Computing the Degenerate Ground Space of Gapped Spin Chains in Polynomial Time

Christopher T. Chubb and Steven T. Flammia

Centre for Engineered Quantum Systems, School of Physics,
The University of Sydney, Sydney, NSW 2006, Australia.

May 12, 2016

Abstract

Given a gapped Hamiltonian of a spin chain, we give a polynomial-time algorithm for finding the degenerate ground space projector. The output is an orthonormal set of matrix product states that approximate the true ground space projector up to an inverse polynomial error in any Schatten norm, with a runtime exponential in the degeneracy. Our algorithm is an extension of the recent algorithm of Landau, Vazirani, and Vidick [1] for the nondegenerate case, and it includes the recent improvements due to Huang [2]. The main new idea is to incorporate the local distinguishability of ground states on the half-chain to ensure that the algorithm returns a complete set of global ground states.

Strongly correlated quantum systems are at the heart of such diverse physical phenomena as frustrated magnets, high-temperature superconductors, and topological quantum phases, to name but a few, and understanding their low-temperature behaviour is a grand challenge at the interface of physics and computer science. The computational tools to study these systems numerically are however inevitably stymied by the exponential growth of the Hilbert space of an n-particle system.

Many numerical tools have been developed to surmount the computational difficulties surrounding the exponential growth of the state space. Tools like quantum Monte Carlo [3, 4], coupled cluster methods [5], and the density-matrix renormalisation group (DMRG) [6] all exploit some underlying physical structure to perform heuristically efficient simulations of interacting quantum systems. However, like all heuristics they often lack basic theoretical guarantees except in certain special cases, and some of them have failure modes that are difficult to diagnose or characterise. This makes it hard to precisely determine which physical features enable efficient simulation and which features might be responsible in the cases where the heuristic fails.

Very recently, a success story has emerged in the rigorous understanding of gapped 1D spin chains. Hastings' proof of the 1D area law [7] established that approximating the ground state of a gapped spin chain is in NP. The proof technique crucially used the idea of matrix product states (MPS) [8], an ansatz for quantum systems that exploits the area law structure of entanglement [9]. The first rigorous nontrivial algorithm for finding ground states, due independently to Schuch and Cirac [10] and Aharonov, Arad, and Irani [11], used dynamic programming to find MPS ground states. The runtime of the algorithm was exponential in the bond dimension, and since Hastings' result used a polynomial bond dimension to achieve a sufficiently accurate approximation, the runtime could only be guaranteed to be exponential for general gapped 1D systems. The first subexponential-time algorithm was found by Arad, Kitaev, Landau, and Vazirani [12]. They used the

same algorithm as Refs. [10,11], but showed that the ground state could be approximated to sufficient accuracy with an MPS of bond dimension $n^{o(1)}$, where n is the length of the chain. This line of research culminated in the breakthrough algorithm of Landau, Vazirani, and Vidick (LVV) [1], who gave the first polynomial-time algorithm for approximating the ground state in the case of a non-degenerate ground space. This was subsequently improved by Huang [2] who gave a deterministic algorithm with an exponentially improved scaling with respect to the gap that also applied to highly frustrated systems.

In this paper, we extend the LVV algorithm and its subsequent improvements by Huang to the case where the Hamiltonian has a degenerate ground space. To give a precise statement of our results, we first introduce some notation.

Consider a 1D chain of d-level quantum systems (*qudits*) with $d = \mathcal{O}(1)$ and a Hamiltonian H that is a sum of local terms on contiguous sites, and let n be the number of sites in the chain. After sufficient coarse-graining, shifting the energy, and dividing by the largest operator norm from the terms in the Hamiltonian, we can take the Hamiltonian to be of the form $H = \sum_{i=1}^{n-1} H_i$, where $0 \leq H_i \leq \mathbb{1}$ and H_i acts non-trivially only on the i and $i+1$ qudits. We additionally require that this transformed Hamiltonian has a constant lower bound $\epsilon_1 - \epsilon_0 \geq \epsilon \geq \Omega(1)$ on the spectral gap, i.e. the difference between the first excited energy (ϵ_1) and the ground state energy (ϵ_0) of H is bounded from below by ϵ independent of n. Our Hamiltonians will have a degenerate spectrum in the ground space with a degeneracy $g = \mathcal{O}(1)$. We say that such a Hamiltonian is in *standard form*, though this is not a unique normal form because of the ambiguity in the coarse-graining step.

Our main result is an approximation of the ground space in the form of an orthonormal basis of matrix product states that approximate the true ground space projector. More precisely, we have the following.

Theorem 1 (Main result). *Given an n-qudit Hamiltonian H in standard form and constant bounds ϵ and g on the gap and degeneracy respectively, then for every sufficiently small approximation error η satisfying $1/\eta = n^{\mathcal{O}(1)}$, Algorithm 1 constructs the orthonormal approximate ground states $\{|\gamma_j\rangle\}_j$ in matrix product form for $1 \leq j \leq g$ in runtime $n^{\mathcal{O}(1)}$. The approximation error gives a bound in the Frobenius norm of*

$$\left\| G - \Upsilon \right\|_F \leq \eta \,,$$

where G is the true ground space projector, and $\Upsilon := \sum_{j=1}^{g} |\gamma_j\rangle\langle\gamma_j|$ is the approximate ground space projector.

The proof of this statement is given in Section 2.7, and a more precise dependence of the runtime on the gap ϵ and the ground space degeneracy g is made explicit in Section 2.8. Note that since both G and Υ have rank g, we can get an approximation error in the trace norm at the cost of a factor of $\sqrt{2g} = \mathcal{O}(1)$ increase in error.

Our result relies heavily on the ideas in LVV [1] and the improvements of Huang [2], so let us recall briefly how this algorithm proceeds. The LVV algorithm proceeds by constructing a *viable set*, which roughly speaking is a small set of matrix product states on a half-chain that is guaranteed to contain the local Schmidt vectors of an approximate ground state on the full chain. In each step of the algorithm, the half-chain of the current viable set is extended by one site to create a new viable set. This initial extension is too large, but trimming away some states with too high an energy yields another suitably small set that is still viable.

The main new idea is an extension of the size trimming step to degenerate systems, based on the concept of *local distinguishability* of ground states. The algorithm works by

2

iteratively sweeping across the chain, and finds the degenerate ground states one-by-one in a manner similar to that of LVV on each iteration. As we run these iterations, we must ensure that the viable set of potential ground states on the half-chain contains a witness that not only has low energy, but also does not simply reduce to something in the span of the previously found ground states. To ensure this, we break the analysis into two cases of *high-* and *low-*distinguishability.

In the low-distinguishability case, the Schmidt vectors on the left half-chain of the previously found ground states and the desired witness have large overlap. By recycling these Schmidt vectors, we therefore get a viable set for this new witness for free.

In the high-distinguishability case however, this recycling would not provide a viable set. In this case we need to trim the viable set in a way that not only ensures the existence of a low-energy witness, but a witness which has low overlap with the existing states. If the relevant Schmidt vectors of the desired witness and the existing states have low overlap, then we can optimise for a state of low energy *and* overlap efficiently.

1 Preliminaries

Let us begin by giving a more precise statement of the LVV algorithm, with the definitions suitably extended to our treatment of the degenerate case.

Given the Hamiltonian $H = \sum_{j=1}^{n-1} H_j$, we will often consider a partition of the system into the first i spins and the rest, we will call this the $i|i+1$ cut. The Hamiltonian on a contiguous subset of spins $\{j, \ldots, k\}$ is denoted $H_{[j,k]} = \sum_{i=j}^{k-1} H_i$. In this notation the local terms can also be expressed as $H_i = H_{[i,i+1]}$, and the full Hamiltonian as $H = H_{[1,n]}$. In the special case that an index i is clear from context, we define $H_L = H_{[1,i]}$ and $H_R = H_{[i+1,n]}$. Note that this implies that $H = H_L + H_i + H_R$. The Hilbert space on which $H_{[j,k]}$ acts is denoted by $\mathcal{H}_{[j,k]}$. Similarly, we have the Hilbert space $\mathcal{H} = \mathcal{H}_{[1,n]}$, and similarly for the other possible subscripts. It will sometimes be convenient to subtract a constant energy shift from a Hamiltonian, in which case we define $H_L' = H_L - \epsilon_L$, where ϵ_L is the least eigenvalue of H_L, and similarly for H_R' and ϵ_R. When a particular state ρ is clear from context, we refer to $\text{Tr}(\rho H_L)$ as the left energy of ρ, and similarly for the right energy.

A *viable set* is simply a set of vectors that can be used to construct the Schmidt vectors of a new approximate ground state on a half-chain, assuming we already have approximate ground states $|\gamma_1\rangle, \ldots, |\gamma_{h-1}\rangle$. The aim is to inductively grow the region of viability to the entire system starting from one end of the chain. The span of such a set gives the desired small domain on which candidate ground states can be found efficiently.

Definition 2. (Viable Set) A set of states S is $(h,i,s,b,\delta$ or $\Delta)$-viable if:

- The states in S are supported on the first $i \leq n$ qudits of the system, i.e. $S \subset \mathcal{H}_{[1,i]}$.

- The cardinality is bounded $|S| \leq s$.

- Each state in S is represented by an MPS with bond dimension at most b.

- There exists a witness state $|v\rangle$ which has, depending on context, either

  - (δ case) overlap with the ground space projector G satisfying $\|G|v\rangle\| \geq 1 - \delta$, or

  - (Δ case) the expected energy satisfying $\langle v|H|v\rangle \leq \epsilon_0 + \Delta\epsilon$.

- The reduced density operator of this witness must be supported on $\text{Span}(S)$.

- The witnesses state $|v\rangle$ is orthogonal to all previously constructed approximate ground states $|\gamma_1\rangle, \ldots, |\gamma_{h-1}\rangle$.

When it is clear from the context (and it almost always will be clear), we will drop the explicit h dependence in the definition of viable set and refer to a (i, s, b, δ)-viable set, and similarly for Δ.

The two error parameters specified, δ or Δ, will be used contextually and we will sometimes refer to them as the overlap error and the energy error, respectively. The two error parameters however can be used interchangeably, though at the cost of some overhead which we would like to avoid. The proof is simple and can be found in [1,2].

Lemma 3 (Interchangeability of δ and Δ [1,2])**.** *For the ground state projector G of a Hamiltonian H, and any $|v\rangle \in \mathcal{H}$, we have*

$$\langle v|H|v\rangle \le \epsilon_0 + \Delta\epsilon \qquad \Longrightarrow \qquad \|G|v\rangle\| \ge 1 - \Delta$$
$$\|G|v\rangle\| \ge 1 - \delta \qquad \Longrightarrow \qquad \langle v|H|v\rangle \le \epsilon_0 + 2\delta n$$

which allows us to convert between the two measures of viability

$$(\cdot, \cdot, \cdot, \cdot, \Delta)\text{-viable} \qquad \Longrightarrow \qquad (\cdot, \cdot, \cdot, \cdot, \delta = \Delta)\text{-viable}$$
$$(\cdot, \cdot, \cdot, \cdot, \delta)\text{-viable} \qquad \Longrightarrow \qquad (\cdot, \cdot, \cdot, \cdot, \Delta = \frac{2n}{\epsilon}\delta)\text{-viable}.$$

For later use, we define the following absolute constants c and c' that are independent of all other parameters, and a simple function f of the ground state degeneracy,

$$c := 6.25 \times 10^{-54} \qquad\qquad c' := 10^{-11} \qquad\qquad f := g(2g + 1),$$

where we note that $200\sqrt[4]{c} = c'$. These values can be used to make all of the constants hidden by the big-O notation in our derivations and procedures completely explicit. However, our final result still relies on certain implicit estimates (in particular, the estimates in Refs. [13,14]) so we are not able to give a completely explicit scaling.

2 The Algorithm

Algorithm 1 works by first running the slight modifications on the existing nondegenerate ground state approximation algorithms of Ref. [1,2], which is discussed in more detail in Section 3.1.1. As discussed earlier, the extended algorithm works by iteratively finding ground states one-by-one, with each guaranteed to be orthogonal to those previously found until all g ground states have been found. It is not necessary to know *a priori* a specific value for g, but unless $g = \mathcal{O}(1)$ holds, then there is no guarantee from the area law alone that all the ground states can be approximated well by the algorithm. Moreover, if the algorithm is iterated more times than the true degeneracy of the system, then it will necessarily produce high energy states which will herald this over-estimation, as we will show in Appendix A. As such we will henceforth assume that the degeneracy g is known.

The STEP/FINALSTEP procedures will be discussed below in Section 2.1, and its four sub-procedures will follow in Sections 2.2-2.5. Finally the APPROXGROUNDSTATE procedure will be discussed in Section 2.6.

Algorithm 1 Ground space approximation

1: **procedure** DegenerateGSA(H, g, ϵ, η)
2:     $|\gamma_1\rangle \leftarrow$ NondegenerateGSA(H, ϵ, η)   ▷ Run non-degenerate algorithm
3:     $E \leftarrow \langle\gamma_1|H|\gamma_1\rangle$ ▷ Approximate the ground energy
4:     **for all** $h \in [2, g]$ **do**
5:         $S_{h,0} \leftarrow \{1\}$ ▷ Initial viable set
6:         **for all** $i \in [1, n-1]$ **do**
7:             $E_i \leftarrow \langle\gamma_1|H_{[1,i]}|\gamma_1\rangle$ ▷ Estimate energy of half-chain
8:             $L_{h,i} \leftarrow$ SchmidtVecs($i, |\gamma_1\rangle, \ldots, |\gamma_{h-1}\rangle$)   ▷ Extract left Schmidt vectors
9:             $S_{h,i} \leftarrow$ Step($S_{h,i-1}, L_{h,i}, H, \epsilon, E, E_i$)   ▷ Step viable set one qudit along
10:         **end for**
11:         $S_{h,n} \leftarrow$ FinalStep($S_{h,n-1}, H, \epsilon, E, \eta$)   ▷ Low-error final step
12:         $|\gamma_h\rangle \leftarrow$ ApproxGroundState($H, S_{h,n}, |\gamma_1\rangle, \ldots, |\gamma_{h-1}\rangle$)   ▷ Extract witness
13:     **end for**
14:     **return** $|\gamma_1\rangle, \ldots, |\gamma_g\rangle$
15: **end procedure**

Algorithm 2 Stepping functions

1: **procedure** Step($S_{h,i-1}, L_{h,i}, H, \epsilon, E, E_i$)
2:     $S_{h,i}^{(1)} \leftarrow$ Extend($S_{h,i-1}$)   ▷ Increment i
3:     $S_{h,i}^{(2)} \leftarrow$ Trim($S_{h,i}^{(1)}, L_{h,i}, H, \epsilon, E_i$)   ▷ Reduce cardinality of set
4:     $S_{h,i}^{(3)} \leftarrow$ Truncate($S_{h,i}^{(2)}, \epsilon$)   ▷ Truncate state bond dimensions
5:     $S_{h,i}^{(4)} \leftarrow$ Reduce($S_{h,i}^{(3)}, H, \epsilon, E$)   ▷ Push back down the error level
6:     **return** $S_{h,i}^{(4)}$
7: **end procedure**

8: **procedure** FinalStep($S_{h,n-1}, H, \epsilon, E, \eta$)
9:     $S_{h,n}^{(1)} \leftarrow$ Extend($S_{h,n-1}$)   ▷ Extend to final qudit
10:     $S_{h,n}^{(2)} \leftarrow$ Trim($S_{h,n}^{(1)}, L_{h,n}, H, \epsilon, E_n$)
11:     $S_{h,n}^{(3)} \leftarrow$ Truncate($S_{h,n}^{(2)}, \epsilon$)
12:     $S_{h,n}^{(4)} \leftarrow$ FinalReduce($S_{h,n}^{(3)}, H, \epsilon, E, \eta$)   ▷ Reduce error to desired level
13:     **return** $S_{h,n}^{(4)}$
14: **end procedure**

2.1 Step

The Step procedure, shown above in Algorithm 2, takes a viable set defined on $i-1$ qudits, and constructs a new set which is viable on i qudits. As well as physically extending the set, this procedure can be efficiently performed such that neither the cardinality nor the error grow. Starting from the trivially $(0,1,1,0)$-viable set $\{1\}$, the Step procedure allows you to inductively construct $(i,s,b,\Delta = c\epsilon^6/f^4)$-viable sets $S_{h,i}$, for all i. Here the parameters s and b are all polynomial functions of n and are given in Table 1. On the final qudit the Step procedure is tweaked in the form of FinalStep, which performs a more powerful error reduction procedure to bring the error down to the desired final level.

By constructing a viable set on the entire system, a polynomial-sized subspace has been constructed on which the Hamiltonian may be efficiently optimised, allowing the

Set	i	s	B	δ	Δ
S_0	0	1	1	0	0
S_1	1	s	b	$c\epsilon^6/f^4$	$c\epsilon^6/f^4$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
S_{i-1}	$i-1$	s	b	$c\epsilon^6/f^4$	$c\epsilon^6/f^4$
S_i $\begin{cases} S_i^{(1)} \\ S_i^{(2)} \\ S_i^{(3)} \\ S_i^{(4)} \end{cases}$	i	ds	b	$c\epsilon^6/f^4$	$c\epsilon^6/f^4$
	i	p_1	$dsb+q^2$	$1/100$	$n/50\epsilon$
	i	p_1+q	p_2	$1/5$	$2n/5\epsilon$
	i	pp_1+pq+q	pp_2	$c\epsilon^6/f^4$	$c\epsilon^6/f^4$
S_{i+1}	$i+1$	s	b	$c\epsilon^6/f^4$	$c\epsilon^6/f^4$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
S_{n-1}	$n-1$	s	b	$c\epsilon^6/f^4$	$c\epsilon^6/f^4$
S_n	n	p_1+2q	p_0p_2+q	$\eta^2/4f$	$\eta^2/4f$

Table 1: The viability parameters of the sets produced during Lines 8–14 of Algorithm 1. Red indicates the desired parameter change in each sub-step, grey the weaker of the two error parameters as given by Lemma 3. Here p is some absolute polynomial $p = n^{\mathcal{O}(1)}$ and the remaining polynomials scale as $p_0, p_1, p_2, q = n^a$ where $a = \mathcal{O}\big(1 + \epsilon^{-1}\sqrt{\log \eta^{-1}/\log n}\big)$. The polynomials $s := pp_1 + pq + q$ and $b := pp_2$ denote the fixed points of the cardinality and bond dimension under the STEP procedure.

previously inaccessible witness state to be harvested, yielding the final desired ground state approximation. This final optimisation is performed in APPROXGROUNDSTATE, which follows the STEP-based viable set construction.

2.2 Extension

By induction, suppose we start with a $(i-1, s, b, \Delta = c\epsilon^6/f^4)$-viable set S_{i-1}. The first step is to take a viable set on $i-1$ qudits and extend it trivially to a viable set on i qudits. To do this we take the element-wise tensor product of the existing viable set with a basis on the ith qudit. It is trivial to see that the set resulting from EXTENSION, $S_{h,i}^{(1)}$, is $(i, ds, b, \Delta = c\epsilon^6/f^4)$-viable; see Ref. [1] for the explicit argument. We simply note that the net result of this step is to increase the cardinality of the viable set from s to ds, while keeping all other parameters fixed.

2.3 Size trimming

For the size-trimming we are going to use two nets to compensate for our ignorance of two properties of the ground states: the local energy differences between them, and their entanglement structure across the $i|i+1$ cut. The former is represented by the difference in the expectation value of H_L between our desired witness and (say) the first ground state found. The latter is represented by the boundary contraction (defined below) of the witness state in question. In both cases we are going to construct a net on the space of possible

values of these unknowns, and perform the necessary optimisations in the neighbourhoods of these elements, thus allowing for error guarantees in spite of this ignorance.

2.3.1 Boundary contraction

Viable sets, like reduced density matrices, capture local information about a given state. As defined, viable sets are solely determined by the left Schmidt vectors of a witness state; the reduced density operator however also contains the Schmidt coefficients across the cut. As such, merely finding vectors of low energy on the half-chain $[1, i]$ in no way guarantees consistency with a low-energy global state, as they might not be consistent with the entanglement structure across the $i|i + 1$ cut. As well as this the system may be frustrated, with a globally low-energy state appearing locally excited. Both of these problems are circumvented by performing the desired optimisations alongside *boundary contraction* constraints.

Given a state of Schmidt rank B and Schmidt decomposition across the $i|i + 1$ cut given by $|v\rangle = \sum_{j=1}^{B} \lambda_j |a_j\rangle |b_j\rangle$, let $U_v : \mathbb{C}^B \to \mathcal{H}_R$ be the partial isometry specified by $U_v |j\rangle = |b_j\rangle$. Then we have the following.

Definition 4 (Boundary Contraction [1]). The *left state* of $|v\rangle$ with respect to the $i|i + 1$ cut is defined as

$$|\mathrm{ls}(v)\rangle := U_v^\dagger |v\rangle = \sum_j \lambda_j |a_j\rangle |j\rangle \in \mathcal{H}_L \otimes \mathbb{C}^B \,.$$

The *boundary contraction* of $|v\rangle$ is defined as

$$\mathrm{cont}(v) := \mathrm{Tr}_{[1,i-1]} \left(|\mathrm{ls}(v)\rangle\langle \mathrm{ls}(v)| \right) = U_v^\dagger \, \mathrm{Tr}_{[1,i-1]} \left(|v\rangle\langle v| \right) U_v \,.$$

Note that the density matrix $\mathrm{cont}(v)$ is supported on $\mathcal{H}_i \otimes \mathbb{C}^B$.

The interpretation of a boundary contraction is the following. The difference in right energies between a candidate state and a valid witness state for the full chain is bounded by how close the boundary contractions of the states are. By extending a state by the as-yet-unknown witness state, we can argue the existence of a low energy solution within our extended viable set.

Lemma 5 ([1,2]). *Let σ be a density matrix on $\mathcal{H}_L \otimes \mathbb{C}^B$ and $|v\rangle = \sum_{j=1}^{B} \lambda_j |a_j\rangle |b_j\rangle \in \mathcal{H}$ be a Schmidt-decomposed state. The density matrix $\sigma' := U_v \sigma U_v^\dagger$ on $\mathcal{H}$ satisfies*

$$\begin{aligned}
\mathrm{Tr}\left(\sigma' H\right) \leq \mathrm{Tr}(\sigma H_L) &+ \langle v|(H_i + H_R)|v\rangle \\
&+ \left\| \mathrm{Tr}_{[1,i-1]}\left(\sigma\right) - \mathrm{cont}(v) \right\|_1 \left(1 + \|H_R \Pi_v\| - \langle v|H_R|v\rangle \right) \,.
\end{aligned}$$

where $\Pi_v = U_v U_v^\dagger$ is the rank B projector from the isometry U_v.

Proof. We give a sketch of the proof and refer the reader to Ref. [1] for more detail. Begin with $\mathrm{Tr}(\sigma' H) = \mathrm{Tr}\left[(\sigma' + |v\rangle\langle v| - |v\rangle\langle v|)H\right]$ for some ground state $|v\rangle$. Expanding $H = H_L + H_i + H_R$ and using the matrix Hölder inequality together with the bound $\|H_i\| \leq 1$ yields the proof. □

2.3.2 Nets

Recall that an ϵ-*net* on a metric space is a set of points such that any point in the space is within distance ϵ of a point in the net. We will use two nets in our argument: one on the interval $[-1, 1 + \eta]$ in the standard absolute value metric and one on the set of matrices with bounded max-norm on $\mathbb{C}^d \otimes \mathbb{C}^B$ in the trace norm.

We denote by $\mathcal{M}_\eta$ the $\eta/2$-net on $[-1, 1+\eta]$ whose points are

$$\mathcal{M}_\eta = \{-1, -1+\eta, \cdots, 1-\eta, 1, 1+\eta\} .$$

The cardinality of $\mathcal{M}_\eta$ is at most $\lceil 2/\eta \rceil + 1$, and for any real value $x \in [-1-\eta, 1+\eta]$ there exists a $y \in \mathcal{M}_\eta$ such that the following *one-sided* error bound holds, $0 \le y - x \le \eta$. We refer to this net as the *energy net*. The choice of limits $[-1, 1+\eta]$ is justified in Appendix C.

Let $\mathcal{N}_\eta$ be the matrices on $\mathbb{C}^d \otimes \mathbb{C}^B$ with entries whose real and imaginary parts are in the set

$$\left\{-1, -1+\eta/B^2 d^2, \cdots, 1-\eta/B^2 d^2, 1\right\}$$

The cardinality of $\mathcal{N}_\eta$ is upper bounded by $(2\lceil Bd/\eta \rceil + 1)^{4Bd}$. For any positive semidefinite matrix ρ on $\mathbb{C}^d \otimes \mathbb{C}^B$ with trace at most 1, there exists a point $X \in \mathcal{N}_\eta$ such that $\|\rho - X\|_1 \le \eta$. We refer to this net as the *boundary contraction net*.

2.3.3 Procedure

Let $L_{h,i}$ be the set of left Schmidt vectors with respect to the $i|i+1$ cut of the states $|\gamma_1\rangle, \ldots, |\gamma_h\rangle$. Furthermore define $\Pi_{h,i}$ to be the projector onto $\mathrm{Span}(L_{h,i})$. Now define parameters $\delta = 8\sqrt{c}\epsilon^3/f^2$, $\eta = 4c'\epsilon/f$, and ξ to be chosen later in Eq. (7).

TRIM: For each boundary contraction $X \in \mathcal{N}_{\xi/2}$ with bond dimension B_δ (to be defined below in Eq. (3)) and left-energy difference $Y \in \mathcal{M}_\eta$, take σ to be a density matrix supported on $\mathrm{Span}\left(S_{h,i}^{(1)} \cup L_{h,i}\right) \otimes \mathbb{C}^B$. We will specifically choose σ to be a solution to the *degenerate size-trimming convex program*:

$$\begin{aligned}
\min \quad & \mathrm{Tr}(\Pi_{h,i}\sigma) && (1)\\
\text{where} \quad & \mathrm{Tr}\big[H_L(\sigma - |\gamma_1\rangle\langle\gamma_1|)\big] \le Y, \\
& \big\|\mathrm{Tr}_{[1,i-1]}(\sigma) - X\big\|_1 \le \xi/2, \\
& \mathrm{Tr}(\sigma) = 1, \quad \sigma \ge 0.
\end{aligned}$$

Now consider the set of vectors $\{|v_k\rangle\}_k$ consisting of all the eigenvectors of σ with eigenvalues at least $10^{-9}/g$. Now define

$$S_{h,i}^{(2)} = \big\{|v_{k,j}\rangle\big\}_{j,k} \cup L_{h,i} , \qquad (2)$$

where $|v_{k,j}\rangle$ is the jth left Schmidt vector of $|v_k\rangle$ across the $i|B$ cut.

The next Claim, whose proof is postponed until the next section, establishes that the viability parameters after each TRIM step remain bounded as advertised.

Claim 6 (Proof in Section 3.2)**.** *The set resulting from* TRIM, $S_{h,i}^{(2)}$, *is* $(i, p_1, dsb + q^2, \delta = 1/100)$-*viable, where*

$$p_1(n), q(n) = n^a \text{ where } a = \mathcal{O}\big(1 + \epsilon^{-1}\sqrt{\log \eta^{-1}/\log n}\big) .$$

2.4 Bond truncation

First, we recall some results on low-rank approximate ground states [12, 14, 15]. Let us define a function

$$B_\delta := 2^{\mathcal{O}(\epsilon^{-1} + \epsilon^{-1/4} \log^{3/4} \delta^{-1})} . \qquad (3)$$

Specifically we will take B_δ to be defined by the following lemma, a corollary to the area law [7, 12, 14].

Lemma 7 ([12, 14]). *For any ground state $|\Gamma\rangle$ and bipartite cut of the system in question, there exists a state $|\gamma\rangle$ with Schmidt rank across that cut bounded by B_δ such that $|\langle\Gamma|\gamma\rangle| \geq 1 - \delta$. Moreover, there exists an MPS approximation $|\gamma'\rangle$ with bond dimension across every cut of at most $B_{\delta/n}$ and for which $|\langle\Gamma|\gamma'\rangle| \geq 1 - \delta$.*

If we know the Schmidt decomposition of a given state across a bipartite cut, then the Eckart-Young theorem [16] guarantees that the *truncation* of that state to Schmidt-rank D is the best approximation with said Schmidt rank.

Definition 8 (Truncation). Consider a state $|v\rangle$ with Schmidt decomposition $|v\rangle = \sum_j \lambda_j |a_j\rangle |b_j\rangle$ across the $i|i+1$ cut, where any degeneracy in the Schmidt coefficients is resolved arbitrarily, and with non-increasing λ_j. Given an integer D, define the *truncation* of $|v\rangle$ by

$$\operatorname{Trunc}_D^i |v\rangle := \sum_{j=1}^{D} \lambda_j |a_j\rangle |b_j\rangle .$$

This approximation is optimal in the following sense [16]. For any state $|v\rangle$, and any integer D, the truncated vector $|v'\rangle := \operatorname{Trunc}_D |v\rangle / \|\operatorname{Trunc}_D |v\rangle\|$ satisfies $\langle v|v'\rangle \geq |\langle v|w\rangle|$ for any $|w\rangle$ of Schmidt rank at most D.

Combining these results, we get a precise statement that low-energy states can also be approximated faithfully with low-rank approximations. See Ref. [2] for a proof.

Lemma 9 ([2]). *For a state $|v\rangle$ with $\langle v|H|v\rangle \leq \epsilon_0 + \Delta\epsilon$ where $\Delta(1+\epsilon) \leq 1/4$, the truncated state $|w\rangle := \operatorname{Trunc}_{B_{\Delta\epsilon}}^i |v\rangle / \|\operatorname{Trunc}_{B_{\Delta\epsilon}}^i |v\rangle\|$ has energy $\langle w|H|w\rangle \leq \epsilon_0 + 48\sqrt{\Delta}$.*

Building on these lemmas, we can define the bond truncation procedure TRUNCATE for our algorithm as follows.

TRUNCATE: Take each vector within $S_{h,i}^{(2)}$, and truncate all bonds on $1, \ldots, i-1$ to $P := 800nB_{1/800n} = n^{1+o(1)}$. Take $S_{h,i}^{(3)}$ to be

$$S_{h,i}^{(3)} := \left\{ \operatorname{Trunc}_P^1 \operatorname{Trunc}_P^2 \cdots \operatorname{Trunc}_P^{i-1} |s\rangle \,\middle|\, |s\rangle \in S_{h,i}^{(2)} \right\} \cup L_{h,i} .$$

As before, the following Claim establishes that the TRUNCATE step preserves the viability with the advertised parameters.

Claim 10 (Proof in Section 3.3). *The set resulting from TRUNCATE, $S_{h,i}^{(3)}$, is $(i, p_1 + q, p_2, \delta = 1/5)$-viable, where*

$$p_2(n) = n^a \quad \text{where } a = \mathcal{O}\left(1 + \epsilon^{-1}\sqrt{\log \eta^{-1}/\log n}\right).$$

2.5 Error reduction

Next step in the algorithm is to push the error level back down. To do this we need to construct operators called approximate ground state projectors (AGSPs). One convenient choice of AGSP would be of the form

$$A := \exp\left[-\frac{x(H - \epsilon_0)^2}{2\epsilon^2}\right],$$

where x is a parameter defined later. We cannot however construct such an operator. Firstly it would require us to exactly know the ground state energy. Secondly it would require exponential time to construct and specify.

We already have a ground state approximation $|\gamma_1\rangle$ which we got from the non-degenerate algorithm with energy at most $\epsilon_0 + \eta^2 \epsilon/4f$. As such, if we define $\epsilon_0' := \langle \gamma_1|H|\gamma_1\rangle$ then for sufficiently small error η we have

$$\left|\epsilon_0 - \epsilon_0'\right| \leq \eta^2 \epsilon/4f \leq \epsilon/2 \,.$$

In order to construct an approximation of A by a more efficient operator K, we define

$$K := \frac{2\epsilon\tau}{\sqrt{2\pi x}} \sum_{j=0}^{\lceil T/\tau \rceil} \exp\left[i\epsilon_0'\tau j - \frac{\epsilon^2\tau^2 j^2}{2x} \right] U_B(\tau j) \,, \tag{4}$$

where $U_B(t)$ is an approximation of $\exp(-iHt)$ with Schmidt rank B across every cut, the construction and analysis of which is given in Ref. [13].

The approximation between A and K is broken into three distinct approximations. The first is that the integral on $(-\infty, \infty)$ is truncated to an integral on $[-T, T]$, with an exponentially small error in T. The second is that this integral is approximated by a Riemann sum, specifically the rectangle rule[1] with a discrete step-size of τ. The third is that $\exp(-iHt)$ is approximated by a unitary $U_B(t)$ that has low Schmidt rank. All three errors are considered in Section 3.4. If we take parameters in Eq. 4 to have the scaling

$$\begin{aligned} x :&= \mathcal{O}(\log \zeta^{-1}) \\ T :&= \mathcal{O}(\epsilon^{-1} \log(1/\zeta)\sqrt{\log(n/\zeta)}) \\ \tau^{-1} :&= \mathcal{O}(n^2 \zeta^{-1}\sqrt{\log n/\zeta}) \\ B :&= (\zeta^{-1})^{\mathcal{O}(1/\epsilon)} \cdot \mathsf{poly}(n/\zeta) \,, \end{aligned}$$

then this approximate AGSP is capable of lowering the energy error down to $\Delta = \zeta$. As such we will henceforth refer to this as a ζ-approximate AGSP. With the definition of K in hand, we can give a precise specification of REDUCE:

REDUCE: Decompose a ζ-approximate AGSP with $\zeta = c\epsilon^6/f^4$ as $K = \sum_j A_j \otimes B_j$. Then return the set given by applying this approximate AGSP, combined with the recycled Schmidt vectors of the previous approximate ground states.

$$S_{h,i}^{(4)} := \left\{ A_j |s\rangle \Big| \forall j, \, |s\rangle \in S_{h,i}^{(3)} \right\} \cup L_{h,i} \,.$$

As we will prove in Section 3.4, the REDUCE step can lower the error of a viable set from $\delta = 1/5$ down to $\Delta = \zeta$, increasing both the bond dimension and cardinality by a factor of

$$(1/\zeta)^{\mathcal{O}(1/\epsilon)} \cdot (n/\zeta)^{\mathcal{O}(1)} \,.$$

Therefore the above construction with $\zeta = c\epsilon^6/f^4$ can send a $(i, p_1 + q, p_2, \delta = 1/5)$-viable set to a $(i, pp_1 + pq + q, pp_2, c\epsilon^6/f^4)$-viable set where

$$p(n) = n^{\mathcal{O}(1)} \,.$$

This forms our next Claim.

Claim 11 (Proof in Section 3.4)**.** *The set resulting from* REDUCE, $S_{h,i}^{(4)}$, *is* $(i, pp_1 + pq + q, pp_2, \Delta = c\epsilon^6/f^4)$-viable.

---

[1]Higher-order approximations such as other Newton-Cotes formulae could be used, though this won't change the overall scaling of the various parameters.

2.5.1 Final error reduction

As for the final step, we now want to reduce the error further than was previously necessary. We do this by simply constructing a stronger version of the previously used approximate AGSP.

FINALREDUCE: Construct a $\eta^2/4f$-approximate AGSP K. Return the set constructed by applying this AGSP, combined with the previous approximate ground states,

$$S_{h,n}^{(4)} := \left\{ K\,|s\rangle \,\middle|\, |s\rangle \in S_{h,n}^{(3)} \right\} \cup \{|\gamma_1\rangle, \cdots, |\gamma_{h-1}\rangle\}\,.$$

Claim 12 (Proof in Section 3.4)**.** *The set resulting from* FINALREDUCE, $S_{h,n}^{(4)}$, *is* $(n, p_1 + q + g, p_0 p_2, \Delta = \eta^2/4f)$-*viable, where*

$$p_0(n) = n^a, \quad \text{where } a = \mathcal{O}\!\left(1 + \epsilon^{-1}\sqrt{\log \eta^{-1}/\log n}\right).$$

2.6 Ground state approximation

Once we have $S_{h,n}$, a viable set on the entire system, we then need to extract the witness state. By construction, a viable set provides a domain on which known optimisation techniques can be applied to efficiently find approximate ground states. We summarise this with the last procedure.

APPROXGROUNDSTATE: Take σ to be a density matrix supported on $\mathrm{Span}(S_{h,n})$, found as the solution to the convex program

$$
\begin{aligned}
\min \quad & \mathrm{Tr}(H\sigma_h) && (5)\\
\text{where} \quad & \langle\gamma_j|\sigma_h|\gamma_j\rangle = 0 \text{ for } 1 \le j < h\,,\\
& \mathrm{Tr}(\sigma_h) = 1\,,\ \sigma_h \ge 0\,.
\end{aligned}
$$

Return $|\gamma_h\rangle$, a leading eigenvector of σ_h.

The objective functions and constraints can both be evaluated efficiently, firstly because our witness states are MPS with small bond dimension, and the Hamiltonian is a matrix product operator (MPO) with small bond dimension, and also because the domain of the optimisation is polynomially large. Since the program takes the form of a convex optimisation, these two properties are sufficient to allow the program to be solved efficiently. By optimality of the above program, the energy of σ_h must be at most that of the witness, $\epsilon_0 + \eta^2\epsilon/4f$. If we take $\eta \le 1/3$ and using Lemma 15 we get that the leading eigenvector, $|\gamma_h\rangle$, has an energy at most $\epsilon_0 + \eta^2\epsilon/4g$, and ground space overlap $\|G\,|\gamma_h\rangle\| \ge 1 - \eta^2/4g$.

2.7 Proof of Theorem 1

We can now turn to our proof of the main result Theorem 1. The proof is conditional on Claims 6, 10, 11, and 12, whose proof is given below in Section 3.

Proof. The Claims 6, 10, 11, and 12 establish that in each iteration of Algorithm 1 a viable set with energy error at most $\eta^2/4f$ is produced, and Lemma 15 establishes that this yields an orthonormal set of vectors such that $\langle\gamma_j|\,G\,|\gamma_j\rangle \ge 1 - \eta^2/2g$ for all $j = 1,\ldots,g$. Define $\Upsilon := \sum_{j=1}^{g} |\gamma_j\rangle\langle\gamma_j|$, and compute

$$\|G - \Upsilon\|_F^2 = \mathrm{Tr}(G - \Upsilon)^2 = 2g - 2\,\mathrm{Tr}(G\Upsilon) \le 2g - 2g(1 - \eta^2/2g) = \eta^2\,.$$

This completes the proof of the main result conditional on the Claims. $\qquad\square$

2.8 Runtime

As all the parameters given in Table 1 are polynomials in n, and all operations required to execute Algorithm 1 (convex optimisation, inner product of MPS etc.) can be performed at polynomial overhead, the overall runtime is also polynomial in n. The leading order term in n gives a runtime of $T = n^a$ where $a = \mathcal{O}\big(1 + \sqrt{\log \eta^{-1}/\log n}\big)$, which reduces to

$$
T = \begin{cases} n^{\mathcal{O}(1/\epsilon)} & \text{where } \eta = n^{-\mathcal{O}(1)}, \\ n^{\mathcal{O}(1)} & \text{where } \eta = n^{-o(1)}. \end{cases}
$$

In Ref. [22], an improved analysis of the FINALREDUCE step (which we will discuss in Section 3.4.2) is presented. The scaling the polynomials p_0, p_1, and p_2 are all improved to $q = n^{\mathcal{O}(1)}$ in the $\eta^{-1} = n^{\mathcal{O}(1)}$ case. This means that the corresponding run-time reduces to $T = n^{\mathcal{O}(1)}$.

One thing to note is that this leading order scaling is independent of the degeneracy. If we now focus on the degeneracy scaling, fixing n and allowing g to vary, we find that the dominant term now comes from the cardinality of the boundary contraction net. As such the leading-order scaling becomes 2^b where $b = g^{\tilde{\mathcal{O}}(1/\epsilon)}$; this is also the leading order scaling in the inverse-gap.

3 Analysis

3.1 Preliminaries

To prove the claims in the previous section, we first state some simple results that are easy to prove. These results were also used in Refs. [1,2], and we state them here without proof.

Lemma 13 (Overlap results [1,2]).

- *For any states for which $|\langle v|w\rangle| \geq 1 - \delta$ and $|\langle v'|w\rangle| \geq 1 - \delta'$, then $|\langle v|v'\rangle| \geq 1 - 2(\delta + \delta')$.*

- *For any two states such that $|\langle v|w\rangle| \geq 1 - \delta$, $|\langle v|O|v\rangle - \langle w|O|w\rangle| \leq 2\sqrt{2\delta}$ for any operator O with $\|O\| \leq 1$.*

- *If $|\langle u|\Gamma\rangle| \leq \omega$ and $|\langle v|u\rangle| \geq 1 - \delta$ then $|\langle v|\Gamma\rangle| \leq \omega + \sqrt{2\delta}$.*

- *For a pair of vectors with bounded overlap $|\langle u_1|u_2\rangle| \leq \omega$, and a second pair $|v_1\rangle$, $|v_2\rangle$ such that $|\langle u_j|v_j\rangle| \geq 1 - \delta$ for $j = 1, 2$, then we have $|\langle v_1|v_2\rangle| \leq \omega + \sqrt{10\delta}$.*

Lemma 14 (Orthogonalising lemma). *Suppose you have two vectors $|u\rangle$ and $|v\rangle$, such that $\langle u|H|u\rangle, \langle v|H|v\rangle \leq \epsilon_0 + \Delta\epsilon$, with bounded overlap $|\langle u|v\rangle| \leq \beta$. Then there exist orthogonal vectors $|u'\rangle, |v'\rangle \in \text{Span}\{|u\rangle, |v\rangle\}$, such that $\langle u'|H|u'\rangle, \langle v'|H|v'\rangle \leq \epsilon_0 + \Delta\epsilon(1+\beta)/(1-\beta)$.*

Proof. The proof is constructive, take the two vectors to be

$$
\big|u'\big\rangle = |u\rangle, \qquad \big|v'\big\rangle = \frac{|v\rangle - |u\rangle\langle u|v\rangle}{\||v\rangle - |u\rangle\langle u|v\rangle\|}.
$$

As $|v'\rangle \propto (I - |u\rangle\langle u|)|v\rangle$ this vector is clearly orthogonal to $|u'\rangle$. As for the energy bound, define $H' := H - \epsilon_0$, then

$$
\langle v'|H'|v'\rangle = \frac{\langle v|H'|v\rangle - 2\,\text{Re}\,\langle v|H'|u\rangle\langle u|v\rangle + \langle v|u\rangle\langle u|H'|u\rangle\langle u|v\rangle}{\||v\rangle - |u\rangle\langle u|v\rangle\|^2}
$$

$$
\leq \frac{1 + |\langle u|v\rangle|}{1 - |\langle u|v\rangle|}\Delta\epsilon \leq \frac{1 + \beta}{1 - \beta}\Delta\epsilon
$$

After reintroducing the ϵ_0 term in H we get the final energy bound. $\qquad\square$

If we apply the same argument to the projector orthogonal to the ground space, the same $(1+\beta)/(1-\beta)$ growth can also be shown to occur in the overlap error. The Schmidt rank of the resulting orthogonalised vectors will in general be the sum of the two original vectors' Schmidt ranks. If we can bound the overlap between a vector and *any* member of a vector space, then that vector can be similarly orthogonalised with respect to that whole space at the same cost.

As expectation values are bilinear in pure states, the optimisations in the TRIM and APPROXGROUNDSTATE are both performed over mixed states such that their objective functions are linear. To aid in the analysis of these optimisations, we will next show that the leading eigenvector of a low energy mixed state must be a low energy pure state.

Lemma 15 (Demixing). *If a mixed state σ has an energy at most $\epsilon_0 + \Delta\epsilon/(2g+1)$ where $\Delta \le 1/3g$, then the leading eigenvector has energy at most $\epsilon_0 + \Delta\epsilon$.*

Proof. First decompose the mixed state as $\sigma = \sum_{k>0} \lambda_k \, |v_k\rangle\langle v_k|$, where $\lambda_k \ge 0$. By the normalisation of σ_k we have $\sum \lambda_k = 1$. Next define the set of indices K as

$$K = \{k \,|\, \langle v_k|H|v_k\rangle \le \epsilon_0 + \Delta\epsilon\} \,.$$

By comparing the energy bound of the mixed state as a whole as well as those on indices in K, we get

$$\begin{aligned}
\epsilon_0 + \Delta\epsilon/(2g+1) &\ge \mathrm{Tr}(\sigma H) \\
&= \sum_{k \in K} \lambda_k \, \langle v_k|H|v_k\rangle + \sum_{k \notin K} \lambda_k \, \langle v_k|H|v_k\rangle \\
&\ge \epsilon_0 + \Delta\epsilon \sum_{k \notin K} \lambda_k
\end{aligned}$$

which in turn gives $\sum_{k \in K} \lambda_k \ge 2g/(2g+1)$, analogous to a Markov inequality. As we required $\Delta \le 1/3g$, each vector corresponding to a index in K has an energy at most $\epsilon_0 + \epsilon/3g$, and thus a ground space overlap at least $1 - 1/3g$. By Lemma A.3 there can only be g such orthonormal vectors, leading us to conclude that $|K| \le g$. Using this we can bound the largest eigenvalue corresponding to an index in K

$$\max_{k \in K} \lambda_k \ge \mathrm{mean}_{k \in K} \lambda_k = \frac{\sum_{k \in K} \lambda_k}{|K|} \ge \frac{2}{2g+1}$$

as well as that *not* in K

$$\max_{k \notin K} \lambda_k \le \sum_{k \notin K} \lambda_k \le \frac{1}{2g+1} \,.$$

As such we can conclude that the largest eigenvalue must belong within K, meaning that the leading eigenvector has energy at most $\epsilon_0 + \Delta\epsilon$. $\qquad\square$

3.1.1 Non-degenerate ground state approximation

Here we will briefly address NONDEGENERATEGSA, which we have claimed is a small modification on the nondegenerate ground state approximation algorithms of Ref. [1, 2]. The analysis of these algorithms largely holds in the degenerate case, appropriately replacing the notions of ground *state* overlap with ground *space* overlap. The only key

difference is the use of a special case of Lemma 15. As Lemma 15 will be invoked during the size-trimming and ground state approximation steps (Sections 2.6, 3.2), the energy levels in these steps grow by a factor of g overlooked in the non-degenerate analysis. By applying a slightly more powerful error reduction, dropping error levels beforehand by g to compensate, the same overall scaling can be proved, at the mere introduction of an extra constant factor. Due to this rather superficial change we will use these algorithms without proof of their validity. Furthermore we will take

$$q(n) := n^a \quad \text{where} \quad a = \mathcal{O}(1 + \epsilon^{-1}\sqrt{\log \eta^{-1}/\log n}) \tag{6}$$

to denote the bond dimension of this first approximate ground state $|\gamma_1\rangle$ given by this algorithm. We will show that the bond dimension of all subsequent ground state approximations follows the same scaling, and thus use $q(n)$ to generally denote the bond dimension of any combination thereof.

3.2 Size trimming

Next we will consider the size trimming sub-step, and prove the associated Claim 6. In Lemma 19 we will show that there exists a low-energy state orthogonal to the existing ground states, which we hope to find. After this we will introduce and define local-distinguishability, and give some intuition about how our size-trimming works in the low- and high-distinguishability regimes. We will then consider how the error levels behave in these two regimes, showing that the size-trimming procedure yields a good viable set for any level of distinguishability. We will conclude this section in with a full proof of Claim 6.

3.2.1 Low-energy state

The optimisation Eq. (1) in TRIM is performed with respect to H_L and not the whole Hamiltonian. There is no reason to assume *a priori* that states which have low energy with respect to the global Hamiltonian H, have low energy with respect to half-chain Hamiltonians H_L or H_R. Using the locality of the Hamiltonian however we will see that a global ground state of H can be excited by an energy at most 1 with respect to $H_L + H_R$.

Lemma 16. *For any ground state $|\Gamma\rangle$ of H,*

$$\langle\Gamma| \left(H'_L + H'_R\right) |\Gamma\rangle \leq 1$$

Proof. Consider the state $|v\rangle := |l\rangle \otimes |r\rangle$, where $|l\rangle$ and $|r\rangle$ are in the 0-eigenspace of H'_L and H'_R respectively, which are non-empty by construction. The energy of $|v\rangle$ is

$$\langle v|H|v\rangle = \epsilon_L + \epsilon_R + \langle v|H_i|v\rangle \ ,$$

so for $|v\rangle$ to not have an energy lower than that of a ground state, we require $\epsilon_0 \leq \epsilon_L + \epsilon_R + \langle v|H_i|v\rangle$. By bounding this further by the normalisation $\|H_i\| \leq 1$, this means the ground expectation of $H'_L + H'_R$ is at most 1. $\qquad\square$

As global ground states have low energy with respect to $H_L + H_R$, there must exist a good approximant of the ground state in the low-energy subspace of $H_L + H_R$. As such we will restrict our analysis to this subspace, which we will refer to as the truncated Hilbert space.

Definition 17 (Truncated Hilbert Spaces)**.** Let P_t be the projection onto $\mathcal{H}_{[1,i]} \otimes \mathcal{H}_{[i+1,n]}^{\leq t}$, where $\mathcal{H}_{[i+1,n]}^{\leq t}$ is the subspace of $\mathcal{H}_{[i+1,n]}$ spanned by eigenstates of H'_R with eigenvalues at most t, that is,

$$\operatorname{im}(P_t) = \operatorname{span}\left\{|v\rangle \big| H'_R |v\rangle = \lambda |v\rangle \; ; \; \lambda \leq t \right\} .$$

Similarly let Q_t be the projection onto $\left(\mathcal{H}_{[1,i]} \otimes \mathcal{H}_{[i+1,n]}\right)^{\leq t}$ which corresponds to the span of the eigenvectors of $H'_L + H'_R$ with eigenvalues at most t, that is,

$$\operatorname{im}(Q_t) = \operatorname{span}\left\{|v\rangle \big| (H'_R + H'_L) |v\rangle = \lambda |v\rangle \; ; \; \lambda \leq t \right\} .$$

As the left and right Hamiltonians commute and are positive, we have $P_t \geq Q_t$.

The following choices for t and ξ ensure that the error of the boundary contraction step remains small. Define $t := 99 \left(\log 99 + 4 \log f - \log c - 6 \log \epsilon\right)$ such that $99 \cdot 2^{-t/99} = c\epsilon^6/f^4$. Furthermore without loss of generality we can take $\epsilon \leq 10^9 \leq \sqrt[6]{70 f^4/c}$ such that

$$t \geq 1 + 4\sqrt{c}\epsilon^3/f^2 \geq \langle \Gamma | (H'_L + H'_R) | \Gamma \rangle + 4\sqrt{c}\epsilon^3/f^2$$

for any ground state $|\Gamma\rangle$. Finally we define the spacing of the boundary contraction net ξ

$$\xi := c'\epsilon/f(1+t) = \Omega((\epsilon/g)/\log(g/\epsilon)) . \tag{7}$$

In general whilst a ground state can be heavily frustrated, containing contributions from states with arbitrarily high local energy, these contributions will be small. Specifically we expect that a given ground state will have a large overlap with P_t and Q_t — this result is known as the *truncation lemma*, and was proven in Refs. [2, 12].

Lemma 18 (Truncation lemma [2, 12])**.** *For any ground state $|\Gamma\rangle$ of H,*

$$\|(1 - P_t) |\Gamma\rangle\| \leq \|(1 - Q_t) |\Gamma\rangle\| \leq 99 \cdot 2^{-t/99} .$$

Given these results, we can now move on to proving the existence of a low-energy witness state which Eq. (1) will approximate.

Lemma 19. *There exists a state $|w\rangle$ in $\operatorname{Span}\{S_i^{(1)}\} \otimes \mathcal{H}_{[i+1,n]}^{\leq t}$ orthogonal to the previously constructed approximate ground states $|\gamma_1\rangle, \ldots, |\gamma_{h-1}\rangle$, of Schmidt rank $B := B_\delta$ where $\delta = 8\sqrt{c}\epsilon^3/f^2$, with an energy at most $\epsilon_0 + c'\epsilon/f$.*

Proof. Let $|v'\rangle$ be a witness of $S_i^{(1)}$, such that it is orthogonal to the previous ground state approximations and has energy at most $\epsilon_0 + c\epsilon^6/f^4$. By Lemma 3 there exists a ground state $|\Gamma\rangle$ such that $|\langle v'|\Gamma\rangle| \geq 1 - c\epsilon^6/f^4$. Using Lemma 13 we get that

$$\langle v'|H_i|v'\rangle \geq \langle \Gamma|H_i|\Gamma\rangle - 2\sqrt{2c}\epsilon^3/f^2 .$$

Using this together with the energy of $|v'\rangle$, and again assuming that $c\epsilon^6/f^4$ is sufficiently small gives

$$\begin{aligned}
\langle v'| \left(H'_L + H'_R\right) |v'\rangle &\leq \langle \Gamma| \left(H'_L + H'_R\right) |\Gamma\rangle + 2\sqrt{2c}\epsilon^3/f^2 + c\epsilon^6/f^4 \\
&\leq \langle \Gamma|(H'_L + H'_R)|\Gamma\rangle + 4\sqrt{c}\epsilon^3/f^2 .
\end{aligned}$$

Decomposing the above using the projector P_t and its orthogonal complement – both of which commute with H'_L and H'_R – then this gives

$$\begin{aligned}
\langle \Gamma|(H'_L + H'_R)|\Gamma\rangle + 4\sqrt{c}\epsilon^3/f^2 &\geq \langle v'|(H'_L + H'_R)|v'\rangle \\
&\geq \langle v'|P_t(H'_L + H'_R)P_t|v'\rangle \\
&\quad + \langle v'|(1 - P_t)(H'_L + H'_R)(1 - P_t)|v'\rangle \\
&\geq \langle v|(H'_L + H'_R)|v\rangle \|P_t|v'\rangle\|^2 + t \|(1 - P_t)|v'\rangle\|^2
\end{aligned}$$

15

where we have defined $|v\rangle := P_t |v'\rangle / \| P_t |v'\rangle \|$, such that $|v\rangle$ now lies in the desired domain $\mathrm{Span}\{S_i^{(1)}\} \otimes \mathcal{H}_{[i+1,n]}^{\leq t}$ by construction. Using $\|P_t|v'\rangle\|^2 + \|(1-P_t)|v'\rangle\|^2 = 1$ as well as the inequality $t \geq \langle\Gamma|(H_L' + H_R')|\Gamma\rangle + 4\sqrt{c}\epsilon^3/f^2$, we can reduce the above simply to

$$\langle v|(H_L' + H_R')|v\rangle \leq \langle\Gamma|(H_L' + H_R')|\Gamma\rangle + 4\sqrt{c}\epsilon^3/f^2 \,. \tag{8}$$

The overlap between this truncated witness and the ground state is bounded

$$\begin{aligned}
|\langle v|\Gamma\rangle| &\geq |\langle v'|P_t|\Gamma\rangle| \\
&\geq |\langle v'|\Gamma\rangle| - |\langle v'|(1-P_t)|\Gamma\rangle| \\
&\geq 1 - c\epsilon^6/f^4 - 99 \cdot 2^{-t/99} \\
&= 1 - 2c\epsilon^6/f^4 \,,
\end{aligned}$$

and so once again applying Lemma 13 gives

$$\langle v|H_i|v\rangle \leq \langle\Gamma|H_i|\Gamma\rangle + 4\sqrt{c}\epsilon^3/f^2 \,. \tag{9}$$

Combining Equations (8) and (9) we can bound the total energy of the witness projected into the truncated Hilbert Space

$$\langle v|H|v\rangle \leq \epsilon_0 + 8\sqrt{c}\epsilon^3/f^2 \,.$$

Next we need to considering trimming the bond dimension. By applying Lemma 9 and recalling the definition of B_δ from Eq. (3), we get that the truncated state

$$|w\rangle := \mathrm{Trunc}_B^i |v\rangle / \left\| \mathrm{Trunc}_B^i |v\rangle \right\|$$

has $\langle w|H|w\rangle \leq \epsilon_0 + 96\sqrt[4]{4c}\epsilon/f \leq \epsilon_0 + 0.75c'\epsilon/f$.

Next we consider the overlap induced with the existing ground states. To bound this we consider the overlap between $|w\rangle$ and $|v'\rangle$. We had that $|\langle v'|\Gamma\rangle| \geq 1 - c\epsilon^6/f^4$ and $|\langle v|\Gamma\rangle| \geq 1 - 2c\epsilon^6/f^4$, Lemma 13 therefore gives

$$|\langle v'|v\rangle| \geq 1 - 6c\epsilon^6/f^4 \,.$$

Next take $|\gamma\rangle := \mathrm{Trunc}_B^i |\Gamma\rangle / \left\| \mathrm{Trunc}_B^i |\Gamma\rangle \right\|$, which has $|\langle\gamma|\Gamma\rangle| \geq 1 - 8\sqrt{c}\epsilon^3/f^2$ by Lemma 7. Lemma 13 once more gives that $|\langle v|\gamma\rangle| \geq 1 - 20\sqrt{c}\epsilon^3/f^2$. As $|w\rangle$ is a truncation of $|v\rangle$, by the Eckart-Young theorem it must have a larger overlap than $|\gamma\rangle$

$$|\langle v|w\rangle| \geq |\langle v|\gamma\rangle| - 20\sqrt{c}\epsilon^3/f^2 \,.$$

Applying Lemma 13 a third time we get finally that

$$|\langle v'|w\rangle| \geq 1 - 52\sqrt{c}\epsilon^3/f^2 \,.$$

As $|\gamma_1\rangle, \ldots, |\gamma_{h-1}\rangle$ are all perpendicular to $|v'\rangle$, the overlap between $|w\rangle$ and any element of $\mathrm{Span}\{|\gamma_1\rangle, \ldots, |\gamma_{h-1}\rangle\}$ is upper bounded by the component of $|w\rangle$ orthogonal to $|v'\rangle$

$$\sqrt{1 - |\langle v'|w\rangle|^2} \leq \sqrt{1 - \left(1 - 52\sqrt{c}\epsilon^3/f^2\right)^2} \leq \sqrt{104}\sqrt[4]{c\epsilon^6}/f \,.$$

Taking $\epsilon \leq 10^7$ without loss of generality this overlap is less than $1/7$, so we can orthogonalise $|w\rangle$ to the existing ground states at a multiplicative energy cost of $4/3$, resulting in a final state with $\langle w|H|w\rangle \leq \epsilon_0 + c'\epsilon/f$, and all other desired properties by construction.

$\square$

3.2.2 Distinguishability

Viable sets capture local information about a witness state, specifically the Schmidt vectors. If two states are globally orthogonal, this can mean that their Schmidt vectors span orthogonal spaces, but they can also span identical regions. Indeed global orthogonality of witness states cannot be locally imposed for this reason.

When considering viable sets, a natural notion of distinguishability is the overlap of the desired witness state, and the Schmidt vectors of the already constructed states. If we take $|w\rangle$ to be a witness state which we wish to capture, $L_{h,i}$ to be the set of Schmidt vectors on $[1, i]$ of the h already constructed ground states, and $\Pi_{h,i}$ the projector onto $\mathrm{Span}(L_{h,i})$, then the distinguishability is defined as

$$D_{h,i} := 1 - \langle w|\Pi_{h,i}|w\rangle \,.$$

From here we will omit the subscripts on L, Π, and D, since they will be clear from context. If we think of Π as being the projector onto the vectors on $[1, i]$ which we already have, then D can be considered a measure of how much of the state is unaccounted for.

Clearly if D is low, then a viable set can be constructed entirely by recycling existing Schmidt vectors. The precise bounds for the quality of viable set constructed in this manner, as a function of D, are given in Lemma 20.

In the case of a high D new vectors need to be found. In this case we know there is little overlap between the Schmidt vectors of our new witness and $\mathrm{Span}(L)$. Our procedure works by restricting to a low-energy subspace, and projecting away from the existing vectors, by minimising the expectation value of Π. In Lemma 21 we will show that this procedure must produce a low energy state. Given that our witness is highly distinguishable, Lemma 22 will show that this new witness after size trimming must also be highly distinguishable. This will culminate in Lemma 23 where we show that this therefore implies that the new witness has small overlap with the previous ground states, producing a good viable set.

We will conclude this subsection by showing in Lemma 24 that by performing both the low- and high-distinguishability procedures and combining the results, we can upper bound the error of our viable set in a distinguishability-independent manner.

3.2.3 Low D

By definition, in the low D case the left Schmidt vectors of $|w\rangle$ have large overlap with L. As a result we can see that L will form a viable set, whose witness is the projection of $|w\rangle$ onto L, and that the error of this viable set is low for low D.

Lemma 20 (Low D Solution). *There exists a state $|v\rangle$ orthogonal to $|\gamma_1\rangle, \ldots, |\gamma_{h-1}\rangle$ with $\mathrm{Tr}_{[i+1,n]} |v\rangle\langle v|$ supported on the span of $L_{h,i}$, such that there exists a ground state $|\Gamma\rangle$ with $|\langle\Gamma|v\rangle| \geq 1 - \delta_{\mathrm{Low}}$ where*

$$\delta_{\mathrm{Low}} = 1 - \frac{1 - c'/f - \sqrt{D}}{\sqrt{1 - D}} \,. \tag{10}$$

Proof. The idea here is to project $|w\rangle$ into $\mathrm{Span}(L_{h,i}) \otimes \mathcal{H}_{[i+1,n]}$ and use the distinguishability to bound the error associated with renormalisation. Specifically take

$$|v\rangle := \frac{\Pi |w\rangle}{\|\Pi |w\rangle\|}$$

17

where we recall that $\langle w|\Pi|w\rangle = 1 - D$. As the existing ground state approximations are all stable under Π by construction, the orthogonality of $|w\rangle$ to these vectors carries over to $|v\rangle$. The energy of $|w\rangle$ is at most $\epsilon_0 + c'\epsilon/f$, by Lemma 3 this means there is a ground state $|\Gamma\rangle$ with

$$|\langle w|\Gamma\rangle| \geq 1 - c'/f \,.$$

Decomposing this with the projector Π, and using both the triangle inequality and Cauchy's ineqality, we get

$$
\begin{aligned}
1 - c'/f &\leq |\langle\Gamma|w\rangle| \\
&\leq |\langle\Gamma|\Pi|w\rangle| + |\langle\Gamma|(I - \Pi)|w\rangle| \\
&\leq \sqrt{1 - D} \cdot |\langle\Gamma|v\rangle| + \sqrt{D} \,.
\end{aligned}
$$

Rearranging this gives the desired ground state overlap stated above. □

3.2.4 High D

Take σ to both be the solution to Program 1 for the X closest to $\mathrm{cont}(w)$ and the smallest Y which is greater than or equal to $\langle w|H_L|w\rangle - \langle\gamma_1|H_L|\gamma_1\rangle$, i.e. the tightest set of constraints satisfied by $|w\rangle\langle w|$. By the nature of the two nets, described in Section 2.3.2, we have that

$$\|X - \mathrm{cont}(w)\|_1 \leq \xi \qquad \text{and} \qquad 0 \leq Y - \langle w|H_L|w\rangle + \langle\gamma_1|H_L|\gamma_1\rangle \leq 4c'\epsilon/f \,.$$

Here we want to argue that if D is not too low, then one of the σ outputs of Eq. (1) has an eigenvector which can be extended to a good witness. To extend the eigenvectors of σ we take the 'boundary-uncontracted' version $\sigma' = U_w\sigma U_w^\dagger$, eigendecomposing these matrices as:

$$\sigma = \sum_k \lambda_k \, |v_k\rangle\langle v_k| \qquad\qquad \sigma' = \sum_k \lambda_k \, |v_k'\rangle\langle v_k'|$$

where $|v_k'\rangle = U_w \, |v_k\rangle$.

In TRIM we kept all eigenvectors whose value exceeded $10^{-9}/g$. We next want to show that there can only exist a finite number of such eigenvectors, and that they represent the vast majority of the spectrum of σ.

Lemma 21. *There exist at most g eigenvectors of σ' with energy at most $\epsilon_0 + 30{,}000c'\epsilon/g$, and the sum of the corresponding eigenvalues is at least $\Lambda = 1 - 1/(10{,}000g + 5000)$.*

Proof. Defining $\sigma' = U_w\sigma U_w^\dagger$, the boundary contraction property Lemma 5 gives an energy bound on σ'.

$$
\begin{aligned}
\mathrm{Tr}(H\sigma') &\leq \mathrm{Tr}(H_L\sigma) + \langle w|(H_i + H_R)|w\rangle \\
&\quad + \left\|\mathrm{Tr}_{[1,i-1]}(\sigma) - \mathrm{cont}(w)\right\|_1 \left(1 + \left\|U_w^\dagger H_R U_w\right\| - \langle w|H_R|w\rangle\right) \\
&\leq \mathrm{Tr}(H_L\sigma) + \langle w|(H_i + H_R)|w\rangle + \left\|\mathrm{Tr}_{[1,i-1]}(\sigma) - \mathrm{cont}(w)\right\|_1 \left(1 + \left\|U_w^\dagger H_R' U_w\right\|\right) \\
&\leq \langle w|H|w\rangle + \left(Y - \langle w|H_L|w\rangle + \langle\gamma_1|H_L|\gamma_1\rangle\right) + \xi(1 + t) \\
&\leq \epsilon_0 + 6c'\epsilon/f \,.
\end{aligned}
$$

By applying an argument similar to Lemma 15, if we increase the excitation not only by a factor of $(2g + 1)$ but also by 5000, we get that there are between 1 and g eigenvectors which have energy at most $\epsilon_0 + 30{,}000c'\epsilon/g$. Moreover we also get that the sum of the corresponding eigenvalues is at least $1 - 1/(10{,}000g + 5000) = \Lambda$. □

Next we rearrange the order of the eigenvectors. Let $1, \ldots, g$ refer to these g vectors[2] described in Lemma 21. In other words:

$$\langle v'_k | H | v'_k \rangle \leq \epsilon_0 + 30{,}000 c' \epsilon / g \quad \text{for } k \leq g$$
$$\langle v'_k | H | v'_k \rangle \geq \epsilon_0 + 30{,}000 c' \epsilon / g \quad \text{for } k > g$$

Next we drop any such vectors with eigenvalues less than the $10^{-9}/g$ threshold[3]. It should first be noted that by applying Lemma 15 to σ we get that the dominant eigenvalue that is at least $2/(2g+1)$ has a corresponding eigenvector in this set, and so this set is non-empty. Next we note that as there are at most $g-1$ such vectors dropped, so decrease in the sum of eigenvalues is at most $10^{-9} \cdot \frac{g-1}{g} \leq 10^{-9}$. As such we can conclude that the sum of the corresponding eigenvalues is still bounded

$$\sum_{k=1}^{g} \lambda_k \geq \Lambda - 10^{-9} \, .$$

Without loss of generality, take $|v'_1\rangle$ to be the member of $\{|v'_1\rangle, \ldots, |v'_g\rangle\}$ for which the expectation value of Π is the lowest. We now want to show that the expectation of Π with respect to $|v'_1\rangle$ not significantly higher than it is for $|w\rangle$ (which is $1 - D$ by definition). As such this will imply that in the high D case $|v'_1\rangle$ has small overlap with L, meaning that we have succeeded in finding a state with low overlap with the existing ground states.

Lemma 22.

$$\langle v'_1 | \Pi | v'_1 \rangle \leq \left[\frac{1}{\Lambda} + 10^{-9} \right] (1 - D)$$

Proof. The specific iteration of Eq. (5) we are referring to is that for which $|w\rangle$ satisfies all of the constraints of the optimisation. As such, by the optimality of the solution σ, we know that $\langle w | \Pi | w \rangle \geq \text{Tr}(\sigma \Pi)$. Decomposing this and using the definition of D:

$$\langle w | \Pi | w \rangle \geq \text{Tr}(\sigma \Pi)$$
$$1 - D \geq \sum_k \lambda_k \langle v_k | \Pi | v_k \rangle \, .$$

As Π acts strictly to the left of the cut and U_w strictly to the right they commute, meaning the expectation value of Π for primed and unprimed eigenvectors is identical.

$$1 - D \geq \sum_k \lambda_k \langle v'_k | \Pi | v'_k \rangle$$
$$\geq \sum_{k=1}^{g} \lambda_k \langle v'_k | \Pi | v'_k \rangle \, .$$

As we defined $|v'_1\rangle$ to have minimal expectation this can be further loosened to

$$1 - D \geq \left(\sum_{k=1}^{g} \lambda_k \right) \langle v'_1 | \Pi | v'_1 \rangle \, .$$

Rearranging this together with the bound for $\sum_{k=1}^{g} \lambda_k$, we get

$$\langle v'_1 | \Pi | v'_1 \rangle \leq \left[\Lambda - 10^{-9} \right]^{-1} (1 - D) \, .$$

After a binomial expansion, followed by a loosening of the inequality, we arrive at our final bound. $\qquad \square$

---

[2]Padding out these indices with zeros as necessary if there are less than g such vectors.

[3]Once again padding out with zero vectors such that there are still g such vectors for convenience.

In general $|v_1'\rangle$ will have some non-trivial overlap with the existing ground state approximations. This low expectation value for Π can however be used to bound this overlap when D is not too small.

Lemma 23. *For any vector $|\gamma\rangle \in \text{Span}\{|\gamma_1\rangle, \dots, |\gamma_{h-1}\rangle\}$ the overlap between $|\gamma\rangle$ and $|v_1'\rangle$ is upper bounded*

$$|\langle\gamma|v_1'\rangle|^2 \leq \left[\frac{1}{\Lambda} + 10^{-9}\right](1 - D).$$

Proof. The main property to use here is that $\Pi|\gamma\rangle = |\gamma\rangle$ by definition. Using this and Cauchy-Schwartz gives

$$|\langle\gamma|v_1'\rangle|^2 = |\langle\gamma|\Pi|v_1'\rangle|^2 \leq \|\Pi|v_1'\rangle\|^2 = \langle v_1'|\Pi|v_1'\rangle \leq \left[\frac{1}{\Lambda} + 10^{-9}\right](1 - D).$$

$\square$

Take $|u\rangle$ to be $|v_1'\rangle$ projected orthogonally to $|\gamma_1\rangle, \dots, |\gamma_{h-1}\rangle$, as per Lemma 14. Then the ground space overlap of this vector can be bounded to be $|\langle\Gamma|u\rangle| \geq 1 - \delta_{\text{High}}$ for some ground state $|\Gamma\rangle$, where

$$\delta_{\text{High}} = \frac{30{,}000c'}{g}\frac{1+\beta}{1-\beta} \quad \text{and} \quad \beta^2 \leq \left[\frac{1}{\Lambda} + 10^{-9}\right](1 - D). \tag{11}$$

3.2.5 All D

Next we can combine the two above errors. We have shown that there exist bounds on the overlap error δ in the low- and high-D cases. By combining these, we can bound δ independent of the unknown D.

Lemma 24. *The combined error $\delta = \min(\delta_{\text{Low}}, \delta_{\text{High}})$ over the total range of distinguishabilities is bounded $\delta \leq 0.01$ for all $g \geq 2$ and $D \in [0, 1]$.*

Proof. Consider the case of $D = 10^{-4}$ and $g = 2$. The two error levels in this case can be explicitly calculated, and are

$$\delta_{\text{Low}}(g = 2, D = 10^{-4}) \approx 0.00995050 < 0.01$$
$$\delta_{\text{High}}(g = 2, D = 10^{-4}) \approx 0.00999947 < 0.01$$

It can also be shown that for a fixed g that δ_{Low} and δ_{High} are monotonically increasing and decreasing in D respectively, and that for a fixed D both decrease monotonically with g. As such we can conclude that $\delta \leq 0.01$ for all $D \in [0, 1]$ and $g \geq 2$. $\square$

Claim 6 The set resulting from TRIM, $S_i^{(2)}$, is $(i, p_1, dsb + q^2, \delta = 1/100)$-viable, where

$$p_1(n), q(n) = n^a \quad \text{where } a = \mathcal{O}(1 + \epsilon^{-1}\sqrt{\log\eta^{-1}/\log n}).$$

Proof. As shown above, marginalising over distinguishability for degeneracies at least two, the overlap error is upper bounded by $1/100$. The cardinality contribution comes in two parts: the results given by Program 1, and the previously obtained left-Schmidt vectors that are recycled. Let the cardinality of the Program 1 results be r where

$$r = \underbrace{g10^9}_{\text{E-values}} \times \underbrace{\mathcal{O}(1/\epsilon)}_{\text{Energy net}} \times \underbrace{2^{\mathcal{O}(g/\epsilon)^{\mathcal{O}(1/\epsilon)}}}_{\text{BC net}} \times \underbrace{2^{\mathcal{O}(\epsilon^{-1} + \epsilon^{-1/4}\log^{3/4}(g/\epsilon))}}_{\text{Schmidt rank}} = 2^{g\tilde{\mathcal{O}}(1/\epsilon)}.$$

The cardinality contribution from recycled Schmidt vectors is bounded by the bond dimension of each previous ground state approximation, as given in Section 3.1.1. This contribution is $q(n) = n^a$ where $a = \mathcal{O}(1 + \epsilon^{-1}\sqrt{\log \eta^{-1}/\log n})$, and dominates over r. The overall cardinality is the summation of these two contributions, and as such is given by $p_1(n) = n^a$ where a has the scaling given above.

The results of Program 1 are linear combinations of vectors in $\text{Span}(S_i^{(1)} \cup L_{h,i})$. The contributions from both spaces to the bond dimension of the worst-case result is the product of their cardinalities and bond dimensions respectively. As such these contributions are dsb for $S_i^{(1)}$, and q^2 for $L_{h,i}$ respectively. □

3.3 Bond truncation

The idea here is to use the existence of low-rank approximate ground states to argue that truncation induces little error due to its optimal nature as a low-rank approximation [16]. First we note an important result [1,2] about truncation, that large overlap with a low-rank vector is maintained under truncation.

Lemma 25 ([1,2]). *Given a vector $|v\rangle$ with Schmidt rank R across $i|i+1$, for any $|u\rangle$*

$$\left| \left\langle \text{Trunc}_{R/\delta}^i(u) \middle| v \right\rangle \right| \geq |\langle u|v\rangle| - \delta .$$

Let $|u\rangle$ denote a witness of $S_i^{(2)}$ such that there exists a ground state $|\Gamma\rangle$ such that

$$|\langle \Gamma|u\rangle| \geq 1 - \frac{1}{100} .$$

Fix $|\Gamma\rangle$ to denote the ground state with which $|u\rangle$ has maximal overlap. Next we invoke Lemma 7, taking $r(n) := B_{1/800n}$ such that there exists an MPS $|\gamma\rangle$ with bond dimension bounded by $r(n)$ such that

$$|\langle \Gamma|\gamma\rangle| \geq 1 - \frac{1}{800} .$$

Applying Lemma 13 we can combine these, showing $|\gamma\rangle$ and $|u\rangle$ have large overlap

$$|\langle \gamma|u\rangle| \geq 1 - 2\left(\frac{1}{800} + \frac{1}{100} \right) = 1 - \frac{18}{800} .$$

Next take $|w\rangle$ to be the renormalised truncation of $|u\rangle$ across every bond

$$|w\rangle := \frac{\text{Trunc}_P^1 \ldots \text{Trunc}_P^{n-1} |u\rangle}{\left\| \text{Trunc}_P^1 \ldots \text{Trunc}_P^{n-1} |u\rangle \right\|} ,$$

where we have taken $P(n) := 800n\, r(n)$.

Applying Lemma 25 across every bond we get that the overlap-error induced by truncation grows additively, giving a final overlap of

$$|\langle \gamma|w\rangle| \geq |\langle \gamma|u\rangle| - n/(800n) \geq 1 - \frac{19}{800} .$$

Once again applying Lemma 13 we get therefore that the ground space overlap of $|w\rangle$, which is lower bounded by the overlap with $|\Gamma\rangle$, can in turn be lower bounded

$$|\langle \Gamma|w\rangle| \geq 1 - 2\left(\frac{1}{800} + \frac{19}{800} \right) = 1 - \frac{1}{20} .$$

Once again we must turn our attention now to orthogonalising $|w\rangle$ with respect to the existing ground states to satisfy the relevant orthogonality condition. The procedure for bounding this once again is to note that because $|u\rangle$ is orthogonal to $\mathrm{Span}\{|\gamma_1\rangle, \ldots, |\gamma_{h-1}\rangle\}$, any overlap $|w\rangle$ has with this space is upper bounded by the component of $|w\rangle$ orthogonal to $|u\rangle$. Applying Lemma 13 once again to their mutual overlap with $|\Gamma\rangle$, we get that

$$|\langle w|u\rangle| \geq 1 - 2\left(\frac{1}{100} + \frac{1}{20}\right) = 1 - \frac{6}{50}$$

and thus that the orthogonal component is bounded

$$\sqrt{1 - |\langle u|w\rangle|^2} \leq \sqrt{1 - 44^2/50^2} \leq 1/2\,.$$

By Lemma 14 then we can orthogonalise $|w\rangle$ with respect to the existing ground state approximations, at most tripling the overlap-error. Thus, after orthogonalisation, there must exist a ground state $|\Gamma'\rangle$

$$\left|\langle \Gamma'|w\rangle\right| \geq 1 - \frac{3}{20} \geq 1 - \frac{1}{5}\,.$$

Claim 10 The set resulting from TRUNCATE, $S_i^{(3)}$, is $(i, p_1 + q, p_2, \delta = 1/5)$-viable, where

$$p_2(n) = n^a \quad \text{where} \quad a = \mathcal{O}(1 + \epsilon^{-1}\sqrt{\log \eta^{-1}/\log n})\,.$$

Proof. The error level $\delta = 1/5$ is proven above. As the truncation procedure is performed element-wise, the cardinality is changed only by the reintroduction of recycled Schmidt vectors, adding q.

After truncation, the bond dimension is bounded by the trimming level

$$P = 800n B_{1/800n} = n^{1+o(1)}\,.$$

The orthogonalisation procedure then will additively increase this bond dimension by q, the bond dimension of the previous approximate ground states. As such the final bond dimension can be bounded by a polynomial p_2 of the same scaling. $\qquad\square$

3.4 Error reduction

Next we need to address the REDUCE and FINALREDUCE procedures. We will do this by considering the construction of an AGSP A, and an approximation thereof K. Both operators are subject to a free parameter ζ, which controls the magnitude of the error reduction they can perform. The only difference between REDUCE and FINALREDUCE will be the final level of ζ utilised. As such we will analyse the procedure for an arbitrary ζ, stating the REDUCE and FINALREDUCE as special cases of a more general analysis.

3.4.1 Exact AGSP

First we consider the exact AGSP construction

$$A := \exp\left[-\frac{x(H - \epsilon_0')^2}{2\epsilon^2}\right]\,,$$

where $\epsilon_0' := \langle \gamma_1|H|\gamma_1\rangle$ is an inverse polynomial approximation to the ground state energy, and x is a ζ-dependent parameter given in Appendix B. As well as specifying the precise

parameters of the AGSP/AAGSP construction, in Appendix B we also show that for each ground state $|\Gamma\rangle$ and excited state $|\Gamma^\perp\rangle$:

$$\| A |\Gamma\rangle \| \geq \frac{19}{20} \qquad\qquad \left\| A |\Gamma^\perp\rangle \right\| \leq \frac{\zeta}{2}$$

It is in this sense that our AGSP approximates a ground state projector.

Take $|u\rangle$ to be a witness of $S_i^{(3)}$, such that there exists a ground state $|\Gamma\rangle$ with

$$|\langle \Gamma | u \rangle| \geq 1 - 1/5 \,.$$

We can decompose this witness in the energy eigenbasis

$$|u\rangle = \sum_{j=0}^{d^n - 1} \sqrt{p_j} |E_j\rangle \,,$$

where $|E_0\rangle$ is a ground state, $\{|E_j\rangle\}_{j>0}$ are all excited states[4] with energy $\epsilon_j \geq \epsilon_0 + \epsilon$. Due to the role of $|u\rangle$ as a witness of $S_i^{(3)}$ we have $p_0 \geq 16/25$, and due to normalisation $\sum_j p_j = 1$. Lastly take $|w\rangle$ and to be the result of applying the AGSP to the witness and renormalising

$$|w\rangle := \frac{A |u\rangle}{\| A |u\rangle \|} \,.$$

We can bound the renormalising denominator term by applying A to $|u\rangle$ in this decomposed form

$$
\begin{aligned}
\| A |u\rangle \| &= \left\| \sum_{j=0}^{d^n - 1} \sqrt{p_j} A |E_j\rangle \right\| \\
&\geq \sqrt{p_0} \| A |E_0\rangle \| - \left\| \sum_{j \geq 1} \sqrt{p_j} A |E_j\rangle \right\| \\
&\geq \frac{4}{5} \cdot \frac{19}{20} - \frac{3}{10} \zeta \,.
\end{aligned}
$$

If we assume $\zeta \leq 1$ then $\| A |u\rangle \| \geq 1/5$. Using this bound we can argue that because A shrinks the high energy terms, yet does not shrink the magnitude all that much, the energy of $|w\rangle$ should be a drastic improvement over that of $|u\rangle$.

Lemma 26.

$$\langle w | H | w \rangle \leq \epsilon_0 + \zeta \epsilon / 80{,}000$$

Proof. We can bound the energy by simply expanding the expectation of $H' := (H - \epsilon_0)$, and seeing how it changes under the application of A using the above decomposition in the energy basis.

$$
\begin{aligned}
\langle w | H' | w \rangle &= \frac{\langle u | A H' A | u \rangle}{\| A |u\rangle \|^2} \\
&\leq 25 \sum_{j,k \geq 0} p_j \langle E_j | A H' A | E_k \rangle \,.
\end{aligned}
$$

---

[4]As we only require that the vectors $|E_j\rangle$ be energy eigenstates up to rephasing, all vectors appearing in this decomposition from the same energy-eigenspace can be merged into a single vector, allowing for this decomposition to be considered non-degenerate without loss of generality.

Because A commutes with H, the double sum collapses and we have

$$\leq 25 \sum_{j \geq 0} p_j \left\langle E_j \middle| AH'A \middle| E_j \right\rangle$$

$$\leq 25 \sum_{j \geq 1} p_j (\epsilon_j - \epsilon_0) \exp(-x(\epsilon_j - \epsilon_0')^2/\epsilon).$$

Using $\epsilon_0' - \epsilon_0 \leq \epsilon/2$, and then using the Hölder inequality, we have

$$\leq 50 \sum_{j \geq 1} p_j (\epsilon_j - \epsilon_0') \exp(-x(\epsilon_j - \epsilon_0')^2/\epsilon)$$

$$\leq 50 \left[\max_{E \geq \epsilon/2} E \exp(-xE^2/\epsilon^2) \right] \left[\sum_{j \geq 1} p_j \right].$$

The maximum is on the boundary because x is sufficiently large, so we find

$$\leq 25\epsilon \, e^{-x/4} \leq \zeta\epsilon/3200$$

Where above we have used the chosen value of $x := 33 - 8\log(\zeta)$ from Appendix B. $\qquad\square$

3.4.2   Approximate AGSP

As well as the bounds on the power of A, Appendix B also gives that we can approximate A by a low bond dimension operator K with

$$\|A - K\| \leq \zeta' := \frac{\zeta\epsilon}{240{,}000n} . \tag{12}$$

For simplicity we can express $K = A + \zeta'O$, where $\|O\| \leq 1$. As with K we can bound the renormalising term for $|u\rangle$,

$$\|K|u\rangle\| = \|(A + \zeta'O)|u\rangle\|$$

$$\geq \|A|u\rangle\| - \zeta' \|O|u\rangle\|$$

$$\geq \frac{1}{5} - \zeta'.$$

Now if we take $\zeta' \leq 1/10$ then we get $\|K|u\rangle\| \geq 1/10$. As such we can replace the state $|w\rangle$ that had an exact AGSP with an approximate version using the AAGSP,

$$|w'\rangle := \frac{K|u\rangle}{\|K|u\rangle\|}$$

and similarly bound the energy.

Lemma 27.

$$\left\langle w' \middle| H \middle| w' \right\rangle \leq \epsilon_0 + \zeta\epsilon/400$$

Proof. Once again we consider the expectation value of H'

$$\begin{aligned}
\langle w'|H'|w'\rangle &= \langle w'|H'|w'\rangle \\
&= \frac{\langle u|KH'K|u\rangle}{\|K|u\rangle\|^2} \\
&\leq 100\langle u|(A+\zeta'O)H'(A+\zeta'O)|u\rangle \\
&\leq 100\langle u|AH'A|u\rangle + 100\zeta'^2\langle u|OH'O|u\rangle \\
&\quad + 100\zeta'\langle u|OH'A|u\rangle + 100\zeta'\langle u|AH'O|u\rangle \\
&\leq 4\langle w|H|w\rangle + 300n\zeta' \\
&\leq \zeta\epsilon/800 + \zeta\epsilon/800 = \zeta\epsilon/400
\end{aligned}$$

Reintroducing ϵ_0 gives the desired energy bound. $\qquad\square$

Finally we must orthogonalise to negate any overlap with the existing vectors this has caused.

Lemma 28. *For any* $|\gamma\rangle \in \mathrm{Span}\{|\gamma_1\rangle,\ldots,|\gamma_h-1\rangle\}$, $|\langle w'|\gamma\rangle| \leq \frac{400-1}{400+1}$.

Proof. The original witness $|u\rangle$ has a large ground state overlap, which in terms of the previous decomposition means

$$\sqrt{p_0} = \langle E_0|u\rangle \geq 4/5.$$

Using the fact that $\|K|u\rangle\| \leq 1$ and that A is diagonal in the energy basis, we can bound the overlap that $|w'\rangle$ in turn has with $|E_0\rangle$. Using $\zeta' \leq 1/100$,

$$\begin{aligned}
|\langle E_0|w'\rangle| &= \frac{|\langle E_0|K|u\rangle|}{\|K|u\rangle\|} \\
&\geq |\langle E_0|K|u\rangle| \\
&\geq \left|\sum_j \sqrt{p_j}\langle E_0|A|E_j\rangle\right| - \zeta' \\
&\geq \sqrt{p_0}\langle E_0|A|E_0\rangle - \zeta' \\
&\geq \frac{4}{5}\frac{19}{20} - \frac{1}{100} = \frac{3}{4}.
\end{aligned}$$

Applying Lemma 13 we therefore get that

$$|\langle u|w'\rangle| \geq 1 - 2\left(\frac{1}{5} + \frac{1}{4}\right) = \frac{1}{10}.$$

As has been argued before, the component of $|w'\rangle$ overlapping with $\mathrm{Span}\{|\gamma_1\rangle,\ldots,|\gamma_{h-1}\rangle\}$ can be bounded by the component perpendicular to $|u\rangle$. This component is bounded

$$\sqrt{1 - |\langle u|w'\rangle|^2} = \frac{3\sqrt{11}}{10} \leq \frac{400-1}{400+1}.$$

$\qquad\square$

By Lemma 14 and the previous lemma, we can thus orthogonalise $|w'\rangle$ to a vector $|w''\rangle$ with a 400-fold increase in the excitation, resulting in an energy of

$$\langle w''|H|w''\rangle \leq \epsilon_0 + \zeta\epsilon.$$

As such setting ζ to the desired final energy error level and applying the local Schmidt operators of K, the energy error can be reduced to ζ.

Claim 11,12 The set resulting from REDUCE, $S_i^{(4)}$, is $(i,\,pp_1 + pq + q,\,pp_2,\,\Delta = c\epsilon^6/f^4)$-viable, where

$$p(n) = n^{\mathcal{O}(1)}.$$

The set resulting from FINALREDUCE, $S_n^{(4)}$, is $(n,\,p_1 + q + g,\,p_0 p_2,\,\Delta = \eta^2/4f)$-viable, where

$$p_0(n) = n^a \quad \text{where} \quad a = \mathcal{O}(1 + \epsilon^{-1}\sqrt{\log \eta^{-1}/\log n}).$$

Proof. The construction described above gives a K with a bond dimension that scales (see Ref. [13]) as $n^{a'}$ where $a' = \mathcal{O}(1 + \epsilon^{-1}\sqrt{\log \zeta^{-1}/\log n})$.

For REDUCE the cardinality growth is caused by the fact that K must be applied locally, i.e. each Schmidt operator must be individually applied. The number of these operators is bounded by the bond dimension of K. As such REDUCE, for which $\zeta = c\epsilon^6/f^4 = \mathcal{O}(1)$, has a multiplicative growth in cardinality of $p(n) = n^{\mathcal{O}(1)}$. On top of this the recycled Schmidt vectors additively increase the cardinality once more by q.

For FINALREDUCE however the whole operator can be applied without Schmidt decomposition, thus leaving cardinality unchanged. The recycled previous ground states grow the cardinality additively by g.

The bond dimension increase is once again multiplicative, and once again bounded by the bond dimension of K as an MPO. For REDUCE this means a growth by a factor of $p(n)$. For FINALREDUCE however $\zeta = \eta^2/4f$ is not necessarily constant, as such the growth in bond dimension has the more general scaling of

$$p_0(n) = n^a \quad \text{where} \quad a = \mathcal{O}(1 + \epsilon^{-1}\sqrt{\log \eta^{-1}/\log n}).$$

$\square$

Here we briefly comment on the improved analysis of Ref. [22]. Suppose we perform the following for our FINALREDUCE: instead applying of a single powerful AGSP, we interleave a mild AGSP with bond trimming multiple times. With sufficient parameters (the details are given in Sec.5.5 of [22]), this can be done such that the bond dimension remains at most $n^{\mathcal{O}(1)}$, and the error level is halved. Performing this procedure the requisite number of times, the scaling of final bond dimension can be improved to $p_0 = n^{\mathcal{O}(1)}$ for $\eta^{-1} = n^{\mathcal{O}(1)}$. Carrying this analysis through, this also reduces the scaling of p_1 and q to both be $n^{\mathcal{O}(1)}$, and therefore a final run-time also of $T = n^{\mathcal{O}(1)}$.

4 Conclusion

We have shown the ground state approximation of Refs. [1, 2] can be extended to degenerate systems with an approximation of the ground space projector that is up to inverse polynomial error in any Schatten norm.

AGSPs have proven powerful tools for relating the structural properties of gapped systems to various computational complexity bounds regarding them. Most AGSP constructions are either functions of the Hamiltonians [2, 12, 14] or parts of the Hamiltonian terms [1, 17–20]. While we utilise such AGSPs in intermediate steps, our algorithm outputs a state-based AGSP, and more thoroughly exploits the underlying one-dimensional structure, allowing for a more powerful overall construction. While previous AGSP constructions have typically been full-rank, our AGSP takes the form of a minimal-rank projector, that moreover has a polynomially bounded bond dimension as a matrix product operator (MPO). It remains an open question for which systems there exist efficiently

constructible AGSPs that are also *low-rank* AGSPs. Note that this is important for applications because while AGSPs that are close to the ground space in operator norm are easy to construct even with small bond dimension in MPO form, they do not generally allow for low-error estimation of expectation values of other MPO observables.

While running in polynomial-time, the specific runtime scaling of the above algorithm is highly suboptimal, leaving wide room for further optimisation. One potential avenue for optimisation is the level of redundancy within the viable sets that are generated. While there are rigorous upper bounds on the cardinality of these sets, no care has been taken to explore to possibility of high levels of linear dependence, opening the door to far better practical runtimes than our analysis would suggest.

Note added in preparation: This work began as the undergraduate Honours thesis of the first author as a project to extend LVV to the degenerate case. During that time, the results of Huang in Ref. [2] (version 1) appeared, and were incorporated into our independent analysis. Near the completion of this manuscript, Huang also extended the results in Ref. [2] to the degenerate case and published this as Ref. [21] (version 3), an updated version of the original manuscript. While our techniques rely on the methods in version 1 (Ref. [2]), our treatment of the degenerate case was made completely independently from version 3 (Ref. [21]).

Our results and the new results in Ref. [21] are essentially the same, but with the following differences. The most notable is that the algorithm of Ref. [21] is single-pass, finding a full set of ground states with a single sweep of the system. In principle our multi-pass algorithm will generate viable sets with far higher levels of redundancy, causing worse scaling in the parameters of Table 1. In the final runtime however this slow-down only influences the various constants hidden under big-O terms, giving an overall scaling that is identical to that of Ref. [21]. The second major difference regards the size-trimming, the most complicated step to generalise. Our algorithm works by applying an energy constraint and minimising the overlap between witnesses, while Ref. [21] constrains the overlap and minimises the energy. This allows one to use boundary contraction to circumvent the issue of local distinguishability, simplifying the analysis of this step. The final difference is that our algorithm does not require knowing a specific bound on the ground state degeneracy g, only that a bound of the form $g = \mathcal{O}(1)$ exists, whereas it is not clear how to perform the algorithm in Ref. [21] without an explicit upper bound.

Subsequent to the journal submission of this manuscript, but prior to its publication, Arad et.al. released a yet further improved algorithm in Ref. [23]. While our algorithm works by iterating along the spin chain, constructing matrix product states, the main new technique in Ref. [23] is to instead apply divide-and-conquer approach, generating states in a tree-like fashion. This method allows for much tighter runtime analysis (e.g. runtime reduction from $\mathsf{poly}(n)$ to $\mathcal{O}(n^{3.38})$ for frustration-free systems), as well as extensions to much larger classes of systems (e.g. a quasi-polynomial-time algorithm for certain gapless models).

References

[1] Z. Landau, U. Vazirani, and T. Vidick, "A polynomial-time algorithm for the ground state of 1D gapped local Hamiltonians," *Proceedings of the 5th Conference on Innovations in Theoretical Computer Science* and *Nature Physics*, **11**, 566–569, arXiv:1307.5143, (2013).

[2] Y. Huang, "A polynomial-time algorithm for approximating the ground state of 1D gapped Hamiltonians," arXiv:1406.6355v1, (2014).

[3] R. R. dos Santos, "Introduction to quantum Monte Carlo simulations for fermionic systems," *Brazilian Journal of Physics*, **33**(1), 36–54, arXiv:cond-mat/0303551, (2003).

[4] W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal, "Quantum Monte Carlo simulations of solids," *Reviews of Modern Physics*, **73**, 33, (2001).

[5] R. J. Bartlett and M. Musiał, "Coupled-cluster theory in quantum chemistry," *Reviews of Modern Physics*, **79**, pp. 291–352, (2007).

[6] U. Schollwöck, "The density-matrix renormalization group in the age of matrix product states," *Annals of Physics*, **326**, 96, arXiv:1008.3477, (2011).

[7] M. Hastings, "An area law for one-dimensional quantum systems," *Journal of Statistical Mechanics*, **2007**, 08, P08024, arXiv:0705.2024, (2007).

[8] F. Verstraete, V. Murg, and J. Cirac, "Matrix product states, projected entangled pair states, and variational renormalization group methods for quantum spin systems," *Advances in Physics*, **57**, 143, arXiv:0907.2796, (2008).

[9] F. Verstraete and J. I. Cirac, "Renormalization algorithms for Quantum-Many Body Systems in two and higher dimensions," arXiv:cond-mat/0407066, (2004).

[10] N. Schuch and J. I. Cirac, "Matrix product state and mean-field solutions for one-dimensional systems can be found efficiently," *Physical Review A*, **82**, 012314, arXiv:0910.4264, (2010).

[11] D. Aharonov, I. Arad, and S. Irani, "Efficient algorithm for approximating one-dimensional ground states," *Physical Review A*, **82**, 012315, arXiv:0910.5055, (2010).

[12] I. Arad, A. Kitaev, Z. Landau, and U. Vazirani, "An area law and sub-exponential algorithm for 1D systems," arXiv:1301.1162, (2013).

[13] T. Osborne, "Efficient Approximation of the Dynamics of One-Dimensional Quantum Spin Systems," *Physical Review Letters*, **97**, 157202, arXiv:quant-ph/0508031, (2006).

[14] Y. Huang, "Area law in one dimension: Degenerate ground states and Renyi entanglement entropy," arXiv:1403.0327, (2014).

[15] F. Verstraete and J. Cirac, "Matrix product states represent ground states faithfully," *Physical Review B*, **73**, 094423, arXiv:cond-mat/0505140, (2006).

[16] C. Eckart and G. Young, "The approximation of one matrix by another of lower rank," *Psychometrika*, **1**(3), (1936).

[17] D. Aharonov, I. Arad, Z. Landau, and U. Vazirani, "The Detectability Lemma and Quantum Gap Amplification," arXiv:0811.3412, (2008).

[18] D. Aharonov, I. Arad, Z. Landau, and U. Vazirani, "Quantum Hamiltonian complexity and the detectability lemma," arXiv:1011.3445, (2010).

[19] I. Arad, Z. Landau, and U. Vazirani, "Improved one-dimensional area law for frustration-free systems," *Physical Review B*, **85**, 195145, arXiv:1111.2970, (2012).

[20] D. Aharonov, I. Arad, Z. Landau, and U. Vazirani, "The 1D Area Law and the Complexity of Quantum States: A Combinatorial Approach," *2011 IEEE 52nd Annual Symposium on Foundations on Computer Science*, 324–333, IEEE, (2011).

[21] Y. Huang, "A polynomial-time algorithm for approximating the ground state of 1D gapped Hamiltonians," arXiv:1406.6355v3, (2014).

[22] Y. Huang, "A polynomial-time algorithm for approximating the ground state of 1D gapped Hamiltonians," arXiv:1406.6355v6, (2014).

[23] I. Arad, Z. Landau, U. Vazirani, and T. Vidick, "Rigorous RG algorithms and area laws for low energy eigenstates in 1D," arXiv:1602.08828, (2016).

A Overlap lemma

In this appendix we want to relax the notion of a basis, and show that key properties thereof are robust to a small amount of allowed error.

Suppose we have some g-dimensional space, represented by a g-rank projector G onto this space. Take $|v_i\rangle$ for $1 \leq i \leq g$ to be an orthonormal set of vectors, each with $\langle v_i|G|v_i\rangle \geq 1 - \delta/g$. First we will prove that all members of $S := \text{Span}\{|v_i\rangle\}$ have non-trivial overlap with G. We will then use this to show that the vectors $G\,|v_i\rangle$ span the space. Finally we will show that any vector orthogonal to $\text{Span}\{|v_i\rangle\}$ has low overlap with the space.

Lemma A.1. *Any vector $|v\rangle \in S$ has $\langle v|G|v\rangle \geq 1 - \delta$.*

Proof. As $|v\rangle$ lies in S, we can write it as a linear combination of the form

$$|v\rangle := \sum_{i=1}^{g} c_i \, |v_i\rangle \ .$$

Next we can bound the expectation value of the complementary projector $I - G$ by using the triangle and Cauchy-Schwartz inequalities.

$$
\begin{aligned}
\langle v|(I-G)|v\rangle &= \sum_{i,j} c_i^* c_j \, \langle v_i|(I-G)|v_j\rangle \\
&\leq \sum_{i,j} |c_i|\,|c_j|\,|\langle v_i|(I-G)|v_j\rangle| \\
&\leq \left[\sum_{i=1}^{g} |c_i|\,\|(I-G)\,|v_i\rangle\|\right]^2 \\
&\leq \delta
\end{aligned}
$$

where we have used the normalisation $\sum_i |c_i|^2 = 1$ and the resulting bound $\sum_i |c_i| \leq \sqrt{g}$. This in turn gives $\langle v|G|v\rangle \geq 1 - \delta$. $\qquad\square$

Lemma A.2 (Basis under projection). *The projections $G\,|v_i\rangle$ form a basis of $\mathrm{im}(G)$.*

Proof. As the number of vectors matches the dimension of the target space, proving linear independence is sufficient to prove the formation of a basis. If the projected vectors were linearly dependent then this would imply there exist a non-zero $|v\rangle \in S$ such that $G\,|v\rangle = 0$. Lemma A.1 however gives $\langle v|G|v\rangle \geq 1 - \delta \geq 0$ for all such vectors, thus proving linear independence by contradiction. $\qquad\square$

Next we want to generalise the concept of a *full* basis, specifically the fact that any vector perpendicular to a full basis is necessarily perpendicular to the entire space, meaning that there is a natural maximum to the size of any given basis.

Lemma A.3 (Fullness). *Any vector $|v'\rangle$ which is orthogonal to S has $\langle v'|G|v'\rangle \leq \delta$.*

Proof. As the projection of this space is spanning, we can pick $|v\rangle \in S$ such that it projects into the same vector as $|v'\rangle$, i.e. $G\,|v\rangle \propto G\,|v'\rangle$. As such we can decompose both states as

$$|v\rangle = \sqrt{\lambda}\,|g\rangle + \sqrt{1-\lambda}\,|h\rangle$$
$$|v'\rangle = \sqrt{\lambda'}\,|g\rangle + \sqrt{1-\lambda'}\,|h'\rangle$$

where $|g\rangle \in \mathrm{im}(G)$ and $|h\rangle\,, |h'\rangle \in \ker(G)$. The orthogonality of $|v\rangle$ and $|v'\rangle$ gives

$$0 = \langle v|v'\rangle = \sqrt{\lambda}\sqrt{\lambda'} + \sqrt{1-\lambda}\sqrt{1-\lambda'}\,\langle h|h'\rangle\ .$$

Using $|\langle h|h'\rangle| \leq 1$ in turn gives $\lambda + \lambda' \leq 1$. Lemma A.1 gives that $\lambda \geq 1 - \delta$, leading us to conclude that $\lambda' \leq \delta$, and thus that $\langle v'|G|v'\rangle \leq \delta$ generically for any vector orthogonal to S. $\qquad\square$

B   Approximate AGSP construction

In this section we explicitly lay out the various errors associated with approximating the AGSP defined in Sections 2.5 and 3.4. The scaling of these errors was stated in less detail in Ref. [2]. First we consider the AGSP itself

$$A := \exp\left[-\frac{x(H - \epsilon_0')^2}{2\epsilon^2}\right] = \frac{\epsilon}{\sqrt{2\pi x}}\int_{-\infty}^{+\infty} \exp\left[-\frac{\epsilon^2 t^2}{2x} - i(H - \epsilon_0')t\right]\,\mathrm{d}t$$

where ϵ_0' is the ground energy approximation given by $|\gamma_1\rangle$. If we take $\eta, \zeta \leq 1/2$ and assume that $\zeta \geq \eta^2/4f$ then

$$\left|\epsilon_0' - \epsilon_0\right| \leq \frac{\eta^2\epsilon}{4f} \leq \epsilon\zeta \leq \frac{\epsilon}{2}\ .$$

The magnitude of this AGSP on the excited subspace, known as the *shrinking factor*, can be bounded

$$\left\|A\left|\Gamma^\perp\right\rangle\right\| \leq \exp\left[-\frac{x(\epsilon/2)^2}{2\epsilon^2}\right] = e^{-x/8}\ .$$

If we define $x := 33 - 8\log\zeta$, then this shrinking factor is at most $\zeta/2$. Further assuming that $\zeta \leq 1/25$, the magnitude of A on the ground space can be lower bounded

$$\|A\,|\Gamma\rangle\| = \exp\left[-\frac{x(\epsilon_0 - \epsilon_0')^2}{2\epsilon^2}\right]$$
$$\geq \frac{19}{20}\ .$$

Next we move on to the approximate AGSP, defined by

$$K := \frac{2\epsilon}{\sqrt{2\pi x}} \sum_{j=0}^{\lceil T/\tau \rceil} \exp\left[i\epsilon_0' \tau j - \frac{\epsilon^2 \tau^2 j^2}{2x}\right] U_D(\tau j).$$

where T, τ and D are parameters and U_D is some unitary operator all of which we shall define below. This definition relies on taking the Fourier decomposition of A, and approximating this by truncating and discretising the integral and approximating the propagators by a known algorithm [13].

B.1 Time Truncation Error

The truncation error is due to the finite range of the integral, controlled by the parameter T. If we let δ_T denote the truncation error (in the operator norm) then it can be bounded

$$
\begin{aligned}
\delta_T &= 2\left\|\frac{\epsilon}{\sqrt{2\pi x}} \int_T^\infty \exp\left[-\frac{\epsilon^2 t^2}{2x} - iHt\right] \mathrm{d}t\right\| \\
&\leq \frac{2\epsilon}{\sqrt{2\pi x}} \int_T^\infty \exp\left[-\frac{\epsilon^2 t^2}{2x}\right] \mathrm{d}t \\
&= \frac{\epsilon}{\sqrt{2x}} \mathrm{erfc}\left(\frac{\epsilon T}{\sqrt{2x}}\right) \\
&\leq \exp\left[-\frac{\epsilon^2 T^2}{2x}\right],
\end{aligned}
$$

where the last inequality holds for sufficiently large T, e.g. $T \geq 1$. For the time-scale

$$T := \frac{\sqrt{2x}}{\epsilon}\sqrt{\log 3/\zeta'} = \mathcal{O}\left(\epsilon^{-1}\log(1/\zeta)\sqrt{\log(n/\zeta)}\right)$$

we can in turn upper bound δ_T by $\delta_T \leq \zeta'/3 = \Omega(\zeta/n)$, where ζ' is defined in Eq. (12).

B.2 Discretisation Error

Next we consider the discretisation error associated with approximating this integral with a Riemann sum. The term being approximately integrated is a Gaussian centred on the origin and so its absolute value is even as a function of t, allowing us to consider only the $t > 0$ error and doubling it. On this range the absolute value is not only even, but monotonically decreasing. As such the concepts of left and right Riemann sums correspond to upper and lower Riemann sums. If we let the discretisation error be δ_D then

$$\delta_D = 2\frac{\epsilon}{\sqrt{2\pi x}}\left\|\sum_{j=0}^{\lceil T/\tau \rceil} \tau \exp\left[-\frac{\epsilon^2 \tau^2 j^2}{2x} - iH\tau j\right] - \int_0^T \exp\left[-\frac{\epsilon^2 t^2}{2x} - iHt\right]\mathrm{d}t\right\|.$$

This error can be upper bounded in turn by the difference between upper and lower Riemann sums which is equal to the left and right Riemann sums owing to the monotonicity of the integrand.

$$\delta_D \leq 2\frac{\epsilon\tau}{\sqrt{2\pi x}}\left\|\sum_{j=0}^{\lceil T/\tau \rceil} \exp\left[-\frac{\epsilon^2 \tau^2 j^2}{2x} - iH\tau j\right] - \sum_{j=1}^{\lceil T/\tau \rceil+1} \exp\left[-\frac{\epsilon^2 \tau^2 j^2}{2x} - iH\tau j\right]\right\|.$$

The difference between these two summations is telescopic. As such it equals the difference between the highest and lowest terms.

$$\delta_D \leq 2\frac{\epsilon\tau}{\sqrt{2\pi x}}\left\|1 - \exp\left[-\frac{\epsilon^2 T^2}{2x} - iHT\right]\right\|.$$

Using the trivial bound on the real exponential term, we get

$$\delta_D \leq 2\frac{\epsilon\tau}{\sqrt{2\pi x}}\left\|1 - \exp\left[-iHT\right]\right\|$$
$$\leq 4\frac{\epsilon\tau}{\sqrt{2\pi x}}.$$

Again this error can also be bounded by $\delta_D \leq \zeta'/3$ if we take the time-step

$$\tau^{-1} := \frac{12}{\zeta'}\sqrt{\frac{\log(3/\zeta')}{\pi}} = \frac{12\epsilon}{\zeta'\sqrt{2\pi x}} = \mathcal{O}\left(\frac{n}{\zeta\sqrt{\log 1/\zeta}}\right).$$

B.3 Bond Truncation Error

To upper bound the total truncation error for the entire approximation by $\zeta'/3$, it is sufficient to bound the truncation error for each propagator $\exp(-iH\tau j)$ by $\zeta'/3$ since the total sum is normalized. Using the construction of Ref. [13] this error can be achieved if the bond dimension of each propagator is

$$D = \exp\left[\mathcal{O}\left(\tau j\right) + \mathcal{O}\left(\log(n^2/\zeta')\right)\right]$$
$$\leq \exp\left[\mathcal{O}\left(T\right) + \mathcal{O}\left(\log n/\zeta\right)\right]$$
$$= 2^{\mathcal{O}(T)}\mathsf{poly}(n/\zeta)$$
$$= 2^{\mathcal{O}\left(\epsilon^{-1}\log(1/\zeta)\sqrt{\log(n/\zeta)}\right)}\mathsf{poly}(n/\zeta).$$

As bond dimension is additive, and there are $2T/(\tau + 1)$ terms, the total bond dimension for the AGSP B is

$$B \leq 2T\tau^{-1}D$$
$$= \mathcal{O}\left(\epsilon^{-1}\log(1/\zeta)\sqrt{\log(n/\zeta)}\right)\mathcal{O}\left(\frac{n}{\zeta\sqrt{\log 1/\zeta}}\right)2^{\mathcal{O}\left(\epsilon^{-1}\log(1/\zeta)\sqrt{\log(n/\zeta)}\right)}\mathsf{poly}(n/\zeta)$$
$$= 2^{\mathcal{O}\left(\epsilon^{-1}\log(1/\zeta)\sqrt{\log(n/\zeta)}\right)}\mathsf{poly}(n/\zeta)$$
$$= n^{\mathcal{O}(1+\epsilon^{-1}\sqrt{\log(1/\zeta)/\log(n)})}$$

where we have assumed $\zeta^{-1} = n^{\mathcal{O}(1)}$.

C Frustration bound

Consider a 1D system with nearest neighbour interactions, were each local Hamiltonian term H_i is bounded $0 \leq H_i \leq \mathbb{1}$. For a cut $i|i+1$ we split the Hamiltonian into a left-Hamiltonian $H_L = \sum_{j=1}^{i-1} H_j$, middle-Hamiltonian $H_M = H_i$, and right-Hamiltonian $H_R = \sum_{j=i+1}^{n} H_j$; we refer to their expectation values as the left/middle/right-energy.

Lemma C.1. *For a Hamiltonian normalised as above, the left-energies relative to any cut for any two states with an energy less than ΔE over the ground, differ by no more than $1 + \Delta E$.*

State	$\langle H \rangle$	$\langle H_L \rangle$	$\langle H_M \rangle$	$\langle H_R \rangle$
$|v_1\rangle$	E	L	$E - L - R - \Delta R$	$R + \Delta R$
$|v_2\rangle$	E	$L + \Delta L$	$E - L - R - \Delta L$	R

Table 2: Local energies of the two ground states across a bipartition.

Proof. Suppose we have two states $|v_1\rangle$ and $|v_2\rangle$ with local-energies given in Table 2. where $E = E_0 + \Delta E$, and E_0 is the ground energy, and we take $\Delta L \geq 0$ without loss of generality. Using the positivity of H_M on the expectations for $|v_2\rangle$ gives

$$L + R \leq E - \Delta L \,.$$

As $|v_1\rangle$ has a left-energy L, it must possess at least one left Schmidt vector $|l\rangle$ with a left-energy at most L, similarly $|v_2\rangle$ has a right Schmidt vector $|r\rangle$ with right-energy at most R. As such we can construct a state $|v\rangle = |l\rangle \otimes |r\rangle$ with energy upper bounded

$$
\begin{aligned}
\langle v|H|v\rangle &= \langle v|H_L|v\rangle + \langle v|H_M|v\rangle + \langle v|H_R|v\rangle \\
&= \langle l|H_L|l\rangle + \langle v|H_M|v\rangle + \langle r|H_R|r\rangle \\
&\leq L + 1 + R \\
&\leq E + 1 - \Delta L \,.
\end{aligned}
$$

As $\langle v|H|v\rangle \geq E_0$, we get that in general the difference between the left-energies is bounded

$$\Delta L \leq 1 + \Delta E \,.$$

A simple example of a perturbed Ising model shows that this bound is tight. $\qquad \square$