variational optimization algorithm using the set of constant-depth circuits (a strict subset of constant-bond-dimension MPS) as the variational ansatz. Additionally, we note that the disentangle-project-reentangle process that we utilize in our proof might be of independent interest as a method for truncating the bond dimension of MPS. We can bound the truncation error of this method when the state has exponentially decaying correlations.

We also consider the question of whether these locally approximate MPS approximations can be rigorously found (à la RRG) more quickly than their globally approximate counterparts (and if they can be found at all in the thermodynamic limit). We prove a reduction for ground states of translationally invariant Hamiltonians showing that finding approximations to local properties to even a fixed $O(1)$ precision implies being able to find an approximation to the ground state energy to $O(1)$ precision with only $O(\log(N))$ overhead. Since strategies for estimating the ground state energy typically involve constructing a globally accurate approximation to the ground state, this observation gives us an intuition that it may not be possible to find the local approximation much more quickly than the global approximation, despite the fact that the bond dimensions required for the two approximations are drastically different.

II. BACKGROUND

A. One-dimensional local Hamiltonians

In this paper we work exclusively with gapped nearest-neighbor 1D Hamiltonians that have a unique ground state. Our physical system is a set of N sites, arranged in one dimension on a line with open boundary conditions (OBC), each with its own Hilbert space \mathcal{H}_i of dimension d . The Hamiltonian H consists of terms $H_{i,i+1}$ that act non-trivially only on \mathcal{H}_i and \mathcal{H}_{i+1} .

$$H = \sum_{i=1}^{N-1} H_{i,i+1} \quad (1)$$

We will always require that $H_{i,i+1}$ be positive semi-definite and satisfy $\|H_{i,i+1}\| \leq 1$ for all i , where $\|\cdot\|$ is the operator norm. When this is not the case it is always possible to rescale H so that it is. We call H translationally invariant if $H_{i,i+1}$ is the same for all i . We will also always assume that H has a unique ground state $|\psi\rangle$ with energy E and an energy gap $\Delta > 0$ to its first excited

state. We let $\rho = |\psi\rangle\langle\psi|$ refer to the (pure) density matrix representation of the ground state. For any density matrix σ and subregion X of the chain, we let σ_X refer to $\text{Tr}_{X^c}(\sigma)$, the reduced density matrix of σ after tracing out the complement X^c of X .

B. Matrix product states and matrix product operators

It is often convenient to describe states with one-dimensional structure using the language of matrix product states (MPS).

Definition 1 (Matrix product state). *A matrix product state (MPS) $|\eta\rangle$ on N sites of local dimension d is specified by Nd matrices $A_j^{(i)}$ with $i = 1, \dots, d$ and $j = 1, \dots, N$. The matrices $A_1^{(i)}$ are $1 \times \chi$ matrices and $A_N^{(i)}$ are $\chi \times 1$ matrices, with the rest being $\chi \times \chi$. The state is defined as*

$$|\eta\rangle = \sum_{i_1=1}^d \dots \sum_{i_N=1}^d A_1^{(i_1)} \dots A_N^{(i_N)} |i_1 \dots i_N\rangle \quad (2)$$

The parameter χ is called the bond dimension of the MPS.

The same physical state has many different MPS representations, although one may impose a canonical form [19] to make the representation unique. The bond dimension of the MPS is a measure of the maximum amount of entanglement across any ‘‘cut’’ dividing the state into two contiguous parts. More precisely, if we perform a Schmidt decomposition on a state $|\eta\rangle$ across every possible cut, the maximum number of non-zero Schmidt coefficients (i.e. Schmidt rank) across any of the cuts is equal to the minimum bond dimension we would need to exactly represent $|\eta\rangle$ as an MPS [20]. Thus to show a state has an MPS representation with a certain bond dimension, it suffices to bound the Schmidt rank across all the cuts. This line of reasoning shows that a product state, which has no entanglement, can be written as an MPS with bond dimension 1. Meanwhile, a general state with any amount of entanglement can always be written as an MPS with bond dimension $d^{N/2}$.

A cousin of matrix product states are matrix product operators (MPO).

Definition 2 (Matrix product operator). *A matrix product operator (MPO) σ on N sites of local dimension d is specified by Nd^2 matrices $A_j^{(i)}$ with $i = 1, \dots, d^2$ and $j = 1, \dots, N$. The matrices $A_1^{(i)}$ are $1 \times \chi$ matrices and $A_N^{(i)}$ are $\chi \times 1$ matrices, with the rest being $\chi \times \chi$. The operator is defined as*

$$\sigma = \sum_{i_1=1}^{d^2} \dots \sum_{i_N=1}^{d^2} A_1^{(i_1)} \dots A_N^{(i_N)} \sigma_{i_1} \otimes \dots \otimes \sigma_{i_N} \quad (3)$$

where $\{\sigma_i\}_{i=1}^{d^2}$ is a basis for operators on a single site. The parameter χ is called the bond dimension of the MPO.

However, MPO representations have the issue that specifying a set of matrices $A_j^{(i)}$ does not always lead to an operator σ that is positive semi-definite, which is a requirement for the MPO to correspond to a valid quantum state. Checking positivity of an MPO in general is NP-hard for chains of length N and undecidable for infinite chains [17].

C. Notions of approximation

We are interested in the existence of an MPS that approximates the ground state $|\psi\rangle$. We will have both a global and a local notion of approximation, which we define here. We will employ two different distance measures at different points in our theorems and proofs, the purified distance and the trace distance.

Definition 3 (Purified distance). *If σ and σ' are two normalized states on the same system, then*

$$D(\sigma, \sigma') = \sqrt{1 - \mathcal{F}(\sigma, \sigma')^2} \quad (4)$$

is the purified distance between σ and σ' , where $\mathcal{F}(\sigma, \sigma') = \text{Tr}(\sqrt{\sigma^{1/2}\sigma'\sigma^{1/2}})$ denotes the fidelity between σ and σ' .

Definition 4 (Trace distance). *If σ and σ' are two normalized states on the same system, then*

$$D_1(\sigma, \sigma') = \frac{1}{2} \|\sigma - \sigma'\|_1 = \frac{1}{2} \text{Tr}(|\sigma - \sigma'|) \quad (5)$$

is the trace distance between σ and σ' .

Lemma 1 ([21]).

$$D_1(\sigma, \sigma') \leq D(\sigma, \sigma') \leq \sqrt{2D_1(\sigma, \sigma')} \quad (6)$$

We also note that $D_1(\sigma, \sigma') = D(\sigma, \sigma')$ if σ and σ' are both pure. If the trace distance between ρ and σ is small then we would say σ is a good global approximation to ρ . We are also interested in a notion of distance that is more local.

Definition 5 (k -local purified distance). *If σ and σ' are two normalized states on the same system, then the k -local purified distance between σ and σ' is*

$$D^{(k)}(\sigma, \sigma') = \max_{X:|X|=k} D(\sigma_X, \sigma'_X) \quad (7)$$

where the max is taken over all contiguous regions X consisting of k sites.

Definition 6 (k -local trace distance). *If σ and σ' are two normalized states on the same system, then the k -local trace distance between σ and σ' is*

$$D_1^{(k)}(\sigma, \sigma') := \max_{X:|X|=k} D_1(\sigma_X, \sigma'_X) \quad (8)$$

where the max is taken over all contiguous regions X consisting of k sites.

Note that these quantities lack the property that $0 = D^{(k)}(\sigma, \sigma') = D_1^{(k)}(\sigma, \sigma')$ implies $\sigma = \sigma'$ [35], but they do satisfy the triangle inequality. It is also clear that taking $k = N$ recovers our notion of global distance: $D^{(N)}(\sigma, \sigma') = D(\sigma, \sigma')$ and $D_1^{(N)}(\sigma, \sigma') = D_1(\sigma, \sigma')$.

Definition 7 (Local approximation). *We say a state σ on a chain of N sites is a (k, ϵ) -local approximation to another state σ' if $D_1^{(k)}(\sigma, \sigma') \leq \epsilon$*

As we discuss in the next subsection, previous results show that $|\psi\rangle$ has a good global approximation $|\tilde{\psi}\rangle$ that is an MPS with bond dimension that scales like a polynomial in N . We will be interested in the question of what bond dimension is required when what we seek is merely a good local approximation.

D. Previous results

1. Exponential decay of correlations and area laws

A key fact shown by Hastings [22, 23] about nearest-neighbor 1D Hamiltonians with a non-zero energy gap is that the ground state $|\psi\rangle$ has exponential decay of correlations.

Definition 8 (Exponential decay of correlations). *A pure state $\sigma = |\eta\rangle\langle\eta|$ on a chain of sites is said to have (t_0, ξ) -exponential decay of correlations if for every $t \geq t_0$ and every pair of regions A and C separated by at least t sites*

$$\begin{aligned} & \text{Cor}(A : C)_{|\eta\rangle} \\ & := \max_{\|M\|, \|N\| \leq 1} \text{Tr}((M \otimes N)(\sigma_{AC} - \sigma_A \otimes \sigma_C)) \\ & \leq \exp(-t/\xi). \end{aligned} \quad (9)$$

We call ξ the correlation length of $|\eta\rangle$.

Lemma 2 ([22, 23]). *If $|\psi\rangle$ is the unique ground state of a Hamiltonian $H = \sum_i H_{i, i+1}$ with spectral gap Δ , then $|\psi\rangle$ has (t_0, ξ) -exponential decay of correlations for some $t_0 = O(1)$ and $\xi = O(1/\Delta)$.*

While the exponential decay of correlations holds for lattice models in any spatial dimension, the other results we discuss are only known to hold in one dimension.

For example, in one dimension it has been shown that ground states of gapped Hamiltonians obey an area law, that is, the entanglement entropy of any contiguous region is bounded by a constant times the length of the boundary of that region, which in one dimension is just a constant. This statement was also first proven by Hastings in [24] where it was shown that for any contiguous region X

$$S(\rho_X) \leq \exp(O(\log(d)/\Delta)) \quad (10)$$

which is independent of the number of sites in X , where S denotes the von Neumann entropy $S(\sigma) = -\text{Tr}(\sigma \log \sigma)$. The area law has since been improved [7, 25] to

$$S(\rho_X) \leq \tilde{O}(\log^3(d)/\Delta) \quad (11)$$

where the \tilde{O} signifies a suppression of factors that scale logarithmically with the quantity stated.

It was also discovered that an area law follows merely from the assumption of exponential decay of correlations in one dimension: if a pure state ρ has (t_0, ξ) -exponential decay of correlations, then it satisfies [26]

$$S(\rho_X) \leq t_0 \exp(\tilde{O}(\xi \log(d))) \quad (12)$$

The area law is closely related to the existence of an efficient MPS approximation to the ground state. To make this implication concrete, one needs an area law using the α -Renyi entropy for some value of α with $0 < \alpha < 1$ [18], where the Renyi entropy is given by $S_\alpha(\rho_X) = -\text{Tr}(\log(\rho_X^\alpha))/(1-\alpha)$. An area law for the von Neumann entropy (corresponding to $\alpha = 1$) is not alone sufficient [27]. However, for all of the previously stated area laws, the methods are strong enough to also give the existence of efficient MPS approximations. Hastings' [24] original area law implied the existence of a global ϵ -approximation $|\tilde{\psi}\rangle$ for $|\psi\rangle$ with bond dimension

$$\chi = e^{\tilde{O}(\frac{\log(d)}{\Delta})} \left(\frac{N}{\epsilon}\right)^{O(\frac{\log(d)}{\Delta})} \quad (13)$$

The improved area law in [7, 25] yields a better scaling for χ which is asymptotically sublinear in N

$$\chi = e^{\tilde{O}(\frac{\log^3(d)}{\Delta})} \left(\frac{N}{\epsilon}\right)^{\tilde{O}(\frac{\log(d)}{(\Delta \log(N/\epsilon))^{1/4}})} \quad (14)$$

Finally, the result implied only from exponential decay of correlations [26] is

$$\chi = e^{t_0 e^{\tilde{O}(\xi \log(d))}} \left(\frac{N}{\epsilon}\right)^{\tilde{O}(\xi \log(d))} \quad (15)$$

Crucially, if the local Hilbert space dimension d and the gap Δ (or alternatively, the correlation length ξ) are taken to be constant, then all three results read $\chi = \text{poly}(N, 1/\epsilon)$.

2. Constant-bond-dimension MPO local approximations

The problem of finding matrix product operator representations that capture all the local properties of a state has been studied before. Huang [15] showed the existence of a positive semi-definite MPO ρ^χ with bond dimension

$$\chi = e^{\tilde{O}\left(\frac{\log^3(d)}{\Delta} + \frac{\log(d) \log^{3/4}(1/\epsilon)}{\Delta^{1/4}}\right)} = (1/\epsilon)^{o(1)} \quad (16)$$

that is a $(2, \epsilon)$ -local approximation to the true ground state ρ , where $o(1)$ indicates the quantity approaches 0 as $1/\epsilon \rightarrow \infty$. Crucially, this is independent of the length of the chain N . Additionally, because the Hamiltonian is nearest-neighbor, we have $\text{Tr}(H\rho^\chi) - \text{Tr}(H\rho) \leq (N-1)\epsilon$, i.e., the energy per site (energy density) of the state ρ^χ is within ϵ of the ground state energy density. Huang constructs this MPO explicitly and notes it is a convex combination over pure states which themselves are MPS with bond dimension independent of N . Thus, one of these MPS must have energy density within ϵ of the ground state energy density. However, it is not guaranteed (nor is it likely) that one of these constant-bond-dimension MPS is also a good local approximation to the ground state; thus our result may be viewed as an improvement on this front as we show the existence not only of a low-energy-density constant-bond-dimension MPS, but also one that is a good local approximation to the ground state.

An alternative MPO construction achieving the same task was later given in [16]. In this case, the MPO is a (k, ϵ) -local approximation to the ground state and has bond dimension

$$\chi = (k/\epsilon)e^{\tilde{O}\left(\frac{\log^3(d)}{\Delta} + \frac{\log(d)\log^{3/4}(k/\epsilon^3)}{\Delta^{1/4}}\right)} = (k/\epsilon)^{1+o(1)}. \quad (17)$$

The idea they use is simple. They break the chain into blocks of size l , which is much larger than k . On each block they construct a constant bond dimension MPO that closely approximates the reduced density matrix of the ground state on that block, which is easy since each block has constant length and they must make only a constant number of bond truncations to the exact state. The tensor product of these MPO will be an MPO on the whole chain that is a good approximation on any length- k region that falls within one of the larger length l blocks, but not on a region that crosses the boundary between blocks. To remedy this, they take the mixture of MPO formed by considering all l translations of the boundaries between the blocks. Now as long as l is much larger than k , any region of length k will only span the boundary between blocks for a small fraction of the MPO that make up this mixture, and the MPO will be a good local approximation.

This same idea underlies our proof of Theorem 1, with the complication that we seek a pure state approximation and cannot take a mixture of several MPS. Instead, we combine the translated MPS in superposition, which brings new but manageable challenges.

III. STATEMENT OF RESULTS

A. Existence of local approximation

Theorem 1. *Let $|\psi\rangle$ be a state on a chain of N sites of local dimension d . For any k and ϵ there exists an MPS $|\tilde{\psi}\rangle$ with bond dimension at most χ such that*

(1) $|\tilde{\psi}\rangle$ is a (k, ϵ) -local approximation to $|\psi\rangle$

(2) $\chi = e^{O(k \log(d)/\epsilon)}$

provided that N is larger than some constant $N_0 = O(k^3/\epsilon^3)$ that is independent of $|\psi\rangle$.

If $|\psi\rangle$ has (t_0, ξ) -exponential decay of correlations, then the bound on the bond dimension can be improved to

(2') $\chi = e^{t_0 e^{\tilde{O}(\xi \log(d))}} (k/\epsilon^3)^{O(\xi \log(d))}$

with $N_0 = O(k^2/\epsilon^2) + t_0 \exp(\tilde{O}(\xi \log(d)))$ and if, additionally, $|\psi\rangle$ is the unique ground state of a nearest-neighbor 1D Hamiltonian H with spectral gap Δ , it can be further improved to

(2'') $\chi = (k/\epsilon)e^{\tilde{O}\left(\frac{\log^3(d)}{\Delta} + \frac{\log(d)}{\Delta^{1/4}} \log^{3/4}(k/\epsilon^3)\right)}$

with $N_0 = O(k^2/\epsilon^2) + \tilde{O}(\log(d)/\Delta^{3/4})$. Here χ is asymptotically equivalent to $(k/\epsilon)^{1+o(1)}$ where $o(1)$ indicates that the quantity approaches 0 as $(k/\epsilon) \rightarrow \infty$.

However, the state $|\tilde{\psi}\rangle$ that we construct in the proof of Theorem 1 is long-range correlated and cannot be generated from a constant-depth quantum circuit. Thus, while $|\tilde{\psi}\rangle$ is a good local approximation to the ground state $|\psi\rangle$ of H , it is not the exact ground state of any gapped local 1D Hamiltonian.

Next, we show that it remains possible to approximate the state even when we require the approximation to be produced by a constant-depth quantum circuit; the scaling of the bond dimension is faster in k and $1/\epsilon$, but it is still polynomial.

Theorem 2. *Let $|\psi\rangle$ be a state on a chain of N sites of local dimension d . If $|\psi\rangle$ has (t_0, ξ) -exponential decay of correlations, then, for any k and ϵ , there is an MPS $|\tilde{\psi}\rangle$ with bond dimension at most χ such that*

(1) $|\tilde{\psi}\rangle$ is a (k, ϵ) -local approximation to $|\psi\rangle$

(2) $\chi = e^{t_0 e^{\tilde{O}(\xi \log(d))}} (k/\epsilon^2)^{O(\xi^2 \log^2(d))}$

(3) $|\tilde{\psi}\rangle$ can be prepared from the state $|0\rangle^{\otimes N}$ by a quantum circuit that has depth $\tilde{O}(\chi^2)$ and consists only of unitary gates acting on neighboring pairs of qubits

If, additionally, $|\psi\rangle$ is the unique ground state of a nearest-neighbor 1D Hamiltonian with spectral gap Δ , then the bound on the bond dimension can be improved to

(2') $\chi = e^{\tilde{O}\left(\frac{\log^4(d)}{\Delta^2}\right)} (k/\epsilon^2)^{O\left(\frac{\log(d)}{\Delta}\right)}$

The sort of constant-depth quantum circuit that can generate the state $|\tilde{\psi}\rangle$ in Theorem 2 is shown in Figure 1. Proof summaries as well as full proofs of Theorems 1 and 2 appear in Section IV.

We also note that, unlike Theorem 1, Theorem 2 does not require that the chain be longer than some threshold N_0 ; the statement holds regardless of the chain length, although this should be considered a technical detail and not an essential aspect of the constructions.

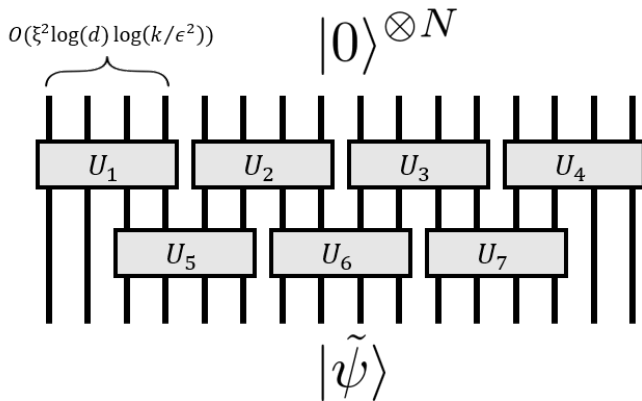


FIG. 1: Constant-depth quantum circuit that constructs the (k, ϵ) -local approximation $|\tilde{\psi}\rangle$ in Theorem 2 starting from the initial state $|0\rangle^{\otimes N}$. It is drawn here as a depth-2 circuit where unitaries U_j act on segments consisting of $O(\xi^2 \log(d) \log(k/\epsilon^2))$ contiguous sites. Each of these unitaries could themselves be decomposed into a sequence of nearest-neighbor gates with depth $\text{poly}(k, 1/\epsilon)$.

B. Reduction from estimating energy density to finding local properties

The previously stated results show that there exists a state that is both a (k, ϵ) -local approximation and an MPS with bond dimension $\text{poly}(k, 1/\epsilon)$. They say nothing of the algorithmic complexity required to *find* a (k, ϵ) -local approximation. The proofs describe how to construct the local approximation from a description of the exact ground state, but following this strategy would require first finding a description of the exact ground state (or perhaps a global approximation to it). One might hope that a different strategy would allow the local approximation to be found much more quickly than the global approximation, since the bond dimension needed to represent the approximation is much smaller. However, the following result challenges the validity of this intuition, at least in the case that the Hamiltonian is translationally invariant, by showing a relationship between the problem of finding a local approximation and the problem of estimating the energy density.

Problem 1 (Estimating energy density). *Given a nearest-neighbor translationally invariant 1D Hamiltonian H on $N \geq 2$ sites and error parameter ϵ , produce an estimate \tilde{u} such that $|u - \tilde{u}| \leq \epsilon$ where $u = E/(N-1)$ is the ground state energy density.*

Problem 2 (Approximating local properties). *Given a nearest-neighbor translationally invariant Hamiltonian H , an error parameter δ , and an operator O whose support is contained within a contiguous region of length k , produce an estimate \tilde{v} such that $|v - \tilde{v}| \leq \delta$, where $v = \langle \psi | O | \psi \rangle / \|O\|$ is the expectation value of the operator $O/\|O\|$ in the ground state $|\psi\rangle$ of H .*

Problem 1 is the restriction of Problem 2 to the case where $k = 2$ and the operator O is the energy interaction term. Thus, there is a trivial reduction from Problem 1 to Problem 2 with $\delta = \epsilon$. However, the next theorem, whose proof is presented in Section IV, states a much more powerful reduction.

Theorem 3. *Suppose one has an algorithm that solves Problem 2 for any single-site ($k = 1$) operator O and $\delta = 0.9$ in $f(\Delta, d, N)$ time, under the promise that the Hamiltonian H has spectral gap at least Δ . Here d denotes the local dimension of H and N the length of the chain. Then there is an algorithm for Problem 1 under the same promise that runs in time*

$$f\left(\frac{\min(2\Delta, (N-1)\epsilon, 2)}{12}, 2d, N\right) O(\log(1/\epsilon)). \quad (18)$$

Estimating the energy density to precision ϵ is equivalent to measuring the total energy to precision $\epsilon(N-1)$, so the quantity $\min(2\Delta, (N-1)\epsilon, 2)$ is equivalent to the global energy resolution, twice the gap, or two, whichever is smallest. Thus, one may take $\epsilon = O(1/N)$ and understand the theorem as stating that finding local properties to within $O(1)$ precision can be done at most $O(\log(N))$ faster than finding an estimate to the total ground state energy to $O(1)$ precision. If local properties can be found in time independent of N (i.e. there is an N -independent upper bound to f), then the ground state energy can be estimated to $O(1)$ precision in time $O(\log(N))$, which would be optimal since the ground state energy scales extensively with N , and $\Omega(\log(N))$ time would be needed simply to write down the output.

Another way of understanding the significance of the theorem is in the thermodynamic limit. Here it states that if one could estimate expectation values of local observables in the thermodynamic limit to $O(1)$ precision in some finite amount of time (for constant Δ and d), then one could compute the ground state energy density of such Hamiltonians to precision ϵ in $O(\log(1/\epsilon))$ time. This would be an exponential speedup over the best-known algorithm for computing the energy density given in [15], which has runtime $\text{poly}(1/\epsilon)$. Taking the contrapositive, if one could show that $\text{poly}(1/\epsilon)$ time is necessary for computing the energy density, this would imply that Problem 2 with $\delta = O(1)$ is in general uncomputable in the thermodynamic limit, even given the promise that the input Hamiltonian is gapped. It is already known that Problem 2 is uncomputable when there is no such promise [28].

It is not clear whether a $O(\log(1/\epsilon))$ time algorithm for computing the energy density is possible. The $\text{poly}(1/\epsilon)$ algorithm in [15] works even when the Hamiltonian is not translationally invariant, but it is not immediately apparent to us how one might exploit translational invariance to yield an exponential speedup.

IV. PROOFS

A. Important lemmas for Theorems 1 and 2

The pair of lemmas stated here are utilized in both Theorem 1 and 2. The first lemma captures the essence of the area laws stated previously, and will be essential when we want to bound the error incurred by truncating a state along a certain cut.

Lemma 3 (Area laws [26], [25]). *If $\sigma = |\psi\rangle\langle\psi|$ has (t_0, ξ) -exponential decay of correlations then for any χ and any region of the chain A , there is a state $\tilde{\sigma}_A$ with rank at most χ such that*

$$D(\sigma_A, \tilde{\sigma}_A) \leq C_1 \exp\left(-\frac{\log(\chi)}{8\xi \log(d)}\right) \quad (19)$$

where $C_1 = \exp(t_0 \exp(\tilde{O}(\xi \log(d))))$ is a constant independent of N .

If σ is the unique ground state of a nearest-neighbor Hamiltonian with spectral gap Δ , then this can be improved to

$$D(\sigma_A, \tilde{\sigma}_A) \leq C_2 \exp\left(-\tilde{O}\left(\frac{\Delta^{1/3} \log^{4/3}(\chi)}{\log^{4/3}(d)}\right)\right) \quad (20)$$

where $C_2 = \exp(\tilde{O}(\log^{8/3}(d)/\Delta))$.

Proof. The first part follows from the main theorem of [26]. The second follows from the 1D area law presented in [25], and $\log(d)$ dependence explicitly stated in [7]. \square

In both proofs we will also utilize the well-known fact that Schmidt ranks cannot differ by more than a factor of d between any two neighboring cuts on the chain.

Lemma 4. *Any state σ_{AB} on a bipartite system AB satisfies relations*

$$\text{rank}(\sigma_{AB})\text{rank}(\sigma_B) \geq \text{rank}(\sigma_A) \quad (21)$$

$$\text{rank}(\sigma_{AB})\text{rank}(\sigma_A) \geq \text{rank}(\sigma_B) \quad (22)$$

$$\text{rank}(\sigma_A)\text{rank}(\sigma_B) \geq \text{rank}(\sigma_{AB}) \quad (23)$$

Proof. We can purify σ_{AB} with an auxiliary system C into the state $|\eta\rangle$. We can let $\sigma = |\eta\rangle\langle\eta|$ and note that $\text{rank}(\sigma_{AB}) = \text{rank}(\sigma_C)$. Thus each of these three equations say the same thing with permutations of A , B , and C . We'll show the first equation. Write Schmidt decomposition

$$|\eta\rangle = \sum_{j=1}^{\text{rank}(\sigma_{AB})} \lambda_j |\nu_j\rangle_{AB} \otimes |\omega_j\rangle_C \quad (24)$$

and then decompose $|\nu_j\rangle$ to find

$$|\eta\rangle = \sum_{j=1}^{\text{rank}(\sigma_{AB})} \sum_{k=1}^{\text{rank}(\sigma_B)} \lambda_j \gamma_{jk} |\tau_{jk}\rangle_A \otimes |\mu_k\rangle_B \otimes |\omega_j\rangle_C \quad (25)$$

where $\{|\mu_k\rangle\}_{k=1}^{\text{rank}(\sigma_B)}$ are the eigenvectors of σ_B . This shows that the support of σ_A is spanned by the set of $|\tau_{jk}\rangle$ and thus its rank can be at most $\text{rank}(\sigma_{AB})\text{rank}(\sigma_B)$. \square

Corollary 1. *If $|\eta\rangle$ is a state on a chain of N sites with local dimension d , and the Schmidt rank of $|\eta\rangle$ across the cut between sites m and $m+1$ is χ , then the Schmidt rank of $|\eta\rangle$ across the cut between sites m' and $m'+1$ is at most $\chi d^{|m'-m|}$.*

Proof. Without loss of generality, assume $m \leq m'$. The reduced density matrix of $|\eta\rangle\langle\eta|$ on sites $[m+1, m']$ has rank at most $d^{|m'-m|}$ since this is the dimension of the entire Hilbert space on that subsystem. Meanwhile the rank of the reduced density matrix on sites $[1, m]$ is χ . So by the previous lemma, the rank over sites $[1, m']$ is at most $\chi d^{|m'-m|}$. \square

B. Proof of Theorem 1

First we state and prove a lemma that will be essential for showing the first part of Theorem 1. Then we provide a proof summary of Theorem 1, followed by its full proof.

Lemma 5. *Given two quantum systems A and B and states τ_A on A and τ_B on B , there exists a state σ_{AB} on the joint system AB such that $\sigma_A = \tau_A$, $\sigma_B = \tau_B$, and $\text{rank}(\sigma_{AB}) \leq \max(\text{rank}(\tau_A), \text{rank}(\tau_B))$*

Proof. We'll apply an iterative procedure. For round 1 let $\alpha_1 = \tau_A$ and $\beta_1 = \tau_B$. In round j write spectral decomposition

$$\alpha_j = \sum_{i=1}^{a_j} \lambda_{j,i} |s_{j,i}\rangle \langle s_{j,i}|_A \quad (26)$$

$$\beta_j = \sum_{i=1}^{b_j} \mu_{j,i} |r_{j,i}\rangle \langle r_{j,i}|_B \quad (27)$$

where a_j and b_j are the ranks of states α_j and β_j , eigenvectors $\{|s_{j,i}\rangle_{i=1}^{a_j}\}$ and $\{|r_{j,i}\rangle_{i=1}^{b_j}\}$ form orthonormal bases of the Hilbert spaces of systems A and B , respectively, and eigenvalues $\{\lambda_{j,i}\}_{i=1}^{a_j}$ and $\{\mu_{j,i}\}_{i=1}^{b_j}$ are non-decreasing with increasing index i (i.e. smallest eigenvalues first). Then define

$$|u_j\rangle = \sum_{i=1}^{\min(a_j, b_j)} \sqrt{\min(\lambda_{j,i}, \mu_{j,i})} |s_{j,i}\rangle_A \otimes |r_{j,i}\rangle_B \quad (28)$$

which may not be a normalized state. Define recursion relation

$$\begin{aligned} \alpha_{j+1} &= \alpha_j - \text{Tr}_B(|u_j\rangle \langle u_j|) \\ \beta_{j+1} &= \beta_j - \text{Tr}_A(|u_j\rangle \langle u_j|) \end{aligned} \quad (29)$$

and repeat until round m when $\alpha_{m+1} = \beta_{m+1} = 0$. Let

$$\sigma_{AB} = \sum_{j=1}^m |u_j\rangle \langle u_j| \quad (30)$$

Clearly $\text{rank}(\sigma_{AB}) \leq m$. We claim that $m \leq \max(\text{rank}(\tau_A), \text{rank}(\tau_B))$. To show this we note the following

$$a_{j+1} + b_{j+1} \leq \max(a_j, b_j). \quad (31)$$

We can see this is true by inspecting the i th term in the Schmidt decomposition of $|u_j\rangle$ in Eq. (28), and noting that either its reduced density matrix on system A is $\lambda_{j,i} |s_{j,i}\rangle \langle s_{j,i}|$ or its reduced density matrix on system B is $\mu_{j,i} |r_{j,i}\rangle \langle r_{j,i}|$ (or both). So when the reduced density matrices of $|u_j\rangle \langle u_j|$ are subtracted from α_j and β_j to form α_{j+1} and β_{j+1} in Eq. (29), each of the $\min(a_j, b_j)$ terms causes the combined rank $a_{j+1} + b_{j+1}$ to decrease by at least one in comparison to $a_j + b_j$.

This alone implies that $m \leq \max(a_1, b_1) + 1$, since by Eq. (31), the sequence $\{a_j + b_j\}_j$ must decrease by at least $\min(a_1, b_1)$ after the first round, and then by at least 1 in every other round, reaching 0 when $j = \max(a_1, b_1) + 1$. However, we can also see that the last round must see a decrease by at least 2, because it is impossible for $a_m = 0$ and $b_m = 1$ or vice versa (since $\text{Tr}(\alpha_j)$ must equal $\text{Tr}(\beta_j)$ for all j). Thus $m \leq \max(a_1, b_1)$.

Moreover, Eqs. (29) and (30) imply that $\sigma_A = \alpha_1 = \tau_A$ and $\sigma_B = \beta_1 = \tau_B$. \square

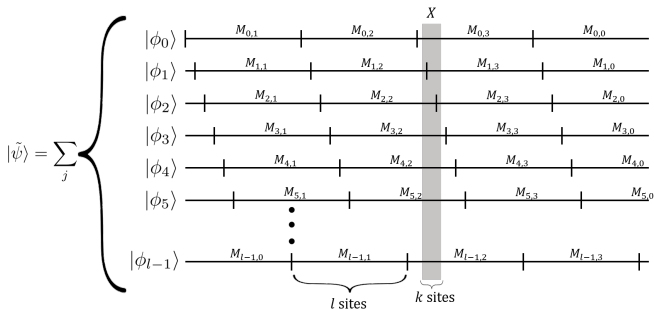


FIG. 2: Basic schematic overview of the proof of Theorem 1. Many states $|\phi_j\rangle$ are constructed with staggered divisions between regions $M_{j,i}$ of length l , then the $|\phi_j\rangle$ are summed in superposition. Properties supported within a length- k region X are faithfully captured by $|\phi_j\rangle$ for values of j such that X does not overlap the boundaries between regions $M_{j,i}$. Most values of j qualify under this criteria as long as l is much larger than k . Additional structure is defined (the regions $B_{j,j'}$ in Part 1, and $B_{j,i}$ in Part 2) in order to force $\langle \phi_j | \phi_{j'} \rangle = \delta_{jj'}$, but this structure is not reflected on the schematic.

Proof summary of Theorem 1. In Ref. [16], an MPO that is a (k, ϵ) -local approximation to a given state $|\psi\rangle$ was formed by dividing the chain into many length- l segments, tensoring together a low-bond-dimension approximation of the exact state reduced to each segment, and then summing over (as a mixture) translations of locations for the divisions between the segments. We follow the same idea but for pure states: for each integer $j = 0, \dots, l-1$, we divide the state into many length- l segments and create a pure state approximation $|\phi_j\rangle$

that captures any local properties that are supported entirely within one of the segments. Then to form $|\tilde{\psi}\rangle$, we sum in superposition over all the $|\phi_j\rangle$, where each $|\phi_j\rangle$ has boundaries between segments occurring in different places (see Figure 2). Thus, for any length- k region X , a large fraction of the terms $|\phi_j\rangle$ in the superposition are individually good approximations on region X . The fact that a small fraction of the terms are not necessarily a good approximation creates additional, but small, error in the local approximation.

In order to avoid interaction between different terms in the superposition, we add additional structure to make the $|\phi_j\rangle$ in the superposition exactly orthogonal to one another. In our construction for general states, this additional structure consists of a set of disjoint, sparsely distributed, single-site regions $B_{j,j'}$, one for each pair of integers $j \neq j'$ with $0 \leq j, j' < l$. We force $|\phi_j\rangle$ to be the pure state $|0\rangle$ and $|\phi_{j'}\rangle$ to be $|1\rangle$ when reduced to $B_{j,j'}$ to guarantee that $\langle \phi_j | \phi_{j'} \rangle = 0$. Our construction for states that have exponential decay of correlations is similar: we define a series of regions $B_{j,i}$ and for each pair (j, j') force $|\phi_j\rangle$ and $|\phi_{j'}\rangle$ to have orthogonal supports when reduced to one of these regions.

Our approach for constructing $|\phi_j\rangle$ differs if $|\psi\rangle$ is a general state (Part 1), or if it is a state that either has exponentially decaying correlations or (additionally) is the ground state of a nearest-neighbor Hamiltonian (Part 2). In the latter case, we examine each length- l segment individually and truncate the bonds of the exact state on all but a few of the rightmost sites within that segment. The area law implies these truncations have minimal effect. We use those few rightmost sites to purify the mixed state on the rest of the segment. Then $|\phi_j\rangle$ is a tensor product over pure states on each of the segments. The bond dimension can be bounded within each segment of $|\phi_j\rangle$ which is sufficient to bound the bond dimension of $|\psi\rangle$.

This does not work for general states because without an area law, bond truncations result in too much error, and without the truncations we do not have enough room to purify the state. For a general state, there is simply too much entropy in the length- l segment to fully absorb with only a few sites at the edge of the segment. Instead, we have the various segments absorb each other's entropy by developing a procedure to engineer entanglement between different length- l segments that exactly preserves the reduced density matrix on each segment and keeps the Schmidt rank constant (albeit exponential in k/ϵ) across any cut. The crux of this procedure is captured in Lemma 5. \square

Proof of Theorem 1. Part 1: Proof of items (1) and (2). First we consider the case of a general state $|\psi\rangle$ and construct an approximation $|\tilde{\psi}\rangle$ satisfying items (1) and (2). A different construction is given in Part 2 to show items (2') and (2''), but it is similar in approach. Throughout this proof, we use a convention where sites are numbered 0 to $N-1$, which differs from the rest of the paper.

We choose integer l to be specified later. We require $l > k$. To construct $|\tilde{\psi}\rangle$, we will first construct states $|\phi_j\rangle$ for $j = 0, \dots, l-1$ and then sum these states in superposition. Any reference to the index j will be taken mod l . Consider a fixed value of j . To construct $|\phi_j\rangle$ we break the chain into $n = \lfloor N/l \rfloor$ blocks where $n-1$ blocks $\{M_{j,i}\}_{i=1}^{n-1}$ have length exactly l and the final block $M_{j,0}$ has length at least l and less than $2l$. We arrange the blocks so that the leftmost site of block $M_{j,1}$ is site j ; thus block $M_{j,i}$ contains the sites $[j+l(i-1), j+li-1]$ for $i = 1, \dots, n-1$, and the last block $M_{j,0}$ “rolls over” to include sites at both ends of the chain: sites $[j+l(n-1), N-1]$ and $[0, j-1]$. A schematic of the arrangement is shown in Figure 2. Any reference to the index i will be taken mod n .

We also define $l^2 - l$ single-site blocks on the chain which we label $B_{j,j'}$ for all pairs $j \neq j'$. $B_{j,j'}$ consists of the site with index $3l^2(j+j') + l(j-j') + 2l^2$. This definition is possible as long as there are sufficiently many sites: $N \geq N_0 = O(l^3)$. It can be verified that since $0 \leq j, j' \leq l-1$, the distance between any $B_{j,j'}$ and $B_{j'',j'''}$ for any distinct pairs (j, j') and (j'', j''') is at least l . For each j , let $B_j = \cup_{j' \neq j} B_{j,j'} \cup B_{j',j}$. For each j, i , let $A_{j,i} = M_{j,i} \setminus M_{j,i} \cap B_j$. Thus, in most cases $A_{j,i} = M_{j,i}$ since B_j is a relatively small set of sites. Let $A_j = \cup_{i=0}^{n-1} A_{j,i} = B_j^c$, the complement of B_j .

The state $|\phi_j\rangle$ will have the following form

$$|\phi_j\rangle = |Q_j\rangle_{A_j} \otimes \bigotimes_{j' \neq j} (|0\rangle_{B_{j,j'}} \otimes |1\rangle_{B_{j',j}}) \quad (32)$$

where $|0\rangle$ and $|1\rangle$ are two of the d computational basis states located on a single site. In other words $|\phi_j\rangle$ is a product state over all single site regions $B_{j,j'}$ with some other (yet to be specified) state $|Q_j\rangle$ on the remainder of the chain A_j .

To construct $|Q_j\rangle$, we apply Lemma 5 iteratively as follows. We let $\sigma_1 = \rho_{A_{j,j+2}}$ (the reduced matrix of the exact state ρ on region $A_{j,j+2}$). We combine σ_1 with $\rho_{A_{j,j+3}}$ using Lemma 5 to form a state σ_2 on region $A_{j,j+2}A_{j,j+3}$ such that $\text{rank}(\sigma_2) \leq \max(\text{rank}(\sigma_1), \text{rank}(\rho_{A_{j,j+3}}))$. For any j, i , the rank of the state on any region $A_{j,i}$ is less than d^{2l} since any region contains at most $2l$ sites. So if we apply this process iteratively, forming σ_{p+1} by combining σ_p and the state on region $A_{j,j+p+2}$ ($j+p+2$ is taken mod n), then we end up with a state σ_{n-2} with rank at most d^{2l} defined over all of A_j except $A_{j,j}$ and $A_{j,j+1}$. Since by construction B_j contains no sites with index smaller than $2l^2$ and $A_{j,j}$ and $A_{j,j+1}$ are contained within the first $2l^2$ sites, we have $A_{j,j} = M_{j,j}$ and $A_{j,j+1} = M_{j,j+1}$ meaning each of these two regions each contain l sites and the total dimension of the Hilbert space over $A_{j,j}A_{j,j+1}$ is at least d^{2l} . Thus we may use regions $A_{j,j}$ and $A_{j,j+1}$ to purify the state σ_{n-2} . We let $|Q_j\rangle$ be any such purification.

The key observation is that the state $|\phi_j\rangle$, as defined by Eq. (32), will get any local properties exactly correct as long as they are supported entirely within a segment

$A_{j,i}$ for some $i \neq j, j+1$. As long as l is large, this will be the case for most length- k regions, but it will not be the case for some regions that cross the boundaries between regions $A_{j,i}$ or for regions that contain one of the single site regions $B_{j,j'}$ or $B_{j',j}$.

To fix this we sum in superposition over the states $|\phi_j\rangle$ for each value of j . The motivation to do this is so that every length- k region will be contained within $A_{j,i}$ for some value of i in most, but not all, of the terms in the superposition. We will show that most is good enough. We let

$$|\tilde{\psi}\rangle = \frac{1}{\sqrt{l}} \sum_{j=0}^{l-1} |\phi_j\rangle \quad (33)$$

We note that $\langle \phi_j | \phi_{j'} \rangle = \delta_{jj'}$ since $|\phi_j\rangle$ is simply $|0\rangle$ when reduced to region $B_{j,j'}$ and $|1\rangle$ when reduced to region $B_{j',j}$, while $|\phi_{j'}\rangle$ is $|1\rangle$ when reduced to region $B_{j,j'}$ and $|0\rangle$ when reduced to $B_{j',j}$. Thus $|\tilde{\psi}\rangle$ is normalized:

$$\langle \tilde{\psi} | \tilde{\psi} \rangle = \frac{1}{l} \sum_{j,j'} \langle \phi_j | \phi_{j'} \rangle = 1 \quad (34)$$

This completes the construction of the approximation. We now wish to show it has the desired properties. To show item (1), we compute the (local) distance from $|\psi\rangle$ to $|\tilde{\psi}\rangle$. Let $\tilde{\rho} = |\tilde{\psi}\rangle\langle\tilde{\psi}|$ and consider an arbitrary length- k region X .

$$\begin{aligned} D_1(\rho_X, \tilde{\rho}_X) &= \frac{1}{2} \left\| \left(\frac{1}{l} \sum_{j=0}^{l-1} \sum_{j'=0}^{l-1} \text{Tr}_{X^c}(|\phi_j\rangle\langle\phi_{j'}|) \right) - \text{Tr}_{X^c}(|\psi\rangle\langle\psi|) \right\|_1 \end{aligned}$$

First we examine terms in the sum for which $j \neq j'$. Since $B_{j,j'}$ and $B_{j',j}$ are separated by at least l sites and $l > k$, X^c must include either $B_{j,j'}$ or $B_{j',j}$ (or both). Since $|\phi_j\rangle$ and $|\phi_{j'}\rangle$ have orthogonal support on both those regions, and at least one of them is traced out, the term vanishes.

Thus we have

$$\begin{aligned} D_1(\rho_X, \tilde{\rho}_X) &= \frac{1}{2} \left\| \frac{1}{l} \sum_{j=0}^{l-1} \text{Tr}_{X^c}(|\phi_j\rangle\langle\phi_j| - |\psi\rangle\langle\psi|) \right\|_1 \\ &\leq \frac{1}{l} \sum_{j=0}^{l-1} D_1(\rho_X, \text{Tr}_{X^c}(|\phi_j\rangle\langle\phi_j|)) \quad (35) \end{aligned}$$

For a particular j , there are two cases. Case 1 includes values of j for which X falls completely within the $A_{j,i}$ for some i with $i \neq j, j+1$. For these values of j the term vanishes because the reduced density matrix of $|\phi_j\rangle\langle\phi_j|$ on X is exactly ρ_X . Case 2 includes all other values of j . For this to be the case, either X spans the boundary between two regions $M_{j,i}$ and $M_{j,i+1}$ (at most $k-1$ different values of j), X contains a site $B_{j,j'}$ or $B_{j',j}$ for some j' (at most 2 values of j , since the separation between sites $B_{j,j'}$ implies only one may lie within X), or

X is contained within $A_{j,j}$ or $A_{j,j+1}$ (at most 2 values of j). In this case, the term will not necessarily be close to zero, but we can always upper bound the trace distance by 1. The number of terms in the sum that qualify as Case 2 is therefore at most $k+3$, and the total error can be bounded:

$$D_1(\rho_X, \tilde{\rho}_X) \leq \frac{k+3}{l} \quad (36)$$

Choosing $l = (k+3)/\epsilon$ shows item (1), that $|\tilde{\psi}\rangle$ is a (k, ϵ) -local approximation to $|\psi\rangle$.

To show item (2), we bound the Schmidt rank of the state $|\tilde{\psi}\rangle$ across every cut that bipartitions the chain into two contiguous regions. Since $|\tilde{\psi}\rangle$ is a superposition over l terms $|\phi_j\rangle$, the Schmidt rank can be at most l times greater than that of an individual $|\phi_j\rangle$. Fix some value of j and some cut of the chain at site s . Since $|\phi_j\rangle$ is a product state between regions A_j and B_j , and moreover it is a product state on each individual site in B_j , we may ignore B_j when calculating the Schmidt rank (it has no entanglement), and focus merely on $|Q_j\rangle_{A_j}$. We constructed the state $|Q_j\rangle$ by building up mixed states σ_p on region $A_{j,j+2} \dots A_{j,j+p+1}$ until $p = n-2$, then purified with the remaining two regions. Each σ_p has $\text{rank}(\sigma_p) \leq d^{2l}$. Now consider an integer b with $1 \leq b \leq n-1$ and $b \neq j, j+1$. Denote $\sigma = \sigma_{n-2}$ and note that

$$\begin{aligned} & \text{rank}(\sigma_{A_{j,1} \dots A_{j,b}}) \\ & \leq \text{rank}(\sigma_{A_{j,j+2} \dots A_{j,b}}) \text{rank}(\sigma_{A_{j,j+2} \dots A_{j,0}}) \\ & = \text{rank}(\sigma_{b-j-1}) \text{rank}(\sigma_{n-j-1}) \leq d^{4l} \end{aligned} \quad (37)$$

where the first inequality follows from Lemma 4.

Moreover, we may choose b such that the cut between $A_{j,b}$ and $A_{j,b+1}$ falls within l sites of site s . We find that the region $A_{j,1} \dots A_{j,b}$ differs from the region containing sites $[0, s]$ by at most l sites at each edge. Thus the Schmidt rank on the region left of the cut can be at most d^{2l} larger than that of $A_{j,1} \dots A_{j,b}$ (Corollary 1), giving a bound of d^{6l} for the the Schmidt rank of $|\phi_j\rangle$. This implies the Schmidt rank of $|\tilde{\psi}\rangle$ is at most ld^{6l} , which proves item (2). This applies whenever $N \geq N_0 = O(l^3) = O(k^3/\epsilon^3)$, a bound which must be satisfied in order for the chain to be long enough to fit all the regions $B_{j,j'}$ as defined above. This completes Part 1.

Part 2: Proof of items (2') and (2'')

This construction is mostly similar to the previous one with a few key differences. We choose integers l, t , and χ' to be specified later. We require t be even and $l \geq 2k$, $l \geq 2t$. We assume $N \geq N_0 = 2l^2$. We also require that $d \geq 4$. If this is not the case, we coarse-grain the system by combining neighboring sites, and henceforth we assume $d \geq 4$.

As in Part 1, to construct $|\tilde{\psi}\rangle$, we will first construct states $|\phi_j\rangle$ for $j = 0, \dots, l-1$ and then sum these states in

superposition. Consider a fixed value of j . To construct $|\phi_j\rangle$ we break the chain into $n = \lfloor N/l \rfloor$ blocks and we arrange them exactly as in Part 1.

Now the construction diverges from Part 1: the state $|\phi_j\rangle$ will be a product state over each of these blocks

$$|\phi_j\rangle = |\phi_{j,0}\rangle_{M_{j,0}} \otimes \dots \otimes |\phi_{j,n-1}\rangle_{M_{j,n-1}} \quad (38)$$

with states $|\phi_{j,i}\rangle$ for $i = 0, \dots, n-1$ that we now specify.

The idea is to create a state $|\phi_{j,i}\rangle$ that has nearly the same reduced density matrix as $|\psi\rangle$ on the leftmost $l-t$ sites of region $M_{j,i}$. It uses the rightmost t sites to purify the reduced density matrix on the leftmost $l-t$ sites. First we denote the leftmost $l-t$ sites of block $M_{j,i}$ by $A_{j,i} = [j+l(i-1), j+li-t-1]$ and the rightmost t sites by $B_{j,i} = [j+li-t, j+li-1]$ (or appropriate ‘‘roll over’’ definitions when $i = 0$). We write $|\psi\rangle$ as an exact MPS with exponential bond dimension and form $|\psi_{j,i}\rangle$ by truncating to bond dimension χ' (i.e. projecting onto the span of the right Schmidt vectors associated with the largest χ' Schmidt coefficients, then normalizing the state) at the cut to the left of region $A_{j,i}$, every cut within $A_{j,i}$, and at the cut to the right of $A_{j,i}$, for a total of at most $2l-t$ truncations (recall the final region $M_{j,0}$ may have as many as $2l-1$ sites). We denote the pure density matrix of this state by $\rho^{(j,i)} = |\psi_{j,i}\rangle \langle \psi_{j,i}|$. We can bound the effect of these truncations using the area law given by Lemma 3:

$$D(\rho^{(j,i)}, \rho) \leq \sqrt{2l-t} \epsilon^{\chi'} \quad (39)$$

where $\epsilon^{\chi'}$ is the cost (in purified distance) of a single truncation, given by the right hand side of Eq. (19), or by Eq. (20) in the case $|\psi\rangle$ is the ground state of a gapped nearest-neighbor 1D Hamiltonian. These truncations were not possible in Part 1 because we could not invoke the area law for general states.

Because of the truncations, we can express the reduced density matrix $\rho_{A_{j,i}}^{(j,i)}$ as a mixture of χ'^2 pure states $\{|\phi_{j,i,z}\rangle\}_{z=0}^{\chi'^2-1}$ each of which can be written as an MPS with bond dimension χ' :

$$\rho_{A_{j,i}}^{(j,i)} = \sum_{z=0}^{\chi'^2-1} p_{j,i,z} |\phi_{j,i,z}\rangle \langle \phi_{j,i,z}| \quad (40)$$

for some probability distribution $\{p_{j,i,z}\}_{z=0}^{\chi'^2-1}$. We now form $|\phi_{j,i}\rangle$ by purifying $\rho_{A_{j,i}}^{(j,i)}$ onto the region $M_{j,i}$ using the space $B_{j,i}$, which contains t sites, as the purifying subspace:

$$|\phi_{j,i}\rangle_{M_{j,i}} = \sum_{z=0}^{\chi'^2-1} \sqrt{p_{j,i,z}} |\phi_{j,i,z}\rangle_{A_{j,i}} \otimes |r_{j,i,z}\rangle_{B_{j,i}} \quad (41)$$

where the set of states $\{|r_{j,i,z}\rangle\}_{z=0}^{\chi'^2-1}$ is an orthonormal set defined on region $B_{j,i}$. This purification will only be possible if the dimension of $B_{j,i}$ is sufficiently large, and

we comment later on this fact, as well as how exactly to choose the set $\{|r_{j,i,z}\}_{z=0}^{\chi'^2-1}$.

The key observation is that state $|\phi_j\rangle$ will get any local properties approximately correct as long as they are supported entirely within a segment $A_{j,i}$ for some i , and as long as χ' is large enough that the $2l-t$ truncations do not have much effect on the reduced density matrix there. Thus we will choose our parameters so that χ' is large (but independent of N), such that l is much larger than t and k (so that most regions fall within a region $A_{j,i}$), and such that t is large enough that it is possible to purify states on $A_{j,i}$ onto $M_{j,i}$. But, as in Part 1, we have the issue that some regions will not be contained entirely within region $A_{j,i}$ for some i .

We again deal with this issue by summing in superposition:

$$|\tilde{\psi}\rangle = \frac{1}{\sqrt{l}} \sum_{j=0}^{l-1} |\phi_j\rangle \quad (42)$$

To complete the construction we also must specify the orthonormal states $\{|r_{j,i,z}\}_{z=0}^{\chi'^2-1}$ defined on the t sites in region $B_{j,i}$. We choose a set that satisfies the following requirements:

- (1) The reduced density matrix of $|r_{j,i,z}\rangle$ on any single site among the leftmost $t/2$ sites of $B_{j,i}$ (recall we have assumed t is even) is entirely supported on basis states $1, \dots, \lfloor d/2 \rfloor$
- (2) The reduced density matrix on any single site among the rightmost $t/2$ sites is entirely supported on basis states $\lfloor d/2 \rfloor + 1, \dots, d$
- (3) Let $j' = j + i \pmod{l}$, and let $i_0 = i \pmod{l}$. If $t \leq i_0 \leq l - t$ then for all z , $|r_{j,i,z}\rangle$ is orthogonal to the support of the reduced density matrix of $|\phi_{j'}\rangle$ on region $B_{j,i}$.

We assess how large t must be for it to be possible to satisfy these three conditions. The third item specifically applies only for values of i that lead to values of j' that are at least t away from j (modulo l) so that the purifying system $B_{j,i}$ does not overlap with $B_{j',i'}$ for any i' . The support of the reduced density matrix of any $|\phi_{j'}\rangle$ on region $B_{j,i}$ has dimension at most χ'^2 . Thus, if the dimension of $B_{j,i}$ is more than $2\chi'^2$ it will always be possible to choose an orthonormal set $\{|r_{j,i,z}\}_{z=0}^{\chi'^2-1}$ satisfying the third condition. The first and second conditions cut the accessible part of the local dimension of the purifying system in half, so a purification that satisfies all three conditions will be possible if $\lfloor d/2 \rfloor^t \geq 2\chi'^2$. Any choice of set that meets all three conditions is equally good for our purposes.

We now demonstrate that the three conditions imply that for any pair (j, j') there are regions of the chain on which the supports of the reduced density matrices of states $|\phi_j\rangle$ and $|\phi_{j'}\rangle$ are orthogonal. If it is the case that

$j - j' \pmod{l} < t$ or $j' - j \pmod{l} < t$, then for every i the region $B_{j,i}$ overlaps with $B_{j',i'}$ for some i' . Because $B_{j,i} \neq B_{j',i'}$, there will be some site that is in the right half of one of the two regions, but in the left half of the other, and items 1 and 2 imply that the two states will be orthogonal when reduced to this site. If this is not the case, then as long as there is some i for which $j' = j + i \pmod{l}$, then item 3 implies the orthogonality of the supports of $|\phi_j\rangle$ and $|\phi_{j'}\rangle$. In fact because $n \geq 2l$, there will be at least 2 such values of i . We conclude that $\langle \phi_j | \phi_{j'} \rangle = \delta_{jj'}$, which implies that $|\tilde{\psi}\rangle$ is normalized as shown by the computation

$$\langle \tilde{\psi} | \tilde{\psi} \rangle = \frac{1}{l} \sum_{j,j'} \langle \phi_j | \phi_{j'} \rangle = 1. \quad (43)$$

We have now shown how to define the approximation $|\tilde{\psi}\rangle$ and discussed the conditions for the parameters t , χ' , and d that make the construction possible. Now we assess the error in the approximation (locally). Let $\tilde{\rho} = |\tilde{\psi}\rangle\langle\tilde{\psi}|$ and consider an arbitrary length- k region X

$$\begin{aligned} D_1(\rho_X, \tilde{\rho}_X) &= \frac{1}{2} \left\| \left(\frac{1}{l} \sum_{j=0}^{l-1} \sum_{j'=0}^{l-1} \text{Tr}_{X^c}(|\phi_j\rangle\langle\phi_{j'}|) \right) - \text{Tr}_{X^c}(|\psi\rangle\langle\psi|) \right\|_1 \end{aligned}$$

For the same reason that led to the conclusion $\langle \phi_j | \phi_{j'} \rangle = \delta_{jj'}$, we can conclude that $\text{Tr}_{X^c}(|\phi_j\rangle\langle\phi_{j'}|) = \delta_{jj'}$, so long as k is smaller than $l/2$. To see that this holds, it is sufficient to show that there is a region lying completely outside of X with the property that $|\phi_j\rangle$ and $|\phi_{j'}\rangle$ share no support on the region. Since $k = |X| \leq l/2$ and $|B_{j,i}| = t \leq l/2$, for any j , X can overlap the region $B_{j,i}$ for at most one value of i . We showed before that for any j there would be at least two values of i for which a subregion of $B_{j,i}$ has this property, implying one of them must lie outside X .

Thus we have

$$\begin{aligned} D_1(\rho_X, \tilde{\rho}_X) &= \frac{1}{2} \left\| \frac{1}{l} \sum_{j=0}^{l-1} \text{Tr}_{X^c}(|\phi_j\rangle\langle\phi_j| - |\psi\rangle\langle\psi|) \right\|_1 \\ &\leq \frac{1}{l} \sum_{j=0}^{l-1} D_1(\rho_X, \text{Tr}_{X^c}(|\phi_j\rangle\langle\phi_j|)) \quad (44) \end{aligned}$$

For a particular j , there are two cases. Case 1 occurs if X falls completely within the $A_{j,i}$ for some i , in which case the only error is due to the $2l-t$ truncations to bond dimension χ' . Since the trace norm D_1 is smaller than the purified distance D (Lemma 1), the contribution for these values of j is at most $\sqrt{2l-t}e^{\chi'}$. Case 2 includes values of j for which X does not fall completely within a region $A_{j,i}$ for any i . In this case, the term will not necessarily be close to zero, but we can always upper bound the trace distance by 1. The number of terms in the sum that qualify as Case 1 is of course bounded by l

(there are only l terms) and the number of Case 2 terms is at most $t+k-1$. Thus the total error can be bounded:

$$\begin{aligned} & D_1(\rho_X, \tilde{\rho}_X) \\ & \leq \frac{1}{l}(l\sqrt{2l-t}\epsilon^{\chi'} + (t+k-1)) \\ & \leq \sqrt{2l}\epsilon^{\chi'} + (k+t)/l \end{aligned} \quad (45)$$

We will choose parameters so this quantity is less than ϵ . Parameters l and t will be related to χ' as

$$t = \log(2\chi'^2)/\log(\lfloor d/2 \rfloor) \quad (46)$$

$$l = 2(k+t)/\epsilon \quad (47)$$

If $|\psi\rangle$ is known only to have exponentially decaying correlations, then we choose

$$\log(\chi') = 16\xi \log(d) \log(16C_1 \sqrt{\xi \log(d)(k+3)}/\epsilon^{3/2}) \quad (48)$$

where $C_1 = \exp(t_0 \exp(\tilde{O}(\xi \log(d))))$ is the constant from Eq. (19). We note that $t \leq 3 \log(\chi')$, so we can bound

$$\begin{aligned} D_1(\rho_X, \tilde{\rho}_X) & \leq \frac{\sqrt{4(k+t)}C_1}{\sqrt{\epsilon}} e^{-\frac{\log(\chi')}{8\xi \log(d)}} + \frac{\epsilon}{2} \\ & \leq \frac{\sqrt{4(k+3)}\log(\chi')C_1}{\sqrt{\epsilon}} e^{-\frac{\log(\chi')}{8\xi \log(d)}} + \frac{\epsilon}{2} \\ & \leq \frac{\sqrt{64\xi \log(d)(k+3)}C_1}{\sqrt{\epsilon}} e^{-\frac{\log(\chi')}{16\xi \log(d)}} + \frac{\epsilon}{2} \\ & \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \end{aligned} \quad (49)$$

where in the third line we have used the (crude) bound $\sqrt{u} \leq e^u$ with $u = \log(\chi')/(16\xi \log(d))$.

In the case that $|\psi\rangle$ is known to be the ground state of a gapped local Hamiltonian, we may choose

$$\log(\chi') = \tilde{O}(\Delta^{-1/4} \log(d) \log^{3/4}(C_2 \sqrt{k+3}/\epsilon^{3/2})) \quad (50)$$

where $C_2 = \exp(\tilde{O}(\log^{8/3}(d)/\Delta))$ is the constant in Eq. (20) and the same analysis will follow. This proves item (1) of the theorem for the construction in Part 2.

Items (2') and (2'') assert that $|\tilde{\psi}\rangle$ can be written as an MPS with constant bond dimension, which we now show. Each state $|\phi_j\rangle$ is a product state with pure state $|\phi_{j,i}\rangle$ on each block $M_{j,i}$, and $|\phi_{j,i}\rangle$ has bond dimension at most $2\chi'^2$. Thus, if we cut the state $|\phi_j\rangle$ at a certain site, the bond dimension will be at most $4\chi'^2$ (recall that block $M_{j,0}$ may have sites at both ends of the chain and can contribute to the bond dimension). Since $|\tilde{\psi}\rangle$ is a sum over $|\phi_j\rangle$ for l values of j , the bond dimension χ of $|\tilde{\psi}\rangle$ is at most $4l\chi'^2$. For the case of exponentially decaying correlations, this evaluates to

$$\begin{aligned} \chi & = 4l(256C_1^2 \xi \log(d)(k+3)/\epsilon^3)^{16\xi \log(d)} \\ & \leq e^{t_0 e^{\tilde{O}(\xi \log(d))}} (k/\epsilon^3)^{O(\xi \log(d))} \end{aligned} \quad (51)$$

proving item (2'), and for the case of ground state of gapped Hamiltonian, we find

$$\begin{aligned} \chi & = 4l \exp(\tilde{O}(\Delta^{-1/4} \log(d) \log^{3/4}(C_2 \sqrt{k+3}/\epsilon^{3/2}))) \\ & \leq (k/\epsilon) e^{\tilde{O}\left(\frac{\log^3(d)}{\Delta} + \frac{\log(d)}{\Delta^{1/4}} \log^{3/4}(k/\epsilon^3)\right)} \end{aligned} \quad (52)$$

proving item (2''), where the second factor is asymptotically $(k/\epsilon)^{o(1)}$. For completeness, we note that if we combined neighboring sites because $d < 4$, we can now uncombine them possibly incurring a factor of 2 or 3 increase in the bond dimension, which has no effect on the stated asymptotic forms for χ .

These results hold as long as $N \geq 2l^2$, which translates to $N \geq O(k^2/\epsilon^2) + t_0 \exp(\tilde{O}(\xi \log(d)))$ in the case of exponentially decaying correlations, and $N \geq O(k^2/\epsilon^2) + \tilde{O}(\log(d)/\Delta^{3/4})$ in the case that $|\psi\rangle$ is a ground state of a local Hamiltonian. This completes the proof. \square

Now we demonstrate that the state $|\tilde{\psi}\rangle$ constructed in the proof of Theorem 1, Part 2, is long-range correlated. Given an integer m , consider the pair of regions $A = [0, l(l-1)-1]$ and $C = [l(l-1+m), N-1]$, which are separated by ml sites. Assume $n \geq 2l+m$, so that A and C both contain at least l^2 sites. Let operators

$$\begin{aligned} Q_1 & = |\phi_{0,1}\rangle \langle \phi_{0,1}| \otimes \dots \otimes |\phi_{0,l}\rangle \langle \phi_{0,l}| \\ Q_2 & = |\phi_{0,l+m}\rangle \langle \phi_{0,l+m}| \otimes \dots \\ & \quad \dots \otimes |\phi_{0,n-1}\rangle \langle \phi_{0,n-1}| \otimes |\phi_{0,0}\rangle \langle \phi_{0,0}| \end{aligned} \quad (53)$$

The operator Q_1 is supported on A and Q_2 is supported on C . Since A and C each contain blocks M_i for at least l values of i , conditions (1), (2), and (3) above imply that $Q_1|\tilde{\psi}\rangle = Q_2|\tilde{\psi}\rangle = |\phi_0\rangle/\sqrt{l}$. Thus

$$\begin{aligned} \text{Cor}(A : C)_{|\tilde{\psi}\rangle} & \geq \text{Tr}((Q_1 \otimes Q_2)(\tilde{\rho}_{AC} - \tilde{\rho}_A \otimes \tilde{\rho}_C)) \\ & = \text{Tr}(Q_1 \otimes Q_2 |\tilde{\psi}\rangle \langle \tilde{\psi}|) \\ & \quad - \text{Tr}(Q_1 |\tilde{\psi}\rangle \langle \tilde{\psi}|) \text{Tr}(Q_2 |\tilde{\psi}\rangle \langle \tilde{\psi}|) \\ & = 1/l - 1/l^2 \end{aligned} \quad (54)$$

The choice of l is independent of the chain length N , so the above quantity is independent of N and independent of the parameter m measuring the distance between A and C . Thus, the correlation certainly does not decay exponentially in the separation between the regions.

C. Proof of Theorem 2

In this section, we first state a pair of lemmas that will be essential for the proof of Theorem 2, then we give a proof summary of Theorem 2, and finally we provide the full proof of Theorem 2.

First, an important and well-known tool we use is Uhlmann's theorem [29], which expresses the fact that if two states are close, their purifications will be equally close up to a unitary acting on the purifying auxiliary space.

Lemma 6 (Uhlmann's theorem [29]). *Suppose τ_A and σ_A are states on system A . Suppose B is an auxiliary system and $|T\rangle_{AB}$ and $|S\rangle_{AB}$ are purifications of τ_A and σ_A , respectively. Then*

$$D(\tau_A, \sigma_A) = \min_U \sqrt{1 - |\langle S|_{AB} (I_A \otimes U) |T\rangle_{AB}|^2} \quad (55)$$

where the min is taken over unitaries on system B .

Second, we prove the following essential statement about states with exponential decay of correlations.

Lemma 7. *If L and R are disjoint regions of a 1D lattice of N sites and the state $\tau = |\eta\rangle\langle\eta|$ has (t_0, ξ) -exponential decay of correlations, then*

$$D(\tau_{LR}, \tau_L \otimes \tau_R) \leq C_3 \exp(-\text{dist}(L, R)/\xi'). \quad (56)$$

whenever $\text{dist}(L, R) \geq t_0$, where $\xi' = 16\xi^2 \log(d)$ and $C_3 = \exp(t_0 \exp(\tilde{O}(\xi \log(d))))$.

If τ is the unique ground state of a gapped nearest-neighbor 1D Hamiltonian with spectral gap Δ , then this can be improved to

$$D(\tau_{LR}, \tau_L \otimes \tau_R) \leq C_4 \exp(-\text{dist}(L, R)/\xi'). \quad (57)$$

whenever $\text{dist}(L, R) \geq \Omega(\log^4(d)/\Delta^2)$, where $\xi' = O(1/\Delta)$ and $C_4 = \exp(\tilde{O}(\log^3(d)/\Delta))$.

For pure states σ , we call σ a Markov chain for the tripartition $L/M/R$ if $\sigma_{LR} = \sigma_L \otimes \sigma_R$. Thus Lemma 7 states that exponential decay of correlations implies that the violation of the Markov condition, as measured by the purified distance (or alternatively, trace distance) decays exponentially with the size of M .

Proof of Lemma 7. The goal is to show that an exponential decay of correlations in $\tau = |\eta\rangle\langle\eta|$ implies that τ_{LR} is close to $\tau_L \otimes \tau_R$. We will do this by truncating the rank of τ on the region L to form σ , arguing that σ_{LR} is close to $\sigma_L \otimes \sigma_R$, and finally using the triangle inequality to show the same holds for τ .

Lemma 3 says that there is a state σ_L with rank χ defined on region L , such that

$$D(\tau_L, \sigma_L) \leq C_1 e^{-\frac{\log(\chi)}{8\xi \log(d)}} \quad (58)$$

In fact, the choice of σ_L of rank χ that minimizes the distance to τ_L is the state $P_L \tau_L / \text{Tr}(P_L \tau_L)$ where P_L is the projector onto the eigenvectors of τ_L associated with the largest χ eigenvalues. Accordingly, we define a normalization constant $q = 1/\text{Tr}(P_L \tau_L)$ and let $|\nu\rangle = \sqrt{q} P_L |\eta\rangle$ and $\sigma = |\nu\rangle\langle\nu| = q P_L \tau P_L$ be normalized states.

We first need to show that

$$\text{Cor}(L : R)_{|\nu\rangle} := \max_{\|A\|, \|B\| \leq 1} \text{Tr}((A \otimes B)(\sigma_{LR} - \sigma_L \otimes \sigma_R)) \quad (59)$$

is small, given only that $\text{Cor}(L : R)_{|\eta\rangle}$ is small. Suppressing tensor product symbols, we can write

$$\begin{aligned} & \text{Tr}((AB)(\sigma_{LR} - \sigma_L \sigma_R)) \\ &= \langle \nu | AB | \nu \rangle - \langle \nu | A | \nu \rangle \langle \nu | B | \nu \rangle \\ &= q \langle \eta | P_L A B P_L | \eta \rangle - q^2 \langle \eta | P_L A P_L | \eta \rangle \langle \eta | P_L B P_L | \eta \rangle \\ &= q \langle \eta | (P_L A P_L) B | \eta \rangle - q^2 \langle \eta | P_L A P_L | \eta \rangle \langle \eta | P_L B | \eta \rangle \\ &= q \langle \eta | (P_L A P_L) B | \eta \rangle \\ &\quad - q^2 \langle \eta | P_L A P_L | \eta \rangle (\langle \eta | P_L B | \eta \rangle - \langle \eta | P_L | \eta \rangle \langle \eta | B | \eta \rangle) \\ &\quad - q^2 \langle \eta | P_L A P_L | \eta \rangle \langle \eta | P_L | \eta \rangle \langle \eta | B | \eta \rangle \\ &= q (\langle \eta | (P_L A P_L) B | \eta \rangle - \langle \eta | P_L A P_L | \eta \rangle \langle \eta | B | \eta \rangle) \\ &\quad - q^2 \langle \eta | P_L A P_L | \eta \rangle (\langle \eta | P_L B | \eta \rangle - \langle \eta | P_L | \eta \rangle \langle \eta | B | \eta \rangle) \end{aligned}$$

from which we can conclude

$$\text{Cor}(L : R)_{|\nu\rangle} \leq (q + q^2) \text{Cor}(L : R)_{|\eta\rangle} \quad (60)$$

The normalization constant q is $1/(1 - D(\tau_L, \sigma_L)^2)$ which will be close to 1 as long as χ is sufficiently large. If we choose $\log(\chi) = 8\xi \log(d)(1 + \log(C_1))$ or larger, then q will certainly be smaller than 2 and $q + q^2 \leq 6$.

The combination of the fact that σ_L has small rank and that σ has small correlations between L and R will allow us to show that σ_{LR} is close to $\sigma_L \otimes \sigma_R$. We do this by invoking Lemma 20 of [26], although we reproduce the argument below. We can express the trace norm as

$$\begin{aligned} \|\sigma_{LR} - \sigma_L \otimes \sigma_R\|_1 &= \max_{\|T\| \leq 1} \text{Tr}(T(\sigma_{LR} - \sigma_L \otimes \sigma_R)) \\ &= \max_{\|T\| \leq 1} \text{Tr}(P_L T P_L (\sigma_{LR} - \sigma_L \otimes \sigma_R)) \end{aligned}$$

where the second equality follows from the fact that P_L fixes the state σ . We can perform a Schmidt decomposition of the operator $P_L T P_L$ into a sum of at most χ^2 terms which are each a product operator across the L/R partition

$$P_L T P_L = \sum_{j=1}^{\chi^2} T_{L,j} \otimes T_{R,j} \quad (61)$$

and also such that $\|T_{L,j}\|, \|T_{R,j}\| \leq 1$ (see Lemma 20 of [26] for full justification of this). Then we may write

$$\begin{aligned} & \|\sigma_{LR} - \sigma_L \otimes \sigma_R\|_1 \\ & \leq \max_{\|T\| \leq 1} \text{Tr} \left(\left(\sum_{j=1}^{\chi^2} T_{L,j} \otimes T_{R,j} \right) (\sigma_{LR} - \sigma_L \otimes \sigma_R) \right) \\ & \leq \sum_{j=1}^{\chi^2} \max_{\|T_{L,j}\|, \|T_{R,j}\| \leq 1} \text{Tr}((T_{L,j} \otimes T_{R,j})(\sigma_{LR} - \sigma_L \otimes \sigma_R)) \\ & \leq \chi^2 \text{Cor}(L : R)_{|\nu\rangle} \\ & \leq 6\chi^2 \text{Cor}(L : R)_{|\eta\rangle} \leq 6\chi^2 \exp(-\text{dist}(L, R)/\xi) \end{aligned}$$

as long as $\chi \geq 8\xi \log(d)(1 + \log(C))$ and $\text{dist}(L, R) \geq t_0$.

Moreover the purified distance is bounded by the square root of the trace norm of the difference (Lemma 1), allowing us to say

$$D(\sigma_{LR}, \sigma_L \otimes \sigma_R) \leq \sqrt{6}\chi \exp(-\text{dist}(L, R)/(2\xi)) \quad (62)$$

Then, by the triangle inequality, we can bound

$$\begin{aligned} & D(\tau_{LR}, \tau_L \otimes \tau_R) \\ & \leq D(\tau_{LR}, \sigma_{LR}) + D(\sigma_{LR}, \sigma_L \otimes \sigma_R) \\ & \quad + D(\sigma_L \otimes \sigma_R, \tau_L \otimes \tau_R) \\ & \leq D(\tau_{LR}, \sigma_{LR}) + D(\sigma_{LR}, \sigma_L \otimes \sigma_R) \\ & \quad + D(\sigma_L, \tau_L) + D(\sigma_R, \tau_R) \\ & \leq 3C_1 \exp\left(-\frac{\log(\chi)}{8\xi \log(d)}\right) + \sqrt{6}\chi \exp\left(-\frac{\text{dist}(L, R)}{2\xi}\right) \end{aligned}$$

We can choose

$$\log(\chi) = 8\xi \log(d)(1 + \log(3C_1)) + \text{dist}(L, R)/(4\xi) \quad (63)$$

Then each term can be bounded so that

$$D(\tau_{LR}, \tau_L \otimes \tau_R) \leq 2\sqrt{6}(3eC_1)^{8\xi \log(d)} e^{-\frac{\text{dist}(L, R)}{32\xi^2 \log(d)}}$$

which proves the first part of the Lemma.

If τ is the unique ground state of a gapped Hamiltonian, then we may use the second part of Lemma 3, and bound

$$\begin{aligned} & D(\tau_{LR}, \tau_L \otimes \tau_R) \\ & \leq 3C_2 e^{-\tilde{O}\left(\frac{\Delta^{1/3} \log^{4/3}(\chi)}{\log^{4/3}(d)}\right)} + \sqrt{6}\chi e^{-O(\Delta \text{dist}(L, R))} \end{aligned}$$

Here we can choose

$$\log(\chi) = O\left(\frac{\log(d)(1 + \log(3C_2))^{3/4}}{\Delta^{1/4}} + \Delta \text{dist}(L, R)\right)$$

and then each term is small enough to make the bound

$$\begin{aligned} & D(\tau_{LR}, \tau_L \otimes \tau_R) \\ & \leq e^{\tilde{O}\left(\frac{\Delta^{5/3} \text{dist}(L, R)^{4/3}}{\log^{4/3}(d)}\right)} \\ & \quad + \sqrt{6}e^{O\left(\frac{\log(d) \log^{3/4}(3eC_2)}{\Delta^{1/4}}\right)} e^{-O(\Delta \text{dist}(L, R))} \\ & \leq e^{\tilde{O}\left(\frac{\log^3(d)}{\Delta}\right)} e^{-O(\Delta \text{dist}(L, R))} \quad (64) \end{aligned}$$

as long as $\text{dist}(L, R) \geq \Omega(\log^4(d)/\Delta^2)$, so that the first term in the second line is dominated by the second term. Also note we have used $\log(C_2) = \tilde{O}(\log^{8/3}(d)/\Delta)$ in the last line. This proves the second part of the lemma. \square

Proof summary for Theorem 2. First, we make the following observation about tripartitions of the chain into contiguous regions L , M , and R : since $|\psi\rangle$ has exponential decay of correlations, the quantity $D(\rho_{LR}, \rho_L \otimes \rho_R)$

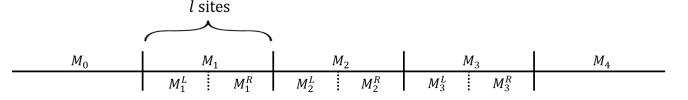


FIG. 3: Schematic for the proof of Theorem 2. The chain is divided into regions M_i of length l , which are themselves divided into left and right halves M_i^L and M_i^R . The state $|\tilde{\psi}\rangle$ is constructed by starting with $|\psi\rangle$, applying unitaries that act only on M_i for each i to disentangle the state across the M_i^L/M_i^R cut, projecting onto a product state across those cuts, and finally applying the inverse unitaries on regions M_i .

is exponentially small in $|M|/\xi'$, where $\xi' = O(\xi^2 \log(d))$. This is captured in Lemma 7 and requires the area law result from [26]. One can truncate ρ_L and ρ_R to bond dimension $d^{|M|/2}$, incurring small error, then purify ρ_L into $|\alpha\rangle$ using the left half of region M as the purifying auxiliary space and ρ_R into $|\beta\rangle$ using the right half. Since $|\alpha\rangle \otimes |\beta\rangle$ and $|\psi\rangle$ are nearly the same state after tracing out M , there is a unitary U acting only on M that nearly disentangles $|\psi\rangle$ across the central cut of M , with $U|\psi\rangle \approx |\alpha\rangle \otimes |\beta\rangle$ (Uhlmann's theorem, Lemma 6).

The proof constructs the approximation $|\tilde{\psi}\rangle$ by applying three steps of operations on the exact state $|\psi\rangle$. First, the chain is broken up into many regions $\{M_i\}_{i=0}^{n+1}$ of length l , and disentangling unitaries U_i as described above are applied to each region M_i in parallel. The state is close to, but not exactly, a product state across the center cut of each region M_i . To make it an exact product state, the second step is to apply rank-1 projectors Π_i onto the right half of the state across each of these cuts, starting with the leftmost cut and working our way down the chain. Then, the third step is to apply the reverse unitaries U_i^\dagger that we applied in step 1.

The projection step is the cause of the error between $|\psi\rangle$ and $|\tilde{\psi}\rangle$. The number of projections is $O(N)$, but the error accrued locally is only constant. This follows from the fact that the projectors are rank 1, so once we apply projector Π_i , region M_j is completely decoupled from $M_{j'}$ when $j < i$ and $j' > i$. Thus any additional projections, which act only on the regions to the right of M_i , have no effect on the reduced density matrix on M_j (except to reduce its norm). Using this logic, we show that the number of errors that actually affect the state locally on a region X is proportional to the number of sites in X , and not the number of sites in the whole chain. To make this error less than ϵ for any region of at most k sites, we can choose $l = O(\xi' \log(k/\epsilon))$.

After step 2, the state is a product state on blocks of l sites each, and in step 3, unitaries are applied that couple neighboring blocks, so the maximum Schmidt rank across any cut cannot exceed d^l . This yields the scaling for χ .

The result is improved when $|\psi\rangle$ is the ground state of a gapped local Hamiltonian by using the improved area law [7, 25] in the proof of Lemma 7 and in the truncation of the states ρ_L and ρ_R before purifying into $|\alpha\rangle$ and $|\beta\rangle$.

Finally, it can be seen that $|\tilde{\psi}\rangle$ is formed by a constant-depth quantum circuit with two layers of unitaries, where each unitary acts on l qubits. The first layer prepares the product state over blocks of length l that is attained after applying projections in step 2, and the second layer applies the inverse unitaries from step 3. Each unitary in this circuit can be decomposed into a sequence of nearest-neighbor gates with depth $\tilde{O}(d^{2l})$. \square

Proof of Theorem 2. We fix an even integer l , which we will specify later, and divide the N sites of the chain into $n + 2$ segments of length l , which we label, from left to right: M_0, M_1, \dots, M_{n+1} . If N does not divide l evenly, then we allow segment M_{n+1} to have fewer than l sites. For $i \in [1, n]$ let L_i be the sites to the left of region M_i and let R_i be the sites to the right of region M_i .

Lemma 7 tells us that, since $|\psi\rangle$ has (t_0, ξ) -exponential decay of correlations, for any $i \in [1, n]$, $\rho_{L_i R_i}$ is close to $\rho_{L_i} \otimes \rho_{R_i}$ when l is much larger than ξ' :

$$D(\rho_{L_i R_i}, \rho_{L_i} \otimes \rho_{R_i}) \leq C_3 \exp(-l/\xi') \quad (65)$$

whenever $l \geq t_0$. We also choose $\chi = d^l$ and for each i define ρ'_{L_i} and ρ'_{R_i} , each with rank at most $\sqrt{\chi}$ by taking $A = L_i$ and $A = R_i$ in Lemma 3. Thus we have

$$D(\rho_{L_i}, \rho'_{L_i}) \leq C_1 \exp(-l/(16\xi)) \quad (66)$$

$$D(\rho_{R_i}, \rho'_{R_i}) \leq C_1 \exp(-l/(16\xi)) \quad (67)$$

Then by the triangle inequality we have

$$\begin{aligned} & D(\rho_{L_i R_i}, \rho'_{L_i} \otimes \rho'_{R_i}) \\ & \leq D(\rho_{L_i R_i}, \rho_{L_i} \otimes \rho_{R_i}) + D(\rho_{L_i} \otimes \rho_{R_i}, \rho'_{L_i} \otimes \rho'_{R_i}) \\ & \quad + D(\rho'_{L_i} \otimes \rho_{R_i}, \rho'_{L_i} \otimes \rho'_{R_i}) \\ & = D(\rho_{L_i R_i}, \rho_{L_i} \otimes \rho_{R_i}) + D(\rho_{L_i}, \rho'_{L_i}) + D(\rho_{R_i}, \rho'_{R_i}) \\ & \leq C \exp(-l/\xi'') \end{aligned} \quad (68)$$

where $C \leq 2C_1 + C_3$, $\xi'' = \max(\xi', 16\xi)$, whenever $l \geq t_0$.

Note that $|\psi\rangle_{L_i M_i R_i}$ can be viewed as a purification of $\rho_{L_i R_i}$ with M_i the purifying auxiliary system. Divide region M_i in half, forming M_i^L (left half) and M_i^R (right half). See Figure 3 for a schematic. Each of these subsystems has total dimension $d^{l/2}$ and thus can act as the purifying auxiliary system for ρ'_{L_i} or ρ'_{R_i} . Let $|\alpha_i\rangle_{L_i M_i^L}$ be a purification of ρ'_{L_i} and $|\beta_i\rangle_{M_i^R R_i}$ be a purification of ρ'_{R_i} . Thus, $|\alpha_i\rangle \otimes |\beta_i\rangle$, which is defined over the entire original chain, is a purification of $\rho'_{L_i} \otimes \rho'_{R_i}$.

Uhlmann's theorem (Lemma 6) shows how these purifications are related by a unitary on the purifying auxiliary system: for each M_i with $i \in [1, n]$, there is a unitary U_i acting non-trivially on region M_i and as the identity on the rest of the chain such that $U_i |\psi\rangle$ is very close to $|\alpha_i\rangle \otimes |\beta_i\rangle$, a product state across the cut between M_i^L and M_i^R . In other words, U_i disentangles L_i from R_i , up to some small error, by acting only on M_i . Formally we say that

$$\left| \langle \alpha_i |_{L_i M_i^L} \otimes \langle \beta_i |_{M_i^R R_i} U_i |\psi\rangle_{L_i M_i R_i} \right| = \sqrt{1 - \delta_i^2} \quad (69)$$

where $\delta_i \leq C \exp(-l/\xi'')$ for all i .

An equivalent way to write this fact is

$$\begin{aligned} U_i |\psi\rangle_{L_i M_i R_i} &= \sqrt{1 - \delta_i^2} |\alpha_i\rangle_{L_i M_i^L} \otimes |\beta_i\rangle_{M_i^R R_i} \\ &\quad + \delta_i |\phi'_i\rangle_{L_i M_i^L M_i^R R_i} \end{aligned} \quad (70)$$

where $|\phi'_i\rangle$ is a normalized state orthogonal to $|\alpha_i\rangle \otimes |\beta_i\rangle$.

We can define the projector

$$\Pi_i = I_{L_i M_i^L} \otimes |\beta_i\rangle \langle \beta_i|_{M_i^R R_i}. \quad (71)$$

whose rank is 1 when considered as an operator acting only on $M_i^R R_i$.

We notice that $\Pi_i U_i |\psi\rangle_{L_i M_i R_i}$ is a product state across the M_i^L/M_i^R cut and has a norm close to 1. Suppose we alternate between applying disentangling operations U_i and projections Π_i onto a product state as we move down the chain. Each Π_i will reduce the norm of the state, but we claim the norm will never vanish completely (and delay the proof for later for clarity of argument).

Claim 1. *If $l \geq \xi'' \log(3C)$, then*

$$\|\Pi_n U_n \dots \Pi_1 U_1 |\psi\rangle\| \neq 0 \quad (72)$$

This allows us to define

$$|\tilde{\phi}\rangle = \frac{\Pi_n U_n \dots \Pi_1 U_1 |\psi\rangle}{\|\Pi_n U_n \dots \Pi_1 U_1 |\psi\rangle\|} \quad (73)$$

Note that, to put our proof in line with what is described in the introduction and proof summary earlier, we may act with all the unitaries prior to the projectors if we conjugate the projectors

$$|\tilde{\phi}\rangle \propto \Pi'_n \dots \Pi'_1 U_n \dots U_1 |\psi\rangle \quad (74)$$

where $\Pi'_i = U_n \dots U_{i+1} \Pi_i U_{i+1}^\dagger \dots U_n^\dagger$, which still only acts on the region R_i .

This can be compared with the state $|\phi\rangle$ defined by applying the disentangling operations without projecting:

$$|\phi\rangle = U_n \dots U_1 |\psi\rangle \quad (75)$$

We claim that $|\tilde{\phi}\rangle$ is a good local approximation for $|\phi\rangle$ (and delay the proof for clarity of argument).

Claim 2. *For any integer k' , $|\tilde{\phi}\rangle$ is a (k', ϵ') -local approximation to $|\phi\rangle$ with $\epsilon' = C\sqrt{k'/l + 3} \exp(-l/\xi'')$.*

Next we can define

$$|\tilde{\psi}\rangle = U_n^\dagger \dots U_1^\dagger |\tilde{\phi}\rangle \quad (76)$$

which parallels the relationship

$$|\psi\rangle = U_n^\dagger \dots U_1^\dagger |\phi\rangle \quad (77)$$

Now suppose X is a contiguous region of the chain of length k . Then there is a region X' of the chain of length

at most $k' = k + 2l$ that contains X and is made up of regions M_j where $j \in [a', b']$.

$$\begin{aligned}
& \|\text{Tr}_{X^c}(|\tilde{\psi}\rangle\langle\tilde{\psi}| - |\psi\rangle\langle\psi|)\|_1 \\
& \leq \|\text{Tr}_{X'^c}(|\tilde{\psi}\rangle\langle\tilde{\psi}| - |\psi\rangle\langle\psi|)\|_1 \\
& = \|\text{Tr}_{X'^c}(|\tilde{\phi}\rangle\langle\tilde{\phi}| - |\phi\rangle\langle\phi|)\|_1 \\
& \leq C\sqrt{k/l+5}\exp(-l/\xi'') \\
& \leq C\sqrt{6k}\exp(-l/\xi'') \tag{78}
\end{aligned}$$

where the third line follows from the fact that $|\phi\rangle$ and $|\psi\rangle$ are related by a unitary that does not couple region X' and region X'^c , and the fourth line follows from Claim 2.

If we choose $l = \max(t_0, \xi'' \log(3C\sqrt{k}/\epsilon))$, then the requirements of Claim 1 are satisfied, and we can see from Eq. (78) that $|\tilde{\psi}\rangle$ is a (k, ϵ) -local approximation to $|\psi\rangle$, item (1) of the theorem.

Item (2) states that $|\tilde{\psi}\rangle$ can be written as an MPS with constant bond dimension. This can be seen by the following logic. The Schmidt rank of $|\tilde{\phi}\rangle$ across any cut M_i^L/M_i^R is 1, as discussed in the proof of Claim 2 (see Eq. (82)), since the projector Π_i projects onto a product state across that cut and unitaries U_j with $j > i$ act trivially across the cut. By Corollary 1, this implies that the Schmidt rank across any cut M_i^R/M_{i+1}^L can be at most $d^{l/2}$. Acting with the inverse unitaries U_j^\dagger on $|\tilde{\phi}\rangle$ to form $|\tilde{\psi}\rangle$ preserves the Schmidt rank across the cut M_i^R/M_{i+1}^L , since none couple both sides of the cut. Because any cut is at most distance $l/2$ from some M_i^R/M_{i+1}^L cut, the Schmidt rank across an arbitrary cut can be at most a factor of $d^{l/2}$ greater, again by Corollary 1, meaning the maximum Schmidt rank across any cut of $|\psi\rangle$ is $\chi = d^l$.

Given our choice of $l = \xi'' \log(C\sqrt{6k}/\epsilon)$, we find that the state can be represented by an MPS with bond dimension

$$\begin{aligned}
\chi & = (\sqrt{6}C)^{\xi'' \log(d)} (k/\epsilon^2)^{\xi'' \log(d)/2} \\
& = e^{\tilde{O}(\xi \log(d))} (k/\epsilon^2)^{O(\xi^2 \log^2(d))}. \tag{79}
\end{aligned}$$

This proves item (2). Note that, in the case that our choice of l exceeds N , it is not possible to form the construction we have described. However, in this case d^l will exceed d^N and we may take $|\tilde{\psi}\rangle = |\psi\rangle$, which is a local approximation for any k and ϵ and has bond dimension in line with item (2) or (2').

Item (2') follows by using the same equations with $C = 2C_2 + C_4 = \exp(\tilde{O}(\log^3(d)/\Delta))$ and $\xi'' = O(1/\Delta)$. For Lemma 7 to apply, we must have $l \geq \Omega(\log^4(d)/\Delta^2)$, but this will be satisfied for sufficiently large choices of k/ϵ^2 . Thus the final analysis yields

$$\chi = e^{\tilde{O}(\log^4(d)/\Delta^2)} (k/\epsilon^2)^{O(\log(d)/\Delta)}. \tag{80}$$

Item (3) states that $|\tilde{\psi}\rangle$ can be formed from a low-depth quantum circuit. In the proof of Claim 2, we show how the state $|\tilde{\phi}\rangle$ is a product state across divisions M_i^L/M_i^R , as in Eq. (82). Thus the state $|\tilde{\phi}\rangle$ can

be created from $|0\rangle^{\otimes N}$ by acting with non-overlapping unitaries on regions $M_i^R M_{i+1}^L$ in parallel. Each of these unitaries is supported on l sites. Then, $|\tilde{\psi}\rangle$ is related to $|\tilde{\phi}\rangle$ by another set of non-overlapping unitaries supported on l sites, as shown in Eq. (76). We conclude that $|\tilde{\psi}\rangle$ can be created from the trivial state by two layers of parallel unitary operations where each unitary is supported on l sites, as illustrated in Figure 1. In [30], it is shown how any l -qudit unitary can be decomposed into $O(d^{2l}) = O(\chi^2)$ two-qudit gates, with no need for ancillas. We can guarantee that these gates are all spatially local by spending at most depth $O(l)$ performing swap operations to move any two sites next to each other, a factor only logarithmic in the total depth. This proves the theorem. \square

Proof of Claim 1. We prove by induction. Let $|\tilde{\phi}_j\rangle = \Pi_j U_j \dots \Pi_1 U_1 |\psi\rangle$. Note that $\Pi_i U_i |\psi\rangle$ is non-zero for all i , so in particular $|\tilde{\phi}_1\rangle$ is non-zero. Furthermore we note that, if it is non-zero, $|\tilde{\phi}_j\rangle$ can be written as a product state $|\alpha'_j\rangle_{L_j M_j^L} \otimes |\beta_j\rangle_{M_j^R R_j}$ for some unnormalized but non-zero state $|\alpha'_j\rangle$, and the reduced density matrix of $|\tilde{\phi}_j\rangle$ on R_j is ρ'_{R_j} . If we assume $|\tilde{\phi}_j\rangle$ is non-zero then we can write

$$|\tilde{\phi}_{j+1}\rangle = \Pi_{j+1} U_{j+1} |\tilde{\phi}_j\rangle = |\alpha'_j\rangle \otimes \Pi_{j+1} U_{j+1} |\beta_j\rangle \tag{81}$$

and

$$\begin{aligned}
\| |\tilde{\phi}_{j+1}\rangle \|^2 & = \| |\alpha'_j\rangle \otimes \Pi_{j+1} U_{j+1} |\beta_j\rangle \|^2 \\
& = \| |\alpha'_j\rangle \|^2 \| \Pi_{j+1} U_{j+1} |\beta_j\rangle \|^2 \\
& = \| |\alpha'_j\rangle \|^2 \text{Tr}(\Pi_{j+1} U_{j+1} |\beta_j\rangle \langle \beta_j| U_{j+1}^\dagger \Pi_{j+1}) \\
& = \| |\alpha'_j\rangle \|^2 \text{Tr}(\Pi_{j+1} U_{j+1} \rho'_{R_j} U_{j+1}^\dagger \Pi_{j+1}) \\
& \geq \| |\alpha'_j\rangle \|^2 \text{Tr}(\Pi_{j+1} U_{j+1} \rho_{R_j} U_{j+1}^\dagger \Pi_{j+1}) \\
& \quad - \| |\alpha'_j\rangle \|^2 \| \rho_{R_j} - \rho'_{R_j} \|_1 \\
& \geq \| |\alpha'_j\rangle \|^2 \text{Tr}(\Pi_{j+1} U_{j+1} \rho_{R_j} U_{j+1}^\dagger \Pi_{j+1}) \\
& \quad - \| |\alpha'_j\rangle \|^2 \| \rho_{R_j} - \rho'_{R_j} \|_1 \\
& \geq \| |\alpha'_j\rangle \|^2 \| \Pi_{j+1} U_{j+1} |\psi\rangle \|^2 \\
& \quad - \| |\alpha'_j\rangle \|^2 \| \rho_{R_j} - \rho'_{R_j} \|_1 \\
& \geq \| |\alpha'_j\rangle \|^2 (1 - C \exp(-l/\xi''))^2 \\
& \quad - \| |\alpha'_j\rangle \|^2 C \exp(-l/\xi'') \\
& > 0
\end{aligned}$$

as long as $l \geq \xi'' \log(3C)$. \square

Proof of Claim 2. First, consider the cut M_i^L/M_i^R during the formation of the state $|\tilde{\phi}\rangle$. When the projector Π_i is applied, the state becomes a product state across this cut. The remaining operators are U_j and Π_j with $j > i$, and thus they have no effect on the Schmidt rank across the M_i^L/M_i^R cut, meaning $|\tilde{\phi}\rangle$ is a product state across

each of these cuts, or in other words

$$|\tilde{\phi}\rangle = |\phi_1\rangle_{M_0 M_1^L} \otimes |\phi_2\rangle_{M_1^R M_2^L} \otimes \dots \\ \dots \otimes |\phi_n\rangle_{M_{n-1}^R M_n^L} \otimes |\phi_{n+1}\rangle_{M_n^R M_{n+1}}. \quad (82)$$

Given an integer k and a contiguous region X of length k , we can find integers a and b such that $Y = M_a^R M_{a+1} \dots M_{b-1} M_b^L$ contains X and $|b-a| \leq k/l + 2$. Then

$$\text{Tr}_{Y^c}(|\tilde{\phi}\rangle\langle\tilde{\phi}|) \\ = |\phi_{a+1}\rangle\langle\phi_{a+1}| \otimes \dots \otimes |\phi_b\rangle\langle\phi_b| \\ \propto \text{Tr}_{L_a M_a^L M_b^R R_b} \left(\Pi_b U_b \dots \Pi_a U_a |\psi\rangle\langle\psi| U_a^\dagger \Pi_a \dots U_b^\dagger \Pi_b \right)$$

The advantage here is that all of the U_i and Π_i for which $i \notin [a, b]$ have disappeared. On the other hand, we have

$$\text{Tr}_{Y^c}(|\phi\rangle\langle\phi|) \\ = \text{Tr}_{Y^c}(U_n \dots U_1 |\psi\rangle\langle\psi| U_1^\dagger \dots U_n^\dagger) \\ = \text{Tr}_{L_a M_a^L M_b^R R_b} \left(U_b \dots U_a |\psi\rangle\langle\psi| U_a^\dagger \dots U_b^\dagger \right) \quad (83)$$

Note that, since

$$U_i |\psi\rangle = \sqrt{1 - \delta_i^2} |\alpha_i\rangle \otimes |\beta_i\rangle + \delta_i |\phi'_i\rangle \quad (84)$$

we can say that

$$\Pi_i U_i |\psi\rangle = U_i |\psi\rangle - \delta_i (I - \Pi_i) |\phi'_i\rangle \quad (85)$$

and thus

$$\Pi_b U_b \dots \Pi_a U_a |\psi\rangle \\ = U_b \dots U_a |\psi\rangle - \\ \sum_{j=a}^b \delta_j (\Pi_b U_b \dots \Pi_{j+1} U_{j+1}) (I - \Pi_j) |\phi'_j\rangle \\ \equiv U_b \dots U_a |\psi\rangle - \delta |\phi'\rangle \quad (86)$$

where $\delta \leq \sqrt{\sum_{j=a}^b \delta_j^2}$ and normalized $|\phi'\rangle$ is normalized. This implies

$$\frac{|\langle\psi| U_a^\dagger \dots U_b^\dagger \Pi_b U_b \dots \Pi_a U_a |\psi\rangle|}{\|\Pi_a \dots U_b^\dagger \Pi_b U_b \dots \Pi_a U_a |\psi\rangle\|} \geq \sqrt{1 - \delta^2} \quad (87)$$

which shows that $D_1(\tau, \tau') \leq \delta$ where

$$\tau = U_b \dots U_a |\psi\rangle\langle\psi| U_a^\dagger \dots U_b^\dagger \\ \tau' = \frac{\Pi_b U_b \dots \Pi_a U_a |\psi\rangle\langle\psi| U_a^\dagger \Pi_a \dots U_b^\dagger \Pi_b}{\|\Pi_a \dots U_b^\dagger \Pi_b U_b \dots \Pi_a U_a |\psi\rangle\|^2} \quad (88)$$

and hence

$$D_1(\text{Tr}_{X^c}(|\phi\rangle\langle\phi|), \text{Tr}_{X^c}(|\tilde{\phi}\rangle\langle\tilde{\phi}|)) \\ \leq D_1(\text{Tr}_{Y^c}(|\phi\rangle\langle\phi|), \text{Tr}_{Y^c}(|\tilde{\phi}\rangle\langle\tilde{\phi}|)) \\ \leq D_1(\tau_Y, \tau'_Y) \leq D_1(\tau, \tau') \leq \delta \\ \leq C \sqrt{k/l + 3} \exp(-l/\xi''). \quad (89)$$

This holds for any region X of length k , so this proves the claim. \square

D. Proof of Theorem 3

First we state a lemma that will do most of the legwork needed for Theorem 3. Then we prove Theorem 3.

Lemma 8. *Suppose, for $j = 0, 1$, $H^{(j)}$ is a translationally invariant Hamiltonian defined on a chain of length N and local dimension d . Further suppose that $|\psi^{(j)}\rangle$ is the unique ground state of $H^{(j)}$ with energy $E_0^{(j)}$, and let $\Delta^{(j)}$ be the spectral gap of $H^{(j)}$. We may form a new chain with local dimension $2d$ by adding an ancilla qubit to each site of the chain. Then there is a Hamiltonian K defined on this chain such that*

(1) *The ground state energy of K is*

$$E_0^K = \frac{1}{3} \min_j E_0^{(j)} \quad (90)$$

(2) *If $E_0^{(0)} < E_0^{(1)}$, then the ground state of K is $|0^N\rangle_A \otimes |\psi^{(0)}\rangle$ and if $E_0^{(1)} < E_0^{(0)}$, then the ground state is $|1^N\rangle_A \otimes |\psi^{(1)}\rangle$, where A refers to the N ancilla registers collectively.*

(3) *If $E_0^{(0)} < E_0^{(1)}$, then the spectral gap of K is at least $\min(\Delta^{(0)}, E_0^{(1)} - E_0^{(0)}, 1)/3$ and if $E_0^{(1)} < E_0^{(0)}$, then the spectral gap of K is at least $\min(\Delta^{(1)}, E_0^{(0)} - E_0^{(1)}, 1)/3$*

Proof of Lemma 8. Note that a variant of this lemma is employed in [28, 31, 32] to show the undecidability of certain properties of translationally invariant Hamiltonians.

Since $H^{(j)}$ is translationally invariant, it is specified by its single interaction term $H_{i,i+1}^{(j)}$:

$$H^{(j)} = \sum_{i=1}^{N-1} H_{i,i+1}^{(j)} \quad (91)$$

The Hamiltonian K will be defined over a new chain where we attach to each site an ancilla qubit, increasing the local Hilbert space dimension by a factor of 2. We refer to the ancilla associated with site i by the subscript A_i , and we refer to the collection of ancillas together with the subscript A . Operators or states without a subscript are assumed to act on the original d -dimensional part of the local Hilbert spaces. Let

$$K = \sum_{i=1}^{N-1} K_{i,i+1} \quad (92)$$

where

$$K_{i,i+1} = \frac{1}{3} H_{i,i+1}^{(0)} \otimes |00\rangle\langle 00|_{A_i A_{i+1}} \quad (93)$$

$$+ \frac{1}{3} H_{i,i+1}^{(1)} \otimes |11\rangle\langle 11|_{A_i A_{i+1}} \quad (94)$$

$$+ I_{i,i+1} \otimes (|01\rangle\langle 01| + |10\rangle\langle 10|)_{A_i A_{i+1}} \quad (95)$$

with $I_{i,i+1}$ denoting the identity operation on sites i and $i+1$. In this form it is clear that K is a nearest-neighbor translationally invariant Hamiltonian and that each interaction term has operator norm 1 (a requirement under our treatment of 1D Hamiltonians). The picture we get is that if two neighboring ancillas are both $|0\rangle$, then $H_{i,i+1}^{(0)}/3$ is applied, if both are $|1\rangle$ then $H_{i,i+1}^{(1)}/3$ is applied, and if the ancillas are different, then $I_{i,i+1}$ is applied. Following this intuition, we can rewrite K as follows

$$K = \sum_{x=0}^{2^N-1} \left(|x\rangle \langle x|_A \otimes \sum_{i=1}^{N-1} K_{x,i} \right) \quad (96)$$

where the first sum is over all settings of the ancillas and the operator $K_{x,i}$ acts on the non-ancilla portion of sites i and $i+1$, with

$$K_{x,i} = \begin{cases} H_{i,i+1}^{(0)}/3 & \text{if } x_i = x_{i+1} = 0 \\ H_{i,i+1}^{(1)}/3 & \text{if } x_i = x_{i+1} = 1 \\ I_{i,i+1} & \text{if } x_i \neq x_{i+1} \end{cases} \quad (97)$$

We analyze the spectrum of K . If $H^{(j)}$ has eigenvalues $E_n^{(j)}$ with corresponding eigenvectors $|\phi_n^{(j)}\rangle$ (where $E_n^{(j)}$ is non-decreasing with increasing integers n), then the states $|0^N\rangle_A \otimes |\phi_n^{(0)}\rangle$ and $|1^N\rangle_A \otimes |\phi_n^{(1)}\rangle$ are eigenstates of K with eigenvalues $E_n^{(0)}/3$ and $E_n^{(1)}/3$, respectively.

Recall that eigenvectors of a Hamiltonian span the whole Hilbert space over which the Hamiltonian is defined. Therefore, the eigenvectors of K listed above span the entire sectors of the Hilbert space associated with the ancillas set to $|0^N\rangle_A$ or $|1^N\rangle_A$. Suppose $|\phi\rangle$ is another eigenvector of K . Since it is orthogonal to all of the previously listed eigenvectors, $|\phi\rangle$ can be written

$$|\phi\rangle = \sum_{x=1}^{2^N-2} \alpha_x |x\rangle_A \otimes |\eta_x\rangle \quad (98)$$

for some set of complex coefficients α_x with $\sum_x |\alpha_x|^2 = 1$ and some set of normalized states $|\eta_x\rangle$. The sum explicitly leaves out the $x=0=0^N$ and $x=2^N-1=1^N$ binary strings because these states lie in the subspace spanned by eigenstates already listed. We wish to lower bound the energy of the state $|\phi\rangle$.

$$\begin{aligned} \langle \phi | K | \phi \rangle &= \sum_{x=1}^{2^N-2} \sum_{y=1}^{2^N-2} \alpha_x^* \alpha_y \langle x|_A \langle \eta_x | K | y \rangle_A | \eta_y \rangle \\ &= \sum_{x=1}^{2^N-2} \sum_{y=1}^{2^N-2} \sum_{z=0}^{2^N-1} \alpha_x^* \alpha_y \langle x|z\rangle \langle z|y\rangle_A \langle \eta_x | \sum_{i=1}^{N-1} K_{z,i} | \eta_y \rangle \\ &= \sum_{x=1}^{2^N-2} |\alpha_x|^2 \langle \eta_x | \sum_{i=1}^{N-1} K_{x,i} | \eta_x \rangle \end{aligned} \quad (99)$$

We would like to show a lower bound for this quantity. We make the following claim:

Claim 3. For any state $|\eta\rangle$, and any $1 \leq a < b \leq N$

$$\langle \eta | \sum_{i=a}^{b-2} H_{i,i+1}^{(j)} | \eta \rangle \geq \frac{b-a}{N-1} E_0^{(j)} - 1 \quad (100)$$

Proof of Claim 3. First we prove it in the case that $M := b-a$ divides N . Let region Y refer to sites $[a, b-1]$, let $\rho = \text{Tr}_{Y^c}(|\eta\rangle \langle \eta|)$, and let $\sigma = \rho \otimes \dots \otimes \rho$ be N/M copies of ρ which covers all N sites. Then

$$\begin{aligned} \text{Tr}(H^{(j)} \sigma) &= \frac{N}{M} \langle \eta | \sum_{i=a}^{b-2} H_{i,i+1}^{(j)} | \eta \rangle \\ &\quad + \sum_{k=1}^{N/M-1} \langle \eta | H_{kM, kM+1}^{(j)} | \eta \rangle \\ &\leq \frac{N}{M} \langle \eta | \sum_{i=a}^{b-2} H_{i,i+1}^{(j)} | \eta \rangle + \left(\frac{N}{M} - 1 \right) \end{aligned} \quad (101)$$

where the last line follows from the fact that the interaction strength $\|H_{i,i+1}^{(j)}\| \leq 1$. Moreover, by the variational principle, $\text{Tr}(H^{(j)} \sigma) \geq E_0^{(j)}$. These observations together yield

$$\langle \eta | \sum_{i=a}^{b-2} H_{i,i+1}^{(j)} | \eta \rangle \geq (M/N)(E_0^{(j)} + 1) - 1 \quad (102)$$

which implies the statement of the claim.

Now suppose M does not divide N . We decompose $N = sM + r$ for non-negative integers s and $r < M$. Let $\sigma = \rho \otimes \dots \otimes \rho \otimes |\nu\rangle \langle \nu|$ where there are s copies of ρ and $|\nu\rangle$ is the exact ground state of $\sum_{i=N-r+1}^{N-1} H_{i,i+1}^{(j)}$, which has energy E_r . Then

$$\text{Tr}(H^{(j)} \sigma) \leq s \langle \eta | \sum_{i=a}^{i_0+M_0-2} H_{i,i+1}^{(j)} | \eta \rangle + E_r + s \quad (103)$$

Here we invoke the variational principle twice. First, note that the expectation value of $\sum_{i=N-r+1}^{N-1} H_{i,i+1}^{(j)}$ in the state $|\phi_0^{(j)}\rangle$ (the exact ground state of the whole chain) is exactly $(r-1)E_0^{(j)}/(N-1)$. Since $|\nu\rangle$ is the exact ground state of that Hamiltonian, E_r must be smaller than this quantity. Second, as before, $\text{Tr}(H^{(j)} \sigma) \geq E_0^{(j)}$. Combining these observations yields

$$\begin{aligned} \langle \eta | \sum_{i=i_0}^{i_0+M-1} H_{i,i+1}^{(j)} | \eta \rangle &\geq \frac{1}{s} E_0^{(j)} \left(1 - \frac{r-1}{N-1} \right) - 1 \\ &= \frac{M}{N-1} E_0^{(j)} - 1 \end{aligned} \quad (104)$$

□

Now we use this claim to complete the proof of Lemma 8. For any binary string x we can associate a sequence of indices $1 = i_0 < i_1 < \dots < i_m < i_{m+1} = N+1$

such that $x_i = x_{i_k}$ for all $k = 1, \dots, m$ and all i in the interval $[i_k, i_{k+1} - 1]$. Moreover we require $x_{i_{k-1}} \neq x_{i_k}$. In other words, x can be decomposed into substrings of consecutive 0s and consecutive 1s, with i_j representing the index of the “domain wall” that separates a substring of 0s from a substring of 1s. The parameter m is the number of domain walls. Using this notation, and letting $E_0^K = \min_j E_0^{(j)}/3$ we can rewrite

$$\begin{aligned} \langle \eta_x | \sum_{i=1}^{N-1} K_{x,i} | \eta_x \rangle &= \sum_{k=0}^m \langle \eta_x | \sum_{i=i_k}^{i_{k+1}-2} \frac{1}{3} H_{i,i+1}^{(x_{i_k})} | \eta_x \rangle \\ &\quad + \sum_{k=1}^m \langle \eta_x | I_{i_{k-1}, i_k} | \eta_x \rangle \\ &\geq \sum_{k=0}^m \left(\frac{i_{k+1} - i_k}{3(N-1)} E_0^{(x_{i_j})} - \frac{1}{3} \right) + m \\ &\geq E_0^K + \frac{2m-1}{3} \end{aligned} \quad (105)$$

For any x other than 0^N and 1^N , there is at least one domain wall and $m \geq 1$. Thus we can say

$$\langle \phi | K | \phi \rangle \geq E_0^K + 1/3 \quad (106)$$

We have shown that any state orthogonal to the states $|0^N\rangle_A \otimes |\phi_n^{(0)}\rangle$ and $|1^N\rangle_A \otimes |\phi_n^{(1)}\rangle$ will have energy at least $1/3$ larger than the lowest energy state of the system. Without loss of generality, suppose $E_0^{(0)} \leq E_0^{(1)}$. Then, the ground state energy is $E_0^K = E_0^{(0)}/3$ and the ground state is $|0^N\rangle_A \otimes |\phi_0^{(0)}\rangle$ (note that in the statement of the Lemma we have $|\psi^{(0)}\rangle = |\phi_0^{(0)}\rangle$). The first excited state is either $|0^N\rangle_A \otimes |\phi_1^{(0)}\rangle$, $|1^N\rangle_A \otimes |\phi_0^{(1)}\rangle$, or lies outside the sector associated with ancillas $|0^N\rangle$ and $|1^N\rangle$, whichever has lowest energy. The three cases lead to spectral gaps of $\Delta^{(0)}/3$, $(E_0^{(1)} - E_0^{(0)})/3$, and something larger than $1/3$ (due to Eq. (106)), respectively. This proves all three items of the Lemma. \square

Proof of Theorem 3. We begin by specifying a family of Hamiltonians, parameterized by $t \in [0, 1/2]$ and defined over a chain of length N with local dimension d .

$$H^Z(t) = \sum_{i=1}^{N-1} I_i \otimes I_{i+1} - (1-t) |0\rangle \langle 0|_i \otimes |0\rangle \langle 0|_{i+1} \quad (107)$$

The ground state of $H^Z(t)$ is the trivial product state $|0\rangle^{\otimes N}$ with ground state energy $t(N-1)$, and thus energy density t . The interaction strength is bounded by 1 and the spectral gap is $1-t \geq 1/2$.

Now we construct an algorithm for Problem 1. We are given H as input, with associated parameters N and d , and a lower bound Δ on the spectral gap. Let the true ground state energy for H be E and let $u = E/(N-1)$. We choose a value of s between 0 and 1, and we apply Lemma 8 to construct a Hamiltonian K combining

Hamiltonians $H/2$ and $H^Z(s/2)$. K acts on N sites, has local dimension $2d$, and has spectral gap at least $\min(\Delta, |s-u|(N-1), 1)/6$. We are given a procedure to solve Problem 2 with $\delta = 0.9$ for a single site, i.e. we can estimate the expectation value of any single site observable in the ground state of K . If $s < u$, the true reduced density matrix of K will have its ancilla bits all set to 1. If $s > u$ the reduced density matrix corresponds to the reduced density matrix of H with all its ancilla bits set to 0. Thus we can choose our single site operator to be the Z_A operator that has eigenvalue 1 for states whose ancilla bit is $|0\rangle$ and eigenvalue -1 for states whose ancilla bit is $|1\rangle$. If we have a procedure to determine $\langle \psi | Z_A | \psi \rangle$ to precision 0.9 then we can determine the setting of one of the ancilla bits in the ground state and thus determine whether u is larger or smaller than s . The time required to make this determination is $f(\min(\Delta, (N-1)|s-u|, 1)/6, 2d, N)$. Because we have control over s , we can use this procedure to binary search for the value of u . We assume we are given a lower bound on Δ but since we do not know u *a priori*, we have no lower bound on $|s-u|$, so we may not know how long to run the algorithm for Problem 2 in each step of the binary search. If our desired precision is ϵ , we will impose a maximum runtime of $f(\min(\Delta, (N-1)\epsilon/2, 1)/6, 2d, N)$ for each step. Thus, if we choose a value of s for which $|s-u| < \epsilon/2$, the output of this step of the binary search may be incorrect. After such a step, our search window will be cut in half and the correct value of u will no longer be within the window. However, u will still lie within $\epsilon/2$ of one edge of the window. Throughout the binary search, some element of the search window will always lie within $\epsilon/2$ of u , so if we run the search until the window has width ϵ and output the value \tilde{u} in the center of the search window, we are guaranteed that $|u - \tilde{u}| \leq \epsilon$. The number of steps required is $O(\log(1/\epsilon))$ and the time for each step is $f(\min(2\Delta, (N-1)\epsilon, 2)/12, 2d, N)$, yielding the statement of the theorem. \square

V. DISCUSSION

Our results paint an interesting landscape of the complexity of approximating ground states of gapped nearest-neighbor 1D Hamiltonians locally. On the one hand, we show all k -local properties of the ground state can be captured by an MPS with only constant bond dimension, an improvement over the poly(N) bond dimension required to represent the global approximation. This constant scales like a polynomial in k and $1/\epsilon$, when parameters like Δ , ξ , and d are taken as constants. On the other hand, we give evidence that, at least for the case where the Hamiltonian is translationally invariant, finding the local approximation may not offer a significant speedup over finding the global approximation: we have shown that the ability to find even a constant-precision estimate of local properties would allow one to learn a constant-precision estimate of the ground state energy

with only $O(\log(N))$ overhead. This reduction does not allow one to learn any global information about the state besides the ground state energy, so it falls short of giving a concrete relationship between the complexity of the global and local approximations. Nonetheless, the reduction has concrete consequences. In particular, at least one of the following must be true about translationally invariant gapped Hamiltonians on chains of length N :

- (1) The ground state energy can be estimated to $O(1)$ precision in $O(\log(N))$ time.
- (2) Local properties of the ground state cannot be estimated to $O(1)$ precision in time independent of N .

In particular, the second item, if true, would seem to imply that, in the translationally invariant case when $N \rightarrow \infty$, local properties cannot be estimated at all.

Indeed, it is when the chain is very long, or when we are considering the thermodynamic limit directly that our results are most relevant. In the translationally invariant case as $N \rightarrow \infty$, our first proof method (Theorem 1) yields a local approximation that is a translationally invariant MPS. However, the MPS is non-injective and the state is a macroscopic superposition on the infinite chain. Thus the bulk tensors alone do not uniquely define the state and specification of a boundary tensor at infinity is also required [12, 14]. Our second proof method (Theorem 2), on the other hand, yields a periodic MPS (with period $O(\log(k/\epsilon^2))$) that is injective and can be constructed by a constant-depth quantum circuit made from spatially local gates. If we allow the locality of the gates to be $O(\log(k/\epsilon^2))$, then the circuit can have depth 2, as in Figure 1. If we require the locality of the gates be only a constant, say 2, then the circuit can have depth $\text{poly}(k, 1/\epsilon)$.

The fact that the approximation is injective perhaps makes the latter method more powerful. Injective MPS are the exact ground states of some local gapped Hamiltonian [19, 33]. Additionally, non-injective MPS form a set of measure zero among the entire MPS manifold, so variational algorithms that explore the whole manifold are most compatible with an injective approximation. In fact, since the approximation can be generated from a constant-depth circuit, the result justifies a more restricted variational ansatz using states of that form.

This ansatz could provide several advantages over MPS in terms of number of parameters needed and ability to quickly calculate local observables, like the energy density. However, algorithms that perform variational optimization of the energy density generally suffer from two issues, regardless of the ansatz they use. First, they do not guarantee convergence to the global minimum within the ansatz set, and second, even when they do find the global minimum, the output does not necessarily correspond to a good local approximation. This stems from the fact that a state that is ϵ -close to the ground state energy density may actually be far, even orthogonal, to the actual ground state. Therefore, even a brute-force optimization over the ansatz set cannot be guaranteed to give any information about the ground state, other than its energy density.

This leaves open many questions regarding the algorithmic complexity of gapped local 1D Hamiltonians. For the general case on a finite chain, can one find a local approximation to the ground state faster than the global approximation? For translationally invariant chains, can one learn the ground state energy to $O(1)$ precision in $O(\log(N))$ time, and can one learn local properties in time independent of the chain length? Relatedly, in the thermodynamic limit, can one learn an ϵ -approximation to the ground state energy density in $O(\log(1/\epsilon))$ time, and can one learn local properties at all? These are interesting questions to consider in future work.

We would like to conclude by drawing the reader's attention to independent work studying the same problem by Huang [34], appearing simultaneously with our own.

Acknowledgments

We thank Thomas Vidick for useful discussions about this work and its algorithmic implications. AMD gratefully acknowledges support from the Dominic Orr Fellowship and the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE1745301. This work was supported by NSF and Samsung. The Institute for Quantum Information and Matter (IQIM) is an NSF Physics Frontiers Center.

-
- [1] S. R. White, "Density matrix formulation for quantum renormalization groups," *Physical review letters*, vol. 69, no. 19, p. 2863, 1992.
 - [2] S. R. White, "Density-matrix algorithms for quantum renormalization groups," *Physical Review B*, vol. 48, no. 14, p. 10345, 1993.
 - [3] A. Klümper, A. Schadschneider, and J. Zittartz, "Matrix product ground states for one-dimensional spin-1 quantum antiferromagnets," *EPL (Europhysics Letters)*, vol. 24, no. 4, p. 293, 1993.
 - [4] A. Weichselbaum, F. Verstraete, U. Schollwöck, J. I. Cirac, and J. von Delft, "Variational matrix-product-state approach to quantum impurity models," *Physical Review B*, vol. 80, no. 16, p. 165117, 2009.
 - [5] M. B. Hastings and T. Koma, "Spectral gap and exponential decay of correlations," *Communications in mathematical physics*, vol. 265, no. 3, pp. 781–804, 2006.
 - [6] Z. Landau, U. Vazirani, and T. Vidick, "A polynomial time algorithm for the ground state of one-dimensional

- gapped local hamiltonians,” *Nature Physics*, vol. 11, no. 7, p. 566, 2015.
- [7] I. Arad, Z. Landau, U. Vazirani, and T. Vidick, “Rigorous rg algorithms and area laws for low energy eigenstates in 1d,” *Communications in Mathematical Physics*, vol. 356, no. 1, pp. 65–105, 2017.
- [8] B. Roberts, T. Vidick, and O. I. Motrunich, “Implementation of rigorous renormalization group method for ground space and low-energy states of local hamiltonians,” *Physical Review B*, vol. 96, no. 21, p. 214203, 2017.
- [9] I. P. McCulloch, “Infinite size density matrix renormalization group, revisited,” *arXiv preprint arXiv:0804.2509*, 2008.
- [10] G. Vidal, “Classical simulation of infinite-size quantum lattice systems in one spatial dimension,” *Physical review letters*, vol. 98, no. 7, p. 070201, 2007.
- [11] J. Haegeman, J. I. Cirac, T. J. Osborne, I. Pizorn, H. Verschelde, and F. Verstraete, “Time-dependent variational principle for quantum lattices,” *Physical review letters*, vol. 107, no. 7, p. 070601, 2011.
- [12] V. Zauner-Stauber, L. Vanderstraeten, M. T. Fishman, F. Verstraete, and J. Haegeman, “Variational optimization algorithms for uniform matrix product states,” *Physical Review B*, vol. 97, no. 4, p. 045145, 2018.
- [13] S. Östlund and S. Rommer, “Thermodynamic limit of density matrix renormalization,” *Physical review letters*, vol. 75, no. 19, p. 3537, 1995.
- [14] L. Vanderstraeten, J. Haegeman, and F. Verstraete, “Tangent-space methods for uniform matrix product states,” *SciPost Physics Lecture Notes*, p. 007, 2019.
- [15] Y. Huang, “Computing energy density in one dimension,” *arXiv preprint arXiv:1505.00772*, 2015.
- [16] N. Schuch and F. Verstraete, “Matrix product state approximations for infinite systems,” *arXiv preprint arXiv:1711.06559*, 2017.
- [17] M. Kliesch, D. Gross, and J. Eisert, “Matrix-product operators and states: Np-hardness and undecidability,” *Physical review letters*, vol. 113, no. 16, p. 160503, 2014.
- [18] F. Verstraete and J. I. Cirac, “Matrix product states represent ground states faithfully,” *Physical Review B*, vol. 73, no. 9, p. 094423, 2006.
- [19] D. Perez-Garcia, F. Verstraete, M. M. Wolf, and J. I. Cirac, “Matrix product state representations,” *arXiv preprint quant-ph/0608197*, 2006.
- [20] G. Vidal, “Efficient classical simulation of slightly entangled quantum computations,” *Physical review letters*, vol. 91, no. 14, p. 147902, 2003.
- [21] M. Tomamichel, R. Colbeck, and R. Renner, “A fully quantum asymptotic equipartition property,” *IEEE Transactions on information theory*, vol. 55, no. 12, pp. 5840–5847, 2009.
- [22] M. B. Hastings, “Lieb-schultz-mattis in higher dimensions,” *Physical review b*, vol. 69, no. 10, p. 104431, 2004.
- [23] M. B. Hastings, “Locality in quantum and markov dynamics on lattices and networks,” *Physical review letters*, vol. 93, no. 14, p. 140402, 2004.
- [24] M. B. Hastings, “An area law for one-dimensional quantum systems,” *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2007, no. 08, p. P08024, 2007.
- [25] I. Arad, A. Kitaev, Z. Landau, and U. Vazirani, “An area law and sub-exponential algorithm for 1d systems,” *arXiv preprint arXiv:1301.1162*, 2013.
- [26] F. G. Brandão and M. Horodecki, “Exponential decay of correlations implies area law,” *Communications in mathematical physics*, vol. 333, no. 2, pp. 761–798, 2015.
- [27] N. Schuch, M. M. Wolf, F. Verstraete, and J. I. Cirac, “Entropy scaling and simulability by matrix product states,” *Physical review letters*, vol. 100, no. 3, p. 030504, 2008.
- [28] J. Bausch, T. Cubitt, A. Lucia, and D. Perez-Garcia, “Undecidability of the spectral gap in one dimension,” *arXiv preprint arXiv:1810.01858*, 2018.
- [29] A. Uhlmann, “The transition probability in the state space of a-algebra,” *Reports on Mathematical Physics*, vol. 9, no. 2, pp. 273–279, 1976.
- [30] G. K. Brennen, S. S. Bullock, and D. P. O’Leary, “Efficient circuits for exact-universal computations with qudits,” *arXiv preprint quant-ph/0509161*, 2005.
- [31] T. S. Cubitt, D. Perez-Garcia, and M. M. Wolf, “Undecidability of the spectral gap,” *Nature*, vol. 528, no. 7581, p. 207, 2015.
- [32] J. Bausch, T. S. Cubitt, A. Lucia, D. Perez-Garcia, and M. M. Wolf, “Size-driven quantum phase transitions,” *Proceedings of the National Academy of Sciences*, vol. 115, no. 1, pp. 19–23, 2018.
- [33] M. Fannes, B. Nachtergaele, and R. F. Werner, “Finitely correlated states on quantum spin chains,” *Communications in mathematical physics*, vol. 144, no. 3, pp. 443–490, 1992.
- [34] Y. Huang, “Approximating local properties by tensor network states with constant bond dimension,” *arXiv preprint arXiv:1903.10048*, 2019.
- [35] To see this consider the simple counter example where $k = 2$, $\sigma = |\eta\rangle\langle\eta|$, $\sigma' = |\nu\rangle\langle\nu|$, with $|\eta\rangle = (|000\rangle + |111\rangle)/\sqrt{2}$, $|\nu\rangle = (|000\rangle - |111\rangle)/\sqrt{2}$. In fact here $\langle\nu|\eta\rangle = 0$. This counterexample can be generalized to apply for any k .