First-order quantum phase transition in adiabatic quantum computation

M. H. S. Amin[1] and V. Choi[1,2]

[1]*D-Wave Systems Inc., 100-4401 Still Creek Drive, Burnaby, British Columbia, Canada V5C 6G9*
[2]*Department of Computer Science, Virginia Tech, Falls Church, Virginia 22043, USA*
(Received 30 April 2009; published 11 December 2009)

We investigate the connection between local minima in the problem Hamiltonian and first-order quantum phase transitions during adiabatic quantum computation. We demonstrate how some properties of the local minima can lead to an extremely small gap that is exponentially sensitive to the Hamiltonian parameters. Using perturbation expansion, we derive an analytical formula that cannot only predict the behavior of the gap, but also provide insight on how to controllably vary the gap size by changing the parameters. We show agreement with numerical calculations for a weighted maximum independent set problem instance.

I. INTRODUCTION

Adiabatic quantum computation (AQC) was first proposed in 2000 by Farhi *et al.* [1] as a means to solve *NP*-hard optimization problems. Later, Aharonov *et al.* [2] proved that AQC is polynomially equivalent to conventional (gate model) quantum computation. It is also believed that AQC is more robust against errors caused by environmental noise [3–5].

In AQC, the system's Hamiltonian, usually written as

$$\mathcal{H}(t) = [1 - \lambda(t)]\mathcal{H}_B + \lambda(t)\mathcal{H}_P, \qquad (1)$$

evolves slowly with time t as $\lambda(t)$ changes monotonically from 0 to 1 within a time t_f. The initial Hamiltonian $\mathcal{H}_B$ is assumed to have an easily accessible ground state into which the system is initialized, while the ground state of the final Hamiltonian $\mathcal{H}_P$ provides the solution to the problem of interest. In order to reach the final ground state with high fidelity, the adiabatic theorem requires $t_f \propto g_{min}^{-\delta}$, where g_{min} is the minimum gap between the two lowest-energy instantaneous eigenstates of $\mathcal{H}$. The power δ can be 1, 2, or possibly some other number depending on the functional form of $\lambda(t)$ and the distribution of the higher energy levels [1,6,7].

Thus, in order to address the efficiency of AQC, one needs to analyze g_{min}, which is unfortunately as hard as solving the original problem if computed directly. The most fundamental problem in AQC is, therefore, how to bound g_{min} analytically. Equally important is how to unveil the quantum evolution blackbox by relating the formation of g_{min} to the structure of the problem, and thus obtain insights for designing efficient algorithms.

Besides a few special cases in which spectral gaps are computed analytically [8], all other known studies have to resort to numerical calculations, e.g., diagonalization [9] or quantum Monte Carlo (QMC) techniques [10]. Unfortunately, these methods are limited to small problem sizes (to date, $N < 30$ for diagonalization and up to 128 for QMC), and tend not to provide much insight into why an extracted minimum gap is large or small.

It has been recognized that during the evolution an adiabatic quantum computer may go through a quantum phase transition (QPT) [11–13]. The transition is first (second) order if the change in the order parameter at the transition point is discontinuous (continuous) [12]. In this article, we investigate the effect of problem structure, specifically the role of local minima [14,15], on the formation of a first-order QPT.

The paper is organized as follows. In the next section, we discuss the relation between AQC, or more specifically, adiabatic quantum optimization, to quantum phase transitions. We qualitatively describe how a first-order QPT may happen during adiabatic evolution due to the presence of a local minimum. In Sec. III, we use perturbation expansion to study such QPT in more detail and obtain analytical formula for the gap position and size. In Sec. IV, we examine our findings via a small size example graph of the (*NP*-hard) weighted maximum independent set (WMIS) problem. We show agreement between our perturbation calculation and numerical diagonalization. We demonstrate controllable variation of g_{min} by 6 orders of magnitude by slightly (25%) changing the Hamiltonian parameters. At the end, we summarize our conclusions in Sec. V.

II. AQC AND QUANTUM PHASE TRANSITIONS

Consider a transverse field Ising Hamiltonian (1) with

$$\mathcal{H}_P = \sum_i h_i \sigma_i^z + \sum_{i,j} J_{ij} \sigma_i^z \sigma_j^z, \quad \mathcal{H}_B = \Delta \sum_i \sigma_i^x, \qquad (2)$$

where $\sigma_i^{x,z}$ are Pauli matrices for the i-th qubit. Let $\mathcal{E}$ denote an energy scale that characterizes Hamiltonian $\mathcal{H}_P$ the same way as Δ characterizes $\mathcal{H}_B$. (Here, we do not specify $\mathcal{E}$, but only mention its existence.) In the interpolation Hamiltonian (1), depending on whether $\mathcal{H}_B$ or $\mathcal{H}_P$ dominates, the system will be localized in one of the two computation (in which σ_i^z is diagonal) or Hadamard (in which σ_i^x is diagonal) bases, hence delocalized in the conjugate basis. We introduce the dimensionless parameter

$$\zeta = \frac{(1 - \lambda)\Delta}{\lambda \mathcal{E}}, \qquad (3)$$

which provides a measure of the relative importance of $\mathcal{H}_B$ and $\mathcal{H}_P$. Let $|\psi_0\rangle$ denote the instantaneous ground state of $\mathcal{H}(t)$. At $\lambda \approx 0$, ζ is very large and $\mathcal{H}_B$ is dominant making $|\psi_0\rangle$ a large superposition in the computation basis or localized in the Hadamard basis. This is the quantum paramag-

netic phase. As ζ is lowered, around $\zeta_c \approx 1$ [or equivalently $\lambda_c \approx \Delta/(\Delta+\mathcal{E})$], the dominance shifts from $\mathcal{H}_B$ to $\mathcal{H}_P$, which favors localization in the computation basis.

The transition from paramagnetic to ordered phase is usually via a QPT. For a large homogeneous system with no field, QPT happens at a well-defined critical point λ_c and is usually continuous, thus second order. The lowest-energy excitations, delocalized over the whole system, are gapped with a gap that shrinks at λ_c, but only polynomially in the system size N [16]. However, this scenario may be significantly altered by inhomogeneity [17,18].

In the absence of any further phase transition, the minimum gap is most likely at λ_c. Despite the localization at λ_c, $|\psi_0\rangle$ will still be a superposition in the computation basis for $\lambda > \lambda_c$, but with a much smaller number of computational states involved. As λ is increased, the total energy of such localized state will change and at some later point (λ^*) it may cross another localized state. The ground state of the system will then make a sudden transition to the new state via a discontinuous (first-order) QPT [24]. The gap at the transition point will be extremely small, making λ^* the new position of the minimum gap. The transition at λ^* is between two ordered phases, in contrast to the order-disorder transition at λ_c. There may even be more than one such transition if the ground state crosses other localized states, but all those transitions can only happen after localization at λ_c. An important question now is what properties of $\mathcal{H}_P$ are responsible for such a first-order QPT. In the next section, we will employ perturbation expansion to answer this question.

III. PERTURBATION EXPANSION

Let us define $\mathcal{H}_0 = \lambda \mathcal{H}_P$ and $\mathcal{H}' = (1-\lambda)\mathcal{H}_B$ as the unperturbed and perturbation Hamiltonians, respectively. We use ζ as defined in Eq. (3) as the dimensionless small parameter. At $\zeta = 0 (\lambda = 1)$, the eigenstates of the system are eigenfunctions of $\mathcal{H}_P$ (computational states). Thus, $|\psi_0\rangle$ is the global minimum of $\mathcal{H}_P$ and the lowest lying excited states are either states that are a few bit flips away from (or neighborhood of) the global minimum, or some low energy local minima and their corresponding neighborhoods. These two types of states behave completely differently as ζ is increased. Since $\mathcal{H}'$ only involves σ_i^x operators, to every order of perturbation, it flips only one qubit. Therefore, the lowest order of perturbation that gives a nonzero off-diagonal element $\mathcal{H}_{mn} \equiv \langle m|\mathcal{H}|n\rangle$ is equal to the number of bit flips f_{mn} (Hamming distance) between states $|m\rangle$ and $|n\rangle$. This gives $\mathcal{H}_{mn} = O(\zeta^{f_{mn}})$, hence at $\lambda \approx 1$, the only nonvanishing $\mathcal{H}_{mn}$ are those between neighboring states for which f_{nm} is small. Essentially, $\mathcal{H}$ becomes block diagonal with every minimum and its neighborhood forming a cluster of states with nonzero off-diagonal elements with each other, but vanishingly small off-diagonal elements with states in other clusters. Upon diagonalization of $\mathcal{H}$, the new eigenstates become superpositions of only the neighboring states. Those neighboring states never cross due to the no-level-crossing theorem [19]. On the other hand, if as λ is decreased, two clusters move as a whole with respect to each other, as their energy levels cross they create anticrossings with very small gaps of $O(\zeta^{f_{mn}})$.
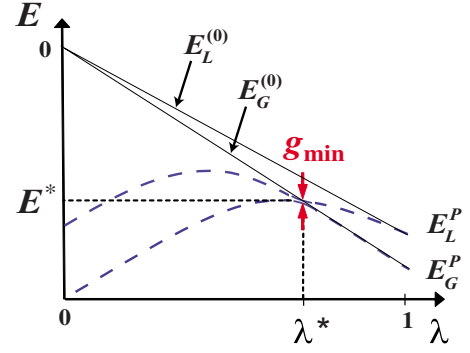

FIG. 1. (Color online) Schematic diagram of crossing global and local minima. At zeroth-order perturbation, levels do not cross (solid lines). Contribution from the second-order perturbation may cause the levels cross if the curvature of the upper level is larger than the lower one (dashed lines).

The lowest of these anticrossings form a first-order QPT.

Let us now make the above observation more quantitative. At $\lambda = 1$, the Hamiltonian $\mathcal{H} = \mathcal{H}_P$ has a global minimum $|G^{(0)}\rangle$ with energy E_G^P. Let $|L^{(0)}\rangle$ represent a low energy local minimum (not necessarily the first excited state) of $\mathcal{H}_P$ with energy E_L^P. For now, we take both the above states to be nondegenerate. We use perturbation expansion to calculate the perturbed states $|\alpha = G, L\rangle$, and their eigenvalues E_α at $\lambda \lesssim 1$. To the second-order perturbation, $E_\alpha = \lambda E_\alpha^P - \chi_\alpha (1-\lambda)^2/\lambda$, where

$$\chi_\alpha = \sum_{n \neq \alpha} \frac{|\langle \alpha^{(0)}|\mathcal{H}_B|n^{(0)}\rangle|^2}{E_n^P - E_\alpha^P} = \frac{\lambda^3}{4}\frac{d^2 E_\alpha}{d\lambda^2}. \qquad (4)$$

The perturbation expansion holds as long as for all states contributing to the sum, $E_n^P - E_\alpha^P \gg (1-\lambda)\Delta/\lambda$. The two perturbed levels cross at (see Fig. 1)

$$\lambda^* = \left(1 + \sqrt{\frac{E_L^P - E_G^P}{\chi_L - \chi_G}}\right)^{-1}. \qquad (5)$$

Since $E_L^P > E_G^P$, if $\chi_L < \chi_G$, then there will be no real solution for λ^* (up to the second order in perturbation). Since the curvature of the energy levels, as shown in Eq. (4), is proportional to χ_α, when $\chi_L < \chi_G$ the local minimum has less curvature than the global minimum. In that case, the lower one of the two curves in Fig. 1 will be more curved than the upper one, hence will not cross it. As a result, no first-order QPT, due to such a local minimum, will occur. Even if Eq. (5) gives a real solution but with $\lambda^* < \lambda_c$, such a solution is not acceptable because perturbation expansion breaks down for $\lambda < \lambda_c(\zeta > 1)$. Therefore, only local minima that yield a $\lambda^* > \lambda_c$ may cause first-order QPT. In the absence of such local minima, there will not be any first-order QPT. Of course, Eq. (5) is a result of second-order perturbation expansion, therefore not accurate. In principle, one can continue to higher orders of perturbation to find a more accurate λ^*, but the computation becomes more complicated.

The degeneracy at the above level crossing is removed by tunneling between the two localized states. To find $g_{\min}$, we need to calculate the effective 2×2 Hamiltonian $\tilde{\mathcal{H}}$ in the
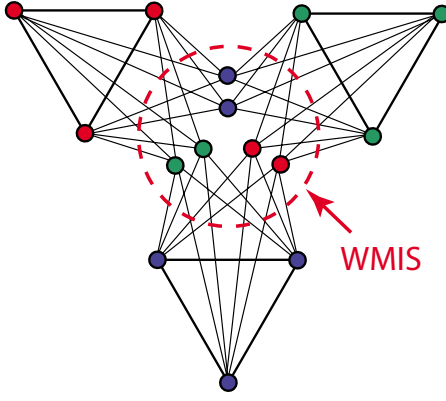
FIG. 2. (Color online) An example graph of WMIS problem. The central six vertices have a weight w_G and the nine outer ones are weighted w_L. For $w_L < 2w_G$, the six central vertices make the WMIS, while every combination of three vertices each from one triangle is a smaller independent set, altogether making 27 degenerate local minima.

FIG. 3. (Color online) (a) The first three energy levels for the example graph of Fig. 2 with $J=2$, $w_G=\Delta=1$, and $w_L=1.8$. (b) Functions S and M, as defined in Eq. (8), for the same problem.

subspace made of the perturbed states $|G\rangle$ and $|L\rangle$. At the degeneracy point, $\widetilde{\mathcal{H}}_{LL}=\widetilde{\mathcal{H}}_{GG}=E^*$ and the off-diagonal elements remove the degeneracy. To the lowest-order perturbation, $g_{\min}=2(\widetilde{\mathcal{H}}_{LG}\widetilde{\mathcal{H}}_{GL})^{1/2}$, where

$$\widetilde{\mathcal{H}}_{LG} = \frac{(1-\lambda^*)^f}{\lambda^{*f-1}} \sum_{\{n_s\}} \frac{\prod_{s=1}^{f} \langle n_{s-1}^{(0)}|\mathcal{H}_B|n_s^{(0)}\rangle}{\prod_{s=1}^{f-1}(E_G^P-E_{n_s}^P)}, \qquad (6)$$

and $L \leftrightarrow G$ for $\widetilde{\mathcal{H}}_{GL}$. Here, $|n_0^{(0)}\rangle=|L^{(0)}\rangle$, $|n_f^{(0)}\rangle=|G^{(0)}\rangle$, f is the Hamming distance between these two states, and $|n_{s\neq 0,f}^{(0)}\rangle$ cover all other states except the above two.

If any of the local or global minima are degenerate, $\widetilde{\mathcal{H}}$ should be written in the subspace of all those degenerate states, hence will not be a 2×2 matrix. However, if there exists a symmetry between those degenerate states, as in the example studied below, one can still use a 2×2 Hamiltonian $\widetilde{\mathcal{H}}$ for the two uniform (lowest-energy) superposition states $|\alpha^{(0)}\rangle=\Sigma_{k\in\mathcal{S}_\alpha}|k^{(0)}\rangle/\sqrt{\mathcal{N}_\alpha}$. Here, $\mathcal{S}_\alpha$ is the subspace made of the $\mathcal{N}_\alpha$ degenerate states contributing to $|\alpha^{(0)}\rangle$. In that case, Eq. (6) can be used, but with the sum over all $|n_s^{(0)}\rangle \notin S_G \cup S_L$.

As is clear from Eq. (6), the gap becomes extremely small as $\lambda^* \to 1$. This limit is reached, according to Eq. (5), when there is a small $E_L^P-E_G^P$ and/or a large $\chi_L-\chi_G$. The former simply means that the local minimum should have an energy close to the global one. To understand the latter, we notice that the denominator in the sum in Eq. (4) is the energy cost of a bit flip away from the local or global minima. Having a large $\chi_L-\chi_G$ means that on average, the cost of bit flips inside the local minimum is smaller than inside the global minimum. In other words, there should be more low energy excitations, close to the bottom of the local minimum compared to the global minimum.
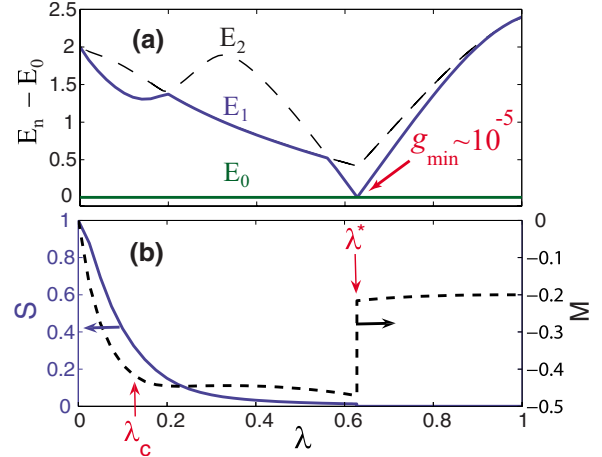
The above observation, although obtained from second-order perturbation, has a wider validity. In the computation basis, $\mathcal{H}_P$ is diagonal, hence plays the role of a potential energy, while $\mathcal{H}_B$ causes bit flips thus is responsible for dynamics. As is clear from Fig. 1, the unperturbed energy levels $E_\alpha^{(0)}$ (straight lines) do not cross at any $\lambda>0$. Thus, the curvature of the levels, which is responsible for their crossing, is a result of the contribution of $\mathcal{H}_B$. Qualitatively speaking, for a local minimum to have a larger curvature than (thus cross) the global minimum, it needs to gain more contribution from $\mathcal{H}_B$. This means the system should have more dynamics, hence more freedom, within the local than the global minimum. A simple example is when the local minimum is at the bottom of a much wider well than the global one. Below, we construct another example (using WMIS) in which there are several degenerate local minima with very small energy barrier between them and a global minimum well separated from all of those.

IV. WEIGHTED MAXIMUM INDEPENDENT SET EXAMPLE

Consider a general graph $\mathcal{G}=[V(\mathcal{G}),E(\mathcal{G})]$, with a set of vortices $V(\mathcal{G})$ and a set of edges $E(\mathcal{G})$ (see Fig. 2 for an example). For each vertex $i \in V(\mathcal{G})=\{1,\ldots,n\}$, one can associate a positive real number w_i (i.e., weight) to obtain a vertex-weighted graph. The weighted maximum independent set problem (WMIS) seeks to find a $\mathcal{S} \subseteq V(\mathcal{G})$, such that $\mathcal{S}$ is independent (i.e., each two vortices in $\mathcal{S}$ are not connected) and the total *weight* of $\mathcal{S}(=\Sigma_{i\in\mathcal{S}}w_i)$ is maximized. WMIS can be solved as an Ising problem with [20]: $\mathcal{H}_P=\Sigma_{i\in V(\mathcal{G})}h_i\sigma_i^z+\Sigma_{i,j\in E(\mathcal{G})}J_{ij}\sigma_i^z\sigma_j^z$, where $h_i=\Sigma_{ij\in E(\mathcal{G})}J_{ij}-2w_i$, with the condition $J_{ij}>\min\{w_i,w_j\}$. Figure 2 is an example graph consisting six (central) vertices of weight w_G and nine (outer) vertices of weight w_L, connected by edges with $J_{ij}=J$. For $w_L<2w_G$, the potential has a deep narrow global minimum and 27 shallow local minima separated from each other by two bit flips (see the figure caption). The barrier height between two neighboring local minima is $\delta U=4(J-w_L)$. Also, one obtains

$$E_L^P - E_G^P = 4(6w_G - 3w_L),$$

$$\chi_G = \frac{1}{4}\left(\frac{6}{w_G} + \frac{9}{4J - w_L}\right),$$

$$\chi_L = \frac{1}{4}\left(\frac{6}{2J - w_G} + \frac{9}{w_L} + \frac{12}{J - w_L}\right). \quad (7)$$

Figure 3(a) shows a plot of the first three energy levels for a chosen set of parameters. The gap shows a first large minimum at $\lambda \approx 0.1$ and a much smaller second minimum at $\lambda \approx 0.6$. To understand what these minima correspond to, we plot in Fig. 3(b) two other quantities

$$S = \frac{1}{2^N}\left[\sum_{n=1}^{2^N} |\langle\psi_0|n\rangle|\right]^2, \quad M = \frac{1}{N}\sum_{i=1}^{N} \langle\psi_0|\sigma_i^z|\psi_0\rangle. \quad (8)$$

S provides a measure for how spread the wave function is in the computation basis. For a wave function that is a uniform superposition of m computational states, $S = m/2^N$, i.e., the fraction of states that participate in the superposition. M, on the other hand, is the normalized total magnetization, which for such a small scale problem can play the role of an order parameter [25]. As we can see in Fig. 3(b), a significant drop of S happens around $\lambda = \lambda_c$, which coincides with the position of the first minimum gap. Since there are only 15 qubits involved, such a transition is not very sharp, as it would be for large scale systems. The order parameter M continuously increases in magnitude around λ_c, but suddenly jumps to another value at λ^*. This is a clear indication of a first-order QPT at λ^*.

As $w_L \rightarrow 2$, $\delta U \rightarrow 0$ and the cost of bit flips within the local minima vanishes, allowing more dynamics within the local minima. Moreover, from Eq. (7), we see $E_L^P - E_G^P \rightarrow 0$. Both of these effects lead to $\lambda^* \rightarrow 1$, and thereby a small $g_{\min}$, as discussed above. Figure 4(a), plots the results of numerical diagonalization and perturbative calculations of λ^* versus w_L. Both curves approach 1 as $w_L \rightarrow 2$.

We have also plotted $g_{\min}$ in Fig. 4(b), using both diagonalization and perturbation. The perturbation calculations are done at the exact λ^* (extracted from diagonalization) as well as its corresponding perturbative value. A striking point is that, by changing w_L by about 25%, the minimum gap changes by 6 orders of magnitude, i.e., 12 orders of magnitude difference in t_f. Such a strong dependence is indeed expected for a first-order QPT, as the tunneling amplitude between the local and global minima is exponentially sensitive to parameters that characterize the barrier between the minima.

V. CONCLUSION

We have shown that extremely small $g_{\min}$ may arise in AQC if the anticrossing is a result of a first-order QPT. Such a phase transition happens if the final Hamiltonian possesses
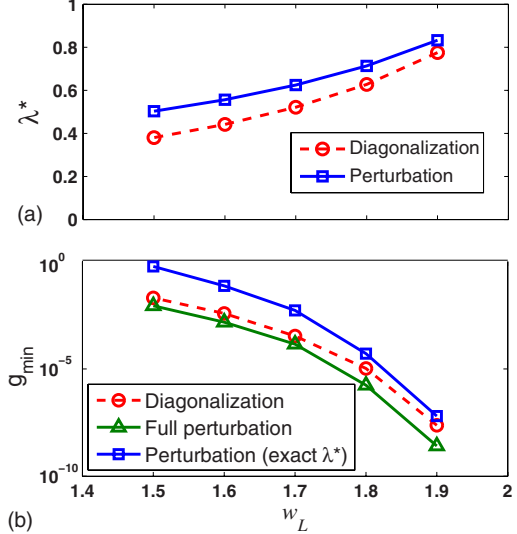

FIG. 4. (Color online) The (a) position and (b) magnitude of the minimum gap for the graph of Fig. 2 with $J = 2$, $w_G = \Delta = 1$. The x axis for both figures are the same.

some low energy local minima with small cost of bit flip within them compared to the global minimum. Our perturbative calculation indicates that the gap depends exponentially on the Hamming distance between the two minima involved in the phase transition. If none of the local minima has the above properties, a first-order QPT and, therefore, an exponentially small minimum gap may not occur. We have also supported our theoretical findings with an example of WMIS problem. A very small $g_{\min}$, extremely sensitive to the Hamiltonian parameters, was observed as a result of a first-order QPT.

Our example, though capturing all the qualitative features of the theory, is not generic. In many instances, a minimum at λ_c may not be visible. Moreover, second-order perturbation expansion may not be adequate to find λ^* with acceptable accuracy for more complicated instances. One should also note that a second-order QPT may also lead to an exponential dependence of $g_{\min}$ on the size if the Hamiltonian has random inhomogeneities [17]. Nevertheless, the physical process leading to the first-order QPT described here is general and not limited to the presented example. Since the first arXiv appearance of this paper, similar ideas have been employed by other groups [21–23] to a different problem and similar results were obtained, confirming the generality of our findings.

[1] E. Farhi, J. Goldstone, S. Gutmann, J. Lapan, A. Lundgren, and D. Preda, Science **292**, 472 (2001).

[2] D. Aharonov, W. van Dam, J. Kempe, Z. Landau, S. Lloyd, and O. Regev, Proceedings of the 45th FOCS, 2004 (unpublished), p. 42.

[3] A. M. Childs, E. Farhi, and J. Preskill, Phys. Rev. A **65**, 012322 (2001); J. Roland and N. J. Cerf, *ibid.* **71**, 032330 (2005); S. Ashhab, J. R. Johansson, and F. Nori, *ibid.* **74**, 052330 (2006); M. Tiersch and R. Schützhold, *ibid.* **75**, 062313 (2007).

[4] M. H. S. Amin, P. J. Love, and C. J. S. Truncik, Phys. Rev. Lett. **100**, 060503 (2008); M. H. S. Amin, D. V. Averin, and J. A. Nesteroff, Phys. Rev. A **79**, 022107 (2009); M. Amin, C. Truncik, and D. Averin, *ibid.* **80**, 022303 (2009).

[5] S. Lloyd, e-print arXiv:0805.2757.

[6] G. Schaller, S. Mostame, and R. Schützhold, Phys. Rev. A **73**, 062307 (2006).

[7] D. Lidar, A. Rezakhani, and A. Hamma, J. Math. Phys. **50**, 102106 (2009).

[8] W. van Dam, M. Mosca, and U. Vazirani, Proceedings of the 42nd FOCS, 2001 (unpublished) pp. 279–287; J. Roland and N. J. Cerf, Phys. Rev. A **65**, 042308 (2002); M. Žnidarič and M. Horvat, *ibid.* **73**, 022329 (2006); R. Schützhold, J. Low Temp. Phys. **153**, 228 (2008); G. Schaller, Phys. Rev. A **78**, 032328 (2008).

[9] M. Žnidarič, Phys. Rev. A **71**, 062305 (2005); G. Schaller and R. Schützhold, e-print arXiv:0708.1882.

[10] A. P. Young, S. Knysh, and V. N. Smelyanskiy, Phys. Rev. Lett. **101**, 170503 (2008).

[11] S. Knysh and V. Smelyanskiy, e-print arXiv:cond-mat/0602257.

[12] R. Schützhold and G. Schaller, e-print arXiv:quant-ph/0608017.

[13] S. Lloyd, Science **319**, 1209 (2008).

[14] B. Reichardt, Proceedings of the 36th STOC, 2004 (unpublished), p. 502.

[15] M. H. S. Amin, Phys. Rev. Lett. **100**, 130503 (2008).

[16] S. Sachdev, *Quantum Phase Transitions* (Cambridge University Press, New York, 1999).

[17] D. Boyanovsky and J. L. Cardy, Phys. Rev. B **26**, 154 (1982); N. Read, S. Sachdev, and J. Ye, *ibid.* **52**, 384 (1995); T. Vojta, J. Phys. A **39**, R143 (2006); D. S. Fisher, Phys. Rev. B **51**, 6411 (1995); J. Dziarmaga, *ibid.* **74**, 064416 (2006); T. Caneva, R. Fazio, and G. E. Santoro, *ibid.* **76**, 144427 (2007).

[18] W. M. Kaminsky and S. Lloyd, presentation at APS March Meeting in Denver, Colorado (2007).

[19] J. J. Sakurai, *Modern Quantum Mechanics* (Addison-Wesley, Reading, MA, 1995).

[20] V. Choi, Quantum Inf. Process. **7**, 193 (2008).

[21] B. Altshuler, H. Krovi, and J. Roland, e-print arXiv:0908.2782.

[22] E. Farhi, J. Goldstone, D. Gosset, S. Gutmann, H. Meyer, and P. Shor, MIT Report No. CTP 4076 (unpublished); E. Farhi, J. Goldstone, D. Gosset, S. Gutmann, H. Meyer, and P. Shor, CERN Report No. PH-TH-2009/175 (unpublished).

[23] A. Young, S. Knysh, and V. Smelyanskiy, e-print arXiv:0910.1378.

[24] In large scale spin glasses, those localized states may not represent two macroscopically distinct phases. Nevertheless, we continue using the phrase "phase transition."

[25] For large scale spin glasses, M is not a good order parameter as it is nearly zero for a large number of states.