Entanglement, correlations, and the energy gap in many-body quantum systems

Henry L. Haselgrove,[1,2,3,*] Michael A. Nielsen,[1,4,2,†] and Tobias J. Osborne[5,‡]
[1]*School of Physical Sciences, The University of Queensland, Brisbane 4072, Australia*
[2]*Institute for Quantum Information, California Institute of Technology, Pasadena, California 91125, USA*
[3]*Information Sciences Laboratory, Defence Science and Technology Organisation, Edinburgh 5111, Australia*
[4]*School of Information Technology and Electrical Engineering, The University of Queensland, Brisbane 4072, Australia*
[5]*School of Mathematics, University of Bristol, University Walk, Bristol BS8 1TW, United Kingdom*
(Received 4 September 2003; published 9 March 2004)

What correlations are present in the ground state of a many-body Hamiltonian? We study the relationship between ground-state correlations, especially entanglement, and the *energy gap* between the ground and first excited states. We prove several general inequalities which show quantitatively that ground-state correlations between systems not directly coupled by the Hamiltonian necessarily imply a small energy gap.

DOI: 10.1103/PhysRevA.69.032303 PACS number(s): 03.67.Mn

I. INTRODUCTION

A central problem in physics is characterizing the ground state of a many-body Hamiltonian. Of particular interest is the problem of understanding the correlations in the ground states of such systems. As an outgrowth of that interest, there has recently been considerable work on understanding the *nonclassical* correlations in the ground state, that is, the *ground-state entanglement*. Some recent work on this problem, with further references, includes [1–13] (see also [14–17], and references therein).

The purpose of the present paper is to prove some general inequalities relating the ground-state correlations and entanglement to the spectrum of the system Hamiltonian. We will prove that the existence of long-range correlations in the ground state implies a small energy gap between the ground and first excited states of the system. Our use of "long range" here is a convenient euphemism; we mean simply correlations between subsystems not directly coupled by the system Hamiltonian.

To be more concrete, let us describe a specific example of our results. Suppose we have a three-body system, with the bodies labeled 1, 2, and 3. We suppose systems 1 and 2 are coupled, and systems 2 and 3 are coupled. Importantly, systems 1 and 3 are *not* directly coupled. This is the only assumption we make about the system Hamiltonian. Suppose ψ is some joint pure state of the three systems, possessing "sufficient correlations" between system 1 and 3, in a sense to be made precise later. Our goal is to relate the energy gap to the overlap $F=|\langle\psi|E_0\rangle|$ between ψ and the ground state $|E_0\rangle$ of the system. Note that throughout the paper we interchange between the two pure-state notations where convenient: with a ket ($|\psi\rangle$) and without (ψ). We will prove that

$$\frac{\Delta E}{E_{\mathrm{tot}}} \leqslant 2(1-F^2), \tag{1}$$

where ΔE is the energy gap, and E_{tot} is the total energy scale for the system, i.e., the difference between the maximal and minimal energies of the Hamiltonian. The ratio of the gap to the total energy scale is an appropriate dimensionless parameter for deciding whether a gap is small or large. Note that rescaling of the Hamiltonian corresponds physically just to a rescaling of time, so one can only expect results in terms of such a dimensionless parameter; it does not make sense to say that a gap is "small" in any absolute sense—one needs to compare it to another relevant energy scale. The inequality (1) tells us that as the overlap F tends to 1, the gap size must vanish, compared to the total energy scale in the system, whenever the state ψ exhibits sufficiently strong correlations between systems 1 and 3.

Equation (1) is just one example of the sort of relation we will prove. We will prove a variety of similar relations, for different situations. In particular, we will analyze more general coupling schemes and consider the relationship of correlations to the energies of low-lying states other than the first excited state.

Our investigations may be placed in several different contexts, including the theory of quantum phase transitions, results from quantum many-body physics such as the Goldstone theorem, and the theory of entanglement developed within the burgeoning field of quantum information science. We now briefly review these connections.

A quantum phase transition [18,19] is a qualitative change in the properties of the ground state of a Hamiltonian $H(g)$ as a parameter g in the Hamiltonian is varied through a *critical point* g_c. The parameter g might, for example, be the value of an external magnetic field applied to a system of spins. Near a critical point, a system undergoing a second-order quantum phase transition usually exhibits two related phenomena. The first phenomenon is truly long-range correlations in the ground state, in the sense of correlations that decay only slowly with distance. The second phenomenon is a vanishing

*Electronic address: hlh@physics.uq.edu.au
†Electronic address: nielsen@physics.uq.edu.au; URL: http://www.qinfo.org/people/nielsen/
‡Electronic address: T.J.Osborne@bristol.ac.uk

energy gap $\Delta E \to 0$. These phenomena are expressed via the relations

$$\frac{1}{\xi} \propto |g - g_c|^{\eta}, \quad \Delta E \propto |g - g_c|^{z\eta}. \tag{2}$$

In these relations, ξ is the characteristic length scale on which correlations occur in the system, ΔE is the energy gap, as before, and η and z are constants known as *critical exponents*. Remarkably, the exact values of η and z do not depend on the particular details of the microscopic interactions in the system, but only on a small number of system parameters, such as dimensionality and symmetry; this phenomenon is known as universality. The exponent z is known as the *dynamical critical exponent*, and relates the way in which the energy gap vanishes to the way long-range correlations emerge near the critical point. In particular, we see that $\Delta E \propto \xi^{-z}$, so that, provided the critical exponent z is positive, the energy gap and the correlation length behave inversely to one another; as the gap becomes small, the correlation length becomes large, and vice versa.

Clearly, the study of the dynamical critical exponent z has much in common with the questions we are pursuing here. However, there are many significant differences. In particular, work on quantum phase transitions usually requires working in the thermodynamic limit of an infinite number of systems and often requires additional symmetry assumptions, such as translational invariance. Although numerous physical examples have suggested that it is generally true that correlations decay exponentially with the size of the energy gap, it is only relatively recently that a general proof of this fact has been provided [20], for systems in the thermodynamic limit.

In contrast, our results apply for any many-body quantum system, whether in the thermodynamic limit or not, and do not require any additional symmetry assumptions, such as translational invariance. Thus, our results complement those obtained in the study of quantum phase transitions.

Another context for our work is a classic result from quantum many-body physics, the nonrelativistic Goldstone theorem [21–23] (see Chap. 9 of [24] for a review), which shows that diverging correlations imply a vanishing energy gap. However, as with the case of work on quantum phase transitions, these results are complementary to ours, in that they rely on having infinite systems and typically require additional symmetry assumptions.

An intriguing aspect of our results is that they make considerable use of techniques developed in the new field of quantum information science,[1] especially techniques developed for the study of entanglement. Thus, our paper illustrates a general idea discussed elsewhere [2,13,27–29] (see also [30]), namely, that quantum information science may provide tools and perspectives for understanding the properties of complex quantum systems, complementary to the existing tools used in quantum many-body physics.

We begin the paper in Sec. II with a simple, easily understood toy model that illustrates many of the main physical ideas of the paper in a heuristic way. Much of the remainder of the paper is devoted to generalizations and formalization of the ideas in Sec. II. Interestingly, the mathematics that arises when generalizing formalizing the results of Sec. II leads in a natural way to other problems of great physical interest, and exploring these connections is a theme of the paper.

The next section of the paper, Sec. III, sets up a general framework for our investigations, introducing a convenient language to describe complex interactions involving many bodies, and precisely framing the questions we address in this language. Section IV is the core of the paper, presenting a series of general results connecting long-range ground-state correlations to the energy gap and other properties of the low-lying states. Section V explores an intriguing connection of our results to the theory of quantum error-correcting codes. Finally, Sec. IV concludes with a discussion of open questions.

II. INVITATION: A TOY MODEL

We begin with a toy model which illustrates in a simple setting many of the important physical ideas developed in more detail later in the paper. Our purpose in presenting these ideas first in a simple form is to keep the underlying physical ideas distinct from some of the mathematical complexities of later sections. Keep in mind, however, that some of these later mathematical complexities reveal surprising connections to other physical problems whose importance may not be apparent in the simplified setting discussed in this section.

Our toy model is a system of three qubits (spin-$\frac{1}{2}$ systems) arranged in a line. We label the qubits 1, 2, and 3. Suppose the qubits are coupled by a Hamiltonian H, which contains only nearest-neighbor interactions and so can be written $H = H_{12} + H_{23}$. Note that single-qubit contributions to the Hamiltonian can be included in the interactions H_{12} and H_{23}. For our purposes all that matters is that there are not couplings between qubits 1 and 3. Suppose the ground state of H, $|E_0\rangle$, is nondegenerate, with corresponding ground-state energy E_0. Suppose the gap to the energy of the first excited state is ΔE.

How entangled are qubits 1 and 3 in the ground state $|E_0\rangle$? We will prove that in order for qubits 1 and 3 to approach maximal entanglement, the gap ΔE must approach zero. We will give only a heuristic argument for now, with general proofs to follow later. Notes, in particular, that while the following argument applies for maximal entanglement between qubits 1 and 3, the results of subsequent sections can be applied to more general types of correlation.

We begin by observing that, since qubits 1 and 3 are nearly maximally entangled, then

$$|E_0\rangle \approx |\psi\rangle \equiv |\mathrm{ME}\rangle_{13} |\phi\rangle_2, \tag{3}$$

where $|\mathrm{ME}\rangle$ is some maximally entangled two-qubit state, $|\phi\rangle$ is a single-qubit state, and subscripts indicate which systems the states are associated with. But since $|E_0\rangle \approx |\psi\rangle$ the expectation energy for $|\psi\rangle$ must also be close to E_0,

[1] See [25,26] for reviews and further references.

$$\langle\psi|H|\psi\rangle \approx E_0. \tag{4}$$

Next, let $|\mathrm{ME}'\rangle$ be a two-qubit maximally entangled state orthogonal to $|\mathrm{ME}\rangle$ and define $|\psi_\perp\rangle \equiv |\mathrm{ME}'\rangle_{13}|\phi\rangle_2$. Note that $|\psi_\perp\rangle$ is orthogonal to $|\psi\rangle$, and in view of Eq. (3) it must be true that $|\psi_\perp\rangle$ is approximately orthogonal to $|E_0\rangle$. It follows that $|\psi_\perp\rangle$ can be expressed, approximately, as a superposition of states with energies E_1 and higher, where E_1 is the energy of the first excited state. Therefore the expectation energy for $|\psi_\perp\rangle$ must be at least E_1:

$$\langle\psi_\perp|H|\psi_\perp\rangle \geqslant E_1 + (\text{small corrections}). \tag{5}$$

These small corrections can, in principle, be negative, and we will see that this *must* be the case, in order to be consistent with the reasoning below.

Next, observe that the expectation energies for $|\psi\rangle$ and $|\psi_\perp\rangle$ are the same,

$$\langle\psi|H|\psi\rangle = \langle\psi_\perp|H|\psi_\perp\rangle. \tag{6}$$

To see this, observe that $\langle\psi|H_{12}|\psi\rangle = \langle\psi_\perp|H_{12}|\psi_\perp\rangle$, since the reduced density matrices for ψ and $\psi_\perp$ are *identical* on the system 12. A similar argument shows that the contribution to the expectation energy from H_{23} is the same from both ψ and $\psi_\perp$. Combining these results gives Eq. (6).

To complete the argument, observe that Eqs. (4)–(6) can be consistent only if $E_0 \approx E_1 +$ small corrections, and thus the energy gap must itself be small.

Summarizing, the presence of nearly maximal ground-state entanglement between sites which do not directly interact allows us to construct a state that (a) is almost orthogonal to the ground state, and thus must have energy of about E_1 or higher; but (b) looks locally very much like the ground state, and thus must have energy approximately E_0. The only way these two facts can simultaneously be true is if the energy gap is comparable in size to the corrections used in our approximations. Making this argument precise, and generalizing it further, is the subject of subsequent sections.

III. FRAMEWORK

This section introduces a framework for generalizing and formalizing the ideas of the previous section. We first introduce some general language for describing interactions in many-body quantum systems, then use this language to precisely state the main questions addressed through the remainder of the paper. we conclude with an overview of our answers to these questions.

In the previous section we considered three interacting qubits, with the restriction that the first and third qubits do not interact. It is helpful to introduce some language to describe more general interactions.

Suppose we have a general many-body system, with components labeled $1,\ldots,N$. We can regard these labels as a set of vertices V for a graph. Given a two-body Hamiltonian for that system, we can naturally associate with each coupling between bodies an (undirected) edge between the corresponding vertices. So, for example, the Hamiltonian[2] $H = XXI + ZIZ$ corresponds to a graph with vertices 1,2,3, and edges $\{1,2\},\{1,3\}$.

More generally, if some terms in the Hamiltonian couple more than two bodies, then we can associate with that Hamiltonian a *hypergraph*. A hypergraph consists of the set V of vertices, together with a collection of *hyperedges* E. Each hyperedge in E is just a subset of V, and represents a coupling term between the corresponding systems. So, for example, the Hamiltonian $H = XXI + ZIZ + YYZ$ corresponds to a hypergraph with vertices 1,2,3 and hyperedges $\{1,2\},\{1,3\}$, and $\{1,2,3\}$.

We call a hypergraph $G=(V,E)$ a *coupling topology* when it is associated with a quantum system in this way. We say that a Hamiltonian H *respects the coupling topology* G if every coupling in H corresponds to a hyperedge in G. We do not require every hyperedge in G to have a corresponding coupling in H. So, for example, the three-qubit Hamiltonian $H = XXI + ZIZ$ respects the coupling topology of the hypergraph with vertices 1, 2, 3 and hyperedges $\{1,2\},\{1,3\}$, and $\{1,2,3\}$, even though there is no term coupling qubits 1, 2, and 3 simultaneously.

Note that there is an apparent ambiguity in this definition, since a given Hamiltonian can be decomposed in more than one way, e.g., $H = XXI + IXX = XXM_+ + M_-XX$, where $M_\pm \equiv I \pm X$. We resolve this ambiguity by saying that H respects the coupling topology G if there is *some* decomposition of H which respects that coupling topology.

With this language we can now give a precise statement of the problem we are interested in. In fact, it is useful to consider two different forms of the problem. The simpler form is as follows.

Exact ground-state problem. Let ψ be a quantum state of some many-body system. We think of ψ as a *target state* that we desire to be the *exact* ground state. Suppose the system Hamiltonian H respects the coupling topology $G=(V,E)$. Given that ψ is an exact ground state of H, what does this imply about the level spacings of H? In particular, do the coupling topology G and the correlations present in ψ imply anything about the level spacings of the system, *independent* of the specific details of H?

We will show that the answer to this question is "yes." An example of the sort of answer we will give is as follows. Suppose ψ us an exact ground state of a Hamiltonian H respecting the coupling topology G. Then the ground state of H is at least m-fold degenerate, where m is an integer determined solely by (a) the coupling topology and (b) the properties of ψ. In particular, we will see that m is closely related to *long-range* correlations in ψ, where by long range we mean correlations between systems not directly coupled by G.

It is important that the degeneracy m is determined solely by properties of G and ψ: the particular details of the Hamiltonian H do not matter, beyond the topology of the interactions. Even given the ability to engineer arbitrary designer Hamiltonians, the fact that ψ is an exact ground state and G

[2]We use I,X,Y,Z to denote the four Pauli matrices and omit tensor product signs for notational brevity.

the coupling topology guarantees an m-fold degeneracy in the ground state.

More interesting and general than the study of exact ground states is the study of how the coupling topology and correlations in ψ affect the ability to *approximate* ψ as a ground state. This question is captured by the following problem.

Approximate ground-state problem. Let ψ be a quantum state of some many-body system. Suppose the system Hamiltonian H respects the coupling topology $G=(V,E)$. Given that the overlap between ψ and the ground state is $F \equiv \sqrt{\langle\psi|P_0|\psi\rangle}$, where P_0 projects onto the ground-state eigenspace, what does this imply about the level spacings of H? In particular, do the coupling topology G, the overlap F, and the correlations present in ψ imply anything about the level spacings of the system, *independent* of the specific details of H?

We will obtain solutions to this problem similar to those obtained for the exact ground-state problem. For example, suppose ψ has overlap F with the ground state of a Hamiltonian H respecting the coupling topology G. We will prove an inequality relating the gap ΔE to the overlap F and a measure C of long-range correlation in the system. This inequality will enable us to prove that as $F \to 1$ the presence of long-range correlations in the system forces the energy gap to vanish.

In the next two sections we will obtain several solutions to the approximate ground-state problem, applicable in different contexts. Interestingly, one of these solutions—in some sense the strongest—involves quantum error-correcting codes, as discussed in Sec. V.

IV. GENERAL THEORY

Suppose H is a Hamiltonian respecting the coupling topology $G=(V,E)$, and ψ is a quantum state having overlap F with the ground state. Our key result is a general theorem, proved in this section, connecting the energy levels of H to the properties of a set we shall define, labeled $R_G(\psi)$. $R_G(\psi)$ is defined to consist of all quantum states, both pure and mixed, which agree with ψ on the hyperedges in E. That is, $R_G(\psi)$ contains all states ρ such that $\mathrm{tr}_{\bar{e}}(\rho)=\mathrm{tr}_{\bar{e}}(|\psi\rangle\langle\psi|)$ for all the hyperedges e in E, where $\bar{e}$ indicates that we trace over all systems *except* those in e.

It is perhaps not obvious why a theorem connecting the energy levels of H to $R_G(\psi)$ should tell us anything about the relationship between those energy levels and long-range correlations. Remarkably, however, the properties of $R_G(\psi)$ are intimately connected with the correlations in ψ, and this fact will enable us to make the desired connections.

Our presentation strategy in this section is to first prove the general theorem and then explore connections between $R_G(\psi)$ and long-range correlations, applying the general theorem to a variety of examples.

A. Connection between the energy levels and $R_G(\psi)$

In this subsection we prove a general theorem connecting the energy levels of a system having ψ as its approximate ground state to $R_G(\psi)$. We begin by specifying some notation and nomenclature.

Recall that P_0 is the projector onto the ground-state eigenspace, and that the overlap between ψ and the ground state is $F \equiv \sqrt{\langle\psi|P_0|\psi\rangle}$. Assuming that $F>0$, we define $|E_0\rangle$ to be the (normalized) ground state onto which ψ projects. Explicitly, we define $|E_0\rangle \equiv P_0|\psi\rangle / \sqrt{\langle\psi|P_0|\psi\rangle}$. It will be convenient to label the energy levels as $E_0 \leqslant E_1 \leqslant \cdots$, and to let $E_{\max}$ be the largest energy level. Note that the energy levels are not assumed to be distinct, so, for example, if the ground state is doubly degenerate then we will have $E_0=E_1$. We choose $|E_1\rangle, |E_2\rangle, \ldots$ so that $|E_0\rangle, |E_1\rangle, \ldots$ forms an orthonormal eigenbasis of energy eigenstates in the obvious way. We let $E_{\mathrm{tot}}=E_{\max}-E_0$ be the total energy scale for the system.

With this nomenclature, we are now ready to proceed to the statement and proof of our main theorem. The key to the proof of the theorem is a lemma from linear algebra. The lemma is easy to state, and the result is rather obvious, yet all the proofs we are aware of make use of surprisingly sophisticated ideas. The result appears to be little known but is useful in many contexts. It appeared as Eq. (133) in a set of unpublished lecture notes [31].

Lemma 1. Let A and B be Hermitian matrices. Then

$$\lambda(\mathbf{A})^{\downarrow} \cdot \lambda(B)^{\uparrow} \leqslant \mathrm{tr}(AB) \leqslant \lambda(A)^{\downarrow} \cdot \lambda(B)^{\downarrow}, \tag{7}$$

where $\lambda(M)$ denotes the vector whose entries are the eigenvalues of the matrix M, $v^{\downarrow}(v^{\uparrow})$ is the vector whose entries are the entries of v rearranged into descending (ascending) order, and $\cdot$ is the Euclidean inner product.

Proof. We work in a basis in which A is diagonal, with its eigenvalues the diagonal entries of the matrix representation in that basis. Then

$$\mathrm{tr}(AB) = \sum_j A_{jj}B_{jj} = \lambda(A) \cdot \mathrm{diag}(B), \tag{8}$$

where $\mathrm{diag}(B)$ is the vector whose entries are the diagonal elements of B in this basis. Elementary results from the theory of majorization imply that $\mathrm{diag}(B) \prec \lambda(B)$, where $\prec$ denotes the majorization relation.[3] Further elementary results from the theory of majorization[4] imply that $\mathrm{diag}(B) = \Sigma p_j P_j \lambda(B)$, where the p_j form a probability distribution, and the P_j are permutation matrices. Substituting into Eq. (8) we obtain

$$\mathrm{tr}(AB) = \sum_j p_j \lambda(A) \cdot P_j \lambda(B). \tag{9}$$

The result now follows from the observation[5] that for any two vectors x and y, $x^{\downarrow} \cdot y^{\uparrow} \leqslant x \cdot y \leqslant x^{\downarrow} \cdot y^{\downarrow}$. ■

We are now in a position to state and prove our main theorem. Note, incidentally, that the proof of the main theorem makes use of only the first inequality in the statement of Lemma 1, not the second inequality. We included both be-

[3] This result appears on p. 218 of [32], as Theorem B.1 in Chap. 9. See, e.g., any of [31–35] for an introduction to majorization and further references.

[4] See p. 113 of [32], Proposition C.1 of Chap. 4.

[5] A proof of this observation may be found as Corollary II.4.4 on p. 49 of [34].

cause both are of interest, appear to be little known, and virtually no extra work is required to obtain the second.

Theorem 1. Let H be a Hamiltonian respecting the coupling topology G. Suppose ψ is a state with overlap F with the ground state of H. Let $\rho \in R_G(\psi)$ have eigenvalues $\rho_0 \geqslant \rho_1 \geqslant \cdots$. Then

$$\sum_{j=1}^{d-1} (E_j - E_0)\rho_j \leqslant (1-F^2)E_{\text{tot}}, \tag{10}$$

where d is the dimension of state space.

It is sometimes convenient to write the sum in a slightly different fashion. Including a $j=0$ term makes no difference, since $E_0 - E_0$ vanishes, so the sum may be rewritten $\Sigma_j (E_j - E_0)\rho_j$, with the sum over all possible indices, j.

Proof. By definition of F as the overlap between ψ and the ground state $|E_0\rangle$, we see that up to an unimportant global phase,

$$|\psi\rangle = F|E_0\rangle + \sqrt{1-F^2}|E_\perp\rangle, \tag{11}$$

where $|E_\perp\rangle$ is orthonormal to $|E_0\rangle$. We now use this expression to evaluate the average energy for the state $|\psi\rangle$. The first term on the right-hand side of Eq. (11) contributes F^2E_0 to the energy, while the second term contributes at most $(1-F^2)E_{\max}$, since the energy of $|E_\perp\rangle$ is no more than $E_{\max}$. It follows that $\langle\psi|H|\psi\rangle \leqslant F^2E_0 + (1-F^2)E_{\max}$. Rewriting this inequality in terms of $E_{\text{tot}} = E_{\max} - E_0$ rather than $E_{\max}$, we obtain

$$\langle\psi|H|\psi\rangle \leqslant E_0 + (1-F^2)E_{\text{tot}}. \tag{12}$$

Furthermore, since ψ and ρ have the same reduced density matrices on hyperedges in the coupling topology, we see that $\text{tr}(\rho H) = \langle\psi|H|\psi\rangle$ and thus

$$\text{tr}(\rho H) \leqslant E_0 + (1-F^2)E_{\text{tot}}. \tag{13}$$

Applying the first inequality of Lemma 1 to the left-hand side of Eq. (13) gives

$$\sum_{j=0}^{d-1} \rho_j E_j \leqslant E_0 + (1-F^2)E_{\text{tot}}. \tag{14}$$

Using the fact that $\Sigma_{j=0}^{d-1}\rho_j = 1$, and doing some elementary algebra and relabeling of indices, we see that this can be rewritten in the form $\Sigma_{j=1}^{d-1}(E_j - E_0)\rho_j \leqslant (1-F^2)E_{\text{tot}}$, as we set out to prove.

An interesting observation related to Theorem 1 is that if G_1 and G_2 are hypergraphs such that the hyperedges of G_1 are a subset of those of G_2, then $R_{G_2}(\psi) \subseteq R_{G_1}(\psi)$. This is true because if ρ and ψ agree on hyperedges in G_2 then they must certainly agree on hyperedges in G_1. It follows that Theorem 1 implies *stronger* constraints on the energy levels for systems whose coupling topology respects G_1 than for systems respecting G_2. Thus, for example, Theorem 1 gives stronger constraints on the energy levels for five spins arranged in a line, with nearest-neighbor interactions, than for the same spins arranged into a circle, again with nearest-neighbor interactions.

B. Example applications

We now explore some applications of Theorem 1, relating the energy spectrum of a system to the presence of long-range correlations in the ground state of that system.

1. Example: Perfect long-range correlations

Suppose we have a three-component system, with subsystems labeled 1,2, and 3. Suppose the coupling topology G is such that systems 1 and 2 may interact, systems 2 and 3 may interact, but systems 1 and 3 cannot interact directly. Note that in this discussion 1,2, and 3 may be aggregates—e.g., systems 1 and 3 might be spins on either end of a long linear chain, with system 2 the collection of all spins in between. Suppose, furthermore, that ψ is some quantum state exhibiting *perfect correlation* between systems 1 and 3. By perfect correlation, we mean that there is a measurement basis in which a measurement outcome of j on system 1 implies, with probability 1, a measurement outcome j on system 3, and conversely.

As an example of such a situation, ψ could be a product $\psi_{13} \otimes \psi_2$. In this case ψ exhibits perfect correlations if measurements are performed in the Schmidt bases for systems 1 and 3, respectively.

Another example is states ψ such that when system 2 is traced out we get a mixed state of the form $\Sigma_j p_j |j\rangle_1\langle j|_1 \otimes |j\rangle_3\langle j|_3$, where $|j\rangle_1$ and $|j\rangle_3$ are orthonormal bases for systems 1 and 3, respectively. It is easy to show that such states must have three-party Schmidt decompositions of the form studied by Thapliyal [36] and Peres [37], i.e., $\psi = \Sigma_j \sqrt{p_j}|j\rangle_1|j\rangle_2|j\rangle_3$, where $|j\rangle_1$, $|j\rangle_2$, and $|j\rangle_3$ are orthonormal bases for the respective systems. An example of such a state is the Greenberger-Horne-Zeilinger (GHZ) state $|\text{GHZ}\rangle = (|000\rangle + |111\rangle)/\sqrt{2}$. Indeed, if we consider an n-qubit linear array, with the first and last qubits considered as system 1 and system 3, with the remaining qubits grouped together as system 2, then we see that the n-party GHZ state $|\text{GHZ}\rangle = (|0\rangle \otimes n + |1\rangle \otimes n)/\sqrt{2}$ is also an example of such a state.

In general, if ψ is any state exhibiting such perfect correlations there must exist normalized, but possibly nonorthogonal, states $|j\rangle_2$ of system 2, such that

$$\psi = \sum_j \sqrt{p_j}|j\rangle_1|j\rangle_2|j\rangle_3. \tag{15}$$

Note that p_j are the probabilities with which the measurement outcome j occurs on systems 1 and 3. Now define

$$\rho \equiv \sum_j p_j |j\rangle_1\langle j|_1 \otimes |j\rangle_2\langle j|_2 \otimes |j\rangle_3\langle j|_3. \tag{16}$$

Observe that $\rho \in R_G(\psi)$, since it has the same reduced density matrices on systems 12 and 23 as does ψ. Note also that ρ has eigenvalues p_j. It will be convenient to assume that the measurement outcomes are labeled $0, 1, \ldots, d-1$ and have been ordered so that $p_0 \geqslant p_1 \geqslant \cdots$. From Theorem 1 we have

$$\sum_{j=1}^{d-1} (E_j - E_0) p_j \leq (1 - F^2) E_{\text{tot}}. \tag{17}$$

Equation (17) tells us that as $F \to 1$ the quantity on the left-hand side gets squeezed toward zero. In particular, if $p_0, \ldots, p_k > 0$, then we conclude that $E_k \to E_0$, as do all the lower energy levels $E_1, \ldots, E_{k-1}$. So, for example, in the scenario of Sec. II, if $\psi = |\text{ME}\rangle_{13} |\phi\rangle_2$, then we have $p_0 = p_1 = \frac{1}{2}$, and Eq. (17) becomes[6] $E_1 \leq E_0 + 2(1 - F^2) E_{\text{tot}}$. Thus, the gap to the first excited state in this example vanishes as $F \to 1$.

Similarly, for an n-qubit linear chain, with systems 1 and 3 the qubits on each end of the chain, if ψ is the n-party GHZ state, then we again conclude that $E_1 \leq E_0 + 2(1 - F^2) E_{\text{tot}}$, and the gap to the first excited state vanishes as $F \to 1$.

Another illuminating—albeit, ultimately trivial—example is when system 13 is in a product state $\psi = |a\rangle_1 |b\rangle_2 |c\rangle_3$. This is a case of the theorem, for the system does exhibit perfect correlation, provided system 1 is measured in a basis including $|a\rangle$, and system 3 is measured in a basis including $|b\rangle$. However, since we have $p_0 = 1$, and all other $p_j = 0$, we see that Eq. (17) gives us only trivial information $0 \leq (1 - F^2) E_{\text{tot}}$ and cannot be used to deduce anything about the spectrum of the system. It is only as the probabilities p_j become mixed that Eq. (17) may be used to duduce interesting information about the spectrum.

2. Example: Imperfect long-range correlations

Let us generalize the previous example so that it applies also to systems with imperfect correlations. Suppose again that we have a three-component system 123, and the coupling topology allows 1 and 2, and 2 and 3 to interact, but not 1 and 3. Suppose ψ is an *exact* ground state for a Hamiltonian respecting this coupling topology. Suppose $|j\rangle$ is an orthonormal basis for system 1, and $|k\rangle$ is an orthonormal basis for system 3. We can expand ψ as

$$\psi = \sum_{jk} \sqrt{p_{j,k}} |j\rangle |e_{jk}\rangle |k\rangle, \tag{18}$$

where $p_{j,k}$ is the probability of getting the measurement outcome j on system 1 and k on system 3, if measurements are performed in the $|j\rangle$ and $|k\rangle$ bases, respectively. The states $|e_{jk}\rangle$ are normalized, but possibly nonorthogonal, states of system 2.

To measure the correlation between the measurement outcomes on systems 1 and 3 we define a correlation measure

$$C \equiv \sum_j p_{j,j}. \tag{19}$$

C is just the probability that the measurement outcome on system 1 is the same as the measurement outcome on system 3. Thus, values of C close to 1 indicate highly correlated measurement outcomes, while values very close to zero indicate a high level of anticorrelation. the choice of this form for C is a matter of convenience in that later results become quite simple because of it. However, the definition is different from other more conventional correlation measures, such as spin-spin correlation functions, or the "correlation coefficient" from statistics.

The definition of C implicitly assumes that the same labels j are being used for measurement outcomes on system 1 and system 3. This need not be the case. For example, system 1 might be a spin-$\frac{1}{2}$ system, with measurement outcomes labeled $\pm\frac{1}{2}$, and system 3 a spin-1 system, with measurement outcomes labeled 0, ± 1. If this is the case we can define an analogous notion of correlation by identifying the outcomes of the spin-$\frac{1}{2}$ measurement with a subset of the spin-1 outcomes, e.g., $1/2 \to 1$, $-1/2 \to -1$, and so $C = p_{1/2,1} + p_{-1/2,-1}$. In general, we can define a measure of correlation by identifying the measurement outcomes for the system with the smaller state space with a subset of the measurement outcomes for the system with the larger state space. The arguments below are easily generalized to this case, but for notational clarity we stick to the case when systems 1 and 3 have identical labelings for their measurements.

Next, we define a normalized and perfectly correlated state ψ' of the joint system by discarding those terms in ψ that leas to the correlations being imperfect, and renormalizing the state appropriately:

$$\psi' \equiv \frac{\sum_j \sqrt{p_{j,j}} |j\rangle |e_{jj}\rangle |j\rangle}{\sqrt{C}}. \tag{20}$$

Note that we must assume $C > 0$ for this definition to be valid. ψ' obviously exhibits perfect correlation between systems 1 and 3, in the sense of the earlier example, and thus we conclude that

$$\sum_j \frac{p_{j,j}}{C} (E_j - E_0) \leq (1 - F^2) E_{\text{tot}}, \tag{21}$$

where F is the overlap between ψ' and the ground state. But we assumed that ψ was a ground state (possibly one of many), so $F \geq |\langle \psi' | \psi \rangle| = \sqrt{C}$, and thus the previous equation may be rewritten

$$\sum_j p_{j,j} (E_j - E_0) \leq C(1 - C) E_{\text{tot}}, \tag{22}$$

provided $C > 0$. Equation (22) tells us that as $C \to 1$, i.e., as we approach perfect correlation, the quantity on the left-hand side must approach zero. Thus, if $p_{0,0}, \ldots, p_{k,k} > 0$ then we conclude that $E_1, \ldots, E_k \to E_0$ as the correlations become perfect.

3. Example: Approximating a state with imperfect long-range correlations

We can generalize the previous two examples still further, to the case where we are trying to *approximate* a state with imperfect correlations as the ground state. Suppose again that we have a three-component system 123, and the coupling topology allows 1 and 2, and 2 and 3 to interact, but not 1 and 3. Suppose ψ is a state with correlation $C = \Sigma p_{j,j} > 0$ in some measurement basis for systems 1 and 3. Suppose there

[6] Compare Eq. (1).

is a Hamiltonian respecting the coupling topology such that the overlap between ψ and the ground state is F. We will prove that the energy levels of the Hamiltonian satisfy

$$\sum_j p_{j,j}(E_j - E_0) \leqslant C(\sqrt{1-C} + \sqrt{1-F^2})^2 E_{\text{tot}}. \quad (23)$$

This result generalizes both the last example, Eq. (22), which corresponds to the case when $F=1$, and the example before that, Eq. (17), which corresponds to the case $C=1$.

Similarly to the previous example, we can write $\psi = \Sigma_{jk}\sqrt{p_{j,k}}|j\rangle|e_{jk}\rangle|k\rangle$, and define $\psi' \equiv \Sigma_j \sqrt{p_{j,j}}|j\rangle|e_{j,j}\rangle|j\rangle/\sqrt{C}$. We now define $F(a,b) \equiv |\langle a|b\rangle|$, the overlap between any two states $|a\rangle$ and $|b\rangle$. It is convenient to note that $\sqrt{1-F(a,b)^2}$ is a metric on projective state space. Recall that $|E_0\rangle = P_0|\psi\rangle/\sqrt{\langle\psi|P_0|\psi\rangle}$ is the normalized state that arises from projecting ψ onto the ground space. From the triangle inequality

$$\sqrt{1-F(\psi',E_0)^2} \leqslant \sqrt{1-F(\psi',\psi)^2} + \sqrt{1-F(\psi,E_0)^2} \quad (24)$$

$$\leqslant \sqrt{1-C} + \sqrt{1-F^2}. \quad (25)$$

But if $F_{\psi'} \equiv \sqrt{\langle\psi'|P_0|\psi'\rangle}$ is the overlap of ψ' with the ground space then we have $F_{\psi'} \geqslant F(\psi',E_0)$ and thus combining with Eq. (25) we have

$$1 - F_{\psi'}^2 \leqslant (\sqrt{1-C} + \sqrt{1-F^2})^2. \quad (26)$$

The result now follows from Eq. (17).

Summarizing, we have proved the following general theorem.

Theorem 2. Let H be a Hamiltonian coupling systems 1 and 2, and 2 and 3, but not 1 and 3. Let $p_{j,k}$ be the joint probability distribution associated with a measurement in some bases for systems 1 and 3, for a state ψ. Label the measurement outcomes $0,1,\ldots$, and so that $p_{0,0} \geqslant p_{1,1} \geqslant \cdots$. Define the correlation measure $C \equiv \Sigma_j p_{j,j}$, and let F be the overlap between ψ and the ground state. Then, provided that $C>0$, the energy levels of H are constrained by the relation

$$\sum_j p_{j,j}(E_j - E_0) \leqslant C(\sqrt{1-C} + \sqrt{1-F^2})^2 E_{\text{tot}}. \quad (27)$$

C. Exact ground states and ground-state degeneracy

We have seen that the properties of $R_G(\psi)$ are closely related to long-range correlations in the state ψ. In this section we make some more specialized observation about $R_G(\psi)$ that can be used to prove results about the ground-state degeneracy of any Hamiltonian with ψ as an *exact* ground state.

We define $N_{\text{rank}}(\psi)$ to be the maximal rank of any density matrix in $R_G(\psi)$. We will see below that $N_{\text{rank}}(\psi)$ is connected to both the long-range correlations in ψ and also to the ground-state degeneracy. We begin with the latter connection.

Theorem 3. Let H be a Hamiltonian respecting the coupling topology G. Suppose ψ is a ground state of H. Then the ground state is at least $N_{\text{rank}}(\psi)$-fold degenerate.

Proof. A direct proof is easily obtained. Let ρ be the state in $R_G(\psi)$ of maximal rank, let ψ_j be the eigenvectors of ρ with nonzero eigenvalues, and argue that all the ψ_j must have energy equal to the ground-state energy. This follows since, if one has energy higher than the ground state, then another must have energy below the ground state—a contradiction—to ensure that $\text{tr}(H\rho)$ is equal to the ground-state energy. Alternatively, observe that this theorem is a special case of Theorem 1, with $F=1$. ■

Example. As an example, suppose we have just three systems 1,2,3, and suppose only couplings between 12 and 23 are involved. Suppose that $\psi = \psi_{13} \otimes \psi_2$, where ψ_{13} is an entangled state of systems 1 and 3, with Schmidt decomposition $\psi_{13} = \Sigma_j \sqrt{p_j}|j\rangle|j\rangle$, and ψ_2 is some state of system 2.

We will analyze this scenario in two different ways. The first method of analysis is similar in spirit to arguments earlier in the paper, such as led to Theorem 2. The second method is from a somewhat different point of view, and we will see that it sometimes leads to stronger results. Our first argument is as follows. Just as argued earlier, $\rho = \Sigma_j p_j |j\rangle\langle j| \otimes |\psi_2\rangle\langle\psi_2| \otimes |j\rangle\langle j|$ is in $R_G(\psi)$. We therefore see, from any one of Theorems 3, 2, and 1, that the ground-state degeneracy is at least equal to the Schmidt number of ψ_{13}, $\text{Sch}(\psi_{13})$, i.e., the number of nonzero coefficients in the Schmidt decomposition. It follows that if $\psi_{13} \otimes \psi_2$ is to be a ground state of the system, then the ground state must be $\text{Sch}(\psi_{13})$-fold degenerate. Of course, the Schmidt number is a well-known entanglement monotone, so in this example we conclude that the ground-state degeneracy is at least as large as the amount of long-range entanglement, as measured by the Schmidt number.

Our second method of analysis takes a state-based, rather than operator-based, point of view. Let $S_G(\psi)$ be the set of *pure* quantum states agreeing with ψ on hyperedges, i.e., it is the subset of $R_G(\psi)$ containing only pure states. Define $N_{\text{span}}(\psi)$ to be the dimension of the linear space spanned by the vectors in $S_G(\psi)$. Observe then that $N_{\text{span}}(\psi) \leqslant N_{\text{rank}}(\psi)$, since given any linearly independent $\psi_1, \ldots, \psi_m \in S_G(\psi)$ we can form $\rho = \Sigma_j |\psi_j\rangle\langle\psi_j|/m \in R_G(\psi)$, which has rank m. Thus, Theorem 3 implies that the ground state is at least $N_{\text{span}}(\psi)$-fold degenerate.

In the scenario studied above, with $\psi = \psi_{13} \otimes \psi_2$, $\psi_{13} = \Sigma_j \sqrt{p_j}|j\rangle|j\rangle$, we see that the states $\Sigma_j \sqrt{p_j} e^{i\theta_j}|j\rangle|j\rangle$ are in $S_G(\psi)$ for any choice of the phases θ_j, and thus $N_{\text{span}}(\psi) \geqslant \text{Sch}(\psi_{13})$, and we conclude, as earlier, that the ground state is at least $\text{Sch}(\psi_{13})$–fold degenerate. However, when the Schmidt coefficients p_j are degenerate, $N_{\text{span}}(\psi)$can actually be somewhat larger than the Schmidt number $\text{Sch}(\psi_{13})$. The following proposition enables us to make a precise evaluation of $N_{\text{span}}(\psi)$.

Proposition 1. Let $\psi = \psi_{13} \otimes \psi_2$, where $\psi_{13} = \Sigma_j \sqrt{p_j}|j\rangle|j\rangle$. Then $N_{\text{span}}(\psi) = \Sigma_k d_k^2$, where the sum is over an index k for distant nonzero Schmidt coefficients, and d_k is the degeneracy of the kth nonzero Schmidt coefficient.

Note that, according to the proposition, when ψ_{13} has nondegenerate Schmidt coefficients, $N_{\text{span}}(\psi)$ is equal to the Schmidt number of ψ_{13}, which is an entanglement monotone. However, using the results of [38] it is easy to construct examples with degenerate Schmidt coefficients that show $N_{\text{span}}(\psi)$ is not, in general, an entanglement monotone.

Proof. It is clear that all states in $S_G(\psi)$ have the form $\phi_{13}\otimes\psi_2$ where ϕ_{13} is a state having the same reduced density matrices on systems 1 and 3 as does ψ_{13}. But it is easy to see that this is the case if and only if $\phi_{13}=e^{i\theta}[(\oplus_k U_k)\otimes I]\psi_{13}$, where θ is a phase factor, U_k is a special unitary operator acting on the subspace of system 1 corresponding to the kth Schmidt coefficient, and $\oplus_k$ denotes the direct sum over those subspaces. The result now follows from the simple observation that in a $d_k\otimes d_k$ space, the dimension spanned by states from the form $(U\otimes I)\Sigma_j|j\rangle|j\rangle$, where $U\in \mathrm{SU}(d_k)$, is d_k^2. ■

This proposition shows us how to evaluate $N_{\mathrm{span}}(\psi)$ for a large class of interesting states and thus to place lower bounds on the ground-state degeneracy. When ψ_{13} is degenerate these results are actually stronger than are obtained using Theorem 2, since $N_{\mathrm{span}}(\psi)$ is strictly larger in this case than the Schmidt number of ψ_{13}. Although the argument leading to Theorem 2 can be modified to give this stronger bound, the modification is not especially natural from a physical point view. Thus, we believe there is some merit in the alternative, state-based point of view taken in the present discussion.

Example. Recall that a state with a multiparty Schmidt decomposition can be written in the form [36,37] $\psi=\Sigma_j\sqrt{p_j}|j\rangle\cdots|j\rangle$. An example of such a state is the n-qubit GHZ state $|\mathrm{GHZ}\rangle=(|0\rangle^{\otimes n}+|1\rangle^{\otimes n})/\sqrt{2}$. Suppose the coupling topology G contains all hyperedges of up to $n-1$ vertices, i.e., the allowed Hamiltonians may couple up to $n-1$ of the systems, but not all n systems simultaneously. It is easy to see that the states $\Sigma_j\sqrt{p_j}e^{i\theta_j}|j\rangle\cdots|j\rangle$ are in $S_G(\psi)$, for any choice of the phases θ_j, and thus $N_{\mathrm{span}}(\psi)\geqslant \mathrm{Sch}(\psi)$, where $\mathrm{Sch}(\psi)$ is the number of terms appearing in the multiparty Schmidt decomposition. It follows that the ground state of H is at least $\mathrm{Sch}(\psi)$- fold degenerate. For example, in the case of the GHZ state, it follows that the ground state is at least twofold degenerate, since the GHZ state has Schmidt number 2.

D. Further development of Theorem 1

Can Theorem 1 be strengthened in any way? We now show that there are physically interesting ways of varying the hypotheses of Theorem 1, in order to reach stronger conclusion. One way of doing this, related to quantum error-correcting codes, is described in detail in Sec. V. We now explain, more briefly, another possible variation.

The basic idea is to amend Theorem 1 so it makes use of information about the relationship between ψ and ρ. Consider two possible cases: (a) ψ is orthogonal to the support of ρ, and (b) ψ is contained in the support of ρ. In the former case, we see that there is a subspace of dimension $\mathrm{rank}(\rho)+1$, spanned by the support of ρ and ψ, in which energies are all approximately equal to E_0, and thus $E_0\approx E_1\approx\cdots E_{\mathrm{rank}(\rho)}$. In the latter case we can conclude only that there is a subspace of dimension $\mathrm{rank}(\rho)$—the support of ρ—in which energies are all approximately equal to E_0, and thus we draw the weaker conclusion that $E_0\approx E_1\approx\cdots E_{\mathrm{rank}(\rho)-1}$.

We have not yet succeeded in obtaining a clean generalization of Theorem 1 incorporating this idea. However, we have obtained a simpler result in this vein, which we now briefly describe.

Proposition 2. Let H be a Hamiltonian respecting the coupling topology G. Suppose ψ is a state having overlap F with the ground state of H. Suppose $\phi\in R_G(\psi)$ is such that $|\langle\psi|\phi\rangle|=\cos(\theta)$. Then

$$E_1-E_0\leqslant\frac{1-F^2}{g(\theta,F)}E_{\mathrm{tot}},\tag{28}$$

where $g(\theta,F)\equiv[1-F\cos(\theta)+\sqrt{1-F^2}\sin(\theta)]^2$.

Note that ϕ plays a role analogous to ρ in Theorem 1. The crucial additional piece of structure in the proposition is the angle θ relating ψ and ϕ. As this angle varies from 0 to $\pi/2$, the bound Eq. (28) varies from the vacuous $E_1-E_0\leqslant E_{\mathrm{tot}}$—as with Theorem 1 we get no information at all in this case—through to $E_1-E_0\leqslant(1-F^2)E_{\mathrm{tot}}/F^2$, which is nontrivial. Note that Theorem 1 can be applied also in this latter case; the strongest bound obtained in this way comes from choosing $\rho=\frac{1}{2}|\psi\rangle\langle\psi|+\frac{1}{2}|\phi\rangle\langle\phi|$, which gives $E_1-E_0\leqslant 2(1-F^2)E_{\mathrm{tot}}$, which is a factor of 2 weaker than Proposition 2, in the $F\to 1$ limit.

Proof. By the same argument that led to Eq. (13), we conclude that

$$\langle\phi|H|\phi\rangle\leqslant E_0+(1-F^2)E_{\mathrm{tot}}.\tag{29}$$

Expressing $|E_0\rangle$ in terms of ψ we have, up to an unimportant global phase, $|E_0\rangle=F|\psi\rangle+\sqrt{1-F^2}|\psi^\perp\rangle$, for some $\psi^\perp$ orthonormal to ψ. Taking the inner product with ϕ gives $|\langle\phi|E_0\rangle|\leqslant F\cos\theta+\sqrt{1-F^2}|\langle\phi|\psi^\perp\rangle|$. Because $\psi^\perp$ is orthonormal to ψ we have $|\langle\phi|\psi^\perp\rangle|\leqslant\sin\theta$, and so

$$|\langle\phi|E_0\rangle|\leqslant F\cos\theta+\sqrt{1-F^2}\sin\theta.\tag{30}$$

We see from this equation that the component of ϕ orthogonal to $|E_0\rangle$ is at least $\sqrt{g(\theta,F)}$, as defined in the statement of the proposition, and thus

$$\langle\phi|H|\phi\rangle\geqslant[1-g(\theta,F)]E_0+g(\theta,F)E_1.\tag{31}$$

Combining this inequality with Eq. (29) and rearranging gives the result. ■

E. Understanding $R_G(\psi)$

The key to applying Theorem 1 is the ability to find states ρ lying in $R_G(\psi)$. To this end, we make a few general remarks on the problem of understanding $R_G(\psi)$.

Our first observation is that $R_G(\psi)$ is a convex set, since a mixture of states, each of which agrees with ψ on hyperedges, also agrees with ψ on hyperedges. Therefore, one might try to understand $R_G(\psi)$ by finding its extreme points. Unfortunately, we do not know what those extreme points are, or even if they are pure or mixed quantum states.

Additional light on $R_G(\psi)$is shed by the work of Linden, Popescu, and Wootters [39], and subsequent work by Linden and Wootters [40]. In [39] it is shown that almost all three-qubit quantum states are uniquely determined by their two-party reduced density matrices. More precisely, given a three-qubit state $\psi=\psi_{123}$, let $\rho_{12},\rho_{13},\rho_{23}$ be that corresponding two-qubit reduced density matrices. They [39] show that unless the state is equivalent, up to local unitaries, to a state

of the form $a|000\rangle+b|111\rangle$, ψ is the unique state, even allowing mixed states, with those reduced density matrices.

Restating in our language, [39] shows that for all ψ except those equivalent to $a|000\rangle+b|111\rangle$ by local unitaries, $R_G(\psi)=\{\psi\}$, when G is the complete graph allowing interactions between any pair of the systems 1, 2, and 3. Thus, Theorem 1 gives only nontrivial information when the state ψ is locally equivalent to $a|000\rangle+b|111\rangle$. Of course, bounds like Theorem 2 apply in general.

The results of [39] were extended in [40], which considered the scenario of n qudits, i.e., d-dimensional quantum systems. Reference [40] proved the existence of constants α and β, $0<\alpha<\beta<1$, such that (a) specifying all reduced density matrices for subsystems containing βn qudits uniquely determined the global state for almost all quantum states, and (b) knowing all the reduced density matrices on up to $n\alpha$ qudits does *not* uniquely determine the state, in general. The estimates they obtained for α and β were of order 1, and depended on the value of d; for details, see [40].

Restating in our language, [40] showed that if G includes all hyperedges involving up to βn vertices, then for almost all $\psi, R_G(\psi)=\{\psi\}$. However, for more physically interesting cases, like when the coupling topology involves only two-body interactions, the results of [40] suggest that $R_G(\psi)$ will typically contain mixed states, and thus the bounds of Theorem 1 become nontrivial.

V. CONNECTION TO QUANTUM ERROR-CORRECTING CODES

There is an interesting way to strengthen the conclusions of the earlier theorems, by making use of stronger hypotheses. Intriguingly, this line of thinking leads to a natural connection with quantum error-correcting codes. We present this material starting with a general theorem connecting the gap to the properties of the ground state, and then explain how those properties are connected to quantum error-correcting codes.

We begin with a little more notation. Let $\overline{S_G(\psi)}$ denote the set of all vectors $\lambda\phi$, where λ is a complex number, and ϕ is a state in $S_G(\psi)$. Let $N_{\text{space}}(\psi)$ be the dimension of the largest vector space which is a subset of $\overline{S_G(\psi)}$. We now prove that $N_{\text{space}}(\psi)$ is connected to the spectral properties of the system.

Theorem 4. Let H be a Hamiltonian respecting the coupling topology G. Suppose ψ is a state with overlap F with the ground state. Then

$$E_0 \leqslant E_1 \leqslant \cdots \leqslant E_{N_{\text{space}(\psi)-1}} \leqslant E_0 + (1-F^2)E_{\text{tot}}. \quad (32)$$

The inequality that is the conclusion of this theorem is substantially stronger than the inequalities proved earlier, such as Theorem 1 and its corollaries. The reason this stronger conclusion is possible is because we use a stronger hypothesis as the basis for our reasoning. The key fact is that every state in the maximal subspace of $\overline{S_G(\psi)}$ is guaranteed to have the same expectation energy for Hamiltonians respecting G. In contrast, in the scenario of Theorem 1, we know the $\rho \in R_G(\psi)$, but this does not imply that all states in the support of $R_G(\psi)$ have the same expectation energy. It is this difference that allows us to draw a stronger conclusion in the present scenario.

Proof. Let V be the maximal vector space which is a subset of $\overline{S_G(\psi)}$. By the Courant-Fischer-Weyl minimax principle (see Chap. 3 of [34]), we have

$$E_{N_{\text{space}}(\psi)-1} \leqslant \max_{\phi \in V, \|\phi\|=1} \langle\phi|H|\phi\rangle. \quad (33)$$

But by the same reasoning that led to Eq. (13) the right-hand side of the previous equation is bounded above by $E_0+(1-F^2)E_{\text{tot}}$, which gives the result.

How can we evaluate $N_{\text{space}}(\psi)$? Insight into this question is provided by noticing an interesting connection, namely, that the maximal vector space contained in $\overline{S_G(\psi)}$ is a type of quantum error-correcting code. To see this, let us recall some basic facts from the theory of quantum error correction [25,26].

Let S be a set whose elements are collections of subsystems of some quantum system. The elements of S represent (collective) subsystems on which errors are allowed to occur, and still be correctable by the code. For example, for a code correcting errors on up to two qubits at a time, S consists of all pairs $\{j,k\}$ of labels for two qubits. A quantum error-correcting code correcting errors on S is vector space W such that

$$PA^\dagger BP \propto P, \quad (34)$$

where P projects onto the code space W, and $A$and B are arbitrary operators that act nontrivially only on subsystems which are elements of S. These conditions, Eq. (34), define what it is to be a quantum error-correcting code correcting errors on S. For more on the physical interpretation of these conditions, see (25) and (26).

We return now to the connection between Theorem 4 and quantum error correction. In one direction, the connection is quite simple. Suppose ψ is a state in a k-dimensional quantum error-correcting code W which corrects errors on a set S. We define a coupling topology on the system, $G=(V,E)$, by specifying that E consists of all hyperedges e such that $e \subseteq s_1 \cup s_2$ for some $s_1, s_2 \in S$. We will use Eq. (34) to show that all states ϕ in the code W must have the same reduced density matrices on any hyperedge e, and thus $W \subseteq \overline{S_G(\psi)}$, and therefore $N_{\text{ortho}}(\psi) \geqslant k$.

To see this, suppose C is an operator that is a tensor product of operators acting on the individual systems in e. It follows that $C=A^\dagger B$ for some operators A and B acting only on the systems in s_1 and s_2. We have, by Eq. (34), $PCP=\gamma P$ for some constant of proportionality γ. It follows that if ϕ is any state in the code then

$$\text{tr}(|\phi\rangle\langle\phi|C) = \gamma. \quad (35)$$

This is true for all ϕ in the code, and because C was an arbitrary tensor product acting on e, we see that the reduced density matrix on e must be the same for all elements ϕ of the code.

The converse statement is also true. Suppose W is the maximal subspace in $\overline{S_G(\psi)}$. Suppose S is any set such that for each pair s_1 and s_2 in S there is a hyperedge e in E satisfying $e \supseteq s_1 \cup s_2$. We will show that W is an error-correcting code correcting errors on S. The proof is similar to but slightly more elaborate than the proof in the previous paragraph. Let A and B be operators acting nontrivially only on subsystems s_1 and s_2. We aim to establish Eq. (34). Because all states ϕ in $S_G(\psi)$ have the same reduced density matrices on e we conclude that

$$\langle \phi | A^\dagger B | \phi \rangle = \gamma, \tag{36}$$

for some constant γ independent of ϕ. This implies

$$|\phi\rangle\langle\phi|A^\dagger B|\phi\rangle\langle\phi| = \gamma|\phi\rangle\langle\phi|. \tag{37}$$

Naively, one might try to establish Eq. (34) by summing over an orthonormal basis of state ϕ for W.Of course, this may not work, because of possible cross terms on the left-hand side of Eq. (37). We will show, however, that these cross terms vanish. To see this, let $|j\rangle$ be an orthonormal basis for W. Then for any pair $j_1 \neq j_2$ we have

$$(\langle j_1| + \langle j_2|)A^\dagger B(|j_1\rangle + |j_2\rangle) = (\langle j_1| - \langle j_2|)A^\dagger B(|j_1\rangle - |j_2\rangle) \tag{38}$$

and

$$(\langle j_1| - i\langle j_2|)A^\dagger B(|j_1\rangle + i|j_2\rangle) = (\langle j_1| + i\langle j_2|)A^\dagger B(|j_1\rangle - i|j_2\rangle. \tag{39}$$

Adding the first of these equations to i times the second equation gives

$$\langle j_2 | A^\dagger B | j_1 \rangle = 0, \tag{40}$$

which establishes that the cross terms vanish, and thus that W is quantum error-correcting code.

We have shown that systems with quantum error-correcting codes as approximate ground states must satisfy especially stringent constraints on their low-lying spectra. It is interesting to compare these results with those of [1], where it was shown that nondegenerate quantum error-correcting codes correcting errors on up to L subsystems cannot be the ground state of any nontrivial L-local Hamiltonian, i.e., a Hamiltonian coupling no more than L subsystems at a time, and not a multiple of the identity. Remarkably, [1] proved a *constant* lower bound on the distance between the ground state and states of code in this scenario. This constant lower bound is much stronger even than the bounds of Theorem 4. However, a critical difference is that the results of [1] applied only to nondegenerate codes, while Theorem 4 is more general in that it applies also to degenerate codes.

Viewed from a slightly different angle, our results provide an amusing counterpoint to [1]. Reference [1] pointed out that no state in a nondegenerate code correcting up to L errors can be a ground state of an L-local Hamiltonian. Theorem 4 implies if one state of a degenerate code correcting L errors is a ground state of an L-local Hamiltonian, then *all states* of that code must be ground states of Hamiltonian. This is because all states in such a code must share the same set of reduced density matrices on collections of up to $2L$ subsystems and thus share the same energy with respect to an n-local Hamiltonian if $n \leqslant 2L$. Physically, this is clear *a priori* — all the states of the code must be energetically indistinguishable, in order to preserve information. However, it seems to us an interesting fact that either all or none of the states of quantum error-correcting code can be ground states. There is nothing in between.

VI. DISCUSSION

We have developed several general results demonstrating that systems exhibiting ground-state entanglement or correlation that is "long range," in the sense of being between subsystems not directly coupled, must necessarily have a small energy gap. These results suggest many interesting avenues for further investigation.

Characterizing the physical properties responsible for the vanishing gap. We have demonstrated several connections linking the energy gap to long-range correlations and entanglement in the ground state. However, many of the connections we have identified only hold for special (albeit still rather general) cases, rather than in the most general case. What are the physical properties responsible for the vanishing of the gap in the most general case?

Characterizing $R_G(\rho)$. Our work has highlighted the importance of understanding the set $R_G(\rho)$, defined to be the set of all density matrices σ with the property that $\mathrm{tr}_{\bar{e}}(\rho) = \mathrm{tr}_{\bar{e}}(\sigma)$ for all sets of systems coupled by the coupling topology G. In physical terms, $R_G(\rho)$ contains all those density matrices σ which are energetically indistinguishable from ρ for any Hamiltonian respecting the coupling topology G. Developing a good mathematical and physical understanding of $R_G(\rho)$ is an extremely challenging and interesting problem in quantum information science. Promising preliminary work on this problem has been done in [39,40], but much remains to be done.

The thermodynamic limit. In the thermodynamic limit of a large number of systems, the energy difference E_{tot} between the maximal and minimal energies in the system typically tends toward infinity. Recall that the results obtained in this paper typically bound $\Delta E / E_{\mathrm{tot}}$ above by some measure of long-range correlation, where ΔE is the energy gap. Since E_{tot} tends to infinity in the thermodynamic limit, it follows that our results do not give interesting information in this limit, except in the case where we require exact ground states, i.e. $F=1$. It would be extremely interesting to develop more powerful results relating the gap to long-range correlations and entanglement in the thermodynamic limit.

Connection between the gap and the range of correlations. We have "long range" to mean entanglement or correlation between parts of a system that are not directly coupled. Of course, we expect there will be substantial differences between a situation where two subsystems are close, e.g., have perhaps a single spin mediating their indirect interaction, and cases where the interaction is much more indirect, e.g., the left- and right-hand ends of a linear chain, with a large block of intermediate spins mediating the interaction

between the two ends. We expect that the latter case will impose much more stringent restrictions on the size of the gap than the former case. Preliminary numerical investigations with the Heisenberg model bear this out, and further investigations are currently underway.

In conclusion, we have used the techniques of quantum information science to develop connections between the energy gap and long-range correlations and entanglement in the ground states of many-body quantum systems. We believe that the techniques of quantum information science will, more generally, be a powerful tool for understanding and predicting the properties of complex quantum systems.

ACKNOWLEDGMENTS

We thank Dave Bacon, Andrew Doherty, Gerardo Oritz, John Preskill, Guifre Vidal and, especially, Nick Bonesteel, for interesting and enjoyable discussions.

[1] H. L. Haselgrove, M. A. Nielsen, and T. J. Osborne, Phys. Rev. Lett. **91**, 210401 (2003).
[2] G. Vidal, e-print quant-ph/0301063.
[3] J. I. Latorre, E. Rico, and G. Vidal, e-print quant-ph/0304098 (2003).
[4] T. Tessier, I. H. Deutsch, A. Delgado, and I. Fuentes-Guridi, e-print quant-ph/0306015 (2003).
[5] T. A. Costi and R. H. McKenzie, e-print quant-ph/0302055 (2003).
[6] A. P. Hines, R. H. McKenzie, and G. J. Milburn, Phys. Rev. A **67**, 013609 (2003).
[7] A. Osterloh, L. Amico, C. Falci, and R. Fazio, Nature (London) **416**, 608 (2002).
[8] T. J. Osborne and M. A. Nielsen, Phys. Rev. A **66**, 032110 (2002).
[9] S. Scheel, J. Eisert, P. L. Knight, and M. B. Plenio, e-print quant-ph/0207120 (2002).
[10] X. Wang and P. Zanardi, Phys. Lett. A **301**, 1 (2002).
[11] K. M. O'Connor and W. K. Wootters, Phys. Rev. A **63**, 052302 (2001).
[12] D. Gunlycke, V. M. Kendon, V. Vedral, and S. Bose, Phys. Rev. A **64**, 042302 (2001).
[13] M. A. Nielsen, Ph.D. thesis, University of New Mexico, 1998, e-print quant-ph/0011036.
[14] D. Aharonov and A. Ta-Shma, in *Proceedings of the 35th STOC*, 2003 (Association for Computing Machinery, New York, 2003), pp. 20–29.
[15] J. Kempe and O. Regev, Quantum Inf. Comput. **3**, 258 (2003).
[16] D. Aharonov and T. Naveh, e-print quant-ph/0210077.
[17] E. Farhi,J. Goldstone, S. Gutmann, J. Lapan, A. Lundgren, and D. Preda, Science **292**, 474 (2001).
[18] S. Sachdev, *Quantum Phase Transitions* (Cambridge University Press, Cambridge, England, 1999).
[19] S. L. Sondhi, S. M. Girvin, J. P. Carini, and D. Shahar, Rev. Mod. Phys. **69**, 315 (1997).
[20] M. B. Hastings, e-print cond-mat/0305505.
[21] R. V. Lange, Phys. Rev. **146**, 301 (1966).
[22] J. Goldstone, Nuovo Cimento **19**, 154 (1961).
[23] J. Goldstone, S. Weinberg, and A. Salam, Phys. Rev. **127**, 965 (1962).
[24] A. Auerbach, *Interacting Electrons and Quantum Magnetism*, Graduate Texts in Contemporary Physics (Springer-Verlag, New York, 1994).
[25] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge England, 2000).
[26] J. Preskill, Physics 229: Advanced Mathematical Methods of Physics—Quantum Computation and Information, California Institute of Technology, Pasadena, CA, 1998 (unpublished), www.theory.caltech.edu/people/preskill/ph229
[27] T. J. Osborne, Ph.D. thesis, The University of Queensland, 2002.
[28] M. A. Nielsen, Sci. Am. **287**, 66 (2002).
[29] J. Preskill, J. Mod. Opt. **47**, 127 (2000).
[30] D. Aharonov, Phys. Rev. A **62**, 062311 (1999).
[31] M. A. Nielsen, California Institute of Technology Technical Report, 1999, available online at www.qinfo.org/talks/
[32] A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and Its Applications* (Academic, New York, 1979).
[33] M. A. Nielsen and G. Vidal, Quantum Inf. Comput. **1**, 76 (2001).
[34] R. Bhatia, *Matrix Analysis* (Springer-Verlag, New York, 1997).
[35] P. M. Alberti and A. Uhlmann, *Stochasticity and Partial Order: Doubly Stochastic Maps and Unitary Mixing* (Dordrecht, Boston, 1982).
[36] A. V. Thapliyal, Phys. Rev. A **59**, 3336 (1999).
[37] A. Peres, Phys. Lett. A **202**, 16 (1995).
[38] M. A. Nielsen, Phys. Rev. Lett. **83**, 436 (1999).
[39] N. Linden, S. Popescu, and W. K. Wootters, Phys. Rev. Lett. **89**, 207901 (2002).
[40] N. Linden and W. K. Wootters, Phys. Rev. Lett. **89**, 277906 (2002).