Commun. Math. Phys. 356, 65–105 (2017)
Digital Object Identifier (DOI) 10.1007/s00220-017-2973-z

Communications in
Mathematical
Physics

# Rigorous RG Algorithms and Area Laws for Low Energy Eigenstates in 1D

**Itai Arad**[1], **Zeph Landau**[2], **Umesh Vazirani**[2], **Thomas Vidick**[3]

[1] Centre for Quantum Technologies (CQT), National University of Singapore, Singapore, Singapore
[2] Electrical Engineering and Computer Sciences, University of California, Berkeley, CA 94720, USA
[3] Department of Computing and Mathematical Sciences, California Institute of Technology, Pasadena, CA, USA. E-mail: vidick@cms.caltech.edu

Received: 4 June 2016 / Accepted: 10 June 2017
Published online: 2 August 2017 – 

**Abstract:** One of the central challenges in the study of quantum many-body systems is the complexity of simulating them on a classical computer. A recent advance (Landau et al. in Nat Phys, 2015) gave a polynomial time algorithm to compute a succinct classical description for unique ground states of gapped 1D quantum systems. Despite this progress many questions remained unsolved, including whether there exist efficient algorithms when the ground space is degenerate (and of polynomial dimension in the system size), or for the polynomially many lowest energy states, or even whether such states admit succinct classical descriptions or area laws. In this paper we give a new algorithm, based on a rigorously justified RG type transformation, for finding low energy states for 1D Hamiltonians acting on a chain of $n$ particles. In the process we resolve some of the aforementioned open questions, including giving a polynomial time algorithm for poly$(n)$ degenerate ground spaces and an $n^{O(\log n)}$ algorithm for the poly$(n)$ lowest energy states (under a mild density condition). For these classes of systems the existence of a succinct classical description and area laws were not rigorously proved before this work. The algorithms are natural and efficient, and for the case of finding unique ground states for frustration-free Hamiltonians the running time is $\tilde{O}(nM(n))$, where $M(n)$ is the time required to multiply two $n \times n$ matrices.

## 1. Introduction

One of the central challenges in the study of quantum systems is their exponential complexity [14]: the state of a system on $n$ particles is given by a vector in an exponentially large Hilbert space, so even giving a classical description (of size polynomial in $n$) of the state is a challenge. The task is not impossible a priori, as the physically relevant states lie in a tiny corner of the Hilbert space. To be useful, the classical description of these states must support the efficient computation of expectation values of local observables. The renormalization group formalism [35] provides an approach to this problem by suggesting that physically relevant quantum states can be coarse-grained at different

length scales, thereby iteratively eliminating the "irrelevant" degrees of freedom. Ideally, by only retaining physically relevant degrees of freedom such a coarse-graining process successfully doubles the length scale while maintaining the total description size constant. This idea lies at the core of Wilson's numerical renormalization group (NRG) approach that successfully solved the Kondo problem [35]. The approach was subsequently improved by White [33,34], to obtain the famous Density Matrix Renormalization Group (DMRG) algorithm [33,34], which is widely used as a numerical heuristic for identifying the ground and low energy states of 1D systems.

Formally understanding the success of DMRG (and NRG) has been extremely challenging, as it touches on deep questions about how non-local correlations such as entanglement arise from Hamiltonians with local interactions. A major advance in our understanding of these questions came through the landmark result by Hastings [16] bounding entanglement for gapped 1D systems with unique ground state. Hasting's work was followed by a sequence of results substantially strengthening the bounds (see e.g. the review article [12]). In addition to the succinct classical description guaranteed by these results, a recent advance [22] gave a polynomial time algorithm to efficiently compute such a description. While the primary goal of this paper is to present rigorous new results about the nature of entanglement in low-energy states of 1D systems, along with efficient classical algorithms for solving such systems, we believe that the techniques we introduce also shed new light on the Renormalization Group (RG) framework.

We let $\mathcal{H} = (\mathbb{C}^d)^{\otimes n}$ denote the Hilbert space of $n$ particles of constant dimension $d$ arranged on a line. We consider the class of local Hamiltonians $H = \sum_i H_i$ where each $H_i$ is a positive semidefinite operator of norm at most 1 acting on the $i$-th and $(i+1)$-st particles. The new algorithms apply to the following classes of 1D Hamiltonians:

1. *Hamiltonians with a degenerate gapped ground space (DG):* $H$ has smallest eigenvalue $\varepsilon_0$ with associated eigenspace of dimension $r = \text{poly}(n)$, and second smallest eigenvalue $\varepsilon_1$ such that $\varepsilon_1 - \varepsilon_0 \geq \gamma$.
2. *Gapless Hamiltonians with a low density of low-energy states (LD):* The dimension of the space of all eigenvectors of $H$ with eigenvalue in the range $[\varepsilon_0, \varepsilon_0 + \eta]$, for some constant $\eta > 0$, is $r = \text{poly}(n)$.

For both classes of Hamiltonians, our results show the existence of succinct representations in the form of matrix product states (MPS; see e.g. [8,28] for background material on MPS and their use in variational algorithms) for a basis of (a good approximation to) the ground space (resp. low energy subspace) of the Hamiltonian. The bond dimension of the MPS is polynomial in $r$ and $n$ and exponential in $\gamma^{-1}$ (under assumption (DG)) or $\eta^{-1}$ (under assumption (LD)). The algorithms return these MPS representations in polynomial time in case (DG), and quasi-polynomial time in case (LD). For the special case of finding unique ground states for frustration-free Hamiltonians the algorithm is particularly efficient, with a running time of $\tilde{O}(nM(n))$, where $M(n)$ is the time required to multiply two $n \times n$ matrices.

Our assumptions are relatively standard in the literature on 1D local Hamiltonians. For an example of the first case, where the system has a spectral gap but the ground space is degenerate with polynomially bounded degeneracy, see e.g. [6,11], who consider a wide class of "natural" frustration-free local Hamiltonians in 1D for which the dimension of the ground space scales linearly with the number of particles. It is also interesting to consider the case of systems that display a vanishing gap (as the number of particles increases), while still maintaining a polynomial density of low-energy eigenstates (see for instance [19]). The assumption of polynomial density arises naturally as one considers local perturbations of gapped Hamiltonians: while conditions under which the existence

of a spectral gap remains stable are known [7], it is expected that as the perturbation reaches a certain constant critical strength the gap will slowly close; in this scenario it is reasonable (though unproven) to expect that low-lying eigenstates should remain amenable to analysis.

Our results should be understood in the context of a substantial body of prior work studying ground state entanglement in 1D systems. The techniques employed in this domain typically break down for low energy and degenerate ground states, and few results were known for these questions: Chubb and Flammia [9] extended the approach from [22] and subsequent improvements by Huang [17] to establish an efficient algorithm (and area law) for gapped Hamiltonians with a constant degeneracy in the ground space. Masanes [24] proves an area law with logarithmic correction under a strong assumption on the density of states, together with an additional assumption on the exponential decay of correlations in the ground state.

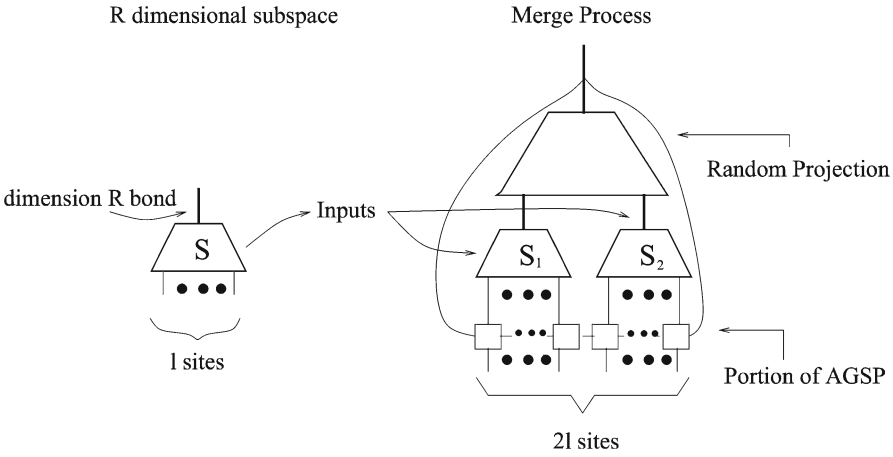
Our algorithm provides a novel perspective on the well known Renormalization Group (RG) formalism within condensed matter physics [35]. Our approach is based on the idea that if our goal is to approximate a subspace $T$ (of low energy states, say) on $n$ qubits, the algorithm can make progress by locally maintaining a small dimensional subspace $S \subset \mathcal{H}_A$ on a set $A$ consisting of $k$ particles, with the property that $T$ is close to $S \otimes \mathcal{H}_B$, where $B$ denotes the remaining $n - k$ particles. A major challenge here is measuring the quality of this partial solution. This is accomplished by a suitable generalization of the definition of a viable set introduced in [22] to the setting of a target subspace $T$, and is one of the conceptual contributions of this paper (Sect. 2). A viable set has two relevant parameters, its dimension $s$ and approximation quality $\delta$ (called the viability parameter). We introduce a number of procedures for manipulating viable sets (see Sect. 2.2). A central procedure is *random projection*. This procedure drastically cuts down the dimension of a viable set, at the expense of degrading its viability $\delta$. Our analysis shows that to a first order, the procedure of random projection achieves a trade-off between sampled dimension and approximation quality that is such that the ratio of the sampled dimension and the overlap $(1 - \delta)$ is invariant (see Lemma 5). A second procedure, *error reduction*, improves the quality of the viable set at the expense of increasing its dimension. This procedure is based on the construction of a suitable class of approximate ground state projections (AGSPs) [3,5]—spectral AGSPs—and improves the dimension-quality trade-off, at the cost of increasing the complexity of the underlying MPS representations. Setting this last cost aside, the two procedures can be combined to achieve what we call *viable set amplification*: a reduction in the dimension of a viable set, while maintaining its viability parameter unchanged (Sect. 3.1). Viable set amplification is key to both the area law proofs and the efficient algorithms given in this paper.

In addition to its dimension as a vector space, another important measure of the complexity of a viable set is the maximum bond dimension of MPS representations for its constituent vectors—this may be thought of as a proxy for the space required to actually write out a basis for the viable set. A final procedure of *bond trimming* helps us keep this complexity in check (Sect. 2.2.4). Bond trimming provides an efficient procedure to replace a viable set with another one of the same dimension and similar viability parameter, but composed of vectors with smaller bond dimension, *provided* that the target subspace $T$ has a spanning set of vectors with small bond dimension—a fact that will follow from our area laws.

The basic building block for the algorithms in this paper combines the above procedures into a process called MERGE. MERGE starts with viable sets defined on adjacent

sets of particles, and combines them into a single viable set by first taking their tensor product. This has the effect of squaring the dimension and slightly degrading the quality of the viable set. Applying viable set amplification restores both dimension and quality (for suitably chosen parameters). Thus MERGE can be used as a building block, starting with viable sets defined on individual sites and iteratively merging results along a binary tree. Since there are only $O(\log(n))$ iterations, and the bond dimension may grow exponentially with the number of iterations, this only yields an $n^{O(\log n)}$ algorithm. To achieve a polynomial time algorithm, each iteration of MERGE is modified into a procedure MERGE' which incorporates a step of bond trimming; we refer to Sect. 3.2 for further discussion.

A tensor network picture of MERGE is provided in the figure below.[1] Beginning with inputs representing subspaces of $\ell$ qubits shown on the left, the MERGE process (shown on the right) outputs a representation of a small subspace on $2\ell$ qubits. The result is a partial isometry that is reminiscent of a MERA [31,32], a more complex tensor network than MPS, which can in some cases arise as part of a renormalization procedure [13]. Completing the MERGE process into the final algorithm, however, requires an additional step of trimming which complicates the tree-like diagram shown in the figure and results in a more complex tensor network that has no direct analogue in the literature. We also note that whereas RG procedures can typically be realized as a tensor network on a binary tree (where each node represents the partial isometry associated with selecting only a small portion of the previous space), the use of the AGSP in our construction allows for selection of the small subspace that can be outside the tensor product of the previous two spaces (in this respect it may be interesting to contrast the advantage gained from AGSPs to the use of disentanglers in MERA).




A major challenge in making the above sketch effective is the construction of appropriate AGSPs. Our new *spectral AGSPs* simultaneously combine the desirable properties that had been achieved previously in different AGSPs. In particular, they are efficiently computable, have tightly controlled bond dimension (the parameter $D$) at two pre-specified cuts, and have bond dimension bounded by a polynomial in $n$ at every other cut. Achieving this requires a substantial amount of technical work, building upon the Chebyshev construction of [3], ideas about soft truncation of Hamiltonians (providing efficient means of achieving similar effects to the hard truncation studied in e.g. [4]), a

[1] We are grateful to Christopher T. Chubb for originally suggesting these pictures to us.

series expansion of $e^{-\beta H}$ known as the *cluster expansion* [15,20], as well as a recent nontrivial efficient encoding of the resulting operator due to [25]. The constructions of spectral AGSP appear in Sect. 4 (non-efficient constructions) and Sect. 5 (efficient constructions).

Our new algorithms could potentially be made very efficient. The main bottlenecks are the complexity of the AGSP and the MPS bond dimension that must be maintained. In the case of a frustration-free Hamiltonian with unique ground state we obtain a running time of $O(2^{O(1/\gamma^2)} n^{1+o(1)} \mathrm{M}(n))$, where $\mathrm{M}(n)$ is the matrix multiplication time. This has an exponentially better scaling in terms of the spectral gap $\gamma$ (due to avoidance of the $\varepsilon$-net argument) and saves a factor of $n/\log n$ (due to the logarithmic, instead of linear, number of iterations) as compared to an algorithm for the same problem considered in [18]. We speculate that it might further be possible to limit the bond dimension of all MPS considered to $n^{o(1)}$ (instead of $n^{1+o(1)}$ currently), which, if true, would imply a nearly-linear time $O(n^{1+o(1)})$ algorithm.

Subsequently to the completion of our work, a heuristic variant of the algorithm described in this paper has been implemented numerically [27]. Although this initial implementation typically suffers from a $\sim 5-10\times$ slowdown compared to the well-established DMRG, it provides encouragingly accurate results, matching those of DMRG in "easy" cases, but also sometimes outperforming DMRG, e.g. in cases where the ground space degeneracy is high (linear in system size) or for some critical systems.

*Organization.* The remainder of the paper is organized as follows. In Sect. 2, we start by introducing viable sets, and provide a comprehensive set of procedures to work them; these procedures form the building blocks of our area laws and algorithms. With these procedures in place, in Sect. 3, we provide an overview of our proof technique; this section may be the best place to start reading the paper for a reader new to our results. The following three sections are devoted to a formal fleshing out of our results. In Sect. 4, we prove our area laws by showing the existence of good AGSP constructions. In Sect. 5, we provide efficient analogues of these AGSP constructions, which are employed in Sect. 6 to derive our efficient algorithms. We conclude in Sect. 7 with a discussion of our results and possible improvements.

## 2. Viable Sets

Our approach starts with the idea that the challenge of finding a solution—a low-energy state—within a Hilbert Space $\mathcal{H}$ of exponential size can be approached by starting with "partial solutions" on small subsystems, and gradually combining those into "solutions" defined on larger and larger subsystems. To implement this approach we need a formal notion of partial solution, as well as techniques for working with them. This is done in the next few subsections where we introduce *viable sets* to capture "partial solutions", and describe procedures to efficiently work with such viable sets.

*2.1. Definition and basic properties.* Given a subset $A$ of particles, we may decompose the full Hilbert space of the system as a tensor product $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$, where $\mathcal{H}_A$ is the Hilbert space associated with particles in $A$ and $\mathcal{H}_B$ is the Hilbert space associated with the remaining particles in the system. Our ultimate goal is to compute (an approximation to) some subspace $T \subset \mathcal{H}$. Towards this we wish to measure partial progress made while processing only particles in the subset $A$. This can be expressed through the sub-goal of

finding a subspace $S \subset \mathcal{H}_A$ with the guarantee that $S_{ext} := S \otimes \mathcal{H}_B$ contains $T$. Since we need to allow the possibility for approximation errors, we are led to the following definition:

**Definition 1** (*Viable set*). Given $0 \leq \delta \leq 1$ and a subspace $T \subseteq \mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$, a subspace $S \subseteq \mathcal{H}_A$ is $\delta$-*viable* for $T$ if

$$P_T P_{S_{ext}} P_T \geq (1-\delta) P_T, \tag{1}$$

where $S_{ext} := S \otimes \mathcal{H}_B$ and $P_T$ (resp. $P_{S_{ext}}$) is the orthogonal projection onto $T$ (resp. $S_{ext}$). We refer to $\delta$ as the *viability* of the set, and $\mu = 1 - \delta$ as its *overlap*.

This definition captures the notion that a reasonable approximation of $T$ can be found within the subspace $S \otimes \mathcal{H}_B$. It generalizes the definition of a viable set from [22], which was specialized to the case where $T$ is a one-dimensional subspace containing a unique ground state. In [9,17] the definition was extended to handle degenerate ground spaces by explicitly requiring that the viable set support orthogonal vectors that are good approximations to orthogonal ground states. Here we avoid making any direct reference to a basis, or families of orthogonal vectors, and instead work directly with subspaces.

While the notion of viable set is quite intuitive for small $\delta$, our arguments also involve viable sets with parameter $\delta$ close to 1 (alternatively, $\mu = 1 - \delta$ close to 0, where $\mu$ is a parameter we will refer to as the *overlap* of the viable set), a regime where there is less intuition. A helpful interpretation of the definition is that it formalizes the fact that for a viable set $S$, the image of the unit ball of $S_{ext}$ when projected to $T$ contains the ball of radius $(1 - \delta)$.

**Lemma 1.** *If $S$ is $\delta$-viable for $T$ for some $0 \leq \delta < 1$ then for every $|t\rangle \in T$ of unit norm, there exists an $|s\rangle \in S_{ext}$ such that $P_T|s\rangle = |t\rangle$ and $\||s\rangle\| \leq \frac{1}{1-\delta}$.*

The proof of Lemma 1 follows directly from the following general operator facts:

**Lemma 2.** *1. If $X$ and $Y$ are positive operators and $X \geq Y$ then range $(Y) \subset$ range $(X)$.*

*2. If $PQP \geq cP$ for projections $P$, $Q$ and $c > 0$ then for every $v \in$ range$(P)$ of unit norm, there exists $w \in$ range$(Q)$, $||w|| = 1$ such that $Pw = c_v v$ for some constant $c_v$ with $|c_v| \geq c$.*

*Proof.* For 1., suppose $y \in \text{range}(Y)$ and let $y = x + x_\perp$, $x \in \text{range}(X)$, $x_\perp \perp \text{range}(X)$ be the orthogonal decomposition. Since $\langle X x_\perp, x_\perp \rangle = 0$ it follows that $\langle Y x_\perp, x_\perp \rangle = 0$ and thus $x_\perp \perp \text{range}(Y)$ as well and hence $x_\perp = 0$ and $y = x \in \text{range}(X)$.

For 2., it follows from 1. that if $PQP \geq cP$ then for any $v \in \text{range}(P)$ there exists an $r \in \text{range}(P)$ such that $PQPr = PQr = v$. So then $\langle PQPr, r \rangle \geq c\langle Pr, r \rangle = c||r||^2$. But $\langle PQPr, r \rangle = \langle v, r \rangle \leq ||r||||v||$. Putting these two inequalities together along with the assumption that $||v|| = 1$ yields $||r|| \leq 1/c$. □

We introduce a notion of proximity between subspaces:

**Definition 2** (*Closeness*). For $0 \leq \delta \leq 1$, a subspace $T$ is $\delta$-*close* to a subspace $T'$ if

$$P_{T'} P_T P_{T'} \geq (1-\delta) P_{T'},$$

where $P_T$ and $P_{T'}$ are the orthogonal projections on $T$ and $T'$ respectively. We say that $T$ and $T'$ are *mutually* $\delta$-*close* if each is $\delta$-close to the other, and denote by $\angle_m(T, T')$ the smallest $\delta$ such that $T$, $T'$ are mutually $\delta$-close.

Closeness of subspaces is approximately transitive in the following sense:

**Lemma 3** (Robustness). *If $T$ is $\delta$-close to $T'$ and $T'$ is $\delta'$-close to $T''$ then $T$ is $2(\delta+\delta')$-close to $T''$. Consequently if $S$ is $\delta$-viable for $T$ and $T$ is $\delta'$-close to $T'$ then $S$ is $2(\delta+\delta')$-viable for $T'$.*

*Proof.* Notice that $P_A P_B P_A \geq (1-\delta) P_A$ is equivalent to the statement that $\|P_B|a\rangle\|^2 \geq (1-\delta)$ for all $|a\rangle \in A$ with $\||a\rangle\| = 1$. It follows for $|t''\rangle \in T'$ of unit norm, $|t'\rangle = P_{T'}|t''\rangle$ has the property that $\||t'\rangle\|^2 \geq (1-\delta')$ and thus $\||t'\rangle - |t\rangle\| \leq \sqrt{\delta'}$. Similarly $|t\rangle = P_T|t'\rangle$ has the property that $\||t\rangle\|^2 \geq \||t'\rangle\|^2(1-\delta)$ and thus $\||t'\rangle - |t\rangle\| \leq \||t'\rangle\|\sqrt{\delta}$. By the triangle inequality, $\||t''\rangle - |t\rangle\| \leq \sqrt{\delta'} + \sqrt{\delta}$ and since $|t\rangle \in T$, this implies that the distance between $|t''\rangle$ and $T$ is at most $\sqrt{\nu} + \sqrt{\delta}$, i.e. $\|P_{S_{ext}}|t\rangle\|^2 \geq 1 - (\sqrt{\nu} + \sqrt{\delta})^2 \geq 1 - 2(\nu+\delta)$. As mentioned, this last statement is equivalent to $T$ being a $2(\nu+\delta)$ close to $T''$. □

*2.2. Procedures.* We introduce a set of procedures that can be performed on viable sets. These procedures will allow us to build viable sets on larger and larger subsystems, while keeping the complexity and size of the viable sets small. They will serve as the core operations for both our area law proofs and our algorithms.

*2.2.1. Tensoring.* The next lemma summarizes the effect of tensoring two viable sets supported on disjoint subsystems.

**Lemma 4** (Tensoring). *Suppose $S_1$, $S_2$ are $\delta_1$-viable and $\delta_2$-viable for $T$ respectively, defined on disjoint subsystems. Then the set $S := S_1 \otimes S_2$ is $(\delta_1+\delta_2)$-viable for $T$.*

*Proof.* Since $S_1, S_2$ are defined on disjoint subsystems, it follows that $P_{S^{(ext)}} = P_{S_1^{(ext)}} P_{S_2^{(ext)}}$, and so

$$P_T P_{S^{(ext)}} P_T = P_T P_{S_1^{(ext)}} P_{S_2^{(ext)}} P_T = P_T P_{S_1^{(ext)}} P_T - P_T P_{S_1^{(ext)}} \big(\mathbb{1} - P_{S_2^{(ext)}}\big) P_T.$$

The definition of a viable set implies that $P_T P_{S_1^{(ext)}} P_T \geq (1-\delta_1) P_T$. In addition,

$$P_T P_{S_1^{(ext)}} \big(\mathbb{1} - P_{S_2^{(ext)}}\big) P_T \leq P_T \big(\mathbb{1} - P_{S_2^{(ext)}}\big) P_T \leq \delta_2 P_T.$$

Therefore, $P_T P_{S^{(ext)}} P_T \geq (1-\delta_1-\delta_2) P_T$. □

*2.2.2. Random sampling.* The following lemma establishes how viability of a set is affected when sampling a random subset.

**Lemma 5** (Random sampling). *Let $T \subseteq \mathcal{H} = \mathcal{H}_L \otimes \mathcal{H}_M \otimes \mathcal{H}_R$ be an $r$-dimensional subspace, and $W$ a $q$-dimensional subspace of $\mathcal{H}_M$ that is $\delta$-viable for $T$. Then a random $s$-dimensional subspace $W'$ of $W$ is $(1-\delta')$-viable for $T$ with probability $1-\eta$, where*

$$\delta' = \frac{(1-\delta)}{8}\frac{s}{q} \quad \text{and} \quad \eta = \left(1 + 4\sqrt{\frac{q}{(1-\delta)s}}\right)^r q e^{-s/16}.$$

*Proof.* Let $|v\rangle$ in $T$ such that $\||v\rangle\| = 1$, and $|w\rangle = P_{W_{ext}}|v\rangle \in W_{ext}$. Using that $W$ is $\delta$-viable for $T$ it follows that $\||w\rangle\|^2 \geq 1-\delta$. Since $W'_{ext} \subseteq W_{ext}$, $P_{W'_{ext}}|v\rangle = P_{W'_{ext}}|w\rangle$. By a standard concentration argument based on the Johnson-Lindenstrauss lemma (see e.g. [10, Theorem 2.1]) it holds that $\|P_{W'_{ext}}|v\rangle\|^2 \geq (1-\delta)\frac{s}{2q}$ with probability at least $1 - qe^{-s/16}$. Let $\nu = \sqrt{(1-\delta)s/8q}$. By a volume argument (see e.g. [29, Lemma 5.2]), there exists a subset $S$ of the Euclidean unit ball of $T$ such that $|S| \leq (1+2/\nu)^r$ and for any unit $|t\rangle \in T$, there is an $|v\rangle \in S$ such that $\||s\rangle - |t\rangle\| \leq \nu$. Applying the preceding argument to each $|v\rangle$ in the net, by the union bound the choice of $\eta$ made in the theorem is with probability at least $1-\eta$, $\|P_{W'_{ext}}|v\rangle\|^2 \geq (1-\delta)s/(2q)$ for all $|v\rangle$ in the net; hence $\|P_{W'_{ext}}|v\rangle\|^2 \geq (1-\delta)s/(8q)$ for all $|v\rangle$ in the unit ball of $T$. □

*2.2.3. Error reduction using approximate ground state projections.* We address the question of how to improve the viability parameter $\delta$ for a given viable set. In previous work this question was addressed for the case of the target space $T$ being one dimensional by introducing the key tool of Approximate Ground State Projections (AGSPs) [3,5]. AGSPs have been used in the context of proofs of the 1D area law for Hamiltonians with a unique ground state as well as in algorithms for finding the ground state of a gapped 1D system [22].

Whereas in previous works an AGSP was primarily constructed to approximate the projector on a *unique* ground state, here our main focus is on the case of a degenerate ground space and low-energy states. We therefore introduce a more general definition of an AGSP as a local operator that increases the norm of eigenvectors in the low part of the spectrum of $H$, while decreasing the norm of eigenvectors in the high energy part of the spectrum. We refer to this object as a *spectral AGSP*.

**Definition 3** (*Spectral AGSP*). Given $\mathcal{H} = \mathcal{H}_L \otimes \mathcal{H}_M \otimes \mathcal{H}_R$, $H$ a Hamiltonian on $\mathcal{H}$ and $\eta_0 < \eta_1$, a positive semidefinite operator $K$ on $\mathcal{H}$ is a $(D, \Delta)$-spectral AGSP for $(H, \eta_0, \eta_1)$ if the following conditions hold:

- $K$ has a decomposition $K = \sum_{i=1}^{D^2} L_i \otimes A_i \otimes R_i$,
- $H$ and $K$ have the same eigenvectors,
- Eigenvalues of $H$ smaller than $\eta_0$ correspond to eigenvalues of $K$ that are larger than or equal to 1,
- Eigenvalues of $H$ larger than $\eta_1$ correspond to eigenvalues of $K$ that are smaller than $\sqrt{\Delta}$.

The advantage of an AGSP, compared to an exact projection operator, lies in the fact that one can often construct a much more *local* operator, i.e., an operator with a much smaller Schmidt rank compared to the exact projector. The existence of an AGSP of small Schmidt rank which greatly shrinks the high energy part of the spectrum can be viewed as a strong characterization of the locality properties of the low-energy space. A favorable scaling between these two competing aspects (in the case of unique ground states) was the key feature in recent proofs of the 1D area law [3,5] via the bootstrapping lemma. The following lemma establishes a lower bound on the quantitative improvement in viability that a spectral AGSP can achieve on a viable set.

**Lemma 6** (Error reduction—Spectral AGSP). *Let $\mathcal{H} = \mathcal{H}_L \otimes \mathcal{H}_M \otimes \mathcal{H}_R$, $H$ a Hamiltonian on $\mathcal{H}$, $\eta_0 < \eta_1$, and $K = \sum_{i=1}^{D^2} L_i \otimes A_i \otimes R_i$ a $(D, \Delta)$-spectral AGSP for $(H, \eta_0, \eta_1)$ where $H$ has ground state energy $\varepsilon_0$ and has no eigenvalues in the interval*

*$(\eta_0, \eta_1)$. Let $S \subseteq \mathcal{H}_M$ be a $\delta$-viable set for $T = H_{[\varepsilon_0,\eta_0]}$ of dimension $s$. Then the space $V = \mathrm{Span}\{A_i S : 1 \le i \le D^2\}$ has dimension at most $D^2 s$ and is $\delta'$-viable for $T$ with*

$$\delta' = \frac{\Delta}{(1-\delta)^2}.$$

*Proof.* The bound on the dimension of $V$ is straightforward. To show $V$ is $\delta'$-viable for $T$, begin with an arbitrary unit norm vector $|v\rangle \in T$. Set $|v'\rangle = \frac{1}{\|K^{-1}|v\rangle\|} K^{-1}|v\rangle$, where $K^{-1}$ is the pseudo-inverse. Then $|v'\rangle$ is also an element of $T$. Since $S$ is $\delta$-viable for $T$, applying Lemma 1 there exists an $|u\rangle \in \mathcal{H}_L \otimes S \otimes \mathcal{H}_R$ whose projection onto $T$ is, up to scaling, precisely $|v'\rangle$; thus $|u\rangle = \alpha|v'\rangle + \sqrt{1-\alpha^2}|v^\perp\rangle$ for some $\alpha \ge 1-\delta$ and unit $|v^\perp\rangle$ that is orthogonal to $T$. In particular $|v^\perp\rangle$ is supported on the span of all eigenvectors of $H$ with eigenvalue outside of $[\varepsilon_0, \eta_1) = [\varepsilon_0, \eta_0) \cup [\eta_0, \eta_1)$ and thus by the property of $K$, $\|K|v^\perp\rangle\|^2 \le \Delta$.

Applying $K$ to $|u\rangle$ yields $K|u\rangle = \alpha'|v\rangle + K|v^\perp\rangle$ with $\alpha' = \alpha \frac{1}{\|K^{-1}|v\rangle\|} \ge \alpha$ (since $|v\rangle$ is supported on eigenvectors of $K$ with corresponding eigenvalue at least 1). Thus

$$\begin{aligned} \left|\left\langle \frac{Ku}{\|K|u\rangle\|} \middle| v \right\rangle\right|^2 &\ge \frac{\alpha'^2}{\alpha'^2 + (1-\alpha'^2)\Delta} \\ &\ge 1 - \frac{1}{(1-\delta)^2}\Delta. \end{aligned}$$

□

*2.2.4. Complexity reduction using trimming.* For a viable set to be efficiently represented it must not only have small dimension but also a basis of states that can be efficiently described, say by polynomial-bond matrix product states. A natural question is, assuming that the target subspace $T$ has a basis of vectors of small bond dimension, whether it is possible to efficiently "trim" any sufficiently good viable set for $T$ into another almost-as-good viable set specified by vectors with comparably small bond dimension.

To achieve this goal we introduce a modified trimming procedure to that of [22]. There the trimming procedure is based on the observation that given a good approximation to a target vector $|v\rangle$ of low bond dimension, trimming the approximating vector by dropping Schmidt vectors associated with the smallest Schmidt coefficients at each cut yields an almost-as-good approximation to $|v\rangle$ with lower bond dimension. In the present scenario the approximating vector is not known: instead we are given a basis for a subspace that contains the approximating vector. A natural idea would be to trim the MPS representations for the basis vectors in a way that guarantees that $|v\rangle$ is still closely approximated by some vector in the span of the resulting set. However, it is not clear if independently trimming each of the basis vectors, as done in [22], works—indeed, the basis vectors themselves could a priori have a very flat distribution of Schmidt coefficients, so that trimming could induce large changes.

We provide a modified procedure which starts with a basis for the viable set and trims the basis vectors collectively at each cut, from the leftmost to the rightmost, as follows (informally): for each cut, project each element of the basis onto the span of the left Schmidt vectors of *any* basis element that is associated with a large Schmidt coefficient.

**Definition 4** (*Trimming*). Let $S \subseteq \mathcal{H}_A$ be a $\delta$-viable set for $T \subseteq \mathcal{H}_A \otimes \mathcal{H}_B$ specified by an orthonormal basis $\{|u_i\rangle,\ i = 1, \ldots s\}$. Suppose $\mathcal{H}_A = \mathcal{H}_A^1 \otimes \cdots \otimes \mathcal{H}_A^\ell$ for some

$\ell \geq 2$. Let $|\psi\rangle = \sum_i |u_i\rangle|i\rangle \in \mathcal{H}_A \otimes \mathbb{C}^s$. For $j$ from 1 to $(\ell-1)$ define $P^j_{\geq\xi}$ inductively as the projection on the subspace of $P^{j-1}_{\geq\xi} \otimes \mathbb{1}_{\mathcal{H}_A^j}$ spanned by the left Schmidt vectors of $P^{j-1}_{\geq\xi} \otimes \mathbb{1}_{\mathcal{H}_A^j \cdots \mathcal{H}_A^\ell} \otimes \mathbb{1}_{\mathbb{C}^s} |\psi\rangle$ across the cut $(j : j+1)$ with associated Schmidt coefficient at least $\xi$.[2] Then the *$\xi$-trimmed set* is

$$\mathrm{Trim}_\xi(S) := \mathrm{Span}\left\{ \left( (P^1_{\geq\xi} \otimes \mathbb{1}_{\mathcal{H}_A^2 \otimes \cdots \otimes \mathcal{H}_A^k}) \cdots (P^{\ell-1}_{\geq\xi} \otimes \mathbb{1}_{\mathcal{H}_A^\ell}) \right) |u_i\rangle,\ i = 1, \ldots, s \right\}. \quad (2)$$

With this notion of trimming, we show that if a set $S$ is a good viable set for a set $T$ whose elements are guaranteed to have low bond dimension then the result of trimming the set $S$ does not degrade the quality of the viable set too much.

**Lemma 7** (Trimming). *Let $S \subseteq \mathcal{H}_A$ be a $\delta$-viable set of dimension $s$ for $T \subseteq \mathcal{H}_A \otimes \mathcal{H}_B$. Suppose $\mathcal{H}_A = \mathcal{H}_A^1 \otimes \cdots \otimes \mathcal{H}_A^\ell$ for some $\ell \geq 2$. Let $b$ be an upper bound on the Schmidt rank of any vector in $T$ across any cut $(j : j+1)$ for $j = 1, \ldots, \ell-1$. Then the $\xi$-trimmed set $\mathrm{Trim}_\xi(S)$ is a $\delta'$-viable set for $T$ for $\delta' \leq \delta + \sqrt{\ell b s}\xi$.*

*Furthermore, a spanning set for $\mathrm{Trim}_\xi(S)$ containing at most $s$ vectors of Schmidt rank at most $s\xi^{-2}$ across any cut can be computed in time $O(\ell \mathrm{M}(dsq))$, where $q$ is an upper bound on the bond dimension of MPS representations for a basis of $S$ and $\mathrm{M}(\cdot)$ denotes matrix multiplication time.*

*Proof.* Let $\{|u_i\rangle,\ i = 1, \ldots, s\}$ denote an orthonormal basis for $S$, and $|v\rangle \in T$ a unit vector. Let $|u\rangle = \sum_i \mu_i |a_i\rangle|b_i\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B$ be a unit vector such that $|\langle u|v\rangle|^2 \geq 1 - \delta$. For $j = 0, \ldots, \ell$, let

$$|u'_j\rangle = (P^1_{\geq\xi} \otimes \mathbb{1}_{\mathcal{H}_A^2 \cdots \mathcal{H}_A^\ell} \otimes \mathbb{1}_{\mathcal{H}_B}) \cdots (P^j_{\geq\xi} \otimes \mathbb{1}_{\mathcal{H}_A^{j+1} \cdots \mathcal{H}_A^\ell} \otimes \mathbb{1}_{\mathcal{H}_B}) |u\rangle,$$

and for $i \in \{1, \ldots, s\}$,

$$|a^j_i\rangle = (P^1_{\geq\xi} \otimes \mathbb{1}_{\mathcal{H}_A^2 \cdots \mathcal{H}_A^\ell} \otimes \mathbb{1}_{\mathcal{H}_B}) \cdots (P^j_{\geq\xi} \otimes \mathbb{1}_{\mathcal{H}_A^{j+1} \cdots \mathcal{H}_A^\ell} \otimes \mathbb{1}_{\mathcal{H}_B}) |a_i\rangle.$$

By definition of the $P^j_{\geq\xi}$ (Definition 4), the Schmidt coefficients of the vector

$$(P^1_{\geq\xi} \cdots P^{j-1}_{\geq\xi} (\mathbb{1} - P^j_{\geq\xi}) \otimes \mathbb{1}) |\psi\rangle,$$

where $|\psi\rangle = \sum |a_i\rangle|i\rangle$, across the cut $(j,\ j+1)$ are all at most $\xi$. Since acting with a local projection (here, $|i\rangle\langle i|$ on $\mathcal{H}_B$) cannot increase the largest Schmidt coefficient, the same holds of the vector $((\mathbb{1} - P^j_{\geq\xi}) \otimes \mathbb{1})|a^{j-1}_i\rangle$. Based on these observations we may upper bound, for any $i$, $j$, and unit $|c\rangle \in \mathcal{H}_A^1 \otimes \cdots \otimes \mathcal{H}_A^j$ and $|d\rangle \in \mathcal{H}_A^{j+1} \otimes \cdots \otimes \mathcal{H}_A^\ell \otimes \mathcal{H}_B$,

$$\left| \langle a^{j-1}_i | \langle v_i | ((\mathbb{1} - P^j_{\geq\xi}) \otimes \mathbb{1}_{\mathcal{H}_A^{j+1} \otimes \cdots \otimes \mathcal{H}_A^\ell} \otimes \mathbb{1}_{\mathcal{H}_B}) |c\rangle |d\rangle \right| \leq \xi,$$

where the inequality follows since we are taking the inner product of a vector with largest Schmidt coefficient at most $\xi$ with another vector of Schmidt rank 1. Using the promised bound on the Schmidt rank of $|v\rangle$ we deduce

[2] Note that we do not re-normalize vectors.

$$\begin{aligned}\left|(\langle u'_j| - \langle u'_{j-1}|)|v\rangle\right| &= \left|\langle a_i^{j-1}|\langle b_i|((\mathbb{1} - P^j_{\geq\xi}) \otimes \mathbb{1}|v\rangle\right| \\ &\leq \xi\sqrt{bs}\left\|(\mathbb{1} - P^j_{\geq\xi}) \otimes \mathbb{1}|v\rangle\right\|.\end{aligned}$$

Using a telescopic sum, and orthogonality of the projections $(\mathbb{1} - P^j_{\geq\xi}) \otimes \mathbb{1}$ for different values of $j$, we get

$$\begin{aligned}\left|(\langle u'_1| - \langle u'_\ell|)|v\rangle\right|^2 &\leq \xi^2 bs\Big(\sum_{j=1}^{\ell}\left\|(\mathbb{1} - P^j_{\geq\xi}) \otimes \mathbb{1}|v\rangle\right\|\Big)^2 \\ &\leq \xi^2 \ell bs,\end{aligned}$$

and the claimed bound on $\delta'$ follows.

For the "furthermore" part, note that $|\psi\rangle$ has at most $s/\xi^2$ Schmidt coefficients larger than $\xi$ across any cut $(j : j+1)$. Thus each $P^j_{\geq\xi}$ has rank at most $s/\xi^2$, so that its application reduces the Schmidt rank across the cut $(j : j+1)$ to at most $s/\xi^2$, while not increasing it to a larger value at any of the previously considered cuts. The left Schmidt vectors of

$$(P^1_{\geq\xi} \otimes \mathbb{1}_{\mathcal{H}_A^2\otimes\cdots\otimes\mathcal{H}_A^\ell}) \cdots (P^{\ell-1}_{\geq\xi} \otimes \mathbb{1}_{\mathcal{H}_A^\ell})\big)|\psi\rangle$$

across the cut specified by the division $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ form a spanning set for $\mathrm{Trim}_\xi(S)$.

In order to compute canonical MPS representations for a basis of $\mathrm{Trim}_\xi(S)$ we first create an MPS representation for $|\psi\rangle$ and reduce it to canonical form (we refer to e.g. the survey [30] for a discussion of basic operations on MPS and their computational efficiency). This costs $O(\ell \mathrm{M}(dsq))$ operations, where $\mathrm{M}(\cdot)$ is matrix multiplication time, and $\mathrm{M}(dsq)$ is the time required to perform required basic operations on tensors of bond dimension $O(dsq)$, such as singular value decompositions. Proceeding from the cut $(\ell - 1, \ell)$ to the $(1, 2)$ cut from right to left, we then set the coefficients of the diagonal tensor matrices $\Lambda_j$ from the MPS representation that are smaller than $\xi$ to zero. The resulting re-normalized state is automatically given in canonical MPS form, and a spanning set for $\mathrm{Trim}_\xi(S)$ can be obtained by cutting the last bond. □

## 3. Overview

In this section we provide an outline of how the procedures introduced in the two previous sections can be put together to yield area laws and efficient algorithms. Our results hinge on our ability to construct AGSPs with good trade-offs between $D$ and $\Delta$. Our goal in this section is to provide a high level picture of how the pieces fit together. For this we assume a very simple, approximate picture of an AGSP. The rigorous results are more intricate, and will be described in the remaining sections of the paper.

Let $H$ be a Hamiltonian with ground state energy $\varepsilon_0$ and no eigenvalues in the interval $(\eta_0, \eta_1)$. We assume that $H$ comes with an associated spectral AGSP $K$ (Definition 3) that satisfies the conditions of Lemma 6. We further assume that the parameters $(D, \Delta)$ associated with $K$ satisfy a sufficiently good trade-off between $D$ and $\Delta$.[3] Our goal is to approximate the low-energy subspace $T = H_{[\varepsilon_0,\eta_0]}$, assumed to be of polynomially bounded dimension.

[3] For our purposes, a tradeoff of the form $D^c\Delta < \frac{1}{2}$, for a large enough constant $c$, will suffice; we refer to later sections for concrete parameters.

*3.1. Viable set amplification and area laws.* As a first step we compose the procedures of random sampling (Lemma 5) and error reduction (Lemma 6) to obtain a procedure that improves the quality of a viable set without increasing its dimension:

**Viable Set Amplification:**
Given is a $\delta$-viable set $W$ of dimension $q$.

1. Generate a **random sample**, as in Lemma 5, to obtain $S \subset W$ of dimension $s$ with viability parameter $\delta' > \delta$.
2. Apply **error reduction** to $S$, as in Lemma 6, using the AGSP $K$, to produce a $\delta$-viable set $W'$ of dimension $q'$.

For a $\delta$-viable set $W$, we refer to $\mu = 1 - \delta$ as its overlap. Random sampling reduces the dimension of the viable set but also proportionately reduces its overlap. The second step (AGSP) increases the overlap at the cost of a comparatively smaller increase in dimension — a favorable trade-off due to the favorable $D - \Delta$ trade-off of the AGSP. With proper setting of parameters, the viable set amplification procedure above reduces the dimension of the viable set while leaving the overlap (and $\delta$) unchanged, as long as the viable set dimension $q > q_0$, for some $q_0$ determined by $\delta$ as well as the parameter $D$ of the AGSP (itself related to parameters of the initial Hamiltonian, including the spectral gap above the low-energy space $T$).

Reasoning by contradiction, the argument implies the existence of a $\delta$-viable set $W_0$ for $T$ of dimension at most $q_0$. The existence of such a $W_0$ in turn implies that any element of $T$ has a $\delta$-approximation by a vector with entanglement rank no larger than $q_0$. An area law follows easily using standard amplification arguments; we give the details in Sect. 4.

*3.2. Merge process and algorithms.* In the argument described in the previous section the parameters were chosen such that a $\delta$-viable set of dimension $q$ was "amplified" to a $\delta$-viable set of dimension $q' < q$. With a slightly more demanding choice of parameters viable set amplification can be made to reduce both the dimension $q \to q' = \sqrt{q}$ and the viability parameter $\delta \to \delta' = \frac{\delta}{2}$. This only requires a slightly more stringent condition on the $D - \Delta$ trade-off provided by the underlying AGSP.

We now explain how viable set amplification can be folded within a larger procedure that we call MERGE. Assume given a decomposition $\mathcal{H} = \mathcal{H}_L \otimes (\mathcal{H}_A \otimes \mathcal{H}_B) \otimes \mathcal{H}_R$ of the $n$-particle Hilbert space. MERGE starts with two viable sets $V_1 \subseteq \mathcal{H}_A$ and $V_2 \subseteq \mathcal{H}_B$ and returns a viable set $V \subseteq \mathcal{H}_A \otimes \mathcal{H}_B$. It does so in a way such that all parameters of the viable set $V$, namely the viability $\delta$, the dimension, and its description complexity, are comparable to those of the original two sets. We proceed to describe MERGE; for expository purposes we set aside considerations on the complexity of representing elements of the viable sets (these will be made formal in subsequent sections).

**MERGE**:
Given are two $\delta$-viable sets $V_1 \subset \mathcal{H}_A$ and $V_2 \subset \mathcal{H}_B$ of dimension $q$.

1. **Tensor** the two sets, as in Lemma 4, to obtain a $2\delta$-viable set $W = V_1 \otimes V_2$ of dimension at most $q^2$.
2. Perform **viable set amplification** to yield a $\delta$-viable set $V \subset \mathcal{H}_1 \otimes \mathcal{H}_2$ of dimension at most $q$.

Our algorithm starts with (easily generated) viable sets defined over small subsets of particles, and iterates MERGE in a tree-like fashion to eventually generate a single viable

set defined over the entire space. With this final viable set in hand, it is not difficult to find low-energy states within the viable set, *provided* we are able to describe its elements using low-complexity representations (e.g. low bond dimension matrix product states). This will not be the case unless explicit constraints are enforced on the complexity of the operators used in the error reduction step of viable set amplification, where the complexity can blow up rapidly due to the application of the AGSP $K$.

To maintain the desired low complexity MPS representations and complete the algorithm we make two modifications to MERGE. The first is within the AGSP construction, where a procedure of soft truncation (Sect. 5.1) leads to the operators used in error reduction having matrix product operator (MPO) representations with polynomial bond dimension. Since these operators are applied a large number of times, however, the complexity of the MPS representations manipulated could still increase to super-polynomial. In order to keep that complexity under control we perform a second modification, which decomposes the viable set amplification procedure into smaller steps of viable set amplification followed by a trimming procedure. The result is the following modified procedure:

**MERGE' (informal)**:
Given are $\delta$-viable sets $V_1 \subset \mathcal{H}_A$ and $V_2 \subset \mathcal{H}_B$ of dimension $q$, each specified by MPS with polynomial bond dimension.

1. **Tensor** the two sets, as in Lemma 4, to obtain a $2\delta$-viable set $W = V_1 \otimes V_2$ of dimension at most $q^2$.
2. Perform **viable set amplification** followed by **trimming** on the viable set to produce a $\delta$-viable set of smaller dimension, again specified by MPS with polynomial bond dimension. Repeat this step until the resulting $\delta$-viable set has dimension $q$.

We note that the correctness of the trimming procedure employed in the second step of MERGE' relies on the area law established using MERGE, as described in the previous section.

The overview given in this section provides an accurate outline of how viable sets can be put together into an efficient algorithm for mapping out the low-energy subspace of a local Hamiltonian. The most important technical ingredient that we have set aside so far is the creation of AGSP with the required parameter trade-off between $D$ and $\Delta$. In Sect. 4 we establish *existence* of the desired AGSP, which lets us formally implement the first part of our results, area laws for local Hamiltonians satisfying assumptions (DG) and (LD) described in the introduction. In order to obtain algorithms we will need to make the AGSP constructions *efficient*: this is achieved in Sect. 5, with the resulting algorithms described in Sect. 6.

## 4. Area Laws

In this section we establish area laws for the ground space and low-energy space of Hamiltonians satisfying assumptions (DG) and (LD) respectively. The proofs are based on the non-constructive bootstrapping argument outlined in Sect. 3.1, which relies on a sufficiently good construction of AGSP. We first review the general Chebyshev-based AGSP construction from [3] in Sect. 4.1. We introduce a scheme of hard truncation for the norm of a Hamiltonian in Sect. 4.2. In Sect. 4.3 we apply the Chebyshev construction to the truncated Hamiltonian to obtain our main AGSP constructions. The AGSP are applied to the proof of the area law under assumption (DG) in Sect. 4.4 and assumption (LD) in Sect. 4.5.

*4.1. The Chebyshev polynomial AGSP.* Given a Hamiltonian $H$ with ground energy $\varepsilon_0$ and a gap parameter $\gamma$, a natural way to define an approximate ground state projection is by setting $K := P_k(H)$, where $P_k$ is a polynomial that satisfies $P_k(\varepsilon_0) = 1$ and $|P_k(x)|^2 \leq \Delta$ for every $\varepsilon_0 + \gamma \leq x \leq \|H\|$. Clearly, $K$ preserves the ground space and reduces the norm of any eigenstate $|\phi\rangle$ of $H$ with eigenvalue at least $\varepsilon_0 + \gamma$ as $\|K|\phi\rangle\|^2 \leq \Delta$. Moreover, the lower the degree of $P_k$, the lower the Schmidt rank of $K$ at every cut. Following [3] we construct such a polynomial based on the use of Chebyshev polynomials. The construction is summarized in the following definition.

**Definition 5** (*The Chebyshev-based AGSP*). Let $H$ be a Hamiltonian and $\eta_0 < \eta_1$ two parameters.[4] For any integer $k > 0$, let $T_k$ be the $k$-th Chebyshev polynomial of the first kind, and $P_k$ the following rescaling of $T_k$:

$$P_k(x) := \frac{1}{\tilde{P}_k(\eta_0)} \tilde{P}_k(x), \quad \text{where} \quad \tilde{P}_k(x) := T_k\left(2\frac{x-\eta_1}{\|H\|-\eta_1} - 1\right). \tag{3}$$

The *Chebyshev AGSP of degree k* for $H$ is $K := P_k(H)$.

The properties of the Chebyshev AGSP are given in the following theorem. Here and throughout we use the convention that a 1D local Hamiltonian on $n$ qudits numbered $1, \ldots, n$ decomposes as $H = \sum_{i=1}^{n-1} h_i$, where $0 \leq h_i \leq \mathbb{1}$ is the local term acting on qudits $\{i, i+1\}$.

**Theorem 1.** *Let $H$ be a Hamiltonian on $n$ qudits, $\eta_0 < \eta_1$ two parameters and $\gamma = \eta_1 - \eta_0$. Suppose that for some $i_1 < i_2 \in \{1, \ldots, n\}$ and $3 \leq \ell \leq (i_2 - i_1)/2$, $H$ can be written as*

$$\begin{aligned} &H_L + h_{i_1-\ell} + \cdots + h_{i_1} + \cdots + h_{i_1+\ell-1} \\ &\quad + H_M + h_{i_2-\ell} + \cdots + h_{i_2} + \cdots + h_{i_2+\ell-1} + H_R, \end{aligned} \tag{4}$$

*where each $h_i$ is a 2-local operator on qudits $\{i, i+1\}$ and $H_L$, $H_M$ and $H_R$ are defined on qudits $J_L = \{1, \ldots, i_1 - \ell\}$, $J_M = \{i_1 + \ell, \ldots, i_2 - \ell\}$ and $J_R = \{i_2 + \ell, \ldots, n\}$ respectively. For any integer $k > 0$ let*

$$\Delta := 4e^{-4k\sqrt{\frac{\gamma}{\|H\|-\eta_0}}}.$$

*Then the degree-$k$ Chebyshev AGSP $K$ is a $(D, \Delta)$ spectral AGSP for $(H, \eta_0, \eta_1)$ such that:*

1. *For any eigenvector $|\psi\rangle$ of $H$ with associated eigenvalue $\lambda$, $|\psi\rangle$ is an eigenvector of $K$ with associated eigenvalue $P_k(\lambda)$.*
2. *If $\lambda \leq \eta_0$ then $P_k(\lambda) \geq 1$, $P_k(\eta_0) = 1$, and if $\lambda \leq \eta_0 + \gamma/k$ then*

$$P_k(\lambda) \geq 1 - O\Big(\frac{k|\lambda - \eta_0|}{\gamma\|H\|}\sqrt{\Delta}\Big).$$

3. *If $\lambda \geq \eta_1$ then $P_k(\lambda) \leq \sqrt{\Delta}$.*
4. *The Schmidt rank of $K$ at all cuts in the region $J_M$ (resp. $J_L$, $J_R$) satisfies $B \leq \tilde{B}^{O(k)}$, where $\tilde{B}$ is an upper bound on the Schmidt rank of $H_M$ (resp. $H_L$, $H_R$) at every cut.*

[4] $\eta_0$ and $\eta_1$ may be chosen as the ground state energy and first excited energy of $H$ respectively, but they need not.

5. *The Schmidt rank of $K$ with respect to the cuts $(i_1, i_1+1)$ and $(i_2, i_2+1)$ satisfies $D \leq (dk)^{O(\ell+k/\ell)}$.*

*Proof.* Item 1. follows from the definition of $K = T_k(H)$ as a polynomial in $H$ (see Definition 5). For item 2. and item. 3 we use the following properties of $T_k$ (see e.g. [3] and [21, Lemma B.1] for a proof):

$$|T_k(x)| \leq 1 \qquad \text{for } |x| \leq 1, \tag{5}$$

$$|T_k(x)| \geq \frac{1}{2}\exp\left(2k\sqrt{\frac{|x|-1}{|x|+1}}\right) \qquad \text{for } |x| \geq 1, \tag{6}$$

$$T_k(x) = \frac{1}{2}\left(x+\sqrt{x^2-1}\right)^k + \frac{1}{2}\left(x-\sqrt{x^2-1}\right)^k \qquad \text{for } |x| \geq 1. \tag{7}$$

The fact that eigenvectors with eigenvalue $\eta_0$ are mapped to fixed points of $K$ follows from $P_k(\eta_0) = 1$. Next suppose $|\psi\rangle$ is an eigenvector of $H$ with eigenvalue $\eta_0 + \delta$ where $|\delta| < \eta_1 - \eta_0$. From (7) we see $|T_k(x+\delta) - T_k(x)| = O(k\delta/\min(x^2-1, x \pm \sqrt{x^2-1}))$ as long as $x, x+\delta \leq -1$. Taking into account the scaling used to define $P_k$,

$$|P_k(\eta_0+\delta) - P_k(\eta_0)| = O\Big(\frac{1}{\tilde{P}_k(\eta_0)}\frac{k\delta}{\gamma\|H\|}\Big) = O\Big(\frac{\delta\, k}{\gamma\|H\|}\Big)e^{-2k\sqrt{\frac{\gamma}{\|H\|-\eta_0}}},$$

where the last inequality uses (6). Item 3 follows by combining (5) and (6).

Item 4. is immediate since $K$ is computed as a linear combination of $j$-th powers of $H$ for $j \leq k$.

Finally, for a proof of item 5 we refer to Proposition 4 in Sect. 5.2 below. □

Theorem 1 provides us with a powerful recipe for constructing good AGSP. To minimize the Schmidt rank at a cut $(i, i+1)$ for $i \in \{i_1, i_2\}$ we should take $k = \Theta\left(\ell^2\right)$, which gives a bound of $D \leq (dk)^{O(\sqrt{k})}$, a much better bound than the naive $d^{O(k)}$. To guarantee a small $\Delta$ we should choose $k$ large enough to ensure that $e^{-4k\sqrt{\gamma/\|H\|}}$ remains small, which requires the Hamiltonian to have a small norm. This is the role of the truncation scheme presented in the following section.

*4.2. Hard truncation.* We introduce a scheme of *hard truncation* that is appropriate (though not efficient) for truncating the norm of an arbitrary local Hamiltonian in a certain region $J$, while preserving its low-energy eigenspace $H_{[\varepsilon_0, \varepsilon_0+\eta]}$. The basic idea is to replace $H \mapsto H\Pi_{\leq\varepsilon_0+t} + (\varepsilon_0+t)\Pi_{>\varepsilon_0+t}$, where $\Pi_{\leq t}$ projects onto the span of eigenvectors of $H$ with eigenvalue less than $t$, $\Pi_{>\varepsilon_0+t} := \mathbb{1} - \Pi_{\leq\varepsilon_0+t}$, and $t$ is chosen to be large enough with respect to $\eta$.

**Definition 6** (*Hard truncation*). Let $t > 0$, $H = H_J + H_{\overline{J}}$ where $H_J = h_{j_0} + h_{j_0+1} + \cdots + h_{j_1-1}$ is a local Hamiltonian acting on a contiguous set of qudits $J = \{j_0, j_0+1, \ldots, j_1\}$, and let $\varepsilon_J$ be the ground energy of $H_J$. Let $\Pi_-$ be the projector onto the span of all eigenvectors of $H_J$ with eigenvalue less than $\varepsilon_J + t$, and $\Pi_+ := \mathbb{1} - \Pi_-$. Then the *hard truncation* of $H_J$ is given by

$$\tilde{H}_J := H_J\Pi_- + (t+\varepsilon_J)\Pi_+ \tag{8}$$

and the *hard-truncated Hamiltonian* $\tilde{H}_t$ associated to the region $J$ is

$$\tilde{H}_t = \tilde{H}_J + H_{\overline{J}}.$$

We now show that truncating a $n$-qubit Hamiltonian on a subset $J$ of the qubits leads to a truncated Hamiltonian whose low-energy space is close to that of the original Hamiltonian. The main tool in proving this result is Theorem 2.6 of [4], a generalization and strengthening of the truncation result that appeared in [3]. Adapted to the current setting it can be formulated as follows.

**Proposition 1** (Adapted from Theorem 2.6 in [4]). *For any $\eta > 0$ let $\Pi_{\leq\eta}$ denote the projector on the span of all eigenvectors of $H$ with eigenvalue at most $\eta$, and similarly $\tilde{\Pi}_{\leq\eta}$ for $\tilde{H}_t$. Let $\varepsilon_0 \leq \varepsilon_1 \leq \varepsilon_2 \leq \ldots$ and $\tilde{\varepsilon}_0 \leq \tilde{\varepsilon}_1 \leq \tilde{\varepsilon}_2 \ldots$ be the sorted eigenvalues of $H$ and $\tilde{H}_t$ respectively, where eigenvalues appear with multiplicity. For any $\eta > 0$, let*

$$\xi = e^{(t-\eta)/8+24}. \tag{9}$$

*Then the following hold:*

1. $\|(H - \tilde{H}_t)\Pi_{\leq\varepsilon_0+\eta}\| \leq \xi$ *and* $\|(H - \tilde{H}_t)\tilde{\Pi}_{\leq\varepsilon_0+\eta}\| \leq \xi$,
2. *For all $j$ for which $\varepsilon_j \leq \varepsilon_0 + \eta$, we have $\varepsilon_j - \xi \leq \tilde{\varepsilon}_j \leq \varepsilon_j$.*

*Proof.* The proposition follows from Theorem 2.6 in Ref. [4] by using $\lambda = \frac{1}{8}$ and the fact that $\varepsilon_0 \leq \tilde{\varepsilon}_0 + 2$ to bound $\Delta\tilde{\varepsilon}$ by $\Delta\varepsilon + 2$. Here we can take $|\partial L| = 2$ since there are two boundary terms connecting the truncated region $J$ and the rest of the system. □

The following lemma summarizes the approximation properties of the hard truncation procedure that will be important for us.

**Lemma 8.** *For any $\eta > 0$, let $T_\eta = H_{[\varepsilon_0,\varepsilon_0+\eta]}$ be the low-energy eigenspace of $H$, $J = \{j_0, \ldots, j_1\}$ a contiguous subset of qudits and $\tilde{H}_t$ the associated hard-truncated Hamiltonian, with corresponding low-energy eigenspace $\tilde{T}_\eta = \tilde{H}_{[\tilde{\varepsilon}_0,\tilde{\varepsilon}_0+\eta]}$. Let $\xi$ be as defined in (9). Then the following hold for any $t > \eta$:*

1. *The ground energy $\tilde{\varepsilon}_0$ of $\tilde{H}_t$ satisfies $\varepsilon_0 - Ce^{-c(t-\eta)} \leq \tilde{\varepsilon}_0 \leq \varepsilon_0$ for some universal constants $C, c$.*
2. *For any $\delta > 0$ there is*

$$\eta' = \eta + \sqrt{\frac{\eta}{\delta}}\, e^{-\Omega(t-\eta)}$$

*such that the subspace $\tilde{T}_{\eta'}$ is $\delta$-close to $T_\eta$, and $T_{\eta'}$ is $\delta$-close to $\tilde{T}_\eta$.*

*Proof.* The first item follows directly from the second item in Proposition 1. For the second item, we prove that $\tilde{T}_{\eta'}$ is $\delta$-close to $T_\eta$, the proof of the second relation being identical. Fix a small width parameter $h$ (to be specified later) and let $|\psi\rangle = \sum_i \beta_i |\psi_i\rangle$ be supported on eigenvectors $|\psi_i\rangle$ of $H$ with eigenvalue $\mu_i \in [\lambda - h, \lambda + h]$ with $\lambda \leq \varepsilon_0 + \eta$. Then $\|H|\psi\rangle - \lambda|\psi\rangle\| \leq h$. Decompose $|\psi\rangle = \sum \alpha_i |\phi_i\rangle$, where for each $i$, $|\phi_i\rangle$ is an eigenvector of $\tilde{H}_t$ with associated eigenvalue $\tilde{\lambda}_i$. Using the first item in Proposition 1,

$$\begin{aligned}\sum_i |\alpha_i|^2 |\lambda - \tilde{\lambda}_i|^2 &\leq \left(\|(H - \tilde{H})|\psi\rangle\| + \|(H - \lambda\mathbb{1})|\psi\rangle\|\right)^2 \\ &\leq \left(e^{-\Omega(t-\eta)} + h\right)^2.\end{aligned}$$

By Markov's inequality it follows that for any $\delta > 0$

$$\left\| \tilde{\Pi}_{>\lambda+\delta} |\psi\rangle \right\| \leq \frac{e^{-\Omega(t-\eta)} + h}{\delta}.$$

Any $|\psi\rangle$ in $T_\eta$ can be written as a linear combination $|\psi\rangle = \sum_j \beta_j |h_j\rangle$ with each $|h_j\rangle$ supported on eigenvectors of $H$ with eigenvalue in a small window of width $2h$, and the number of terms at most $\lceil \frac{\eta-\varepsilon_0}{2h} \rceil$. Thus

$$\begin{aligned} \left\| \tilde{\Pi}_{>\varepsilon_0+\eta'} |\psi\rangle \right\| &\leq \sum_j |\beta_j| \left\| \tilde{\Pi}_{>\varepsilon_0+\eta'} |h_j\rangle \right\| \\ &\leq \sqrt{\frac{\eta}{2h}} \frac{e^{-\Omega(t-\eta)} + h}{\eta' - \eta}. \end{aligned}$$

Choosing $h = e^{-\Theta(t)}$, we see that the choice of $\eta'$ made in the statement of the lemma suffices to ensure that this quantity is at most $\sqrt{\delta}$, as desired. □

*4.3. The AGSP constructions.* The combination of Theorem 1, Proposition 1, and Lemma 8 yield a construction that starts with a local Hamiltonian $H$, produces a truncated Hamiltonian $\tilde{H}$ with low energy space close to that of $H$ along with a spectral AGSP $K$ for $\tilde{H}$ with a good trade-off between the parameters $D$ and $\Delta$.

**Corollary 1.** *Let $H$ be a 1D local Hamiltonian with ground energy $\varepsilon_0$, and $\mathcal{H} = \mathcal{H}_L \otimes \mathcal{H}_M \otimes \mathcal{H}_R$ a decomposition of the $n$-qudit space in contiguous regions. For any integer $\ell \geq 1$ and $t > 0$ there exists a Hamiltonian $\tilde{H}$ such that for any $\varepsilon_0 < \eta_0 < \eta_1$ there is a $(D, \Delta)$ spectral AGSP $K$ for $(\tilde{H}, \eta_0, \eta_1)$ with the following properties.*

1. $D = (d\ell)^{O(\ell)}$ *and* $\Delta = e^{-\Omega(\frac{\ell^2}{\sqrt{t+\ell}}\sqrt{\eta_1-\eta_0})}$,
2. *There are universal constants $C, c > 0$ such that for $i \in \{0, 1\}$*

$$0 \leq \varepsilon_i - \tilde{\varepsilon}_i \leq C e^{-c(t-\varepsilon_0)} \tag{10}$$

   *where $\varepsilon_i$, $\tilde{\varepsilon}_i$ are the $i$-th smallest (counted with multiplicity) eigenvalues of $H$, $\tilde{H}$ respectively.*
3. *The space $H_{[\varepsilon_0,\eta_1]}$ is $\delta$-close to $\tilde{H}_{[\tilde{\varepsilon}_0,\eta_0]}$ and $\tilde{H}_{[\tilde{\varepsilon}_0,\eta_1]}$ is $\delta$-close to $H_{[\varepsilon_0,\eta_0]}$, for*

$$\delta = \Theta\Big(\frac{\eta_0 - \varepsilon_0}{(\eta_1 - \eta_0)^2}\Big) e^{-\Omega(t-(\eta_0-\varepsilon_0))}. \tag{11}$$

*Proof.* Let $L = \{1, \ldots, i_1\}$, $M = \{i_1 + 1, \ldots, i_2\}$ and $R = \{i_2 + 1, \ldots, n\}$ be the set of qudits contained in $\mathcal{H}_L$, $\mathcal{H}_M$ and $\mathcal{H}_R$ respectively. We define the truncated Hamiltonian $\tilde{H}$ by applying the hard truncation transformation described in Definition 6 thrice, to the regions $J_L = \{1, \ldots, i_1 - \ell - 1\}$, $J_M = \{i_1 + \ell + 1, \ldots, i_2 - \ell - 1\}$ and $J_R = \{i_2 + \ell + 1, \ldots, n\}$ respectively (provided each region is non-empty). The resulting truncated Hamiltonian $\tilde{H} = \tilde{H}_t$ has norm $O(\ell + t)$.

Applying Lemma 8 thrice in sequence, for the three truncations performed, it follows that the sorted eigenvalues of $\tilde{H}$ satisfy (10). Eq. (11) similarly follows from item 2. in Lemma 8.

Finally we define the AGSP $K$ by applying the Chebyshev polynomial construction from Definition 5 to $\tilde{H}$ with a choice of $k = \ell^2$. The bounds on $\Delta$ and $D$ follow directly from item 3. and 5. from Theorem 1 respectively. □

From the corollary follow our two main AGSP contructions, which hold under assumptions (DG) and (LD) respectively.

**Theorem 2** (Existence of AGSP, (DG)). *Let $H$ be a local Hamiltonian satisfying Assumption (DG), and $\mathcal{H} = \mathcal{H}_L \otimes \mathcal{H}_M \otimes \mathcal{H}_R$ a decomposition of the $n$-qudit space in three contiguous blocks. There exists a collection of $D^2$ operators $\{A_i\}_{i=1}^{D^2}$ acting on $\mathcal{H}_M$ along with a subspace $\tilde{T} \subseteq \mathcal{H}$ such that:*

- *$H_{[\varepsilon_0, \varepsilon_0+\eta_0]}$ and $\tilde{T}$ are mutually .005-close;*
- *$D = e^{\tilde{O}\left(\frac{1}{\gamma}\log^3 d\right)}$,*
- *There is $\Delta > 0$ such that $D^{12}\Delta \leq 10^{-5}$ and whenever $S \subseteq \mathcal{H}_M$ is $\delta$-viable for $\tilde{T}$ then $S' = \mathrm{Span}\{\cup_i A_i S\}$ is $\delta'$-viable for $\tilde{T}$, with $\delta' = \frac{\Delta}{(1-\delta)^2}$.*

*Proof.* Let $\eta_0 = \varepsilon_0 + \gamma/10$ and $\eta_1 = \varepsilon_0 + 9\gamma/10$. Provided the implied constants are chosen large enough, setting $\ell = \Theta(\gamma^{-1}\log\gamma^{-1})$, $t = \tilde{O}(\ell)$ and $t > \tilde{O}(\frac{1}{\gamma}\log^2(d/\gamma))$ in Corollary 1 gives $D^{12}\Delta < 10^{-5}$. Due to the gap assumption it holds that $T = H_{[\varepsilon_0,\eta_0]} = H_{[\varepsilon_0,\eta_1]}$. The choice of $t$ above also ensures that the right-hand side of (10) is smaller than $\frac{1}{10}\gamma$ and the right hand side of (11) is smaller than .005, in which case $\tilde{H}$ has a spectral gap between $\eta_0$ and $\eta_1$, so that $\tilde{H}_{[\tilde{\varepsilon}_0,\eta_0]} = \tilde{H}_{[\tilde{\varepsilon}_0,\eta_1]}$. Then item 2 in the corollary implies that $\tilde{H}_{[\tilde{\varepsilon}_0,\eta_0]}$ and $T$ are mutually .005-close, giving the first condition in the theorem with $\tilde{T} = \tilde{H}_{[\tilde{\varepsilon}_0,\tilde{\varepsilon}_0+\frac{1}{10}\gamma]}$.

The operators $\{A_i\}$ are defined from a decomposition $K = \sum_{i=1}^{D^2} L_i \otimes A_i \otimes R_i$ associated to the factorization $\mathcal{H} = \mathcal{H}_L \otimes \mathcal{H}_M \otimes \mathcal{H}_R$ of the AGSP from Corollary 1. Lemma 6 gives the desired quantitative tradeoff between the increase in dimension of a viable set and its increase in overlap, when acted upon by the $\{A_i\}$. □

**Theorem 3** (Existence of AGSP, (LD)). *Let $\mu > 0$ be a constant, $H$ a local Hamiltonian satisfying Assumption (LD), and $\mathcal{H} = \mathcal{H}_L \otimes \mathcal{H}_M \otimes \mathcal{H}_R$ a decomposition of the $n$-qudit space in three contiguous blocks. For any $\eta \geq \eta_1 \geq 2\frac{\mu}{\log n}$ there exists a collection of $D^2$ operators $\{A_i\}_{i=1}^{D^2}$ acting on $\mathcal{H}_M$ along with two subspaces $\tilde{T}_- \subseteq \tilde{T} \subseteq \mathcal{H}$ such that:*

- *$H_{[\varepsilon_0,\varepsilon_0+\eta_1]}$ is .005-close to $\tilde{T}$,*
- *$\tilde{T}_-$ is .005-close to $H_{[\varepsilon_0,\varepsilon_0+\eta_1-\frac{\mu}{\log n}]}$,*
- *$D = e^{\tilde{O}\left(\frac{\log n}{\mu}\log^3 d\right)}$,*
- *There is a $\Delta > 0$ such that $D^{12}\Delta < 10^{-5}$ and for any $S \subseteq \mathcal{H}_M$ that is $\delta$-viable for $\tilde{T}$ it holds that $S' = \mathrm{Span}\{\cup_i A_i S\}$ is $\delta'$ -viable for $\tilde{T}_-$ with $\delta' = \frac{\Delta}{(1-\delta)^2}$.*

*Proof.* The main difference with the proof of Theorem 2 is that the parameter corresponding to the gap $\gamma$ is replaced by the quantity $\frac{\mu}{\log n}$. The proof of the first two items claimed in the theorem then closely mirrors that of Theorem 2.

It only remains to verify the third item. Despite having the desired AGSP, unlike in the gapped case we cannot hope to improve the quality of the viable set $S$ for all of $\tilde{T} = \tilde{H}_{[0,\eta'-\frac{\mu}{3\log n}]}$ by the application of the AGSP $K^k$. However, if we view $S$ as a viable set for the smaller $\tilde{T}_- = \tilde{H}_{[0,\eta'-\frac{2\mu}{3\log n}]} \subseteq \tilde{T}$, we now have an effective AGSP with respect to $\tilde{T}_-$ and the orthogonal complement of the larger $\tilde{T}$ and we can proceed as if

in the presence of a small spectral gap of $\frac{\mu}{3\log n}$. To see this, fix any vector $|\psi\rangle \in \tilde{T}_-$. Lemma 1 shows that there exists a $|w\rangle \in S$ such that $|w\rangle = c|\psi\rangle + |\psi^\perp\rangle$ for some $|\psi^\perp\rangle$ orthogonal to $\tilde{T}$, and $c \geq (1-\delta)$. This brings us in line with the proof of Lemma 6 and we can use the same analysis to show that applying $K$ improves the parameter of the viable set $S$ from $\delta$ to the desired $\delta' = \frac{\Delta}{(1-\delta)^2}$. □

*4.4. Area law for degenerate Hamiltonians.*

**Theorem 4** (Area law for degenerate gapped Hamiltonians). *Let $H$ be a 1D local Hamiltonian acting on $n$ qudits of local dimension $d$ such that $H$ satisfies Assumption (DG). For any fixed cut and any $\delta = \text{poly}^{-1}(n)$, for every unit $|\psi\rangle \in T$ there is an approximation $|\psi'\rangle$ such that $|\langle\psi|\psi'\rangle| \geq 1-\delta$ and $|\psi'\rangle$ has Schmidt rank no larger than*

$$s(\delta) = r\, e^{\tilde{O}\left(\frac{1}{\gamma}\log^3 d + \frac{1}{\gamma^{1/4}}\log^{3/4}(\frac{1}{\delta})\log d\right)}$$

*at that cut, and an MPS representation with bond dimension bounded by*

$$r\, e^{\tilde{O}\left(\frac{1}{\gamma}\log^3 d + \frac{1}{\gamma^{1/4}}\log^{3/4}(\frac{n}{\delta})\log d\right)}.$$

*Moreover, every state $|\psi\rangle \in T$ has entanglement entropy*

$$S(|\psi\rangle\langle\psi|) \leq \ln r + \tilde{O}\left(\frac{1}{\gamma}\log^3 d\right).$$

The proof of the theorem proceeds in two steps. First we use a "bootstrapping argument" to show the existence of a viable set of constant error for the ground space, such that all states in the viable set have low Schmidt rank. The existence of arbitrarily good approximations with increasing Schmidt rank, as well as the bound on the entanglement entropy, follow by the application of a suitable AGSP. We state the bootstrapping step as the following proposition. The proposition can be understood as an analysis of the effect of a single application of the MERGE procedure introduced in Sect. 3.2 with the initial tensoring step omitted. (The connection will be made more formal once we analyze algorithms in Sect. 6.)

**Proposition 2.** *Let $H$ be a local Hamiltonian satisfying assumption (DG), and $J \subseteq \{1,\ldots,n\}$. Then there exists a subspace $W \subseteq \mathcal{H}_J$ of dimension $q = re^{\tilde{O}\left(\frac{1}{\gamma}\log^3 d\right)}$ that is .015-viable for the ground space $T$ of $H$.*

*Proof.* Let $W \subseteq \mathcal{H}_J$ be a subspace of minimal dimension $q$ among all .015-viable subspaces for $T$. Let $\{A_i\}_{i=1}^{D^2}$ be AGSP operators guaranteed by Theorem 2 for the Hamiltonian $H$ and region $M = J$, and $\tilde{T}$ the associated subspace. The first condition in the theorem together with Lemma 3 establishes that $W$ is .04-viable for $\tilde{T}$. Let $s = q/(2D^2)$ and $W' \subseteq W$ a random subspace of dimension $s$. By Lemma 5, $W'$ is $(1-\delta')$-viable for $T$ with $\delta' = s/(16q) = 1/(32D^2)$ with positive probability provided

$$s = \Omega\left(\log q + r\log\left(\frac{q}{s}\right)\right) \tag{12}$$

for a large enough implied constant, as this will suffice to guarantee that $\eta$ stated in the lemma is strictly less than 1.

Let $S = \cup_{i=1}^{D^2} A_i W'$. Then given our choice of $s$, $S$ has cardinality at most $q/2$, and by Lemma 6 is $(32D^2)^2\Delta$-viable for $\tilde{T}$. The condition $D^{12}\Delta < 10^{-5}$ implies $(32D^2)^2\Delta + 0.005 \leq 0.1 \leq 0.15$, giving a contradiction with the minimality of $q$. The contradiction holds as long as the condition (12) holds, which given the bound on $D$ from Theorem 2 will be the case as long as $q = re^{\tilde{\Omega}(\gamma^{-1}\log^3 d)}$ for a large enough implied constant in the exponent. □

Given the proposition, the proof of Theorem 4 follows by application of an AGSP derived from Corollary 1.

*Proof (of Theorem 4).* Fix a cut $\mathcal{H} = \mathcal{H}_L \otimes \mathcal{H}_R$ as in the theorem. Let $V_L$ and $V_R$ be 0.015-viable sets of minimal dimension for regions $J = L$ and $J = R$ respectively, and let $q$ be an upper bound on their dimension. Proposition 2 guarantees that we may take $q = re^{\tilde{O}\left(\frac{1}{\gamma}\log^3 d\right)}$. By Lemma 4 the set $W = V_L \otimes V_R$ is .03-viable for $T$. The tensor product structure ensures that every element of $W$ has Schmidt rank no larger than $q$. Apply Corollary 1 to $H$, with $\eta_0 = \varepsilon_0 + \gamma/10$, $\eta_1 = \varepsilon_0 + 9\gamma/10$, $t = \Theta(\gamma^{-1/4}\log\delta^{-1})$ and $\ell = \Theta(t^{3/4})$. This gives a spectral AGSP $K$ with

$$D = e^{\tilde{O}\left(\gamma^{-1/4}\log^{3/4}(\frac{1}{\delta})\log d\right)} \quad \text{and} \quad \Delta \leq \delta/2,$$

for a Hamiltonian $\tilde{H}$ such that $\tilde{T} = \tilde{H}_{[\tilde{\varepsilon}_0, \tilde{\varepsilon}_0 + \frac{1}{10}\gamma]}$ and $T$ are mutually $(\delta/2)$-close. Applying Lemma 6 the set $W' = KW$ is $(\delta/2)$-viable for $\tilde{T}$ and every element within it has Schmidt rank no larger than $qD$. Since $\tilde{T}$ and $T$ are $(\delta/2)$-close, $W'$ is $\delta$-viable for $T$.

This proves the first statement in the theorem. The second follows by setting $\delta = \delta'/n$ in the above and noticing that the error made at each cut will add up linearly. The proof of the last statement is standard and follows from the bound on $s(\delta)$ as in [3]: we bound

$$S\big(|\psi\rangle\langle\psi|\big) \leq \ln\left(re^{\tilde{O}\left(\frac{1}{\gamma}\log^3 d\right)}\right) + \sum_{i=3}^{\infty} 2^{-i}\log\big(s(2^{-(i+1)})\big),$$

which is dominated by the first term. □

*4.5. Area law for low-density Hamiltonians.*

**Theorem 5** (Area law for low-density Hamiltonians). *Let $H$ be a 1D local Hamiltonian acting on $n$ qudits of local dimension $d$ such that $H$ satisfies Assumption (LD), $\mu < \eta \log n$ any positive constant and $T = H_{[\varepsilon_0, \varepsilon_0 + \eta - \mu/\log n]}$. For any fixed cut and any $\delta = \mathrm{poly}^{-1}(n)$, for every unit $|\psi\rangle \in T$ there is an approximation $|\psi'\rangle$ such that $|\langle\psi|\psi'\rangle| > 1 - \delta$ and $|\psi'\rangle$ has Schmidt rank no larger than*

$$s(\delta) = r\, e^{\tilde{O}\left(\frac{\log n}{\mu}\log^3 d + (\frac{\log n}{\mu})^{1/4}\log^{3/4}(\frac{1}{\delta})\log d\right)}$$

*at that cut, and $|\psi'\rangle$ has an MPS representation with bond dimension bounded by*

$$r\, e^{\tilde{O}\left(\frac{\log n}{\mu}\log^3 d + (\frac{\log n}{\mu})^{1/4}\log^{3/4}(\frac{n}{\delta})\log d\right)}.$$

*Moreover, every state $|\psi\rangle \in T$ has entanglement entropy*

$$S(|\psi\rangle\langle\psi|) \leq \ln r + \tilde{O}\left(\frac{\log n}{\mu}\log^3 d\right)$$

As for Theorem 4, the proof of Theorem 5 follows from a bootstrapping argument. We establish the analogue of Proposition 2 below. Just as for Theorem 4, the theorem then follows by application of a suitable AGSP, and we omit that part of the proof.

**Proposition 3.** *Let $H$ be a local Hamiltonian satisfying assumption (LD), for some $\eta > 0$. Let $J \subseteq \{1, \ldots, n\}$ and $\mu > 0$. Then there exists a subspace $W \subseteq \mathcal{H}_J$ of dimension $q = re^{\tilde{O}\left(\frac{\log n}{\mu} \log^3 d\right)}$ that is .015-viable for the low-energy space $T_\mu = H_{[\varepsilon_0, \varepsilon_0 + \eta - \mu/\log n]}$.*

*Proof.* For fixed $d$ and $n$, let $C = C(d, n)$ be a constant such that the bound $D \leq e^{C\frac{\log n}{\mu} \log^c(\frac{\log n}{\mu})}$ holds in Theorem 3 for all $\mu > 0$, where $c > 0$ is a universal constant implied by the $\tilde{O}$ notation. For any $\mu > 0$ let $q(\mu)$ be the smallest dimension of a subspace $W_\mu \subset \mathcal{H}_J$ that is .015-viable for $T_\mu$. Note that $q(\mu)$ is a non-increasing function of $\mu$. For $\mu > 0$, let $r(\mu) = re^{C'\frac{\log n}{\mu} \log(\log n/\mu)}$, where $C' = 3C$. For any $\mu$, let $i_0$ be the smallest power of two such that $q(\mu/2^{i_0}) \leq r(\mu/2^{i_0})$. Note that $i_0$ is finite as $q(\mu) \leq d^n$ for all $\mu > 0$. If $i_0 = 0$ then the proposition is proven. Suppose $i_0 > 0$, and let $\mu_0 = \mu/2^{i_0 - 1}$. Let $W = W_{\mu_0/2}$ be a subspace of dimension $q = q(\mu_0/2)$ that is 0.15-viable for $T_{\mu_0/2}$. Let $\{A_i\}_{i=1}^{D^2}$ be AGSP operators guaranteed by Theorem 3 for the Hamiltonian $H$, region $M = J$, and parameters $\eta_1 = \eta - \mu_0/(2 \log n)$ and $\mu = \mu_0/2$. Let $\tilde{T}$ and $\tilde{T}_-$ be the resulting subspaces. The first condition in the theorem, together with Lemma 3, establishes that $W$ is .04-viable for $\tilde{T}$.

Let $s = q(\mu_0)/D^2$ and $W' \subseteq W$ a random subspace of dimension $s$. By Lemma 5 and the definition of $i_0$, $W'$ is $(1 - \delta')$-viable for $\tilde{T}$ with

$$\delta' = \frac{s}{16q} = \frac{q(\mu_0)}{16q(\mu_0/2)D^2} \geq \frac{r(\mu_0)}{16r(\mu_0/2)D^2} \tag{13}$$

with positive probability provided

$$s = \Omega\left(\log q + r \log\left(\frac{q}{s}\right)\right) \tag{14}$$

for a large enough implied constant, as this will suffice to guarantee that $\eta$ stated in the lemma is strictly less than 1.

Let $S = \cup_{i=1}^{D^2} A_i W'$. Then $S$ has cardinality at most $q(\mu_0)$, and by Lemma 6 is $\Delta/(\delta')^2$-viable for $\tilde{T}_-$, itself 0.005-close to $W_{\mu_0}$. The condition $D^{12}\Delta < 10^{-5}$, together with (13) implies

$$\frac{\Delta}{(\delta')^2} < 16^2 \cdot 10^{-5} D^{-8} \left(\frac{r(\mu_0/2)}{r(\mu_0)}\right)^2 \leq 10^{-2} e^{(2C' - 8C)\frac{\log n}{\mu_0} \log^c\left(\frac{\log n}{\mu_0}\right)}.$$

Provided $C' \leq 4C$ this is at most $10^{-2}$, leading to a contradiction with the minimality of $i_0$. The contradiction holds as long as the condition (14) holds, which will be the case provided $C' > 2C$. Choosing $C' = 3C$ satisfies both conditions. □

## 5. Efficient AGSP Constructions

This section is devoted to the construction of efficiently computable, and efficiently implementable [as polynomial-size matrix product operators (MPO)], analogues of the existential AGSP constructions obtained in Sect. 4. The first step in obtaining efficient

constructions consists in replacing the method of hard truncation considered in Sect. 4.2 with a method of "soft truncation". This method, described in Sect. 5.1, is somewhat less effective than hard truncation, but has the advantage that it can be made efficient; this is essential for its use in the algorithms presented in Sect. 6. In Sect. 5.2 we show that the Chebyshev polynomial AGSP introduced in Sect. 4.1 can also be made efficient. Our efficient AGSP constructions for the (DG) and (LD) cases are provided in Sect. 5.3. We conclude in Sect. 5.4 with a more efficient construction specialized to the (FF) case; this last construction replaces the intricate AGSP constructions with a much simpler one based on the detectability lemma [1]. (The reader new to AGSP constructions may wish to start with the latter section.)

*5.1. Soft truncation.* We introduce a scheme of *soft truncation* that reduces the norm of a local Hamiltonian $H$ in a certain region $J$ in a way that the truncated operator can be well-approximated by an MPO with small bond dimension. In hard truncation (Definition 6) the operator $\Pi_{\leq\varepsilon_0+t} H + (\varepsilon_0 + t)\Pi_{>\varepsilon_0+t})$ is used. This can be written as $g_t(H)$, where $g_t(x)$ is defined by $g_t(x) := x$ for $x \leq \varepsilon_0 + t$ and $g_t(x) := t$ for $x > \varepsilon_0 + t$. The main idea of soft truncation is to replace this non-smooth function by the infinitely differentiable function

$$h_{t',t}(x) := t\Big(f_t(x) + \frac{f_t(x)^2}{2} + \cdots + \frac{f_t(x)^{t'}}{t'}\Big), \quad \text{where} \quad f_t(x) := 1 - e^{-x/t}, \tag{15}$$

which results in an operator $h_{t',t}(H)$ that closely approximates the hard-truncated Hamiltonian. Moreover, $h_{t',t}(H)$ can be given an efficient representation as an MPO by leveraging the truncated cluster expansion [15,20] and its matrix product operator (MPO) representation from [25, Section IV].

The following are basic properties of $h_{t',t}$.

**Lemma 9.** *For any integers $t', t \geq 1$ and $x \geq 0$,*

$$\big|h_{t',t}(x) - x\big| \leq \frac{t}{t'}\Big(\frac{x}{t}\Big)^{t'}, \qquad \text{and} \qquad \big|h_{t',t}(x)\big| \leq t\ln(t').$$

*Proof.* Let $g_t(y) = -t\ln(1-y)$, so that $g_t(f_t(x)) = x$ for any $x \in [0,\infty)$. The function $h_{t',t}$ contains the first $t'$ terms of the Taylor expansion of $g_t$ around 0, applied to $f_t(x)$, and the first inequality follows from Taylor's theorem and $f_t(x) \leq x$ for all $x$. The second inequality follows since $f_t(x) \leq 1$ for all $x$. □

Recall our convention that a 1D local Hamiltonian acting on $n$ qudits numbered $1,\ldots,n$ decomposes as $H = \sum_{i=1}^{n-1} h_i$, where $0 \leq h_i \leq \mathbb{1}$ is the local term acting on qudits $\{i, i+1\}$. In addition to the truncation parameters $t$ and $t'$ the soft truncation construction is parametrized by a region $J \subseteq \{1,\ldots,n\}$ which specifies the set of local terms on which truncation is to be performed, and an energy $\varepsilon'_J$ which is meant to be an approximation to the ground state energy of the restriction $H_J$ of $H$ to $J$.

**Definition 7** (*Soft truncation*). Let $H = H_J + H_{\overline{J}}$ be a 1D Hamiltonian, where $H_J = h_{j_0} + \cdots + h_{j_1-1}$ acts on a contiguous set $J = \{j_0,\ldots,j_1\}$ of qudits. Let $\varepsilon_J$ be the ground energy of $H_J$, and $\varepsilon'_J$ an approximation to $\varepsilon_J$ satisfying $\varepsilon_J - 10 \leq \varepsilon'_J \leq \varepsilon_J$. For given truncation parameters $t \geq t' \geq 1$, the *soft truncation* of $H_J$ is given by

$$\tilde{H}_J := \varepsilon'_J\mathbb{1} + h_{t',t}(H_J - \varepsilon'_J\mathbb{1}),$$

and the *soft-truncated Hamiltonian* $H$ associated to region $J$ is

$$\tilde{H}_{t',t} := \tilde{H}_J + H_{\overline{J}}.$$

The following lemma shows that for sufficiently large $t$ and $t'$, $\tilde{H}_{t',t}$ provides a good approximation to the lower part of the spectrum of $H$.

**Lemma 10.** *Let $H = H_J + H_{\overline{J}}$ be a local 1D Hamiltonian. Given truncation parameters $t \geq t' \geq 2$, the soft-truncated Hamiltonian $\tilde{H}_{t',t}$ satisfies $\tilde{H}_{t',t} \leq H$ and for any eigenvector $|\psi\rangle$ of $H$ with energy $\lambda$ (resp. $|\phi\rangle$ of $\tilde{H}$ with energy $\mu \leq t$) it holds that*

$$\lambda - O\Big(\frac{(\lambda-\varepsilon)^{t'}}{t' t^{t'-1}}\Big) \leq \langle\psi|\tilde{H}_{t',t}|\psi\rangle \leq \lambda \quad \text{and} \quad \mu \leq \langle\phi|H|\phi\rangle \leq \mu + O\Big(\frac{(2(\mu-\varepsilon))^{t'}}{t' t^{t'-1}}\Big), \tag{16}$$

*where $\varepsilon = \varepsilon_{\overline{J}} + \varepsilon'_J$. In addition, if $H$ is gapped with gap $\gamma$ then provided $t = \Omega(\gamma^{-1})$, $\tilde{H}_{t',t}$ is gapped with gap $\gamma/2 \leq \tilde{\gamma} \leq 2\gamma$.*

*For $\eta > 0$ let $T_\eta = H_{[\varepsilon_0, \varepsilon_0+\eta]}$ (resp. $\tilde{T}_\eta = \tilde{H}_{[\tilde{\varepsilon}_0, \tilde{\varepsilon}_0+\eta]}$) be the span of all eigenvectors of $H$ (resp. $\tilde{H}_{t',t}$) with associated eigenvalue in the indicated range. Then for any $\eta, \delta > 0$ there is*

$$\eta' = \eta + O\Big(\Big(\frac{\eta+10}{t}\Big)^{t'-1} \frac{1}{t'\sqrt{\delta}}\Big)$$

*such that the subspace $\tilde{T}_{\eta'}$ is $\delta$-close to $T_\eta$ and $T_{\eta'}$ is $\delta$-close to $\tilde{T}_\eta$.*

*Proof.* From Definition 7,

$$\tilde{H}_{t',t} - H = h_{t',t}(H_J - \varepsilon'_J \mathbb{1}) - (H_J - \varepsilon'_J \mathbb{1}). \tag{17}$$

Using the first bound from Lemma 9, we get that for any vector $|\psi\rangle$,

$$\big|\langle\psi|\tilde{H}_t|\psi\rangle - \langle\psi|H|\psi\rangle\big| \leq \frac{1}{t' t^{t'-1}} \langle\psi|(H_J - \varepsilon'_J \mathbb{1})^{t'}|\psi\rangle. \tag{18}$$

Furthermore,

$$\begin{aligned} H_J - \varepsilon'_J \mathbb{1} &\leq H_{\overline{J}} - \varepsilon_{\overline{J}} \mathbb{1} + H_J - \varepsilon'_J \mathbb{1} \\ &= H - (\varepsilon_{\overline{J}} + \varepsilon'_J)\mathbb{1}, \end{aligned}$$

which combined with (18) and $H_J - \varepsilon'_J \mathbb{1} \geq 0$ proves the first two inequalities in (16); the other two are obtained in the same way using in addition $x \leq 2h_{t',t}(x)$ for $0 \leq x \leq t$. The relations between the spectral gaps of $H$ and $\tilde{H}_{t',t}$ follow from these inequalities.

Starting from (17), squaring both sides and using (the square of) the first bound from Lemma 9 we get the operator inequality

$$(\tilde{H}_{t',t} - H)^2 \leq \frac{1}{(t')^2 t^{2t'-2}} (H_J - \varepsilon'_J \mathbb{1})^{2t'}. \tag{19}$$

Let $\bar{H}_{\overline{J}} = H_{\overline{J}} - h_{j_0-1} - h_{j_1}$, so that $\bar{H}_{\overline{J}}$ and $H_J$ commute. Using $\bar{H}_{\overline{J}} + (2 - \varepsilon_{\overline{J}})\mathbb{1} \geq 0$,

$$(H_J - \varepsilon'_J \mathbb{1})^{2t'} \leq (H_J - \varepsilon'_J \mathbb{1} + \bar{H}_{\overline{J}} + (2 - \varepsilon_{\overline{J}})\mathbb{1})^{2t'} \leq ((H - \varepsilon)\mathbb{1} + 10\mathbb{1})^{2t'}. \tag{20}$$

Let $|\psi\rangle$ be supported on eigenvectors of $H$ with eigenvalues in the range $[\lambda - h, \lambda + h]$ with $\lambda \leq \varepsilon_0 + \eta$ and $h$ a small width parameter. Decompose $|\psi\rangle = \sum_i \alpha_i |\phi_i\rangle$, where for each $i$, $|\phi_i\rangle$ is an eigenvector of $\tilde{H}_{t',t}$ with associated eigenvalue $\tilde{\lambda}_i$. Thus

$$\begin{aligned}
\Big(\sum_i |\alpha_i|^2 |\lambda - \tilde{\lambda}_i|^2\Big)^{1/2} &\leq \Big\| \sum_i \alpha_i (\lambda - \tilde{\lambda}_i) |\phi_i\rangle \Big\| + h \\
&= \big\| (\tilde{H}_{t',t} - H)|\psi\rangle \big\| + h \\
&= \langle \psi | (\tilde{H}_{t',t} - H)^2 |\psi\rangle^{1/2} + h \\
&\leq \frac{1}{t' t^{t'-1}} \langle \psi | ((H - \varepsilon)\mathbb{1} + 10\mathbb{1})^{2t'} |\psi\rangle^{1/2} + h \\
&\leq \frac{1}{t' t^{t'-1}} (\eta + 10)^{t'} + h,
\end{aligned}$$

where the inequality before last follows by combining (19) and (20). Applying Markov's inequality it follows that for any $\delta > 0$

$$\big\| \tilde{\Pi}_{>\lambda+\delta} |\psi\rangle \big\| \leq \frac{\frac{1}{t' t^{t'-1}} (\eta + 10)^{t'} + h}{\delta}.$$

Any $|\psi\rangle$ in $T_\eta$ can be written as a linear combination $|\psi\rangle = \sum_j \beta_j |h_j\rangle$ with each $|h_j\rangle$ supported on eigenvectors of $H$ with eigenvalue in a small window of width $2h$, and the number of terms is at most $\lceil \frac{\eta - \varepsilon_0}{2h} \rceil$. Thus

$$\begin{aligned}
\big\| \tilde{\Pi}_{>\varepsilon_0+\eta'} |\psi\rangle \big\| &\leq \sum_j |\beta_j| \big\| \tilde{\Pi}_{>\varepsilon_0+\eta'} |h_j\rangle \big\| \\
&\leq \sqrt{\frac{\eta}{2h}} \frac{\frac{1}{t' t^{t'-1}} (\eta + 10)^{t'} + h}{\eta' - \eta}.
\end{aligned}$$

Chosing $h = \frac{1}{t' t^{t'-1}} (\eta + 10)^{t'}$, we see that the choice of $\eta'$ made in the statement of the lemma suffices to ensure that this quantity is at most $\delta$, as desired. □

We end this section by showing that the soft-truncated Hamiltonian $\tilde{H}_{t',t}$ can be approximated by an operator with polynomial bond dimension which can be computed efficiently. Our construction is based on the cluster expansion from [15,20] in the 1D case, with some small adjustments. We first state the result.

**Lemma 11.** *Let $t$ and $t' < (\ln(2)/2)t$ be truncation parameters and $H$ a $n$-qudit local Hamiltonian. For any $\xi > 0$ there is an MPO representation $\tilde{H}'$ for the truncated Hamiltonian $\tilde{H} = \tilde{H}_{t,t'}$ such that $\|\tilde{H} - \tilde{H}'\| \leq \xi$ and $\tilde{H}'$ has bond dimension $\mathrm{poly}(t' 2^{t'} n/\xi)$ across all bonds. Such an MPO can be constructed in time polynomial in its size.*

*Proof.* The truncation $h_{t',t}(H)$ can be expressed as a linear combination of $O(t' 2^{t'})$ terms of the form $e^{-\beta H}$ for values of $\beta$ in $\{1/t, \ldots, t'/t\}$; moreover the coefficients of the linear combination are at most $O(t' 2^{t'})$ each. Using Theorem 6 and the assumption

$t'/t \leq \ln(2)/2$ each $e^{-\beta H}$ can be approximated, in the operator norm, by an MPO of the form $M_r(H)$ with error less than $\xi/(t'2^{t'})^2$ as long as $r = \Omega(\ln((t')^2 2^{2t'} n^2/\xi))$. Finally, Theorem 7 states that such an MPO with the claimed bond dimension can be found efficiently. □

Let $H = \sum_{i=1}^{n-1} h_i$ be a 1D, 2-local Hamiltonian on $n$ qudits of dimension $d$, with $\|h_i\| \leq 1$ (but the $h_i$ are not necessarily non-negative), and let $\beta > 0$ be an inverse temperature. We write the cluster expansion $e^{-\beta H} = \sum_w f(w)$, where $w$ runs over all words on $\{1, \ldots, n-1\}$ and $f(w) := \frac{(-\beta)^{|w|}}{|w|!} h_w$ with $h_w := \prod_{i \in w} h_i$. For an integer $r > 0$, let $S_{<r}$ be the set of all those $w$ such that the support of $w$, the set of qudits on which $h_w$ acts non-trivially, consists of connected components of size smaller than $r$. Let $M_r(H) := \sum_{w \in S_{<r}} f(w)$ be the "truncated cluster expansion" of $e^{-\beta H}$. The following theorem follows from the proof of Lemma 2 in [20]; we give the proof in Appendix 7.

**Theorem 6.** *Let $\beta$ be such that $e^{\beta} - 1 < 1$. Then the following approximation holds in the operator norm:*

$$\|e^{-\beta H} - M_r(H)\| \leq e^{n^2 (e^{\beta}-1)^r} - 1.$$

The next theorem states that the operator $M_r(H)$ can be written efficiently as an MPO. This encoding also shows that the operator $M_r(H)$ has a low Schmidt rank. The proof, which is given in Appendix 7, follows very closely the ideas of [25, Section IV].

**Theorem 7.** *The $r$th order cluster expansion $M_r(H)$ of the operator $e^{-\beta H}$ can be written as an MPO of bond dimension $\leq r^2 d^r$ which can be computed in time $n d^{O(r)}$.*

*5.2. The Chebyshev polynomial.* For algorithmic purposes it is important that the Chebyshev AGSP can be constructed efficiently once one is given MPO representations for the truncated part of the Hamiltonian. The following proposition states that this is possible.

**Proposition 4.** *Let $H$ be a Hamiltonian having a decomposition of the form described in (4), $k$ an integer, and $K = P_k(H)$ the associated degree-$k$ Chebyshev AGSP as defined in Definition 5. Assume that $H_M$ (but not necessarily $H_L$ or $H_R$) is specified by an MPO with bond dimensions at most $\tilde{B}$.*

*Then there exists $D \leq (dk)^{O(\ell + k/\ell)}$ such that a family of $D^2$ MPO $\{A_1, \ldots, A_{D^2}\}$ of bond dimension at most $\tilde{B}^k$ each such that there exists $B_1, \ldots, B_{D^2}$ with $K = \sum A_i \otimes B_i$ can be computed in time $n D^2 \tilde{B}^{O(k)}$. Here the $A_i$ act on qudits $\{i_1, i_1 + 1, \ldots, i_2\}$ and the $B_i$ on the remaining qudits. This computation does not require knowledge of $\eta_0, \eta_1$.*

*Furthermore, if $H_L$ and $H_R$ are also given as MPO with bond dimension at most $\tilde{B}$ then the $B_i$ can be computed as well.*

*Proof.* The proof follows from a close examination of the proof of [3, Lemma 4.2].[5] Adapting to our setting (where there are two cuts to consider simultaneously) the argument made in [3] shows that in order to obtain an MPO for $K$ it suffices to include in the set $\{A_1, \ldots, A_{D^2}\}$ MPO representations for operators $P_{u_1 u_2, k j_1 j_2}(Z)$ where $u_1 \in \{i_1 - \ell, \ldots, i_1 + \ell - 1\}$, $u_2 \in \{i_2 - \ell, \ldots, i_2 + \ell - 1\}$, $j_1, j_2 \in \{0, \ldots, k + 2\ell - 2\}$

[5] To follow the ensuing argument it may be helpful to translate the notation used for the indices in [3, Lemma 4.2] to the notation used here as follows: $s \to 2\ell - 2$, $\ell \to k$, $k \to j$.

and $Z$ is an $(4\ell-4)$-tuple of complex variables which takes on $\binom{k-j_1+2\ell-2}{2\ell-2}\binom{k-j_2+2\ell-2}{2\ell-2}$ possible values. For our purposes, a random choice of such values, e.g. distributed uniformly on the unit circle, will lead to a correct construction with probability 1 (i.e. only depending on the number of digits of accuracy). We argue below that for each $P_{u_1u_2,kj_1j_2}(Z)$ one can efficiently construct an explicit set of MPO $\{A_\alpha\}$, where $1 \le i \le \binom{k+j_1+1}{2j_1+1}\binom{k+j_2+1}{2j_2+1}d^{2(j_1+j_2)+4\ell}$, such that there exists $B_i$ for which $\sum A_i \otimes B_i$ is an MPO for $P_{u_1u_2,kj_1j_2}(Z)$. This will lead to the claimed bounds as

$$\sum_{u_1=i_1-\ell}^{i_1+\ell-1}\sum_{u_2=i_2-\ell}^{i_2+\ell-1}\sum_{j_1,j_2=0}^{\lfloor k/2\ell\rfloor} d^{2(j_1+j_2)+4\ell}\cdot\binom{k-j_1+2\ell-2}{2\ell-2}\binom{k-j_2+2\ell-2}{2\ell-2}\cdot\binom{k+j_1+1}{2j_1+1}\binom{k+j_2+1}{2j_2+1}$$

can be crudely bounded by $(dk)^{O(\ell+k/\ell)}$.

Fix $u_1, u_2$ and recall that $P_{u_1u_2,kj_1j_2}(Z)$ is defined as the sum of those terms in the expansion of $(H_L + \cdots + H_{i_1} + \cdots + H_{i_2} + \cdots + H_R)^k$ which contain exactly $j_1$ (resp. $j_2$) occurrences of $H_{u_1}$ (resp. $H_{u_2}$). There are $\binom{r+j_1+1}{2j_1+1}\binom{r+j_2+1}{2j_2+1}$ such terms. By cutting to the left of $u_1$ and right of $u_2$ we can efficiently construct at most $d^{2(j_1+j_2)}$ MPO which, properly combined, would give an MPO for the corresponding product. Finally we cut these MPO further so as to make the separation be to the left of $i_1$ and right of $i_2$ (or complete them appropriately, depending on whether $u_1 \le i_1$ or $u_1 > i_1$, and similarly for $u_2$ with respect to $i_2$). This last step multiplies the number of MPO by at most $d^{4\ell}$ (where we use $|i_1 - u_1|, |i_2 - u_2| \le \ell$), giving the claimed bound. □

*5.3. Efficient AGSP constructions.* We combine the soft truncation scheme with the Chebyshev polynomial AGSP to show that matrix product operator representations for operators $\{A_i\}$ satisfying the conditions of Theorem 2 and Theorem 3 can be computed efficiently (in polynomial and quasi-polynomial time respectively). The same procedure, GENERATE, underlies both constructions, merely requiring a different choice of parameters in the two cases. The procedure is summarized in Fig. 1 (it is implicit that the procedure is passed as an argument which assumption $H$ satisfies). We state its properties for the (DG) case in Theorem 8, and for the (LD) case in Theorem 9. For the case of a Hamiltonian satisfying assumption (DG) with a ground energy $\varepsilon_0$ and a unique ground state (assumption (FF) of frustration-freeness) the procedure can be made even more efficient, and the result is stated in Theorem 11.

**Theorem 8** (Efficient AGSP, (DG)). *Let $H$ be a local Hamiltonian satisfying Assumption (DG), $\{1,\ldots,n\} = L \cup M \cup R$, where $L = \{1,\ldots,i_1\}$, $M = \{i_1+1,\ldots,i_2\}$, and $R = \{i_2+1,\ldots,n\}$, a partition of the $n$-qudit space, and $\varepsilon'_M$ an estimate for the minimal energy $\varepsilon_M$ of the restriction of $H$ to $\mathcal{H}_M$ such that $|\varepsilon_M - \varepsilon'_M| \le 10$. Then the procedure* GENERATE$(H, M, \varepsilon'_M)$ *described in Fig. 1 returns*

- *MPO representations for a collection of $D^2$ operators $\{A_i\}_{i=1}^{D^2}$ acting on $\mathcal{H}_M$ and of bond dimension at most $n^{\tilde{O}(\gamma^{-2})}$ such that there exists a subspace $\tilde{T}$ for which the conclusions of Theorem 2 are satisfied;*
- *An MPO for an operator $\tilde{H}_M$ such that $\|\tilde{H}_M\| = O(\gamma^{-1}\log\gamma^{-1})$ and the minimal energy $\tilde{\varepsilon}_M$ of $\tilde{H}_M$ restricted to $\tilde{T}$ satisfies $|\varepsilon_M - \tilde{\varepsilon}_M| < 1/2$.*

GENERATE$(H, M, \varepsilon'_M, (\eta_1, \mu))$: $H$ a Hamiltonian, $M = \{i_1 + 1, \ldots, i_2\}$ a subset of qudits, $\varepsilon'_M$ an energy estimate for $H_M$, and $(\eta_1, \mu)$ energy parameters used only in the (LD) case.

1. **Soft truncation:** Set $\ell$ as in (34) in case (FF), (22) in case (DG), and (28) in case (LD). Set $J_M$ as in (23). In case (FF), construct an MPO for the truncated Hamiltonian as in Definition 8. In case (DG) and (LD), construct an MPO for the soft-truncated Hamiltonian $\tilde{H}_M$ via the cluster expansion (see Definition 7 and Lemma 11).
2. **Chebyshev polynomial:** Compute MPO representations for operators $\{A_i\}$ acting on $M$ using the decomposition of the Chebyshev polynomial provided in Proposition 4, using energy parameters specified in (35) in case (FF), (25) in case (DG), and (29) in case (LD).

Return the MPO representations for $\tilde{H}_M$ and for the $\{A_i\}$.

**Fig. 1.** The GENERATE procedure

*Moreover,* GENERATE$(H, M, \varepsilon'_M)$ *runs in time* $n^{\tilde{O}(\gamma^{-2})}$.[6]

*Proof.* We construct an AGSP $K$ from which the operators $\{A_i\}$ claimed in the theorem will be derived. The construction follows very closely the one employed in the proof of Theorem 2, replacing the use of hard truncation by soft truncation.

The first step in GENERATE consists in truncating the Hamiltonian associated to each of the three regions. For this, introduce truncation parameters

$$t = \Theta(\ell), \qquad t' = 4, \tag{21}$$

a width parameter

$$\ell = \Theta(\gamma^{-1} \log \gamma^{-1}), \tag{22}$$

and define a Hamiltonian $\tilde{H} = \tilde{H}_{t',t}$ by applying the soft truncation transformation described in Definition 7 thrice, to the regions

$$J_L = \{1, \ldots, i_1 - \ell - 1\}, \quad J_M = \{i_1 + \ell + 1, \ldots, i_2 - \ell - 1\}, \quad J_R = \{i_2 + \ell + 1, \ldots, n\} \tag{23}$$

respectively (provided each region is non-empty). The resulting truncated Hamiltonian $\tilde{H}$ has norm $O(\ell + t \log t') = O(\ell)$. Note that the computation of the complete Hamiltonian $\tilde{H}$ requires estimates for the ground energies of the restriction of $H$ to each of the three regions that are being truncated. We will only need to efficiently compute an MPO for $\tilde{H}_M$, for which a rough estimate for the ground state energy of $H_M$, as provided as input to GENERATE, will be sufficient.

The second step is to apply the Chebyshev polynomial from Definition 5 to $\tilde{H}$ to obtain the AGSP $K$. For this we make a choice of degree

$$k = \ell^2 \tag{24}$$

and set the energy parameters $\eta_0$ and $\eta_1$ to

$$\eta_0 = \varepsilon_0 + \gamma/10, \qquad \eta_1 = \varepsilon_0 + 9\gamma/10. \tag{25}$$

[6] Here and in all our estimates on running times we suppress dependence on the local dimension $d$, which is treated as a constant.

We first verify that $K$ as defined is a spectral AGSP with the required properties, and then we show how it can be computed efficiently. By item 2. from Theorem 1 the scaling parameter $\Delta$ is given by

$$\Delta := 4e^{-4k\sqrt{\frac{8\gamma}{10(\|\tilde{H}\|-(\varepsilon_0+\gamma/10))}}} = e^{-\Omega\left(k\sqrt{\frac{\gamma}{(\ell+t)}}\right)}. \tag{26}$$

Furthermore, applying Theorem 1 twice, once for the region centered at $i_1$ and once for the region centered at $i_2$, the bond parameter $D$ of $K$ across each of the cuts $(i_1 : i_1 + 1)$ and $(i_2 : i_2 + 1)$ is bounded by

$$D \leq (dk)^{O(\ell+k/\ell)} = e^{\tilde{O}\left(\frac{1}{\gamma}\log^3 d\right)}, \tag{27}$$

as desired. Moreover,

$$D^{12}\Delta = e^{\gamma^{-1}\tilde{O}(\log(\gamma^{-1}))}e^{-\Omega(\gamma^{-1}\log^{3/2}(\gamma^{-1}))}$$

can be made smaller than $10^{-5}$ by choosing the implicit constants appropriately.

Next we apply Lemma 10 to evaluate the closeness between the low-energy subspaces of $H$ and $\tilde{H}$. Since $H$ has a spectral gap the subspace $T_{\gamma/20} = H_{[\varepsilon_0,\varepsilon_0+\gamma/20]}$ is the ground space $T$ of $H$. Setting $\delta = 0.05$ the lemma implies that $\tilde{H}_{[\tilde{\varepsilon}_0,\tilde{\varepsilon}_0+\gamma/10]}$ is $\delta$-close to $T$ as long as the constant implied in the definition (21) of the truncation parameter $t$ is large enough. Conversely, we can write $T = T_{\gamma/2} = H_{[\varepsilon_0,\varepsilon_0+9\gamma/10]}$, in which case the lemma implies that $T$ is $\delta$-close to $\tilde{H}_{[\tilde{\varepsilon}_0,\tilde{\varepsilon}_0+\gamma/10]}$. Thus the two spaces are $\delta$-close. The claim on the ground state energies of $H_M$ and $\tilde{H}_M$ follows directly from Lemma 10 and our choice of $t$.

Finally we turn to efficiency, and verify that in time $n^{O(k)} = n^{\tilde{O}(\gamma^{-2})}$ one can construct a set of at most $D^2$ MPO $A_1, \ldots, A_{D^2}$ acting on $\mathcal{H}_M$ such that there exists $B_1, \ldots, B_{D^2}$ acting on $\mathcal{H}_L \otimes \mathcal{H}_R$ such that the AGSP $K$ can be represented as $K = \sum A_i \otimes B_i$. For this we first need to construct MPO representations for the truncated terms in the Hamiltonian. This is provided by Lemma 11 (applied to $H_M - \varepsilon'_M \mathbb{1}$), which given our choice of parameters $t, t'$ guarantees that an MPO providing inverse polynomial approximation (in the operator norm) to $\tilde{H}_M$ can be efficiently computed that has polynomial bond dimension across all cuts. Proposition 4 shows that an efficient construction of MPO for the $A_i$ follows. □

**Theorem 9** (Efficient AGSP, (LD)). *Let $H$ be a local Hamiltonian satisfying Assumption (LD), parameters $\eta_1 \leq \eta$ and $\mu > 0$, $\{1, \ldots, n\} = L \cup M \cup R$, where $L = \{1, \ldots, i_1\}$, $M = \{i_1 + 1, \ldots, i_2\}$, and $R = \{i_2 + 1, \ldots, n\}$, a partition of the $n$-qudit space, and $\varepsilon'_M$ an estimate for the minimal energy $\varepsilon_M$ of the restriction of $H$ to $\mathcal{H}_M$ such that $|\varepsilon_M - \varepsilon'_M| \leq 10$. Then the procedure* GENERATE$(H, M, \varepsilon'_M, (\eta_1, \mu))$ *described in Fig. 1 returns*

- *MPO representations for a collection of $D^2$ operators $\{A_i\}_{i=1}^{D^2}$ acting on $\mathcal{H}_M$ and of bond dimension at most $e^{\tilde{O}(\log^3 n)}$ each such that there exists subspaces $\tilde{T}, \tilde{T}_-$ for which the conclusions of Theorem 3 are satisfied;*
- *An MPO for an operator $\tilde{H}_M$ such that $\|\tilde{H}_M\| = \tilde{O}(\log(n)/\mu)$ and the minimal energy $\tilde{\varepsilon}_M$ of $\tilde{H}_M$ restricted to $\tilde{T}_-$ satisfies $|\varepsilon_M - \tilde{\varepsilon}_M| < 1/2$.*

*Moreover,* GENERATE$(H, M, \varepsilon'_M, (\eta_1, \mu))$ *runs in time $e^{\tilde{O}(\log^3 n)}$.*

*Proof.* The proof is similar to Theorem 8, and the construction of $\tilde{H}$ and $K$ are the same except for a different choice of parameters. Here we choose

$$\ell = \Theta\Big(\frac{\log n}{\mu} \log \frac{\log n}{\mu}\Big), \qquad k = \ell^2 \qquad \text{and} \qquad t = \Theta(\ell), \qquad t' = 4. \tag{28}$$

The truncated Hamiltonian $\tilde{H} = \tilde{H}_{t,t'}$ is obtained as in the proof of Theorem 8, by applying the soft truncation transformation described in Definition 7 thrice. The AGSP $K$ is obtained by applying the Chebyshev polynomial from Definition 5 to $\tilde{H}$, with the energy parameters $\eta_0'$ and $\eta_1'$ defined as

$$\eta_0' = \varepsilon_0 + \eta_1 - \frac{\mu}{2\log n}, \qquad \eta_1' = \varepsilon_0 + \eta_1 \tag{29}$$

respectively. As a result the parameters $D$ and $\Delta$ satisfy

$$D^{12}\Delta = e^{\frac{\log n}{\mu}\tilde{O}\left(\log \frac{\log n}{\mu}\right)} e^{-\Omega\left(\frac{\log n}{\mu} \log^{1.5}\left(\frac{\log n}{\mu}\right)\right)} = o(1),$$

which can be made less than $10^{-5}$ by a proper choice of implied constants. The conditions on closeness of $T$, $T_-$ and $\tilde{T}$, $\tilde{T}_-$ follow from an application of Lemma 10, observing that our choice of truncation parameters $t, t'$ is sufficient to conclude closeness of the appropriate subspaces. The claim on the ground state energies of $H_M$ and $\tilde{H}_M$ follows directly from Lemma 10 as well.

Finally, applying Proposition 4 and Lemma 11 we see that an MPO for the part of $K$ acting on region $M$ can be computed in time $n^{O(k)} = e^{\tilde{O}(\log^2 n)}$. □

*5.4. The frustration-free case.* In this section we give a simpler construction of AGSP specialized to the case of a local Hamiltonian $H = \sum_i h_i$ that is frustration-free with a spectral gap $\gamma > 0$ and a unique ground state $|\Gamma\rangle$. Replacing each $h_i$ by the projection on its range preserves the ground state and, given our usual normalization assumption $0 \le h_i \le \mathbb{1}$, can only increase the spectral gap; thus we may without loss of generality assume that each $h_i$ is a projection.

We define a truncated version of $H$ based on the detectability lemma from [1] as follows.

**Definition 8** (*Truncated Hamiltonian in the frustration-free case*). Suppose given a local Hamiltonian $H$ such that $H = H_J + H_{\overline{J}}$ where $H_J = h_{j_0} + h_{j_0+1} + \cdots + h_{j_1-1}$ is a local Hamiltonian acting on a contiguous set of qudits $J = \{j_0, j_0+1, \ldots, j_1\}$. Let $J_e$ (resp. $J_o$) denote the subset of indices $i \in J$ that are even (resp. odd). Define $H_{J,e} := \sum_{i \in J_e} h_i$ and $H_{J,o} := \sum_{j \in J_o} h_i$. Then the truncation of $H_J$ is given by $\tilde{H}_J := \tilde{H}_{J,e} + \tilde{H}_{J,o}$, where

$$\tilde{H}_{J,e} := \mathbb{1} - \otimes_{i \in J_e}(\mathbb{1} - h_i), \qquad \tilde{H}_{J,o} := \mathbb{1} - \otimes_{i \in J_o}(\mathbb{1} - h_i). \tag{30}$$

The *truncated Hamiltonian* $\tilde{H}$ associated to region $J$ is given by

$$\tilde{H} := \tilde{H}_J + H_{\overline{J}}. \tag{31}$$

Clearly, $\tilde{H}_{J,e}$ and $\tilde{H}_{J,o}$ are projectors and hence their norm is 1. In addition, they are the sum of the identity operator and a product of non-overlapping local terms, and as such, their Schmidt rank is at most $d^2+1$ across any cut. We show that $\tilde{H}$ has the same ground state as $H$, as well as a large spectral gap. This is done through the detectability lemma and its converse stated below.

**Definition 9** (*The detectability lemma operator in 1D*). Let $H = h_1 + \cdots + h_{n-1}$ be a 1D nearest-neighbor Hamiltonian such that each $h_i$ is a projector. Then the *DL operator* of $H$ is defined by

$$\mathrm{DL}(H) := \otimes_i (\mathbb{1} - h_{2i}) \otimes_i (\mathbb{1} - h_{2i+1}).$$

Note that the operator $\mathrm{DL}(H)$ is in general not Hermitian. The usefulness of the definition comes primarily from the detectability lemma:

**Lemma 12** (The detectability lemma). *Let $h_1, \ldots, h_m$ be projectors such that each $h_i$ commutes with all but at most $g$ other $h_j$, and let $H := \sum_i h_i$. For any state $|\psi\rangle$ let $|\phi\rangle := \prod_i (\mathbb{1} - h_i)|\psi\rangle$, where the product is taken in any order. Then*

$$\||\phi\rangle\|^2 \leq \frac{1}{\varepsilon_\phi / g^2 + 1}, \qquad \text{where} \qquad \varepsilon_\phi := \frac{1}{\||\phi\rangle\|^2}\langle\phi|H|\phi\rangle. \tag{32}$$

The version of the detectability lemma stated above is stronger and more general than the one appearing in [1]. It also has a much simpler proof, which is given in [2]. In addition to the detectability lemma, we will use a converse statement which gives a lower bound on the norm of $\mathrm{DL}(H)|\psi\rangle$. The converse, and its proof, appear in [2].

**Lemma 13** (Converse of detectability lemma). *Let $H = \sum_{i=1}^{n-1} h_i$ be a 1D nearest-neighbor Hamiltonian such that each $h_i$ is a projector. Then for any eigenvector $|\psi\rangle$ of $H$,*

$$\|\mathrm{DL}(H)|\psi\rangle\|^2 \geq 1 - 4\varepsilon'_\psi, \quad \text{where} \quad \varepsilon'_\psi := \langle\psi|H|\psi\rangle. \tag{33}$$

With these two lemmas at hand we show the following.

**Theorem 10.** *The truncated Hamiltonian $\tilde{H}$ from Definition 8 satisfies the following:*

1. *$\tilde{H}$ is frustration free and has the same ground state $|\Gamma\rangle$ as $H$.*
2. *The Schmidt rank of $\tilde{H}$ at every cut is at most $d^2+2$.*
3. *$\tilde{H}$ has a spectral gap $\tilde{\gamma} = \Omega(\gamma)$.*

*Proof.* Property 1. follows from the definition. For property 2. note first that the Schmidt rank of every operator on two $d$-dimensional qudits is at most $d^2$. This implies that the Schmidt rank of $\tilde{H}$ at every cut in $\overline{J}$ is at most $d^2+2$: we get a $d^2$ contribution from the local term that is defined on the cut and the extra 2 comes from terms to the right/left of the cut. Consider now a cut between $i, i+1$ for an even $i$ that is in $J$. Since $i$ is even $\tilde{H}_{J,e}$ will contribute at most $d^2$, and $\tilde{H}_{J,o}$ at most 1. The terms in $H_{\overline{J}}$ contribute at most 1 as well, giving the claimed bound of $d^2+2$.

To prove 3. let $|\psi\rangle$ be orthogonal to $|\Gamma\rangle$. By the detectability lemma applied to $H$, $\|\mathrm{DL}(H)|\psi\rangle\| \leq \frac{1}{\gamma/4+1}$. By the converse of the detectability lemma applied to $\tilde{H}$, $\|\mathrm{DL}(\tilde{H})|\psi\rangle\| \geq 1 - 4\tilde{\gamma}$. Since by construction $\mathrm{DL}(H) = \mathrm{DL}(\tilde{H})$, this implies

$$\tilde{\gamma} \geq \frac{1}{4}\Big(1 - \frac{1}{\gamma/4+1}\Big),$$

from which the claim follows. □

The following is an analogue of Theorem 8 which provides a more efficient construction for the frustration-free case.

**Theorem 11** (Efficient AGSP, (FF)). *Let $H$ be a local Hamiltonian satisfying Assumption (FF) and $\{1,\ldots,n\} = L \cup M \cup R$, where $L = \{1,\ldots,i_1\}$, $M = \{i_1+1,\ldots,i_2\}$, and $R = \{i_2+1,\ldots,n\}$ a partition of the $n$-qudit space. Then the procedure* GENERATE$(H, M)$ *returns MPO representations for a collection of $D^2$ operators $\{A_i\}_{i=1}^{D^2}$ acting on $\mathcal{H}_M$ such that the following hold:*

- $D = 2^{\tilde{O}(\gamma^{-1}\log^3 d)}$;
- *There is $\Delta > 0$ such that $D^{12}\Delta < 10^{-5}$ and for any $S \subseteq \mathcal{H}_M$ that is $\delta$-viable for $\{|\Gamma\rangle\}$ it holds that $S' = \mathrm{Span}\{\cup_i A_i S\}$ is $\delta'$-viable for $T$ with $\delta' = \frac{\Delta}{(1-\delta)^2}$;*
- *Each $A_i$ has bond dimension at most $2^{\tilde{O}(\gamma^{-2}\log^5 d)}$.*

*Moreover, for constant $d$ and $\gamma > 0$ the procedure* GENERATE$(H, M)$ *runs in time $n^{(1+o(1))}$.*

*Proof.* We construct a suitable AGSP $K$ from which the operators $A_i$ will be derived. The first step consists in truncating the Hamiltonian associated to each of the three regions. For this, introduce a width parameter

$$\ell = \tilde{\Theta}(\gamma^{-1}\log^2 d), \tag{34}$$

and define a Hamiltonian $\tilde{H}$ by applying the truncation scheme described in Definition 8 thrice, to the regions $J_L = \{1,\ldots,i_1-\ell-1\}$, $J_M = \{i_1+\ell+1,\ldots,i_2-\ell-1\}$ and $J_R = \{i_2+\ell+1,\ldots,n\}$ respectively (provided each region is non-empty). Based on Theorem 10 the resulting truncated Hamiltonian $\tilde{H}$ has norm $O(1)$, the same ground state as $H$, and a spectral gap $\tilde{\gamma} = \Theta(\gamma)$.

$K$ is obtained by applying Definition 5 to $\tilde{H}$ with

$$\eta_0 = 0, \qquad \eta_1 = \tilde{\gamma} \tag{35}$$

and $k = \Theta(\ell^2)$. The bound on $D$ follows from Theorem 1, using which one can also verify that the desired trade-off $D^{12}\Delta < 10^{-5}$ will be achieved provided the right choice of constants is made in the choice of $\ell$.

By Theorem 10 $\tilde{H}$ can be represented as an MPO with bond dimension $O(d^2)$, from which it follows that we can compute a decomposition $K = \sum L_i \otimes A_i \otimes R_i$ where each $A_i$ has bond dimension $O(d^k) = e^{\tilde{\Theta}(\gamma^{-2}\log^5 d)}$.

The claim on the running time follows from the estimates provided in Proposition 4.

□

## 6. Algorithms

Equipped with the efficient construction of AGSP described in Sect. 5, we are ready to turn MERGE into an efficient algorithm. The algorithm, LOW- SPACE, follows the outline given in Sect. 3.2, but requires additional ingredients. The first is the use of the procedure GENERATE described in Fig. 1, which creates MPO representations for the spectral AGSP required to perform error reduction. The second is an additional step of *energy estimation*, which computes an energy estimate required by GENERATE.

LOW-SPACE$(H, \delta, (\eta, \mu))$: $H$ a local Hamiltonian acting on $\otimes_{i=1}^{n}\mathbb{C}^d$, $n$ a power of two; $\delta$ an accuracy parameter; $(\eta, \mu)$ energy parameters for the (LD) case.

1. **Initialization:** For $j \in \{1, 2, \ldots n\}$ let $V_j^0$ contain a family of MPS representations for an (arbitrary) basis of $\mathbb{C}^d$, and $\varepsilon'_{0,j} = 0$.
2. **Iteration:** For $i$ from 1 to $\log n$ do:

   For all $j \in \{1, \ldots, n/2^i\}$ do:
   - **Generate:** Let $M = \{(j-1)2^i, (j-1)2^i + 1, \ldots, j2^i - 1\}$ and $\varepsilon'_M = \varepsilon'_{i-1,2j-1} + \varepsilon'_{i-1,2j}$. Set $(\{A_i\}, \tilde{H}_M)$= GENERATE$(H, M, \varepsilon'_M)$ in the (FF) and (DG) cases, and $(\{A_i\}, \tilde{H}_M)$= GENERATE$(H, M, \varepsilon'_M, \eta - (i-1)\mu/\log n, \mu)$ in the (LD) case.
   - **Merge:** Set $V_j^i = \text{MERGE}'(V_{2j-1}^{i-1}, V_{2j}^{i-1}, \{A_i\}, s, (k, \xi)) \subseteq \mathcal{H}_{[(j-1)2^i+1, j2^i]}$, where $(s, k)$ are specified in (36) and $\xi$ should satisfy (37) for the case (DG) and (FF); in case (LD) the procedure MERGE can be used instead.
   - **New Energy Estimation:** Form the subspace $V = \{A_i\}^t \cdot (V_{2j-1}^{i-1} \otimes V_{2j}^{i-1})$, where $t = 4\lceil \log \gamma^{-1} \rceil$. Compute the smallest eigenvalue $\varepsilon'_{i,j}$ of the restriction of $\tilde{H}_M$ to $V$. (This step is not needed in case (FF).)
3. **Final step:** Set $K = (\mathbb{1} - H/\|H\|)$ and $\tau = 10\|H\|\gamma^{-1}\log(1/\delta)$. Choose an orthonormal basis $\{|y_i^{(0)}\rangle\}$ for $V_1^{\log n}$. Repeat for $t = 1, \ldots, \tau$:
   - Set $\{|y_i^{(t)}\rangle\} = \text{Trim}_\xi(\text{Span}\{K|y_i^{(t-1)}\rangle\})$,

   where $\xi$ is as previously in cases (DG) and (FF), and as in (38) in case (LD).

Return $\{|z_i\rangle\}$, the smallest $r$ eigenvectors of $H$ restricted to $W = \text{Span}\{|y_i^{(\tau)}\rangle\}$.

**Fig. 2.** The LOW- SPACE algorithm

The complete algorithm is described in Fig. 2. It takes as input a local Hamiltonian satisfying assumptions (FF), (DG) or (LD) (we assume the algorithm is told which assumption holds) and a precision parameter $\delta$, and returns MPS representations for a viable set that is $\delta$-close to the low-energy space $T$ of $H$.[7]

We note that the LOW- SPACE algorithm described in Fig. 2 already incorporates the modified procedure MERGE' sketched in Sect. 3.2. As described in that section, MERGE' differs from MERGE by adding a step of bond trimming. The reason for the modification is that due to the logarithmic number of iterations, successive applications of MERGE may, even if the $\{A_i\}$ can be applied efficiently, lead to MPS whose bond dimension eventually becomes super-polynomial. The procedure MERGE' is described and analyzed in detail in Sect. 6.1. In Sects. 6.2, 6.3 and 6.4 we build on the analysis of MERGE' and the efficient AGSP constructions from the previous section to show that LOW- SPACE leads to an efficient algorithm under assumptions (DG), (FF) and (LD) respectively.

*6.1. A modified* MERGE *procedure.* The procedure MERGE' is described in Fig. 3. It takes additional trimming parameters $k$ and $\xi$ as input ($k$ and $\xi$ will usually be of order $\log(n)$ and $\text{poly}^{-1}(n)$ respectively).

Correctness of MERGE' (for an appropriate choice of $\xi$) relies on the area laws proven in Sect. 4 and on the analysis of the trimming procedure given in Sect. 2.2.4. We give the analysis for the case of Hamiltonians satisfying assumption (DG) in the next section, for frustration-free Hamiltonians in Sect. 6.3, and for Hamiltonians satisfying assumption (LD) in Sect. 6.3.

[7] The algorithm should also be provided a lower estimate for the gap $\gamma$. If not, it can iterate for different values and return the lowest-energy states found.

MERGE' $(V_1, V_2, \{A_i\}, s, (k, \xi))$: Subsets $V_1 \subseteq \mathcal{H}_A$, $V_2 \subseteq \mathcal{H}_B$ of vectors (represented as MPS), operators $A_i$ acting on $\mathcal{H}_1 \otimes \mathcal{H}_2$ (represented as MPO), $s$ a dimension bound, $k \in \mathbb{N}$ and $\xi > 0$ parameters for the trimming subroutine.

1. **Tensoring:** Set $W$ to be a set of MPS representations for an orthonormal basis for the space $\mathrm{Span}\{V_1 \otimes V_2\}$.
2. **Random Sampling:** Let $W' \subseteq W$ be a random $s$-dimensional subspace of $W$ obtained by applying a random orthogonal transformation to the vectors in $W$ and returning the first $s$ vectors obtained.
3. **Error Reduction:** Set $V = W'$. Repeat $k$ times:
   - Set $V = \mathrm{Trim}_\xi(\mathrm{Span}\{\cup_i A_i W'\})$, where the trimming procedure Trim is described in Definition 4.

Return MPS representations for the vectors in $V$.

**Fig. 3.** The MERGE' procedure

*6.2. Degenerate Hamiltonians.* The following theorem proves the correctness of algorithm LOW- SPACE for the case where the input Hamiltonian satisfies assumption (DG).

**Theorem 12.** *Let $H$ be a local Hamiltonian satisfying Assumption (DG), $T$ its ground space, $r = \dim(T)$ and $\delta \geq \mathrm{poly}^{-1}(n)$. Then with probability at least $1 - \frac{1}{n}$ the set of MPS returned by* LOW- SPACE$(H, \delta)$ *is $\delta$-viable for $T$.*[8] *The running time of the algorithm is $n^{\tilde{O}(\gamma^{-2})}$.*

*Proof.* The proof is based on the same ingredients as the proof of the area law given in Theorem 4. There are two main differences: we must show that the addition of the trimming step in MERGE' does not affect the quality of the viable set returned, and we must verify that the energy estimation step is sufficiently accurate.

We show by induction on $i = 0, \ldots, \log n$ that for all $j \in \{1, \ldots, n/2^i\}$, (i) the set $V_j^{i-1}$ is .015-viable for $T$ and satisfies $|V_j^{i-1}| \leq Ds^2$, for $D$ and $s$ to be specified below, and (ii) $\varepsilon'_{i-1,j}$ is within an additive $\pm 3$ of its true value (the ground state energy of the restriction of $H$ to the corresponding spaces).

Both conditions are satisfied for $i = 0$: for each $j \in \{1, \ldots, n\}$, $V_j^0$ is 0-viable for $T$ with $|V_j^0| = d$, and the energy estimate is accurate since the restriction of the Hamiltonian to a single qudit is identically 0.

Suppose the induction hypothesis verified for $i - 1$, fix $j \in \{1, \ldots, n/2^i\}$, and let $M$ be the region defined in the algorithm. Correctness of the energy estimates $\varepsilon'_{i-1,2j-1}$ and $\varepsilon'_{i-1,2j}$ at step $(i - 1)$ implies that $\varepsilon'_M$ is within $\pm 7$ of the correct value $\varepsilon_M$. By Theorem 8, GENERATE returns a set of $D^2$ operators $\{A_i\}$ with the properties stated in Theorem 2.

At this stage we are exactly in the same setting as for the proof of Proposition 2, except for the additional trimming step in MERGE'. Following that proof we conclude that, prior to the trimming step, the merged set $V_j^i$ is .01-viable for $T$ with probability $1 - e^{-\Omega(s)} \geq 1 - \frac{1}{n^2}$ provided $s = \Omega(r \log(q/s) + \log n)$. We choose

$$s \geq 1600 r(\log r + 1) \qquad \text{and} \qquad k = \frac{1}{2}\lceil \log_D(s) \rceil. \tag{36}$$

[8] The probability of success can be improved to $1 - \mathrm{poly}^{-1}(n)$ by scaling the parameter $s$ used in the algorithm by an appropriate constant.

This choice of $k$ ensures $s^2 \leq D^{2k}s \leq Ds^2$, so that the bound on the dimension of $V_j^i$ required to establish the induction hypothesis holds.

It remains to verify the quality of $V_j^i$ as a viable set. Note first that Theorem 4 allows us to bound the bond dimension $b$ of any vector in $T$ by a polynomial, at the expense of replacing $T$ by a set that is $10^{-4}$-close to $T$. Then the analysis given in Lemma 7 shows that the effect of the trimming can be incorporated by replacing the error reduction parameter $\Delta$ associated with the $\{A_i\}$ by $(\Delta + \sqrt{nrb}\xi)$. Choosing $\xi$ such that

$$\sqrt{nrb}\xi < 10^{-4}\delta D^{-12}, \tag{37}$$

the remaining calculation applies and yields that $V_j^i$ is .015-viable for $T$.

Once this has been established, an application of the third item from Theorem 2 shows that given the choice of $t$ made in the algorithm the subspace $V$ obtained after the energy estimation step is $O(\gamma^2)$-viable for $\tilde{T}$. Using that $\|\tilde{H}_M\| = O(\gamma^{-1}\log\gamma^{-1})$ it follows that $\varepsilon'_{i,j}$ is within an arbitrarily small constant of the minimal energy of $\tilde{H}_M$ restricted to $\tilde{T}$. Using the guarantee from Theorem 8, $\varepsilon'_{i,j}$ is within $\frac{3}{2}$ of the minimal energy $\varepsilon_M$ of the restriction of $H$ to $\mathcal{H}_M$. This completes the inductive step.

We have shown that the iterative step succeeds with probability at least $1 - 1/n^2$; since there are a total of $n$ such merging steps, applying a union bound the set $V_1^{\log n}$ is .015-viable with probability at least $1 - \frac{1}{n}$.

To conclude it remains to analyze the final error improvement step. Let $|\psi\rangle$ be an eigenvector of $H$ with eigenvalue $\varepsilon_0$, and $|v\rangle \in V_1^{\log n}$ such that $|v\rangle = \alpha|\psi\rangle + \sqrt{1-|\alpha|^2}|v^\perp\rangle$, where $\alpha \geq 0.9$ and $|v^\perp\rangle$ is supported on eigenvectors of $H$ with eigenvalue at least $\varepsilon + \gamma$. Following the same analysis as given in the proof of Lemma 6 it follows that after renormalization the overlap of $K|v\rangle/\|K|v\rangle\|$ with $|v\rangle$ has improved from $\alpha$ to

$$\frac{\alpha^2}{\alpha^2 + (1-\alpha^2)(1-\gamma/\|H\|)} = \frac{\alpha^2}{1-\gamma(1-\alpha^2)/\|H\|} \geq \alpha^2\Big(1 + \frac{\gamma}{2\|H\|}\Big).$$

Thus the set $K\{|y_i^{(1)}\rangle\}$ is $0.9(1 + \gamma/(2\|H\|))$-viable for $T$. Assuming $\xi$ is chosen small enough (satisfying (37) suffices), by Lemma 7 the set $\{|y_i^{(2)}\rangle\}$ will remain $0.9(1 + \gamma/(3\|H\|))$-viable for $T$. Repeating this procedure $\tau$ times yields a set $W$ that is $\delta$-viable for $T$. Finally, each of the $r$ vectors $|z_i\rangle$ returned by the algorithm must have energy at most $\varepsilon_0 + \delta\gamma$, which using the spectral gap condition implies that $\mathrm{Span}\{|z_i\rangle\}$ and $T$ are mutually $\delta$-close.

The algorithm requires only a polynomial number of operations on MPS representations of vectors. Due to trimming, all these vectors have polynomial bond dimension and thus each operation can be implemented in polynomial time. The complexity is dominated by the complexity of the procedure GENERATE and the application of the operators $A_i$, which is $n^{\tilde{O}(\gamma^{-2})}$. □

*6.3. Frustration-free Hamiltonians with a unique ground state.* The most computation-intensive step of the LOW- SPACE algorithm is the construction, via GENERATE, and subsequent application in MERGE', of the set of operators $\{A_i\}$. In the special case where the Hamiltonian $H$ satisfies Assumption (FF), i.e. $H$ is frustration-free and has a spectral gap, the operators $\{A_i\}$ can be constructed very efficiently, yielding improved

bounds on the running time. The overall algorithm remains as described in Fig. 2, with GENERATE instantiated with the efficient procedure described in Theorem 11.

**Theorem 13.** *Let $H$ be a local Hamiltonian satisfying Assumption (FF), $|\Gamma\rangle$ the unique ground state of $H$, and $\delta = n^{-\omega(1)}$. With probability at least $1 - \frac{1}{n}$ the lowest-energy vector $|z\rangle$ returned by* LOW- SPACE$(H, \delta)$ *satisfies $|\langle z|\Gamma\rangle| \geq 1 - \delta$. Moreover the algorithm runs in time $O(n^{1+o(1)}\mathrm{M}(n))$, where $\mathrm{M}(n) = O(n^{2.38})$ denotes matrix multiplication time.*

*Proof.* The proof follows very closely the proof of Theorem 12, and we only indicate the main differences. To ensure the algorithm is efficient, it is important to choose the trimming parameter $\xi$ to be as large as possible. It follows from the area law for 1D gapped systems [3] (see also Theorem 4 for $r = 1$) that the ground state $|\Gamma\rangle$ of $H$ can be approximated up to accuracy $\mathrm{poly}^{-1}(n)$ by a matrix product state with sub-linear bond dimension. Thus by Lemma 7, using that $r, s$ are both constant, and treating $d, \gamma$ as constants, a choice of $\xi = n^{-(1/2+\omega(1))}$ satisfying (37) will suffice to ensure the error remains negligible, while also maintaining the property that all MPS manipulated have quasi-linear bond dimension. The essential operations on such vectors required in the algorithm, such as multiplication by an MPO $A_i$ of constant bond dimension, or writing in canonical form, can all be computed in time $\tilde{O}(n\mathrm{M}(B))$ where $\mathrm{M}(B)$ is matrix multiplication time for $B \times B$ matrices and $B$ is an upper bound on the bond dimension of the MPS being manipulated; $\mathrm{M}(B)$ corresponds to the cost of performing individual singular value decompositions on the tensors that form each of the MPS. The claim on the running time follows since the number of iterations of the algorithm is logarithmic. □

*6.4. Gapless Hamiltonians.* We extend the analysis of the LOW- SPACE algorithm to the case of gapless Hamiltonians satisfying the (LD) assumption. The main obstacle, of course, consists in dealing with a gapless system. What makes it possible to tackle this case are the strong properties of a viable set. Suppose that $S$ is a viable set for $T$, the set of states of energy at most $\eta$. Then $S$ is also a viable set for $T'$, the set of states of energy at most $\eta - \mu$ for an arbitrary choice of $\mu$. Now, if we apply an AGSP which amplifies the norm of states with energy less than $\eta - \mu$, and decreases the norm of states with energy greater than $\eta$, this is guaranteed to improve the quality of the viable set. This is because by Lemma 1, for each state in $T'$ the viable set $S$ contains an approximation to that state that is guaranteed to have no projection onto the orthogonal complement of $T'$ in $T$. In this sense, regarding $S$ as a viable set for $T'$ creates a virtual spectral gap $\mu > 0$.

Due to the absence of a constant spectral gap, and our introduction of an "artificial" gap of order $1/\log n$, the procedure now runs in quasipolynomial time $e^{\tilde{O}(\log^3 n)}$. The result is stated in the following theorem.

**Theorem 14.** *Let $H$ be a local Hamiltonian satisfying Assumption (LD), $\eta > 0$ the associated energy parameter, $\mu = \Omega(1/\log n)$ and $\delta \geq \mathrm{poly}^{-1}(n)$. With probability at least $1 - \frac{1}{n}$ the set $\{|z_i\rangle\}$ returned by* LOW- SPACE$(H, \delta, (\eta, \mu))$ *is an orthonormal set of $r$ states each having energy at most $\varepsilon_0 + \eta - \mu + \delta$ with respect to $H$. The algorithm runs in time $2^{\mathrm{poly}\log(n)}$.*

*Proof.* The proof follows closely that of Theorem 12 with the following simple modifications: Theorem 8 is replaced by Theorem 9, there is no need to introduce MERGE'

(since the final running time we are obtaining is already $n^{\text{poly}\log n}$ anyways), and finally Theorem 4 is replaced by Theorem 5; as a consequence any choice of $\xi$ for the final step that is of order

$$\xi = e^{-\log^{1+\omega(1)} n} \tag{38}$$

will suffice to guarantee that trimming induces an error that is negligible compared to $\delta = \text{poly}^{-1}(n)$. □

We note that we cannot make the stronger conclusion that the $r$ vectors $|z_i\rangle$ returned by the algorithm are low-energy eigenstates; while it does hold that each must have energy at most $\varepsilon_0 + \eta - \mu + \delta$ (since the closest vectors to $H_{[\varepsilon_0, \varepsilon_0+\eta-\mu]}$ in $W$ will have this property), in the absence of a spectral gap for $H$ the $|z_i\rangle$ may still be constituted of a mixture of low-energy eigenstates with energy slightly higher than $\varepsilon_0 + \eta - \mu + \delta$.

## 7. Discussion

We have introduced a framework for designing algorithms (and proving area laws) by combining procedures for efficiently manipulating viable sets. The scope and efficiency of the resulting algorithms depend upon the efficiency of these procedures. The central limiting component is the efficiency of the underlying AGSP constructions: any substantial improvement of the parameters of our constructions would almost automatically lead to improved area laws, faster algorithms, possibly for scenarios that we are currently unable to handle. This naturally leads to a program of determining the ultimate limits for these parameters and efficiency bounds, and in particular to the following questions:

1. What is the best $D - \Delta$ trade-off achievable for an AGSP, depending on assumptions placed on the local Hamiltonian from which it originates? Currently our trade-offs take the form $2^{\log^{3/2-\mu} D}\Delta < 1$ for arbitrarily small constant $\mu$. Is a better tradeoff achievable, with a larger exponent than $3/2$? Note that currently we only make use of trade-offs of the form $D^c\Delta < 1$ for constant $c$, which is already implied by the above with exponent 1 instead of $3/2$. Improving the exponent could help make progress towards an area law for 2D systems. For a given trade-off, a related question asks for the smallest value of $D$ for which $D\Delta < 1$: this value is important for the efficiency of the algorithm, and also directly enters the parameters obtained for the area law by the bootstrapping argument. Currently, our constructions achieve $D \simeq \exp(\log^3 d/\gamma)$.
2. The soft truncation procedure used for the AGSP construction for our algorithms achieves a $\text{poly}(n)$ bond dimension at all cuts. Could that dimension be lowered, perhaps to polylogarithmic at all cuts?
3. Is it possible to construct an AGSP with a favorable $D - \Delta$ trade-off, not only at one, two, or a constant number of pre-specified cuts, but simultaneously at every (or a constant fraction of) cut?
4. Our trimming procedure for viable sets is not completely satisfactory, and its dependence on the number of cuts as well as on the viable space dimension could potentially be improved. Could the more simple trimming procedure of [22] also be applied in this setting?
5. For the case of an MPS approximation to a unique ground state, the parallel trimming procedure used in [22] yields a bound on the trimmed bond dimension that depends inverse-linearly on the desired approximation error, multiplied by the number of bonds trimmed. It is not impossible that the same procedure would be more effective

than proven, with a cost that does not scale with the number of bonds. Such a procedure could yield a nearly-linear time algorithm for the frustration-free case.

6. What implications can be drawn from our results for the challenging scenario of ground states of local Hamiltonians in higher dimension—e.g. on 2D lattices? Difficulties such as the efficiency of contracting 2D PEPS networks present significant obstacles to any algorithmic procedure; nevertheless, it could be that our bootstrapping arguments could be ported to yield mild area laws in higher dimensions.
7. The tensor network picture of our algorithm may have an interesting interpretation in terms of the bulk-boundary correspondence in AdS/CFT (see e.g. [23,26]). Specifically, the physical qubits would constitute the "boundary", and are acted on directly by the AGSP, while the bulk degrees of freedom are the ones that are subject to the random sampling.

*Acknowledgements.* We thank Andras Molnar for comments on an earlier draft of this paper, and Christopher T. Chubb for comments and the permission to include the suggestive pictures representing the tensor network structure of the isometry produced by our algorithms. We thank the anonymous referees for valuable comments that greatly helped improve the presentation of the paper. I. Arad's research was partially performed at the Centre for Quantum Technologies, funded by the Singapore Ministry of Education and the National Research Foundation, also through the Tier 3 Grant random numbers from quantum processes. Z. Landau and U. Vazirani acknowledge support by ARO Grant W911NF-12-1-0541, NSF Grant CCF-1410022 and Templeton Foundation Grant 52536 T. Vidick was partially supported by the IQIM, an NSF Physics Frontiers Center (NFS Grant PHY-1125565) with support of the Gordon and Betty Moore Foundation (GBMF-12500028).



## Appendix A: Constructing an MPO for the cluster expansion

In this appendix we provide the proof of Theorem 6 and Theorem 7 from Sect. 5.1.

*Proof (Proof of Theorem 6).* For an integer $m \geq 1$ we let $\rho_m$ be the summation of $f(w)$ over all words $w$ such that there exists $m$ disjoint intervals, each of length at least $r$, such that the support of $w$ contains each interval but does not contain the two qudits that lie immediately to the left and right of the interval (we call these two qudits the "boundary" of the interval). Using the inclusion-exclusion principle one can verify that

$$e^{-\beta H} - M_r(H) = -\sum_{m=1}^{\infty} (-1)^m \rho_m. \tag{39}$$

We bound the operator norm of each $\rho_m$ individually. Write $\rho_m = \sum_{I=\{I_1,\ldots,I_m\}} \rho_I$, where the summation is over all $m$-tuples of disjoint intervals $I_1, \ldots, I_m$ of length at least $r$, and each $\rho_I$ contains all those $h_w$ for which the support of $w$ contains each of the intervals $I_i$ but not its boundary and is arbitrary everywhere else. Very roughly, the summation is over at most $n^{2m}/(m!)$ terms. Using that the boundaries are excluded, it is not hard to see that $\rho_I = e^{-\beta H_{\bar{I}}} \prod_{j=1}^{m} \eta(I_j)$, where $H_{\bar{I}}$ contains all terms in the Hamiltonian that do not act on the qudits in the boundary of $I_j$ and $\eta(I_j)$ is the sum of all $f(w)$ such that the support of $w$ is exactly $I_j$. Using $\|e^{-\beta H_{\bar{I}}}\| \leq 1$ we can bound

$$\|\rho_I\| \leq \prod_j \|\eta(I_j)\|$$

$$\leq \prod_j \Big( \sum_{w:\, supp(w)=I_j} \frac{(-\beta)^{|w|}}{|w|!} \Big)$$
$$= \left(e^{\beta} - 1\right)^{\sum_j |I_j|}.$$

Combining with (39),

$$\begin{aligned} \|e^{-\beta H} - M_r(H)\| &\leq \sum_{m=1}^{\infty} \|\rho_m\| \\ &\leq \sum_{m=1}^{\infty} \sum_{I=\{I_1,\ldots,I_m\}} \|\rho_I\| \\ &\leq \sum_{m=1}^{\infty} \frac{n^{2m}}{m!} \left(e^{\beta} - 1\right)^{mr} \\ &= e^{n^2 (e^{\beta}-1)^r} - 1\,, \end{aligned}$$

where for the third line we used that $\beta$ is such that $e^{\beta} - 1 < 1$. □

*Proof (Proof of Theorem 7).* The $r^{\text{th}}$ expansion of $e^{-\beta H}$ is given by

$$M_r(H) := \sum_{w \in S_r} \frac{(-\beta)^{|w|}}{|w|!} h_w\,,$$

where $w$ is a word on the alphabet of local Hamiltonian terms $\{1, \ldots, n-1\}$, $h_w := \prod_{i \in w} h_i$, and $S_r$ is the set of words in which all connected components have a support of size at most $r-1$. Let $\underline{I} = (I_1, I_2, \ldots, I_m)$ be a collection of disjoint segments on the line, and $\max(\underline{I})$ denote the length of the largest segment in $\underline{I}$. We write $w \in \underline{I}$ to mean that the connected components of $w$ matches the segments specified by $\underline{I}$. Using this notation, $M_r(H)$ can be rewritten as

$$M_r(H) = \sum_{\max(\underline{I}) < r} \sum_{w \in \underline{I}} \frac{(-\beta)^{|w|}}{|w|!} h_w.$$

A rather straightforward combinatorial argument shows that for a given $\underline{I} = (I_1, \ldots, I_m)$,

$$\sum_{w \in \underline{I}} \frac{(-\beta)^{|w|}}{|w|!} h_w = \prod_{j=1}^{m} \sum_{w \in I_j} \frac{(-\beta)^{|w|}}{|w|!} h_w\,,$$

where the notation $w \in I_j$ means that the support of the word $w$ has a single connected component whose support is $I_j$. Therefore, if we define for each segment $I$

$$\rho_I := \sum_{w \in I} \frac{(-\beta)^{|w|}}{|w|!} h_w\,, \tag{40}$$

then

$$M_r(H) = \sum_{\max(\underline{I}) < r} \rho_{I_1} \otimes \rho_{I_2} \otimes \cdots \otimes \rho_{I_m}. \tag{41}$$

We use (41) as the basis for an efficient MPO representation of $M_r(H)$.

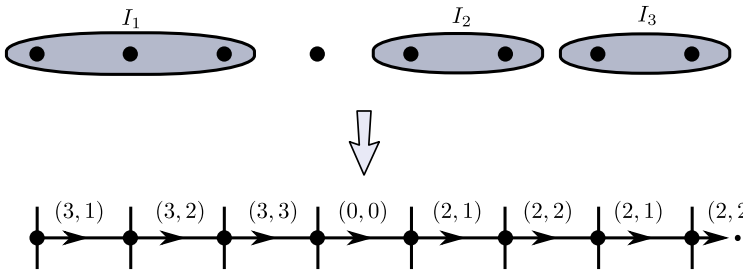



**Fig. 4.** An example of the $(\ell, k)$ indices that give rise to the configuration of segments $I_1 = (1, 2, 3)$; $I_2 = (5, 6)$; $I_3 = (7, 8)$

*1st step: creating a table of $\rho_I$.* The first step is a pre-processing step, which can be run performed before the start of the algorithm. Its goal is to create a table of MPO representations of *all* $\rho_I$ that appear in (41). This can be done in $nd^{O(r)}$ time. Indeed, note first that the total number of intervals $I$ to consider is at most $nr$. The associated MPO can be computed iteratively, starting with $I = \emptyset$ for which $\rho_\emptyset = \mathbb{1}$. Assuming all $\rho_I$ with $|I| < s$ have been determined, compute an MPO for $\rho_I$, for any $I$ such that $|I| = s$, as follows. Clearly,

$$\rho_I = e^{-\beta H_I} - \sum_{\underline{I}'} \rho_{I'_1} \otimes \rho_{I'_2} \cdots \otimes \rho_{I'_m},$$

where the summation runs over all disjoint subsets $\underline{I}' = (I'_1, I'_2, \ldots I'_m)$ included in $I$ and with $m \geq 2$. An MPO for the first term can be obtained in time $d^{O(s)}$ by direct matrix exponentiation. The second term is expressed as the sum of most $2^s$ terms, for each of which an MPO was computed in a previous iteration. Altogether $\rho_I$ can therefore be computed in time $d^{O(s)}$ and stored in memory as an MPO of bond dimension at most $d^r$.

*2nd step: creating the MPO of $M_r(H)$.* We follow the expansion (41), using a signaling mechanism through which every site tells the site to its right to which $\rho_I$ it belongs. This ensures that every non-vanishing contraction of the virtual indices corresponds to exactly one product $\rho_{I_1} \otimes \rho_{I_2} \otimes \cdots$ from (41).

Virtual bonds are indexed by triples $(\ell, k, \alpha)$. The virtual bond across sites $a, a+1$ describes the segment $I$ to which $a$ belongs: $\ell \in \{0, 1, \ldots, r-1\}$ denotes the width of $I$, $k \in \{1, \ldots, r-1\}$ denotes the position of the site $a$ within $I$, and $\alpha$ corresponds to the index of the virtual bond in the MPO expansion of $\rho_I$. For example, suppose that site $a$ is in third position in the support of $\rho_I$, where $|I| = 8$. Then it transmits to site $a+1$ the indices $\ell = 8, k = 3$. Site $a+1$ will then transmit to $a+2$ the indices $\ell = 8, k = 4$ and so on. When the last site in $\rho_I$ is reached, in our example site $a+5$, it transmits to $a+6$ the indices $(k = 8, \ell = 8)$. Then $a+6$ could either be an empty site, transmitting $\ell = k = 0$ to the right, or start a new segment $I$ with any $\ell > 0$. See Fig. 4 for an illustration of this signaling mechanism.

To write a formal definition of the MPO, let us use $[A^{(a)}(I)]^{i,j}_{\alpha_1,\alpha_2}$ to denote the tensor associated with $\rho_I$ at site $a \in I$. In order to simplify notation, when the site $a$ is the left-most (resp. right-most) site in $I$ we use the convention that $[A^{(a)}(I)]^{i,j}_{\alpha_1,\alpha_2}$ is non-vanishing only when $\alpha_1 = 1$ (resp. $\alpha_2 = 1$). Finally, we denote each segment $I$ by $I(\ell, a)$ where $\ell$ is the width of the segment and $a$ is its first site. For a non-extremal site $a$, the tensor $A^{(a)}$ of $M_r(H)$ is given by

$$[A^{(a)}]^{i,j}_{(\ell_1,k_1,\alpha_1),(\ell_2,k_2,\alpha_2)}$$
$$:= \begin{cases} \left[A^{(a)}\big(I(\ell_1, a-k_1+1)\big)\right]^{i,j}_{\alpha_1,\alpha_2} & \text{for } k_1 < \ell_1 \text{ and } \ell_1 = \ell_2, \text{ and } k_2 = k_1+1, \\ \left[A^{(a)}\big(I(\ell_2, a)\big)\right]^{i,j}_{\alpha_1,\alpha_2} & \text{for } k_1 = \ell_1 \text{ and } 0 < \ell_2 \le n-a+1 \text{ and } k_2 = 1, \\ \delta_{i,j} & \text{for } k_1 = \ell_1 \text{ and } \ell_2 = k_2 = 0 \text{ and } \alpha_1 = \alpha_2 = 1, \\ 0 & \text{otherwise.} \end{cases} \tag{42}$$

The first case corresponds to a site $a$ in the interior of the segment $I = I(\ell_1, a-k_1+1)$. The second case corresponds to an $a$ that is the first site of a new segment $I = I(\ell_2, a)$. Note that the condition $\ell_2 \le n-a+1$ guarantees that this segment does not exceed the right side of the chain. Finally, the third case corresponds to an empty site $a$.

To complete the definition it remains to specify $A^{(1)}$ and $A^{(n)}$. Just as the tensors for $\rho_I$, we keep both left and right indices but make them non-zero only when $\ell = k = 0$ and $\alpha = 1$. Then $A^{(1)}$ is defined as $A^{(a)}$ with the additional requirement that it is non-vanishing only when $\ell_1 = k_1 = 0$ and $\alpha_1 = 1$. The tensor $A^{(n)}$ is defined directly by (42). In that case, for every $(\ell_1, p_1, \alpha_1)$ there is at most one triple $(\ell_2, p_2, \alpha_1)$ for which $A^{(n)}$ is non-vanishing, and so without loss of generality we can map it to $\ell_2 = k_2 = 0$ and $\alpha_2 = 1$.

To finish the proof note that the virtual bond dimension is bounded by $r(r-1)d^r < r^2 d^r$, and therefore the second step can be done in time $nd^{O(r)}$ since it only involves local assignments. □

## References


1. Aharonov, D., Arad, I., Vazirani, U., Landau, Z.: The detectability lemma and its applications to quantum hamiltonian complexity. New J. Phys. **13**(11), 113043 (2011)
2. Anshu, A., Arad, I., Vidick, T.: Simple proof of the detectability lemma and spectral gap amplification. Phys. Rev. B **93**(20), 205142 (2016)
3. Arad, I., Kitaev, A., Landau, Z., Vazirani, U.: An area law and sub-exponential algorithm for 1D systems. In: Proceedings of the 4th Innovations in Theoretical Computer Science (ITCS) (2013)
4. Arad, I., Kuwahara, T., Landau, Z.: Connecting global and local energy distributions in quantum spin models on a lattice. Technical report, J. Stat. Mech. **3**, 033301 (2016)
5. Arad, I., Landau, Z., Vazirani, U.: Improved one-dimensional area law for frustration-free systems. Phys. Rev. B **85**, 195145 (2012)
6. Bravyi, S., Gosset, D.: Gapped and gapless phases of frustration-free spin-1 2 chains. J. Math. Phys. **56**(6), 061902 (2015)
7. Bravyi, S., Hastings, M.B., Michalakis, S.: Topological quantum order: stability under local perturbations. J. Math. Phys. **51**(9), 093512 (2010)
8. Bridgeman, J.C., Chubb, C.T.: Hand-waving and interpretive dance: An introductory course on tensor networks. J. Phys. A Math. Theor. **50**, 223001 (2017)
9. Chubb, C.T., Flammia, S.T.: Computing the degenerate ground space of gapped spin chains in polynomial time. Chic. J. Theor. Comput. **9**, 1–35 (2016)
10. Dasgupta, S., Gupta, A.: An elementary proof of a theorem of Johnson and Lindenstrauss. Random Struct. Algorithms **22**(1), 60–65 (2003)
11. de Beaudrap, N., Osborne, T.J., Eisert, J.: Ground states of unfrustrated spin hamiltonians satisfy an area law. New J. Phys. **12**(9), 095007 (2010)
12. Eisert, J., Cramer, M., Plenio, M.B.: Colloquium: area laws for the entanglement entropy. Rev. Mod. Phys. **82**(1), 277–306 (2010)
13. Evenbly, G., Vidal, G.: Tensor network renormalization yields the multiscale entanglement renormalization ansatz. Phys. Rev. Lett. **115**(20), 200401 (2015)
14. Feynman, R.P.: Simulating physics with computers. Int. J. Theor. Phys. **21**(6), 467–488 (1982)
15. Hastings, M.B.: Solving gapped hamiltonians locally. Phys. Rev. B **73**(8), 085115 (2006)
16. Hastings, M.B.: An area law for one-dimensional quantum systems. J. Stat. Mech. Theory Exp. **2007**(08), P08024 (2007)

17. Huang, Y.: A polynomial-time algorithm for the ground state of one-dimensional gapped Hamiltonians (2014). arXiv:1406.6355
18. Huang, Y.: A simple efficient algorithm in frustration-free one-dimensional gapped systems (2015). arXiv:1510.01303
19. Keating, J., Linden, N., Wells, H.: Spectra and eigenstates of spin chain hamiltonians. Commun. Math. Phys. **338**(1), 81–102 (2015)
20. Kliesch, M., Gogolin, C., Kastoryano, J., Riera, M.A., Eisert, J.: Locality of temperature. Phys. Rev. X **4**, 031019 (2014)
21. Kuwahara, T., Arad, I., Amico, L., Vedral, V.: Local reversibility and entanglement structure of many-body ground states. Quantum Sci. Technol. **2**, 015005 (2017)
22. Landau, Z., Vazirani, U., Vidick, T.: A polynomial time algorithm for the ground state of one-dimensional gapped local Hamiltonians. Nat. Phys. (2015). doi:10.1038/nphys3345
23. Maldacena, J.: Eternal black holes in anti-de sitter. J. High Energy Phys. **2003**(04), 021 (2003)
24. Masanes, L.: Area law for the entropy of low-energy states. Phys. Rev. A **80**(5), 052104 (2009)
25. Molnar, A., Schuch, N., Verstraete, F., Cirac, J.I.: Approximating gibbs states of local hamiltonians efficiently with projected entangled pair states. Phys. Rev. B **91**, 045138 (2015)
26. Pastawski, F., Yoshida, B., Harlow, D., Preskill, J.: Holographic quantum error-correcting codes: Toy models for the bulk/boundary correspondence. JHEP **06**, 149 (2015)
27. Roberts, B., Vidick, T., Motrunich, O.I.: Rigorous renormalization group method for ground space and low-energy states of local Hamiltonians. arXiv:1703.01994 (2017)
28. Schollwöck, U.: The density-matrix renormalization group in the age of matrix product states. Ann. Phys. **326**(1), 96–192 (2011)
29. Vershynin, R.: Introduction to the non-asymptotic analysis of random matrices. In: Eldar, Y. et al. (eds.) Compressed Sensing, Theory and Applications. Cambridge University Press. arXiv:1011.3027 (2012)
30. Verstraete, F., Murg, V., Cirac, J.I.: Matrix product states, projected entangled pair states, and variational renormalization group methods for quantum spin systems. Adv. Phys. **57**(2), 143–224 (2008)
31. Vidal, G.: Class of quantum many-body states that can be efficiently simulated. Phys. Rev. Lett. **101**, 110501 (2008)
32. Vidal G.: Entanglement renormalization: an introduction. arXiv:0912.1651 (2009)
33. White, S.R.: Density matrix formulation for quantum renormalization groups. Phys. Rev. Lett. **69**, 2863–2866 (1992)
34. White, S.R.: Density-matrix algorithms for quantum renormalization groups. Phys. Rev. B **48**, 10345–10356 (1993)
35. Wilson, K.G.: The renormalization group: Critical phenomena and the kondo problem. Rev. Mod. Phys. **47**(4), 773 (1975)

Communicated by M. M. Wolf