Quantum Simulation of Many-Body Hamiltonians Using Perturbation Theory with Bounded-Strength Interactions

Sergey Bravyi,[1] David P. DiVincenzo,[1] Daniel Loss,[2] and Barbara M. Terhal[1]
[1]IBM Watson Research Center, P.O. Box 218, Yorktown Heights, New York 10598 , USA
[2]Department of Physics, University of Basel, Klingelbergstrasse 82, CH-4056 Basel, Switzerland
(Received 2 April 2008; published 15 August 2008)

We show how to map a given n-qubit target Hamiltonian with bounded-strength k-body interactions onto a simulator Hamiltonian with two-body interactions, such that the ground-state energy of the target and the simulator Hamiltonians are the same up to an extensive error $O(\epsilon n)$ for arbitrary small ϵ. The strength of the interactions in the simulator Hamiltonian depends on ϵ and k but does not depend on n. We accomplish this reduction using a new way of deriving an effective low-energy Hamiltonian which relies on the Schrieffer-Wolff transformation of many-body physics.

DOI: 10.1103/PhysRevLett.101.070503 PACS numbers: 03.67.Ac, 31.15.am, 89.70.Eg

A variety of theoretical models studied in condensed matter physics and quantum computation theory deals with ground states of spin Hamiltonians involving *many-body* interactions, i.e., interactions affecting three or more spins at a time. For example, quantum loop models [1] describing topological quantum order require Hamiltonians with four-spin interactions. The 2D cluster state used as a resource for the measurement-based quantum computation can be represented as the ground state of a Hamiltonian with five-spin interactions [2]. Another group of examples includes Hamiltonians for adiabatic quantum computation [3] and constructions of quantum-computationally-hard Hamiltonians [4–6] which rely on the mapping between quantum circuits and Hamiltonians proposed by Kitaev [4].

Given this multitude of interesting models, an important question is whether it is possible to design realistic quantum Hamiltonians with only two-spin interactions whose ground-state properties (such as the ground-state energy, 2-point correlation functions, spectral gap above the ground state, etc.) simulate one of the more complicated Hamiltonians with multispin interactions. Outstanding recent examples of using perturbative techniques for engineering interesting Hamiltonians are Kitaev's honeycomb lattice model [7], or the use of perturbation theory in deriving effective Hamiltonians for cold atoms or molecules in optical lattices (see, e.g., [8]). A demonstration of how such simulations can be achieved rigorously is due to Kempe, Kitaev, and Regev [5] who developed a technique known as perturbation theory gadgets (PTGs).

While in theoretical physics perturbation theory is used to derive an effective low-energy Hamiltonian starting from the ''full'' Hamiltonian (including high-energy degrees of freedom), PTGs use the perturbation theory in the reverse direction. One starts from a target Hamiltonian H_{target} chosen for some interesting ground-state property. Then a high-energy simulator Hamiltonian H is to be designed such that H_{target} can be obtained from H by perturbation theory as a low-energy effective Hamiltonian. The main objective of PTGs is to make the simulator Hamiltonian H as simple and realistic as possible while retaining the interesting ground-state properties of H_{target}. For example, the PTGs have been used in [9] to prove universality of quantum adiabatic computation with local two-body Hamiltonians on a 2D square lattice. For more recent developments, see [10].

Although PTGs are a flexible and powerful technique, it has a major shortcoming which we will address in this Letter. This shortcoming is the unphysical scaling of parameters of the simulator Hamiltonian, such as the norm of spin-spin interactions growing polynomially with the system size n. Technically this unphysical scaling arises from the well-known convergence criterion for perturbative series: the norm of a perturbation $\|V\|$ must be small compared to the spectral gap Δ of the unperturbed Hamiltonian. The convergence criterion forces Δ to scale with system size n, since $\|V\|$ typically scales with n; this is what has been done in [5,9,10].

If the applicability of perturbation theory were limited by convergence of the perturbative series, its common use in physics would be unwarranted. The lack of convergence for quantum electrodynamics was argued on physical grounds by Dyson [11]. There is a widespread belief that the general perturbative series for quantum field theory and many-body physics do not converge but are to be viewed as an asymptotic series, meaning that their lowest-order terms are a good approximation to the quantity of interest while inclusion of higher-order terms may actually give a worse result (see, e.g., [12]).

In this Letter we make these beliefs more rigorous by developing a formalism that justifies application of perturbation theory in the regime $\|V\| \gg \Delta$. This formalism is applicable to many-body Hamiltonians that possess certain locality properties. A Hamiltonian H describing a system of n qubits will be called ''k local'' if and only if it is represented as a sum of local interactions $H = \sum_i H_i$ such that each operator H_i acts on some subset of k or less qubits. Throughout this Letter we assume that k is a

constant independent of n. The interaction strength of a k-local Hamiltonian is defined as the largest norm of the local interactions, $J = \max_i \|H_i\|$. In addition, the Hamiltonians in this Letter will have the important property that, for a growing system size, the number of interactions in which any one-qubit participates does not grow but stays constant. For convenience we shall refer to this condition as a bounded degree. It is generically fulfilled for Hamiltonians with short-range interactions studied in physics. An example is the standard (2-local) Heisenberg Hamiltonian on a lattice $H = -\sum_{i,j} J(r_{ij})(X_iX_j + Y_iY_j + Z_iZ_j)$; the fact that the coupling $J(r_{ij})$ is bounded range, $J(r_{ij} > r) = 0$, ensures that each spin is acted upon by a constant number of terms. The following simulation theorem is the main result of this Letter.

Theorem 1.—Let H_{target} be a k-local Hamiltonian acting on n qubits with interaction strength J and bounded degree. For any fixed precision ϵ one can efficiently construct a 2-local simulator Hamiltonian H acting on $O(n)$ qubits with interaction strength $O(J)$ such that the ground-state energy of H approximates the ground-state energy of H_{target} with an absolute error at most ϵJn.

Theorem 1 eliminates the essential shortcoming of the earlier PTG constructions [5,9,10] mentioned above, namely, the unphysical scaling of the interaction strength in the simulator Hamiltonian H. In our construction the interaction strength of H is bounded by a constant (depending only on k, ϵ, and J) which makes it more physical and, in principle, implementable in a lab. The price we pay for this improvement is that we are able to reproduce the ground-state energy only up to an extensive error ϵJn. For all realistic physical Hamiltonians, the ground-state energy itself is generically proportional to nJ, and thus the relative error can be made arbitrarily small. Note that this price, i.e., the presence of an extensive simulation error, is not a function of what perturbative expansion or what perturbation gadgets one uses, but is a generic feature of perturbation theory in the regime when $\|V\| \gg \Delta$. This can be easily understood by considering an example of n disconnected systems, each of which is individually analyzed using perturbation theory. The smallest eigenvalue of the total system will pick up the error in each individual perturbative expansion and hence will be $O(\epsilon n)$, where ϵ is the perturbation parameter.

Techniques.—Let us sketch the proof of the theorem. The simulator Hamiltonian H will act on a Hilbert space describing two species of qubits: n system qubits that are the qubits of the target Hamiltonian, and $O(n)$ mediator qubits that mediate interactions between system qubits. The simulator H is constructed using the perturbation theory gadgets introduced in [9]. "Gadget" is a technical term used broadly in theoretical computer science; in the present application, a gadget is simply a mediator qubit, and a Hamiltonian coupling the mediator qubit with some small subset of system qubits. For every mediator qubit u we define a projector onto its low-energy subspace $P^u = |0\rangle\langle 0|_u$ and its high-energy subspace $Q^u = |1\rangle\langle 1|_u$. The purpose of a gadget is to simulate some particular k-spin interaction H^u_{target} in the decomposition $H_{\text{target}} = \sum_u H^u_{\text{target}} + H_{\text{else}}$. Here H_{else} are additional terms in the target Hamiltonian that we do not wish to treat using perturbation theory (since they are already 2 local, for example).

The simulation is achieved by applying perturbation theory to each gadget individually. A gadget's simulator Hamiltonian is $H^u = H^u_0 + V^u$, where $H^u_0 = \Delta Q^u$ penalizes the mediator qubit for being in the state $|1\rangle$, and V^u is a perturbation. With the proper choice of V^u the effective Hamiltonian on the low-energy subspace, in which the mediator qubit u is in the state $|0\rangle$, approximates H^u_{target} with an error ϵ. Furthermore, this effective Hamiltonian can be obtained from H^u via an approximate Schrieffer-Wolff transformation—a unitary transformation e^{S^u} that brings the Hamiltonian H^u into a block-diagonal form (up to a small error); that is, $P^u e^{S^u} H^u e^{-S^u} Q^u = O(\epsilon J)$ and $P^u e^{S^u} H^u e^{-S^u} P^u = H^u_{\text{target}} + O(\epsilon J)$. The transformation e^{S^u} is generated by some anti-Hermitian operator S^u having block-off-diagonal form; i.e., $P^u S^u P^u = Q^u S^u Q^u = 0$. (The Schrieffer-Wolff transformation proposed originally in [13] refers to a particular perturbative method of deriving a low-energy effective Hamiltonian.)

By combining the local gadgets together, we get as a candidate for the simulator Hamiltonian $H = H_0 + V + H_{\text{else}}$, with $H_0 = \sum_u H^u_0$ and $V = \sum_u V^u$. Let $\lambda(H)$ and $\lambda(H_{\text{target}})$ be the ground-state energy of H and H_{target}, respectively. We prove that $\lambda(H)$ approximates $\lambda(H_{\text{target}})$ with a small extensive error by constructing a global unitary transformation e^S mapping H to H_{target} (with a small extensive error); that is, $H_{\text{target}} = Pe^S H e^{-S} P + O(\epsilon nJ)$, where $P = \bigotimes_u P^u$ projects onto the subspace in which every mediator qubit is in the state $|0\rangle$. Given such a transformation one immediately gets an upper bound $\lambda(H) = \lambda(e^S H e^{-S}) \le \lambda(Pe^S H e^{-S} P) = \lambda(H_{\text{target}}) + O(\epsilon nJ)$. Here we have taken into account that restricting a Hamiltonian on a subspace can only increase its ground-state energy [14]. Making a natural choice $S = \sum_u S^u$ we prove that $Pe^S H e^{-S} P$ contains the desired term H_{target} and some cross-gadget terms where S^u acts on $H^v_0 + V^v$ with $u \neq v$. Using the "independence" properties of the gadgets, the block-off diagonality of S^u, and Lemmas 1 and 2 stated below, we are able to show that the contribution of these cross-gadget terms is small enough to be absorbed into the error term $O(n\epsilon J)$ (see Sec. IIb).

In order to prove a matching lower bound, choose any gadget u and consider the transformed Hamiltonian $\tilde{H}^u = e^{S^u} H^u e^{-S^u}$, where $H^u = H^u_0 + V^u$. We prove an operator inequality $\tilde{H}^u \ge I^u \otimes H^u_{\text{target}} + O(\epsilon J)$. Here I^u is the identity operator acting on the mediator qubit u, and $O(\epsilon J)$ stands for some operator with norm $O(\epsilon J)$. Intuitively one should expect this inequality to be true since the P block of $\tilde{H}^u$ approximates H^u_{target} with an error $O(\epsilon J)$, the Q block

of $\tilde{H}^u$ contains a large energy penalty Δ, and the off-diagonal blocks $P\tilde{H}^u Q$ are small by the definition of the Schrieffer-Wolff transformation. Using the unitarity of e^{S^u} and the smallness of S^u, we transform the above inequality into $H^u \geq I^u \otimes H^u_{\text{target}} + O(\epsilon J)$, which implies $H \geq I \otimes H_{\text{target}} + O(n\epsilon J)$ and thus gives $\lambda(H) \geq \lambda(H_{\text{target}}) + O(n\epsilon J)$. These arguments are fleshed out in Sec. IIa. Our results will be stated for two different mappings $H_{\text{target}} \rightarrow H$ corresponding to two different gadgets, the one reducing the locality parameter k by a factor of 2, and the other reducing $k = 3$ to $k = 2$. By composing these mappings we arrive at Theorem 1.

Now we state the two Lemmas used in the proof. The two Lemmas together can be regarded as an infinitesimal version of the Lieb-Robinson bound that governs time evolution of a local observable under a local Hamiltonian (see, e.g., [15]).

Lemma 1.—Let S be an anti-Hermitian operator. Define a superoperator L such that $L(X) = [S, X]$ and $L^0(X) = X$. For any operator H and integer k define $r_k(H) = \|e^S H e^{-S} - \sum_{p=0}^{k-1} \frac{1}{p!} L^p(H)\|$ if $k \geq 1$ and $r_0(H) = \|e^S H e^{-S}\| = \|H\|$. Then for any $k \geq 0$ one has $r_k(H) \leq \frac{1}{k!}\|L^k(H)\|$.

Using this Lemma one can show the following.

Lemma 2.—Let S and H be any $O(1)$-local operators acting on n qubits with a bounded degree and interaction strengths J_S and J_H, respectively. Then for any $k = O(1)$ one has $\|L^k(H)\| = O(n \cdot J_S^k J_H)$.

The proofs are rather elementary and can be found in [16].

We shall use two types of gadgets proposed in [9], namely, the subdivision gadget and the 3-to-2-local gadget. The former will be used to break k-local interactions down to 3-local interactions, while the latter breaks 3-local interactions into 2-local interactions. Note that in the description of these two gadgets we shall often omit the label u.

Subdivision gadget.—Let the target Hamiltonian be a single k-qubit interaction $H_{\text{target}} = JAB$, where A, B act on nonoverlapping subsets of $\lceil k/2 \rceil$ or less qubits and $\|A\|, \|B\| \leq 1$. Introduce one mediator qubit u, choose a parameter $\Delta \gg J$, and define the simulator Hamiltonian $H = H_0 + V$, with

$$H_0 = \Delta|1\rangle\langle 1|_u, \quad V = \sqrt{\Delta J/2} X_u \otimes (-A + B) + V_{\text{extra}}. \quad (1)$$

Here $V_{\text{extra}} = (J/2)(A^2 + B^2)$ acts trivially on the mediator qubit. Note that H contains only $(\lceil k/2 \rceil + 1)$-body interactions. The purpose of the term V is to induce transitions $|0\rangle_u \rightarrow |1\rangle_u \rightarrow |0\rangle_u$ in the second order of perturbation theory, such that the corresponding effective Hamiltonian is proportional to $(-A + B)^2$ containing the desired term AB and unwanted terms A^2, B^2, which we cancel by V_{extra}. Next we define an approximate Schrieffer-Wolff transformation:

$$S = -iJ^{1/2}(2\Delta)^{-1/2} Y_u \otimes (-A + B). \quad (2)$$

A straightforward calculation utilizing Lemma 1 shows that

$$e^S H e^{-S} = \left(H + [S, H] + \frac{1}{2}[S, [S, H]]\right) + O(J^{3/2}\Delta^{-1/2}) = \begin{bmatrix} H_{\text{target}} & 0 \\ 0 & \Delta I + O(J) \end{bmatrix} + O(J^{3/2}\Delta^{-(1/2)}), \quad (3)$$

where we used that $[S, V_{\text{extra}}] = 0$. The upper and lower blocks correspond to the subspaces P and $Q = I - P$, respectively. Thus, $Pe^S H e^{-S} P$ is close to $H_{\text{target}} = JAB$, as desired; the error can be made $O(\epsilon J)$ by choosing $\Delta = J\epsilon^{-2}$.

3-to-2-local gadget.—Let the target Hamiltonian be a single 3-body interaction, $H_{\text{target}} = JABC$, where A, B, C are one-qubit operators acting on different qubits and $\|A\|$, $\|B\|$, $\|C\| \leq 1$. Introduce one mediator qubit u, choose $\Delta \gg J$, and define the simulator Hamiltonian $H = H_0 + V$ with $H_0 = \Delta|1\rangle\langle 1|_u$, $V = V_d + V_{od} + V_{\text{extra}}$ with

$$V_d = -\Delta^{2/3} J^{1/3} |1\rangle\langle 1|_u \otimes C, \; V_{\text{od}} = \frac{\Delta^{2/3} J^{1/3}}{\sqrt{2}} X_u \otimes (-A + B), \quad (4)$$

and $V_{\text{extra}} = \Delta^{1/3} J^{2/3}(-A + B)^2/2 + J(A^2 + B^2)C/2$. Note that H contains only 2-body interactions. The purpose of the term V is to induce transitions $|0\rangle_u \rightarrow |1\rangle_u \rightarrow |1\rangle_u \rightarrow |0\rangle_u$ in the third order of perturbation theory, such that the corresponding contribution to the effective Hamiltonian is proportional to $(-A + B)^2 C$ which coincides with ABC up to some unwanted terms which are canceled by V_{extra}. We define an approximate Schrieffer-Wolff transformation $S = -ix2^{-1/2}O$, where

$$O = Y_u \otimes (-A + B)\{I + xC + x^2[C^2 - \tfrac{2}{3}(-A + B)^2]\} \quad (5)$$

with $x \equiv J^{1/3}\Delta^{-1/3}$. We calculate the effective Hamiltonian $Pe^S H e^{-S} P$. Let us first estimate an error resulting from cutting off the expansion; see Lemma 1. Recalling that $L = [S, \cdot]$ one gets

$$\|Pe^S H e^{-S} P - P[H + L(H) + \tfrac{1}{2}L^2(H)]P\| \leq \tfrac{1}{6}\|PL^3(H)P\| + \tfrac{1}{24}\|L^4(H)\|. \quad (6)$$

We note that $PL^3(H)P = PL^3(V_{\text{od}})P$ since each application of S flips the mediator qubit and a nonzero contribution comes only from the terms with an even number of flips. Using a bound $\| S \| = O(x)$, see Eq. (5), one can upper-bound the right-hand side of Eq. (6) as $O(J^{4/3}\Delta^{-1/3})$. A direct but lengthy calculation shows that

$$e^S H e^{-S} = H + L(H) + \frac{1}{2}L^2(H) + \frac{1}{6}L^3(H) + O(J^{4/3}\Delta^{-1/3}) = \begin{bmatrix} H_{\text{target}} & 0 \\ 0 & \Delta I + O(\Delta^{2/3}J^{1/3}) \end{bmatrix} + O(J^{4/3}\Delta^{-1/3}), \quad (7)$$

where $H_{\text{target}} = JABC$, as desired. The error $O(J^{4/3}\Delta^{-1/3})$ in H_{target} can be made $O(\epsilon J)$ by choosing $\Delta = J\epsilon^{-3}$.

IIa: Combining the gadgets together.—Let $H_{\text{target}} = \sum_u H^u_{\text{target}} + H_{\text{else}}$, where each term H^u_{target} can be dealt with using one of the gadgets described above. The simulator Hamiltonian is $H = \sum_u H^u + H_{\text{else}}$, where H^u is the simulator constructed for a gadget u as above. Using Eqs. (3) and (7) one gets the inequality $e^{S^u} H^u e^{-S^u} \geq I^u \otimes H^u_{\text{target}} + O(\epsilon J)$, which yields $H^u \geq e^{-S^u}(I^u \otimes H^u_{\text{target}})e^{S^u} + O(\epsilon J)$. Applying Lemma 1, one gets $\|e^{-S^u}(I^u \otimes H^u_{\text{target}})e^{S^u} - I^u \otimes H^u_{\text{target}}\| \leq \|[S^u, I^u \otimes H^u_{\text{target}}]\| = O(\epsilon J)$; see Eqs. (2) and (5). It follows that $H^u \geq I^u \otimes H^u_{\text{target}} + O(\epsilon J)$. Summing up these inequalities over all gadgets one arrives at $H \geq I \otimes H_{\text{target}} + O(\epsilon nJ)$, where I acts on the mediator qubits. Thus

$$\lambda(H) \geq \lambda(H_{\text{target}}) + O(\epsilon nJ). \quad (8)$$

IIb: Bounding cross-gadget contributions.—Let $P = \bigotimes_u P^u$ be the projector on the subspace in which every mediator qubit is in the state $|0\rangle$. Define $S = \sum_u S^u$ where S^u is constructed using Eqs. (2) and (5). Note that for both gadgets S is a $O(1)$-local operator with bounded degree and interaction strength $O(\epsilon)$. We shall prove that

$$\|Pe^S H e^{-S} P - H_{\text{target}}\| = O(\epsilon nJ). \quad (9)$$

Then one can get an upper bound on $\lambda(H)$ by observing that

$$\begin{aligned} \lambda(H) &= \lambda(e^S H e^{-S}) \\ &\leq \lambda(Pe^S H e^{-S} P) \\ &= \lambda(H_{\text{target}}) + O(\epsilon nJ). \end{aligned}$$

Combining Eq. (8) with this upper bound, one gets $|\lambda(H) - \lambda(H_{\text{target}})| = O(\epsilon nJ)$, which is the desired result. It remains to prove Eq. (9). First, the contribution of all commutators of S and H_{else} to $e^S H e^{-S}$ can be bounded by $O(\epsilon nJ)$; see Lemmas 1 and 2. To bound the cross-gadget terms we exploit two important properties (valid for both gadgets): (i) S^u always flips a mediator qubit u; (ii) For $u \neq v$ one has $[P^u, H^v] = 0$, $[P^u, S^v] = 0$, and $[H^u_0, S^v] = 0$. This latter property essentially captures the independent action of the local gadgets. Using the properties (i) and (ii) one can identify the two dominant cross-gadget contributions to $Pe^S H e^{-S} P$, namely, $O'_{\text{cg}} \sim P[S^u, [S^u, V^v_{\text{extra}}]]P$ and $O''_{\text{cg}} \sim P[S^u, [S^u, V^v_d]]P$ (the term O''_{cg} appears only for the 3-to-2-local gadget) with $u \neq v$. Using the explicit form of V^v_d, see Eq. (4), one concludes that $O''_{\text{cg}} = 0$. The norm of every term O'_{cg} can be bounded as $O(\epsilon^2 J)$ and $O(\epsilon J)$ for the subdivision and the 3-to-2-local gadgets, respectively. Finally, the number of the terms O'_{cg} is $O(n)$ because of the bounded degree assumption. It proves Eq. (9). The details of this derivation which are otherwise not very insightful can be found at [16].

S. B., D. P. D., and B. M. T. acknowledge support by DTO through ARO Contract No. W911NF-04-C-0098. D. L. acknowledges support from the Swiss NSF.

[1] A. Kitaev, Ann. Phys. (N.Y.) **303**, 2 (2003); P. Fendley, arXiv.org:0711.0014.
[2] R. Raussendorf, D. Browne, and H. Briegel, Phys. Rev. A **68**, 022312 (2003).
[3] D. Aharonov *et al.*, SIAM J. Comput. **37**, 166 (2007).
[4] A. Kitaev, A. Shen, and M. Vyalyi, *Classical and Quantum Computation*, Graduate Studies in Mathematics Vol. 47 (AMS, Providence, RI, 2002).
[5] J. Kempe, A. Kitaev, and O. Regev, SIAM J. Comput. **35**, 1070 (2006).
[6] D. Aharonov *et al.*, in *Proceedings of the 48th Annual Symposium on FOCS* (IEEE, Washington D.C., 2007), pp. 373–383.
[7] A. Kitaev, Ann. Phys. (N.Y.) **321**, 2 (2006).
[8] L. -M. Duan, E. Demler, and M. D. Lukin, Phys. Rev. Lett. **91**, 090402 (2003); A. Micheli, G. K. Brennen, and P. Zoller, Nature Phys. **2**, 341 (2006).
[9] R. I. Oliveira and B. M. Terhal, arXiv:quant-ph/0504050 [Quantum Inf. Comput. (to be published)].
[10] S. Bravyi *et al.*, Quantum Inf. Comput. **8**, 361 (2008); S. Jordan and E. Farhi, Phys. Rev. A **77**, 062329 (2008); N. Schuch and F. Verstraete, Phys. Rev. A **78**, 012352 (2008); J. Biamonte and P. Love, arXiv:0704.1287.
[11] F. J. Dyson, Phys. Rev. **85**, 631 (1952).
[12] M. Reed and B. Simon, *Methods of Modern Mathematical Physics IV: Analysis of Operators* (Academic Press, San Diego, CA, 1978).
[13] J. R. Schrieffer and P. A. Wolff, Phys. Rev. **149**, 491 (1966).
[14] By abuse of notation, $Pe^S He^{-S}P$ denotes the restriction of the operator $e^S He^{-S}$ onto the subspace P.
[15] M. B. Hastings and T. Koma, Commun. Math. Phys. **265**, 781 (2006); S. Bravyi, M. B. Hastings, and F. Verstraete, Phys. Rev. Lett. **97**, 050401 (2006).
[16] S. Bravyi, D. P. DiVincenzo, D. Loss, and B. M. Terhal, arXiv:0803.2686.