Preparing Ground States of Quantum Many-Body Systems on a Quantum Computer

David Poulin[1,*] and Pawel Wocjan[2]

[1]*Département de Physique, Université de Sherbrooke, Sherbrooke, Québec, Canada*

[2]*School of Electrical Engineering and Computer Science, University of Central Florida, Orlando, Florida, USA*

(Received 17 September 2008; published 3 April 2009)

Preparing the ground state of a system of interacting classical particles is an *NP*-hard problem. Thus, there is in general no better algorithm to solve this problem than exhaustively going through all N configurations of the system to determine the one with lowest energy, requiring a running time proportional to N. A quantum computer, if it could be built, could solve this problem in time $\sqrt{N}$. Here, we present a powerful extension of this result to the case of interacting *quantum* particles, demonstrating that a quantum computer can prepare the ground state of a quantum system as efficiently as it does for classical systems.

DOI: 10.1103/PhysRevLett.102.130503 PACS numbers: 03.67.Ac, 03.67.Lx, 24.10.Cn

The simulation of quantum many-body systems is a notoriously hard problem in condensed matter physics, but it could easily be handled by a quantum computer [1]. There is, however, one catch: while a quantum computer can naturally implement the dynamics of a quantum system, i.e., solve Schrödinger's equation, there was until now no general method to initialize the computer in a physically relevant state of the simulated system.

For most physical applications, we are interested in the low-energy eigenstates of the Hamiltonian H because they describe the most interesting phases of matter, e.g., ferromagnetism, superconductivity, quantum Hall effect, and Bose-Einstein condensation to name a few. Unfortunately, preparing low-energy states is already a very difficult task even when H describes a classical system.

Indeed, this problem is an archetype of the complexity class *NP*. This class contains all decision problems, i.e., problems of the form "Does x satisfy the property $\mathcal{L}$?," such that when the answer is yes, there exists a witness w that can be used to prove this answer efficiently. More precisely, for each x there exists a polynomial-size verification circuit V_x such that (1) when $x \in \mathcal{L}$ there exists a witness w that will cause V_x to output 1, and (2) when $x \notin \mathcal{L}$, all witnesses cause V_x to output 0.

Consider for instance a local Ising model $H(\{\sigma\}) = \sum_{i<j} J_{ij}\sigma_i\sigma_j + \sum_i h_i\sigma_i$ on n spins $\sigma_i \in \{0, 1\}$. "Is there a spin configuration σ of energy less than E?" is a problem in *NP*. Indeed, when the answer is yes, the configuration σ that achieves this low energy can serve as a witness. Verifying the answer boils down to computing the energy, which requires at most n^2 operations. Finally, when the answer is no, there is no configuration that can cause the verification procedure to accept. Clearly, an algorithm that solves this problem can be used to determine the ground state energy with little overhead.

The Ising problem is in fact *NP* complete, meaning that it is the hardest problem in the class [2]. Even more surprising is the fact that approximating the energy of the system (imbedded not in any finite dimension but on a sparse interaction graph) with an error that increases with the system size n is just as hard as the exact case—it is also *NP* complete. This is a consequence of a famous theorem on probabilistically checkable proofs [3,4]. Although some special cases can be solved efficiently [5–7], there is in general no better algorithm to solve the Ising problem than systematically going through all $N = 2^n$ spin configurations to determine the one with lowest energy.

In addition to the immediate physical context, finding ground states provides a very natural setting for studying combinatorial optimization problems—problems that consist in minimizing an objective function H (playing the role of energy) over some configuration space. Optimization problems play a vital role in almost every branch of science, from computer science to statistical physics and computational biology [8]. Determining a solution by exhaustive search is, in general, computationally prohibitive because the size N of the search space grows exponentially with the input size. Given the practical importance of optimization problems, more efficient methods are highly desirable.

A common strategy to solve optimization problems is simulated annealing [9]. As its name suggests, this method imitates the process undergone by a metal that is heated to a high temperature and then slowly cooled to its configuration of lowest energy. If the cooling process is too fast, the system can become trapped in a local minimum, resulting in a failure of the algorithm. When the cooling is sufficiently slow, however, thermal fluctuations should prevent this phenomenon from occurring. Thus, simulated annealing requires a detailed knowledge of the energy landscape and therefore cannot be applied to all minimization problems. It was shown recently that a simulated annealing algorithm operated on a quantum computer achieves a quadratic speed-up over classical annealing [10]. Whether the method can minimize the energy of a quantum system as efficiently is unknown.

Adiabatic quantum computation is another method to tackle this class of problems with a quantum computer. The

adiabatic theorem asserts that a system prepared in the instantaneous ground state of a Hamiltonian that varies slowly in time will remain in the ground state. The ground state of H can thus be prepared by choosing a time-dependent Hamiltonian with a simple initial ground state and slowly changing it to H. This algorithm was applied to randomly generated instances of an NP-complete problem [11]. The algorithm worked well for the small examples that could be simulated on a classical computer. It was later shown [12], however, that the particular interpolation scheme suggested [11] fails for satisfiability problems, and the best known upper bound is a running time of N [12] (polylogarithmic corrections are ignored throughout). In principle, this technique can also be applied to minimize the energy of a quantum Hamiltonian, but little is known about its performances in that case.

Finally, Grover's algorithm [13] can find the ground state of a classical system in $\sqrt{N}$ steps. Given a projector R and a state ψ with $\|R|\psi\rangle\|^2 = q > 0$, Grover's algorithm consists of a sequence of two reflections, $I - 2R$ and $I - 2|\psi\rangle\langle\psi|$. Repeating this sequence $1/\sqrt{q}$ times has the effect of projecting ψ onto the image of R plus a small correction. The correction can be completely ignored if we end Grover's algorithm by a measurement that distinguishes R from its orthogonal complement, and start over when the complement is obtained. This increases the running time only by a constant factor. Note that the value of q must be approximately known, and this can be achieved by quantum counting [14]. Choosing R to be the projector on $H < E$ and ψ a uniform superposition of all spin configurations yields, after at most $\sqrt{N}$ iterations, a state of energy less than E. The ground state is obtained by "sweeping" the value of E. Although this remains an exponential scaling, it is significantly faster than a brute force search, and there are indications that this scaling is optimal [15].

At first sight, it seems like this last technique could be used to find the ground state of a quantum many-body system just as well. All that is needed is a method to implement a projector R onto the low-energy states of the system, i.e., $H < E$ for some given E. Combining this method with Grover's algorithm on an initial random state would create the desired outcome with high probability. In fact, it is not necessary to initialize the system in a truly random state, but instead it can be randomly selected among all *stabilizer* states. These have all the essential properties of random states and, most importantly, can be prepared with at most n^2 operations [16].

Unfortunately, there is no known procedure to implement the projector on $H < E$ exactly. Using the phase estimation algorithm [17], however, we can obtain an approximation of the projector. Recall that the phase estimation can be used to estimate the energy of any local Hamiltonian with a polynomial small error and failure probability. To describe this method, it is convenient to assume that H has been normalized such that $\|H\| < 1/2$ and to consider its spectral decomposition, $H|a\rangle = \varphi_a|a\rangle$. The phase estimation algorithm uses k auxiliary qubits initially in the state $|0_k\rangle$. These qubits are placed in a uniform superposition of all bases by Hadamard transform and are used to control the evolution time of the system, thereby mapping $|a\rangle \otimes \frac{1}{\sqrt{2^k}}\sum_j |j\rangle$ to $|a\rangle \otimes |\varphi_a\rangle$ via phase kickback, where

$$|\varphi_a\rangle = \frac{1}{\sqrt{2^k}} \sum_j e^{-i2\pi\varphi_a j}|j\rangle. \tag{1}$$

These are "momentum" states, so the value of φ_a can be estimated via inverse Fourier transform. Hence, we can implement an approximation R of the projector on $H < E$ by running the phase estimation algorithm and projecting the auxiliary qubits onto the subspace of low momentum. Combining this method with Grover's algorithm should thus yield a good approximation of the ground state.

However, a detailed analysis (see Appendix C [18]) of this "naive" approach reveals a failure. The problem is that the projector R constructed from phase estimation is only an approximation of $H < E$ and errors can build up during the amplification procedure. There are two sources of errors. First, the quantum computer cannot exactly reproduce the dynamics of the many-body system. This is not a problem, however, since a $1/\mathrm{poly}(n)$ accuracy can be achieved using a Trotter-Suzuki decomposition at a polynomial cost [1], and this error does not build up (see Appendix A [18] for a detailed proof). We will henceforth safely ignore this source of error.

Second, the inverse Fourier transform is discrete while the energies φ_a take value from a continuum, producing unavoidable round-off errors. Even when the energy φ_a associated with $|a\rangle$ is well above the acceptance threshold E, there is a small probability that phase estimation will diagnose it as being smaller than E. It is these imperfections that cause the algorithm to fail because they build up during amplification. Detailed knowledge of the energy landscape—such as the presence of an energy gap—could be used to circumvent this effect, but in general the method will fail.

We will now present our algorithm that works for all local Hamiltonians. We proceed by making two modifications to the naive algorithm. A detailed analysis is presented in Appendix D of Ref. [18]. The first modification is to run the algorithm backward: we initialize the system qubits in a random state $|\psi\rangle = \sum_a \alpha_a|a\rangle$, the auxiliary qubits in a low momentum state $|\mu\rangle$ [cf. Eq. (1)], and execute the inverse of the phase kickback circuit followed by Hadamard transform. This produces the state

$$|\Phi\rangle = \sum_a \alpha_a \langle\varphi_a|\mu\rangle|a\rangle \otimes |0_k\rangle + \cdots, \tag{2}$$

where the ellipsis represents terms where the auxiliary qubits are in a state orthogonal to $|0_k\rangle$. The factor $|\langle\varphi_a|\mu\rangle|$ is a function of $\mu - \varphi_a$ peaked at 0 with a width 2^{-k}. Thus, we can use Grover's algorithm to amplify the all-zero state of the auxiliary qubits and obtain a state that

is mostly a superposition of those eigenstates of H with eigenvalues close to μ; i.e., the amplitude of each term in the superposition gets reweighted by $|\langle\varphi_a|\mu\rangle|$. This procedure truly acts as a filter, suppressing the amplitude of eigenstates outside its bandwidth for benefit of the eigenstates inside the bandwidth. Moreover, the auxiliary qubits are systematically returned to $|0_k\rangle$ as desired.

Unfortunately, this is still not sufficient for our purpose because the filter has a heavy tail. There are exponentially many states with energy outside the bandwidth, so unless their amplitude is exponentially suppressed, they can significantly shift the energy of the state. The filter we have constructed offers a polynomial suppression; we need a filter that drops more abruptly outside its bandwidth.

This requires a second modification to the naive algorithm and is realized by repeating the phase estimation η times, using a total of ηk auxiliary qubits. We obtain the same state Φ as above [cf. Eq. (2)], except that the factor $\langle\varphi_a|\mu\rangle$ is now raised to the ηth power. For those φ_a that are within $2^{-k}/\sqrt{\eta}$ of μ, this factor is at least $1/2$. Thus, because ψ is a random stabilizer state, the overlap of Φ with the projector $Q = I_n \otimes |0_k\rangle\langle 0_k|^{\otimes\eta}$ will typically be $\|Q|\Phi\rangle\|^2 \geq \frac{m}{2N}$, where m is the number of eigenstates of H in the bandwidth of the filter.

Thus, Grover's algorithm can be used to amplify this overlap to nearly 1 in a time at most $\sqrt{N}$. When the overlap of the state with the filter is too small, i.e., if $\|Q|\Phi\rangle\|^2 < 1/N^2$ say, this step will fail and the algorithm will abort. Choosing $k \sim \log_2(\frac{1}{\epsilon})$ and $\eta \sim n$ yields, after a successful application of Grover's amplification, a superposition of eigenstates of energy $\mu \pm \epsilon$ as desired (see Appendix D [18]). To summarize, this algorithm acts as a filter on the energy. The position μ and width $\epsilon \sim 1/\mathrm{poly}(n)$ of the filter are specified by the user. When no eigenstates of H lie within the filter's bandwidth, the algorithm aborts as desired.

Note that the method can be adapted in a straightforward way to produce thermal distributions of the system at any temperature $T \geq 1/\mathrm{poly}(n)$. We could in a first step combine our method with quantum counting [14] to estimate the density of states $\mathcal{D}(E) = \sum_a \delta(\varphi_a - E)$ with a $1/\mathrm{poly}(n)$ resolution. We could then choose an energy scale E at random according to the distribution $P(E) \sim e^{E/k_B T}\mathcal{D}(E)$ and use our algorithm to prepare a state of energy close to E. Combined with the random choice of initial state, this procedure reproduces the statistical properties of the thermal state.

Analogously to the classical case, determining the ground state energy of a local quantum many-body system within accuracy $1/\mathrm{poly}(n)$ is a complete problem for the complexity class known as quantum Merlin and Artur (QMA) [19]. Whether the problem remains complete when an extensive error is tolerated is unknown, but would be a natural quantum extension of the probabilistically checkable proofs theorem. Indeed, QMA is a natural generalization of NP: it is defined similarly except that both the witness and the verification circuit V_x are quantum mechanical. In addition to the n witness qubits on which it operates, V_x can also make use of $h \sim \mathrm{poly}(n)$ auxiliary qubits initialized in the state 0 that serve as a scratchpad during the computation. The output of the verification procedure is obtained by measuring the first output qubit of the circuit. Because of the intrinsic randomness of quantum mechanics, this procedure is probabilistic: (1) when $x \in \mathcal{L}$, there exists a witness w that will cause V_x to output 1 with probability greater than u, and (2) when $x \notin \mathcal{L}$, all witnesses cause V_x to output 1 with probability less than v where $u - v > 1/\mathrm{poly}(n)$.

The completeness of the local Hamiltonian problem for the class QMA suggests that our algorithm can be used to solve all these problems and prepare the relevant witness in a time $\sqrt{2^n}$. This is not quite right because the mapping to the local Hamiltonian problem does not preserve the size of the witness. The algorithm we suggest instead is a small variation of the previous algorithm, but uses the defining circuit V_x directly instead of reducing to a local Hamiltonian.

A good witness for V_x is a $n + h$ qubit state ψ with (1) all h ancillary qubits in the state 0—summarized by $Q|\psi\rangle = |\psi\rangle$ where Q is the projector onto the all-zero state of the auxiliary qubits—and (2) a probability at least v of outputting 1 at the verification procedure—summarized by $\|R|\psi\rangle\|^2 \geq v$ where R is the projector associated with the verification procedure V_x. Thus, the tasks of preparing a good witness boils down to producing a state that maximizes the overlap with two projectors Q and R. When $[Q, R] = 0$, this task can be accomplished straightforwardly with Grover's algorithm, but additional efforts are required in the general case. Indeed, if we start, say, with a random state in the image of Q and amplify the projector R, we will typically obtain a state that is mostly supported outside the image of Q, unless $[Q, R] = 0$.

The method we propose to solve this problem is a modification of our algorithm that builds on the work of Marriott and Watrous [20]. The main modification is to replace the phase estimation algorithm by the circuit of Fig. 1. It consists of a sequence of k (odd) alternating measurements of R and Q whose outcomes are coherently recorded on k auxiliary qubits initially in the state 0. The behavior of this circuit on an eigenstate $|a\rangle$ of the operator

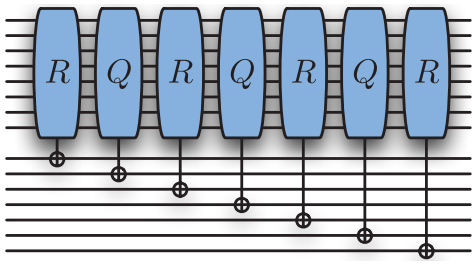

FIG. 1 (color online). The circuit consists of a sequence of measurements of R and Q. The results are coherently imprinted on k auxiliary qubits.

QRQ with eigenvalue p_a can be analyzed using a result of Jordan [21] (see Appendix B [18]). The state of the k auxiliary qubits becomes a superposition of all sequences of 0 and 1, and the amplitude of consecutive distinct outcomes, i.e., the amplitude associated to each ''switches'' from 0 to 1 or vice versa, is $\sqrt{1-p_a}$. Thus, counting the number of switches in the measurement outcomes allows us to estimate the eigenvalue p_a of the state.

The situation is therefore analogous to phase estimation, except that the eigenvalue p_a is not encoded in a momentum state but in a state with a certain number of switches between the outcomes 0 and 1. Accordingly, we must replace the momentum state used in our algorithm by a state with the right distribution of switches $|\mu\rangle = \sum_{j\in\{0,1\}^k} (\sqrt{\mu})^{k-s(j)} (\sqrt{1-\mu})^{s(j)} (-1)^{\ell(j)} |j\rangle$, where s denotes the number of switches and ℓ the number of pairs of consecutive 0's. The bandwidth ϵ is adjusted by setting $k = 2\mu(1-\mu)/\epsilon^2$. One important advantage of this type of filter state is that it drops very abruptly outside its bandwidth; $|\langle\mu|p\rangle|$ is essentially proportional to a normal distribution centered at $p = \mu$ and of variance $2\mu(1-\mu)/k^2$. Thus, there is no need for multiple copies of the filter state and the rest of the algorithm proceeds as before (see Appendix E [18] for a detailed analysis).

This more general algorithm does not perform as well as the algorithm used for local Hamiltonians because it searches over a larger Hilbert space: the space of the witness and the scratchpad. This is to be expected since it makes no assumption about the structure of the verification procedure V_x. Note however that all known ''natural'' problems in QMA—e.g., nonidentity check [22], consistency of quantum states [23], N representability [24], and zero-error capacity of quantum channels [25]—use a scratchpad of only logarithmic size, so in those cases the running time is the same as for local Hamiltonians. It is tempting to conjecture that the scratchpad of all problems in QMA can be reduced to this size.

To summarize, we have presented a method to prepare ground and thermal states of quantum many-body systems on a quantum computer. The time required by our algorithm is equal to the square root of the Hilbert space dimension of the system—the same time required to prepare the ground state of a classical many-body system. This represents a quadratic speed-up and an exponential memory reduction over Lanczos method, which in general is the most efficient available technique to accomplish this task on a classical computer. It is perhaps surprising that this task cannot be accomplished by a straightforward combination of phase estimation and Grover's algorithm, but our analysis of this strategy reveals an important failure, and more elaborate methods were required. A quantum computer, if it could be built, could serve as an efficient simulator of quantum many-body systems. The method we have presented would complement this simulation by initializing the computer in a low-energy state of the simulated system.

With some modifications, our algorithm can be used to solve and prepare relevant witnesses of all problems in the complexity class QMA, the quantum generalization of NP. In that case, the physical task consists of preparing a state that has a large overlap with two projectors. Problems in NP form a special case where those projectors commute and can be solved straightforwardly using Grover's algorithm. However, in the general case the projectors do not commute and more sophisticated techniques were required.

We acknowledge Andrew Childs, Peter Høyer, Dominik Janzing, John Preskill, and John Watrous for stimulating discussions. This work began while P. W. was visiting D. P. at California Institute of Technology.

*To whom correspondence should be addressed.
David.Poulin@USherbrooke.ca

[1] S. Lloyd, Science **273**, 1073 (1996).
[2] F. Barahona, J. Phys. A **15**, 3241 (1982).
[3] S. Arora and S. Safra, J. ACM **45**, 70 (1998).
[4] S. Arora *et al.*, J. ACM **45**, 501 (1998).
[5] I. Bieche *et al.*, J. Phys. A **13**, 2553 (1980).
[6] F. Barahona *et al.*, J. Phys. A **15**, 673 (1982).
[7] A. Hartmann and H. Rieger, *Optimization Algorithms in Physics* (Wiley-VCH, Berlin, 2001).
[8] W. Cook *et al.*, *Combinatorial Optimization* (Wiley and Sons, New York, 1998).
[9] S. Kirkpatrick *et al.*, Science **220**, 671 (1983).
[10] R. Somma *et al.*, Phys. Rev. Lett. **101**, 130504 (2008).
[11] E. Farhi *et al.*, Science **292**, 472 (2001).
[12] W. van Dam and U. Varizani, http://www.eecs.berkeley.edu/~vazirani/pubs/qao.ps.
[13] L. Grover, in *Proceedings of the 28th Annual ACM Symposium on the Theory of Computation* (ACM, New York, 1996), p. 212.
[14] G. Brassard *et al.*, in *Proceedings of the International Colloquium on Automata, Languages and Programming (ICALP'98)*, edited by K. G. Larsen, S. Skyum, and G. Winskel (Springer-Verlag, Berlin, 1998).
[15] C. H. Bennett *et al.*, SIAM J. Comput. **26**, 1510 (1997).
[16] C. Dankert *et al.*, arXiv:quant-ph/0606161.
[17] A. Kitaev, arXiv:quant-ph/9511026.
[18] See EPAPS Document No. E-PRLTAO-102-043916 for appendices. For more information on EPAPS, see http://www.aip.org/pubservs/epaps.html.
[19] A. Y. Kitaev, A. H. Shen, and M. N. Vyalyi, *Classical and Quantum Computation* (AMS, Providence, RI, 2002).
[20] C. Marriott and J. Watrous, Comput. Complex. **14**, 122 (2005).
[21] C. Jordan, Bull. Soc. Math. Fr. **3**, 103 (1875).
[22] D. Janzing *et al.*, Int. J. Quantum. Inform. **3**, 463 (2005).
[23] Y.-K. Liu, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, Lecture Notes in Computer Science Vol. 4110/2006 (Springer, Berlin, 2006), p. 438–449.
[24] Y.-K. Liu *et al.*, Phys. Rev. Lett. **98**, 110503 (2007).
[25] S. Beigi and P. Shor, arXiv:0709.2090.