The detectability lemma and its applications to quantum Hamiltonian complexity

Dorit Aharonov\textsuperscript{1}, Itai Arad\textsuperscript{1}, Umesh Vazirani\textsuperscript{2} and Zeph Landau\textsuperscript{2}

\textsuperscript{1} School of Computer Science and Engineering, The Hebrew University, Jerusalem, Israel
\textsuperscript{2} Department of Computer Science, UC Berkeley, Berkeley, CA, USA
E-mail: dorit.aharonov@gmail.com, arad.itai@gmail.com, zeph.landau@gmail.com and vazirani@eecs.berkeley.edu

\textit{New Journal of Physics} 13 (2011) 113043 (21pp)
Received 16 September 2011
Published 30 November 2011
Online at \url{http://www.njp.org/}
doi:10.1088/1367-2630/13/11/113043

Abstract. Quantum Hamiltonian complexity, an emerging area at the intersection of condensed matter physics and quantum complexity theory, studies the properties of local Hamiltonians and their ground states. In this paper we focus on a seemingly specialized technical tool, the detectability lemma (DL), introduced in the context of the quantum PCP challenge (Aharonov et al 2009 arXiv:0811.3412), which is a major open question in quantum Hamiltonian complexity. We show that a reformulated version of the lemma is a versatile tool that can be used in place of the celebrated Lieb–Robinson (LR) bound to prove several important results in quantum Hamiltonian complexity. The resulting proofs are much simpler, more combinatorial and provide a plausible path toward tackling some fundamental open questions in Hamiltonian complexity. We provide an alternative simpler proof of the DL that removes a key restriction in the original statement (Aharonov et al 2009 arXiv:0811.3412), making it more suitable for the broader context of quantum Hamiltonian complexity. Specifically, we first use the DL to provide a one-page proof of Hastings’ result that the correlations in the ground states of gapped Hamiltonians decay exponentially with distance (Hastings 2004 \textit{Phys. Rev. B} 69 104431). We then apply the DL to derive a simpler and more intuitive proof of Hastings’ seminal one-dimensional (1D) area law (Hastings 2007 \textit{J. Stat. Mech.} (2007) P8024) (both these proofs are restricted to frustration-free systems). Proving the area law for two and higher dimensions is one of the most important open questions in the field of Hamiltonian complexity, and the combinatorial nature of the
1. Introduction

A striking difference between quantum and classical systems is the number of parameters required to describe them. Whereas a classical system of \( n \) particles can generally be described by \( \mathcal{O}(n) \) parameters, the state of a similar quantum system would generally require \( 2^{\mathcal{O}(n)} \) parameters. A fundamental question is the following: for what class of states is this exponential gap inherent? This is the central issue in quantum Hamiltonian complexity, an emerging field at the intersection of condensed matter physics and quantum computational complexity theory.

Structurally, this exponential gap is directly related to the phenomenon of entanglement, which is the fundamental obstacle in the quest to simulate quantum systems on a classical computer, and has far-reaching implications for the power of quantum computers. The central object of study in quantum Hamiltonian complexity is ground states of local Hamiltonians. What are the conditions under which these ground states can be described efficiently, and what are the conditions under which such description does not exist? There is a beautiful sequence of papers using structures based on tensor networks: MPS [6–9], PEPS [10], TN [11] and MERA [12], which provide efficient descriptions in certain cases. Indeed, a necessary condition for these descriptions to be efficient is that the quantum states must have small entanglement.

Area laws constitute one of the most important tools for bounding entanglement in such systems. Consider the interaction graph (hypergraph) associated with a local Hamiltonian—it
has a vertex for each particle and an edge for each term of the Hamiltonian. Intuitively and very roughly, an area law says that entanglement is local in this interaction graph in the following sense: consider a subset of particles $L$. Then the entanglement between $L$ and $\bar{L}$ in the ground state is locally ‘concentrated’ along the edges between $L$ and $\bar{L}$; more precisely, the area law states that the entanglement entropy across the cut is big-Oh of the number of edges crossing between $L$ and $\bar{L}$. This is clearly a very strong restriction on the entropy, which in the general case would be of the order of the number of particles (nodes) in $L$. Proving area laws for typical classes of Hamiltonians is thus a holy grail in quantum Hamiltonian complexity.

A few years ago, in a seminal paper [3], Hastings proved that the area law holds for one-dimensional (1D) systems (i.e. when the interaction graph is a path), for gapped Hamiltonians—that is, Hamiltonians whose overall spectral gap is of order $O(1)$. In this case, the area law says that ground state entanglement across any contiguous cut is bounded by a constant. From this, one can deduce that the ground state of such systems can be described efficiently (by an MPS of polynomial bond dimension—see [3]). The question of whether ground states in two and higher dimensions obey an area law is still open.

Hastings’ proof of the 1D area law, and many other proofs related to entanglement and correlations in ground states, use sophisticated analytic methods. Perhaps the most important of these is the famous Lieb–Robinson bound (LR bound) [2, 13], which bounds the velocity at which disturbances propagate in quantum local systems; Fourier analysis, and other techniques are important players too.

In this paper, we introduce a combinatorial tool for analyzing correlations and entanglement in ground states of local Hamiltonians. This is a simple, basic version of the detectability lemma (DL) of [1]. We demonstrate that when the system is frustration-free, some of the important results that rely on the traditional analytic tools can be obtained in a much simpler, direct and intuitive way using this tool; we argue that the DL in this form constitutes a useful tool for the study of local Hamiltonians and their ground states.

### 1.1. Results

#### 1.1.1. The detectability lemma (DL).

Our starting point is the DL introduced in [1]. The DL was developed in order to sensibly make a statement of the form ‘If the ground state energy is at least $k$ then the probability that it violates (or frustrates) at least $ck$ terms of the Hamiltonian is bounded below by a constant’. The motivation for developing such a tool in [1] was quite specific: to help translate certain classical results about constraint satisfaction problems to quantum results about local Hamiltonians. In particular, it was used to prove a quantum analogue of gap amplification, a component of Dinur’s proof of the PCP theorem [14], which is a cornerstone in classical computational complexity. The work [1] holds under the assumption that each particle participates in a bounded number of terms of the Hamiltonian (an assumption that holds for most systems of interest including interactions on a lattice embedded in any constant dimensions). The work [1] also required an additional technical assumption, that the number of distinct types of terms of the Hamiltonian is bounded. As we shall see, this latter assumption is no longer needed.

Here, we reformulate the DL and put it into a different context, that of understanding local properties of ground states. More precisely, our reformulation of the DL asks the following question. Consider a gapped frustration-free local Hamiltonian $H = \sum_{i=1}^{m} H_{i}$ with $0 \leq H_{i} \leq 1$, in which the ground energy is 0 and the spectral gap is $\epsilon = O(1)$. Denote the ground state

---

*New Journal of Physics 13 (2011) 113043 (http://www.njp.org/)*
by $|\Omega\rangle$. We would like to understand the properties of the ground state, such as the behavior of correlations between distant particles in the ground state and the propagation in space of perturbations. To this end, we would like to approximate the projection on the ground state, $\Pi_{gs} = |\Omega\rangle\langle\Omega|$, by a ‘local’ operator (for some notion of locality to be clarified soon).

One way of approaching this question is to ask: is there a local operator that fixes the ground state but shrinks all other eigenvectors significantly? A natural first guess at such an operator is the operator $G = (1 - \frac{1}{m}H)$. It fixes the ground space and shrinks its orthogonal complement by a factor that can be as large as $(1 - \epsilon/m)$. As $m$ increases, the factor approaches 1 and the shrinking becomes negligible. This is not strong enough for various purposes, as we will see; our challenge is to do better. Suitably reformulated, the DL provides such an improvement: an operator that fixes the ground state but shrinks all other eigenstates by a constant factor, independent of $m$.

For simplicity of the current discussion, let us provide the exact statement of the DL, and in particular the definition of the DL operator, under the assumption that the particles are on a 1D chain and that $H_i$ are projections $Q_i$ that act on adjacent particles, so that $H = \sum_{i=1}^{n-1} Q_i$.

The projections into the common ground spaces of these layers are given by $\Pi^{\text{odd}} = P_1 P_3 P_5 \cdots$ and $\Pi^{\text{even}} = P_2 P_4 P_6 \cdots$. The DL states that the operator

$$A \overset{\text{def}}{=} \Pi^{\text{even}} \Pi^{\text{odd}}$$

(1)

is the ‘local’ operator we want.

**Lemma 1.1 (DL in 1D).** Let $A$ be the operator defined above in (1), and let $\mathcal{H}'$ be the orthogonal complement of the ground space. Then

$$\|A|_{\mathcal{H}'}\| \leq \frac{1}{(\epsilon/2 + 1)^{1/3}}.$$  (2)

This statement follows from the DL of [1]; we give here a new simple proof of this reformulation of the DL, in the process dropping the assumption of [1] about the number of distinct types of terms of the Hamiltonian.

We can now explain what we mean by ‘locality’ of $A$. When $A$ is applied $\ell$ times to some local perturbation $B$ that acts on the ground state $|\Omega\rangle$, there is a pyramid-shaped ‘causality cone’ of projections that is defined by $B$. These are simply all terms that are graph-connected to the operator in question (see figure 2). All the projections outside that cone commute with $B$ and can therefore be absorbed in the ground state, leaving us with a local operator of size $O(\ell)$.  

---

**Figure 1.** Partitioning the local terms of the 1D, nearest-neighbor Hamiltonian $H = \sum_{i=1}^{n-1} H_i$ into two subsets (layers): an even layer, consisting of the terms $H_i$ that work on particles $i, i + 1$ with even $i$, and a complementary odd layer.
Figure 2. The ‘causality cone’ of projections that is defined by a local operator \( B \) working on \( \ell = 4 \) layers. These are the local terms in the layers that are graph-connected to \( B \) (shaded ovals).

The DL implies that \( A^\ell \) can be viewed as an excellent approximation to the ground state projection \( \Pi_{\text{gs}} \). Indeed, as \( A \) shrinks all the states that are perpendicular to the ground space by a constant factor, it follows that

\[
\| A^\ell - \Pi_{\text{gs}} \| \leq e^{-O(\ell)}.
\]

Using \( A^\ell \), we can deduce various properties of the ground state.

1.1.2. Exponential decay of correlations in the ground state. As a first example of how the DL can be used to handle correlations in the ground state, we provide a one-page proof of Hastings’ result \([2]\) that the correlations in the ground states of gapped Hamiltonians decay exponentially with distance. This applies to \( D \)-dimensional grids for any \( D \). More precisely, consider two local observables \( X \) and \( Y \), which act on sets of particles that are of distance \( m \) on the grid; the decay of correlations means that the expectation value of their product is (almost) equal to the product of their expectation, up to an error that decays exponentially with the grid distance between those operators.

1.1.3. Simplifying the proof of Hastings’ one-dimensional (1D) area law. Our main application of the new formulation of the DL is to give a significant simplification of the proof of Hastings’ area law in 1D \([3]\). The proof uses the DL in two key points, completely bypassing the analytic methods, which are now replaced by combinatorial constructions. By this we hope to make this important result accessible to a wider audience, as well as possibly extendable to higher dimensions. The outline of the proof still follows that of Hastings and, indeed, it helps to clarify and shed light on the essence of this important proof; we defer the explanation of the proof and how the DL enters the picture to section 5.

We mention that at first sight, one might connect the exponential decay of correlations to an intuition that entanglement between a region \( L \) and its surrounding is ‘located’ only close to the boundary of \( L \), and thus scales as the area rather than as the volume. Although an appealing intuition, such an implication of exponential decay of correlation to area laws is not known, and indeed quantum expanders provide a counter-example to a naïve application of this logic \([15]\).

1.1.4. The DL versus the LR bound. In both the applications described above, the DL replaces a combination of the LR bound and other analytic tools. In fact, there is a more general outline for which both applications can be viewed as examples. We illustrate how exactly the DL is related to the analytic methods and how it can be used to replace them, with a more general toy application comparing the use of the LR bound to the alternative route offered by the DL. This is done in section 6.
1.2. Discussion

1.2.1. Further improvements to the 1D area law. Soon after the first publication of this paper on the arXiv [16], we used the DL to provide a completely new proof for the 1D area law [5]. Whereas the 1D proof in this paper follows the same outline as Hastings’ proof, though with the DL replacing the analytical tools, the proof in [5] follows a very different route. There, the main idea is to start with the DL operator $A$ and ‘dilute’ it to achieve an operator that remains a good approximation to the ground state projection, yet creates much less entanglement. As a result, the bound it achieves on the entanglement entropy is exponentially lower (i.e. better) than the bound in Hastings’ 1D proof (and the version presented here). To be specific, for a 1D system $H = \sum_{i=1}^{n-1} H_i$ of $n$ particles of dimension $d$ with a spectral gap $\epsilon > 0$ and an interaction strength $\|H_i\| \leq J$, one defines the parameter $X \equiv J\log d / \epsilon$. Then Hastings’ entropy bound (reproved here) is given by $S \leq e^{O(X)}$, whereas the proof in [5] gives $S \leq e^{O(\log^2 X)}$.

Shortly after the submission of [5], a major simplification in the proof of [5] was discovered. This led to a third paper [4] that is not only much simpler than [5], but also gives a much better bound: $S \leq O(1) \cdot X^3 \log^8 X$. We mention that this bound is only within a polynomial factor from the best known lower bounds for $S$ which scale as $S \geq O(\epsilon^{-1/4})$ [17, 18].

1.2.2. The frustration-free restriction. Our results hold in the frustration-free case. We note, however, that there are various frustration-free systems that are interesting from a physics and a computational point of view: for example the ferromagnetic XXZ model, the AKLT model [20] and stabilizer codes such as the Toric code [21]. In addition, many of the quantum phenomena are revealed already in the context of frustration-free Hamiltonians, and the major open problems in quantum Hamiltonian complexity (e.g. quantum PCP and 2D area law) are open already for this case. Much has to be learned from studying frustration-free Hamiltonians before we proceed to the more general case. Moreover, Hastings’ original proof [3] essentially reduces the frustrated case into an approximately frustration-free system by coarse graining, and a similar approach may work to extend the results presented here to the general case.

1.2.3. Other related works. The DL seems to be connected to various diverse scientific areas. The connection to the LR bound and other analytic tools used in condensed matter physics is discussed extensively in section 6; another connection is to view the DL operator $A$ as a special instance of the general method of alternating projections (MAP), which was first studied by von Neumann [22]. In that method, one applies a fixed sequence of projections in order to approach the intersection subspace. In the general setting, the projections are not assumed to be local, nor is the Hilbert space assumed to be of finite dimension. In recent results [23], the convergence rate is given as a function of the Fridriechs angle, which is not easily related to a physical quantity. The DL, on the other hand, is a MAP under the special assumption that the projections are local, associated with a frustration-free $k$-local Hamiltonian, with a convergence rate that is given as a function of the spectral gap. It would be interesting to see whether more insights can be derived from these connections.

The examples in these papers are of frustrated Hamiltonians, whereas our results apply to frustration-free systems. However, the construction in [18] is translation invariant. It can be easily modified to be made frustration-free [19] if the translation-invariance requirement is dropped, yielding a system in which the entanglement entropy $S$ is lower-bounded by $\epsilon^{-1/6}$, where $\epsilon > 0$ is the spectral gap of the system.
In recent years, much attention has been given to a quantum algorithm that, given a local Hamiltonian, uses a process involving random measurements of the energies of the local terms to approach the ground state efficiently (for certain cases) [24, 25]. The algorithm discussed in those papers carries similarities to the situation we are handling here, despite the fact that measurements are applied rather than projections, and also that the terms are chosen randomly, rather than in some fixed order. It seems that the DL lemma, and the energy-norm trade-off, could potentially also be useful for the analysis of such algorithms. In particular, it would be very interesting to see a version of the DL that applies to the case in which the terms are chosen randomly.

1.2.4. Paper organization. We start with notations and preliminaries in section 2. We then provide the statement and proof of the new version of the DL in section 3. We proceed to the one-page proof of the exponential decay in section 4 and then provide the 1D area law proof in section 5. In section 6, we conclude with the toy example comparing the LR bound approach to the DL one.

2. Notations and preliminaries

We consider a \(k\)-local Hamiltonian \(H\) acting on \(\mathcal{H} = (\mathbb{C}^d)^\otimes n\), the space of \(n\) particles of dimension \(d\). \(H = \sum_i H_i\), where each \(H_i\) is a non-negative and bounded operator that acts nontrivially on a constant number of \(k\) qubits (hence the term local Hamiltonian). We assume that \(H\) has a ground space of energy 0, which must therefore also be a common zero eigenspace of all terms \(H_i\). This means that \(H\) is frustration free. We also assume that \(H\) is ‘gapped’, meaning that its lowest eigenvalue is 0 (the ground energy) and all the next are equal to or larger than some constant \(\epsilon > 0\). We denote by \(\mathcal{H}'\) \(\subset \mathcal{H}\) the orthogonal complement ground space of \(H\). Thus, \(\mathcal{H}'\) is an invariant subspace for \(H\), and

\[ H|_{\mathcal{H}'} \geq \epsilon \mathbb{1}. \]  

Most of these assumptions, except for perhaps the frustration-free assumption, are very often used in condensed matter physics.

Throughout this paper, we further assume that the \(H_i\)s are projections and hence would be denoted by \(Q_i\). We define \(P_i\) to be the projection on the ground space of \(Q_i\), \(P_i \overset{\text{def}}{=} \mathbb{1} - Q_i\). The assumption that \(H\) is made of projections is not actually a restriction because we can reduce any frustration-free, bounded and gapped system into that case. Specifically, for \(H = \sum_i H_i\) with \(\|H_i\| \leq K\) and a spectral gap \(\tau > 0\), we first add an appropriate constant to each \(H_i\) such that their ground energy is 0. Then, for every \(i\) we define \(Q_i\) as the projection into the space where the energy of \(H_i\) is greater than 0 and \(P_i \overset{\text{def}}{=} \mathbb{1} - Q_i\) as the projection to the ground space of \(H_i\). Finally, we define the auxiliary Hamiltonian \(H' = \sum_i Q_i\). This system is frustration-free because the original ground states would also be ground states in \(H'\) with a vanishing energy. Moreover, for any state \(|\psi\rangle \in \mathcal{H}'\) and every \(H_i\),

\[ \langle \psi \perp | H_i | \psi \perp \rangle = \langle \psi \perp | Q_i H_i Q_i | \psi \perp \rangle \leq K \langle \psi \perp | Q_i | \psi \perp \rangle, \]

and therefore the gap in \(H'\) is \(\epsilon \geq \tau / K\). It follows that all of our results can be applied to bounded frustration-free Hamiltonians by replacing the gap \(\epsilon\) in DL with the scaled version \(\tau / K\).
Given a state $|\phi\rangle$ and a partition of the qubits to two nonintersecting sets, $R$ and $L$, with the corresponding Hilbert spaces $\mathcal{H}_L$, $\mathcal{H}_R$, we can consider the Schmidt decomposition of the state along this cut: $|\phi\rangle = \sum_j \alpha_j |L_j\rangle \otimes |R_j\rangle$. Here $\alpha_1 \geq \alpha_2 \geq \cdots$ are the Schmidt coefficients. Their squares are equal to the nonzero eigenvalues of the reduced density matrices to either side of the cut $\rho_L(\phi)$ and $\rho_R(\phi)$, which we denote by $\lambda_1 \geq \lambda_2 \geq \cdots$. The Schmidt rank of $|\phi\rangle$ is then the number of nonzero eigenvalues $\lambda_j$ (or the Schmidt coefficients $\alpha_j$), and the entanglement entropy is the entropy of the set $\{\lambda_j\}$ or, equivalently, the von Neumann entropy of the matrix $\rho_L(\phi)$. A straightforward corollary of the Eckart–Young theorem [26] is then that the truncated Schmidt decomposition provides the best approximation to a vector in the following sense.

**Fact 2.1.** Let $|\phi\rangle$ be a vector on $\mathcal{H}_L \otimes \mathcal{H}_R$ and let $\lambda_1 \geq \lambda_2 \geq \cdots$ be the eigenvalues of its reduced density matrix. The largest inner product between $|\phi\rangle$ and a norm one vector with the Schmidt rank $r$ is $\sqrt{\sum_{j=1}^r \lambda_j}$.

Throughout the paper, the function $\log(\cdot)$ denotes the logarithmic function of base 2.

### 3. The detectability lemma (DL)

For clarity of presentation, we will prove the DL in the case stated in the introduction: where the particles are set on a line and the local terms are two-local involving nearest neighbors. This proof contains all the necessary ingredients for the proof of the more general DL in the case where the Hamiltonian has $k$ local terms that can be partitioned into $g$ layers; we make the precise statement of the more general case at the end of this section.

We begin with a simple lemma that quantifies the norm-energy trade-off in the simple case of two projections $X, Y$: we show that if the application of $XY$ does not move a vector very much, then the energy of that vector with respect to $I - Y$ must be small.

**Lemma 3.1.** Given arbitrary projections $X, Y$ and $|v\rangle$ of norm 1, if $\|XYv\|^2 = 1 - \epsilon$, then

$$\|((I - Y)XYv\|^2 \leq \epsilon(1 - \epsilon). \tag{5}$$

The proof consists of a few lines of simple algebra.

**Proof.** Set $|w\rangle \overset{\text{def}}{=} |v\rangle/\|v\|$. Then

$$\|((I - Y)XYv\|^2 = \|v\|^2 \cdot \|(I - Y)Xw\|^2.$$ 

By definition, $|w\rangle$ is a normalized vector inside the support of $Y$ and therefore for every vector $|\psi\rangle$, we have $\|(I - Y)\psi\| \leq \|(I - |w\rangle \langle w|)\psi\|$. Plugging this to the equality above, we find that

$$\|((I - Y)XYv\|^2 \leq \|v\|^2 \cdot \|(I - |w\rangle \langle w|)Xw\|^2$$

$$= \|v\|^2 \cdot \langle w|X(I - |w\rangle \langle w|)X|w\rangle$$

$$= \|v\|^2 \cdot \|Xw\|^2 \cdot (1 - \|Xw\|^2)$$

$$= \|XYv\|^2 \cdot (1 - \|Xw\|^2)$$

$$\leq \|XYv\|^2 \cdot (1 - \|XYv\|^2) = (1 - \epsilon)\epsilon,$$

where the last inequality follows from the fact that $\|Xw\| \geq \|XYv\|$.

□
Let us now proceed to prove the DL in 1D.

**Proof of lemma 1.1.** Suppose $|\psi\rangle \in \mathcal{H}$ is a norm 1 state that is orthogonal to the ground space, and define $|\phi\rangle \overset{\text{def}}{=} A|\psi\rangle$. Note that for every ground state $|\Omega\rangle$, $\langle \Omega|A|\psi\rangle = 0$, and so $|\phi\rangle$ is orthogonal to the ground space. We would like to show that

$$
\|\phi\| \leq \frac{1}{(\epsilon/2 + 1)^{1/3}}.
$$

We will find both a lower and an upper bound for the energy of $|\phi\rangle$, $\langle \phi|H|\phi\rangle$, which will give us an inequality for $\|\phi\|^2$, from which (6) will follow.

The lower bound is straightforward since $|\phi\rangle$ is orthogonal to the ground state, and so

$$
\langle \phi|H|\phi\rangle \geq \epsilon \|\phi\|^2.
$$

We shall now upper bound the energy $\langle \phi|H|\phi\rangle$ by carefully upper bounding the contributions of the individual terms $\langle \phi|Q_i|\phi\rangle$. We begin by noting that these terms are equal to 0 for $i$ odd since $A = \Pi_{\text{odd}}\Pi_{\text{even}}$ and $Q_i\Pi_{\text{odd}} = 0$ for any odd $i$ (recall that $\Pi_{\text{even}}, \Pi_{\text{odd}}$ are products of the projections $P_i = |\Omega\rangle\langle \Omega|$. We now want to bound the contributions coming from the even terms.

For this purpose, we present $A$ in a convenient form, by reordering its terms. We call the triplet product of projections $(P_1 P_3 P_2), (P_5 P_7 P_6), \ldots$ as *pyramids* and denote them by $\Delta_i = P_{4i-3}P_{4i-1}P_{4i-2}$. The remaining terms are combined with the operator $R \overset{\text{def}}{=} P_4 P_8 \ldots$. See figure 3 for an illustration of this structure. Note that by just using the fact that $P_i$ and $P_j$ commute when $i$ and $j$ are not consecutive, we can write

$$
A = \Delta_1 \Delta_2 \cdots \Delta_m R,
$$

where $m$ is the number of pyramids, which is approximately $n/4$.

We will use this reordering to bound the energy contribution of the terms $Q_2, Q_6, \ldots$; a symmetric argument will bound the remaining even terms $Q_4, Q_8, \ldots$. The energy contribution of $Q_{4i-2}$ will be related to the amount of movement produced by the $\Delta_i$ portion of the operator $A$.

The key point in providing this bound is this. We view the transformation of $|\psi\rangle \rightarrow A|\psi\rangle = |\phi\rangle$ as a series of steps given by the application of the pyramids $\Delta_i$. Specifically, letting $|v_i\rangle = \Delta_i\Delta_{i-1} \cdots \Delta_1 R|\psi\rangle$, we consider the transformation $|\psi\rangle \rightarrow R|\psi\rangle \rightarrow |v_1\rangle \rightarrow |v_2\rangle \rightarrow \cdots \rightarrow |v_m\rangle = A|\psi\rangle$. The square of the norm of the first state, after applying $R$, is $a_1 \overset{\text{def}}{=} \|R|\psi\|^2$.

Let $a_i \overset{\text{def}}{=} \|v_i\|^2/\|v_{i-1}\|^2$ be the ‘shrinkage’ of the square norm (or movement) resulting from the application of the $i$th pyramid, for $1 \leq i < m$.

---

*New Journal of Physics* **13** (2011) 113043 (http://www.njp.org/)
We shall now prove, using lemma 3.1, that the shrinkage \( a_i \) is related to the energy of the operator \( Q_{4i-2} \), which is at the top of the pyramid \( \Delta_i \):

\[
\langle \phi | Q_{4i-2} | \phi \rangle \leq 1 - a_i.
\]

We write

\[
\langle \phi | Q_{4i-2} | \phi \rangle = \| (1 - P_{4i-2}) A \psi \|^2
\]

\[
= \| \Delta_m \cdots \Delta_{i+1} \cdot (1 - P_{4i-2}) \cdot \Delta_i v_{i-1} \|^2
\]

\[
\leq \| (1 - P_{4i-2}) \Delta_i v_{i-1} \|^2,
\]

and recall that \( \Delta_i = P_{4i-3} P_{4i-1} P_{4i-2} \). We can therefore apply lemma 3.1 to \((1 - P_{4i-2}) \times \Delta_i \frac{1}{\|v_{i-1}\|} |v_{i-1}\rangle \) (with \( Y = P_{4i-2} \) and \( X = P_{4i-3} P_{4i-1} \)) and conclude that

\[
\langle \phi | Q_{4i-2} | \phi \rangle \leq (1 - \| \Delta_i v_{i-1} \|^2 / \| v_{i-1} \|^2) \cdot (\| \Delta_i v_{i-1} \|^2 / \| v_{i-1} \|^2) \cdot \| v_{i-1} \|^2 = (1 - a_i) \| v_i \|^2 \leq 1 - a_i. \]

Consequently,

\[
\langle \phi | Q_{4i-2} | \phi \rangle \leq 1 - a_i.
\]

Using this upper bound gives an upper bound for the energy contribution for \( Q_i, i \in \{2, 6, 10, \ldots\} \):

\[
\langle \phi | (Q_2 + Q_6 + \cdots) | \phi \rangle \leq \sum_i (1 - a_i),
\]

with the constraint on the norm \( \| \phi \|^2 \leq \prod_i a_i \). The right-hand side is maximized when all the \( a_i \)'s are equal to each other, i.e. \( a_i = \| \phi \|^2 / \| \phi \|^2 \), and therefore we are left with an upper bound of the energy coming from \( Q_2 + Q_6 + \cdots \) as

\[
\langle \phi | (Q_2 + Q_6 + \cdots) | \phi \rangle \leq m [1 - \| \phi \|^{2/m}] \leq \frac{1 - \| \phi \|^2}{\| \phi \|},
\]

where the last inequality follows from the fact that for every \( x \in [0, 1] \), we have \( m [1 - x^{1/m}] \leq \frac{1 - x}{\sqrt{x}} \).

For the energy of \( Q_4 + Q_8 + \cdots \), a similar decomposition to \( A = (P_1 P_3 P_5) (P_1 P_3 P_5) \cdots (P_2 P_6 \cdots) \) can be made, thereby upper bounding the energy by \( 2^{1-\| \phi \|^2 / \| \phi \|^2} \). Combining the energy upper and lower bounds we therefore obtain

\[
\epsilon \| \phi \|^2 \leq \langle \phi | H | \phi \rangle \leq 2 \frac{1 - \| \phi \|^2}{\| \phi \|}
\]

and so

\[
\frac{\epsilon}{2} \| \phi \|^3 \leq 1 - \| \phi \|^2 \leq 1 - \| \phi \|^3.
\]

which gives

\[
\| \phi \| \leq \frac{1}{(\epsilon/2 + 1)^{1/3}}.
\]

The above proof can easily be generalized to other geometries. In the general case, in accordance with section 2, we assume that we have a \( k \)-local, frustration-free Hamiltonian \( H = \sum_i Q_i \) that is made of projections and has a spectral gap \( \epsilon > 0 \). We further assume that each particle participates in a constant number of projections, and therefore the \( Q_i \) can be

\[
\text{New Journal of Physics 13 (2011) 113043 (http://www.njp.org/)}
\]
partitioned into a constant number of \( g \) layers; each layer is made of projections that do not intersect each other and are therefore commuting.

Then for each layer we define the projection \( \Pi_i \) as the product of all \( P_j = 1 - Q_j \) that are in the layer, and define the DL operator \( A \) by

\[
A \equiv \Pi_g \cdots \Pi_1.
\]

Finally, we define \( f(k, g) \) to be the number of sets of pyramids that are necessary to estimate the energy contribution of all the \( Q_i \) terms. In the 1D case that we proved, we had \( f(k, g) = 2 \), because only the even layer contributed energy and we needed two sets of pyramids to cover that layer. In the general case it is easy to see that \( f(g, k) \) can be crudely bounded by \( f(g, k) \leq (g - 1)k^g \).

Using the above definitions, the general DL is as follows.

**Lemma 3.2 (the DL).** Consider the local Hamiltonian system that is described above. Then

\[
\|A|_{H'}\| \leq \frac{1}{[\epsilon / f(k, g) + 1]^{1/3}}.
\]

**4. Exponential decay of correlations**

Consider now a Hamiltonian \( H = \sum_i Q_i \) which is \( k \)-local and set on a \( d \)-dimensional grid. Once again, in accordance with section 2, we assume that \( Q_i \) are projections and that \( H \) is frustration-free with a unique ground state \( |\Omega\rangle \) (i.e. \( Q_i|\Omega\rangle = 0 \)) and a spectral gap \( \epsilon > 0 \). We wish to show

**Theorem 4.1 (decay of correlations in ground states of gapped Hamiltonians on a \( d \)-Dim grid).** Consider a setting as described above. Let \( X, Y \) be two local observables whose distance on the grid from each other is \( m \). Denote \( \bar{X} \equiv \langle \Omega|X|\Omega\rangle \), \( \bar{Y} \equiv \langle \Omega|Y|\Omega\rangle \). Then

\[
|\langle \Omega|(X - \bar{X})(Y - \bar{Y})|\Omega\rangle| = |\langle \Omega|XY|\Omega\rangle - \bar{X}\bar{Y}| \leq \|X\| \cdot \|Y\| \cdot e^{-O(m)}.
\]

**Proof.** Let us now consider two operators: \( P_{in}, P_{out} \): \( P_{in} \) is defined by applying the DL \( \ell \) times to \( Y \) and discarding all projections outside the causality cone of \( Y \). \( \ell \) is chosen such that the resulting cone will not overlap with \( X \) (see figure 4). Therefore \( \ell \propto m \), with the proportionality constant that is a geometrical factor. \( P_{out} \) is the complement of \( P_{in} \), i.e. it is the product of projections one gets by applying the DL \( m \) times, excluding the projections that are inside the causality cone of \( Y \). Together, we have \( P_{in} \cdot P_{out} = A^\ell \)—see figure 4 for an illustration in 1D.

\( P_{in}, P_{out} \) leave the ground-state invariant. In addition, they commute with \( X \) and \( Y \), respectively; hence

\[
\langle \Omega|X = \langle \Omega|X P_{in}, Y|\Omega\rangle = P_{out} Y|\Omega\rangle,
\]

and therefore

\[
\langle \Omega|XY|\Omega\rangle = \langle \Omega|XP_{in} P_{out} Y|\Omega\rangle = \langle \Omega|X A^\ell Y|\Omega\rangle.
\]

We now recall that \( A^\ell \) is in fact an approximation of the ground-state projection \( \Pi_{gs} \) (see (26) in section 6),

\[
A^\ell = \Pi_{gs} + e^{-O(\ell)} = \Pi_{gs} + e^{-O(m)}
\]

\( \text{New Journal of Physics 13 (2011) 113043 (http://www.njp.org/)} \)
and so
\[ |\langle \Omega | X Y | \Omega \rangle - \langle \Omega | X \Pi_{gs} Y | \Omega \rangle| \leq \|X\| \cdot \|Y\| \cdot e^{-O(m)}. \]

Assuming that the ground state is unique, \( \Pi_{gs} = |\Omega\rangle \langle \Omega| \), and therefore
\[ |\langle \Omega | X Y | \Omega \rangle - X Y| \leq \|X\| \cdot \|Y\| \cdot e^{-O(m)}. \]

5. The area law in one dimension using the DL

We consider a two-local Hamiltonian on a chain of \( n \) particles of dimension \( d \), \( H = \sum_{i=1}^{n-1} Q_i \) and, as before, we assume that the \( Q_i \) are projections and that \( H \) is frustration-free with a unique ground state \( |\Omega\rangle \), and a spectral gap \( \epsilon \). Then according to (2),
\[ \Delta = \frac{1}{(\epsilon/2 + 1)^{3/2}} \] is the DL shrinking exponent of the square of the norm. In order to keep the presentation simple we will assume that \( \epsilon \leq 3 \), and therefore \( \frac{1}{2} < \Delta < 1 \). We prove the following version of a 1D area law.

**Theorem 5.1 (area law for frustration-free Hamiltonians in 1D).** For any contiguous cut along the chain, the entanglement entropy of the ground state \( |\Omega\rangle \) across the cut is bounded by a constant which depends on the dimensionality of the particles \( d \) and on the spectral gap \( \epsilon \); specifically,
\[ S \leq O(1) \cdot \frac{\log^2 d}{\epsilon} d^{40/\epsilon} = e^{O(X)}, \] (15)
for \( X \equiv \frac{\log d}{\epsilon} \). We note that in accordance with section 2, for 1D frustration-free systems \( H = \sum_i H_i \), with a spectral gap \( \epsilon > 0 \) and a bounded interaction strength \( \|H_i\| \leq J \), one should replace \( \epsilon \mapsto \epsilon / J \) and consequently \( X \mapsto J \log d / \epsilon \).

The proof relies on two main lemmas. The first shows that for any cut along the line, there is a product state \( |\phi_1\rangle \otimes |\phi_2\rangle \) that has a constant inner product with the ground state.

**Lemma 5.2 (constant overlap with a product state).** For every cut, there is a product state \( |\phi_1\rangle \otimes |\phi_2\rangle \) such that \( |\langle \phi_1 \otimes \phi_2 | \Omega \rangle| \geq \mu \equiv d^{-\ell_0} \Delta^{\ell_0/8} \), with \( \ell_0 \equiv d^{40/\epsilon} \).
The second lemma shows that if there exists a product state with a constant overlap with the ground state $|\Omega\rangle$, then $|\Omega\rangle$ has finite entanglement entropy.

**Lemma 5.3 (constant overlap with a product state implies finite entropy).** If for some cut there exists a product state $|\phi_1\rangle \otimes |\phi_2\rangle$ such that $|\langle \phi_1 \otimes \phi_2 | \Omega \rangle| \geq \mu$, then the entanglement entropy of $|\Omega\rangle$ across the cut is bounded by

$$S \leq O(1) \cdot \frac{\log d}{\epsilon} \log \mu^{-1}.$$  

(16)

Theorem 5.1 then follows easily by combining the two lemmas and using the facts that $\log \mu^{-1} = \ell_0 (\log d + \frac{1}{8} \log \Delta^{-1}) \leq \ell_0 (\log d + \epsilon/2) = O(\ell_0 \log d)$. We prove lemma 5.2 in section 5.2 and lemma 5.3 in section 5.1.

### 5.1. Constant overlap implies finite entropy (proof of lemma 5.3)

In this section we prove lemma 5.3. The DL is clearly the right tool for the task, since it provides a ‘local’ operator that can be repeatedly applied to the promised product state $|\phi_1\rangle \otimes |\phi_2\rangle$ without increasing its entanglement rank much, while exponentially decreasing its distance from the ground state.

The only thing that is not entirely clear is how to get a constant bound on the entanglement entropy of the ground state, since a straightforward argument would mean applying the operator non-constant number of times to get arbitrarily close to the ground state. The key is to observe that after $\ell$ applications of the DL we get a state with a bounded Schmidt rank that is close to the ground state, and by Fact 2.1, this gives us a bound on the sum of the largest $d^2 \ell$ Schmidt coefficients of the ground state. With these bounds we can find a pessimistic constant upper bound on the entanglement entropy. We can now proceed to the more detailed proof.

Consider then a cut in the line between the particles $i_0$ and $i_0 + 1$, and let $Q_{i_0}$ be the local term in $H$ that involves $i_0$, $i_0 + 1$. Assume that along that cut, the product state $|\phi_0\rangle \overset{\text{def}}{=} |\phi_1\rangle \otimes |\phi_2\rangle$ has a constant projection $\mu$ on the ground state $|\Omega\rangle$:

$$|\phi_0\rangle = \mu |\Omega\rangle + (1 - \mu^2)^{1/2} |\Omega^\perp\rangle,$$

(17)

where $|\Omega^\perp\rangle \in H'$. We now apply the operator DL operator $A$ $\ell$ times on $|\phi_0\rangle$. We obtain

$$A^\ell |\phi_0\rangle = \mu |\Omega\rangle + \delta_\ell |\Omega^\perp\rangle,$$

where by the DL, $|\Omega^\perp\rangle \in H'$ and $\delta_\ell^2 \leq \Delta^\ell$. Let $|v_\ell\rangle$ be the normalized version of $A^\ell |\phi_0\rangle$. Then

$$|\langle v_\ell |\Omega\rangle| \geq \frac{\mu}{\sqrt{\mu^2 + \Delta^\ell}} = 1 - O(\Delta^\ell).$$

(18)

This means that $|v_\ell\rangle$ are exponentially close to the ground state, as a function of $\ell$.

How entangled are those states? We note that at each application of $A$, the entanglement rank of the state can only increase by a multiplicative factor of $d^2$: for every $i \neq i_0$, the projection term $P_i$ in $A$ works entirely left to the cut or entirely right to the cut, thereby not increasing the Schmidt rank of the state. The only projection in $A$ that may increase the rank is $P_{i_0}$, and as it
is a two-local projection that works on $d$-dimensional particles, it can at most increase it by a factor of $d^2$. Consequently, the Schmidt rank of $|v_i\rangle$ is at most $d^{2\ell}$.

We have therefore obtained a family of states $\{|v_i\rangle\}$ with Schmidt ranks bounded by $d^{2\ell}$, which are closer and closer to $|\Omega\rangle$. Then using Fact 2.1, together with (18), it follows that the eigenvalues of the reduced density matrix of $|\Omega\rangle$ along the cut, $\lambda_1 \geq \lambda_2 \geq \cdots$, must satisfy

$$\frac{\mu}{\sqrt{\mu^2 + \Delta^\ell}} \leq |\langle v_\ell | \Omega \rangle| \leq \left( \sum_{j=1}^{d^{2\ell}} \lambda_j \right)^{1/2},$$

which implies the following series of inequalities:

for every $\ell \geq 1$:

$$\sum_{j \geq d^{2\ell}+1} \lambda_j \leq \frac{1}{\mu^2} \Delta^\ell. \quad (20)$$

From here, the desired upper bound on the entropy can be deduced by choosing the distribution of maximal entropy which still satisfies the inequalities in (21). The following lemma, whose proof can be found in appendix A, gives one such bound.

**Lemma 5.4.** Consider a probability distribution $\{\lambda_j\}$ whose values are ordered in a non-increasing fashion, $\lambda_1 \geq \lambda_2 \geq \cdots$, and let $D \geq 2$ be an integer and $K \geq 4$, $1/2 < \Delta < 1$ some constants such that

for every $\ell \geq 1$:

$$\sum_{j \geq d^{2\ell}+1} \lambda_j \leq K \Delta^\ell. \quad (21)$$

Then the entropy of $\{\lambda_j\}$ is upper bounded by

$$S \leq O(1) \cdot \frac{\log K \cdot \log D}{\log \Delta^{-1}}. \quad (22)$$

Substituting $D = d^2$ and $K = \mu^{-2}$ gives

$$S \leq O(1) \cdot \frac{\log \mu^{-1} \cdot \log d}{\log \Delta^{-1}}.$$

But looking at (14), we see that as long as $\epsilon \leq 3$ (which is our working assumption), $\log \Delta^{-1} \geq \epsilon/5$, and so

$$S \leq O(1) \cdot \frac{\log d}{\epsilon} \log \mu^{-1}.$$

### 5.2. A product state having constant overlap with $|\Omega\rangle$ (proof of lemma 5.2)

A natural candidate for a tensor product state with a constant overlap with $|\Omega\rangle$ is the mixed state $\rho_L \otimes \rho_R$, where $\rho_L$ is the reduced density matrix of $|\Omega\rangle$ to the left of the cut, and $\rho_R$ is the reduced density matrix to the right.

---

$^5$ Given a vector $|v\rangle = \sum_{j=1}^d a_j |L_j\rangle \otimes |R_j\rangle$ of the Schmidt rank $r$ and $i < i_0$, $P_i |v\rangle = \sum_{j=1}^{d^{2\ell}} a_j (P_i |L_j\rangle) \otimes |R_j\rangle$ and thus has Schmidt rank bounded by $r$ (the symmetric argument shows the same result for $i > i_0$). For $i = i_0$, we can decompose $P_{i_0} = \sum_{k=1}^{d^{2\ell}} X_k \otimes Y_k$ with $X_k$ acting on the $i_0$th particle and $Y_k$ acting on the $(i_0 + 1)$th particle. Consequently, $P_{i_0} |v\rangle = \sum_{j=1}^{d^{2\ell}} a_j (X_k |L_j\rangle) \otimes (Y_k |R_j\rangle)$ has Schmidt rank of at most $rd^2$. 

New Journal of Physics 13 (2011) 113043 (http://www.njp.org/)
Let us assume for contradiction that the overlap between $|\Omega\rangle$ and $\rho_L \otimes \rho_R$, and in fact with any tensor product state along a certain cut, is less than $\Delta^{\ell/8}$ for some sufficiently large constant $\ell$. If the overlap is small, then there is a measurement that distinguishes $|\Omega\rangle$ from $\rho_L \otimes \rho_R$ with a probability of at least $1 - \Delta^{\ell/4}$; this is simply the projection on the ground state, $\Pi_{gs}$.

The challenge is to show that there is such a local measurement, i.e. a measurement confined to a local window, which distinguishes these two states almost as well. Using the DL, we shall now find such a local measurement that distinguishes with a slightly worse probability $1 - 2\Delta^{\ell/4}$.

Let us denote by $\rho_L^\ell$ (respectively $\rho_R^\ell$) the reduced density matrix of $|\Omega\rangle$ restricted to the $\ell$ particles on the left (respectively right) of the cut. Also, let $\rho^{2\ell}$ be $\rho = |\Omega\rangle\langle\Omega|$ restricted to the $2\ell$ particles, $\ell$ on each side of the cut. We refer to the state $\rho_L^\ell \otimes \rho_R^\ell$ as the ‘disentangled’ version of the state $\rho^{2\ell}$. The following lemma shows that under the assumption that $|\Omega\rangle$ has low overlap with every product state $|\phi_1 \otimes |\phi_2\rangle$ (along a given cut), there exists a measurement confined to the window of $2\ell$ particles around the cut, that with high probability distinguishes $\rho^{2\ell}$ from $\rho_L^\ell \otimes \rho_R^\ell$.

**Lemma 5.5 (existence of a distinguishing measurement).** Assuming that the overlap of the ground state with any product state satisfies $\langle\phi_1 \otimes \phi_2| \Omega\rangle \leq \Delta^{\ell/8}$, there is a measurement that distinguishes $\rho^{2\ell}$ from $\rho_L^\ell \otimes \rho_R^\ell$ with probability $1 - 2\Delta^{\ell/4}$.

The DL ensures that by applying the layers one by one, we converge to the projection on the ground state quickly, and it is this projection that is exactly the distinguishing measurement we want to approximate. We can thus apply $A$ only $\ell/2$ times, approximating the projection on the ground space; now, following the intuition explained in the introduction (and in the example of section 6), only the causality cone of the cut should be used in this measurement, and the rest of the operators in those layers are swallowed by the state being measured; this amounts to a measurement which is restricted to the $-\ell, \ell$ interval and still distinguishes well enough. The detailed proof can be found in the appendices.

The fact that such a measurement exists, distinguishing the original state confined to the $2\ell$ window from its ‘disentangled’ version, with high probability, must somehow indicate that there is a lot of entanglement along the cut, whose disentanglement caused this distinguishability. This can be made precise using an information-theoretical argument.

**Lemma 5.6 (distinguishing measurement implies large difference in entropies).** If there is a $2\ell$-local measurement that distinguishes $\rho^{2\ell}$ from $\rho_L^\ell \otimes \rho_R^\ell$ with a probability of at least $1 - 2\Delta^{\ell/4}$, then

$$S(\rho_L^\ell) + S(\rho_R^\ell) - S(\rho^{2\ell}) \geq \frac{\epsilon}{20} \ell - 1.$$  

The lemma implies that the entropy in $S(\rho_L^\ell) + S(\rho_R^\ell)$ is significantly larger than $S(\rho^{2\ell})$, implying that disentangling along the cut has introduced a lot of new entropy. The proof is simple, based on relative entropy; essentially, all it uses is the fact that a measurement that distinguishes with high probability two states implies high relative entropy between the results of the measurements. Once again, details can be found in the appendices.

To finish the proof of lemma 5.2, we now need to derive a contradiction. Denote by $S(2\ell)$ the value of $S(\rho^{2\ell})$, with $2\ell$ being the segment centered around the cut that provides our contradictory assumption (namely, that any tensor product state has less than $\Delta^{\ell/8}$ inner product with the ground state). Under these conditions, lemma 5.5 applies, and hence also
the conditions of lemma 5.6 apply to this segment. Applying lemma 5.6 we conclude that
\[ S(2\ell) \leq 2S(\ell) - \frac{20}{\epsilon} \ell + 1. \]
We now want to recursively apply this inequality for the \( \ell \) long segments on both sides of the cut and then for the \( \ell/2 \)-long segments within those segments and so on. The problem is that the cuts now move to different locations within the \( 2\ell \) long window, and so our assumption no longer applies for these cuts. However, if the inner product state with any tensor product state is small along the original cut, it can be shown to be quite small also along nearby cuts, and so all the above arguments can be applied for those cuts too. This can be formalized in the following claim, whose easy proof can once again be found in the appendices.

**Claim 5.7.** If \( |\langle \phi_1 \otimes \phi_2 | \Omega \rangle| \leq \mu \) for all product states across the cut \((k, k+1)\), then \( |\langle \chi_1 \otimes \chi_2 | \Omega \rangle| \leq \mu d^{\ell} \) for all product states across any cut \((k+j, k+j+1)\), with \(-\ell \leq j \leq \ell\).

We therefore assume by contradiction that the inner product along a given cut is smaller than \( \mu \), such that \( \mu d^{\ell} = \Delta^{\ell/8} \), and so along all the cuts in the \( 2\ell \) window we have that the inner product of the states is at most \( \Delta^{\ell/8} \), and hence our assumptions apply. We can therefore use the same argument recursively. Since \( S(1) \leq \log(d) \), we get (for \( \ell \) a power of 2), \( S(\ell) \leq \ell \log(d) - \frac{20}{\epsilon} \ell \log \ell + \ell \leq \ell(\log(d) + 1) - \frac{20}{\epsilon} \ell \log \ell \). Choosing \( \ell = d^{40/\epsilon} \) then yields \( \frac{20}{\epsilon} \log \ell \geq \log d + 1 \), giving us the contradiction \( S(\ell) < 0 \).

### 6. Comparison of the Lieb–Robinson bound approach and the DL approach

In this section, we compare the DL with a standard method used in many of the seminal results in quantum Hamiltonian complexity, such as Hastings’ areas law for 1D gapped systems [3], and Hastings’ exponential decay of correlations proof [2]. The method combines the use of the Lieb–Robinson bound (LR bound) with a Fourier analysis and the existence of a gap, to reveal the locality properties of the ground state. Specifically, the method uses these tools to approximate expressions that involve the projection operator to the ground state, \( \Pi_{\text{gs}} \), by local operators.

To understand how this is done, let us concentrate on a simple example of locality in the ground state, and derive it using both the DL and the LR bound.

We focus on \( \Pi_{\text{gs}} \), the projection on the ground space of \( H \). On the surface, this projector seems very far from being local in any sense. Nevertheless, in gapped systems it does possess some locality properties that are crucial to the analysis of correlations and entanglement in the ground state. To see this, one standardly considers an approximation of \( \Pi_{\text{gs}} \) by another operator:

\[
P_q \overset{\text{def}}{=} \frac{1}{\sqrt{2\pi q}} \int_{-\infty}^{\infty} dt \ e^{-t^2/2q} \ e^{-iHt}. \tag{23}
\]

Here \( q \) is a free parameter to be chosen as appropriate. For an eigenvector \( |E\rangle \) of \( H \) with eigenvalue \( E \), we have \( P_q |E\rangle = e^{-qE^2/2} |E\rangle \). Consequently, if the system has a constant spectral gap \( \epsilon > 0 \), \( P_q \) approximates \( \Pi_{\text{gs}} \) well:

\[
\| P_q - \Pi_{\text{gs}} \| \leq e^{-\epsilon^2/2}. \]

We now want to argue regarding the local nature of \( P_q \) in various contexts. Let us illustrate the LR bound approach with a simple example: consider the expression \( \Pi_{\text{gs}} B|\Omega\rangle \approx P_q B|\Omega\rangle \),
where $|\Omega\rangle$ is a ground state of the system with zero energy, and $B$ is some local perturbation. It is easy to see that

$$P_q B|\Omega\rangle = \frac{1}{\sqrt{2\pi q}} \int dt \, e^{-t^2/2q} e^{-iHt} B e^{iHt}|\Omega\rangle = \frac{1}{\sqrt{2\pi q}} \int dt \, e^{-t^2/2q} B(t)|\Omega\rangle. \quad (24)$$

$B(t) \overset{\text{def}}{=} e^{-iHt} B e^{iHt}$ is the time evolution of the perturbation $B$. The key point now is to use the famous LR bound, to approximate it by a ‘local’ operator, i.e. an operator that acts only on the ‘neighborhood’ of the particles on which $B$ acts. The following is an immediate corollary of the original LR bound, which we omit for the sake of brevity. The full statement of the LR bound, together with the proof of this corollary, can be found in [27].

**Theorem 6.1 (Lieb–Robinson bound, adapted from [27]).** Given a local Hamiltonian $H = \sum_i H_i$ on $n$ particles, there exists a constant velocity $v$ s.t. $B(t)$ can be approximated by an operator denoted by $B(t)$ whose support is inside a ball of radius $\ell = vt$ around the support of $B$, s.t.

$$\| B(t) - B(t) \| \lesssim \| B \| \cdot e^{-\mathcal{O}(\ell)}. \quad (25)$$

Given a length scale $\ell > 0$, we may now set $q = \ell$ in (24) and obtain

$$\Pi_{gs} B|\Omega\rangle \approx P_t B|\Omega\rangle \approx \frac{1}{\sqrt{2\pi \ell}} \int_{|t| \leq \ell} dt \, e^{-t^2/2\ell} B(t)|\Omega\rangle \approx \frac{1}{\sqrt{2\pi \ell}} \int_{|t| \leq \ell} dt \, e^{-t^2/2\ell} B(t)|\Omega\rangle. \quad (26)$$

In the above series of approximations, $\approx$ implies an approximation of up to an error of $e^{-\mathcal{O}(\ell)}$. The first approximation follows from the assumption of the constant gap and (23). The second approximation is due to the exponential decay of the filter function $e^{-t^2/2\ell}$ and the third one is due to the LR bound. We therefore obtain an exponentially (in $\ell$) good approximation to $\Pi_{gs} B$ in the expression $\Pi_{gs} B|\Omega\rangle$ by an operator which is $\ell$-local.

Let us now derive the same result for the frustration-free case using the DL. First, we approximate the ground space projection $\Pi_{gs}$, by applying the DL operator $A$ for $m$ times. By (12), $A$ leaves the ground space invariant while shrinking the orthogonal space by a constant factor. Therefore

$$\Pi_{gs} = A^m + e^{-\mathcal{O}(m)}. \quad (26)$$

We now write

$$\Pi_{gs} B|\Omega\rangle = A^m B|\Omega\rangle + \| B \| \cdot e^{-\mathcal{O}(m)}$$

and consider the expression $A^m B|\Omega\rangle$. By assumption, the system is frustration-free, and therefore every local projection operator $P_i$ that appears in $A$ leaves $|\Omega\rangle$ invariant: $P_i |\Omega\rangle = |\Omega\rangle$.

We now consider the ‘causality cone’ of projections in $A^m$ that are defined by $B$. These are simply the projections that are graph-connected to $B$ when all the projections in $A^m$ are arranged in consecutive $gm$ layers (see figure 2). The main observation is that all the projections outside this causality cone commute with $B$, and can therefore be absorbed by $|\Omega\rangle$. We are therefore left only with the projections of the causality cone, whose support size $\ell$ is proportional to $m$. In other words, just as in the LR bound method, we found an exponentially (in $\ell$) good approximation to $\Pi_{gs} B$ in the expression $\Pi_{gs} B|\Omega\rangle$ by an operator which is $\ell$-local.
Acknowledgments

We are grateful to Matt Hastings, Tobias Osborne and Bruno Nachtergaele for stimulating discussions on the present and related topics. IA acknowledges support from Julia Kempe’s ERC Starting Grant QUOCO and Julia Kempe’s Individual Research Grant 759/07 of the Israeli Science Foundation. ZL and UV were supported in part by ARO grant W911NF-09-1-0440 and NSF grant CCF-0905626.

Appendix A. Upper bound on the entropy

In this section we prove lemma 5.4, which is required in order to finish the proof of lemma 5.3.

Proof. Define

\[ p_\ell \overset{\text{def}}{=} K \Delta^\ell, \]

and choose \( \ell_0 = \left\lfloor \frac{\log K}{\log \Delta} \right\rfloor = \mathcal{O}(1) \cdot \frac{\log K}{\log \Delta} \), so that \( p_{\ell_0} < 1 \). For \( \ell \geq 2\ell_0 + 1 \), we will upper bound the entropy of the distribution with the bounds

\[ \sum_{j=D^{2(\ell+1)}}^{D^{2(\ell)+1}} \lambda_j^2 \leq p_{2\ell} \leq p_\ell. \]

This is maximized by spreading out the probability mass uniformly in each interval \([D^{2\ell}+1, D^{2(\ell+1)}]\), which results in an upper bound on the contribution of the entropy on the interval \([D^{2\ell}+1, D^{2(\ell+1)}]\) of

\[ p_\ell \log \frac{D^{2(\ell+1)} - D^{2\ell}}{p_\ell} \leq p_{\ell_0} \log \frac{D^{2(\ell+1)} - D^{2\ell}}{p_{\ell_0}} \leq \Delta^{\ell-\ell_0} \log \frac{D^{2(\ell+1)} - D^{2\ell}}{D^{2(\ell+1)} - D^{2\ell}} \leq \Delta^{\ell-\ell_0} (\ell - \ell_0) \log \frac{D^{2(\ell+1)} - D^{2\ell}}{D^{2(\ell+1)} - D^{2\ell}} \Delta. \]  

(A.1)

Above, the second inequality follows from \( p_{\ell_0} < 1 \) along with the fact that the function \( x \log \frac{D^{2(\ell+1)} - D^{2\ell}}{x} \) is increasing for \( x \leq 1 \).

Summing (A.1) over \( \ell \geq 2\ell_0 + 1 \) then yields an upper bound of the entropy contribution of this tail as

\[ \sum_{\ell \geq 2\ell_0 + 1} \Delta^{\ell-\ell_0} (\ell - \ell_0) \log \frac{D^{2(\ell+1)} - D^{2\ell}}{D^{2(\ell+1)} - D^{2\ell}} \Delta \leq \sum_{\ell' \geq \ell_0 + 1} \ell' \Delta^{\ell'} \log \frac{D^4}{\Delta} \leq \Delta (1 - \Delta)^2 \log \frac{D^4}{\Delta}. \]

Above, the first inequality follows from first noting that for the choice of \( \ell \), the exponent on the power of \( D \) is upper bounded by 4, then making the substitution of \( \ell' = \ell - \ell_0 \). The second follows from using the series equality \( \sum_{j \geq 1} j r^j = \frac{r}{(1-r)^2} \).

Combining this estimate with the maximal contribution of entropy over the first \( D^{2(\ell_0+1)} \) terms, namely \( (2\ell_0 + 1) \log D = \mathcal{O}(1) \cdot \frac{\log(\mu - 1)}{\log \frac{1}{\Delta}} \log D \), gives the following bound on the entropy:

\[ S \leq \mathcal{O}(1) \cdot \frac{\log(\mu - 1)}{\log(\Delta^{-1})} \log D + \Delta (1 - \Delta)^2 \log \frac{D^4}{\Delta}. \]

6 The second equality is justified because \( K \geq 4 \) and \( 1/2 \leq \Delta < 1 \), hence \( \frac{\log K}{\log \Delta} \geq 2 \).
Figure B.1. An illustration of the identity $\Pi A^{\ell/2}(\rho_L \otimes \rho_R) = \Pi (\rho_L \otimes \rho_R)$. The left, right and middle sets of projections in figure (b) are inside the invariant space of $\rho_L, \rho_R$ and $\Pi$, respectively.

To simplify this expression, note that for any integer $k \geq 1$, we can use only the inequalities for $\ell = k, 2k, 3k, \ldots$. This amounts to considering the original problem with $(D, \Delta)$ replaced by $(D^k, \Delta^k)$. Then choosing $k = \lceil \frac{1}{2} \log(\Delta + 1) \rceil$ ensures that $\frac{1}{4} \leq \Delta^k \leq \frac{1}{2}$. Then substituting the parameters $D^k$ for $D$ and $\Delta^k$ for $\Delta$ in the above expression yields

$$S \leq O(1) \cdot \left[ \log(\mu^{-1}) \log D^k + \log D^k + 1 \right]$$

$$\leq O(1) \cdot k \log(\mu^{-1}) \log D = O(1) \cdot \frac{\log(\mu^{-1})}{\log(\Delta^{-1})} \log D.$$  \hfill \Box

Appendix B. Proof of existence of distinguishing measurement, lemma 5.5

Proof of lemma 5.5. Let $Q = \{ Q_i : Q_i$ acts only on particles in the $2\ell$ interval$\}$. Let $\Pi$ be a projection onto the ground space of all the operators in $Q$. We will show that $\{ \Pi, 1 - \Pi \}$ is the desired distinguishing measurement.

Clearly, $\text{Tr}(\Pi \rho^{2\ell}) = 1$. We would now like to prove that $\text{Tr}(\Pi \rho_L^k \otimes \rho_R^k)$ is at most $2\Delta^{\ell/4}$.

We start by considering, instead of $\Pi$, the applications of the DL operator $A$.

We can write $\rho_L \otimes \rho_R$ as a convex combination of rank 1 density matrices of product states of the form $|\phi_1\rangle \otimes |\phi_2\rangle$. By assumption, the overlap with the ground state is $|\langle \phi_1 \otimes \phi_2 | \Omega \rangle| \leq \Delta^{\ell/8}$. Therefore, since we assume a unique ground state, $|\phi_1 \otimes \phi_2 = c|\Omega\rangle + (1 - c^2)^{1/2}|\Omega^\perp\rangle$, with $c \leq \Delta^{\ell/8}$ and $|\Omega^\perp\rangle$ perpendicular to the ground space. Then by the DL, applying $A$ for $\ell/2$ times, we obtain

$$\text{Tr}(A^{\ell/2}|\phi_1\rangle\langle \phi_1| \otimes |\phi_2\rangle\langle \phi_2|) = \| A^{\ell/4}|\phi_1\otimes \phi_2\| \leq c^2 + \Delta^{\ell/4} \leq 2\Delta^{\ell/4},$$

and this remains true when we take convex combinations:

$$\text{Tr}(A^{\ell/2}\rho_L \otimes \rho_R) \leq 2\Delta^{\ell/4}.$$  \footnote{Again, we can always do that because we assumed that $1/2 \leq \Delta < 1$.}
Thus $\text{Tr}(A^{\ell/2} \rho) = 1$, and $\text{Tr}(A^{\ell/2} \rho_L \otimes \rho_R) \leq 2\Delta^{\ell/4}$; this establishes that applying $A$ for $\ell/2$ times distinguishes between $\rho$ and $\rho_L \otimes \rho_R$ with the desired probability. However, we would like the measurement to be confined to a short interval. Since we know that $\text{Tr}(\Pi \rho^{2\ell}) = \text{Tr}(\Pi \rho) = 1$, the proof will follow from showing that

$$\text{Tr}(\Pi \rho_L^{2\ell} \otimes \rho_R) = \text{Tr}(\Pi \rho_L \otimes \rho_R) \leq \text{Tr}(A^{\ell/2}(\rho_L \otimes \rho_R)).$$

The first equality holds since $\Pi$ only acts on the $2\ell$ particles in $\rho_L^{2\ell} \otimes \rho_R^{2\ell}$ and the last inequality is trivial; it is the middle equality which uses the structure of $A^{\ell/2}$. Indeed, let us write $A^{\ell/2} = A_M A_L A_R$ where

$$A_M = \prod_{(\text{groups})} (P_{k-2} P_k P_{k+2})(P_{k-1} P_{k+1}) P_k$$

is a ‘pyramid’ of terms centered at the cut $k$, and $A_L$ and $A_R$ are the terms to the left and right of the pyramid, respectively, as in figure B.1. Then $\Pi A^{\ell/2}(\rho_L \otimes \rho_R) = \Pi A_M A_L A_R(\rho_L \otimes \rho_R)$. But since we applied $A$ for exactly $\ell/2$ times, every $P_i$ projection in $A_M$ is also in the $2\ell$ window of $\Pi$. Therefore $A_M P_i = A_M$ and consequently $\Pi A_M = \Pi$. Similarly, $A_L A_R(\rho_L \otimes \rho_R) = \rho_L \otimes \rho_R$, and therefore $\Pi A^{\ell/2}(\rho_L \otimes \rho_R) = (\rho_L \otimes \rho_R)$, implying the desired equality. \hfill \square

Appendix C. Proof of information theoretical bound, lemma 5.6

Proof of lemma 5.6. Let $X$ and $Y$ be $\{0, 1\}$ random variables that result from applying the measurement $\Pi$ on $\rho$ and $\rho_L \otimes \rho_R$, respectively. Then by the Lindblad–Uhlmann theorem [28, 29],

$$S(\rho_L^\ell) + S(\rho_R^\ell) - S(\rho^{2\ell}) = S(\rho^{2\ell} || \rho_L^\ell \otimes \rho_R^\ell) \geq S(X||Y) = \sum_{i \in \{0, 1\}} x_i \log \frac{x_i}{y_i}.$$  

In this case, we have $X = 1$ with probability $1$ and $Y$ is $1$ with probability $\alpha \leq 2\Delta^{\ell/4}$. Thus, using straightforward analysis, $\sum_i x_i \log \frac{x_i}{y_i} = \log(\frac{x}{y}) \geq \log(1/2) + \frac{\epsilon}{4} \log(\Delta^{-1})$. But by (14), as long as $\epsilon \leq 3$ (which is our working assumption), $\log(\Delta^{-1}) \geq \epsilon/5$, and so $\sum_i x_i \log \frac{x_i}{y_i} \geq \frac{\epsilon}{20} \ell - 1$, and the result follows. \hfill \square

Appendix D. Proof that close cuts behave similarly: claim 5.7

Proof of Claim 5.7. Assume for contradiction that $|\chi_1 \otimes \chi_2\rangle$ is a product state across the cut $(k + j, k + j + 1)$ with $j > 0$ such that $|\langle \chi_1 \otimes \chi_2 | \Omega \rangle| > \mu d^j$. Schmidt-decompose $|\chi_1\rangle = \sum_{i=1}^{d^k} \alpha_i |L_i\rangle \otimes |R_i\rangle$ where the cut is between the first $k$ particles and the $j$ particles between $k + 1$ and $k + j$. By simple algebra, there exists at least one $i$ such that $|\langle L_i \otimes R_i \otimes \chi_2 | \Omega \rangle| \geq \mu$ which violates the hypothesis. \hfill \square

References

[1] Aharonov D, Arad I, Landau Z and Vazirani U 2009 The detectability lemma and quantum gap amplification STOC ’09: Proc. 41st Annu. ACM Symp. on Theory of Computing (New York) (New York: ACM) pp 417–26
[2] Hastings M B 2004 Lieb–Schultz–Mattis in higher dimensions Phys. Rev. B 69 104431
[3] Hastings M 2007 An area law for one dimensional quantum systems J. Stat. Mech. (2007) P08024

New Journal of Physics 13 (2011) 113043 (http://www.njp.org/)
Arad I, Landau Z and Vazirani U 2011 An improved 1D area law for frustration-free systems Phys. Rev. B submitted (arXiv:1111.2970)

Aharonov D, Arad I, Landau Z and Vazirani U 2011 The 1D area law and the complexity of quantum states: a combinatorial approach FOCS ’11: Proc. 52nd Annu. IEEE Symp. on Foundations of Computer Science (FOCS 2011)

Östlund S and Rommer S 1995 Thermodynamic limit of density matrix renormalization Phys. Rev. Lett. 75 3537–40

Rommer S and Östlund S 1997 Class of ansatz wave functions for one-dimensional spin systems and their relation to the density matrix renormalization group Phys. Rev. B 55 2164–81

Vidal G 2004 Efficient simulation of one-dimensional quantum many-body systems Phys. Rev. Lett. 93 040502

Verstraete F, Porras D and Cirac J I 2004 Density matrix renormalization group and periodic boundary conditions: a quantum information perspective Phys. Rev. Lett. 93 227205

Verstraete F and Cirac J I 2004 Renormalization algorithms for quantum-many body systems in two and higher dimensions arXiv:cond-mat/0407066

Shi Y, Duan L and Vidal G 2006 Classical simulation of quantum many-body systems with a tree tensor network Phys. Rev. A 74 022320

Vidal G 2007 Entanglement renormalization Phys. Rev. Lett. 99 220405

Lieb E and Robinson D 1972 The finite group velocity of quantum spin systems Commun. Math. Phys. 28 251–7

Dinur I 2007 The pcp theorem by gap amplification J. ACM 54 12

Hastings M B 2007 Entropy and entanglement in quantum ground states Phys. Rev. B 76 035114

Aharonov D, Arad I, Landau Z and Vazirani U 2010 Quantum Hamiltonian complexity and the detectability lemma arXiv:1011.3445

Gottesman D and Hastings M B 2010 Entanglement versus gap for one-dimensional spin systems New J. Phys. 12 025002

Iziri S 2010 Ground state entanglement in one-dimensional translationally invariant quantum systems J. Math. Phys. 51 022101

Iziri S 2011 private communication

Affleck I, Kennedy T, Lieb E and Tasaki H 1987 Rigorous results on valence-bond ground states in antiferromagnets Phys. Rev. Lett. 59 799–802

Kitaev A Y 2003 Fault-tolerant quantum computation by anyons Ann. Phys. 303 2–30

von Neumann J 1950 The Geometry of Orthogonal Spaces vol 2 (Princeton, NJ: Princeton University Press)

Badea C, Grivaux S and Müller V 2010 A generalization of the Friedrichs angle and the method of alternating projections C. R. Math. 348 53–6

Verstraete F, Wolf M and Cirac J 2009 Quantum computation and quantum-state engineering driven by dissipation Nat. Phys. 5 633–6

Farhi E, Gosset D, Hassidim A, Lutomirski A, Nagaj D and Shor P 2010 Quantum state restoration and single-copy tomography for ground states of Hamiltonians Phys. Rev. Lett. 105 190503

Eckart C and Young G 1936 The approximation of one matrix by another of lower rank Psychometrika 1 211–8

Hastings M B 2010 Locality in quantum systems arXiv:1008.5137

Lindblad G 1975 Completely positive maps and entropy inequalities Commun. Math. Phys. 40 147–51

Uhlmann A 1977 Relative entropy and the Wigner–Yanase-Dyson–Lieb concavity in an interpolation theory Commun. Math. Phys. 54 21–32