Entropy and entanglement in quantum ground states

M. B. Hastings
Center for Nonlinear Studies and Theoretical Division, Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA
(Received 8 January 2007; revised manuscript received 30 May 2007; published 19 July 2007)

We consider the relationship between correlations and entanglement in gapped quantum systems, with application to matrix product state representations. We prove that there exist gapped one-dimensional local Hamiltonians such that the entropy is exponentially large in the correlation length, and we present strong evidence supporting a conjecture that there exist such systems with arbitrarily large entropy. However, we then show, under an assumption on the density of states which is believed to be satisfied by many physical systems such as the fractional quantum Hall effect, that an efficient matrix product state representation of the ground state exists in any dimension. Finally, we comment on the implications for numerical simulation.

DOI: 10.1103/PhysRevB.76.035114 PACS number(s): 03.65.Ud, 05.30.−d, 05.10.Cc

I. INTRODUCTION

Finding the ground state of a local quantum Hamiltonian is one of the basic problems in physics. Efficient numerical techniques, however, only exist for certain special cases. Quantum Monte Carlo, for example, is most effective for systems which lack a sign problem. Fortunately, in one dimension, the extremely powerful density matrix renormalization group[1] algorithm is available, at least for systems with an excitation gap. This algorithm is based ultimately on knowledge of the structure of the ground state: it is not an arbitrary wave function, but rather has the special form of a matrix-product state.[2] Recently, a very promising algorithm capable of finding higher-dimensional matrix-product states to approximate the ground state of fairly arbitrary Hamiltonians has been developed,[3] and appears to offer, at least for certain systems, the possibility of studying systems that cannot be attacked by any other numerical technique.

All this work raises the following important question: Are ground states of such gapped, local quantum systems, indeed, close to matrix product states? A basic consequence of such a description is an *area law*: the reduced density matrix of the ground state wave function on some subvolume of the entire system has an entropy that is bounded by some constant times the surface area of that volume, while critical systems in one dimension may exhibit logarithmic corrections.[4] One advance[5,6] was to show that for all such Hamiltonians, the Hamiltonian could be written as a sum of local terms such that the ground state wave function was close to an eigenvector of each term separately, allowing one to use certain generalized matrix-product states. Unfortunately, these states required a hierarchical construction, joining blocks of the system together at successively longer length scales, and hence, do *not* provide a construction of the desired local matrix-product states. Further, this work left open the question of the number of such states required in each block, precisely the question of whether an area law holds or not.

Let X,Y be sets of sites in the system. An area law would follow from the assumption that the reduced ground state density matrix of the system $\rho_{X\cup Y}$ on the set of sites $X\cup Y$ approximately factorizes into a product of density matrices, $\rho_X\otimes\rho_Y$, if X,Y are separated by some distance from each other, as discussed in the Appendix. As a partial result to showing this factorization, it was shown that a gapped system with a local Hamiltonian has exponentially decaying correlations,[7,8] and so for any operators O_X,O_Y with support on X,Y we have $\mathrm{Tr}(O_XO_Y\rho_{X\cup Y})\approx\mathrm{Tr}(O_XO_Y\rho_X\otimes\rho_Y)$, up to exponentially small corrections in the distance between X and Y. However, the exponential decay of correlation functions is not directly useful in showing the factorization of density matrices, as there are so-called data hiding states on bipartite systems for which correlations are very small despite a high degree of entanglement.[9,10]

In this paper, we provide two results relating entanglement entropy, correlations, and a spectral gap. The first result is, in a sense, a negative result, showing that even in a system with short-range correlations, the entanglement entropy may be large. The second result is a positive result, giving a condition obeyed by many systems which guarantees that the ground state can be approximately represented in a matrix-product form, which is in many ways stronger than simply having an area law especially in higher dimensions.

To show the first result, we provide two examples of systems with large entropy compared to the correlation length. The first example proves that there exist one-dimensional systems for which the entropy is exponentially large in the correlation length. The second example provides strong evidence to support a conjecture that for any, arbitrarily large, S there exist one-dimensional systems with an entanglement entropy equal to S, and with a Hilbert space dimension on each site D equal to 3 or 4 and a correlation length bounded above by some S-independent constant of order unity. This result contradicts the naive expectation that a system with short-range correlations obeys an area law and, in particular, has an entanglement entropy proportional to the correlation length in one dimension. One handwaving argument for this naive expectation is that only the degrees of freedom close to the surface of the subvolume can be correlated with the degrees of freedom on the other side. However, the problem with this naive argument is that there can be truly many-body entanglement of a form that cannot be detected by measuring correlation functions of just a few degrees of freedom.

However, all these systems involve long-range interactions so that the entropy is still of order $\log(D)$ times the Lieb-Robinson group velocity[13] divided by the energy gap.

Thus, this result leaves open the question of whether the entropy obeys an area law in a system with short-range interactions. These results point out one important fact in any attempt to find an area law: one must consider more than two-point correlation functions, and must instead consider measures which are more sensitive to entanglement, such as the trace norm distance $\mathrm{Tr}(|\rho_{XY}-\rho_X\otimes\rho_Y|)$.

The second result shows that under a certain assumption on the density of states believed to be satisfied by many important physical systems such as the 1/3 fractional Hall effect (in this case, defined for a lattice system of electrons), we show that systems in arbitrary dimensions can be efficiently represented as higher-dimensional matrix-product states. This result is based on a recent result[5] that the density matrix at a nonzero temperature can be represented as a matrix-product density matrix.[11]

II. ONE-DIMENSIONAL MODEL SYSTEM AND THE EXPANDER GRAPH STATE

In this section, we construct the one-dimensional model systems described above. Following Refs. 2 and 12, we construct the system by first writing its ground state as a matrix product and then defining the Hamiltonian as a sum of projection operators. We write the ground state wave function for this V site system as

$$\Psi(s_1,s_2,\ldots,s_V)=\mathrm{Tr}(A(s_1)A(s_2)\cdots A(s_V)), \tag{1}$$

where $1\leqslant s_i\leqslant D$ is the state of the system on site i and where the $A(s_i)$ are $k\times k$ dimensional matrices, with k denoting the *dimension* of the matrix-product state. We normalize so that

$$\sum_{s=1}^{D}A(s)A^{\dagger}(s)=1. \tag{2}$$

This state (1) is the ground state of a Hamiltonian, $\mathcal{H}=\Sigma_i P_{i,i+1,\ldots,i+l}$, where $P_{i,i+1,\ldots,i+l}$ projects onto the set of states $|\Psi_{\alpha,\beta}\rangle$ on sites $i,i+1,\ldots,i+l$ defined by $\Psi_{\alpha,\beta}(s_i,s_{i+1},\ldots,s_l)=\Sigma_{\{\gamma\}}A(s_i)_{\alpha\gamma_1}A(s_{i+1})_{\gamma_1\gamma_2}\cdots A(s_{i+l})_{\gamma_l\beta}$, where $A(s_i)_{\alpha\gamma}$, with $1\leqslant\alpha,\gamma\leqslant k$, are matrix elements of $A(s_i)$. Here, the interaction length l is of order $\log_D(k)$,[2,12] and hence, the Lieb-Robinson[13] group velocity v of this system is of order l.

A sufficient condition[2,12] for this Hamiltonian to have a unique ground state with a gap and to have a finite correlation length is that the linear map from $k\times k$ matrices to $k\times k$ matrices,

$$\mathcal{E}(M)=\sum_{s=1}^{D}A(s)MA^{\dagger}(s), \tag{3}$$

have one nondegenerate eigenvalue equal to unity, and then have a gap to the next largest (in absolute value) eigenvalue. From Eq. (2), the eigenvector with unit eigenvalue is proportional to the unit matrix.

We now propose a specific choice of $A(s)$ such that $\mathcal{E}(M)$ has a gap in its spectrum and such that k is exponentially large in D. We pick

$$A(s)=\frac{1}{\sqrt{D}}U(s), \tag{4}$$

where $U(s)$ is a unitary matrix depending on s. Thus, Eq. (2) is automatically satisfied. We then pick $U(s)$ following two different rules, one rule for $1\leqslant s\leqslant D/2$ and one for $D/2<s\leqslant D$.

For $1\leqslant s\leqslant D/2$, we pick $U(s)$ to be a diagonal matrix with matrix elements

$$U(s)_{\alpha\beta}=\delta_{\alpha\beta}F(\alpha,s), \tag{5}$$

where $F(\alpha,s)$ is some function such that $F(\alpha,s)=\pm 1$ for all α,s. We pick $F(\alpha,s)$ such that for any two α,β with $\alpha\neq\beta$,

$$\left|\sum_{s=1}^{D/2}F(\alpha,s)F(\beta,s)\right|\leqslant D/4. \tag{6}$$

The following question arises: For which values of k and D is it possible to satisfy Eq. (6)? Define the vector v_α to be a vector in a $D/2$ dimensional vector space by $v_\alpha=[F(\alpha,1)/\sqrt{D/2},F(\alpha,2)/\sqrt{D/2},\ldots,F(\alpha,s)/\sqrt{D/2}]$. Then, to satisfy Eq. (6), we must find k vectors v_α in a $D/2$ dimensional vectors space such that the inner product between v_α and v_β is less than 1/2 for $\alpha\neq\beta$, and hence, the angle between the vectors v_α,v_β is greater than $\phi=\cos^{-1}(1/2)$. It is known[14,15] that for large D, it is possible to do this for $k\leqslant\exp(cD/2)$ such vectors for some constant $c>0$, and hence, k may be exponentially large in D.

For $D/2<s\leqslant D$, we pick $U(s)$ to be the matrix with matrix elements

$$U(s)_{\alpha\beta}=\frac{1}{k}\sum_{a=0}^{k-1}\exp\left[2\pi i a\frac{(\alpha-\beta)}{k}\right]\hat{F}(a,s), \tag{7}$$

and we pick $\hat{F}$ such that for any two a,b with $a\neq b$,

$$\left|\sum_{s=D/2+1}^{D}\hat{F}(a,s)\hat{F}(b,s)\right|\leqslant D/4. \tag{8}$$

Again, this is possible to do so as long as $k\leqslant\exp(cD/2)$.

Equation (6) implies that off-diagonal elements of M are reduced by the map $\mathcal{E}(M)$, while Eq. (6) implies that off-diagonal elements of M in the Fourier basis are also reduced by this map. We now use this idea to show that the map $\mathcal{E}(M)$ does, indeed, have a gap between the unit eigenvalue and the next largest eigenvalue. The completely positive map $\mathcal{E}(M)$ is a Hermitian linear operator from $k\times k$ matrices, to $k\times k$ matrices, and hence, the eigenvalue which is second largest in absolute value can be found by taking the maximum over all traceless matrices M with $\mathrm{Tr}(M^{\dagger}M)=1$ of $|\mathrm{Tr}(M^{\dagger}\mathcal{E}(M))|$. Let $\mathcal{E}_1(M)=\Sigma_{s=1}^{D/2}A(s)MA^{\dagger}(s)$ and let $\mathcal{E}_2(M)=\Sigma_{s=D/2+1}^{D}A(s)MA^{\dagger}(s)$ so that $\mathcal{E}(M)=\mathcal{E}_1(M)+\mathcal{E}_2(M)$. Let M_d denote the diagonal part of such a matrix M and let $P_d=\mathrm{Tr}(M_d^{\dagger}M_d)$. Then, from Eq. (6), we have $|\mathrm{Tr}(M^{\dagger}\mathcal{E}_1(M))|\leqslant P_d/2+(1-P_d)/4$. Let M_f be the matrix with elements $M_f^{ab}=(1/k)\Sigma_{\alpha\beta}M_{\alpha\beta}\exp[-2\pi i(a\alpha-b\beta)/k]$ and let M_{df} de-

note the diagonal components of this matrix M_f with $P_{df}=\mathrm{Tr}(M_{df}^\dagger M_{df})$. Then, from Eq. (8), we have $|\mathrm{Tr}(M^\dagger \mathcal{E}_2(M))| \leq P_{df}/2+(1-P_{df})/4$.

We now rewrite P_d, P_{df} as sums of squares of traces of $k\times k$ matrices. Let $M_{d,x}$, for $x=1,\ldots,k-1$, be traceless, diagonal matrices such that

$$\mathrm{Tr}(M_{d,x}^\dagger M_{d,y}) = \delta_{x,y}. \tag{9}$$

Let $M_{df,x}$, for $x=1,\ldots,k$, be matrices with matrix elements $(M_{df,x})_{\alpha\beta}=(1/k)\exp[2\pi i x(\alpha-\beta)/k]$ so that

$$\mathrm{Tr}(M_{df,x}^\dagger M_{df,y}) = \delta_{x,y}. \tag{10}$$

Note that

$$\mathrm{Tr}(M_{d,x}^\dagger M_{df,y}) = 0. \tag{11}$$

Now, for a traceless matrix M, $P_d=\Sigma_{x=1}^{k-1}|\mathrm{Tr}(M_{d,x}^\dagger M)|^2$, while $P_{df}=\Sigma_{x=1}^{k}|\mathrm{Tr}(M_{df,x}^\dagger M)|^2$. Now consider the space of $k\times k$ matrices as a vector space, with an inner product given by the trace of two matrices; Eqs. (9)–(11) imply that the matrices $M_{d,x}, M_{df,y}$ form a set of orthonormal vectors in this vector space. While these are not a complete set of vectors, it still follows that $\mathrm{Tr}(M^\dagger M)\geq \Sigma_{x=1}^{k-1}|\mathrm{Tr}(M_{d,x}^\dagger M)|^2 + \Sigma_{x=1}^{k}|\mathrm{Tr}(M_{df,x}^\dagger M)|^2$, so that for a traceless matrix M with $\mathrm{Tr}(M^\dagger M)=1$, we have $P_{df}+P_d\leq 1$. Therefore, for any traceless matrix M, we have $|\mathrm{Tr}(M\mathcal{E}(M))|\leq 3/4$, showing the existence of a gap as claimed.

In the case of this construction, we showed the existence of a gap in the spectrum of $\mathcal{E}(M)$, which implies[2,12] that the correlation length is bounded by a constant, independent of D and k. Since the eigenvector of the map with unit eigenvalue is proportional to the identity matrix, $\mathbb{1}$, for a sufficiently long chain, the entanglement entropy between two halves of the chain is equal to $-\mathrm{Tr}((1/k)\mathbb{1})\log_e((1/k)\mathbb{1}) = \log_e(k)$. This construction allows us to take k exponentially large in D and still have a gap in the spectrum of $\mathcal{E}(M)$, as claimed, and hence, the wave function has short-range correlations with a correlation length of order unity. Alternately, one can represent each site on this system with D states by $\log_2(D)$ sites with 2 states on each site. In this case, the correlation length of the system is of order $\log_2(D)$, while the entropy is of order $\log_e(\exp(cD/2))=cD/2$, which is exponentially large in the correlation length as claimed. While this is surprising, since one might believe that the entropy should be proportional to the correlation length, it must be noted that this construction, as far as we know, does require a Hamiltonian with a long interaction range.

We now present an alternative construction which improves on the first construction, since it gives an arbitrarily large k for a *fixed* D, along with evidence that it also leads to a gap in $\mathcal{E}(M)$. Hence, this provides a system with a Hilbert space dimension on each site of order unity, a correlation length of order unity, and yet arbitrarily large entropy. We take $D=4$, and for $s=1,2$, we take the matrices $U(s)$ to be of the form

$$U(s) = P(s)\sigma(s), \tag{12}$$

where $\sigma(s)$ is a diagonal matrix with entries equal to ± 1, and $P(s)$ is a permutation matrix (that is, it has exactly one in each row and column, with all other entries equal to zero). We then take $U(3)=U(1)^\dagger$ and $U(4)=U(2)^\dagger$. We set $P(3)=P(1)^\dagger$ and $P(4)=P(2)^\dagger$. We now fix k at an arbitrary value, and then argue that by an appropriate choice of P and σ, it is possible to have a gap in $\mathcal{E}(M)$.

Let $P(s,\alpha)$ denote the value of j such that the matrix element $P(s)^{\alpha\beta}=1$. That is, $P(s,\alpha)$ is the result of applying the given permutation to α. The operator $\mathcal{E}(M)$ is block diagonal: by acting on a diagonal matrix, it produces a diagonal matrix, and by acting on an off-diagonal matrix, it produces an off-diagonal matrix. We study the spectrum of $\mathcal{E}(M)$ in each block separately, starting with the action on a diagonal matrix. Let $M(\alpha,\beta)$ be the matrix with matrix elements $M(\alpha,\beta)_{\sigma\tau}=\delta_{\alpha\sigma}\delta_{\beta\tau}$. We have

$$\mathcal{E}(M(\alpha,\alpha)) = \frac{1}{4}\sum_{s=1}^{4} M(P(s,\alpha),P(s,\alpha)). \tag{13}$$

Equation (13) can be represented by a diffusion process on a graph. The graph has vertices labeled by $\alpha=1,\ldots,k$, and with an undirected edge from vertex α to β if for some $s=1,2$ we have $P(s,\alpha)=\beta$ or $P(s,\beta)=\alpha$. This graph has fixed coordination number equal to 4, although some vertices may have more than one edge connecting them if for some $s\neq t$ we have $P(s,\alpha)=P(t,\alpha)$.[17] Then, Eq. (13) implies that in the first block, $\mathcal{E}(M)$ has the same spectrum as $1/4$ times the adjacency matrix of the given graph. However, it is known that it is possible to find graphs with k vertices and fixed coordination number q, for any $q>2$, such that $1/q$ times the adjacency matrix has a unit eigenvalue and then a gap to the next eigenvalue which is bounded below by some k-independent constant.[16] Indeed, it has been shown that the gap between the unit eigenvalue and the next eigenvalue of the adjacency matrix of a *random* graph generated using the above procedure with a random choice of permutations P is, with probability tending to unity as the size k of the graph tends to infinity, bounded below by some k-independent constant.[18] Graphs with such a gap in the adjacency matrix spectrum are called expander graphs, and we refer to the state with the appropriate choice of permutation matrices and diagonal matrices $\sigma(s)$ discussed below as the *expander graph state*.

Now we consider the spectrum of $\mathcal{E}(M)$ in the second block. We have

$$\begin{aligned}\mathcal{E}(M(\alpha,\beta)) = \frac{1}{4}\Bigg[&\sum_{s=1}^{2} M(P(s,\alpha),P(s,\beta))\sigma(s)_{\alpha\alpha}\sigma(s)_{\beta\beta} \\ &+ \sum_{s=3}^{4} M(P(s,\alpha),P(s,\beta))\sigma(s)_{P(s,\alpha)P(s,\alpha)} \\ &\times \sigma(s)_{P(s,\beta)P(s,\beta)}\Bigg].\end{aligned} \tag{14}$$

While it is possible to prove the existence of a gap in the spectrum of $\mathcal{E}(M)$ in the first block for suitable matrices P, we have to apply some physical intuition to show that the second block has all eigenvalues separated from unity by a

constant. If all the $\sigma(s)^{\alpha\alpha}$ were equal to +1, then the second block would have one unit eigenvalue, with eigenvector proportional to a matrix with all off-diagonal entries equal to unity and all diagonal entries equal to zero. In this case, Eq. (14) would describe two correlated random walks on the given graph, one for each index α and β. Then, for a random choice of the permutations P, we would expect that, for similar reasons to the existence of a gap in the diagonal sector, there would typically be a gap to the next eigenvalue in the off-diagonal sector as any correlations between α and β would be short lived under this random process. We instead randomly set each $\sigma(s)^{\alpha\alpha}$ equal to ± 1, independently for each α. In this case, we expect that random choices of the σ and P will, due to the random signs, cause all eigenvalues in the second block to be separated from unity by at least some k-independent gap with probability tending to unity as k tends to infinity.

We have tested this numerically by generating random permutations and random σ. We picked permutations with the additional restriction that each permutation had exactly one cycle of length k. We performed tests on systems with k up to 50 and found that there was always one unit eigenvalue and a gap in the rest of the spectrum, with the most negative eigenvalue separated by a gap from -1, and further, that the spectrum away from the unit eigenvalue exhibited a scaling collapse, such that the density of eigenvalues appeared to be roughly equal to k^2 times some k-independent function.

In the above, we considered $D=4$ so that we could take the matrices $U(3)=U(1)^\dagger$ and $U(4)=U(2)^\dagger$ and arrive at a real spectrum for $\mathcal{E}(M)$. If we take $D=3$, and choose $U(s)=P(s)\sigma(s)$ for $s=1,2,3$ with $P(1)$, $P(2)$, and $P(3)$ random permutations and $\sigma(1)$, $\sigma(2)$, and $\sigma(3)$ random diagonal matrices with entries ± 1, then for the first block of $\mathcal{E}(M)$ we arrive at the adjacency matrix of a *directed* graph of degree 3. In this case, $\mathcal{E}(M)$ may have a complex spectrum. However, we still expect there to be a gap.

III. MATRIX-PRODUCT STATES FROM THERMAL DENSITY MATRICES

In this section, we build on the result in Ref. 5 that it is possible for local Hamiltonians to approximate the thermal density matrix by a matrix-product operator to show, subject to an assumption on the density of states, that the ground state is close to a matrix-product state. We start by considering the case of a unique ground state and a gap ΔE to the next excited state, and then generalize to multiple ground states. The idea is as follows: we approximate the thermal density matrix, $Z^{-1}\exp[-\beta H]$ for a Hamiltonian H, with $Z=\mathrm{Tr}(\exp[-\beta H])$, by a matrix-product density operator up to some small error in *trace norm*. For large enough β, the thermal density matrix becomes a good approximation in *operator norm* to $\rho_0=|\Psi_0\rangle\langle\Psi_0|$, the projector onto the ground state Ψ_0 of the given Hamiltonian. We then make an assumption on the number of low energy states of the Hamiltonian, which then allows to show that the thermal density matrix is a good approximation in trace norm to the projector, and hence, that our matrix-product operator is a good approximation in trace norm to the projector. This means that, in any complete orthonormal basis, there must exist some state such that the matrix-product operator acting on the basis is close to the ground state; picking this basis to be a factorized basis gives us a matrix-product state which is close to the ground state.

We consider a regular lattice (this condition can be weakened to include other lattices) of V sites labeled 1,2,…, V, with a metric $d(i,j)$ between sites i and j, and we consider a Hamiltonian H which is a sum of terms H_i supported on the sites within distance R of i for some range R and with operator norm $\|H_i\|$ bounded by some constant J. Then, it was shown that for any ϵ and for any regular d-dimensional lattice of V sites, there exists a matrix-product density operator approximation to the thermal density matrix of the form

$$\rho(\beta,l_{proj})=\sum_{\{\alpha_k\}}\rho_1(\alpha_1)\rho_2(\alpha_2)\cdots\rho_V(\alpha_V)\times F_1(\{\alpha_j\})F_2(\{\alpha_j\})\cdots F_V(\{\alpha_j\}), \tag{15}$$

where $\mathrm{Tr}(\rho(\beta,l_{proj}))=1$, where each operator $\rho_i(\alpha_i)$ acts only on site i, where α_j are some set of indices, with range $1\leq\alpha_j\leq\alpha_{max}$ for all j, and where each function F_i depends only on α_j with $d(i,j)\leq l_{proj}$, and such that

$$\mathrm{Tr}(|\rho(\beta,l_{proj})-Z^{-1}\exp[-\beta H]|)\leq\epsilon. \tag{16}$$

$\mathrm{Tr}(|\cdots|)$ denotes the trace norm and

$$l_{proj}\sim R\log(V\beta/J\epsilon), \tag{17}$$

with the constant of proportionality depending on the exact lattice structure, and with

$$\alpha_{max}\sim D^{l_{proj}^d\beta/J}. \tag{18}$$

Now, we want to introduce a condition on the density of states of a given Hamiltonian that will lead to a bound on how accurately the thermal density matrix approximates the ground state projection operator. To motivate this bound, suppose that we have a Hamiltonian of the following particularly simple form: $H=\Delta E\Sigma_i O_i$, where O_i is an operator on site i with one zero eigenvalue and $D-1$ eigenvalues equal to unity. Then, there exists one state with all spins down and with energy 0. There exist $(D-1)V$ states, each with one spin up and with energy ΔE. There exist $(D-1)^2V(V-1)/2$ states with energy $2\Delta E$, and so on. Thus, we have the bound that $\rho(m)$, which we define to be the number of states with energy E_i with $m\Delta E\leq E_i<(m+1)\Delta E$, obeys

$$\rho(m)\leq(cV)^m/m!, \tag{19}$$

for $c=D-1$.

In general, however, many gapped systems will obey assumption (19) for some constant c. All free fermion systems with a gap in the single particle spectrum will obey this assumption. Physically, we expect that a lattice realization of the fractional Hall effect at an incompressible filling fraction will also obey this. In the rest of this section, we show that if a system does obey this assumption, then it is possible to write the ground state as a higher-dimensional matrix-

product state, providing bounds on the error as a function of the dimension of the matrix-product state. Using this assumption,

$$\mathrm{Tr}(|Z^{-1}\exp[-\beta H]-\rho_0|) \leq 2\sum_{m=1}^{\infty}\frac{(\exp[-\beta\Delta E]cV)^m}{m!} = 2(\exp\{\exp[-\beta\Delta E]cV\}-1), \tag{20}$$

and hence, for any ϵ, for $\beta \geq \log(cV/\log(1+\epsilon/2))/\Delta E$, we have

$$\mathrm{Tr}(|Z^{-1}\exp[-\beta H]-\rho_0|) \leq \epsilon. \tag{21}$$

The operator $\rho(\beta, l_{proj})$ can be written in the form $\rho(\beta, l_{proj})=O^\dagger O$ for some operator O. We write an orthonormal basis of states for the system by vectors $|\Psi_{\vec{v}}\rangle$, where $\vec{v}=(v_1, v_2, \ldots, v_V)$ with $1 \leq v_i \leq D$. The vector $|\Psi_{\vec{v}}\rangle$ denotes the state where site i is in state v_i. This is a factorized basis of states for the system. Then, $\rho(\beta, l_{proj})=\Sigma_{\{\vec{v}\}}O^\dagger|\Psi_{\vec{v}}\rangle\langle\Psi_{\vec{v}}|O$. Let

$$p^0_{\vec{v}}=\langle\Psi_{\vec{v}}|O\rho_0 O^\dagger|\Psi_{\vec{v}}\rangle \tag{22}$$

and

$$p^{>}_{\vec{v}}=\langle\Psi_{\vec{v}}|OO^\dagger|\Psi_{\vec{v}}\rangle - p^0_{\vec{v}}. \tag{23}$$

Thus, $p^0_{\vec{v}}$ is equal to $|O^\dagger\Psi_{\vec{v}}|^2$ times the probability that the state $O^\dagger\Psi_{\vec{v}}/|O^\dagger\Psi_{\vec{v}}|$ is in the ground state Ψ_0, with $p^{>}_{\vec{v}}$ equal to $|O^\dagger\Psi_{\vec{v}}|^2$ times the probability that the state is not in the ground state. Then $\Sigma_{\vec{v}}(p^0_{\vec{v}}+p^{>}_{\vec{v}})=1$ and $\epsilon \geq \mathrm{Tr}(|\rho-\rho_0|) \geq (1-\Sigma_{\vec{v}}p^0_{\vec{v}})+\Sigma_{\vec{v}}p^{>}_{\vec{v}}$, so $\Sigma_{\vec{v}}p^{>}_{\vec{v}} \leq \epsilon/2$. Hence,

$$\frac{\sum_{\vec{v}} p^{>}_{\vec{v}}}{\sum_{\vec{v}} p^0_{\vec{v}}} \leq \frac{\epsilon/2}{1-\epsilon/2}. \tag{24}$$

Hence, there must exist some $\vec{w}$ such that

$$\frac{p^{>}_{\vec{w}}}{p^0_{\vec{w}}} \leq \frac{\epsilon/2}{1-\epsilon/2}. \tag{25}$$

Then,

$$|\Psi_{mps}\rangle \equiv \frac{1}{\sqrt{p^0_{\vec{w}}+p^{>}_{\vec{w}}}}O^\dagger|\Psi_{\vec{w}}\rangle \tag{26}$$

is a normalized matrix-product state by the assumption that O is a matrix-product operator. Further,

$$|\langle\Psi_0|\Psi_{mps}\rangle|^2 = \frac{p^0_{\vec{w}}}{p^0_{\vec{w}}+p^{>}_{\vec{w}}} \geq \frac{1}{[1+(\epsilon/2)]/[1-(\epsilon/2)]} = 1-\frac{\epsilon}{2}+\mathcal{O}(\epsilon^2), \tag{27}$$

so that Ψ_{mps} is close to the state Ψ_0.

Thus, for any ϵ, we can find a matrix-product state $|\Psi_{mps}\rangle$ such that Eq. (27) holds and such that l_{proj} grows logarithmically in V, with α_{max} growing as an exponential of $\log(V)^{d+1}$. This is a matrix-product state with bond variables that connect sites separated by a distance up to l_{proj}. By introducing auxiliary variables, this can be written as a state with bond variables that connect nearest neighbor sites only, with a bond variable scaling as $\alpha_{max}^{l_{proj}^d}$, which is exponentially large in $\log(V)^{2d+1}$. Although this is a rather high power of $\log(V)$, it is still significantly better than the number of bond variables required to represent an arbitrary state, which scales exponentially in V. Further, we have imposed a very strict requirement on our matrix-product state: it must be able to approximate Ψ_0 with an overlap of order unity. In many applications, one is interested only in approximating local observables of Ψ_0 such as the energy, and hence, a much smaller number of bond variables may suffice.

This derivation can be readily generalized to the case of multiple ground states, labeled $|\Psi_0^{(a)}\rangle$, $a=0, \ldots, n-1$. Let $\rho_0=n^{-1}\Sigma_{a=1}^n|\Psi_0^a\rangle\langle\Psi_0^a|$, and assume that there exists a $\rho=O^\dagger O$ with $\mathrm{Tr}(\rho)=1$ and $\mathrm{Tr}(|\rho-\rho_0|) \leq \epsilon$. Then, define $p^0_{\vec{v}}$ as above and again find an appropriate $|\Psi_{\vec{w}_0}\rangle$, so that $|\Psi_{mps,0}\rangle=(1/\sqrt{p^0_{\vec{w}_0}+p^{>}_{\vec{w}_0}})O^\dagger|\Psi_{\vec{w}_0}\rangle$ is close to some state $\Psi_0^{(0)}$ in the ground state subspace. Then, define $\rho_1=\rho_0-n^{-1}|\Psi_0^{(0)}\rangle\langle\Psi_0^{(0)}|$ and define $p^1_{\vec{v}}=\langle\Psi_{\vec{v}}|O\rho_1 O^\dagger|\Psi_{\vec{v}}\rangle$. One can then find a state $|\Psi_{\vec{w}_1}\rangle$ such that $|\Psi_{mps,1}\rangle=(1/\sqrt{p^1_{\vec{w}_1}+p^{>}_{\vec{w}_1}})O^\dagger|\Psi_{\vec{w}_1}\rangle$ is close to some state $|\Psi_0^{(1)}\rangle$ in the ground state subspace, with $\langle\Psi_0^{(1)}|\Psi_0^{(0)}\rangle=0$. Repeating this n times, one can then approximate a set of states that span the ground state subspace with a set of matrix product states, $|\Psi_{mps,a}\rangle$, $a=0, \ldots, n-1$.

We note that in the procedure discussed in this section, the thermal density matrix is introduced simply as a means of approximating the projection operator by a local operator. Other approximate projection operators could have been used, such as

$$\frac{1}{\sqrt{2\pi}t_q}\int_{-\infty}^{\infty}dt\,\exp[iHt]\exp\left[-\frac{(t/t_q)^2}{2}\right], \tag{28}$$

for some t_q, where the approximation becomes more accuracte, but also less local, as t_q increases. In this case, given the entropy assumption (19), we would need to take t_q of order $\log(V)/\Delta E$ to ensure that the approximate projection operator was close in trace norm to the exact projection operator.

IV. DISCUSSION

We have shown that it is possible to find systems for which the entanglement entropy between a given subvolume and the rest of the system is much larger than expected from the correlation length. However, we have also shown that, subject to assumption (19) which is satisfied by many physical systems, matrix products give a good representation of the ground state even in higher-dimensional systems, providing additional reason to consider the use of these states as a numerical technique.

Equation (19) has an interesting physical interpretation. If there is a gap and this equation is satisfied, then the system is not "glassy" in that at temperatures of order $\Delta E/\log(V)$ the

system is in its ground state with probability of order unity. Consider, however, the following system, related to an example of Terhal and DiVincenzo: a one-dimensional system with $D=3$, represented by a spin-1 degree of freedom on each of V different sites. The Hamiltonian is $-\Sigma_i[(S_i^z)^2-1/2][(S_{i+1}^z)^2-1/2]+(1/V)\Sigma_i(S_i^z)^2$. The ground state has spins with $S^z=0$ and energy equal to $-V/4$. However, there are 2^V states, with each spin having $S^z=\pm 1$ and with energy $-V/4+1$. Thus, this system does not satisfy Eq. (19) for any fixed, V-independent c. However, the related Hamiltonian $-\Sigma_i[(S_i^z)^2-1/2][(S_{i+1}^z)^2-1/2]+\Sigma_i(S_i^z)^2$ has the same ground state and does satisfy Eq. (19). In general, for any classical system, meaning that the Hamiltonian is a sum of operators which are diagonal in some factorized basis, the ground state will trivially be a matrix-product state. Thus, we conjecture that for any gapped local Hamiltonian, it is possible to find a related Hamiltonian close to the same ground state and which satisfies Eq. (19), and hence, has an approximate matrix-product ground state.

ACKNOWLEDGMENTS

I thank F. Verstraete for useful discussions. This work was carried out under the auspices of the NNSA of the U.S. DOE at LANL under Contract No. DE-AC52-06NA25396.

APPENDIX: DENSITY MATRIX FACTORIZATION AND AN AREA LAW

The goal in this section is to relate the factorization of the density matrix to the entropy of the reduced density matrix. Suppose X is an arbitrary set and let Y be the set of all sites i such that $\mathrm{dist}(X,i)\geq l$ for some l. Let B denote the set of sites j which are neither in X nor in Y. In this section, we assume that for some state Ψ_0 there is a bound

$$\mathrm{Tr}(|\rho_{XY}-\rho_X\otimes\rho_Y|)\leq\epsilon \tag{A1}$$

for some ϵ for the given X,Y and derive a bound on the entropy of the reduced density matrix ρ_X. The case that $\epsilon=0$ was considered is in Ref. 9, where it was shown that this implies that the density matrix ρ_X has at most $D^{|B|}$ nonzero eigenvalues. In this appendix, we consider the case of a nonzero ϵ.

The wave function Ψ_0 can be written as

$$|\Psi_0\rangle=\sum_\alpha A(\alpha)|\Psi_B^\alpha\rangle\otimes|\Psi_{XY}^\alpha\rangle, \tag{A2}$$

where $|\Psi_{XY}\rangle$ is a wave function on $X\cup Y$ and $|\Psi_B\rangle$ is a wave function on the set B, where B is the set of all sites i such that $i\notin X$ and $i\notin Y$, and where $\langle\Psi_B^\alpha|\Psi_B^\beta\rangle=\langle\Psi_{XY}^\alpha|\Psi_{XY}^\beta\rangle=\delta_{\alpha,\beta}$. Then,

$$\rho_{XY}=\sum_\alpha|A(\alpha)|^2|\Psi_{XY}^\alpha\rangle\langle\Psi_{XY}|. \tag{A3}$$

Then, ρ_{XY} is equal to the weighted sum of at most $D^{|B|}$ different density matrices $|\Psi_{XY}^\alpha\rangle\langle\Psi_{XY}|$, each of which corresponds to a pure state. Define P to be the projection operator onto the space spanned by these $D^{|B|}$ different states Ψ_{XY}^α. Then, from Eq. (A1), we have $\mathrm{Tr}((\rho_{XY}-\rho_X\otimes\rho_Y)P)\leq\epsilon$ so, therefore, $\mathrm{Tr}(P\rho_X\otimes\rho_Y)\geq 1-\epsilon$.

The operator ρ_X may be diagonalized, so that

$$\rho_X=\sum_\alpha\rho_X(\alpha)|\Psi_X^\alpha\rangle\langle\Psi_X^\alpha|, \tag{A4}$$

where $\langle\Psi_X^\alpha|\Psi_X^\beta\rangle=\delta_{\alpha,\beta}$. The index α ranges from 1 to at most $D^{|X|}$. Let us order the eigenvalues, so that $\rho_X(\alpha)\geq\rho_X(\beta)$ if $\alpha>\beta$. Similarly, we write $\rho_Y=\Sigma_\alpha\rho_Y(\alpha)=\Sigma_\alpha|\Psi_Y^\alpha\rangle\langle\Psi_Y^\alpha|$ and again order the eigenvalues so that $\rho_Y(\alpha)\geq\rho_Y(\beta)$ if $\alpha>\beta$. Finally, we denote the eigenvalues of $\rho_X\otimes\rho_Y$ by $\rho(\gamma)$ and also order these eigenvalues. Note that for each γ, $\rho(\gamma)=\rho(\alpha)\rho(\beta)$ for some α,β with $\alpha\leq\gamma$. Since P projects onto a space with dimension at most $D^{|B|}$, we have

$$1-\epsilon\leq\mathrm{Tr}(P\rho_X\otimes\rho_Y)\leq\sum_{\gamma=1}^{D^{|B|}}\rho(\gamma)\leq\sum_{\alpha=1}^{D^{|B|}}\rho_X(\alpha). \tag{A5}$$

Therefore,

$$\sum_{\alpha=D^{|B|}+1}^{\alpha=D^{|X|}}\rho_X(\alpha)\leq\epsilon. \tag{A6}$$

Equation (A6) means that the total probability of the system being in any state on set X *other* than some given set of $D^{|B|}$ states is bounded by ϵ. This is very close to a result for the entropy. Indeed, ρ_X has at most $D^{|X|}$ nonzero eigenvalues. Thus, the entropy is bounded by

$$S(\rho_X)\equiv-\sum_\alpha\rho_X(\alpha)\log(\rho_\alpha)=-\sum_{\alpha=1}^{D^{|B|}}\rho_X(\alpha)\log(\rho_\alpha)-\sum_{\alpha=D^{|B|}+1}^{D^{|X|}}\rho_X(\alpha)\log(\rho_\alpha)\leq|B|\log(D)+\epsilon|X|(\log(D)+\log(\epsilon)). \tag{A7}$$

[1] S. R. White, Phys. Rev. Lett. **69**, 2863 (1992).
[2] M. Fannes, B. Nachtergaele, and R. F. Werner, Commun. Math. Phys. **144**, 443 (1992).
[3] F. Verstraete and J. I. Cirac, arXiv:cond-mat/0407066.
[4] G. Vidal, J. I. Latorre, E. Rico, and A. Kitaev, Phys. Rev. Lett. **90**, 227902 (2003); F. Verstraete, M. M. Wolf, D. Perez-Garcia, and J. I. Cirac, *ibid.* **96**, 220601 (2006).
[5] M. B. Hastings, Phys. Rev. B **73**, 085115 (2006).
[6] A. Kitaev, Ann. Phys. (N.Y.) **321**, 2 (2006).
[7] M. B. Hastings, Phys. Rev. B **69**, 104431 (2004).

[8] M. B. Hastings, Phys. Rev. Lett. **93**, 140402 (2004).
[9] F. Verstraete and J. I. Cirac, Phys. Rev. B **73**, 094423 (2006).
[10] P. Hayden, D. Leung, P. W. Shor, and A. Winter, Commun. Math. Phys. **250**, 371 (2004).
[11] The operators we will consider are related to those in F. Verstraete, J. J. García-Ripoll, and J. I. Cirac, Phys. Rev. Lett. **93**, 207204 (2004), although we consider them in arbitrary dimension, and hence, rather than a matrix product, they are represented by a sum over bond variables.
[12] D. Pérez-García, F. Verstraete, M. M. Wolf, and J. I. Cirac, Quantum Inf. Comput. **7**, 401 (2007).
[13] E. H. Lieb and D. W. Robinson, Commun. Math. Phys. **28**, 251 (1972); M. B. Hastings and T. Koma, *ibid.* **265**, 781 (2006); B. Nachtergaele and R. Sims, *ibid.* **265**, 119 (2006).
[14] J. H. Conway and N. J. A. Sloane, *Sphere Packing, Lattices and Groups* (Springer-Verlag, New York, 1988).
[15] C. E. Shannon, Bell Syst. Tech. J. **38**, 611 (1959).
[16] N. Alon, Combinatorica **6**, 83 (1986).
[17] With the random choice of permutations below, typically only $O(1)$ nodes will have a coordination number different from 4.
[18] J. Friedman, Conference Proceedings of the Annual ACM Symposium on Theory of Computing, 2003, p. 720; arXiv:cs.DM/0405020.