Adiabatic quantum computation

Tameem Albash

Information Sciences Institute, University of Southern California,
Marina del Rey, California 90292, USA,
Department of Physics and Astronomy, University of Southern California,
Los Angeles, California 90089, USA,
and Center for Quantum Information Science & Technology,
University of Southern California, Los Angeles, California 90089, USA

Daniel A. Lidar

Department of Physics and Astronomy, University of Southern California,
Los Angeles, California 90089, USA,
Center for Quantum Information Science & Technology, University of Southern California,
Los Angeles, California 90089, USA,
Department of Electrical Engineering, University of Southern California,
Los Angeles, California 90089, USA,
and Department of Chemistry, University of Southern California,
Los Angeles, California 90089, USA

Adiabatic quantum computing (AQC) started as an approach to solving optimization problems and has evolved into an important universal alternative to the standard circuit model of quantum computing, with deep connections to both classical and quantum complexity theory and condensed matter physics. This review gives an account of the major theoretical developments in the field, while focusing on the closed-system setting. The review is organized around a series of topics that are essential to an understanding of the underlying principles of AQC, its algorithmic accomplishments and limitations, and its scope in the more general setting of computational complexity theory. Several variants are presented of the adiabatic theorem, the cornerstone of AQC, and examples are given of explicit AQC algorithms that exhibit a quantum speedup. An overview of several proofs of the universality of AQC and related Hamiltonian quantum complexity theory is given. Considerable space is devoted to stoquastic AQC, the setting of most AQC work to date, where obstructions to success and their possible resolutions are discussed.

CONTENTS

I. INTRODUCTION

Quantum computation (QC) originated with Benioff's proposals for quantum Turing machines (Benioff, 1980, 1982) and Feynman's ideas for circumventing the difficulty of simulating quantum mechanics by classical computers (Feynman, 1982). This led to Deutsch's proposal for universal QC in terms of what has become the "standard" model: the circuit or gate model of QC (Deutsch, 1989). Adiabatic quantum computation (AQC) is based on an idea that is quite distinct from the circuit model. Whereas in the latter a computation may in principle evolve in the entire Hilbert space and is encoded into a series of unitary quantum logic gates, in AQC the computation proceeds from an initial Hamiltonian whose ground state is easy to prepare, to a final Hamiltonian whose ground state encodes the solution to the computational problem. The adiabatic theorem guarantees that the system will track the instantaneous ground state provided the Hamiltonian varies sufficiently slowly. It turns out that this approach to QC has deep connections to condensed matter physics, computational complexity theory, and heuristic algorithms.

In its first incarnation, the idea of encoding the solution to a computational problem in the ground state of a quantum Hamiltonian appeared as early as 1988, in the context of solving classical combinatorial optimization problems, where it was called quantum stochastic optimization (Apolloni, Carvalho, and de Falco, 1989).[1] It was renamed quantum annealing (QA) by Apolloni, Cesa-Bianchi, and de Falco (1988) and reinvented several times (Somorjai, 1991; Amara, Hsu, and Straub, 1993; Finnila *et al.*, 1994; Kadowaki and Nishimori, 1998).[2] These early papers emphasized that QA was to be understood as an *algorithm* that exploits simulated quantum (rather than thermal) fluctuations and tunneling, thus providing a quantum-inspired version of simulated annealing (SA) (Kirkpatrick, Gelatt, and Vecchi, 1983). The first direct comparison between QA and SA (Kadowaki and Nishimori, 1998) suggested that QA can be more powerful.

A very different approach was taken via an *experimental* implementation of QA in a disordered quantum ferromagnet (Brooke *et al.*, 1999; Brooke, Rosenbaum, and Aeppli, 2001). This provided the impetus to reconsider QA from the perspective of quantum computing, i.e., to consider a dedicated device that solves optimization problems by exploiting quantum evolution. Thus was born the idea of the quantum adiabatic algorithm (QAA) (Farhi *et al.*, 2000, 2001) [also referred to as adiabatic quantum optimization (AQO) (Smelyanskiy, Toussaint, and Timucin, 2001; Reichardt, 2004)], wherein a physical quantum computer

---

[1]Even though Apolloni, Carvalho, and de Falco (1989) was published in 1989, it was submitted in 1988, before Apolloni, Cesa-Bianchi, and de Falco (1988), which referenced it.

[2]It was called the "quasiquantal method" by Somorjai (1991), the "imaginary-time algorithm" by Amara, Hsu, and Straub (1993), and quantum annealing by Finnila *et al.* (1994) and Kadowaki and Nishimori (1998). The latter term has become widely accepted.

solves a combinatorial optimization problem by evolving adiabatically in its ground state. The term adiabatic quantum computation we use here was introduced by van Dam, Mosca, and Vazirani (2001), although the context was still optimization.[3]

Adiabatic quantum algorithms for optimization problems typically use "stoquastic" Hamiltonians, characterized by having only nonpositive off-diagonal elements in the computational basis. Adiabatic quantum computation with nonstoquastic Hamiltonians is as powerful as the circuit model of quantum computation (Aharonov *et al.*, 2007). In other words, nonstoquastic AQC and all other models for universal quantum computation can simulate one another with at most polynomial resource overhead. For this reason the contemporary use of the term AQC typically refers to the general, nonstoquastic setting, thus extending beyond optimization to any computational problem. When discussing the case of stoquastic Hamiltonians we use the term "stoquastic AQC" (StoqAQC).[4]

For most of this review we essentially adopt the definition of AQC from Aharonov *et al.* (2007), as this definition allows for the proof of the equivalence with the circuit model and is thus used to establish the universality of AQC. Interestingly, this proof builds on one of the first QC ideas due to Feynman, which was later shown to allow a general purpose quantum computation to be embedded in the ground state of a quantum system (Feynman, 1985; Kitaev, Shen, and Vyalyi, 2000). A related ground state embedding approach was independently pursued by Mizel, Mitchell, and Cohen (2001, 2002) around the same time as the original development of the QAA. To define AQC, we first need the concept of a *k*-local Hamiltonian, which is a Hermitian matrix H acting on the space of *p*-state particles that can be written as $H = \sum_{i=1}^{r} H_i$ where each H_i acts nontrivially on at most *k* particles, i.e., $H_i = h \otimes \mathbb{1}$ where h is a Hamiltonian on at most *k* particles, and $\mathbb{1}$ denotes the identity operator.

Definition 1 (adiabatic quantum computation). A *k*-local adiabatic quantum computation is specified by two *k*-local Hamiltonians H_0 and H_1 acting on *n* *p*-state particles, $p \geq 2$. The ground state of H_0 is unique and is a product state. The output is a state that is ε close in ℓ_2 norm to the ground state of H_1. Let $s(t) \colon [0, t_f] \mapsto [0, 1]$ (the "schedule") and let t_f be the smallest time such that the final state of an adiabatic evolution generated by $H(s) = (1 - s)H_0 + sH_1$ for time t_f is ε close in ℓ_2 norm to the ground state of H_1.

Several comments are in order. (1) A uniqueness requirement was imposed on the ground state of H_1 by Aharonov *et al.* (2007), but this is not necessary. For example, in the setting where H_1 represents a classical optimization problem, multiple final ground states do not pose a problem as any of

the final states represent a solution to the optimization problem. (2) Sometimes it is beneficial to consider adiabatic quantum computation in an excited state (see, e.g., Sec. VI.C). (3) As already noted by Aharonov *et al.* (2007), it is useful to allow for more general "paths" between H_0 and H_1, e.g., by introducing an intermediate "catalyst" Hamiltonian that vanishes at $s = 0$, 1 (see, e.g., Sec. VII.E).

A crucial question that will occupy us throughout this review is the cost of running an algorithm in AQC. In the circuit model the cost is equated with the number of gates, so one cost definition would be to count the number of gates needed to simulate the equivalent adiabatic process. This cost definition presupposes that the circuit model is fundamental, which may be unsatisfactory. In AQC one might be tempted to just use the run time t_f, but in order for this quantity to be meaningful it is necessary to define an appropriate energy scale for the Hamiltonian. Aharonov *et al.* (2007) defined the cost of the adiabatic algorithm to be the dimensionless quantity

$$\text{cost} = t_f \max_s \|H(s)\|, \tag{1}$$

in order to prevent the cost from being made arbitrarily small by changing the time units or distorting the scaling of the algorithm by multiplying the Hamiltonians by some size-dependent factor.[5] From here on we focus on the run time t_f, which should be compared to the circuit depth of analogous circuit model algorithms, whereas the full cost in Eq. (1) should be compared to the circuit gate count.

The run time t_f of an adiabatic algorithm scales at worst as $1/\Delta^3$, where Δ is the minimum eigenvalue gap between the ground state and the first excited state of the Hamiltonian of the adiabatic algorithm (Jansen, Ruskai, and Seiler, 2007). If the Hamiltonian is varied sufficiently smoothly, one can improve this to $O(1/\Delta^2)$ up to a polylogarithmic factor in Δ (Elgart and Hagedorn, 2012). While these are useful sufficient conditions, they involve bounding the minimum eigenvalue gap of a complicated many-body Hamiltonian, a notoriously difficult problem. This is one reason that AQC has generated so much interest among physicists: it has a rich connection to well-studied problems in condensed matter physics. For example, because of the dependence of the run time on the gap, the performance of quantum adiabatic algorithms is strongly influenced by the type of quantum phase transition the same system would undergo in the thermodynamic limit (Latorre and Orus, 2004).

Nevertheless, a number of examples are known where the gap analysis can be carried out. For example, adiabatic quantum computers can perform a process analogous to Grover search (Grover, 1997) and thus provide a quadratic

---

[3]The first documented use of the term "adiabatic quantum computation" was by Averin (1998), but the context was an adiabatic implementation of a quantum logic gate in the circuit model.

[4]Quantum annealing, like StoqAQC, currently usually involves stoquastic Hamiltonians; we differentiate between them as we restrict StoqAQC purely to the case of closed-system evolutions, whereas QA refers to a (not necessarily adiabatic) evolution in an open system.

[5]Unless stated otherwise, we always use $\| \cdot \|$ to denote the operator norm for operators:

$$\|A\| = \sup \{\|A|\psi\rangle\| \colon |\psi\rangle \in \mathcal{H} \quad \text{with} \quad \langle\psi|\psi\rangle = 1\}$$

(i.e., the largest singular value of the operator A), and the Euclidean vector norm for vectors. Which one is used will be clear by the context.

speedup over the best possible classical algorithm for the Grover search problem (Roland and Cerf, 2002). Other examples are known where the gap analysis can be used to demonstrate that AQC provides a speedup over classical computation, including adiabatic versions of some of the keystone algorithms of the circuit model. However, much more common is the scenario either where the gap analysis reveals no speedup over classical computation or where a clear answer to the speedup question is unavailable. In fact, the least is known about adiabatic quantum speedups in the original setting of solving classical combinatorial optimization problems. This remains an area of very active research, partly due to the original (still unmaterialized) hope that the QAA would deliver quantum speedups for nondeterministic polynomial (NP-)complete problems (Farhi et al., 2001), and partly due the availability of commercial quantum annealing devices such as those manufactured by D-Wave Systems Inc. (Johnson et al., 2011), designed to solve optimization problems using stoquastic Hamiltonians.

The goal of this article is to review the field of AQC from its inception, with a focus on the *closed-system* case. That is, we omit the topic of AQC in open systems coupled to an environment. This includes all experimental work on AQC, and all work on quantum error correction and suppression methods for AQC, as these topics deserve a separate review (Albash and Lidar, 2018) and including them here would limit our ability to cover the many years of work on AQC in closed systems, an extremely rich topic with many results. For the same reasons we also omit the related fields of holonomic QC (Zanardi and Rasetti, 1999), topological QC (Nayak et al., 2008), and adiabatic state preparation for quantum simulation (Babbush, Love, and Aspuru-Guzik, 2014). To achieve our goal we organized this review around a series of topics that are essential to an understanding of the underlying principles of AQC, its algorithmic accomplishments and limitations, and its scope in the more general setting of computational complexity theory.

We begin by reviewing the adiabatic theorem in Sec. II. The adiabatic theorem forms the backbone of AQC: it provides a sufficient condition for the success of the computation and in doing so provides the run time of a computation in terms of the eigenvalue gap Δ of the Hamiltonian and the Hamiltonian time derivative. In fact there is not one single adiabatic theorem, and we review a number of different variants that provide different run-time requirements, under different smoothness and differentiability assumptions about the Hamiltonian.

Next we review in Sec. III the handful of *explicit* algorithms for which AQC is known to give a speedup over classical computation. The emphasis is on "explicit," since Sec. IV provides several proofs for the universality of AQC in terms of its ability to efficiently simulate the circuit model, and vice versa. This means that every quantum algorithm that provides a speedup in the circuit model [many of which are known (Jordan, 2016b)] can in principle be implemented with up to polynomial overhead in AQC. That the number of explicit AQC algorithms is still small is therefore likely to be a reflection of the relatively modest amount of effort that has gone into establishing such results compared to the circuit model. However, there is also a real difficulty, in that

performing the gap analysis in order to establish the actual scaling (beyond the polynomial-time equivalence) is, as already mentioned, in many cases highly nontrivial. A second nontrivial aspect of establishing a speedup by AQC is that when such a speedup is polynomial, relying on universality is insufficient, since the polynomial overhead involved in implementing the transformation from the circuit model to AQC can then swamp the speedup. A good example is the case of Grover's algorithm, where direct use of the equivalence to the circuit model does not suffice; instead, what is required is a careful analysis and choice of the adiabatic schedule $s(t)$ in order to realize the quantum speedup.

In Sec. V we go beyond universality into Hamiltonian quantum complexity theory. This is an active contemporary research area that started with the introduction of the complexity class "quantum Merlin-Arthur" (QMA) as the natural quantum generalization of the classical complexity classes NP and MA (Kitaev, Shen, and Vyalyi, 2000). The theory of QMA completeness deals with decision problems that are efficiently checkable using quantum computers. It turns out that these decision problems can be formulated naturally in terms of k-local Hamiltonians of the same type that appear in the proofs of the universality of AQC. Thus universality and Hamiltonian quantum complexity studies are often pursued hand in hand, and a reduction of k as well as the dimensionality p of the particles appearing in these constructions is one of the main goals. For example, already $k = 2$ and $p = 2$ leads to both universal AQC and QMA-complete Hamiltonians in 2D, while in 1D $p > 2$ is needed for both.[6]

We turn our attention to StoqAQC in Sec. VI. This is the setting of the vast majority of AQC work to date. The final Hamiltonian H_1 is assumed to be a classical Ising model Hamiltonian, typically (but not always) representing a hard optimization problem such as a spin glass. The initial Hamiltonian H_0 is typically assumed to be proportional to a transverse field, i.e., $\sum_i \sigma_i^x$, whose ground state is the uniform superposition state in the computational basis. AQC with stoquastic Hamiltonians is probably less powerful than universal quantum computation, but examples can be constructed which show that it may nevertheless be more powerful than classical computation. Moreover, if we relax the definition of AQC to allow for computation using excited states, it turns out that stoquastic Hamiltonians can even be QMA complete and support universal AQC. We first review examples where it is known that StoqAQC does not outperform classical computation (essentially because the eigenvalue gap Δ decreases rapidly with problem size but classical algorithms do not suffer a slowdown), then discuss examples where StoqAQC offers a quantum scaling advantage over simulated annealing in the sense that it outperforms classical simulated annealing but not necessarily other classical algorithms, and finally point out examples where it is currently not known whether StoqAQC offers a quantum speedup, but one hopes that it does. We also discuss the role of

---

[6]We say that H is a dD (d-dimensional) Hamiltonian if the particles are arranged on a d-dimensional grid and the summands of H couple only pairs of nearest-neighbor particles. Note that being dD implies that the Hamiltonian is 2-local.

potential quantum speedup mechanisms, in particular, tunneling and entanglement.

The somewhat bleak picture regarding StoqAQC should not necessarily be a cause for pessimism. Some of the obstacles in the way of a quantum speedup can be overcome or circumvented, as discussed in Sec. VII. In all cases this involves modifying some aspect of the Hamiltonian, either by optimizing the schedule $s(t)$ or by adding certain terms to the Hamiltonian such that small gaps are avoided. This can result in a nonstoquastic Hamiltonian whose final ground state is the same as that of the original Hamiltonian, with an exponentially small gap (often corresponding to a first-order quantum phase transition) changing into a polynomially small gap (often corresponding to a second-order phase transition). Another type of modification is to give up adiabatic evolution itself and allow for diabatic transitions. While this results in giving up the guarantee of convergence to the ground state provided by the adiabatic theorem, it can be a strategy that results in better run time scaling for the same Hamiltonian than an adiabatic one.

We conclude with an outlook and discussion of future directions in Sec. VIII. Various technical details are provided in the Appendix.

II. ADIABATIC THEOREMS

The origins of the quantum adiabatic approximation date back to Einstein's "Adiabatenhypothese": "If a system be affected in a reversible adiabatic way, allowed motions are transformed into allowed motions" (Einstein, 1914). Ehrenfest was the first to appreciate the importance of adiabatic invariance, guessing—before the advent of a complete quantum theory—that quantum laws would only allow motions which are invariant under adiabatic perturbations (Ehrenfest, 1916). The more familiar, modern version of the adiabatic approximation was put forth by Born and Fock (1928) for the case of discrete spectra, after the development of the Born-Oppenheimer approximation for the separation of electronic and nuclear degrees of freedom a year earlier (Born and Oppenheimer, 1927). Kato (1950) put the approximation on a firm mathematical foundation in 1950 and arguably proved the first quantum adiabatic theorem.

The adiabatic approximation states, roughly, that for a system initially prepared in an eigenstate (e.g., the ground state) $|\varepsilon_0(0)\rangle$ of a time-dependent Hamiltonian $H(t)$, the time evolution governed by the Schrödinger equation

$$i \frac{\partial |\psi(t)\rangle}{\partial t} = H(t)|\psi(t)\rangle \qquad (2)$$

(we set $\hbar \equiv 1$ from now on) will approximately keep the actual state $|\psi(t)\rangle$ of the system in the corresponding instantaneous ground state (or other eigenstate) $|\varepsilon_0(t)\rangle$ of $H(t)$, provided that $H(t)$ varies "sufficiently slowly." Quantifying the exact nature of this slow variation is the subject of the adiabatic theorem (AT), which exists in many variants. In this section we provide an overview of these variants of the AT, emphasizing aspects that are pertinent to AQC. We discuss the "folklore" adiabatic condition that the total evolution time t_f should be large on the time scale set by

the square of the inverse gap, and the question of how to ensure a high fidelity between the actual state and the ground state. We then discuss a variety of rigorous versions of the AT, emphasizing different assumptions and consequently different performance guarantees. Throughout this discussion, it is important to keep in mind that ultimately the AT provides only an *upper bound* on the evolution time required to achieve a certain fidelity between the actual state and the target eigenstate of $H(t)$.

A. Approximate versions

Let $|\varepsilon_j(t)\rangle$ ($j \in \{0, 1, 2, \ldots\}$) denote the instantaneous eigenstate of $H(t)$ with energy $\varepsilon_j(t)$ such that $\varepsilon_j(t) \leq \varepsilon_{j+1}(t) \, \forall \, j, t$, i.e., $H(t)|\varepsilon_j(t)\rangle = \varepsilon_j(t)|\varepsilon_j(t)\rangle$ and $j = 0$ denotes the (possibly degenerate) ground state. Assume that the initial state is prepared in one of the eigenstates $|\varepsilon_j(0)\rangle$.

The simplest as well as one of the oldest traditional versions of the adiabatic approximation states that a system initialized in an eigenstate $|\varepsilon_j(0)\rangle$ will remain in the same instantaneous eigenstate $|\varepsilon_j(t)\rangle$ (up to a global phase) for all $t \in [0, t_f]$, where t_f denotes the final time, provided (Messiah, 1962)

$$\max_{t \in [0, t_f]} \frac{|\langle \varepsilon_i | \partial_t \varepsilon_j \rangle|}{|\varepsilon_i - \varepsilon_j|} = \max_{t \in [0, t_f]} \frac{|\langle \varepsilon_i | \partial_t H | \varepsilon_j \rangle|}{|\varepsilon_i - \varepsilon_j|^2} \ll 1 \quad \forall \, i \neq j. \quad (3)$$

This version has been critiqued (Marzlin and Sanders, 2004; Tong *et al.*, 2005; Du *et al.*, 2008; Wu *et al.*, 2008) on the basis of arguments and examples involving a separate, independent time scale. Indeed, if the Hamiltonian includes an oscillatory driving term then the eigenstate population will oscillate with a time scale determined by this term, that is, independent of t_f, even if the adiabatic criterion (3) is satisfied.[7]

A more careful statement of the adiabatic condition that excludes such additional time scales is thus required. The first step is to assume that the Hamiltonian $H_{t_f}(t)$ in the Schrödinger equation $\partial |\psi_{t_f}(t)\rangle / \partial t = -i H_{t_f}(t)|\psi_{t_f}(t)\rangle$ can be written as $H_{t_f}(st_f) = H(s)$, where $s \equiv t/t_f \in [0, 1]$ is the dimensionless time, and $H(s)$ is t_f independent. This includes the "interpolating" Hamiltonians of the type often considered in AQC, i.e., $H(s) = A(s)H_0 + B(s)H_1$ [where $A(s)$ and $B(s)$ are monotonically decreasing and increasing, respectively] and excludes cases with multiple time scales.[8] The Schrödinger equation then becomes

---

[7]For example, it is easily checked that when $H(t) = a\sigma^z + b \sin(\omega t)\sigma^x$, the adiabatic condition (3) reduces to $|b\omega| \ll a^2$. However, even if this condition is satisfied the population can oscillate between the two eigenstates: at resonance (when $\omega \approx 2a$) the system undergoes Rabi oscillations with period $\pi/|b|$, a time scale that is independent of t_f.

[8]For example, a case such as $H(t) = a\sigma^z + b \sin(\omega t)\sigma^x$ is now excluded since after a change of variables we have $H(s) = a\sigma^z + b \sin(\omega t_f s)\sigma^x$ and evidently $H(s)$ still depends on t_f.

$$\frac{1}{t_f}\frac{\partial|\psi_{t_f}(s)\rangle}{\partial s} = -iH(s)|\psi_{t_f}(s)\rangle, \tag{4}$$

which is the starting point for all rigorous adiabatic theorems.

A more careful adiabatic condition subject to this formulation is given by (Amin, 2009)

$$\frac{1}{t_f}\max_{s\in[0,1]}\frac{|\langle\varepsilon_i(s)|\partial_s H(s)|\varepsilon_j(s)\rangle|}{|\varepsilon_i(s)-\varepsilon_j(s)|^2} \ll 1 \quad \forall\, j\neq i. \tag{5}$$

The conditions (3) and (5) give rise to the widely used criterion that the total adiabatic evolution time should be large on the time scale set by the minimum of the square of the inverse spectral gap $\Delta_{ij}(s) = \varepsilon_i(s) - \varepsilon_j(s)$. In most cases one is interested in the ground state, so that $\Delta_{ij}(s)$ is replaced by

$$\Delta \equiv \min_{s\in[0,1]}\Delta(s) = \min_{s\in[0,1]}\varepsilon_1(s) - \varepsilon_0(s). \tag{6}$$

However, arguments such as those leading to Eqs. (3) and (5) are approximate in the sense that they do not result in strict inequalities and do not result in bounds on the closeness between the actual time-evolved state and the desired eigenstate. We discuss this next.

B. Rigorous versions

The first rigorous adiabatic condition is due to Kato (1950) and was followed by numerous alternative derivations and improvements giving tighter bounds under various assumptions; see, e.g., Nenciu (1993), Avron and Elgart (1999), Hagedorn and Joye (2002), Teufel (2003), Ambainis and Regev (2004), Reichardt (2004), Jansen, Ruskai, and Seiler (2007), O'Hara and O'Leary (2008), Lidar, Rezakhani, and Hamma (2009), Cheung, Høyer, and Wiebe (2011), Elgart and Hagedorn (2012), and Ge, Molnár, and Cirac (2015). All these rigorous results are more severe in the gap condition than the traditional criterion, and they involve a power of the norm of time derivatives of the Hamiltonian, rather than a transition matrix element.

We summarize a few of these results here and refer the interested reader to the original literature for their proofs. For simplicity we always assume that the system is initialized in its ground state and that the gap is the ground state gap (6). We also assume that for all $s \in [0, 1]$ the Hamiltonian $H(s)$ has an eigenprojector $P(s)$ with eigenenergy $\varepsilon_0(s)$, and that the gap never vanishes, i.e., $\Delta > 0$.[9] The ground state, and hence the projector $P(s)$, is allowed to be (even infinitely) degenerate. $P(s)$ represents the "ideal" adiabatic evolution.

Let $P_{t_f}(s) = |\psi_{t_f}(s)\rangle\langle\psi_{t_f}(s)|$. This is the projector onto the time-evolved solution of the Schrödinger equation, i.e., the "actual" state. Adiabatic theorems are usually statements about the "instantaneous adiabatic distance" $\|P_{t_f}(s) - P(s)\|$ between the projectors associated with the actual and ideal evolutions, or the "final-time adiabatic distance" $\|P_{t_f}(1) - P(1)\|$. Typically, adiabatic theorems give a bound

[9]There is a weaker form of the AT, where one does not require a nonvanishing gap (Avron and Elgart, 1999). In this case, as in Theorem 2, the estimate on the error term is $o(1)$ as $t_f \to \infty$.

of the form $O(1/t_f)$ for the instantaneous case, and a bound of the form $O(1/t_f^n)$ for any $n \in \mathbb{N}$ for the final-time case. After squaring, these projector-distance bounds immediately become bounds on the transition probability, defined as $|\langle\psi_{t_f}^\perp(s)|\psi_{t_f}(s)\rangle|^2$, where $|\psi_{t_f}^\perp(s)\rangle = Q_{t_f}(s)|\psi_{t_f}(s)\rangle$, with $Q = I - P$.

1. Inverse cubic gap dependence with generic $H(s)$

Kato's work on the perturbation theory of linear operators (Kato, 1950) introduced techniques based on resolvents and complex analysis that have been widely used in subsequent work. Jansen, Ruskai, and Seiler (2007) proved several versions of the AT that build upon these techniques and that rigorously establish the gap dependence of t_f, without any strong assumptions on the smoothness of $H(s)$. Their essential assumption is that the spectrum of $H(s)$ has a band associated with the spectral projection $P(s)$ which is separated by a nonvanishing gap $\Delta(s)$ from the rest. Here we present one of their theorems.

Theorem 1. Suppose that the spectrum of $H(s)$ restricted to $P(s)$ consists of $m(s)$ eigenvalues separated by a gap $\Delta(s) = \varepsilon_1(s) - \varepsilon_0(s) > 0$ from the rest of the spectrum of $H(s)$, and that $H(s)$ is twice continuously differentiable. Assume that H, $H^{(1)}$, and $H^{(2)}$ are bounded operators, an assumption that is always fulfilled in finite-dimensional spaces.[10] Then for any $s \in [0, 1]$,

$$\|P_{t_f}(s) - P(s)\| \leq \frac{m(0)\|H^{(1)}(0)\|}{t_f\Delta^2(0)} + \frac{m(s)\|H^{(1)}(s)\|}{t_f\Delta^2(s)} + \frac{1}{t_f}\int_0^s\left(\frac{m\|H^{(2)}\|}{\Delta^2} + \frac{7m\sqrt{m}\|H^{(1)}\|^2}{\Delta^3}\right)dx. \tag{7}$$

The numerator depends on the norm of the first or second time derivative of $H(s)$, rather than the matrix element that appears in the traditional versions of the adiabatic condition.

Ignoring the m dependence for simplicity, this result shows that the adiabatic limit can be approached arbitrarily closely if (but not only if)

$$t_f \gg \max\left\{\max_{s\in[0,1]}\frac{\|H^{(2)}(s)\|}{\Delta^2(s)}, \max_{s\in[0,1]}\frac{\|H^{(1)}(s)\|^2}{\Delta^3(s)},\right.$$
$$\left.\max_{s\in[0,1]}\frac{\|H^{(1)}(s)\|}{\Delta^2(s)}\right\}. \tag{8}$$

Similar techniques based on Kato's approach can be used to prove a rigorous adiabatic theorem for open quantum systems, where the evolution is generated by a non-Hermitian Liouvillian instead of a Hamiltonian (Venuti *et al.*, 2016).

2. Rigorous inverse gap squared

A version of the AT that yields a scaling of t_f with the inverse of the gap squared (up to a logarithmic correction) was

[10]We use the notation $H^{(k)}(s) \equiv (\partial/\partial x)^k H(x)|_s$ throughout.

given by Elgart and Hagedorn (2012). All other rigorous AT versions to date have a worse gap dependence (cubic or higher). The proof introduces assumptions on $H(s)$ that go beyond those of Theorem 1. Namely, it is assumed that $H(s)$ is bounded and infinitely differentiable, and the higher derivatives cannot have a magnitude that is too large, or more specifically, that $H(s)$ belongs to the Gevrey class G^α.

Definition 2 (Gevrey class). $H(s) \in G^\alpha$ if $dH(s)/ds \neq 0 \,\forall\, s \in [0, 1]$ and there exist constants $C, R > 0$, such that for all $k \geq 1$,

$$\max_{s\in[0,1]} \|H^{(k)}(s)\| \leq CR^k k^{\alpha k}. \qquad (9)$$

An example is $H(s) = [1 - A(s)]H_0 + A(s)H_1$, where $A(s) = c \int_{-\infty}^{s} \exp[-1/(x - x^2)]dx$ if $s \in (0, 1)$, and $A(s) = 0$ if $s \notin [0, 1]$. The constant c is chosen so that $A(1) = 1$. For this family $\|H^{(k)}(s)\| = |A^{(k)}(s)|\|H_1 - H_0\| \leq Ck^{2k}$, so that $H(s) \in G^2$.

The AT due to Elgart and Hagedorn (2012) can now be stated as follows.

Theorem 2. Assume that $H(s)$ is bounded and belongs to the Gevrey class G^α with $\alpha > 1$, and that $\Delta \ll h$, where $h \equiv \|H(0)\| = \|H(1)\|$. If

$$t_f \geq \frac{K}{\Delta^2}|\ln(\Delta/h)|^{6\alpha} \qquad (10)$$

for some Δ-independent constant $K > 0$ (with units of energy), then the distance $\|P_{t_f}(s) - P(s)\|$ is $o(1) \,\forall\, s \in [0, 1]$.

This result is remarkable in that it rigorously gives an inverse gap squared dependence, which is essentially tight due to the existence of a lower bound of the form $t_f = O(\Delta^{-2}/|\ln \Delta|)$ for Hamiltonians satisfying rank$H(1) \ll \dim(\mathcal{H})$ (Cao and Elgart, 2012). However, the error bound is not tight, and we address this next.

3. Arbitrarily small error

Building on work originating with Nenciu (1993) [see also Hagedorn and Joye (2002)], Ge, Molnár, and Cirac (2015) proved a version of the AT that results in an exponentially small error bound in t_f. The inverse gap dependence is cubic.

Assume for simplicity that $\varepsilon_0(s) = 0$ and choose the phase of $|\varepsilon_0(s)\rangle$ so that $\langle \dot{\varepsilon}_0(s)|\varepsilon_0(s)\rangle = 0$, where the dot denotes ∂_s.

Theorem 3. Assume that all derivatives of the Hamiltonian $H(s)$ vanish at $s = 0, 1$, and moreover that it satisfies the following Gevrey condition: there exist constants $C, R, \alpha > 0$ such that for all $k \geq 1$,

$$\max_{s\in[0,1]} \|H^{(k)}(s)\| \leq CR^k \frac{(k!)^{1+\alpha}}{(k+1)^2}. \qquad (11)$$

Then the adiabatic error is bounded as

$$\min_{\theta}\|\|\psi_{t_f}(1)\rangle - e^{i\theta}|\varepsilon_0(1)\rangle\|\| \leq c_1 \frac{C}{\Delta} e^{-(c_2\Delta^3 t_f/C^2)^{1/(1+\alpha)}}, \qquad (12)$$

where $c_1 = eR(8\pi^2/3)^3$ and $c_2 = (3/4\pi^2)^5/(4eR^2)$.

Thus, as long as $t_f \gg C^2/\Delta^3$, the adiabatic error is exponentially small in t_f.

The idea of using vanishing boundary derivatives dates back at least to Garrido and Sancho (1962). It was also used by Lidar, Rezakhani, and Hamma (2009) for a different class of functions than the Gevrey class: functions that are analytic in a strip of width 2γ in the complex time plane and have a finite number V of vanishing boundary derivatives, i.e., $H^{(v)}(0) = H^{(v)}(1) = 0 \,\forall\, v \in [1, V]$. The adiabatic error is then upper bounded by $(V + 1)^{\gamma+1}q^{-V}$ as along as $t_f \geq (q/\gamma)V\max_s\|H_V^{(1)}(s)\|^2/\Delta^3$, where $q > 1$ is a parameter that can be optimized given knowledge of $\|H_V^{(1)}\|$. Thus, the adiabatic error can be made arbitrarily small in the number of vanishing derivatives, while the scaling of t_f with V is encoded into $\|H_V^{(1)}\|$.[11] An example of a function whose first V derivatives vanish at the boundaries $s = 0$, 1 is the regularized β function (Rezakhani, Pimachev, and Lidar, 2010)

$$A(s) = \frac{\int_0^s x^V(1 - x)^V dx}{\int_0^1 x^V(1 - x)^V dx}.$$

It is possible to further reduce the error quadratically in t_f using an interference effect that arises from imposing an additional boundary symmetry condition (Wiebe and Babcock, 2012).

Note that an important difference between Theorems 2 and 3 is that the former applies for all times $s \in [0, 1]$ ("instantaneous AT"), while the latter applies only at the final time $s = 1$ ("final-time AT"), which typically gives rise to tighter error bounds.

Also note that Landau and Zener already showed that the transition probability out of the ground state is $O(e^{-C\Delta^2 t_f})$ (Landau, 1932; Zener, 1932) [see Joye (1994) for a rigorous proof for analytic Hamiltonians], thus combining an inverse gap square dependence with an exponentially small error bound. However, this result holds only for two-level systems.

4. Lower bound

Let $H(s)$ with $s \in [0, 1]$ be a given continuous Hamiltonian path and $|\varepsilon(s)\rangle$ the corresponding nondegenerate eigenstate path (eigenpath). In the so-called black-box model the only assumption is to be able to evolve with $H[s(t)]$ for some schedule $s(t)$ (here s is allowed to be a general function of t),

---

[11]This corrects an omission by Lidar, Rezakhani, and Hamma (2009), where the dependence of $\|H^{(1)}\|$ on V was ignored since the supremum of $\|H^{(1)}(s)\|$ was taken over $s \in [0, 1]$ instead of over the region of analyticity of $H(s)$, as noted by Ge, Molnár, and Cirac (2015).

without exploiting the unknown structure of $H(s)$. Define the path length L as

$$L = \int_0^1 \||\dot{\varepsilon}(s)\rangle\| ds, \qquad (13)$$

where the dot denotes ∂_s. Assuming, without loss of generality, that the phase of $|\varepsilon(s)\rangle$ is chosen so that $\langle \varepsilon(s)|\dot{\varepsilon}(s)\rangle = 0$, then L is the only natural length in projective Hilbert space (up to irrelevant normalization factors).

It was shown by Boixo and Somma (2010) that there is a lower bound on the time required to prepare $|\varepsilon(1)\rangle$ from $|\varepsilon(0)\rangle$ with bounded precision

$$t_f > O(L/\Delta). \qquad (14)$$

Since an upper bound on L is $\max_s \|\dot{H}(s)\|/\Delta$,[12] one obtains the estimate $t_f \sim O[\max_s \|\dot{H}(s)\|/\Delta^2]$, reminiscent of the approximate versions of the adiabatic condition [Eq. (5)]. The proof of the lower bound is essentially based on the optimality of the Grover search algorithm.

The lower bound is nearly achievable using a "digital," nonadiabatic method proposed by Boixo, Knill, and Somma (2010) that does not require path continuity or differentiability. The time required scales as $O[(L/\Delta) \log(L/\epsilon)]$, where ϵ is a specified bound on the error of the output state $|\varepsilon(1)\rangle$. L is the angular length of the path and is suitably defined to generalize Eq. (13) to the nondifferentiable case.

Armed with an arsenal of adiabatic theorems we are now well equipped to start surveying AQC algorithms.

III. ALGORITHMS

In this section we review the algorithms which are known to provide quantum speedups over classical algorithms. However, to make the idea of a quantum speedup precise we need to draw distinctions among different types of speedups, as several such types will arise in the course of this review. Toward this end we adopt a classification of quantum speedup types proposed by Rønnow et al. (2014). The classification is the following, in decreasing order of strength.

- A "provable" quantum speedup is the case where there exists a proof that no classical algorithm can outperform a given quantum algorithm. The best known example is Grover's search algorithm (Grover, 1997), which, in the query complexity setting, exhibits a provable quadratic speedup over the best possible classical algorithm (Bennett et al., 1997).
- A "strong" quantum speedup was originally defined by Papageorgiou and Traub (2013) by comparing a quantum algorithm against the performance of the best classical algorithm, whether such a classical algorithm is explicitly known or not. This aims to capture computational complexity considerations allowing for the

existence of yet-to-be discovered classical algorithms. Unfortunately, the performance of the best possible classical algorithm is unknown for many interesting problems (for example, for factoring).
- A "quantum speedup" (unqualified, without adjectives) is a speedup against the best available classical algorithm [for example, Shor's polynomial-time factoring algorithm (Shor, 1994)]. Such a speedup may be tentative in the sense that a better classical algorithm may eventually be found.
- Finally, a "limited quantum speedup" is a speedup obtained when compared specifically with classical algorithms that "correspond" to the quantum algorithm in the sense that they implement the same algorithmic approach, but on classical hardware. This definition allows for the existence of other classical algorithms that are already better than the quantum algorithm. The notion of a limited quantum speedup will turn out to be particularly useful in the context of StoqAQC.

A refinement of this classification geared at experimental quantum annealing was given by Mandrà et al. (2016).

Using this classification, this section collects most of the adiabatic quantum algorithms known to give a provable quantum speedup (Grover, Deutsch-Jozsa, Bernstein-Vazirani, and glued trees), or just a quantum speedup (PAGERANK).[13]

Many other adiabatic algorithms have been proposed, and we review a large subset of these in Sec. VI. In a few of these cases there is a scaling advantage over classical simulated annealing, while in some cases there are definitely faster classical algorithms.

A. Adiabatic Grover

The adiabatic Grover algorithm (Roland and Cerf, 2002) is perhaps the hallmark example of a provable quantum speedup using AQC, so we review it in detail. As in the circuit model Grover algorithm (Grover, 1997), informally the objective is to find the marked item (or possibly multiple marked items) in an unsorted database of N items by accessing the database as few times as possible. More formally, one is allowed to call a function $f: \{0,1\}^n \mapsto \{0,1\}$ (where $N = 2^n$ is the number of bit strings) with the promise that $f(m) = 1$ and $f(x) = 0 \,\forall\, x \neq m$, and the goal is to find the unknown index m in the smallest number of calls. This is an oracular problem (Nielsen and Chuang, 2000), in that the algorithm can make queries to an oracle that recognizes the marked items. The oracle remains a black box, i.e., the details of its implementation and its complexity are ignored. This allows for an uncontroversial determination of the complexity of the algorithm in terms of the number of queries to the oracle.

For a classical algorithm, the only strategy is to query the oracle until the marked item is found. Whether the classical algorithm uses no memory, i.e., the algorithm does not keep

---

[12]See Appendix A.1 as well as Appendix G of Boixo, Knill, and Somma (2009).

[13]The glued trees case is strictly not an adiabatic quantum algorithm, since it explicitly makes use of excited states. Also, in the PAGERANK case the evidence for a quantum speedup is numerical.

track of items that have already been checked, or uses an exponential amount of memory (in n) to store all the items that have been checked, the classical algorithm will have an average number of queries that scales linearly in N.

In the AQC algorithm we denote the marked item by the binary representation of m. The oracle is defined in terms of the final Hamiltonian $H_1 = \mathbb{1} - |m\rangle\langle m|$, where $|m\rangle$ is the marked state associated with the marked item. In this representation, the binary representations give the eigenvalues under σ^z, i.e., $\sigma^z|0\rangle = +|0\rangle$ and $\sigma^z|1\rangle = -|1\rangle$. The marked state is the ground state of this Hamiltonian with energy 0, and all other computational basis states have energy 1.

1. Setup for the adiabatic quantum Grover algorithm

We use the initial Hamiltonian $H_0 = \mathbb{1} - |\phi\rangle\langle\phi|$, where $|\phi\rangle$ is the uniform superposition state,

$$|\phi\rangle = \frac{1}{\sqrt{N}} \sum_{i=0}^{N-1} |i\rangle = |+\rangle^{\otimes n}, \qquad (15)$$

where $|\pm\rangle = (1/\sqrt{2})(|0\rangle \pm |1\rangle)$. We take the time-dependent Hamiltonian to be an interpolation:

$$\begin{aligned} H(s) &= [1 - A(s)]H_0 + A(s)H_1 \\ &= [1 - A(s)](\mathbb{1} - |\phi\rangle\langle\phi|) + A(s)(\mathbb{1} - |m\rangle\langle m|), \end{aligned} \qquad (16)$$

where $s = t/t_f \in [0, 1]$ is the dimensionless time, t_f is the total computation time, and $A(s)$ is a schedule that can be optimized. For simplicity, we first consider a linear schedule $A(s) = s$. Note that H_1 is n local.

If the initial state is initialized in the ground state of $H(0)$, i.e., $|\psi(0)\rangle = |\phi\rangle$, then the evolution of the system is restricted to a two-dimensional subspace, defined by the span of $|m\rangle$ and

$$|m^\perp\rangle = \frac{1}{\sqrt{N-1}} \sum_{i \neq m}^{N-1} |i\rangle.$$

In this two-dimensional subspace $H(s)$ can be written as

$$[H(s)]_{|m\rangle,|m^\perp\rangle} = \frac{1}{2}\mathbb{1}_{2\times 2} - \frac{\Delta(s)}{2} \begin{pmatrix} \cos\theta(s) & \sin\theta(s) \\ \sin\theta(s) & -\cos\theta(s) \end{pmatrix}, \qquad (17)$$

where

$$\Delta(s) = \sqrt{(1-2s)^2 + \frac{4}{N} s(1-s)}, \qquad (18a)$$

$$\cos\theta(s) = \frac{1}{\Delta(s)} \left[1 - 2(1-s)\left(1 - \frac{1}{N}\right) \right], \qquad (18b)$$

$$\sin\theta(s) = \frac{2}{\Delta(s)} (1-s) \frac{1}{\sqrt{N}} \sqrt{1 - \frac{1}{N}}. \qquad (18c)$$

The eigenvalues and eigenvectors in this subspace are then given by

$$\varepsilon_0(s) = \tfrac{1}{2}[1 - \Delta(s)], \qquad \varepsilon_1(s) = \tfrac{1}{2}[1 + \Delta(s)], \qquad (19a)$$

$$|\varepsilon_0(s)\rangle = \cos\frac{\theta(s)}{2} |m\rangle + \sin\frac{\theta(s)}{2} |m^\perp\rangle, \qquad (19b)$$

$$|\varepsilon_1(s)\rangle = -\sin\frac{\theta(s)}{2} |m\rangle + \cos\frac{\theta(s)}{2} |m^\perp\rangle. \qquad (19c)$$

The remaining $N - 2$ eigenstates of $H(s)$ have eigenvalue 1 throughout the evolution. The minimum gap occurs at $s = 1/2$ and scales exponentially with n:

$$\Delta_{\min} = \Delta(s = 1/2) = \frac{1}{\sqrt{N}} = 2^{-n/2}. \qquad (20)$$

(This can be viewed as a special case of Lemma 1.)

In our discussion of the adiabatic theorem we saw that without special assumptions on $s(t)$ except that it is twice differentiable, the adiabatic condition is inferred from Eq. (7), which requires setting

$$t_f \gg 2\max_s \|\partial_s H(s)\|/\Delta^2(s) + \int_0^1 \|\partial_s H(s)\|^2/\Delta^3(s) ds,$$

where we have accounted for the boundary conditions and used the positivity of the integrand to extend the upper limit to 1.[14] Differentiating Eq. (17) yields

$$\partial_s H(s) = \begin{pmatrix} -(1 - \frac{1}{N}) & \frac{1}{\sqrt{N}} \sqrt{1 - \frac{1}{N}} \\ \frac{1}{\sqrt{N}} \sqrt{1 - \frac{1}{N}} & 1 - \frac{1}{N} \end{pmatrix}, \qquad (21)$$

which has eigenvalues $\pm\sqrt{1 - 1/N}$ so that $\|\partial_s H\| \leq 1$. The other integrand in Eq. (7), involving $\|\partial_s^2 H(s)\|/\Delta^2(s)$, vanishes after differentiating Eq. (21). The ground state degeneracy $m(s) = 1$ throughout. Since

$$\int_0^s 1/\Delta^3(x) dx = \frac{N}{2} - \frac{N^{3/2}(1 - 2s)}{2\sqrt{N(1 - 2s)^2 + 4(1 - s)s}},$$

which is a monotonically increasing function of s that approaches $N = \Delta_{\min}^{-2}$ as $s \to 1$, the adiabatic condition becomes

$$t_f \gg 2\max_s \frac{1}{\Delta^2(s)} + \int_0^1 ds \frac{1}{\Delta^3(s)} = \frac{3}{\Delta_{\min}^2}. \qquad (22)$$

This suggests the disappointing conclusion that the quantum adiabatic algorithm scales in the same way as the classical algorithm.

However, by imposing the adiabatic condition globally, i.e., to the entire time interval t_f, the evolution rate is constrained throughout the whole computation, while the gap becomes small only around $s = 1/2$. Thus, it makes sense to use a schedule $A(s)$ that adapts and slows down near the minimum gap, but speeds up away from it (van Dam, Mosca, and Vazirani, 2001;

---

[14]Whenever we use the $\gg$ symbol we mean that the larger quantity should be larger by some large multiplicative constant, such as 100.

Roland and Cerf, 2002) [this is related to the idea of rapid adiabatic passage, which has a long history in nuclear magnetic resonance (Powles, 1958)]. By doing so the quadratic quantum speedup can be recovered as we address next.

2. Quadratic quantum speedup

Consider again the adiabatic condition (7), which we can rewrite as

$$t_f \gg 2\max_s \frac{\|\partial_s H(s)\|}{\Delta^2(s)} + \int_0^1 \left(\frac{\|\partial_s^2 H\|}{\Delta^2} + \frac{\|\partial_s H\|^2}{\Delta^3} \right) ds, \quad (23)$$

where now H and Δ depend on a schedule $A(s)$. Let us now use the ansatz (Roland and Cerf, 2002; Jansen, Ruskai, and Seiler, 2007)

$$\partial_s A = c\Delta^p[A(s)], \qquad A(0) = 0, \qquad p, c > 0. \quad (24)$$

This schedule slows down as the gap becomes smaller as desired. The normalization constant $c = \int_0^1 \Delta^{-p}[A(s)]\partial_s A ds = \int_{A(0)}^{A(1)} \Delta^{-p}(u) du$ [using $u = A(s)$] is chosen to ensure that $A(1) = 1$.

It follows that

$$\int_0^1 \left(\frac{\|\partial_s^2 H[A(s)]\|}{\Delta^2[A(s)]} + \frac{\|\partial_s H[A(s)]\|^2}{\Delta^3[A(s)]} \right) ds$$

$$\leq 4c \int_0^1 \Delta^{p-3}(u) du \quad (25)$$

(the proof is given in Appendix A.2). Finally, the boundary term in Eq. (23) yields

$$2\max_s \frac{\|\partial_s H(s)\|}{\Delta^2(s)} \leq 4c\Delta_{\min}^{p-2}.$$

The case $p = 2$ serves to illustrate the main point. In this case the boundary term is $4c$ and evaluating the integrals yields

$$c = \int_0^1 \Delta^{-2}(u) du$$

$$= \frac{N}{\sqrt{N-1}} \tan^{-1} \sqrt{N-1} \to \frac{\pi}{2} \sqrt{N},$$

$$\int_0^1 \Delta^{-1}(u) du$$

$$= \frac{\log \left[\sqrt{N-1}\sqrt{N} + N - 1/\sqrt{N-1}\sqrt{N} - (N-1) \right]}{2\sqrt{N-1}/N}$$

$$\to \log(2N)/2,$$

where the asymptotic expressions are for $N \gg 1$. Substituting this into Eq. (25) yields the adiabatic condition

$$t_f \gg 2\pi\sqrt{N}[1 + \log(2N)], \quad (26)$$

which is a sufficient condition for the smallness of the adiabatic error and nearly recovers the quadratic speedup expected from Grover's algorithm.

The appearance of the logarithmic factor latter is actually an artifact of using bounds that are not tight.[15] The quadratic speedup, i.e., the scaling of t_f with $\sqrt{N}$, can be fully recovered by solving for the schedule from Eq. (24) in the $p = 2$ case (Roland and Cerf, 2002). We first rewrite Eq. (24) in dimensional time units as $\partial_t A = c'\Delta^2[A(t)]$, with the boundary conditions $A(0) = 0$ and $A(t_f) = 1$. To solve this differential equation we rewrite it as

$$t = \int_0^t dt = \int_{A(0)}^{A(t)} dA/[c'\Delta^2(A)].$$

After integration we obtain

$$t = \frac{N}{2c'\sqrt{N-1}} \{\tan^{-1}(\sqrt{N-1}[2A(t) - 1]) + \tan^{-1}\sqrt{N-1}\}. \quad (27)$$

Evaluating Eq. (27) at t_f gives

$$t_f = \frac{N}{c'\sqrt{N-1}} \tan^{-1}\sqrt{N-1} \to \frac{\pi}{2c'}\sqrt{N}, \quad (28)$$

which is the expected quadratic quantum speedup.

One may be tempted to conclude that t_f can be made arbitrarily small since so far c' is arbitrary and can be chosen to be large. However, the adiabatic error bound (26) shows that this is not the case: while it is not tight, it suggests that if t_f scales as $\sqrt{N}$ then c' must scale as $1/\log(2N)$ in order to keep the adiabatic error small. Thus, the general conclusion is that increasing c' results in a larger adiabatic error.[16]

Inverting Eq. (27) for $A(t)$ [or, equivalently, solving Eq. (24) for $p = 2$] gives the locally optimized schedule

$$A(s) = \frac{1}{2} + \frac{1}{2\sqrt{N-1}} \tan\left[(2s - 1)\tan^{-1}\sqrt{N-1} \right], \quad (29)$$

where we replaced t/t_f [with t_f given by Eq. (28)] with s. As expected, this schedule rises rapidly near $s = 0$, 1 and is nearly flat around $s = 1/2$, i.e., it slows down near the minimum gap.

Since the choice in Eq. (24) is not unique, we wonder if there exists a schedule that gives an even better scaling. Given that Grover's algorithm is known to be optimal in the circuit model setting (Bennett et al., 1997; Zalka, 1999), this is, not surprisingly, not the case, and a general argument to that effect which applies to any Hamiltonian quantum computation was given by Farhi and Gutmann (1998). We review this argument in the AQC setting in Appendix B.

---

[15]A detailed analysis of the adiabatic Grover algorithm along with tighter error bounds than given here was presented by Rezakhani, Pimachev, and Lidar (2010).

[16]Note that the scaling conclusion $t_f \sim \sqrt{N}$ reported by Roland and Cerf (2002) is based on the interpretation of Eq. (24) as a heuristic "local" adiabatic condition and does not constitute a proof that the adiabatic error is small. The evidence that the rigorous bound (26) is not tight and that $t_f \sim \sqrt{N}$ suffices to achieve a small adiabatic error for the schedule (29) is numerical.

3. Multiple marked states

The present results generalize easily to the case where we have $M \geq 1$ marked states, for which Grover's algorithm is known to also give a quadratic speedup in the circuit model (Boyer et al., 1998; Biham et al., 1999). The final Hamiltonian can be written as

$$H_1 = \mathbb{1} - \sum_{m \in \mathcal{M}} |m\rangle\langle m|, \qquad (30)$$

where $\mathcal{M}$ is the index set of the marked states. Let

$$|m_\perp\rangle = \frac{1}{\sqrt{N-M}} \sum_{i \notin \mathcal{M}} |i\rangle\langle i|. \qquad (31)$$

Instead of evolving in a two-dimensional subspace, the system evolves in an $(M+1)$-dimensional subspace spanned by $(\{|m\rangle\}_{m \in \mathcal{M}}, |m_\perp\rangle)$, and instead of Eq. (17), the Hamiltonian can be written in this basis as

$$H(s) = \begin{pmatrix} (1-s)(1-\frac{1}{N}) & -\frac{1-s}{N} & \cdots & -(1-s)\frac{\sqrt{N-M}}{N} \\ -\frac{1-s}{N} & (1-s)(1-\frac{1}{N}) & -\frac{1-s}{N} \cdots & -(1-s)\frac{\sqrt{N-M}}{N} \\ \vdots & & \ddots & \vdots \\ -(1-s)\frac{\sqrt{N-M}}{N} & -(1-s)\frac{\sqrt{N-M}}{N} & \cdots & s+(1-s)(1-\frac{N-M}{N}) \end{pmatrix}. \qquad (32)$$

This Hamiltonian can be easily diagonalized, and one finds that there are $M-1$ eigenvalues equal to $1-s$, and two eigenvalues

$$\lambda_\pm = \frac{1}{2} \pm \frac{1}{2}\sqrt{(1-2s)^2 + \frac{4M}{N}s(1-s)} \qquad (33)$$

that determine the relevant minimum gap

$$\Delta(s) = \sqrt{(1-2s)^2 + \frac{4M}{N}s(1-s)}.$$

The remaining $N-M-1$ eigenvalues of the unrestricted Hamiltonian are equal to 1. Comparing the $M=1$ case Eq. (18a) to the present case, the only difference is the change from $1/N$ to M/N. Therefore our discussion from earlier goes through with only this modification.

In closing, we note that an experimentally realizable version of the adiabatic Grover search algorithm using a single bosonic particle placed in an optical lattice was recently proposed by Hen (2017).

B. Adiabatic Deutsch-Jozsa algorithm

Given a function $f : \{0,1\}^n \mapsto \{0,1\}$ which is promised to be either constant or balanced [i.e., $f(x) = 0$ on half the inputs and $f(x) = 1$ on the other half], the Deutsch-Jozsa problem is to determine which type the function is. There exists a quantum circuit model algorithm that solves the problem in a single f query (Deutsch and Jozsa, 1992). Classically, the problem requires $2^{n-1} + 1$ f queries in the worst case, since it is possible that the first 2^{n-1} queries return a constant answer, while the function is actually balanced. It is important to note that the quantum advantage requires a deterministic setting, since the classical error probability is exponentially small in the number of queries.

An adiabatic implementation of the Deutsch-Jozsa algorithm using unitary interpolation was given by Sarandy and Lidar (2005) and an implementation using a linear interpolation was given by Wei and Ying (2006). These algorithms match the speedup obtained in the circuit model [for an earlier example where this is not the case see Das, Kobes, and Kunstatter (2002)], and we proceed to review both. We note that, just like the adiabatic Grover's algorithm, the adiabatic Deutsch-Jozsa algorithm requires n-local Hamiltonians. We also note that both the unitary interpolation and linear interpolation strategies are not unique to the Deutsch-Jozsa problem and apply equally well to any depth-one quantum circuit. Thus they should be viewed in this more general context and are used here with a specific algorithm for illustrative purposes.

1. Unitary interpolation

The initial Hamiltonian is chosen such that its ground state is the uniform superposition state $|\phi\rangle$ [Eq. (15)] and $N = 2^n$, i.e., $H(0) = \omega \sum_{i=1}^{n} |-\rangle_i\langle-|$, where ω is the energy scale. The Deutsch-Jozsa problem can be solved by a single computation of the function f through the unitary transformation $U|x\rangle = (-1)^{f(x)}|x\rangle$ ($x \in \{0,1\}^n$) (Collins, Kim, and Holton, 1998), so that in the $\{|x\rangle\}$ (computational) basis U is represented by the diagonal matrix $U = \text{diag}[(-1)^{f(0)}, ..., (-1)^{f(2^n-1)}]$. An adiabatic implementation requires a final Hamiltonian $H(1)$ such that its ground state is $|\psi(1)\rangle = U|\psi(0)\rangle$. This can be accomplished via a unitary transformation of $H(0)$, i.e., $H(1) = UH(0)U^\dagger$. Then the final Hamiltonian encodes the solution of the Deutsch problem in its ground state, which can be extracted via a measurement of the qubits in the $\{|+\rangle, |-\rangle\}$ basis (note that this is compatible with the definition of AQC, Definition 1, which does not restrict the measurement basis). A suitable unitary interpolation between $H(0)$ and $H(1)$ can be defined by $H(s) = \tilde{U}(s)H(0)\tilde{U}^\dagger(s)$, where $\tilde{U}(s) = \exp[i(\pi/2)sU]$, for which $\tilde{U}(1) = iU$. Since a unitary transformation of $H(0)$ preserves its spectrum, it does not change the ground state gap which remains ω. The run time of the algorithm can be determined from the adiabatic condition (8), and what remains is the numerator

$$\|H^{(1)}(s)\| = \left\| i\frac{\pi}{2}[U, H(s)] \right\| \leq \pi\|H(0)\| = \pi$$

[and similarly for $\|H^{(2)}(s)\|$]. This yields $t_f \gg 1/\omega$. This result is independent of n so the adiabatic run time is $O(1)$.

2. Linear interpolation

Unitary interpolations, introduced by Siu (2005), are somewhat less standard. Therefore we also present the standard linear interpolation method as an alternative. Consider the usual initial Hamiltonian $H_0 = \mathbb{1} - |\phi\rangle\langle\phi|$ over n qubits, where once again $|\phi\rangle$ is the uniform superposition state. Let the final Hamiltonian be $H_1 = \mathbb{1} - |\psi\rangle_f\langle\psi|$, where

$$|\psi\rangle_f = \frac{\mu_f}{\sqrt{N/2}}\sum_{i=0}^{N/2-1}|2i\rangle + \frac{1-\mu_f}{\sqrt{N/2}}\sum_{i=0}^{N/2-1}|2i+1\rangle, \quad (34)$$

and

$$\mu_f = \left|\frac{1}{N}\sum_{x\in\{0,1\}^n}(-1)^{f(x)}\right|. \quad (35)$$

Clearly, $\mu_f = 1$ or 0 if f is constant or balanced, respectively. Therefore $|\psi\rangle_f$ is a uniform superposition over all even or odd index states if f is constant or balanced, respectively, and a measurement of the ground state of H_1 in the computational basis reveals whether f is constant or balanced, depending on whether the observed state belongs to the even or odd sector, respectively. However, we note that one may object to the reasonableness of the final Hamiltonian H_1. Namely, preparing the state $|\psi\rangle_f$ involves precomputing the quantity μ_f, which directly encodes whether f is constant or balanced and so may be thought to represent an oracle that is too powerful.[17] Indeed, H_1 in this construction is not of the standard form $\sum_x f(x)|x\rangle\langle x|$, where each oracle call $f(x)$ corresponds to a query about a single basis state $|x\rangle$. Therefore, there is no classical analog to this oracle in the computational basis.

Setting aside this concern in the present version of the algorithm due to Wei and Ying (2006), it remains to determine the adiabatic run time for the adiabatic Hamiltonian $H(s) = (1-s)H_0 + sH_1$. The following Lemma (Aharonov and Ta-Shma, 2003) comes in handy:

Lemma 1. Let $|\alpha\rangle$ and $|\beta\rangle$ be two states in some subspace of an N-dimensional Hilbert space $\mathcal{H}$, and let $H_\alpha = \mathbb{1} - |\alpha\rangle\langle\alpha|$, $H_\beta = \mathbb{1} - |\beta\rangle\langle\beta|$. For any convex combination $H_\eta = (1-\eta)H_\alpha + \eta H_\beta$, where $\eta \in [0,1]$, the ground state gap $\Delta(H_\eta) \geq |\langle\alpha|\beta\rangle|$.

Proof. Expand $|\beta\rangle = a|\alpha\rangle + b|\alpha^\perp\rangle$ where $\langle\alpha|\alpha^\perp\rangle = 0$, and complete $\{|\alpha\rangle, |\alpha^\perp\rangle\}$ to an orthonormal basis for $\mathcal{H}$. Writing H_η in this basis yields

$$H_\eta = \begin{pmatrix} \eta|b|^2 & \eta ab^* \\ \eta a^*b & \eta|a|^2 + 1 - \eta \end{pmatrix} \oplus \mathbb{1}_{(N-2)\times(N-2)}. \quad (36)$$

The eigenvalues of this matrix are all 1 for the identity matrix block and the difference between the eigenvalues in the 2×2 block is $\Delta(H_\eta) = \sqrt{1 - 4\eta(1-\eta)|b|^2}$. This is minimized for $\eta = 1/2$, where it equals $|a|$.   ∎

Applying this lemma, we see that $\Delta[H(s)] \geq |\langle\phi|\psi_f\rangle| = 1/\sqrt{2}$. Since $\|H^{(1)}(s)\| = \|H_1 - H_0\| \leq 2$ and $\|H^{(2)}(s)\| = 0$, it follows from the adiabatic condition (8) that t_f is independent of n, i.e., the adiabatic run time is $O(1)$ as in the circuit model depth.

3. Interpretation

As mentioned, a classical probabilistic algorithm that simply submits random queries to the oracle will fail with a probability that is exponentially small in the number of queries. One might thus be concerned that the adiabatic algorithms are no better (Hen, 2014a), since they are probabilistic in the sense that there is a nonzero probability of ending in an excited state. However, for the linear interpolation adiabatic algorithm reviewed, measuring the energy of the final state returns 0 in the ground state or 1 in an excited state. In the latter "inconclusive" case, the algorithm needs to be repeated until an energy of 0 is found and only then is a computational basis measurement performed. If an even (or odd) index state is measured, the corresponding constant (or balanced) result is guaranteed to be correct. Moreover, an excited state outcome can (and should) be made exponentially unlikely using a smooth schedule as per Theorem 3. Thus, the adiabatic algorithms improve upon a classical probabilistic algorithm in the following sense: In the adiabatic case, to know with certainty that the function is constant or balanced (an even or odd index measurement result) happens with probability $p = 1 - q$ where $q \sim e^{-t_f}$, where the run time t_f is independent of n. Therefore, the expected number of runs r to certainty in the adiabatic case is

$$\langle r \rangle = \sum_{r=1}^{\infty} pq^{r-1}r = \frac{1}{1-q} \approx 1 + q \sim 1 + e^{-t_f}. \quad (37)$$

On the other hand, classically, to know with certainty that the function is constant requires $N = 2^n/2 + 1$ runs or queries (all yielding identical outcomes).

Finally, we note that there exists a nonadiabatic Hamiltonian quantum algorithm that solves the Deutsch-Jozsa problem in constant time with a deterministic guarantee of ending up with the right answer (i.e., in the ground state) (Hen, 2014a). This algorithm is based on finding a fine-tuned schedule $s(t)$.

C. Adiabatic Bernstein-Vazirani algorithm

The Bernstein and Vazirani (1993) problem is to find an unknown binary string $a \in \{0,1\}^n$ with as few queries as possible of the function (or oracle)

$$f_a(w) = w \odot a \in \{0,1\}, \quad (38)$$

---

[17]The only efficient way known to compute a quantity similar to μ_f involves running the Deutsch-Jozsa algorithm in the gate model, where $(1/N)\sum_{x\in\{0,1\}^n}(-1)^{f(x)}$, without the absolute value, appears as the amplitude of the state $|0\rangle^{\otimes n}$ (Nielsen and Chuang, 2000).

where $\odot$ denotes the bitwise inner product modulo 2, and $w \in \{0,1\}^n$ as well. In the quantum circuit model, it can be shown that a can be determined with $O(1)$ queries (Bernstein and Vazirani, 1993) whereas classical algorithms require n queries (the classical algorithm tries all n w's with a single 1 entry to identify each bit of a). This is a polynomial quantum speedup.

Before presenting the adiabatic algorithm we point out the following useful observation. For an initial state

$$|\Psi(0)\rangle = \sum_{w \in \{0,1\}^n} c_w |w\rangle_A \otimes |\psi_w(0)\rangle_B, \qquad \sum_{w \in \{0,1\}^n} |c_w|^2 = 1 \tag{39}$$

that undergoes an evolution according to the time-dependent Hamiltonian of the form

$$H(s) = \sum_{w \in \{0,1\}^n} |w\rangle_A \langle w| \otimes H_w(s), \tag{40}$$

we have

$$|\Psi(t)\rangle = \sum_{w \in \{0,1\}^n} c_w |w\rangle_A \otimes |\psi_w(t)\rangle_B, \tag{41}$$

where

$$\begin{aligned}|\psi_w(t)\rangle_B &= |\psi_{t_f, w}(s)\rangle_B \\ &= \mathrm{Texp}\left[-it_f \int_0^s d\sigma H_w(\sigma)\right]|\psi_w(0)\rangle_B, \quad\end{aligned} \tag{42}$$

and Texp denotes the time-ordered exponential. To see this, simply expand the formal solution:

$$\begin{aligned}|\Psi(t)\rangle &= \mathrm{Texp}\left[-it_f \int_0^s d\sigma H(\sigma)\right]|\Psi(0)\rangle \\ &= \sum_{w \in \{0,1\}^n} c_w |w\rangle_A \otimes \mathrm{Texp}\left[-it_f \int_0^s d\sigma H_w(\sigma)\right]|\psi_w(0)\rangle_B.\end{aligned} \tag{43}$$

Thus for each state $|w\rangle$ in subsystem A, there is an independently evolving state in subsystem B. In particular, note that adiabaticity in subsystem B does not depend on the size of system A.

The adiabatic algorithm (Hen, 2014b) encodes the action of $f_a(w)$ in a Hamiltonian acting on two subsystems A and B comprising n qubits and 1 qubit, respectively,

$$H_1 = \sum_{w \in \{0,1\}^n} h_w, \tag{44a}$$

$$h_w \equiv -\tfrac{1}{2}|w\rangle_A \langle w| \otimes [\mathbb{1}_B + (-1)^{f_a(w)}\sigma_B^z]. \tag{44b}$$

The initial Hamiltonian is chosen to be

$$\begin{aligned}H_0 &= \frac{1}{2}[\mathbb{1}_A \otimes (\mathbb{1}_B - \sigma_B^x)] \\ &= \frac{1}{2}\sum_{w \in \{0,1\}^n} |w\rangle_A \langle w| \otimes (\mathbb{1}_B - \sigma_B^x).\end{aligned} \tag{45}$$

Any state of the form

$$|\Psi(0)\rangle = \sum_{w \in \{0,1\}^n} c_w |w\rangle_A \otimes |+\rangle_B \tag{46}$$

is a ground state of H_0, with eigenvalue 0. We assume that the initial state is prepared as the uniform superposition state, i.e., $c_w = 2^{-n/2} \; \forall \, w$.

The total Hamiltonian is thus given by

$$H(s) = (1-s)H_0 + sH_1 = \sum_{w \in \{0,1\}^n} |w\rangle_A \langle w| \otimes H_w(s), \tag{47}$$

where

$$H_w(s) = \frac{1-s}{2}(\mathbb{1}_B - \sigma_B^x) - \frac{s}{2}[\mathbb{1}_B + (-1)^{f_a(w)}\sigma_B^z]. \tag{48}$$

The adiabatic algorithm proceeds, after preparation of the initial state, by adiabatic evolution to the final state

$$\begin{aligned}&|\Psi(t_f)\rangle \\ &= \frac{1}{2^{n/2}}\sum_{w \in \{0,1\}^n} |w\rangle_A \otimes \exp\left[-it_f \int_0^1 \varepsilon_{0,w}(s)ds\right]|f_a(w)\rangle_B,\end{aligned} \tag{49}$$

where $\varepsilon_{0,w}(s)$ is the instantaneous ground state energy of $H_w(s)$, and we used the general argument from Eqs. (39)–(42). Finally an x measurement on subsystem B is performed. Since we can write

$$|f_a(w)\rangle = \frac{1}{\sqrt{2}}[|+\rangle + (-1)^{f_a(w)}|-\rangle], \tag{50}$$

the state collapses to either of the following states with equal probability:

$$|\Psi_+\rangle = \frac{1}{2^{n/2}}\sum_{w \in \{0,1\}^n} |w\rangle_A \otimes |+\rangle_B = |\Psi(0)\rangle, \tag{51a}$$

$$|\Psi_-\rangle = \frac{1}{2^{n/2}}\sum_{w \in \{0,1\}^n} |w\rangle_A \otimes (-1)^{f_a(w)}|-\rangle_B. \tag{51b}$$

Note that since $f_a(w)$ counts the number of 1 agreements between a and w, we can write

$$\sum_{w \in \{0,1\}^n} (-1)^{f_a(w)}|w\rangle_A = \overset{n-1}{\underset{k=0}{\otimes}}[|0\rangle_k + (-1)^{a_k}|1\rangle_k]_A, \tag{52}$$

so that

$$|\Psi_-\rangle = \frac{1}{2^{n/2}} \overset{n-1}{\underset{k=0}{\otimes}} [|0\rangle_k + (-1)^{a_k} |1\rangle_k]_A \otimes |-\rangle_B. \quad (53)$$

If the measurement gives $+1$ [i.e., Eq. (51a)], then the measured state is the initial state and no information is gained and the process must be repeated. If the measurement gives -1, then the resulting state in the A subspace encodes all the bits of a, since if the kth qubit is in the $|+\rangle_k$ state, then $a_k = 0$ and if it is in the $|-\rangle_k$ state, then $a_k = 1$. The probability of failure after m tries is 2^{-m} so it is exponentially small and n independent.

The run time of the algorithm is also n independent, since only a single qubit (system B) is effectively evolving.

In conclusion, the adiabatic Bernstein-Vazirani algorithm finds the unknown binary string a in $O(1)$ time, matching the circuit model depth. Using a similar technique, Hen (2014b) presented a quantum adiabatic version of Simon's exponential-speedup period finding algorithm (Simon, 1997) [a precursor to Shor's factoring algorithm (Shor, 1997)], again matching the circuit model depth scaling. An important aspect of these quantum adiabatic constructions is that they go beyond the general-purpose (and hence suboptimal) polynomial-equivalence prescription of universality proofs that map circuit-based algorithms into quantum adiabatic ones (see Sec. IV). That equivalence does not necessarily preserve a polynomial quantum speedup, whereas the construction in Hen (2014b) discussed here does.

D. The glued-trees problem

Consider two binary trees, each of depth n. Each tree has $\sum_{j=0}^{n} 2^j = 2^{n+1} - 1$ vertices, for a total of $N = 2^{n+2} - 2$ vertices, each labeled by a randomly chosen $2n$-bit string. The two trees are randomly glued as shown in Fig. 1. More specifically, choose a leaf on the left end at random and connect it to a leaf on the right end chosen at random. Then connect the latter to a leaf on the left chosen randomly among the remaining ones, and so on, until every leaf on the left is connected to two leaves on the right and vice versa. This creates a random cycle that alternates between the leaves of the two trees. The problem is, starting from the left root, to find a path to the right root in the smallest possible number of
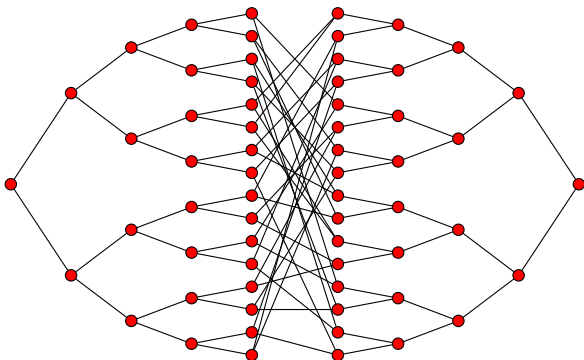

FIG. 1. A glued tree with $n = 4$. The labeling j from Eq. (55) is depicted on top of the tree.

steps, while traversing the tree as in a maze, i.e., keeping a record of one's moves is allowed, but at any given vertex one can only see the adjacent vertices. More formally, an oracle outputs the adjacent vertices of a given input vertex (note that the roots of the trees are the only vertices with adjacency two, so it is easy to check if the right root was found). The problem is, given the name of the left root and access to the oracle, to find the name of the right root in the smallest number of queries. Classical algorithms require at least a subexponential in n number of oracle calls, but there exists a polynomial-time quantum algorithm based on quantum walks for solving this problem (Childs *et al.*, 2003). A polynomial-time quantum almost-adiabatic algorithm was given by Somma, Nagaj, and Kieferová (2012). The qualifier "almost" is important: the algorithm is not adiabatic during the entire evolution, since it explicitly requires a transition from the ground state to the first excited state and back. We now review the algorithm, which so far provides the only example of a (sub)exponential almost-adiabatic quantum speedup.

Let us denote the bit-string corresponding to the first root by a_0 and the second root by a_{N-1}. Define the diagonal (in the computational basis) Hamiltonians

$$H_0 = -|a_0\rangle\langle a_0|, \qquad H_1 = -|a_{N-1}\rangle\langle a_{N-1}|, \quad (54)$$

and the states

$$|c_j\rangle = \frac{1}{\sqrt{N_j}} \sum_{i \in j\text{th column}} |a_i\rangle, \quad (55)$$

which are a uniform superposition over the vertices in the jth column with $N_j = 2^j$ for $0 \le j \le n$ and $N_j = 2^{2n+1-j}$ for $n+1 \le j \le 2n+1$. Note that $|c_0\rangle = |a_0\rangle$ and $|c_{2n+1}\rangle = |a_{N-1}\rangle$. Let us define the Hamiltonian A associated with the oracle as having the following nonzero matrix elements:

$$\langle c_j|A|c_{j+1}\rangle = \begin{cases} \sqrt{2} & j = n, \\ 1 & \text{otherwise}. \end{cases} \quad (56)$$

We then pick as our interpolating Hamiltonian

$$H(s) = (1-s)\alpha H_0 - s(1-s)A + s\alpha H_1, \quad (57)$$

where $\alpha \in (0, 1/2)$ is a constant (independent of n) and $s(t)$ is the schedule. Note that a unitary evolution according to this Hamiltonian will keep a state within the subspace spanned by $\{|c_j\rangle\}$ if the state is initially within that subspace. Since the instantaneous ground state at $s = 0$ ($|a_0\rangle$) is in this subspace, it suffices to consider only this subspace. Because of the form of the Hamiltonian, the eigenvalue spectrum is symmetric about $s = 1/2$.

In this subspace, at $s_\times = \alpha/\sqrt{2}$ (and by symmetry at $1 - s_\times$), the energy gap between the ground state and the first excited state closes exponentially in n. This is depicted in Fig. 2, where s_1, s_2 represent the region around $s_\times$ and s_3, s_4 represent the region around $1 - s_\times$. In the regions $s \in [0, s_1)$,
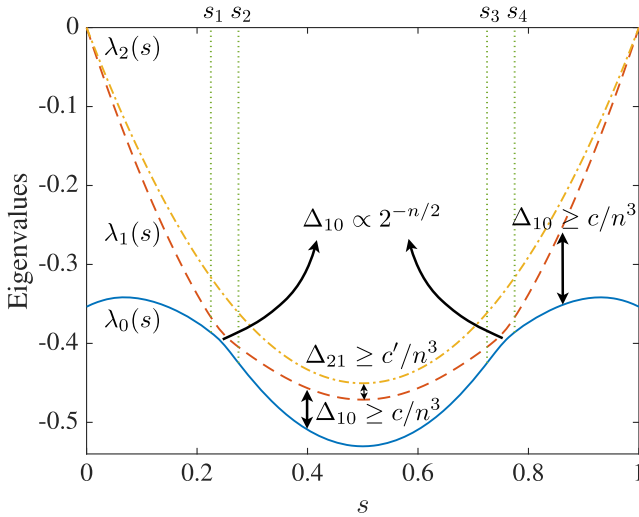
FIG. 2. The ground state [$\lambda_0(s)$, solid blue curve], first excited state [$\lambda_1(s)$, dashed red curve], and second excited state [$\lambda_2(s)$ dot-dashed yellow curve] of the glued-trees Hamiltonian (57) for $\alpha = 1/\sqrt{8}$ and $n = 6$. Inside the region $[s_1, s_2]$ and $[s_3, s_4]$, the gap between the ground state and first excited state Δ_{10} closes exponentially with n. In the region $[s_2, s_3]$, the gap between the ground state and first excited state Δ_{10} and the gap between the first excited state and second excited state Δ_{21} are bounded by n^{-3}. Similarly, in the region $[s_4, 1]$, the gap between the ground state and first excited state Δ_{10} is bounded by n^{-3}.

$s \in [s_2, s_3)$, and $s \in [s_4, 1]$, the energy gap between the ground state and first excited state is lower bounded by c/n^3. The gap between the first and second excited states is lower bounded by c'/n^3 throughout the evolution. Both $c, c' > 0$.

The proposed evolution exploits the symmetry and gap structure of the spectrum as follows. A schedule is chosen that guarantees adiabaticity only if the energy gap scales as $1/n^3$. Then during $s \in [0, s_1)$, the desired evolution is sufficiently adiabatic that it follows the instantaneous ground state. During $s \in [s_1, s_2)$, the evolution is nonadiabatic (since the gap scales as $1/e^n$) and a transition to the first excited state occurs with high probability. During $s \in [s_2, s_3)$, the evolution is again sufficiently adiabatic that it follows the instantaneous first excited state. During $s \in [s_3, s_4)$, the evolution is again nonadiabatic and a transition from the first excited state back to the ground state occurs with high probability. During $s \in [s_4, 1]$, the evolution is again adiabatic and follows the instantaneous ground state.

Since $t_f = \int_0^1 ds (ds/dt)^{-1} \sim n^6$, we conclude that $|a_{N-1}\rangle$ can be found in polynomial time.

E. Adiabatic PAGERANK algorithm

We review the adiabatic quantum algorithm from Garnerone, Zanardi, and Lidar (2012) that prepares a state containing the same ranking information as the PAGERANK vector. The latter is a central tool in data mining and information retrieval at the heart of the Google search engine (Brin and Page, 1998). Using the adiabatic algorithm, the extraction of the full PAGERANK vector cannot, in general, be

done more efficiently than when using the best classical algorithms known. However, there are particular graph topologies and specific tasks of relevance in the use of search engines (such as finding just the top-ranked entries) for which the quantum algorithm, combined with other known quantum protocols, may provide a polynomial or even exponential quantum speedup. Note that unlike the previous algorithms we reviewed in this section, which all provided a provable quantum speedup, the current algorithm provides a "regular" quantum speedup in the sense that it outperforms all currently known classical algorithms, but better future classical algorithms have not been ruled out.

1. Google matrix and PAGERANK

PAGERANK can be seen as the stationary distribution of a random walker on the web graph, which spends its time on each page in proportion to the relative importance of that page (Langville and Meyer, 2006).

To model this define the transition matrix P_1 associated with the (directed) adjacency matrix A of the graph

$$P_1(i, j) = \begin{cases} 1/d(i) & \text{if } (i, j) \text{ is an edge of } A, \\ 0 & \text{else,} \end{cases} \tag{58}$$

where $d(i)$ is the out degree of the ith node.

The rows having zero matrix elements, corresponding to dangling nodes, are replaced by the vector $\vec{e}/n$ whose entries are all $1/n$, where n is the number of pages or nodes, i.e., the size of the web graph. Call the resulting (right) stochastic matrix P_2. However, there could still be subgraphs with in links but no out links. Thus one defines the Google matrix G as

$$G := \alpha P_2^T + (1 - \alpha) E, \tag{59}$$

where $E \equiv |\vec{v}\rangle\langle\vec{e}|$. The personalization vector $\vec{v}$ is a probability distribution; the typical choice is $\vec{v} = \vec{e}/n$. The parameter $\alpha \in (0, 1)$ is the probability that the walker follows the link structure of the web graph at each step, rather than hop randomly between graph nodes according to $\vec{v}$ (Google reportedly uses $\alpha = 0.85$). By construction, G is irreducible and aperiodic, and hence the Perron-Frobenius theorem (Horn and Johnson, 2012)[18] ensures the existence of a unique eigenvector with all positive entries associated to the maximal eigenvalue 1. This eigenvector is precisely the PAGERANK $\vec{p}$. Moreover, the modulus of the second eigenvalue of G is upper bounded by α (Nussbaum, 2003). This is important for the convergence of the power method, the standard computational technique employed to evaluate $\vec{p}$. It uses the fact that, for any probability vector $\vec{p}_0$,

$$\vec{p} = \lim_{k \to \infty} G^k \vec{p}_0. \tag{60}$$

---

[18]This theorem states that if all elements of a real symmetric square matrix A are non-negative, then the largest eigenvalue of A is real; furthermore, the components of the corresponding eigenvector can be chosen to be all non-negative.

The power method computes $\vec{p}$ with accuracy ν in a time that scales as $O[sn\log(1/\nu)/\log(1/\alpha)]$, where s is the sparsity of the graph (maximum number of nonzero entries per row of the adjacency matrix). The rate of convergence is determined by α.

2. Hamiltonian and gap

Consider the following nonlocal final Hamiltonian associated with a generic Google matrix G (in this section we use H and h for local and nonlocal Hamiltonians, respectively):

$$h_1 = h(G) \equiv (\mathbb{1} - G)^\dagger (\mathbb{1} - G). \tag{61}$$

Since $h(G)$ is positive semidefinite, and 1 is the maximal eigenvalue of G associated with $\vec{p}$, it follows that the ground state of $h(G)$ is given by $|\pi\rangle \equiv \vec{p}/\|\vec{p}\|$. The initial Hamiltonian has a similar form, but it is associated with the Google matrix G_c of the complete graph

$$h_0 = h(G_c) \equiv (\mathbb{1} - G_c)^\dagger (\mathbb{1} - G_c). \tag{62}$$

The ground state of h_0 is the uniform superposition state $|\psi(0)\rangle = \sum_{j=1}^n |j\rangle/\sqrt{n}$. The basis vectors $|j\rangle$ span the n-dimensional Hilbert space of $\log_2 n$ qubits. The interpolating adiabatic Hamiltonian is

$$h(s) = (1-s)h_0 + sh_1. \tag{63}$$

Equations (61)–(63) completely characterize the adiabatic quantum PAGERANK algorithm, apart from the schedule $s(t)$.

By numerically simulating the dynamics generated by $h(s)$, Garnerone, Zanardi, and Lidar (2012) showed that for typical random graph instances generated using the "preferential attachment model" (Barabasi and Albert, 1999; Bollobás et al., 2001) and "copying model" (Kleinberg et al., 1999) (both of which yield sparse random graphs with small-world and scale-free features) the typical run time of the adiabatic quantum PAGERANK algorithm scales as

$$t_f \sim (\log\log n)^{b-1}(\log n)^b, \tag{64}$$

where $b > 0$ is some small integer that depends on the details of the graph parameters. The numerically computed gap scales as $(\log n)^{-b}$, which Garnerone, Zanardi, and Lidar (2012) found to be due to the power-law distribution of the out-degree nodes $d(i)$.[19]

3. Speedup

We next discuss two tasks for which this adiabatic quantum ranking algorithm offers a speedup.

The best currently known classical Markov chain Monte Carlo technique used to evaluate the full PAGERANK vector requires a time [in the bulk synchronous parallel computational model (Valiant, 1990)] which scales as

$O[\log(n)]$ (Das Sarma et al., 2015). The algorithm launches $\log n$ random walks from each node of the graph in parallel [for a total of $n\log(n)$ walkers], with each node communicating $O[\log(n)]$ bits of data to each of its connected neighbors after each step. After $O[\log(n)]$ steps, the total number of walkers that have visited a node is used to estimate the PAGERANK of that node. In the absence of synchronization costs [synchronization and communication are known to be important issues for networks with a large number of processors (Awerbuch, 1985; Kumar et al., 2003; Bisseling, 2004; Rauber and Rünger, 2010)], the classical cost can be taken to be $O[n\log(n)^2]$, i.e., the number of parallel processes multiplied by the duration of each process.[20]

At the conclusion of the adiabatic evolution generated by the Hamiltonian in Eq. (63), the PAGERANK vector $\vec{p} = \{p_i\}$ is encoded into the quantum PAGERANK state $|\pi\rangle = \sum_{i=1}^n \sqrt{\pi_i}|i\rangle$ of a $(\log_2 n)$-qubit system, where $|i\rangle$ denotes the ith node in the graph G. The probability of measuring node i is $\pi_i = p_i^2/\|\vec{p}\|^2$. One can estimate π_i by repeatedly sampling the expectation value of the operator σ_i^z in the final state. The number of measurements M needed to estimate π_i is given by the Chernoff-Hoeffding bound (Hoeffding, 1963), allowing one to approximate π_i with an additive error e_i and with $M = \mathrm{poly}(e_i^{-1})$. A nontrivial approximation requires $e_i \leq p_i$ and these are typically $O(1/n)$.

The fact that the amplitudes of the quantum PAGERANK state are $\{\sqrt{\pi_i} = p_i/\|\vec{p}\|\}$, rather than $\{\sqrt{p_i}\}$ is a virtue: the number of samples needed to estimate the rank π_i with additive error $e_i \sim \pi_i$ scales as $O[n^{2\gamma_i-1}]$, so the total quantum cost is $O[n^{2\gamma_i-1}\mathrm{polylog}(n)]$.[21] Thus, for the combined task of state preparation and rank estimation there is a polynomial quantum speedup whenever $\gamma_i < 1$, namely, $O[n^{2\gamma_i-1}\mathrm{polylog}(n)]$ vs $O[n\,\mathrm{polylog}(n)]$; simulations reported by Garnerone, Zanardi, and Lidar (2012) show that this is indeed the case for the top-ranked $\log(n)$ entries, and in applications one is most often interested in the top entries. We emphasize that this holds in the average (not worst) case and is not a provable speedup; the evidence for the scaling is numerical, and it is unknown whether a classical algorithm for the preparation of π rather than $\vec{p}$ may give a similar scaling to the quantum scaling, although if that is the case one could consider quantum preparation of $\{\pi_i^2/\|\pi\|^2\}$, etc.

Another context for useful applications is comparing successive PAGERANKs, or more generally "q sampling" (Aharonov and Ta-Shma, 2003). Suppose one perturbs the

---

[19]The gap becomes too small for a quantum advantage, i.e., scales as $1/\mathrm{poly}(n)$, for graphs with only in-degree power-law distribution or when the out degrees are equal to the in degrees. This was studied in more detail by Frees et al. (2013).

[20]This analysis improves upon the estimates of the classical cost presented by Garnerone, Zanardi, and Lidar (2012) and accounts for the critique presented by Moussa (2013).

[21]It was observed numerically by Garnerone, Zanardi, and Lidar (2012) that $p_i \propto 1/n^{\gamma_i}$, where $\gamma_i \in (0.6, 1]$, and that $\|\vec{p}\|_2^2 \propto 1/n$. Let e_i denote the additive error corresponding to $\pi_i = p_i^2/\|\vec{p}\|_2^2 \sim np_i^2$. It follows from the Chernoff-Hoeffding inequality that the number of samples $M(x)$ from the distribution x, where $x = \pi = \{\pi_i\}$ (output of the quantum algorithm), required for a given, fixed additive estimation error is proportional to the inverse of the additive error $M(\pi) \sim 1/e_i$. Assuming $e_i \sim \pi_i$, it follows that $M(\pi) \sim 1/\pi_i \sim 1/(np_i^2) \sim n^{2\gamma_i-1}$. The total cost required to prepare the sample in the quantum case is $\mathcal{O}[\mathrm{polylog}(n)]$.

web graph. The adiabatic quantum algorithm can provide, in time $O[\text{polylog}(n)]$, the pre- and postperturbation states $|\pi\rangle$ and $|\tilde{\pi}\rangle$ as input to a quantum circuit implementing the SWAP test (Buhrman *et al.*, 2001). To obtain an estimate of the fidelity $|\langle\pi|\tilde{\pi}\rangle|^2$ one needs to measure an ancilla $O(1)$ times, the number depending only on the desired precision. In contrast, deciding whether two probability distributions are close classically requires $O[n^{2/3}\log n]$ samples from each (Batu *et al.*, 2000). Whenever some relevant perturbation of the previous quantum PAGERANK state is observed, one can decide to run the classical algorithm again to update the classical PAGERANK.

IV. UNIVERSALITY OF AQC

What is the relation between the computational power of the circuit model and the adiabatic model of quantum computing? It turns out that they are equivalent up to polynomial overhead. It is well known that the circuit model is universal for quantum computing, i.e., that there exist sets of gates acting on a constant number of qubits each that can efficiently simulate a quantum Turing machine (Deutsch, 1985; Yao, 1993). A set of gates is said to be universal for QC if any unitary operation may be approximated to arbitrary accuracy by a quantum circuit involving only those gates (Nielsen and Chuang, 2000). The analog of such a set of gates in AQC is a Hamiltonian. An operational definition of universal AQC is thus to efficiently map any circuit to an adiabatic computation using a sufficiently powerful Hamiltonian. Formally:

Definition 3 (universal adiabatic quantum computation). A time-dependent Hamiltonian $H(t)$, $t \in [0, t_f]$ is universal for AQC if, given an arbitrary quantum circuit U operating on an arbitrary initial state $|\psi\rangle$ of n p-state particles and having depth L, the ground state of $H(t_f)$ is equal to $U|\psi\rangle$ with probability greater than $\epsilon > 0$, the number of particles $H(t)$ operates on is poly(n) $\forall t$, and $t_f = \text{poly}(n, L)$.

The stipulation that the ground state of $H(t_f)$ is equal to the final state at the end of the circuit ensures that the circuit and the adiabatic computation have the same output. We note that it is possible and useful to relax the ground state requirement and replace it with another eigenstate of $H(t)$ (see, e.g., Sec. VI.C). The requirement that the number of particles and time taken by the adiabatic computation are polynomial in n and L ensures that the resources used do not blow up.

We begin, in Sec. IV.A, by showing that the circuit model can efficiently simulate AQC. The real challenge is to show the other direction, i.e., that AQC can efficiently simulate the circuit model, which is what we devote the rest of this section to. Along the way, this establishes the universality of AQC. We present several proofs starting in Sec. IV.B with a detailed review of the history state construction of Aharonov *et al.* (2007), who showed in addition that six-state particles in two dimensions suffice for universal adiabatic quantum computation. This was improved upon by Kempe, Kitaev, and Regev (2006), using perturbation-theory gadgets, who showed that qubits can be used instead of six-state particles, and that adiabatic evolution with 2-local Hamiltonians is quantum universal. A 2-local model of universal AQC in 2D, which we

reviewed in Sec. IV.C, was proposed by Mizel, Lidar, and Mitchell (2007) using fermions. Universal AQC using qubits on a two-dimensional grid was accomplished by Oliveira and Terhal (2008). Further simplifications of universal AQC in 2D were presented by Breuckmann and Terhal (2014), Gosset, Terhal, and Vershynina (2015), and Lloyd and Terhal (2016), using the space-time circuit model, which we review in Sec. IV.D. The ultimate reduction in spatial dimensionality was accomplished by Aharonov *et al.* (2009), who showed that universal AQC is possible with 1D nine-state particles as reviewed in Sec. IV.E. Finally, in Sec. IV.F we review a construction that allows one to quadratically amplify the gap of any Hamiltonian used in AQC (satisfying a frustration-freeness property), although this requires the computation to take place in an excited state.

A. The circuit model can efficiently simulate AQC

That the circuit model can efficiently simulate the adiabatic model is relatively straightforward and was first shown by Farhi *et al.* (2000). Assume for simplicity a linear schedule, i.e., an AQC Hamiltonian of the form

$$H(t) = \left(1 - \frac{t}{t_f}\right)H_0 + \frac{t}{t_f}H_1.$$

The evolution of a quantum system generated by the time-dependent Hamiltonian $H(t)$ is governed by the unitary operator

$$U(t_f, 0) = \text{Texp}\left[-i\int_0^{t_f} dt H(t)\right]. \quad (65)$$

If t_f satisfies the condition for adiabaticity, $U(t_f, 0)$ will map the ground state at $t = 0$ to the ground state at t_f. Therefore it suffices to show that the circuit model can simulate $U(t_f, 0)$. To do so, we approximate the evolution by a product of unitaries involving time-independent Hamiltonians $H'_m \equiv H(m\Delta t)$:

$$U(t_f, 0) \mapsto U'(t_f, 0) = \prod_{m=1}^{M} U'_m = \prod_{m=1}^{M} e^{-i\Delta t H'_m}, \quad (66)$$

where $\Delta t = t_f/M$. The error incurred by this approximation is (van Dam, Mosca, and Vazirani, 2001)

$$\|U(t_f, 0) - U'(t_f, 0)\| \in O\left[\sqrt{t_f \text{poly}(n)/M}\right]. \quad (67)$$

We now approximate each individual term in the product in Eq. (66) using the Baker-Campbell-Hausdorff formula (Klarsfeld and Oteo, 1989) by

$$U'_m \mapsto U''_m = e^{-i\Delta t(1 - m\Delta t/t_f)H_0}e^{-i\Delta t(m\Delta t/t_f)H_1}, \quad (68)$$

which incurs an error $\|e^{A+B} - e^A e^B\| \in O(\|AB\|)$ due to the neglected leading order commutator term $[A, B]/2$, i.e.,

$$\|U'_m - U''_m\| \in O\left(\frac{t_f^2}{M^2}\|H_0 H_1\|\right). \quad (69)$$

Therefore, accounting for the M terms in the product and observing that the error in Eq. (67) is subdominant, the total error is (van Dam, Mosca, and Vazirani, 2001)

$$\left\| U(t_f, 0) - \prod_{m=1}^{M} U''_m \right\| \in O[\text{poly}(n)t_f^2/M]. \qquad (70)$$

This means we can approximate $U(t_f, 0)$ with a product of $2M$ unitaries provided that M scales as $t_f^2 \text{poly}(n)$.

Depending on the form of H_0 and H_1, they may need further decomposition in order to write the terms in Eq. (68) in terms of few-qubit unitaries. For example, for the standard initial Hamiltonian $H_0 = -\sum_i \sigma_i^x$, which is a sum of commuting single-qubit operators, we can write $e^{-i\Delta t(1-t/t_f)H_0/K}$ as a product of n one-qubit unitaries. Likewise, assuming that H_1 is 2-local, we can write $e^{-i(m\Delta t/t_f)\Delta t H_1/M}$ as a product of up to n^2 two-qubit unitaries within the same order of approximation as Eq. (68). Thus, $U(t_f, 0)$ can be approximated as a product of unitary operators each of which acts on a few qubits. The scaling of t_f required for adiabatic evolution is inherited by the number of few-qubit unitary operators in the associated circuit version of the algorithm.

A more efficient method was proposed by Boixo, Knill, and Somma (2009), building upon the ideas explained in Sec. II.B.4. This "eigenpath traversal by phase randomization" method applies the Hamiltonian $H(t_j)$ in piecewise continuous manner at random times t_j. Each interval $[t_j, t_{j+1}]$ corresponds to a unitary $e^{-iH(t_j)}$, which then needs to be decomposed into one- and two-qubit gates. The randomization introduces an effective eigenstate decoupling in the Hamiltonian eigenbasis (similarly to the effect achieved by projections in the Zeno effect), so that if the initial state is the ground state, the evolution will follow the ground state throughout as required for AQC. The algorithmic cost of this randomization method is defined as the average number of times the unitaries are applied, and it can be shown that the cost is $O(L^2/\varepsilon\Delta)$, where ε is the desired maximum error of the final state compared to the target eigenstate, and L is the path length [Eq. (13)]. Since $L \leq \max_s \|\dot{H}(s)\|/\Delta$ as seen in Sec. II.B.4, the worst-case bound on the cost is $\max_s \|\dot{H}(s)\|^2/(\varepsilon\Delta^3)$, up to logarithmic factors.

B. AQC can efficiently simulate the circuit model: History state proof

The goal is, given an arbitrary n-qubit quantum circuit, to design an adiabatic computation whose final ground state is the output of the quantum circuit described by a sequence of L one- or two-qubit unitary gates $U_1, U_2, \ldots, U_L$. This adiabatic simulation of the circuit should be efficient, i.e., it may incur at most polynomial overhead in the circuit depth L. In this section we review the proof presented by Aharonov et al. (2007). This was the first complete proof of the universality of AQC, and many of the ideas and techniques introduced therein inspired subsequent proofs, remaining relevant today.

Let us assume that the n-qubit input to the circuit is the $|0\cdots0\rangle$ state. After the ℓth gate, the state of the quantum

circuit is given by $|\alpha(\ell)\rangle$. To proceed, we use the "circuit-to-Hamiltonian" construction (Kitaev, Shen, and Vyalyi, 2000), where the final Hamiltonian will have as its ground state the entire history of the quantum computation. This "history state" is given by

$$|\eta\rangle = \frac{1}{\sqrt{L+1}} \sum_{\ell=0}^{L} |\gamma(\ell)\rangle, \qquad (71a)$$

$$|\gamma(\ell)\rangle \equiv |\alpha(\ell)\rangle \otimes |1^\ell 0^{L-\ell}\rangle_c, \qquad (71b)$$

where $|1^\ell 0^{L-\ell}\rangle_c$ denotes the Feynman clock (Feynman, 1985) register composed of $L+1$ qubits. The notation means that we have ℓ ones followed by $L - \ell$ zeros to denote the time after the ℓth gate. We want to construct a Hamiltonian H_{init} with ground state $|\gamma(0)\rangle$ and a Hamiltonian H_{final} with ground state $|\eta\rangle$. Let

$$H_{\text{init}} = H_{\text{c-init}} + H_{\text{input}} + H_c, \qquad (72a)$$

$$H_{\text{final}} = \tfrac{1}{2}H_{\text{circuit}} + H_{\text{input}} + H_c, \qquad (72b)$$

$$H_{\text{circuit}} = \sum_{\ell=1}^{L} H_\ell. \qquad (72c)$$

The full time-independent Hamiltonian $H(s)$ is given by (Aharonov et al., 2007)

$$\begin{aligned} H(s) &= (1-s)H_{\text{init}} + sH_{\text{final}} \\ &= H_{\text{input}} + H_c + (1-s)H_{\text{c-init}} + \frac{s}{2}H_{\text{circuit}}. \quad (73) \end{aligned}$$

The various terms are chosen so that the ground state always has energy 0:

- H_c: This term should ensure that the clock's state is always of the form $|1^\ell 0^{L-\ell}\rangle_c$. Therefore, we energetically penalize any clock-basis state that has the sequence 01:

$$H_c = \sum_{\ell=1}^{L-1} |0_\ell 1_{\ell+1}\rangle_c \langle 0_\ell 1_{\ell+1}|, \qquad (74)$$

where $|0_\ell 1_{\ell+1}\rangle_c$ denotes a 0 on the ℓth clock qubit and 1 on the $(\ell+1)$th clock qubit. Any illegal clock state will have an energy ≥ 1. Any legal clock state will have energy 0.

- $H_{\text{c-init}}$: Ensures that the initial clock state is $|0^L\rangle_c$,

$$H_{\text{c-init}} = |1_1\rangle_c \langle 1_1|. \qquad (75)$$

Note that we need to specify only the first clock qubit to be in the zero state. For a legal clock state, Eqs. (74) and (75) imply that the rest are in the zero state as well.

- H_{input}: Ensures that if the clock state is $|0^L\rangle$, then the computation qubits are in the $|0^n\rangle$ state,

$$H_{\text{input}} = \sum_{i=1}^{n} |1_i\rangle\langle 1_i| \otimes |0_1\rangle_c\langle 0_1|. \qquad (76)$$

- H_ℓ: Ensures that the propagation from $\ell - 1$ to ℓ corresponds to the application of U_ℓ,

$$
\begin{aligned}
H_1 = &\mathbb{1} \otimes |0_1 0_2\rangle_c\langle 0_1 0_2| - U_1|1_1 0_2\rangle_c\langle 0_1 0_2| \\
&- U_1^\dagger|0_1 0_2\rangle_c\langle 1_1 0_2| + \mathbb{1} \otimes |1_1 0_2\rangle_c\langle 1_1 0_2|, \qquad (77a)
\end{aligned}
$$

$$
\begin{aligned}
H_{2 \le \ell \le L-1} = &\mathbb{1} \otimes |1_{\ell-1} 0_\ell 0_{\ell+1}\rangle_c\langle 1_{\ell-1} 0_\ell 0_{\ell+1}| \\
&- U_\ell|1_{\ell-1} 1_\ell 0_{\ell+1}\rangle_c\langle 1_{\ell-1} 0_\ell 0_{\ell+1}| \\
&- U_\ell^\dagger|1_{\ell-1} 0_\ell 0_{\ell+1}\rangle_c\langle 1_{\ell-1} 1_\ell 0_{\ell+1}| \\
&+ \mathbb{1} \otimes |1_{\ell-1} 1_\ell 0_{\ell+1}\rangle_c\langle 1_{\ell-1} 1_\ell 0_{\ell+1}|, \qquad (77b)
\end{aligned}
$$

$$
\begin{aligned}
H_L = &\mathbb{1} \otimes |1_{L-1} 0_L\rangle_c\langle 1_{L-1} 0_L| \\
&- U_L|1_{L-1} 1_L\rangle_c\langle 1_{L-1} 0_L| \\
&- U_1^\dagger|1_{L-1} 0_L\rangle\langle 1_{L-1} 1_L| \\
&+ \mathbb{1} \otimes |1_{L-1} 1_L\rangle_c\langle 1_{L-1} 1_L|. \qquad (77c)
\end{aligned}
$$

Note that the first and last terms leave the state unchanged. The second term propagates the computational state and clock register forward, while the third term propagates the computational state and clock register backward.

It turns out that the state $|\gamma(0)\rangle = |\alpha(0)\rangle \otimes |0^L\rangle$ is the ground state of H_{init} with eigenvalue 0, and $|\eta\rangle$ is the ground state of H_{final} with eigenvalue 0. Let $\mathcal{S}_0$ be the subspace spanned by $\{|\gamma(\ell)\rangle\}_{\ell=0}^{L}$. The state $|\alpha(0)\rangle$ is the input to the circuit, so it can be taken to be the $|0\cdots 0\rangle$ state, i.e., the initial ground state is an easily prepared state. Since the initial state $|\gamma(0)\rangle \in \mathcal{S}_0$, the dynamics generated by $H(s)$ keep the state in $\mathcal{S}_0$. It turns out that the ground state is unique for $s \in [0, 1]$. By mapping the Hamiltonian within $\mathcal{S}_0$ to a stochastic matrix, it is possible to find a polynomial lower bound on the gap from the ground state within $\mathcal{S}_0$:

$$\Delta(H_{\mathcal{S}_0}) \ge \frac{1}{4}\left(\frac{1}{6L}\right)^2. \qquad (78)$$

It is also possible to bound the global gap (i.e., not restricted to the $\mathcal{S}_0$ subspace) as

$$\Delta(H) \ge \Omega(1/L^3). \qquad (79)$$

A measurement of the final state will find the final outcome of the quantum circuit $|\gamma(L)\rangle$ with probability $1/(L+1)$. This can be amplified by inserting identity operators at the end of the circuit, hence causing the history state to include a greater superposition of the final outcome of the circuit. Together these results show that there is an efficient implementation of any given quantum circuit using the adiabatic algorithm with

$H(s)$. Here and elsewhere "efficient" means up to polynomial overhead, i.e., where t_f scales as a polynomial in L.

The proof techniques used to obtain these results are instructive and of independent interest, so we review additional technical details in Appendix C.

To conclude this section, we mention additional results supporting the equivalence of the circuit and adiabatic approach in terms of state preparation. Aharonov and Ta-Shma (2003) showed (Theorem 2) that any quantum state that can be efficiently generated in the circuit model can also be efficiently generated by an adiabatic approach, and vice versa, for the same initial state. The proof relies on two important lemmas, the "sparse Hamiltonian lemma" and the "jagged adiabatic path" lemma. The former gives conditions under which a Hamiltonian is efficiently simulatable in the circuit model, and the latter provides conditions under which a sequence of Hamiltonians, defining a path, can have a non-negligible spectral gap.

C. Fermionic ground state quantum computation

A model of ground state quantum computation (GSQC) using fermions was independently proposed by Mizel, Mitchell, and Cohen (2001) [see also Mizel, Mitchell, and Cohen (2002), Mizel (2004), and Mao (2005a, 2005b)] around the same time as AQC. In GSQC, one executes a quantum circuit by producing a ground state that spatially encodes the entire temporal trajectory of the circuit, from input to output. Mizel, Lidar, and Mitchell (2007) showed how to adiabatically reach the desired ground state, thus providing an alternative to history state-type constructions for universal AQC. One of the differences between the GSQC and history state constructions is that instead of relying on the Feynman "global clock particle" idea, particles are synchronized *locally* (via CNOT gates), an idea that traces back to Margolus (1990) and was later adopted in some of the space-time circuit-to-Hamiltonian constructions (Breuckmann and Terhal, 2014).

Consider a quantum circuit with n qubits and depth L. We associate $2(L+1)$ fermionic modes with every qubit q, via creation operators $a_{q,\ell}^\dagger$ and $b_{q,\ell}^\dagger$, where $\ell = 0, \ldots, L$. One can view these $2n(L+1)$ modes as the state space of n spin-1/2 fermions, where each fermion can be localized at sites on a 1D (time) line of length $L+1$. To illustrate this with a concrete physical system, imagine a two-dimensional array of quantum dots with $L+1$ columns and two rows per qubit, corresponding to the $|0\rangle$ and $|1\rangle$ basis states of that qubit. A total of n electrons is placed in the array. The state of each qubit determines the spin state of the corresponding electron, which in term determines which of the two rows it is in, while the clock of each qubit is represented by which column the electron is in.

It is convenient to group creation operators into row vectors $C_{q,\ell}^\dagger = (a_{q,\ell}^\dagger b_{q,\ell}^\dagger)$. Then for each single-qubit gate $U_{q,\ell}^{(1)}$ we introduce a term

$$H_{q,\ell}^{(1)}(s) = [C_{q,\ell}^\dagger - sC_{q,\ell-1}^\dagger (U_{q,\ell}^{(1)})^\dagger](C_{q,\ell} - sC_{q,\ell-1} U_{q,\ell}^{(1)}) \qquad (80)$$

into the circuit Hamiltonian H_{circuit}. The off-diagonal terms represent hopping or tunneling of the qth electron from site

$\ell - 1$ to ℓ and vice versa, while $U_{q,\ell}^{(1)}$ acts on the electron's spin. The diagonal terms $C_{q,\ell}^{\dagger} C_{q,\ell}$ and $C_{q,\ell-1}^{\dagger} C_{q,\ell-1}$ ensure that $H_{q,\ell}^{(1)} \geq 0$. The parameter $s \in [0,1]$ controls the interpolation from the initial, simple to prepare ground state at $s = 0$ when there is no tunneling and every electron is frozen in place, to the full realization of all the gates $U_{q,\ell}^{(1)}$ when $s = 1$.

One can similarly define CNOT Hamiltonian terms between electrons or fermions, whose form can be found in Mizel, Lidar, and Mitchell (2007) [see also Breuckmann and Terhal (2014)]. These 2-local terms can be understood as a sum of an identity and NOT term. For such two-qubit gates, the fermions corresponding to the control and target qubits tunnel forward or backward and the internal spin state of the target fermion changes depending on the internal state of the control fermion. An important additional ingredient is the addition of a penalty term that imposes an energy penalty on states in which one qubit has gone through the CNOT gate without the other. Instead of the Feynman clock used in the history state construction, there are many local clocks, one per qubit. The synchronization mechanism takes place via the CNOT Hamiltonian. Moreover, the entire construction naturally involves only 2-local interactions between fermions in 2D.

While the fermionic GSQC model proposed by Mizel, Lidar, and Mitchell (2007) was shown there to be universal for AQC, its gap analysis was incomplete.[22] This was fixed by Childs, Gosset, and Webb (2013), which proved the "null space projection lemma" that was implicitly assumed by Mizel, Lidar, and Mitchell (2007). This lemma is interesting in its own right, so we reproduce it here.

Lemma 2: Null space projection lemma. Let $\Delta(A)$ denote the smallest nonzero eigenvalue of the positive semi-definite operator A. Let H_0 and H_1 be positive semidefinite and assume the nullspace S of H_0 is nonempty. Also assume that $\Delta(H_1|_S) \geq c > 0$ and $\Delta(H_0) \geq d > 0$. Then

$$\Delta(H_0 + H_1) \geq \frac{cd}{c + d + \|H_1\|}. \tag{81}$$

As shown by Breuckmann and Terhal (2014), the fermionic GSQC model can be unitarily mapped onto the space-time circuit-to-Hamiltonian model for qubits in 2D, where the gap analysis is more convenient. Using the same mapping, Breuckmann and Terhal (2014) also showed that the fermionic model of Mizel, Lidar, and Mitchell (2007) is in fact QMA complete. We thus proceed to discuss the space-time model next.

---

[22]Mizel, Lidar, and Mitchell (2007), p. 4 claimed that "$\langle Z|H|Z \rangle \geq \mathcal{E}O(1/N^2)$" ($N$ is L in our notation), implying a lower bound on the spectral gap of the total Hamiltonian H. This claim was based on Mizel, Mitchell, and Cohen (2002), but was in fact not proven there. Here $\mathcal{E}$ is the energy scale of the CNOT terms, the total Hamiltonian is $H = H_0 + H_1$, where H_0 contains all of the single-qubit terms, H_1 includes all of the CNOT terms, and $|Z\rangle$ denotes the known ground state of H_0. As pointed out by Breuckmann and Terhal (2014), the missing step is essentially to exclude zero-energy, invalid time configurations.

(a)   (b)   (c)
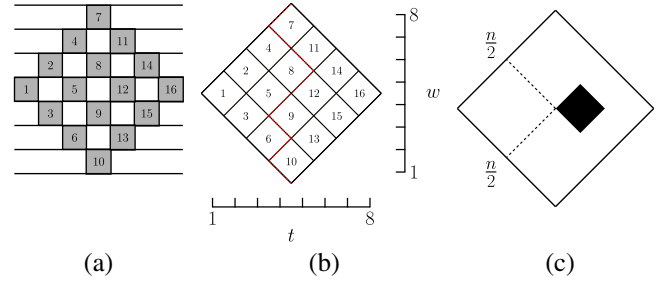
FIG. 3. (a) A $2n = 8$ qubit quantum circuit, where each gray square ($n^2 = 16$ in total) corresponds to a two-qubit gate. (b) An equivalent representation of the quantum circuit in (a) in terms of a rotated grid. The dashed red line corresponds to an allowed string configuration for the particles. (c) The circuit is constrained such that the majority of the gates are identity except in a $k \times k$ subgrid (shown in black), located such that its left vertex is at the center of the rotated grid. A successful computation requires the t positions of the $2k$ particles with w positions that cross the interaction region to lie to the right of the interaction region. See also Fig. 1 in Gosset, Terhal, and Vershynina (2015).

D. Space-time circuit-to-Hamiltonian construction

Here we review another construction that realizes universal adiabatic quantum computation (Gosset, Terhal, and Vershynina, 2015; Lloyd and Terhal, 2016). This builds on the so-called space-time circuit-to-Hamiltonian construction (Breuckmann and Terhal, 2014), which in turn is based on the Hamiltonian computation construction of Janzing (2007). We consider the $2n$-qubit quantum circuit with n^2 two-qubit gates, arranged as shown in Fig. 3(a). This form is sufficient for universal quantum computation (Janzing, 2007). An equivalent representation of the circuit is given in Fig. 3(b), where the n^2 gates are arranged in a rotated $n \times n$ grid. Each plaquette p is associated with a gate U_p, of which the majority are identity gates. Only a $k \times k$ subgrid of the $n \times n$ grid with $k = \sqrt{n}/16$ has nonidentity gates, with the subgrid located as shown in Fig. 3(c). This region is referred to as the interaction region.

The circuit is mapped to a Hamiltonian $H(\lambda)$, with $\lambda \in [0,1]$. The Hamiltonian describes the evolution of particles that live on the edges of the rotated $n \times n$ grid. The positions of the particles are given in terms of the coordinates $(t,w) \in \{1,...,2n\}^2$ as shown in Fig. 3(b). Each particle has 2 internal degrees of freedom in order to encode the qubits of the circuit. Let $a_{t,s}[w]$ denote the annihilation operator which annihilates a particle with internal state $s \in \{0,1\}$ on the edge (t,w). The number operator is defined as $n_{t,s}[w] = a_{t,s}^{\dagger}[w]a_{t,s}[w]$, which counts the number of particles (which will be either 0 or 1) at position (t,w) with state s. Let $\mathbf{n}_t[w] = n_{t,0}[w] + n_{t,1}[w]$.

We focus on configurations of particles that form connected segments starting at the top and ending at the bottom [an example is shown in Fig. 3(b)], referred to as consistent connected string configurations. For a fixed w (i.e., a horizontal line on the rotated grid), there is only one occupied edge. We can describe such configurations in terms of $2n$ bits, denoted by z. Specifically, let the bit value 0 correspond to an edge going down and left and 1 correspond to an edge going

down and right. The Hamming weight of such configurations must be n, since they start and end in the middle of the grid and so must go left and right an equal number of times.

We are now ready to describe the Hamiltonian

$$H(\lambda) = H_{\text{string}} + H_{\text{circuit}}(\lambda) + H_{\text{input}}. \quad (82)$$

- H_{input}: This term ensures that the ground state has the internal state of all particles set to $s = 0$ when the string lies on the left-hand side of the grid by energetically penalizing all states (on the left-hand side) with $s = 1$. It is given by

$$H_{\text{input}} = \sum_{w=1}^{2n} \sum_{t \leq n} n_{t,1}[w]. \quad (83)$$

- H_{string}: This term ensures that the ground state is in the subspace of connected strings. Consider a single vertex v in the grid with incident edges labeled by (t, w), $(t + 1, w)$, $(t, w + 1)$, and $(t + 1, w + 1)$. We can associate a Hamiltonian H_{string}^v to each vertex,

$$H_{\text{string}}^v = \mathbf{n}_t[w] + \mathbf{n}_{t+1}[w] + \mathbf{n}_t[w + 1] + \mathbf{n}_{t+1}[w + 1]$$
$$- 2(\mathbf{n}_t[w] + \mathbf{n}_{t+1}[w])(\mathbf{n}_t[w + 1] + \mathbf{n}_{t+1}[w + 1]) \quad (84)$$

(for vertices at the boundary of the grid with two or three incident edges, the definition of H_{string}^v needs to be modified accordingly) such that $H_{\text{string}} = \sum_v H_{\text{string}}^v$. For connected string configurations, the energy due to this Hamiltonian is zero, while disconnected strings with L string segments have a higher energy $2L - 2$.

- $H_{\text{circuit}}(\lambda) = \sum_p H_{\text{gate}}^p(\lambda) + \sqrt{1 - \lambda^2} H_{\text{init}}$: Define for each plaquette p with borders given by the edges $\{(t, w), (t + 1, w), (t, w + 1), \text{ and } (t + 1, w + 1)\}$

$$H_{\text{gate}}^p(\lambda) = \mathbf{n}_t[w]\mathbf{n}_t[w + 1] + \mathbf{n}_{t+1}[w]\mathbf{n}_{t+1}[w + 1]$$
$$+ \lambda H_{\text{prop}}^p, \quad (85)$$

where

$$H_{\text{prop}}^p = -\sum_{\alpha,\beta,\gamma,\delta} (\langle \beta, \delta | U_p | \alpha, \gamma \rangle a_{t+1,\beta}^\dagger a_{t,\alpha}[w]$$
$$\times a_{t+1,\delta}^\dagger[w + 1] a_{t,\gamma}[w + 1]) + \text{H.c.} \quad (86)$$

The term H_{prop}^p allows for a pair of particles located on the left (right) edges of a plaquette to hop together such that they both are located on the right (left) edges of a plaquette, with their internal states changed according to U_p ($U_p^\dagger$). Note that this move preserves the connectedness of the string. Furthermore, the term $\sum_p H_{\text{gate}}^p(0)$ is minimized by a configuration lying either entirely on the left border (corresponding to the bit string $z = 0^n 1^n$) or entirely on the right border ($z = 1^n 0^n$), which in conjunction with H_{init}, given by

$$H_{\text{init}} = \mathbf{n}_{n+1}[w = 1] + \mathbf{n}_{n+1}[w = 2n], \quad (87)$$

ensures that the ground state of $H(0)$ is such that all particles lie along the left boundary of the grid. Including the effects of $H_{\text{circuit}}(0)$ and H_{input}, the ground state of $H(0)$ is given by $|0^{2n}\rangle|0^n 1^n\rangle$ with eigenvalue 1. This is an easily prepared ground state.

It can be shown that the ground state of $H(\lambda)$ [Eq. (82)] along $\lambda \in [0, 1]$ is unique and the energy gap above the ground state is lower bounded by $1/\text{poly}(n)$ for all $\lambda \in [0, 1]$ [Theorem 1 in Gosset, Terhal, and Vershynina (2015)]. To measure the output of the quantum circuit, we measure the t positions of the $2k$ particles for the w values that cross the interaction region (recall that there will always be one particle per horizontal w line). If we find that all $2k$ particles lie to the right of the interaction region, then their internal states must encode the output of the quantum circuit. For the choice $k = \sqrt{n}/16$, this occurs with a probability lower bounded by a positive constant. Together these properties allow for an efficient (up to polynomial overhead) simulation of the quantum circuit using the adiabatic algorithm generated by $H(\lambda)$.

However, this implementation requires four-body interactions [see, for example, the product term in Eq. (84)]. Lloyd and Terhal (2016) presented improvements to this construction with only 2-local interactions using a first-order perturbation gadget and a quadratic increase in the number of qubits from the original quantum circuit. The use of only first-order perturbation theory is particularly significant, since effective interactions obtained in the kth order degenerate perturbation theory with perturbative coupling g and gap Δ of the unperturbed Hamiltonian scale in strength as $g(g/\Delta)^{k-1}$, leading to a correspondingly small gap of the effective Hamiltonian. In addition, multiple uses of higher-order perturbation theory can increase qubit overhead and complexity.

E. Universal AQC in 1D with nine-state particles

The constructions of universal AQC reviewed so far are all spatially two dimensional (2D). It was unclear for some time whether universal AQC is possible in 1D, with some suggestive evidence to the contrary, such as the success of density matrix renormalization group techniques in calculating ground state energies and other properties of a variety of 1D quantum systems (Schollwöck, 2005). Moreover, classical 1D systems are generally "easy"; e.g., a 1D restriction of MAX-2-SAT with p-state variables can be solved by dynamic programming and hence is in the complexity class P. In addition, the area law implies that 1D systems with a constant spectral gap can be efficiently simulated classically (Hastings, 2009). All this implies that adiabatic evolution with 1D Hamiltonians is not useful for universal QC unless certain conditions are met, in particular, a spectral gap that tends to zero.

This was accomplished by Aharonov *et al.* (2009), who proved that it is possible to perform universal AQC using a 1D quantum system of nine-state particles. The striking

qualitative difference between the quantum and the classical 1D versions of the same problem seems surprising. However, the k-local Hamiltonian problem allows for the encoding of an extra dimension (time), by making the ground state a superposition of states corresponding to different times. This means that the correct analog of the quantum 1D local Hamiltonian problem is 2D classical MAX-k-SAT, which is NP complete.

The proof presented by Aharonov *et al.* (2009) builds heavily on the history state construction reviewed in Sec. IV.B. However, there are a couple of important differences. As in the history state construction, the starting point is a quantum circuit U_x acting on n qubits (where x is the classical input to the function implemented by the circuit in the universal AQC case). A 1D p-state Hamiltonian is designed which will verify correct propagation according to this circuit. Then this is used as the final Hamiltonian for the adiabatic evolution. The problem with directly realizing this in the 1D case is that only the particles nearest to the clock would be able to take advantage of it in order to check correct propagation in time. To overcome this, the circuit U_x is first modified into a new circuit $\tilde{U}_x$ with a distributed clock. The history state construction relies on the ability to copy qubits from one column to the next in order to move to the next block of gates in the computation, so a new strategy is needed in 1D. For the modified circuit $\tilde{U}_x$, the qubits are instead placed in a block of n adjacent particles. One set of gates is performed, and then all of the qubits are moved over n places to advance time in the original circuit U_x. More states per particle are needed to accomplish this than in the 2D case. The second main new idea that is needed is related to ensuring that the state of the system had a valid structure. In the 2D case local constraints were used to check that there are no two-qubit states in adjacent columns. However, using only local constraints, there is no way to check that there are exactly n-qubit data states in an unknown location in a 1D system, since there are only a constant number of local rules available, which are therefore unable to count to an arbitrarily large n. Instead, it is ensured that, under the transition rules of the system, any invalid configurations will evolve in polynomial time into a configuration which can be detected as illegal by local rules. Thus, for every state which is not a valid history state, either the propagation is wrong, which implies an energy penalty due to the propagation Hamiltonian, or the state evolves to an illegal configuration which is locally detectable, which implies an energy penalty due to the local check of illegal configurations. For additional technical details required to complete the proof see Aharonov *et al.* (2009). A 20-state translation-invariant modification of the construction from Aharonov *et al.* (2009) for universal AQC in 1D was given by Nagaj and Wocjan (2008), improving on a 56-state construction by Janzing, Wocjan, and Zhang (2008).

F. Adiabatic gap amplification

In all universality constructions the run time of the adiabatic simulation of a quantum circuit depends on the inverse minimum gap of the simulating Hamiltonian. Therefore it is of interest to develop a general technique for amplifying this gap as was done by Somma and Boixo (2013).

Consider a Hamiltonian H with ground state $|\phi\rangle$. The goal is to construct a new Hamiltonian H' that has $|\phi\rangle$ as an eigenstate (not necessarily the ground state) but with a larger spectral gap. A quadratic spectral gap amplification is possible when H is frustration free (Bravyi and Terhal, 2009). *Definition 4 (frustration freeness)*. A Hamiltonian $H \in \mathbb{C}^N \times \mathbb{C}^N$ is frustration free if it can be written as a sum over positive semidefinite operators $H = \sum_{k=1}^{L} a_k \Pi_k$, with $a_k \in [0, 1]$ and $L = \text{polylog}(N)$. Further, if $|\phi\rangle$ is the ground state of H then it is a ground state (i.e., zero eigenvector) of every term in the decomposition of H, i.e., $\Pi_k |\phi\rangle = 0 \, \forall \, k$.

Somma and Boixo (2013) took Π_k as projectors. The quadratic amplification is optimal for frustration-free Hamiltonians in a suitable black-box model, and no spectral gap amplification is possible, in general, if the frustration-free property is removed. An important caveat is that the construction replaces ground state evolution by evolution of a state that lies in the middle of the spectrum; thus it does not fit the strict definition of AQC (Definition 1). We will have another occasion to relax the definition in the same sense in Sec. VI.C.

We now review the construction by Somma and Boixo (2013) in some detail. To place it in context, note that the universality results reviewed thus far can be summarized as follows: Any quantum circuit specified by unitary gates $U_1, ..., U_Q$ can be simulated by an adiabatic quantum evolution involving frustration-free Hamiltonians $H(s) = \sum_{k=1}^{L} a_k(s) \Pi_k(s)$. The ground state of the final Hamiltonian $H(1)$ has a large overlap with the output state of the quantum circuit. Moreover, L is polynomial in Q, and Π_k denotes nearest-neighbor, two-body interactions between spins of corresponding many-body systems in one- or two-dimensional lattices. The inverse minimum gap of $H(s)$ is polynomial in Q, and hence so is the duration of the adiabatic simulation.

Now, consider a frustration-free Hamiltonian

$$H(s) = \sum_{k=1}^{L} a_k \Pi_k(s), \qquad (88)$$

where each $\Pi_k(s)$ is a projector for all $s \in [0, 1]$ and is a local operator. Denote the eigenvalues of this Hamiltonian by $\{\lambda_j\}$, where $\lambda_1 = 0$ is the ground state energy. Then, take the Hamiltonian

$$\bar{H}(s) = \sum_{k=1}^{L} \sqrt{a_k} \, \Pi_k(s) \otimes (|k\rangle\langle 0| + |0\rangle\langle k|), \qquad (89)$$

where $|0\rangle$ and $|k\rangle$ are ancilla registers defined over 1 and $\log_2(L)$ qubits, respectively. It can be shown that $\bar{H}(s)$ has the desired properties, i.e., if $|\psi(s)\rangle$ was the ground state of H, then $|\psi(s)\rangle|10\cdots0\rangle$ is a (degenerate) zero-eigenvalue eigenstate of $\bar{H}$ and the eigenvalues of $\bar{H}(s)$ are $\{\pm\sqrt{\lambda_j}\}$ [the proof is given in Appendix B of Somma and Boixo (2013)]. Thus, the gap has been quadratically amplified, and one can evolve with $\bar{H}$ to transform eigenstates at $s = 0$ to eigenstates at $s = 1$ and simulate the original quantum circuit with a quadratic speedup over the simulation involving H.

In general $\bar{H}(s)$ will be $\log_2(L)$ local due to the appearance of $|k\rangle$. To avoid these many-body interactions one can represent $|k\rangle$ using a unary encoding, i.e., $|k\rangle \mapsto |0\cdots010\cdots0\rangle$ (with 1 at the kth position). In this single-particle subspace the new Hamiltonian becomes

$$\bar{H}(s) = \sum_{k=1}^{L} \sqrt{a_k}\, \Pi_k(s) \otimes (\sigma_k^+ \sigma_0^- + \sigma_k^- \sigma_0^+), \quad (90)$$

where $\sigma^{\pm} = (\sigma^x \pm i\sigma^y)/2$ are Pauli raising and lowering operators. Note that since each $\Pi_k(s)$ interacts with the same qubit 0 of the new register, if the original H was geometrically local, then $\bar{H}$ is not, i.e., it has a central spin geometry.

One more issue that needs to be dealt with is the degeneracy of the zero eigenvalue. To remove this degeneracy from contributions within the single-particle subspace one can add a penalty term $\frac{1}{4}\sqrt{\Delta}(\mathbb{1} + \sigma_0^z)$ to $\bar{H}(s)$, which penalizes all states with qubit 0 in $|0\rangle$; the relevant spectral gap in the single-particle subspace is then still of order $\sqrt{\Delta}$. To remove additional degeneracy from the many-particle subspaces one can add penalties for states that belong to such subspaces. Adding $Z = (L-2)\mathbb{1} - \sum_{k=0}^{L} \sigma_k^z$ achieves this since it acts as a penalty that grows with the Hamming weight a of states in the a-particle subspace. Thus

$$H'(s) = \frac{1}{L^{1/d}}\left[\sum_{k=1}^{L} \sqrt{a_k}\, \Pi_k(s) \otimes (|k\rangle\langle 0| + |0\rangle\langle k|) \right.$$
$$\left. + \frac{1}{4}\sqrt{\Delta}(\mathbb{1} + \sigma_0^z) \right] + Z \quad (91)$$

has $|\psi_0\rangle|10\cdots0\rangle$ as a unique eigenstate of eigenvalue 0, and all other eigenvalues are at a distance of at least $\sqrt{\Delta}/L^{1/d}$ if $d \geq 2$.[23] This is the desired quadratic gap amplification result.

How far can gap amplification methods go? Schaller (2008) showed that for the one-dimensional transverse field quantum Ising model, and for the preparation of cluster states (Raussendorf and Briegel, 2001), it is possible to use a series of straight-line interpolations in order to generate a schedule along which the gap is always greater than a constant independent of the system size, thus avoiding the quantum phase transition. However, there exists an efficient method to compute the ground state expectation values of local operators of 2D lattice Hamiltonians undergoing exact adiabatic evolution, and this implies that adiabatic quantum algorithms based on such local Hamiltonians, with unique ground states, can be simulated efficiently if the spectral gap does not scale with the system size (Osborne, 2007).

V. HAMILTONIAN QUANTUM COMPLEXITY THEORY AND UNIVERSAL AQC

In this section we review Hamiltonian quantum complexity theory from the perspective of QMA completeness. This theory naturally incorporates decision problems of the type that motivate AQC. Essentially, it concerns a problem involving the ground state of a local Hamiltonian, whose ground state energy is promised to either be below a threshold a or above another threshold $b > a$, and where $b - a$ is polynomially small in the system size. In some cases this problem is easy and in other cases it turns out to be so hard that we do not hope to solve it efficiently even on a quantum computer. Characterizing which types of local Hamiltonians fall into the latter category is the subject of QMA completeness.

Hamiltonian quantum complexity theory is an extremely rich subject that is rapidly advancing and has already been reviewed a number of times, so we will only touch upon it and highlight some aspects that are relevant to AQC. Perhaps the most direct connection is the fact that 2-local Hamiltonians of a form that naturally appears in AQC are QMA complete. Additionally, some of the technical tools that played an important role in QMA-completeness locality reductions, such as perturbative gadgets, have also found great use in proofs of the universality of AQC with different Hamiltonians.

The reviews by Aharonov and Naveh (2002), Gharibian (2013), and Gharibian et al. (2015) are excellent resources for additional perspectives and details on Hamiltonian quantum complexity theory.

A. Background

1. Boolean satisfiability problem: k-SAT

Consider a Boolean formula Φ that depends on n literals $x_i \in \{0,1\}$ (with 0 and 1 representing false and true, respectively) or their negations. The problem is to decide whether there exists an assignment of values to the literals that satisfies the Boolean formula, i.e., such that $\Phi = 1$. If there exists such an assignment then the formula is satisfiable; otherwise, it is unsatisfiable.

The Boolean formula is typically written in conjunctive normal form: it is written in terms of a conjunction (AND—$\wedge$) of r clauses, where each clause contains the disjunction (OR—$\vee$) of k literals (variables) or their negation (NOT—$\neg$). A literal and its negation are often referred to as positive and negative literals. The Boolean formula is written as

$$\Phi = C_1 \wedge C_2 \wedge \cdots \wedge C_r, \quad (92)$$

where $C_i = x_{i_1} \vee x_{i_2} \cdots \vee x_{i_k}$ and x_{i_j} is the jth positive or negative literal in the ith clause. The question of Boolean satisfiability, or k-SAT, is whether there exists a choice $X = (x_1, \ldots, x_n)$ such that $\Phi(X) = 1$. Note that it requires only $O(kr)$ steps to check whether X is a satisfying assignment, yet there are 2^n possible choices for X.

For $k = 3$, the Boolean satisfiability problem, called 3-SAT, is NP complete. Let us explain what this means.

2. NP, NP complete, and NP hard

Informally, problems in NP are those whose verification can be done efficiently (e.g., checking whether a Boolean formula is satisfied). An important conjecture, called the exponential time hypothesis (Impagliazzo and Paturi, 2001), states that there are problems in NP that take exponentially long to solve.

---

[23]The factor $L^{1/d}$ in Eq. (91) is introduced so that the eigenvalues coming from the many-particle subspaces will not mix with the eigenvalues of the single-particle subspace.

Formally, a decision problem Q is in NP if and only if there is an efficient algorithm V, called the verifier, such that for all inputs η (e.g., in the case of SAT, this would be the clauses) of the problem:

- If $Q(\eta) = 1$, then there exists a witness X such that $V(\eta, X) = 1$.
- If $Q(\eta) = 0$, then for all witnesses X we have $V(\eta, X) = 0$.

In both cases, we typically take $|X| = \text{poly}(|\eta|)$, where $|\eta|$ is the number of bits in the binary string associated with the input η. The verifier is efficient in the sense that its cost scales as $\text{poly}(|X|)$. In SAT the witness X would be our test assignment.

A decision problem Q is NP complete if

- Q is in NP.
- Every problem in NP is reducible to Q in polynomial time.

Here reducibility means that given a problem A in NP and a problem B that is NP complete, A can be solved using a hypothetical polynomial-time algorithm that solves for B. A commonly used reduction is the polynomial-time many-to-one reduction (Karp, 1972), whereby the inputs of A are mapped into the inputs to B such that the output of B matches the output of A. The hypothetical algorithm then solves B to get the answer to A.

A decision problem Q is NP hard if every problem in NP is reducible to Q in polynomial time. (Note that unlike the NP-complete case, Q does not need to be in NP.) Clearly, NP complete $\subseteq$ NP hard.

3. The k-local Hamiltonian problem

The history state construction of Sec. IV.B relies on a 5-local Hamiltonian. Such a Hamiltonian belongs to an important class of decision problems known as the k-local Hamiltonian problem of which a complete complexity classification was given by Cubitt and Montanaro (2016) subject to restrictions on the set of local terms from which the Hamiltonian can be composed [see also Bravyi and Hastings (2014)]. Recall that a k-local Hamiltonian is a Hermitian matrix that acts nontrivially on at most k p-state particles.

The k-local Hamiltonian problem is defined on n qubits, with the following input:

- A k-local Hamiltonian $H = \sum_{i=1}^{r} H_i$ with $r = \text{poly}(n)$. Each H_i is k local and satisfies $\|H_i\| = \text{poly}(n)$ and its nonzero entries are specified by $\text{poly}(n)$ bits.
- Two real numbers a and b specified with $\text{poly}(n)$ bits of precision, such that

$$b - a > \frac{1}{\text{poly}(n)}. \qquad (93)$$

The output (0 or 1) answers the question: Is the smallest eigenvalue of H smaller than a (output is 1), or are all eigenvalues larger than b (output is 0)? We are promised that the ground state eigenvalue cannot be between a and b.[24]

We can map 3-SAT to the 3-local Hamiltonian problem as follows. For every clause C_i (which involves three literals), we

---

[24]The quantity $b - a$ is sometimes called the "promise gap" and is distinct from the spectral gap.

can define a 3-local projector H_i onto all the unsatisfying assignments of C_i. Because H_i is a projector, it has eigenvalues 0 and 1, where the 0 eigenvalue is associated with satisfying assignments and the 1 eigenvalue with unsatisfying assignments. Therefore,

$$H|X\rangle = \sum_{i=1}^{r} H_i |X\rangle = q|X\rangle, \qquad (94)$$

where q is the number of unsatisfied assignments by X. Thus 3-SAT is equivalent to the following 3-local Hamiltonian problem: is the smallest eigenvalue of H zero (the 3-SAT problem is satisfiable) or is it at least 1 (the 3-SAT problem is unsatisfiable)?

4. Motivation for adiabatic quantum computing

Adiabatic evolution seems well suited to tackling the k-local Hamiltonian problem. By initializing an n-qubit system in an easily prepared ground state, we can in principle evolve the system with a time-dependent Hamiltonian whose end point is the k-local Hamiltonian. If the evolution is adiabatic, then we are guaranteed to be in the ground state of the k-local Hamiltonian with high probability. By measuring the state of the system, we can determine the energy eigenvalue of the state (which hopefully is the ground state energy) and hence determine the answer to an NP-complete problem such as 3-SAT. This motivated early work on the quantum adiabatic algorithm (Farhi et al., 2001).

Another possibility is to try to use AQC as the verifier. However, the quantum algorithm gives us the answer only probabilistically, so we must first define a probabilistic analog of NP and then a quantum version. These new complexity classes are MA and QMA (Kitaev, Shen, and Vyalyi, 2000).

B. MA and QMA

Informally, MA can be thought of as a probabilistic analog of NP, allowing for two-sided errors. Formally, a decision problem Q is in MA iff there is an efficient probabilistic verifier V such that for all inputs η of the problem:

- If $Q(\eta) = 1$, then there exists a witness X such that $\Pr[V(\eta, X) = 1] \geq 2/3$ (completeness).
- If $Q(\eta) = 0$, then for all witnesses X we have $\Pr[V(\eta, X) = 1] \leq 1/3$ (soundness).

Again we take $|X| = \text{poly}(|\eta|)$. MA is typically viewed as an interaction between two parties, Merlin and Arthur. Merlin provides Arthur with a witness X on which Arthur runs V. If $Q(\eta) = 0$, Merlin should never be able to fool Arthur with a witness X into believing that $Q(\eta) = 1$ with probability $> 1/3$.

Note that there is nothing special about the probabilities $(2/3, 1/3)$. We can generalize our description to $\text{MA}(c, s)$.

Claim 1. $\text{MA}(c, c - 1/|\eta|^g) \subseteq \text{MA}(2/3, 1/3) = \text{MA}(1 - e^{-|\eta|^g}, e^{-|\eta|^g})$, where g is a constant, $c > 0$, and $c - 1/|\eta|^g < 1$.

The proof of this "amplification lemma" (Marriott and Watrous, 2005; Goldreich, 2008; Nagaj, Wocjan, and Zhang, 2009) is interesting since it invokes the Chernoff bound, a widely used tool. We thus present it in Appendix D for pedagogical interest.

The complexity class QMA can be viewed as the quantum analog of MA. Thus, QMA is informally the class of problems that can be efficiently checked on a quantum computer given a "witness" quantum state related to the answer to the problem. Formally, define a quantum verifier V (a quantum circuit) that takes η and a quantum witness state $|X\rangle \in (\mathbb{C}^2)^{\otimes \text{poly}(|\eta|)}$ as inputs and probabilistically outputs a binary number. The decision problem Q is said to be in QMA if and only if there exists an efficient (polynomial time) V for all inputs η of the problem that satisfies the following:

- If $Q(\eta) = 1$, then there exists a witness $|X\rangle$ such that $\Pr[V(\eta, |X\rangle) = 1] \geq 2/3$ (completeness).
- If $Q(\eta) = 0$, then for all witnesses $|X\rangle$ we have $\Pr[V(\eta, |X\rangle) = 1] \leq 1/3$ (soundness).

The amplification lemma applies here as well. The definition of QMA also allows V to have $\text{poly}(|\eta|)$ ancilla qubits each initialized in the $|0\rangle$ state (Gharibian *et al.*, 2015).

One can also define the class QCMA, which is similar to QMA except that $|X\rangle$ is a classical state (Aharonov and Naveh, 2002). Since the quantum verifier can force Merlin to send him a classical witness by measuring the witness before applying the quantum algorithm, we have MA $\subseteq$ QCMA $\subseteq$ QMA.

C. The general relation between QMA completeness and universal AQC

The class of efficiently solvable problems on a quantum computer is bounded error quantum polynomial time (BQP) (Bernstein and Vazirani, 1993), which consists of the class of decision problems solvable by a uniform family of polynomial-size quantum circuits with error probability bounded below 1/2. Because of the polynomial equivalence between AQC and the circuit model, BQP is also the class of efficiently solvable problems on a universal adiabatic quantum computer. Its classical analog is the class bounded error probabilistic polynomial time (BPP), and as expected BPP $\subseteq$ BQP (Bernstein and Vazirani, 1997). In addition, BQP $\subseteq$ QCMA (Aharonov and Naveh, 2002). Another interesting characterization is that BQP = QMA$_{\log}$, where QMA$_{\log}$ is the same as QMA except that the quantum proof has $O(\log |\eta|)$ qubits instead of $\text{poly}(|\eta|)$ (Marriott and Watrous, 2005).

This motivates the study of QMA, and, in particular, QMA completeness, as a tool for understanding universality. Indeed, it is often the case that whenever adiabatic universality can be proven for some class of Hamiltonians, then the local Hamiltonian problem with roughly the same class can be shown to be QMA complete and vice versa. Note, however, that there is no formal implication from either of those problems to the other (Aharonov *et al.*, 2009). On the one hand, proving QMA completeness is in general substantially harder than achieving universal AQC, where we can choose the initial state to be any easily prepared state that will help us solve the problem, so we can choose to work in any convenient subspace that is invariant under the Hamiltonian. Indeed, in the history state construction, we introduce penalty terms to guard against illegal clock states [recall Eq. (74)]. For QMA, the states we work with are chosen adversarially from the *full* Hilbert space, and we must be able to check, using only local Hamiltonian terms, that they are of the correct (clock-state) form. On the other hand, proving adiabatic universality involves analyzing the spectral gap of the continuous sequence of Hamiltonians over the *entire* duration of the computation, whereas QMA-completeness proofs are concerned with only one Hamiltonian.

D. QMA completeness of the *k*-local Hamiltonian problem and universal AQC

To prove that a promise problem is QMA complete, one needs to prove that it is contained in QMA and that it is QMA hard. The k-local Hamiltonian problem belongs to QMA for any constant k, and in fact even for $k = O(\log n)$ (Kitaev, Shen, and Vyalyi, 2000). For pedagogical proofs see Aharonov and Naveh (2002), Gharibian (2013), and Gharibian *et al.* (2015).

The first example of a QMA-hard problem was the k-local Hamiltonian problem for $k \geq 5$ (Kitaev, Shen, and Vyalyi, 2000), so that, in particular, the 5-local Hamiltonian problem is QMA complete. This was reduced to 3-local (Kempe and Regev, 2003; Nagaj and Mozes, 2007) and then to 2-local (Kempe, Kitaev, and Regev, 2006). Note that the 1-local Hamiltonian problem is in the complexity class P, since one can simply optimize for each 1-local term independently. Various simplifications of QMA completeness for the 2-local case followed. In order to describe these, we first need to define a class of Hamiltonians

$$
\begin{aligned}
H_1 = & \sum_{(i,j) \in \mathcal{E}} J_{ij}^x X_i X_j + J_{ij}^y Y_i Y_j + J_{ij}^z Z_i Z_j \\
& + \sum_{i \in \mathcal{V}} h_i^x X_i + h_i^y Y_i + h_i^z Z_i,
\end{aligned}
\tag{95}
$$

where $\mathcal{V}$ and $\mathcal{E}$ are the vertex and edge sets of a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, and all local fields $\{h_i^\alpha\}$ and couplings $\{J_{ij}^\alpha\}$ ($\alpha \in \{x, y, z\}$) are real. The Heisenberg model corresponds to $J_{ij}^x = J_{ij}^y = J_{ij}^z$, the XY model to $J_{ij}^x = J_{ij}^y$ and $J_{ij}^z = 0$, and the Ising model to $J_{ij}^x = J_{ij}^y = 0$. When $J_{ij}^\alpha < 0$ (> 0) the interaction between qubits i and j is ferromagnetic (antiferromagnetic). When we write "fully" below we mean that all interactions have the same sign. Unless explicitly mentioned otherwise we assume that the local fields are all zero.

Most of the simplifications of QMA completeness are special cases of Eq. (95):

- Geometrical locality: nearest-neighbor interactions with $\mathcal{G}$ being a 2D square lattice (Oliveira and Terhal, 2008) or a triangular lattice (Piddock and Montanaro, 2015).
- Simple interactions in 2D: $ZZXX$ and ZX models (Biamonte and Love, 2008) [defined in Eqs. (98) and (99)], fully ferromagnetic and fully antiferromagnetic Heisenberg models with local fields (Schuch and Verstraete, 2009), antiferromagnetic Heisenberg and XY models without local fields (Piddock and Montanaro, 2015).

Some of the simplifications of QMA completeness use other types of Hamiltonians:

- Interacting fermions in 2D and the space-time construction (Breuckmann and Terhal, 2014).
- Multistate particles in 1D (Nagaj, 2008; Aharonov *et al.*, 2009; Hallgren, Nagaj, and Narayanaswami, 2013).
- Nontranslationally invariant 1D systems (all two-particle terms identical but position-dependent one-particle terms) (Kay, 2008).

- Translationally invariant 1D systems for which finding the ground state energy is complete for QMA_{exp}[25] (Gottesman and Irani, 2013).

The 1D case is interesting since, as remarked in Sec. IV.E, the 1D restriction of MAX-2-SAT with p-state variables is in P, yet Aharonov *et al.* (2009) showed that for 12-state particles the problem of approximating the ground state energy of a 1D system is QMA complete. This result was improved to 11-state particles in (Nagaj (2008), and then to 8-state particles in Hallgren, Nagaj, and Narayanaswami (2013), who also pointed out a small error in Aharonov *et al.* (2009) that could be fixed by using 13-state particles. Whether and at which point a further Hilbert space dimensionality reduction becomes impossible remains an interesting open problem.

The reduction from 5-local to 2-local is done using perturbative gadgets (Kempe, Kitaev, and Regev, 2006; Biamonte and Love, 2008; Bravyi, DiVencenzo, Loss, and Terhal, 2008; Jordan and Farhi, 2008; Oliveira and Terhal, 2008; Cao *et al.*, 2015). The goal of the gadget is to approximate some target Hamiltonian H^{T} of n qubits (e.g., the 5-local Hamiltonian from the history state construction at any time s) by a gadget Hamiltonian H^{G} acting on the same n qubits as well as an additional poly(n) ancilla qubits. The gadget Hamiltonian is typically written as

$$H^{\text{G}} = H^{\text{A}} + \lambda V, \qquad (96)$$

where H^{A} is an unperturbed Hamiltonian (also called the penalty Hamiltonian), acting only on the ancilla space, and λV is a perturbation that acts between the qubits of H^{T} and the ancilla qubits. Using perturbation theory, which we review in Appendix E, one can show that the lowest 2^n eigenvalues of H^{G} differ from those of H^{T} by at most ϵ and the corresponding eigenstates have an overlap of at least $1 - \epsilon$.

Completeness of the 2-local Hamiltonian problem means that every problem in QMA is reducible to the 2-local Hamiltonian decision problem in polynomial time. Since this reduction involves perturbative gadgets that preserve the spectrum of the original 5-local Hamiltonian, this means that the 2-local Hamiltonian derived from the 5-local Hamiltonian appearing in the universality proof of Sec. IV.B will also have an energy gap that is an inverse polynomial in the circuit length, and that the computation remains in the ground subspace with illegal clock states gapped away by the (now 2-local) penalty Hamiltonian. In the remainder of this section we discuss a particularly simple form of 2-local Hamiltonians that is universal for AQC.

The QMA completeness of general 2-local Hamiltonians can be extended to show that a more restricted set of 2-local Hamiltonians composed of real-valued sums of the following pairwise products of Pauli matrices are QMA complete (Biamonte and Love, 2008):

$$\{IX, XI, IZ, ZI, ZX, XZ, ZZ, XX\}. \qquad (97)$$

The two basic steps to do this are as follows: (1) Using the result of Bernstein and Vazirani (1997) that any quantum circuit can be represented using real-valued unitary gates operating on real-valued wave functions in the proof of the QMA completeness of the 5-local Hamiltonian of the previous section, the Hamiltonian terms are all real valued. This therefore extends QMA completeness to 5-local real Hamiltonians. (2) The same gadgets used by Kempe, Kitaev, and Regev (2006) and Oliveira and Terhal (2008) can be used to reduce the locality from five to two.

This can be further simplified to show that "$ZZXX$ Hamiltonians" that are linear combinations with real coefficients of only

$$\{IX, XI, IZ, ZI, ZZ, XX\} \qquad (98)$$

are QMA complete. This is done by showing, using perturbation theory, that such Hamiltonians can be used to approximate the $\sigma^z \otimes \sigma^x$ and $\sigma^x \otimes \sigma^z$ terms. Similarly, perturbation theory can be used to show that "ZX Hamiltonians" that are linear combinations with real coefficients of only

$$\{IX, XI, IZ, ZI, ZX, XZ\} \qquad (99)$$

are QMA complete (Biamonte and Love, 2008; Bravyi and Hastings, 2014; Cubitt and Montanaro, 2016).

VI. STOQUASTIC ADIABATIC QUANTUM COMPUTATION

In this section we focus on the special class of "stoquastic Hamiltonians" [originally introduced by Bravyi, DiVencenzo, Oliveira, and Terhal (2008)] that often arise in the context of quantum optimization.

Definition 5 (stoquastic Hamiltonian). A Hamiltonian H is called stoquastic (Bravyi and Terhal, 2009) with respect to a basis $\mathcal{B}$ iff H has real nonpositive off-diagonal matrix elements in the basis $\mathcal{B}$.

For example, a Hamiltonian is stoquastic in the computational basis iff

$$\langle x|H|x'\rangle \le 0 \quad \forall \, x, x' \in \{0, 1\}^n \qquad x \neq x'. \quad (100)$$

The computational basis is often singled out since it plays the role of the basis in which the final Hamiltonian is measured, which sometimes coincides with the basis in which that Hamiltonian is diagonal. The term stoquastic was introduced due to the similarity to stochastic matrices, such as arise in the theory of classical Markov chains.

Restricting to any basis still leaves some freedom in the definition. For example, a Hamiltonian $H = -\sum_i \sigma_i^x + H_Z$, where H_Z is diagonal in the computational basis, is clearly stoquastic. However, applying a unitary transformation $U = \prod_i \sigma_i^z$ to the Hamiltonian gives $H' = \sum_i \sigma_i^x + H_Z$, which according to Definition 5 is not stoquastic in the computational basis. Applying a local unitary basis transformation should not change the complexity of the problem. Therefore, for clarity we fix the basis such that the standard initial Hamiltonian always carries a minus sign, i.e., $-\sum_i \sigma_i^x$. From this point forward, we restrict our discussion of stoquasticity

---

[25]QMA_{exp} is the same as QMA but with exponential size (in the input) witness and verification circuit, whereas both are polynomial for QMA.

to the standard computational basis. With this in mind, the class of stoquastic Hamiltonians includes the fully ferromagnetic Heisenberg and *XY* models, and the quantum transverse field Ising model [recall Eq. (95)].

Given the restriction of the Hamiltonian, one may ask whether there is a complexity class for which the *k*-local stoquastic Hamiltonian problem is complete. This led to the introduction of the class StoqMA, for which the $k \geq 2$-local stoquastic Hamiltonian is StoqMA complete (Bravyi, Bessen, and Terhal, 2006). This can be further refined to the result that the transverse Ising model on degree-3 graphs is StoqMA complete (Bravyi and Hastings, 2017). Rather than give the formal and rather involved definition of StoqMA, we note that the only difference between StoqMA and MA is that a stoquastic verifier in StoqMA is allowed to do the final measurement in the $\{|+\rangle, |-\rangle\}$ basis, whereas a classical coherent verifier in MA can do a measurement only in the standard $\{|0\rangle, |1\rangle\}$ basis.[26] Unlike MA and QMA, the threshold probabilities in StoqMA have an inverse polynomial rather than constant separation; this prevents amplification of the gap between the threshold probabilities based on repeated measurements with majority voting. Finally, it is known that MA $\subseteq$ StoqMA $\subseteq$ QMA (Bravyi, Bessen, and Terhal, 2006).

To capture the important class of problems that are characterized by stoquastic evolution with the constraint of adiabatic evolution, we first introduce the following definition of a model of computation:

Definition 6 (StoqAQC). Stoquastic adiabatic quantum computation (StoqAQC) is the special case of AQC (Definition 1) restricted to *k*-local (*k* fixed) stoquastic Hamiltonians.

Because we defined StoqAQC as a special case of AQC, the computation must proceed in the ground state. However, recall that the algorithm for the glued-trees problem (Sec. III.D) is not subject to this ground state restriction and hence is not in StoqAQC. In Sec. VI.C we consider another model of stoquastic computation that is not subject to the ground state restriction.

StoqAQC has generated considerable interest since experimental implementations of stoquastic Hamiltonians are quite advanced (Bunyk *et al.*, 2014; Weber *et al.*, 2017). To characterize its computational power, we introduce a natural promise problem based on StoqAQC and modeled after the *k*-local Hamiltonian problem[27]:

Definition 7 (StoqAQCEval). The StoqAQCEval problem is defined on *n* qubits, with the following input:

- A continuous family of $(k \geq 2)$-local (*k* fixed) stoquastic Hamiltonians $H(s) = \sum_{i=1}^{r} H_i(s)$ with $r = \text{poly}(n)$ and parametrized by $s \in [0, 1]$. For all *i* and all *s*, the nonzero entries of $H_i(s)$ are specified by poly(n) bits of precision, and $\|H_i(s)\| = \text{poly}(n)$. The ground state
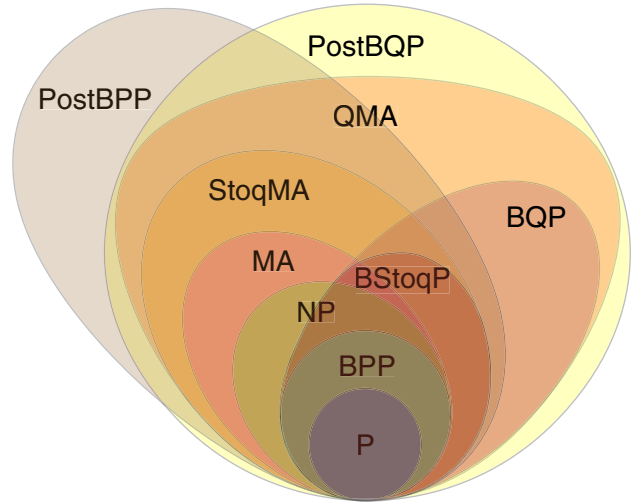

FIG. 4. Known relations between complexity classes relevant for AQC. The BStoqP class defined here (Definition 8) lies in the intersection StoqMA and BQP and includes BPP.

energy gap $\Delta[H(s)]$ satisfies $\Delta[H(s)] \geq 1/\text{poly}(n)$ for all *s*.
- Two real numbers *a* and *b* specified with poly(n) bits of precision, and $b - a > 1/\text{poly}(n)$.

The output (0 or 1) answers the question: Is the smallest eigenvalue of $H(s = 1)$ smaller than *a* (output is 1), or are all eigenvalues larger than *b* (output is 0)? Just as in the local Hamiltonian problem, we are promised that the outcome in which the ground state energy is between *a* and *b* is not possible.

This allows us to informally define the complexity class that captures StoqAQC.

Definition 8 (BStoqP). BStoqP is the set of problems that are polynomial-time reducible to StoqAQCEval.

The StoqAQCEval problem is clearly in StoqMA, because the $(k \geq 2)$-local stoquastic Hamiltonian problem is StoqMA complete (Bravyi, Bessen, and Terhal, 2006). Hence BStoqP $\subseteq$ StoqMA, as depicted in Fig. 4, which summarizes the relations between many of the complexity classes discussed.[28] NP and MA are unlikely to be subsets of BStoqP, since StoqAQC would not be expected to solve NP-complete problems in polynomial time. The tightest inclusion in a classical complexity class we know of is in AM,[29] since the latter includes StoqMA (Bravyi, DiVencenzo, Oliveira, and Terhal, 2008). It is clear that BPP $\subseteq$ BStoqP, since 5-local StoqAQCEval is BPP hard (using a classical reversible circuit for universal AQC, with stoquastic gate terms and a 5-local stoquastic clock Hamiltonian). Finally, we know that BStoqP $\subseteq$ BQP, since StoqEvalAQC is in BQP by using

---

[26]MA has an alternative quantum definition as a restricted version of QMA in which the verifier is a coherent classical computer (Bravyi, DiVencenzo, Oliveira, and Terhal, 2008).

[27]We are indebted to Elizabeth Crosson for help in formulating the StoqAQCEval problem, the BStoqP class, and working out the relations of BStoqP to other complexity classes.

[28]As far as we know the related term StoqP was informally introduced by Stephen Jordan in a talk presented at the AQC 2016 conference (Jordan, 2016a), showing that StoqP is not equal to BQP unless BQP is in the third level of the polynomial hierarchy.

[29]Like MA, the class AM (Arthur Merlin) is a probabilistic generalization of NP. See https://complexityzoo.uwaterloo.ca/Complexity_Zoo for definitions of the complexity classes mentioned here.

the same proof that the adiabatic model in general can be simulated by the circuit model.

A. Why it might be easy to simulate stoquastic Hamiltonians

In this section we summarize the complexity-theoretic evidence obtained so far that suggests that the StoqAQC setting is less powerful than universal quantum computation. Let us start with a lemma that characterizes the "classicality" of ground states of stoquastic Hamiltonians.

Lemma 3. The ground state $|\psi\rangle$ of a stoquastic Hamiltonian H can always be expressed using only real non-negative amplitudes $|\psi\rangle = \sum_{x\in\{0,1\}^n} a_x|x\rangle$, where $a_x \geq 0 \,\forall\, x$.

Proof. It follows directly from the stoquastic property that the corresponding Gibbs density matrix $\rho = \exp(-\beta H)/\text{Tr}[\exp(-\beta H)]$ has non-negative matrix elements in the computational basis for any $\beta > 0$. In particular, if H is stoquastic then for sufficiently small β, $\mathbb{1} - \beta H$ has only non-negative matrix elements. The largest eigenvalue of $\mathbb{1} - \beta H$ corresponds to the ground state energy of H. Thus, by the Perron-Frobenius theorem (see Sec. III.E.1) the ground state of H can be chosen to have non-negative amplitudes. ∎

Consequently, if the Hamiltonian is stoquastic, a classical probability distribution can be associated with the ground state. This raises the question as to whether StoqAQC is a model that is capable of quantum speedup over classical algorithms. Following is the evidence regarding this question.

(1) The ground state energy of the fully ferromagnetic transverse field Ising model can be found to a given additive error in polynomial time with a classical algorithm on any graph, with or without a transverse magnetic field (Bravyi and Gosset, 2016).

(2) Bravyi, DiVencenzo, Oliveira, and Terhal (2008) showed that for any fixed k the stoquastic k-local Hamiltonian is contained in the complexity class AM. Thus, unless QMA ⊆ AM (which is believed to be unlikely), the stoquastic k-local Hamiltonian is not QMA complete.

(3) Bravyi, DiVencenzo, Oliveira, and Terhal (2008) also showed that gapped StoqAQC can be simulated in PostBPP, the complexity class described by a polynomial-time classical randomized computer with the ability to postselect on some subset of the bits after the algorithm is run.[30] That is, it suffices to call an oracle for problems in PostBPP a polynomial number of times to efficiently sample from the ground state of a gapped stoquastic Hamiltonian.[31]

Suppose that StoqAQC could be used to perform universal quantum computation. Since gapped StoqAQC can be simulated in PostBPP, this implies that SampBQP ⊆ SampPostBPP.[32] In other words, this implies that polynomial-time quantum algorithms can be simulated classically in polynomial time using postselection. This implies that PostBPP = PostBQP which in turn would collapse the polynomial hierarchy. Thus it is unlikely that StoqAQC is universal for AQC.

(4) Bravyi and Terhal (2009) showed that adiabatic evolution along a path composed entirely of stoquastic frustration-free Hamiltonians (recall Definition 4) may be simulated by a sequence of classical random walks, i.e., is contained in BPP.

With this evidence for the potential limitations of stoquastic Hamiltonians, the question arises if they are worthy of pursuit, either theoretically or experimentally. However, it is important to remember that the weakness of stoquastic Hamiltonians arises when one assumes that they generate an evolution that occurs in the ground state. Indeed, we will see in Sec. VI.C that *excited state* stoquastic evolution can be as powerful as universal AQC. Moreover, in the next section we review counterexamples to the claim that stoquastic Hamiltonians are necessarily easy to simulate using heuristic classical algorithms.

B. Why it might be hard to simulate stoquastic Hamiltonians

A general theorem does not exist that rules out a quantum speedup of StoqAQC over all possible classical algorithms. However, it is often stated that Monte Carlo simulations of StoqAQC do not suffer from the sign problem and will therefore simulate StoqAQC without a slowdown. Specifically, the conjecture is that if the Monte Carlo simulation starts at $s = 0$ in the equilibrium state, and if s changes by a small amount ϵ from one step to the next, where ϵ is polynomially small in the system size n, the inverse temperature β, and the spectral gap Δ, then the Monte Carlo simulation stays close to the equilibrium state along the path. For sufficiently large β, this corresponds to following the instantaneous ground state. In this section we review theoretical evidence that such a conjecture is not always true. We focus on two of the most direct classical competitors to StoqAQC: path integral quantum Monte Carlo (PI-QMC), and diffusion quantum Monte Carlo (D-QMC).

1. Topological obstructions

Hastings and Freedman (2013) gave examples of StoqAQC with a polynomially small eigenvalue gap, but where PI-QMC takes an exponential time to converge. Loosely, the failure of convergence was due to topological obstructions around

---

[30]See also Farhi and Harrow (2016), where gapped StoqAQC was called stoquastic gapped adiabatic evolution.

[31]PostBPP, also known as BPP$_{\text{path}}$, contains NP. For example, consider the Grover problem with two registers, a bit string x for the input and a second register where $f(x)$ is stored. Now if we pick x at random and postselect on the second register being 1, we find a marked item. PostBPP is known to be contained in the third level of the polynomial hierarchy (Han, Hemaspaandra, and Thierauf, 1993).

[32]In sampling problems we are given an input $x \in \{0, 1\}^n$, and the goal is to sample (exactly or approximately) from some probability distribution over poly(n)-bit strings. SampBQP and SampPostBPP are the classes of sampling problems solvable on quantum computers and probabilistic classical computers with postselection, respectively, to within ϵ error in total variation (or trace-norm) distance, in time polynomial in n and $1/\epsilon$ (Aaronson, 2010).

which the worldlines (trajectories in imaginary time) can get tangled.

The simplest of the examples can be understood intuitively as follows. A sombrerolike potential is constructed for a single particle with a deep circular minimum of radius r. The worldline of the particle in PI-QMC with closed boundary conditions is some closed path that follows this circle in imaginary time. Because of the depth of the potential at the minimum the distribution of trajectories has very small probability to include any point with radius larger than r and it takes an exponential time in the winding number to transition from one winding number sector to another. Therefore, if an appropriate dimensionless combination of the radius r or the mass of the particle is changed sufficiently fast then PI-QMC fails to equilibrate. At the same time it can be shown that for this example the gap closes polynomially and so one expects that adiabatic evolution requires only polynomial time to find the ground state.

While this example uses winding numbers to construct a protocol for which PI-QMC takes exponential time to equilibrate, PI-QMC can still find the ground state. To observe a more dramatic effect where not only equilibration is hampered but also the probability of finding the ground state is low, one can introduce stronger topological effects and additionally exploit the discrepancy between L_1- and L_2-normalized wave functions. This was first done in the "bouquet of circles" example introduced by Hastings and Freedman (2013), which shows that PI-QMC can fail to converge even when using open boundary conditions. The example was designed so that the majority of the amplitude ψ lies within an expander graph, although the majority of the probability $|\psi|^2$ does not. Because the end points of the wordlines are distributed according to ψ and not $|\psi|^2$, this effectively "pins" them to the expander graph. This pinning means that even though the worldline is in principle open, the worldline is nevertheless prevented from changing its topological sector within the bouquet of circles. This then causes failure of convergence.

A more general method using perturbative gadgets was explained by Hastings and Freedman (2013) that allows one to map between continuous variables and spins and applies to all the examples given there.

2. Nontopological obstructions

Diffusion Monte Carlo algorithms should not be affected by topological obstructions that depend on closed boundary conditions, since they do not exhibit periodicity in the imaginary-time direction. Rather than use topological obstructions, it is possible to rely entirely on the discrepancy between L_1 and L_2 normalization to design examples where Monte Carlo methods have differing convergence from AQC. This discrepancy was used by Jarret, Jordan, and Lackey (2016) to ensure that the walkers in a D-QMC algorithm never "learn" about a potential well that contains the solution, causing D-QMC to take exponential time to converge. Since the gap for the adiabatic process is large, QA takes only polynomial time.

Let $H(s)$ be some stoquastic Hamiltonian acting on a Hilbert space whose basis states can be equated with the vertices V of some graph. Let $\psi_s(x)\colon V \mapsto \mathbb{C}$ denote the ground state of $H(s)$. Define probability distributions $p_s^{(1)}(x) = \psi_s(x)/\sum_{y \in V} \psi_s(y)$ and $p_s^{(2)}(x) = \psi_s^2(x)$. The stoquasticity of $H(s)$ ensures that $\psi_s(x) \geq 0$, so that $p_s^{(1)}(x)$ is a valid probability distribution.

D-QMC algorithms perform random walks designed to ensure that a population of random walkers converges to $p_s^{(1)}(x)$. However, in exponentially large Hilbert spaces there can be vertices such that the distribution associated with the L_2-normalized wave function $p_s^{(2)}(x)$ is polynomial, but the distribution associated with the L_1-normalized wave function $p_s^{(1)}(x)$ is exponentially small. The idea behind the examples by Jarret, Jordan, and Lackey (2016) is to exploit this discrepancy to design polynomial-time stoquastic adiabatic processes that the corresponding D-QMC simulations will fail to efficiently simulate.

The main example given by Jarret, Jordan, and Lackey (2016) is the stoquastic Hamiltonian $H(s) = (1/n)[L + b(s)W] - c(s)P$, where L is the graph Laplacian of the n-bit hypercube, W is the Hamming weight operator (i.e., $W|x\rangle = |x||x\rangle$ where $|x|$ is the Hamming weight of the bit string x), and $P = |0\cdots0\rangle\langle0\cdots0|$. In terms of Pauli matrices this Hamiltonian can be written, up to an overall constant, as

$$H(s) = -\frac{1}{n}\sum_{j=1}^{n}\left(X_j + \frac{1}{2}b(s)Z_j\right) - c(s)P. \quad (101)$$

The schedules $b(s)$ and $c(s)$ are

$$b(s) = \begin{cases} 2sb, \\ b, \end{cases} \quad c(s) = \begin{cases} 0, & s \in [0, 1/2), \\ (2s-1)c, & s \in [1/2, 1]. \end{cases} \quad (102)$$

For $s \in [0, 1/2)$ this is a Hamiltonian of n noninteracting qubits whose gap is easily seen to be $(2/n)\sqrt{1 + (sb/2)^2}$, minimized at $s = 0$ where it equals $2/n$. The ground state is given by $|\psi(\theta)\rangle^{\otimes n}$, where $|\psi(\theta)\rangle = \cos(\theta/2)|0\rangle + \sin(\theta/2)|1\rangle$ and $\theta = \tan^{-1}(2/sb)$. For $s \in [1/2, 1]$ it can be shown that the minimum gap is attained at $s = 1/2$, where it equals $1/\sqrt{2n} + O(n^{-3/2})$. Thus the overall minimum gap is polynomial $(2/n)$ and the StoqAQC process converges to the ground state $|0\cdots0\rangle$ in polynomial time. By choosing b so that at $s = 1$ we have $\cos(\theta/2) = 1 - 1/4n$, it is easy to show from the analysis of the noninteracting problem that the probability of ending up in the ground state is $p_{s=1}^{(2)}(0\cdots0) = \cos^{2n}(\theta_{s=1}/2) \to e^{-1/2}$ in the limit $n \to \infty$.

On the other hand, for the noninteracting problem (when $s \in [0, 1/2)$) the D-QMC process[33] samples from the distribution $p_s^{(1)}(x) = \sin(\theta/2)^{|x|}\cos(\theta/2)^{n-|x|}/Z_s$, where

$$Z_s = \sum_{x \in \{0,1\}^n} \sin(\theta/2)^{|x|}\cos(\theta/2)^{n-|x|} = [\sin(\theta/2) + \cos(\theta/2)]^n,$$

---

[33]Here D-QMC refers to the "substochastic Monte Carlo" algorithm introduced by Jarret, Jordan, and Lackey (2016).

so that for large n we have $Z_1 \approx (1 + 1/\sqrt{2n})^n \to e^{\sqrt{n/2}}$ for the same choice of b. Thus, for D-QMC the probability of being in the ground state at $s = 1/2$ is

$$p^{(1)}_{s=1/2}(0 \cdots 0) = \cos^n(\theta_{s=1}/2)/Z_{s=1} \to e^{-1/4} e^{-\sqrt{n/2}}.$$

Since at $s = 1/2$ the random walkers that diffuse in the D-QMC process have a probability to be at the all-zeros string that is of order $e^{-\sqrt{n/2}}$, with high likelihood, no walkers will land on the all-zeros string until the number of time steps times the number of walkers approaches $e^{\sqrt{n/2}}$. Until this happens it is impossible for the distribution of walkers to be affected by the change in the potential at the all-zeros string that is occurring from $s = 1/2$ to 1; no walkers have landed there, and the D-QMC algorithm has therefore never queried the value of the potential at that site. Only after allowing for this exponential cost, and by appropriately choosing c, does the D-QMC algorithm find the ground state with high probability.[34]

C. QMA-complete problems and universal AQC using stoquastic Hamiltonians with excited states

Our definition of StoqAQC (Definition 6) stipulates that the computation must proceed in the ground state. It turns out that if this condition is relaxed, computation with stoquastic Hamiltonians is as powerful as AQC, i.e., it is universal. Here we review a construction by Jordan, Gosset, and Love (2010) of a 3-local stoquastic Hamiltonian that, by allowing for excited state evolution, is both QMA complete and universal for AQC.

We start with the QMA-complete Hamiltonian introduced in Sec. V.D that can be written as

$$H_{ZZXX} = \sum_i d_i X_i + h_i Z_i + \sum_{i \leq j} J^x_{ij} X_i X_j + J^z_{ij} Z_i Z_j, \qquad (103)$$

where d_i, h_i, J^x_{ij}, and J^z_{ij} are arbitrary real coefficients. The key idea is to eliminate the negative matrix elements in each term. Toward this end the Hamiltonian is written as

$$H_{ZZXX} = -\sum_k \alpha_k T_k, \qquad (104)$$

where $T_k \in \{\pm X_i, \pm Z_i, \pm X_i X_j, \pm Z_i Z_j\}$ and such that $\alpha_k > 0$. For an n-qubit system, the operators T_k are represented by $2^n \times 2^n$ symmetric matrices with entries taking values $+1, -1$, and 0. We use the regular representation of the Z_2 group to make the replacement

$$1 \to \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad -1 \to \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad 0 \to \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

$$(105)$$

in T_k to define a new operator $\tilde{T}_k$. The matrix representation of $\tilde{T}_k$ is of size $2^{n+1} \times 2^{n+1}$, and since the original T_k was either

---

[34] For $c = 2$ one finds that $p^{(1)}_{s=1}(0 \cdots 0) = 1/2 + O(n^{-1/2})$.

1-local or 2-local acting on n qubits, we can interpret $\tilde{T}_k$ as being 2-local or 3-local acting on $n + 1$ qubits. Note that T_k is such that it only has one nonzero entry per row and column; hence with the substitution in Eq. (105), the $\tilde{T}_k$'s are permutation matrices. We can write the following Hamiltonian acting on $n + 1$ qubits:

$$\tilde{H}_{ZZXX} = -\sum_k \alpha_k \tilde{T}_k \qquad (106)$$

which is a linear combination of permutation matrices with negative coefficients. This makes $\tilde{H}_{ZZXX}$ a 3-local stoquastic Hamiltonian. We can write it as

$$\tilde{H}_{ZZXX} = H_{ZZXX} \otimes |-\rangle\langle -| + \bar{H}_{ZZXX} \otimes |+\rangle\langle +|, \qquad (107)$$

where $\bar{H}_{ZZXX} = -\sum_k \alpha_k |T_k|$ and $|T_k|$ is the entrywise absolute value of T_k. To see why this is the case, first consider a positive element $(T_k)_{ij}$. Then

$$-\alpha_k (T_k)_{ij} \otimes |-\rangle\langle -| - \alpha_k (T_k)_{ij} \otimes |+\rangle\langle +| = -\alpha_k (T_k)_{ij} \otimes \mathbb{1}$$

corresponding to the first replacement in Eq. (105). For a negative element, we have

$$-\alpha_k (T_k)_{ij} \otimes |-\rangle\langle -| + \alpha_k (T_k)_{ij} \otimes |+\rangle\langle +| = -\alpha_k (T_k)_{ij} \otimes \sigma^x$$

corresponding to the second replacement in Eq. (105). The spectrum of $\tilde{H}_{ZZXX}$ separates into two sectors $\mathcal{L}_\pm$. The sector $\mathcal{L}_-$ is spanned by $|\varepsilon_j\rangle \otimes |-\rangle$, where $|\varepsilon_j\rangle$ are the eigenstates of H_{ZZXX}, while the sector $\mathcal{L}_+$ is spanned by $|\bar{\varepsilon}_j\rangle \otimes |+\rangle$, where $|\bar{\varepsilon}_j\rangle$ are the eigenstates of $\bar{H}_{ZZXX}$. Because the Hamiltonian does not couple the two sectors (there are no interactions that take the ancilla qubit from $|\pm\rangle$ to $|\mp\rangle$), a closed-system evolution initialized in the $\mathcal{L}_-$ sector will remain in that sector.

Because the spectrum in the $\mathcal{L}_-$ sector is identical to that of H_{ZZXX}, which is capable of universal adiabatic quantum computation, universal adiabatic quantum computation can be performed in the $\mathcal{L}_-$ sector. However, the lowest energy state in $\mathcal{L}_-$ may not necessarily be the ground state of $\tilde{H}_{ZZXX}$. Therefore, this establishes universal AQC using a stoquastic Hamiltonian only if we do not restrict ourselves to the ground state of the Hamiltonian. Attempting to make the lowest energy state in $\mathcal{L}_-$ be the ground state requires introducing a sufficiently large term proportional to $\mathbb{1} \otimes |+\rangle\langle +|$ to the Hamiltonian $\tilde{H}_{ZZXX}$, but such a term would make the new Hamiltonian nonstoquastic since it would introduce positive off-diagonal elements. Therefore this method does not establish universal adiabatic quantum computation using the ground state of a stoquastic Hamiltonian.

D. Examples of slowdown by StoqAQC

It should not come as a surprise that AQC with an arbitrary final Hamiltonian, which is essentially a black-box approach, does not guarantee quantum speedups. It can be vulnerable to the same sorts of locality traps confronted by heuristic classical algorithms such as simulated annealing.

A slowdown, or failure of AQC to provide a speedup, is a scenario wherein a more efficient classical algorithm is known. All the known examples that fall into this category arise when the gap closes "too fast" in the problem size. However, it is important to note that adiabatic theorems provide only upper bounds on run time, not lower bounds. Thus, an exponentially small gap does not strictly imply an exponentially long run time. The inverse gap is often treated as a proxy for run time, but a claim such as an equal scaling of the inverse gap and the run time does not hold as a general theorem.

With this caveat in mind, in this section we review such "small-gap" examples in increasing order of generality or difficulty of analysis, which all arise in the StoqAQC context. However, we must first note another important caveat. Namely, in some of the examples we present numerical evidence that is based, in necessity, on finite size calculations. One is then often tempted to extrapolate such evidence to the asymptotic scaling. Of course, any such extrapolations based purely on numerics are conjectures. For example, a claim of exponential scaling can never be proven based on numerics alone, as any finite set of data points can always be perfectly fit by a polynomial of sufficiently high degree. Nevertheless, numerics-driven conjectures about scaling can be quite useful, especially if supported by other, analytical arguments.

Subsequently, we will see in Sec. VII that there are various methods for circumventing slowdowns, e.g., via the introduction of nonstoquastic terms.

1. Perturbed Hamming weight problems with exponentially small overlaps

The plain Hamming weight problem is described by

$$H_{\mathrm{HW}}(s) = (1-s)\frac{1}{2}\sum_i (1-\sigma_i^x) + s\sum_x |x||x\rangle\langle x|. \quad (108)$$

Its cost function is simply the Hamming weight $|x|$ of the binary bit string x, which is trivially minimized at $x = 0^n$. Consider the following perturbation of the plain Hamming weight problem (van Dam, Mosca, and Vazirani, 2001):

$$h(x) = \begin{cases} |x| & \text{if } |x| < n, \\ -1 & \text{if } |x| = n. \end{cases} \quad (109)$$

This is a toy problem that is designed to be hard for classical algorithms based on local search: its global optimum lies in a narrow basin, while there is a local optimum with a much larger basin. An algorithm such as simulated annealing with single spin updates would require exponential time to find the global minimum.

Let us write the corresponding StoqAQC Hamiltonian to make the perturbation explicit

$$H(s) = H_{\mathrm{HW}}(s) - s(n+1)|1^n\rangle\langle 1^n|, \quad (110)$$

where $|1^n\rangle$ is the all-one state. Denote the instantaneous eigenstates of $H_{\mathrm{HW}}(s)$ by $\{|v_i(s)\rangle\}$ (v_0 denotes the ground state). Note that the overlap of the all-one state with the instantaneous ground state of the plain Hamming weight algorithm is always exponentially small:

$$\langle 1^n|v_0(s)\rangle \le \frac{1}{\sqrt{2^n}}. \quad (111)$$

We show that this fact causes the adiabatic algorithm as defined in Eq. (110) to take exponential time because it leads to an exponentially small gap.

Define a matrix $A(s)$ with elements

$$A_{ij} = \langle v_i(s)|H(s)|v_j(s)\rangle. \quad (112)$$

Note that $A(0)$ is diagonal and $A_{00}(0) = 0$, equal to the ground state eigenvalue. Also $A(1)$ is diagonal, but now $A_{2^n-1,2^n-1}(1) = -1$ is equal to the ground state eigenvalue. Define a matrix B in the same basis as

$$B_{ij}(s) = \begin{cases} A_{00}(s) & i = j = 0, \\ 0 & i = 0, j > 0, \\ 0 & i > 0, j = 0, \\ A_{ij}(s) & \text{otherwise.} \end{cases} \quad (113)$$

The matrix B always has A_{00} as an eigenvalue. By construction, we know that at $s = 1$ the matrix B has -1 as its ground state eigenvalue (located in the $2^{n-1} \times 2^{n-1}$ submatrix). Because the matrix transforms continuously between these two extremes, there cannot be a jump in the ground state eigenvalue, so there must be a critical value of s, which we denote by s_c, where B has a vanishing gap.

The optimal matching distance between A and B expresses how close their eigenvalue spectra are

$$d(A, B) = \min_\pi \max_{1 \le j \le 2^n} |\lambda_j - \mu_{\pi(j)}|, \quad (114)$$

where π denotes a permutation. Since A and B are Hermitian, this is upper bounded by $\|A - B\|_2$ (Bhatia, 1997). The matrix $A - B$ only has nonzero entries $(A-B)_{0,j>0} = A_{0,j>0}$ and $(A-B)_{j>0,0} = A_{j>0,0} = A_{0,j>0}^*$, with $A_{0,j>0} = -s(n+1)\langle v_1(s)|1^n\rangle\langle 1^n|v_j(s)\rangle$. Therefore,

$$\|A - B\|_2$$
$$= s(n+1)|\langle v_1(s)|1^n\rangle|\sqrt{\sum_{j=1}^{2^n-1} \langle 1^n|v_j(s)\rangle\langle v_j(s)|1^n\rangle}$$
$$= s(n+1)|\langle v_1(s)|1^n\rangle|\sqrt{1 - |\langle 1^n|v_0(s)\rangle|^2}$$
$$\le s(n+1)|\langle v_1(s)|1^n\rangle| \le \frac{s(n+1)}{\sqrt{2^n}}. \quad (115)$$

Thus, the gap of A [and hence of $H(s)$] is always upper bounded by the gap of B plus twice $\|A - B\|_2$. Since at $s = s_c$ the gap of B is zero, it follows that the gap of $H(s_c)$ is $\le s_c(n+1)/\sqrt{2^{n-2}}$ (van Dam, Mosca, and Vazirani, 2001). Therefore, the exponentially small overlap between the unperturbed instantaneous ground state and the perturbed final ground state results in the adiabatic algorithm requiring exponential time to reach the final ground state. Informally, this can also be viewed as the inability of local quantum search

(fluctuations induced by the local initial Hamiltonian) to explore the entire (nonlocal) energy landscape effectively.

2. 2-SAT on a ring

In this section we review the "2-SAT on a ring" problem introduced in the seminal work of Farhi *et al.* (2000), which launched the field of AQC. This example is instructive because of its use of the Jordan-Wigner and Fourier transformation techniques and is also of historical interest. It also serves to illustrate that even a polynomially small gap does not guarantee a quantum speedup. We thus review it in detail.

Consider an n-bit SAT problem with n clauses. Each clause acts only on adjacent bits, i.e., the clause C_j acts only on bits j and $j+1$, where we identify bit $n+1$ with bit 1. Let each clause be of only two forms: "agree" clauses where 00 and 11 are satisfying assignments, and "disagree" clauses where 01 and 10 are satisfying assignments. Since an odd number of satisfied disagree clauses means that the first bit of the first disagree clause is the opposite of the second bit of the last disagree clause, yet bits 1 and $n+1$ must agree, there must be an even number of disagree clauses in order for a satisfying assignment to exist. The classical computational cost of finding a satisfying assignment is at most n: given the list of clauses, a satisfying assignment is found (assuming an even number of disagree clauses) simply by going around the ring and satisfying each clause one at a time. Note that if $\{w_i\}_{i=1}^n$ is a satisfying assignment then so is $\{\neg w_i\}_{i=1}^n$.

Let us now define the final Hamiltonian $H_1 = \sum_{i=1}^n H_{C_i}$ associated with the SAT problem, where each clause is represented by

$$H_{C_i} = \tfrac{1}{2}[1 - (-1)^{x_i}\sigma_i^z \sigma_{i+1}^z],$$
$$x_i = 0 \ (1) \quad \text{if } C_i \text{ is an agree (disagree) clause.} \quad (116)$$

The ground states of H_P are then given by $|0\rangle_1 \otimes_{i=2}^n |w_i\rangle_i$ and $\otimes_{i=1}^n |\neg w_i\rangle_i$, where $w_i = \bigoplus_{j=1}^{i-1} x_j$ ($i \geq 2$ and addition modulo 2). It is possible to gauge away all the disagree clauses. To see this, let U be the unitary transformation defined such that

$$U|z_i\rangle = \begin{cases} |\neg z_i\rangle & \text{if } w_i = 1, \\ |z_i\rangle & \text{if } w_j = 0. \end{cases} \quad (117)$$

Under this unitary transformation we have

$$H_1' = UH_1 U^\dagger = \sum_i \frac{1}{2}(1_i - \sigma_i^z \sigma_{i+1}^z), \quad (118)$$

i.e., the new final Hamiltonian is a sum of just agree clauses. Note that this unitary transformation requires us to know the ground state, but H_1' and H_1 are isospectral, so we can use it for convenience in our gap analysis. The adiabatic computation procedure will be governed by the following time-dependent Hamiltonian:

$$H(s) = (1-s)H_0 + sH_1', \quad 0 \leq s \leq 1, \quad (119)$$

with the initial Hamiltonian $H_0 = \sum_i 1_i - \sigma_i^x$. We diagonalize $H(s)$ in order to find its ground state gap. First, define the negation operator $G = \prod_{i=1}^n \sigma_i^x$ such that

$$G(\otimes_{i=1}^n |z_i\rangle) = \otimes_{i=1}^n |\neg z_i\rangle, \quad (120)$$

which commutes with $H(s)$. The uniform superposition state, which is the ground state of $H(0)$, is invariant under G, i.e., it has eigenvalue $+1$ under G. Therefore, the unitary dynamics will keep the state in the sector with $G = +1$ if it starts in the ground state of $H(0)$. Let us then write $H(s)$ purely in the $G = +1$ sector. Second, define the Jordan-Wigner transformation

$$b_j = \sigma_1^x \sigma_2^x \cdots \sigma_{j-1}^x \sigma_j^-, \quad (121a)$$

$$b_j^\dagger = \sigma_1^x \sigma_2^x \cdots \sigma_{j-1}^x \sigma_j^+, \quad (121b)$$

where $\sigma_j^\pm = (1/2)(\sigma_j^z \pm i\sigma_j^y)$. These are fermionic operators that satisfy

$$\{b_j, b_k\} = 0 \quad (\text{amounts to } \{\sigma_j^-, \sigma_k^-\} = 0), \quad (122a)$$

$$\{b_j, b_k^\dagger\} = \delta_{jk} \quad (\text{amounts to } \{\sigma_j^-, \sigma_k^+\} = \delta_{jk}). \quad (122b)$$

Note that

$$b_j^\dagger b_j = \tfrac{1}{2}(1_j - \sigma_j^x), \quad j = 1, \ldots, n, \quad (123a)$$

$$(b_j^\dagger - b_j)(b_{j+1}^\dagger + b_{j+1}) = \sigma_j^z \sigma_{j+1}^z, \quad j = 1, \ldots, n-1, \quad (123b)$$

$$(b_n^\dagger - b_n)(b_1^\dagger + b_1) = -G\sigma_n^z \sigma_1^z. \quad (123c)$$

In order to make Eqs. (123b) and (123c) consistent in the $G = +1$ sector, we take $b_{n+1} \equiv -b_1$. Using this, we have

$$H(s)|_{G=+1} = \sum_{j=1}^n \left[2(1-s)b_j^\dagger b_j + \frac{s}{2}[1_j - (b_j^\dagger - b_j)(b_{j+1}^\dagger + b_{j+1})] \right]. \quad (124)$$

Third, since this Hamiltonian is invariant under translations $j \mapsto j+1$, define Fourier operators β_p,

$$\beta_p = \frac{1}{\sqrt{n}} \sum_{j=1}^n e^{i\pi p j/n} b_j, \quad p = \pm 1, \pm 3, \ldots, \pm(n-1), \quad (125)$$

where for simplicity it is assumed that n is even. Equivalently,

$$b_j = \frac{1}{\sqrt{n}} \sum_{p=\pm 1,\ldots} e^{-i\pi p j/n} \beta_p, \quad (126)$$

where we used the fact that $\sum_{p=\pm 1,\ldots} e^{i\pi p(j-j')/n} = n\delta_{j,j'}$. Furthermore, note that

$$\{\beta_{2a-1},\beta_{2b-1}\} = \frac{1}{n}\sum_{j,j'}e^{i\pi((2a-1)j+(2b-1)j')/n}\{b_j,b_{j'}\}$$

$$= 0, \tag{127a}$$

$$\{\beta_{2a-1},\beta_{2b-1}^\dagger\} = \frac{1}{n}\sum_{j,j'}e^{i\pi((2a-1)j-(2b-1)j')/n}\{b_j,b_{j'}^\dagger\}$$

$$= \frac{1}{n}\sum_{j=1}^{n}e^{2\pi i(a-b)/n} = \delta_{a,b}, \tag{127b}$$

so the set $\{\beta_p\}$ comprises valid fermionic operators. Writing the Hamiltonian in terms of this set, we have

$$H(s) = \sum_{p=1,3,\dots}\left\{2(1-s)(\beta_p^\dagger\beta_p + \beta_{-p}^\dagger\beta_{-p})\right.$$

$$+ s\left[\mathbb{1} - \cos\left(\frac{\pi p}{n}\right)(\beta_p^\dagger\beta_p - \beta_{-p}\beta_{-p}^\dagger)\right.$$

$$\left.\left.+ i\sin\left(\frac{\pi p}{n}\right)(\beta_{-p}^\dagger\beta_p^\dagger - \beta_p\beta_{-p})\right]\right\} \tag{128a}$$

$$\equiv \sum_{p=1,3,\dots}A_p(s). \tag{128b}$$

Now that $H(s)$ has finally been written as a sum of commuting operators ($[A_p, A_{p'}] = 0$ for $p \neq p'$), we can diagonalize each summand separately. For a given p, let us denote by $|\Omega_p\rangle$ the state that is annihilated by β_p and β_{-p}, i.e., $\beta_p|\Omega_p\rangle = \beta_{-p}|\Omega_p\rangle = 0$. Note that $A_p(s=0)|\Omega_p\rangle = 0$, so $|\Omega_p\rangle$ is the ground state of A_p at $s = 0$ (recall that we already knew that the ground state energy at $s = 0$ was zero). Let $|\Sigma_p\rangle = \beta_{-p}^\dagger\beta_p^\dagger|\Omega_p\rangle$. $A_p(s)$ keeps states in the subspace spanned by $|\Omega_p\rangle$ and $|\Sigma_p\rangle$ in the same subspace; the initial state is in this subspace, so we can restrict our attention to it. Let us write $A_p(s)$ in the $\{|\Omega_p\rangle, |\Sigma_p\rangle\}$ basis,

$$A_p(s) = \begin{pmatrix} s + s\cos\left(\frac{\pi p}{n}\right) & is\sin\left(\frac{\pi p}{n}\right) \\ -is\sin\left(\frac{\pi p}{n}\right) & 4 - 3s - s\cos\left(\frac{\pi p}{n}\right) \end{pmatrix}. \tag{129}$$

Diagonalizing this, we find for the energies,

$$E_p^\pm(s) = 2 - s \pm \left\{(2-3s)^2 + 4s(1-s)\left[1 - \cos\left(\frac{\pi p}{n}\right)\right]\right\}^{1/2}. \tag{130}$$

The instantaneous ground state energy of $H(s)$ is thus given by $\sum_{p=1,3,\dots}E_p^-(s)$. The first excited state energy is given by $E_1^+(s) + \sum_{p=3,\dots}E_p^-(s)$. The energy gap $\Delta(s)$ is therefore given by

$$\Delta(s) = E_1^+(s) - E_1^-(s)$$

$$= 2\left\{(2-3s)^2 + 4s(1-s)\left[1 - \cos\left(\frac{\pi p}{n}\right)\right]\right\}^{1/2}. \tag{131}$$

The minimum occurs at

$$s^* = \frac{2(2 + \cos\pi/n)}{5 + 4\cos\pi/n} \to 2/3$$

as $n \to \infty$. Therefore, the minimum gap is given by

$$\Delta(s^*) = 4\left|\sin\frac{\pi}{n}\right|\frac{1}{\sqrt{5 + 4\cos\pi/n}} \to \frac{4\pi}{3n}, \tag{132}$$

which implies a polynomial run time for the adiabatic algorithm. As mentioned, the classical computational cost of finding a satisfying assignment is at most n. Therefore, despite the polynomially small gap in this example, there is no quantum speedup. This illustrates that a StoqAQC slowdown need not necessarily be associated with an exponentially small gap.

3. Weighted 2-SAT on a chain with periodicity

We now discuss another problem, proposed by Reichardt (2004), that combines 2-SAT with an exponential slowdown of StoqAQC. It can thus be viewed as exhibiting aspects of the two previous problems we discussed.

Consider a weighted 2-SAT problem on a chain with agree clauses between bits $i, i + 1$ for $i = 1, \dots, N - 1$ with weights

$$J_i = \begin{cases} w & \text{if } \lceil\frac{i}{n}\rceil \text{ is odd}, \\ 1 & \text{if } \lceil\frac{i}{n}\rceil \text{ is even}, \end{cases} \tag{133}$$

where n is the period and $w > 1$. As for the previous 2-SAT problem, we can map this to a spin chain with ferromagnetic couplings with strength given by J_i. The adiabatic Hamiltonian is given by

$$H(s) = -(1-s)\sum_{i=1}^{N}\sigma_i^x - s\sum_{i=1}^{N-1}J_i\sigma_i^z\sigma_{i+1}^z. \tag{134}$$

This chain has coefficients that alternate between w and 1 in sectors of size n each, with the $b + 1$ odd-numbered sectors being "heavy" ($J_i = w > 1$), b even-numbered sectors being "light" ($J_i = 1$), and where the total number of sectors is $(N - 1)/n = 2b + 1$. Since the chain is ferromagnetic, the ground state of $H(1)$ is trivially the all-0 or all-1 computational-basis state. The problem is thus classically easy and can be solved by inspection or in time $O(N)$ by a heuristic classical algorithm such as simulated annealing, by simply traversing the chain and updating one spin at a time.

Note that at $s = 0$ there is a unique ground state, while at $s = 1$ the ground state is doubly degenerate. Therefore, the relevant quantum ground state gap Δ is not the gap to the first excited state (since at the end of the evolution, this merges with the ground state), but to the second excited state.

It turns out this gap is exponentially small in the sector size n across a constant range $s \in (1/(1+w), 1/2)$. Moreover, there are exponentially many (in $\sqrt{N}$) exponentially small excitations above the ground state for $n \sim \sqrt{N}$.

More precisely, let $\mu_w = sw/(1-s)$. Theorem 4 in Reichardt (2004) states the following:

(1) For any fixed $s > 1/(1+w)$, i.e., $\mu_w < 1$, $H(s)$ has one eigenvalue only $O(\mu_w^n)$ above the ground state energy. This means that the gap Δ is exponentially decreasing with the sector size n.

(2) For $s \in (1/(1+w), 1/(1+\sqrt{w})]$ (i.e., again $\mu_w < 1$), $H(s)$ has $2^{b+1} - 1$ eigenvalues only $O(b\mu_w^n)$ above the ground state energy. This means that there are exponentially many (in the number of odd sectors $b + 1$) excited states that likewise have an exponentially small (in n) gap from the ground state. Note that $b = [(N-1)/n + 1]/2$, so $b \sim \sqrt{N}$ when $n \sim \sqrt{N}$.

(3) For $s \in [1/(1+\sqrt{w}), 1/2)$, where $\mu_1 = s/(1-s) > 1$, $H(s)$ has $2^{b+1} - 1$ eigenvalues $O(b\mu_1^{-n})$ above the ground state energy. This again means an exponentially large number (in b) of excited states with an exponentially small (in n) gap.

The proof uses a Jordan-Wigner transformation to diagonalize $H(s)$, similarly to the technique in Sec. VI.D.2. The spectral gaps of the $2^N \times 2^N$ matrix $H(s)$ are the square roots of the eigenvalues of an $N \times N$ symmetric, tridiagonal matrix. The Sturm sequence of the principal leading minors of this matrix is then analyzed to bound the eigenvalue gaps of $H(s)$.

Why might we expect this problem to be hard for the adiabatic algorithm? Within any given light or heavy sector, the problem (at fixed s) is that of a uniform transverse field Ising chain. Consider the thermodynamic limit $N \gg 1$, and also let $n \gg 1$. In this limit the transverse field Ising chain encounters a phase transition separating the disordered phase and the ordered phase when $1 - s = sJ_i$ (Sachdev, 2001), i.e., at $s = 1/(1 + J_i)$. [The boundary of the chain adds only $O(1)$ energy, so it does not impact this intuitive argument in the thermodynamic limit.] This means that the heavy sectors encounter the phase transition at $s = 1/(1 + w)$ whereas the light sectors encounter the phase transition at $s = 1/2$, i.e., the light sectors order after the heavy ones. At $s = 1/(1 + w)$ each heavy sector orders in either the all-0 or all-1 state, and different heavy sectors are separated by light sectors that have not ordered yet. Since the initial Hamiltonian generates only local spin flips, the algorithm is likely to get stuck in a local minimum with a domain wall in one or more disordered sectors, if run for less than exponential time in n. This mechanism in which large local regions order before the whole is well known in disordered, geometrically local optimization problems, gives rise to a Griffiths phase (Fisher, 1995).

4. Topological slowdown in a dimer model or local Ising ladder

Another interesting example of a local spin model that leads StoqAQC astray was given by Laumann et al. (2012). They showed that a translation-invariant quasi-1D transverse field Ising model with nearest-neighbor interactions only, the ground state of which is readily found by inspection, results in exponentially long run times for StoqAQC. The model can be understood as either a dimer model on a two-leg ladder of even length L or, using a duality transformation, a two-leg frustrated Ising ladder of the same length in a uniform magnetic field, the ground states of which map onto the dimer states. The frustrated Ising ladder Hamiltonian is

$$H_1 = -\sum_{\langle i,j \rangle} J_{ij}\sigma_i^z\sigma_j^z - K\sum_{i \in \text{upper}} \sigma_i^z + \frac{1}{2}U\sum_{i \in \text{lower}} \sigma_i^z, \quad (135)$$

where upper and lower refer to the legs of the ladder, $J_{ij} = -K$ for the upper-leg couplings and $J_{ij} = K$ for all other (lower-leg and rungs) couplings.

Quantum dynamics is introduced via a standard initial Hamiltonian $-\Gamma(t)\sum_i \sigma_i^x$, where $\Gamma(0) \gg \|H_1\|$ and $\Gamma(t_f) = 0$. The dimer model exhibits a first-order quantum phase transition with an exponentially small gap when $K \gg U$, which is inherited by the frustrated Ising ladder model. Namely, the system prefers the sector with exponentially many ground states, while any degeneracy-lifting interaction favors another containing only $O(1)$ states. StoqAQC selects the wrong sector, tunneling out of which becomes exponentially slow as Γ is reduced.

More specifically, for $K \gg U$, the Hilbert space is spanned by an orthonormal basis of hard core dimer coverings ("perfect matchings") of the ladder. These fall into three sectors which are topological in that they are not connected by any local rearrangement of the dimers. The sectors are labeled by a winding number w, the difference between the number of dimers on the top and bottom rows (on any fixed plaquette). The model assigns extensive energy $\propto L$ to every state in the $w = 0$ sector while leaving the two staggered states $w = \pm 1$ as ground states with energy 0. On the other hand, at large Γ (strong transverse field) the $w = 0$ sector is favored. Intuitively, slowly turning the transverse field off by reducing Γ does not help change the topological sector since any off-diagonal term in the dimer Hilbert space involving only a finite number of rungs in the ladder leaves the winding number w invariant. This is depicted in Fig. 5. Numerical analysis of the Ising ladder confirms this picture by revealing that the gap is exponentially small in L when $K > U$. The critical point is found to be at $\Gamma_c \approx U/b + U^2/4Kb^3$, where $b \approx 0.6$ (from exact diagonalization numerics).
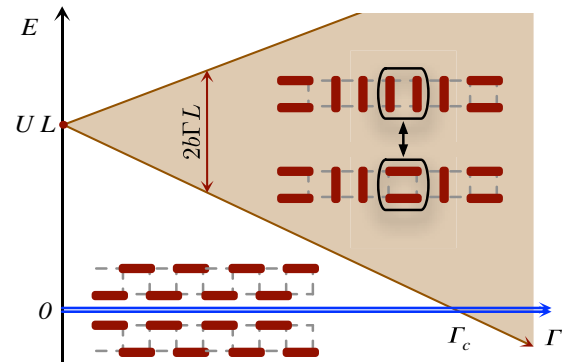

FIG. 5. Energy spectrum of the dimer model on an even length periodic ladder, with the dimer configurations illustrated. The $w = \pm 1$ states are at energy $E = 0$, while the $w = 0$ sector splits into a band for $\Gamma > 0$. For sufficiently large Γ, the $w = 0$ sector contains the ground state of the system. An unavoided level crossing (first-order quantum phase transition) occurs at $\Gamma = \Gamma_c$, which is responsible for the quantum slowdown. From Laumann et al., 2012.

5. Ferromagnetic mean-field models

The quantum ferromagnetic p-spin model is given by

$$H = -\frac{1}{n^{p-1}}\left(\sum_{i=1}^{n}\sigma_i^z\right)^p - \Gamma\sum_{i=1}^{n}\sigma_i^x. \qquad (136)$$

By inspection it is clear that when p is even, the ground state at $\Gamma = 0$ is either of the two fully aligned ferromagnetic states, while when p is odd, the unique ground state at $\Gamma = 0$ is the fully aligned spin-down state. As Γ is tuned from a large value toward zero, the system encounters a first-order phase transition for $p > 2$. This can be readily shown by employing the Suzuki-Trotter decomposition and the static approximation (Chayes *et al.*, 2008; Krzakala *et al.*, 2008; Jörg, Krzakala, Kurchan *et al.*, 2010; Suzuki, Inoue, and Chakrabarti, 2013) to calculate the partition function Z in the large n limit, where

$$Z = \int dm e^{-\beta n F(\beta,\Gamma,m)}. \qquad (137)$$

Here β is the inverse temperature, m is the Hubbard-Stratonovich field (Hubbard, 1959), and F is the free energy density given by

$$F = (p-1)m^p - \frac{1}{\beta}\log\left[2\cosh\left(\beta\sqrt{\Gamma^2 + p^2 m^{2p-2}}\right)\right]. \qquad (138)$$

The dominant contribution to F comes from the saddle point of the partition function Z, which provides consistency equations for the field m:

$$m = pm^{p-1}\frac{\tanh\left(\beta\sqrt{\Gamma^2 + p^2 m^{2(p-1)}}\right)}{\sqrt{\Gamma^2 + p^2 m^{2(p-1)}}}. \qquad (139)$$

Solving Eq. (139) numerically for $p > 2$ reveals a discontinuity in the value of m that minimizes the free energy as Γ is tuned through the phase transition point. At this critical point, the free energy exhibits a degenerate double-well potential, and an instantonic calculation on this potential gives an exponentially small energy gap with system size (Jörg, Krzakala, Kurchan *et al.*, 2010).

6. 3-Regular 3-XORSAT

All the problems discussed so far were amenable to a classical solution "by inspection" (i.e., the solution is obvious from the form of the cost function). Some problems were even easy for classical heuristic algorithms performing a local search. We now discuss a problem that is nontrivial in this respect, i.e., classically yields only in polynomial time to a tailored approach.

In 3-XORSAT, each clause involves 3 bits, and there are M clauses and n bits in total. A clause is satisfied if the sum of the 3 bits (mod 2) is a specified value; it can be 0 or 1 depending on the clause. For 3-regular 3-XORSAT, every bit is in exactly three clauses and $M = n$. This problem is associated with a spin-glass phase but is "glassy without being hard to solve" (Franz *et al.*, 2001; Ricci-Tersenghi,

2010): the problem of finding a satisfying assignment can be solved in polynomial time using Gaussian elimination because the problem involves only linear constraints (mod 2) (Haanpaa *et al.*, 2006).

A final Hamiltonian involving n spins can be written such that each satisfied clause gives energy 0 and each unsatisfied clause gives energy 1

$$H_1 = \sum_{c=1}^{n}\left(\frac{\mathbb{1} - J_c\sigma_{i_1,c}^z\sigma_{i_2,c}^z\sigma_{i_3,c}^z}{2}\right). \qquad (140)$$

Here the index (i_k, c) denotes the 3 bits associated with clause c, and $J_c \in \{\pm 1\}$ depending on whether the clause is satisfied if the sum of its bits (mod 2) is 0 or 1. The StoqAQC Hamiltonian is then given as usual by $H(s) = -(1-s)\sum_{i=1}^{n}\sigma_i^x + sH_1$. The median minimum gap for random 3-regular 3-XORSAT has numerically been shown to be exponentially small in the system size up to $n = 24$ (Jörg, Krzakala, Semerjian, and Zamponi, 2010) and $n = 40$ (Farhi *et al.*, 2012) [using both the quantum cavity method (Krzakala *et al.*, 2008; Laumann, Scardicchio, and Sondhi, 2008) and QMC simulations], with a first-order quantum phase transition at $s = 1/2$. Thus, the numerical evidence suggests that StoqAQC takes exponential time to solve this problem. The same is true for classical heuristic local-search algorithms such as WALKSAT (Guidetti and Young, 2011).

Note that since the Hamiltonian gap is not a thermodynamic quantity, one must be careful not to automatically associate a first-order quantum phase transition with an exponentially small gap. While the examples presented in this review agree with this rule [for additional examples, see Dusuel and Vidal (2005), Knysh and Smelyanskiy (2006), Jörg *et al.* (2008), Jörg, Krzakala, Kurchan *et al.* (2010), and Bapst and Semerjian (2012)], counterexamples wherein a first-order quantum phase transition is associated with a polynomially small gap are known (Cabrera and Jullien, 1987; Laumann *et al.*, 2012, 2015; Tsuda, Yamanaka, and Nishimori, 2013).

7. Sherrington-Kirkpatrick and two-pattern Gaussian Hopfield models

The Sherrington-Kirkpatrick model, the prototypical spin-glass model, is NP hard, yet its quantum transverse field Ising model version (Ishii and Yamamoto, 1985; Usadel, 1986; Ray, Chakrabarti, and Chakrabarti, 1989; Das, Chakrabarti, and Stinchcombe, 2005) exhibits a second-order phase transition separating the paramagnetic phase from the spin-glass phase (Miller and Huse, 1993; Ye, Sachdev, and Read, 1993). The model is defined via the final Hamiltonian

$$H_1 = \sum_{i_1 < i_2} J_{i_1 i_2}\sigma_{i_1}^z\sigma_{i_2}^z, \qquad (141)$$

where the couplings $J_{i_1 i_2}$ are zero mean, independent and identically distributed random variables (e.g., Gaussian or bimodal, i.e., $J_{i_1 i_2} = \pm 1$) and every spin is coupled to every other spin. The adiabatic computation proceeds via

$$H(t) = H_1 - \Gamma(t) \sum_{i=1}^{n} \sigma_i^x, \tag{142}$$

where Γ is adiabatically reduced to zero.

The polynomial closing of the gap at this phase transition appears promising for AQC. However, a spin glass is dominated by a rough free energy landscape with many local minima forming bottlenecks for classical heuristic local-search algorithms (Mezard, Parisi, and Virasoro, 1987; Nishimori, 2001).

To gain insight into this phenomenon, and, in particular, its impact on StoqAQC, Knysh (2016) studied another fully connected model with a vanishing classical gap: the Gaussian Hopfield model, generally defined via

$$J_{i_1 i_2 \cdots i_p} = \frac{1}{n^{p-1}} \sum_{\mu=1}^{r} \xi_{i_1}^{(\mu)} \cdots \xi_{i_p}^{(\mu)} \tag{143}$$

(the Hebb rule), where $\xi_i^{(\mu)}$ are zero-mean independent and identically distributed random variables of unit variance. By focusing on the analytically more tractable Hopfield model, Knysh (2016) rigorously analyzed for $r = 2$ (the two-pattern case) and $p = 2$ (two-local interactions) the properties of local minima away from the global minimum.

The main insight gained from the theoretical analysis of Knysh (2016) is that the complexity of the model is not determined by the phase transition, but rather by the existence of small-gap bottlenecks in the spin-glass phase. Namely, after the occurrence of the polynomially closing gap associated with the second-order phase transition separating the paramagnetic and glass phases, there are $O(\log n)$ additional gap minima in the spin-glass phase appearing in an approximate geometric progression, a phenomenon that can be attributed to the self-similar properties of the free energy landscape in a Γ interval bounded by the appearance of the spin-glass phase. At these bottlenecks, the gaps scale as a stretched exponential $e^{-c\Gamma_m^{3/4} n^{3/4}}$, where Γ_m is the location of the mth minimum. This is illustrated in Fig. 6. Nevertheless this means that StoqAQC suffers a (stretched) exponential slowdown, since the two-pattern Gaussian Hopfield model admits an efficient classical solution based on angle sorting and an exhaustive search that scales as $O[n \log(n)]$ (Knysh, 2016). Thus, this is another case where a StoqAQC algorithm is too generic to exploit problem structure, and consequently a tailored classical algorithm has exponentially better scaling.

E. StoqAQC algorithms with a scaling advantage over simulated annealing

A substantial effort is underway to develop problems that may exhibit any form of a quantum speedup (recall the classification given in Sec. III). One approach has been to develop "tunneling gadgets," i.e., small toy Hamiltonians that exhibit tunneling (Boixo et al., 2016), and use these gadgets to construct larger problems (Denchev et al., 2016). An alternative approach has been to develop instances that are believed to exhibit "small-and-thin" energy barriers in their classical energy landscape (Katzgraber et al., 2015) in the
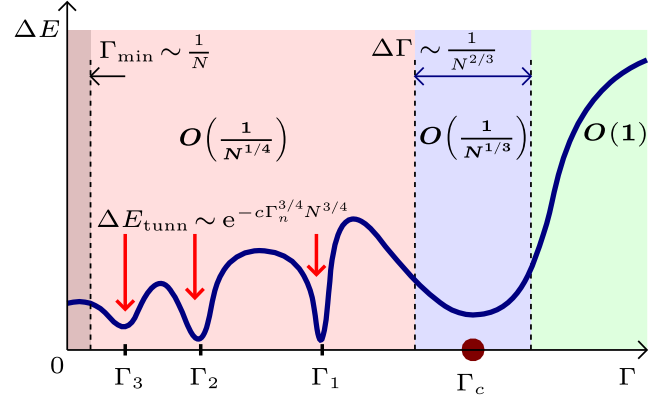

FIG. 6. The gap behavior in the $r = 2$, $p = 2$ Gaussian Hopfield model. The paramagetic-spin-glass transition occurs at Γ_c, with $\Gamma < \Gamma_c$ denoting the spin-glass phase. The typical gap is denoted using the big-O notation. The spin-glass phase contains $\log n$ additional minima in the gap (indicated by red arrows). $\Gamma_{\min}$ corresponds to the lowest energy scale of the classical Hamiltonian, which in this case scales as $1/N$, where N was used to represent the variable we denote by n. From Knysh, 2016.

hope that such barriers persist in the quantum energy landscape where tunneling occurs. These approaches have been used primarily to assess the performance of the D-Wave devices and are based on numerical analysis, which makes extrapolation and conclusions about asymptotic scaling rather challenging (Brady and Dam, 2016; Mandrà et al., 2016).

In this section we consider several examples of StoqAQC with a demonstrable quantum scaling advantage over SA. While none of the examples are demonstrations of an unqualified quantum speedup, these examples are illustrative in that they reveal important qualitative differences between SA, where thermal fluctuations are used to explore the energy landscape, and StoqAQC, where quantum fluctuations are used to explore a different energy landscape. Still, these results are based on a comparison with SA that uses only single-spin updates. SA-like algorithms with cluster-spin updates can be significantly more efficient (Swendsen and Wang, 1987; Wolff, 1989; Houdayer, 2001; Zhu, Ochoa, and Katzgraber, 2015; Zintchenko, Hastings, and Troyer, 2015; Mandrà et al., 2016), and their performance relative to StoqAQC is largely an open question. The same is true for parallel tempering (also known as exchange Monte Carlo) (Katzgraber et al., 2006; Swendsen and Wang, 1986; Marinari and Parisi, 1992; Hukushima and Nemoto, 1996; Earl and Deem, 2005).

1. Spikelike perturbed Hamming weight problems

We start with a problem for which there is no quantum speedup, in order to set up the more interesting problems that follow. Consider a cost function $f(x)$ to be minimized with $x \in \{0, 1\}^n$ an n-bit string. The final Hamiltonian can generically be written as

$$H_1 = \sum_x f(x) |x\rangle \langle x|. \tag{144}$$

We first consider the cost function of the "plain" Hamming weight problem

$$f(x) = |x|, \qquad (145)$$

where $|x|$ denotes the Hamming weight of the n-bit string x [as in Eq. (108)]. This problem is equivalent to a system of n noninteracting spins in a global (longitudinal) field, which is of course a trivial problem that can be solved in time $O(1)$, e.g., by parallelized SA running with a a single thread for each spin. The scaling of the time needed by the quantum algorithm is $O(n^{1/2})$, and the full cost of the quantum algorithm is $O(n^{3/2})$ according to Eq. (1), since it requires $O(n)$ single-qubit terms in the Hamiltonian. A fairer comparison is to an SA algorithm that is ignorant of the structure of the problem. In this case one can show that the cost for single-spin update SA with random spin selection is lower bounded by $\mathcal{O}(n \log n)$ (Muthukrishnan, Albash, and Lidar, 2016).

Next we consider a more interesting problem, referred to as the "spike," first studied by Farhi, Goldstone, and Gutmann (2002a). The cost function is given by

$$f(x) = \begin{cases} n & \text{if } |x| = n/4, \\ |x| & \text{otherwise.} \end{cases} \qquad (146)$$

Since the barrier scales with n, we can expect that single-spin-update SA will take $\exp(n)$ time to traverse the barrier. However, it can be shown that the quantum gap scales as $\Omega(n^{-1/2})$ (Farhi, Goldstone, and Gutmann, 2002a; Kong and Crosson, 2015), so the adiabatic algorithm takes only polynomial time.

This type of "perturbed" Hamming weight problem can be generalized, while still retaining an advantage over single-spin-update SA. For cost functions of the form

$$f(x) = \begin{cases} |x| + h(n) & \text{if } l(n) < |x| < u(n), \\ |x| & \text{otherwise,} \end{cases} \qquad (147)$$

satisfying $h[(u-l)/\sqrt{l}] = o(1)$, the minimum gap of the quantum algorithm is lower bounded by a constant (Reichardt, 2004) [see Appendix A of Muthukrishnan, Albash, and Lidar (2016) for a pedagogical review of the proof]. The SA run time, on the other hand, scales exponentially in $\max_n h(n)$.

Similarly, consider barriers with width proportional to n^{α} and height proportional to n^{β}, i.e.,

$$f(x) = \begin{cases} |x| + n^{\alpha} & \text{if } \frac{n}{4} - \frac{1}{2}n^{\beta} < |x| < \frac{n}{4} + \frac{1}{2}n^{\beta}, \\ |x| & \text{otherwise.} \end{cases} \qquad (148)$$

When α and β satisfy $\alpha + \beta \geq 1/2$, $\alpha < 1/2$, and $2\alpha + \beta \leq 1$, the minimum gap scales polynomially as $n^{1/2-\alpha-\beta}$ (Brady and van Dam, 2016a, 2016b), while the SA run time scales exponentially in n^{α}.

2. Large plateaus

The previous examples have relied on energy barriers in the classical cost that scale with problem size to foil single-spin-update SA. This agrees with the intuition that a StoqAQC advantage over SA is associated with tall and thin barriers (Ray, Chakrabarti, and Chakrabarti, 1989; Das and Chakrabarti, 2008). However, somewhat counterintuitively, it is also possible to foil SA by having very large plateaus in the classical cost function. Specifically, consider

$$f(x) = \begin{cases} u - 1 & \text{if } l < |x| < u, \\ |x| & \text{otherwise,} \end{cases} \qquad (149)$$

where $l, u = \mathcal{O}(1)$ [a special case of Eq. (147)]. SA with single-spin updates and random spin selection has run time $O(n^{u-l-1})$, where $u - l - 1$ is the plateau width (Muthukrishnan, Albash, and Lidar, 2016). This polynomial scaling arises because the energy landscape provides no preferred direction and SA then behaves as a random walker on the plateau. Numerical diagonalization shows that the quantum minimum gap is constant and the adiabatic run time is only $O(n^{1/2})$, where the scaling with n arises from the numerator of the adiabatic condition (Muthukrishnan, Albash, and Lidar, 2016). Thus StoqAQC has a polynomial scaling advantage over SA in this case.

A natural question is whether these potential quantum speedup results for StoqAQC relative to SA survive when other algorithms are considered. The answer is negative. Muthukrishnan, Albash, and Lidar (2016) showed that a diabatic evolution is more efficient than the adiabatic evolution to solve these problems, and a similar efficiency is achieved using classical spin-vector dynamics. There is also a growing body of numerical (Crosson and Deng, 2014; Brady and van Dam, 2016a; Denchev et al., 2016; Muthukrishnan, Albash, and Lidar, 2016) and analytical (Crosson and Harrow, 2016; Isakov et al., 2016) research that shows that quantum Monte Carlo methods exhibit similar or even identical advantages over SA for many spikelike perturbed Hamming weight problems.

F. StoqAQC algorithms with undetermined speedup

In this section we focus on examples where it is currently unknown whether there is a quantum speedup or slowdown for StoqAQC.

1. Number partitioning

The number partitioning problem is a canonical NP-complete problem (Garey and Johnson, 1979) that is defined as follows: given a set of n positive numbers $\{a_i\}_{i=1}^n$, the objective is to find a partition $\mathcal{P}$ of this set that minimizes the partition residue E defined as

$$E = \left| \sum_{j \in \mathcal{P}} a_j - \sum_{j \notin \mathcal{P}} a_j \right|. \qquad (150)$$

The problem exhibits an easy-hard phase transition at the critical value $b/n = 1$, where b is the number of bits used to represent the set $\{a_i\}$ (Mertens, 1998; Borgs, Chayes, and Pittel, 2001). In the hard phase it roughly corresponds to finding the minimum in a set of 2^n numbers (Mertens, 2001). To translate it into Ising spin variables let $s_j = 1$ when $j \in \mathcal{P}$ and $s_j = -1$ otherwise, so that

$$E = \left| \sum_{j=1}^n a_j s_j \right|, \qquad (151)$$

which can then be turned into a Mattis-like Ising Hamiltonian whose ground state is the minimizing partition:

$$H_1 = \sum_{i,j=1}^{n} a_i a_j s_i s_j. \tag{152}$$

The energy landscape of this final Hamiltonian is known to be extremely rugged in the hard phase (Smelyanskiy, Toussaint, and Timucin, 2002; Stadler, Hordijk, and Fontanari, 2003), and the asymptotic behavior can already be seen for small sizes n. While SA effectively requires the searching of all possible bit configurations with a run time $\propto 2^{0.98n}$ (Stadler, Hordijk, and Fontanari, 2003), numerical simulations of StoqAQC exhibit a slightly better run time $\propto 2^{0.8n}$ (Denchev *et al.*, 2016). State-of-the-art classical algorithms have scalings as low as $2^{0.291n}$ (Becker, Coron, and Joux, 2011).

It should be noted that number partitioning is known as the "easiest hard problem" (Hayes, 2002) due to the existence of efficient approximation algorithms that apply in most (although of course not all) cases, e.g., a polynomial-time approximation algorithm known as the differencing method (Karmarkar and Karp, 1982). It should further be noted that if all the a_j's are bounded by a polynomial in n, then integer partitioning can be solved in polynomial time by dynamic programming (Mertens, 2003). The NP hardness of the number partitioning problem requires input numbers of size exponentially large in n or, after division by the maximal input number, of exponentially high precision. This is problematic since the $\{a_j\}$'s are used as coupling coefficients in the adiabatic Hamiltonian (152) and suggests that a different encoding will be needed in order to allow AQC to meaningfully address number partitioning.

2. Exact cover and its generalizations

We review the adiabatic algorithm for exact cover, which initiated and sparked a tremendous interest in the power of AQC when it was first studied by Farhi *et al.* (2001). While the optimistic claim made that "the quantum adiabatic algorithm worked well, providing evidence that quantum computers (if large ones can be built) may be able to outperform ordinary computers on hard sets of instances of NP-complete problems" turned out to be premature, the historical impact of this study was large, and it led to the avalanche of work that forms the core of this review.

The exact cover 3 (EC3) problem is an NP-complete problem that is a particular formulation of 3-SAT (recall Sec. V.A) whereby each clause C (composed of 3 bits $x_{C_1}, x_{C_2}, x_{C_3}$ that are taken from the set of variables $\{x_i \in \{0,1\}\}_{i=1}^{n}$) is satisfied if $x_{C_1} + x_{C_2} + x_{C_3} = 1$. There are only three satisfying assignments: (1,0,0), (0,1,0), and (0,0,1). A 3-local Hamiltonian H_C can be associated with each clause that assigns an energy penalty to the unsatisfying assignments (Farhi *et al.*, 2001; Latorre and Orus, 2004):

$$
\begin{aligned}
H_C = \tfrac{1}{8}[&(1+\sigma_{C_1}^z)(1+\sigma_{C_2}^z)(1+\sigma_{C_3}^z) \\
&+ (1-\sigma_{C_1}^z)(1-\sigma_{C_2}^z)(1-\sigma_{C_3}^z) \\
&+ (1-\sigma_{C_1}^z)(1-\sigma_{C_2}^z)(1+\sigma_{C_3}^z) \\
&+ (1-\sigma_{C_1}^z)(1+\sigma_{C_2}^z)(1-\sigma_{C_3}^z) \\
&+ (1+\sigma_{C_1}^z)(1-\sigma_{C_2}^z)(1-\sigma_{C_3}^z)].
\end{aligned}
\tag{153}
$$

A 2-local alternative is (Young, Knysh, and Smelyanskiy, 2010)

$$H_C = \tfrac{1}{4}(\sigma_{C_1}^z + \sigma_{C_2}^z + \sigma_{C_3}^z - 1)^2. \tag{154}$$

The final Hamiltonian is then given by $H_1 = \sum_C H_C$. If the ground state energy is 0, then an assignment exists that satisfies all clauses. The adiabatic algorithm is given as usual by $H(s) = (1-s)H_0 + sH_1$, with $H_0 = \sum_{i=1}^{n}(1/2)(1-\sigma_i^x)$.

For instances with a unique satisfying assignment, while the initial (small n) scaling of the typical minimum gap (median) is consistent with a polynomial (Farhi *et al.*, 2001; Latorre and Orus, 2004), the true (large n) scaling is exponential and can be associated with a first-order phase transition (Young, Knysh, and Smelyanskiy, 2008, 2010) occurring at intermediate $s = s_c > 0$. The fraction of instances with this behavior increases with increasing problem size (Young, Knysh, and Smelyanskiy, 2010). This illustrates the perils of extrapolating the asymptotic scaling from studies based on small problem sizes.

A natural generalization of the exact cover problem is to have the sum of K variables sum to 1 for a clause to be satisfied, which defines the problem known as "1-in-K SAT." Another is to have the clause satisfied unless all the variables are equal, which defines the problem "K-not-all-equal-SAT." Both of these are NP complete and have been shown analytically to exhibit a first-order phase transition for sufficiently large K (Smelyanskiy, Knysh, and Morris, 2004). Numerical results for locked 1-in-3 SAT and locked 1-in-4 SAT, where "locked' is the additional requirement that every variable is in at least two clauses and that one cannot get from one satisfying assignment to another by flipping a single variable (Zdeborová and Mézard, 2008a, 2008b), have been shown to exhibit an exponentially small gap at the satisfiability transition (Hen and Young, 2011).

Since all these problems are NP complete, there is no polynomial-time classical algorithm known for their worst-case instances. Using StoqAQC has, in all cases that have been studied to date, resulted in exponentially small gaps. Thus, whether these problems can be sped up (even polynomially) is at this time still an open problem.

3. 3-Regular MAXCUT

For 3-regular MAXCUT, the problem is to find the assignment that gives the maximum number of satisfied clauses, where each bit appears in exactly three clauses. Each clause involves only 2 bits and is satisfied if and only if the sum of the 2 bits (mod 2) is 1. The number of clauses is $M = 3n/2$. The final Hamiltonian can be written as

$$H_1 = \sum_{c=1}^{3n/2} \left(\frac{\mathbb{1} + \sigma_{i_1,c}^z \sigma_{i_2,c}^z}{2} \right), \tag{155}$$

where the index (i_k, c) denotes the 2 bits associated with clause c. This model can also be viewed as an antiferromagnet on a 3-regular random graph. Because the random graph in general has loops of odd length, it is not possible to satisfy all of the clauses. This problem is NP hard.
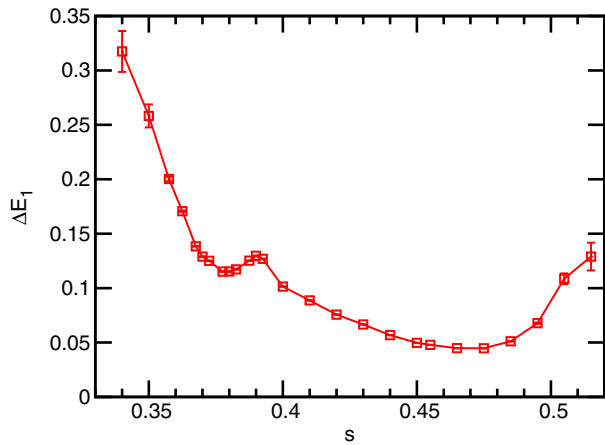
FIG. 7. The gap to the first even excited state for an instance of size $n = 128$ exhibiting two minima. The lower minimum occurs well within the spin-glass phase, while the higher minimum is associated with the second-order phase transition. From Farhi *et al.*, 2012.

For random instances of this problem, where there is a doubly degenerate ground state (the smallest possible because of the Z_2 symmetry) and with a specified energy of $n/8$, the standard adiabatic Hamiltonian $H(s) = -(1-s)\sum_{i=1}^{n}\sigma_i^x + sH_1$ exhibits, for sufficiently large sizes of up to $n = 160$, two minima in the energy gap (Farhi *et al.*, 2012) (see Fig. 7 for an example). The first minimum, at $s \approx 0.36$, is associated with a second-order phase transition from paramagnetic to glassy, and the gap closes polynomially with system size. The second minimum occurs inside the spin-glass phase, with a gap that closes exponentially (or possibly a stretched exponential). Therefore, while the first minimum does not pose a problem for the adiabatic algorithm (although it has been shown that the quantum algorithm with a linear interpolating schedule does not pass through the associated glass phase transition faster than SA (Liu, Polkovnikov, and Sandvik, 2015)), the second minimum implies an exponential run time.

4. Ramsey numbers

An adiabatic algorithm for the calculation of the Ramsey numbers $R(k, l)$ was proposed by Gaitan and Clark (2012). $R(k, l)$ is the smallest integer r such that every graph on r or more vertices contains either a k clique or an l-independent set.[35] Computing them by brute force is doubly exponential in $N = \max\{k, l\}$ [note that $R(k, l) = R(l, k)$] using graph coloring techniques as follows: Try every one of the $2^{N(N-1)/2}$ colorings of the edges of the complete graph K_N with the colors blue and red. For every coloring, check whether or not there is an induced subgraph on k vertices

with only blue edges, or an induced subgraph on l vertices with only red edges. If every coloring contains at least one of the desired subgraphs, we are done. Otherwise, increment N by 1 and repeat. Except for certain special values of k and l, no better algorithm is currently known.

The idea of Gaitan and Clark (2012) was to construct a cost function $h(G)$ for a graph G where

$$h(G) = \mathcal{C}(G) + \mathcal{I}(G), \qquad (156)$$

where $\mathcal{C}(G)$ counts the number of m cliques in the graph G and $\mathcal{I}(G)$ counts the number of l-independent sets in the graph G. The cost $h(G)$ equals zero only if there does not exist a k clique or an l-independent set. This will occur only if $R(k, l) > N$. The algorithm then proceeds as follows. By mapping $h(G)$ over K_N to a final Hamiltonian H_1, the adiabatic algorithm $H(s) = -(1-s)\sum_{i=1}^{n}\sigma_i^x + sH_1$ is performed and the final energy of the state is measured. If $h(G) = 0$, then N is incremented by 1 and the experiment is repeated. This process continues until the first occurrence of $h(G) > 0$, in which case $N = R(k, l)$. Thus the algorithm is essentially an adiabatic version of the graph coloring method. It is unknown whether its StoqAQC version improves upon the classical brute force $2^{N(N-1)/2}$ scaling. The adiabatic quantum algorithm was simulated by Gaitan and Clark (2012) and shown to correctly determine the Ramsey numbers $R(3, 3)$ and $R(2, s)$ for $5 \leq s \leq 7$. It was also shown there that Ramsey number computation is in QMA.

An adiabatic algorithm for generalized Ramsey numbers (where the induced subgraphs are trees rather than complete graphs) was presented by Ranjbar *et al.* (2016). Whether this results in a quantum speedup is also unknown. We also remark that Ising formulations for many NP-complete and NP-hard problems, including all of Karp's 21 NP-complete problems (Karp, 1972), are known (Lucas, 2014), but it is unknown whether they are amenable to a quantum speedup.

5. Finding the largest cliques in random graphs

The fastest algorithm known to date for the NP-hard problem of finding the largest clique in a graph runs in time $O(2^{0.249n})$ for a graph with n vertices (Robson, 2001).[36] For random graphs, a superpolynomial time is required to find cliques larger than $\log n$ using the Metropolis algorithm, while the maximum clique is likely to be of a size very close to $2\log n$ (Jerrum, 1992). One of the earliest papers on the quantum adiabatic algorithm was concerned with the largest clique problem for random graphs (Childs *et al.*, 2002), although the algorithm presented there works for general graphs. The results were numerical and showed, by fixing the desired success probability, that the median time required by the adiabatic algorithm to find the largest clique in a random graph are consistent with quadratic growth for graphs of up to 18 vertices. These results on small graphs probably do not capture the asymptotic behavior of the algorithm (the

---

[35]A k clique is a subset of k vertices such that every two distinct vertices are adjacent. Equivalently, the subgraph induced by the clique is a complete graph. An independent set is a subset of the vertices no two of which are adjacent. $R(k, l)$ can be phrased as the "party problem": What is the smallest number of guests one can invite to a party such that there is always either a group of k guests that all know each other or a group of l guests, none of whom know each other? Such a threshold number always exists (Ramsey, 1930).

[36]As stated, this is actually an algorithm for the complementary maximum independent set (MIS) problem, but this is sufficient since $\text{MIS}(G) = \max -\text{clique}(\bar{G})$ for any graph G and its complement $\bar{G}$, and the algorithm applies for arbitrary G.

coefficients grow rapidly and have alternating sign), which is likely to be dominated by exponentially small gaps [however, to the extent that these are due to perturbative crossings, they can be avoided by techniques we discuss in Sec. VII.G, in particular, as related to the maximum independent set problem (Choi, 2011)].

6. Graph isomorphism

In the graph isomorphism problem, two N-vertex graphs G and G' are given, and the task is to determine whether there exists a permutation of the vertices of G such that it preserves the adjacency and transforms G to G', in which case the graphs are said to be isomorphic. If and only if the graphs are isomorphic does there exist a permutation matrix σ that satisfies

$$A' = \sigma A \sigma^T, \qquad (157)$$

where A and A' are the adjacency matrices of G and G', respectively. An adiabatic algorithm to determine whether a pair of graphs are isomorphic was first proposed by Hen and Young (2012), mostly for strongly regular graphs, and generalized to arbitrary graphs by Gaitan and Clark (2014), which also showed how to determine the permutation(s) that connect a pair of isomorphic graphs, and the automorphism group of a given graph. The final Hamiltonian formulated by Gaitan and Clark (2014) is such that when the ground state energy vanishes the graphs are isomorphic and the bit string $s = (s_0, \ldots, s_{N-1})$ associated with the ground state gives an $N \times N$ permutation matrix $\sigma(s)$ to perform the transformation:

$$\sigma(s)_{ij} = \begin{cases} 0 & \text{if } s_j > N - 1, \\ \delta_{i,s_j} & \text{if } 0 \le s_j \le N - 1. \end{cases} \qquad (158)$$

The computational complexity of these adiabatic algorithm is currently unknown. However, a recent breakthrough gave a quasipolynomial $(\exp[(\log n)^{O(1)}])$ time classical algorithm for graph isomorphism (Babai, 2015). It seems unlikely that this can be improved upon by using StoqAQC without deeply exploiting problem structure.

7. Machine learning

Quantum machine learning is currently an exciting and rapidly moving frontier in the context of the circuit model (Lloyd, Mohseni, and Rebentrost, 2014; Rebentrost, Mohseni, and Lloyd, 2014; Wiebe, Kapoor, and Svore, 2014), although it must be evaluated carefully (Aaronson, 2015). One StoqAQC approach is to find a quantum version of the classical method of boosting, wherein multiple weak classifiers (or features) are combined to create a single strong classifier (Freund, Schapire, and Abe, 1999; Meir, 2003). The task is to find the optimal set of weights of the weak classifiers so as to minimize the training error of the strong classifier on a training data set. After this training step, the strong classifier is then applied to a test data set. This optimization problem can be mapped to a quadratic unconstrained binary optimization (QUBO) problem, which can then be trivially turned into an Ising spin Hamiltonian suitable for adiabatic quantum optimization, where the binary variables represent the weights. This idea was implemented by Neven et al. (2008, 2009), Neven, Rose, and Macready (2008), Denchev et al. (2012), Pudenz and Lidar (2013), and Babbush et al. (2014), where the ground states found by the adiabatic algorithm encode the solution for the weights.

Another idea is to learn the weights of a Boltzmann machine or, after the introduction of a hidden layer, a reduced Boltzmann machine (Hinton, Osindero, and Teh, 2006). The latter forms the basis for various modern methods of deep learning. StoqAQC approaches for this problem were developed by Adachi and Henderson (2015), Amin et al. (2016), and Benedetti et al. (2016).

Neither the classical nor the quantum computational complexity is known in this case, but scaling of the solution time with problem size is not the only relevant criterion: classification accuracy on the test data set is another crucial metric. It is possible, although at this point entirely speculative, that the quantum method will lead to better classification performance. This can come about in the case of ground state degeneracy, if the weights are reconstructed via ground state solutions and if quantum and classical heuristics for solving the QUBO problem find different ground states (Matsuda, Nishimori, and Katzgraber, 2009; Azinović et al., 2016; Mandrà, Zhu, and Katzgraber, 2017; Zhang et al., 2017).

G. Speedup mechanisms?

While the universality of AQC suggests that similar speedup mechanisms are at play as in the circuit model of quantum computing, the situation is less clear regarding StoqAQC. Here we discuss two potential mechanisms, tunneling and entanglement, that might be thought to endow StoqAQC with an advantage over classical algorithms.

1. The role of tunneling

It is often stated that an advantage of StoqAQC over classical heuristic local-search algorithms is that the quantum system has the ability to tunnel through energy barriers, which can provide an advantage over classical algorithms such as simulated annealing that only allow probabilistic hopping over the same barriers. Indeed, such a qualitative picture motivated some of the early research on quantum annealing (Finnila et al., 1994). However, this statement requires a careful interpretation as it has the potential to be misleading. Whereas only the final cost function—which generates the energy landscape that the classical random walker explores—matters for the classical algorithm, this energy landscape does not become relevant for the quantum evolution until the end. Therefore, tunneling does not occur on the energy landscape defined by the final cost function alone, if it occurs at all. A different notion of tunneling is at work, which we now explain.

The standard notion of tunneling from single-particle quantum mechanics involves a semiclassical potential where classically allowed and classically forbidden regions can be defined. Starting from a many-body Hamiltonian, there is no unique way to take the semiclassical limit. Consider one such limit, based on the spin-coherent path integral formalism (Klauder, 1979):

$$\langle\Omega(t_f)|\mathrm{Texp}\left[-i\int_0^{t_f}d\tau H(\tau)\right]|\Omega(0)\rangle = \int \mathcal{D}\Omega(t)e^{(i/\hbar)S[\Omega(t)]}, \tag{159}$$

where the action $S[\Omega(t)]$ is given by

$$S[\Omega(t)] = \int_0^{t_f}dt[i\hbar\langle\Omega(t)|\partial_t|\Omega(t)\rangle - \langle\Omega(t)|H(t)|\Omega(t)\rangle], \tag{160}$$

and

$$\begin{aligned}|\Omega\rangle &\equiv |\theta,\varphi\rangle\\ &\equiv \otimes_{j=1}^n\left[\cos(\theta_j/2)|0\rangle_j + e^{i\varphi_j}\sin(\theta_j/2)|1\rangle_j\right]\end{aligned} \tag{161}$$

is the spin-coherent state (Arecchi *et al.*, 1972).

Despite the absence of a true kinetic term, we can identify the semiclassical potential as

$$V_{\mathrm{SC}}(\{\theta_j\},\{\varphi_j\},t) = \langle\Omega|H(t)|\Omega\rangle. \tag{162}$$

This form for V_{SC} has been used (Farhi, Goldstone, and Gutmann, 2002a; Schaller and Schützhold, 2010; Boixo *et al.*, 2016; Muthukrishnan, Albash, and Lidar, 2016) to capture many of the relevant features of StoqAQC problems endowed with qubit-permutation symmetry; this symmetry often allows for analytical and numerical progress.[37]

We illustrate this approach with the Grover Hamiltonian [Eq. (16)]. Recall that the final Hamiltonian is $H_1 = \mathbb{1} - |m\rangle\langle m|$, where $|m\rangle$ is the marked state associated with the marked item. As a cost function, this is the antithesis of the "tall and narrow" potential that is often associated with a classical speedup $\langle x|H_1|x\rangle = 1 - \delta_{x,m}$, i.e., the potential is flat everywhere, except for a well of constant depth at the marked state. Nevertheless, we now show that following the instantaneous ground state will involve the tunneling of $O(n)$ qubits.

Without loss of generality we assume that the "marked" state is the all-zero bit string. Setting $\theta_j \equiv \theta$ and $\varphi_j \equiv \varphi \,\forall\, j$ in Eq. (161), the Hamiltonian can be written succinctly as

$$\begin{aligned}H(s) = &(1-s)[\mathbb{1} - |\Omega(\pi/2,0)\rangle\langle\Omega(\pi/2,0)|]\\ &+ s[\mathbb{1} - |\Omega(0,0)\rangle\langle\Omega(0,0)|].\end{aligned} \tag{163}$$

The semiclassical potential for the Grover problem is then

$$\begin{aligned}V_{\mathrm{SC}}(\theta,0) = &(1-s)\left(1 - \frac{1}{2^n}(1+\sin\theta)^n\right)\\ &+ s\left(1 - \frac{1}{2^n}(1+\cos\theta)^n\right).\end{aligned} \tag{164}$$

The locations of the two degenerate minima at $s = 1/2$ are given by the pair of transcendental equations:

$$\frac{1-\cos\theta+\sin\theta}{1+\cos\theta-\sin\theta} = \left(\frac{1+\sin\theta}{1+\cos\theta}\right)^n, \tag{165a}$$

$$\frac{1+\cos\theta-\sin\theta}{1-\cos\theta+\sin\theta} = \left(\frac{1+\cos\theta}{1+\sin\theta}\right)^n, \tag{165b}$$

which in the limit of $n \to \infty$ have solutions 0 and $\pi/2$, respectively. This equation is invariant under $\theta \to \pi/2 - \theta$, which corresponds to the two minima. Since the semiclassical potential in Eq. (164) at $s = 1/2$ is also invariant under $\theta \to \pi/2 - \theta$, the local minima have identical structure. Using the Hamming weight operator defined as

$$\mathrm{HW} = \frac{1}{2}\sum_{i=1}^n(\mathbb{1} - \sigma_i^z), \tag{166}$$

this potential suggests that in the large n limit we can expect that $n/2$ spins need to be flipped in order to move from the $\theta \approx \pi/2$ minimum to the $\theta \approx 0$ minimum, i.e.,

$$\langle\Omega(\pi/2,0)|\mathrm{HW}|\Omega(\pi/2,0)\rangle - \langle\Omega(0,0)|\mathrm{HW}|\Omega(0,0)\rangle = n/2. \tag{167}$$

The instantaneous ground state, as it passes through the minimum gap at $s = 1/2$, indeed exhibits this behavior as shown in Fig. 8.

However, the more general role of tunneling in providing quantum speedups is not by any means evident. This topic was studied in detail by Muthukrishnan, Albash, and Lidar (2016), who showed that tunneling is neither necessary nor sufficient for speedups in the class of perturbed Hamming weight optimization problems with qubit-permutation symmetry.

Our discussion here has been restricted to coherent tunneling, and compelling arguments have been presented by Boixo *et al.* (2016), Denchev *et al.* (2016), and Andriyash and Amin (2017) that incoherent, thermally assisted tunneling plays a computational role in quantum annealing. However, this mechanism is in the open-system setting, which is outside the scope of this review. Moreover, its role in Boixo *et al.* (2016) and Denchev *et al.* (2016) was limited to a prefactor and does not translate into a scaling advantage, i.e., it does not qualify as a speedup according to the classification of Rønnow *et al.* (2014).

2. The role of entanglement

The role that entanglement plays in quantum computation with pure states in the circuit model depends on the entanglement measure used. On the one hand, it is well known that for any circuit model quantum algorithm operating on pure states the presence of multipartite entanglement quantified via the Schmidt rank (with a number of parties that increases unboundedly with input size) is necessary if the quantum algorithm is to offer an exponential speedup over classical computation (Jozsa and Linden, 2003). On the other hand, universal quantum computation can be achieved in the standard pure-state circuit model while the entanglement entropy (or any other suitably continuous entanglement measure) of every bipartition is small in each step of the
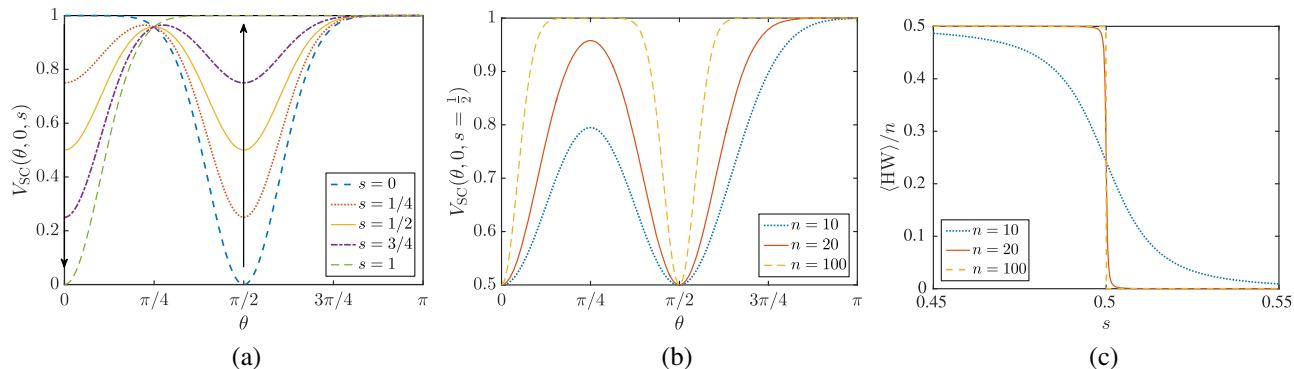
FIG. 8. Analysis of tunneling in the Grover problem. (a) The semiclassical potential for $n = 20$ at different dimensionless times s. The arrows indicate the behavior of the local minima as s increases. There is a discrete jump in the position of the global minimum at $s = 1/2$, where it changes from being at $\theta \approx \pi/2$ to $\theta \approx 0$, corresponding to a first-order quantum phase transition. (b) The behavior of the potential when the two minima are degenerate at $s = 1/2$. As n grows, both the barrier height grows (and saturates at 1) and the curvature of the local minima grows. (c) The expectation value of the Hamming weight operator [defined in Eq. (166)] of the instantaneous ground state as n grows. This is to be interpreted as the system requiring $O(n)$ spins to tunnel in order to follow the instantaneous ground state as the system crosses the minimum gap at $s = 1/2$.

computation (Van den Nest, 2013). The corresponding role of entanglement in the computational efficiency of AQC remains an open question. Partly this is because the connection between entanglement and spectral gaps is not yet very well understood, and partly this is because even if entanglement is present, its computational role in AQC is unclear.

The area law asserts that for any subset S of particles, the entanglement entropy between S and its complement is bounded by the surface area of S rather than the trivial bound of the volume of S. While generic quantum states do not obey an area law (Hayden, Leung, and Winter, 2006), and there are 1D systems for which there is exponentially more entanglement than suggested by the area law (Movassagh and Shor, 2016), a sweeping conjecture in condensed matter physics is that in a *gapped* system the entanglement spreads only over a finite length, which leads to area laws for the entanglement entropy (Eisert, Cramer, and Plenio, 2010).[38] For example, the area law for gapped 1D systems, proved by Hastings (2007), states that for the ground state, the entanglement of any interval is upper bounded by a constant independent of the size of the interval. While this leaves open the question of the general dependence of the upper bound on the spectral gap Δ, this means that the ground state of such systems is accurately described by polynomial-size matrix product states (MPSs) (White, 1992; White and Noack, 1992; Östlund and Rommer, 1995). Gottesman and Hastings (2010) showed that for a certain 1D system the entanglement entropy in some regions can be as high as poly$(1/\Delta)$. This demonstrates that the entanglement entropy can become large as the gap becomes small. Two other important recent results are the existence of a polynomial-time algorithm for the ground state of 1D gapped local Hamiltonians with constant ground state energy (Huang, 2014; Landau, Vazirani, and Vidick, 2015), and the fact that 1D quantum many-body

states satisfying exponential decay of correlations always fulfill an area law (Brandao and Horodecki, 2013).

However, the connection between entanglement entropy and gaps is not nearly as clear in higher dimensional systems, even though entanglement close to quantum phase transitions is a well-developed subject (Osborne and Nielsen, 2002; Osterloh *et al.*, 2002; Vidal *et al.*, 2003; Wu, Sarandy, and Lidar, 2004; Amico *et al.*, 2008).

It is not surprising that entanglement is necessary for the computation to succeed if the intermediate ground states that the system must follow are entangled. This was explicitly verified by Bauer *et al.* (2015), where the quantum state was represented by an MPS and projected entangled-pair states (PEPS) (Verstraete and Cirac, 2004; Verstraete, Murg, and Cirac, 2008). This work showed that the probability of finding the ground state of an Ising spin glass on either a planar or nonplanar two-dimensional graph increases with the amount of entanglement in the MPS state or PEPS state. Furthermore, even a small amount of entanglement gives improved success probability over a mean-field model. However, this does not resolve the role entanglement plays in generating a speedup.

In an attempt to address this, the entanglement entropy for the adiabatic Grover algorithm was studied, and it was found to be bounded (≤ 1) throughout the evolution (Orús and Latorre, 2004). This was also observed numerically for systems with 10 qubits (Wen and Qiu, 2008). In an effort to check whether more entanglement may help the Grover speedup, Wen, Huang, and Qiu (2009) considered adding an additional term to the Hamiltonian to make the ground state more entangled, to reach an $O(1)$ scaling in a Grover search task. However, since it is impossible to achieve a better-than-quadratic speedup in the Grover search problem without introducing an explicit dependence on the marked state (Bennett *et al.*, 1997), this result is not conclusive in linking entanglement with enhanced computational efficiency. Furthermore, a two-dimensional path for the Grover problem using the quantum adiabatic brachistochrone approach (see Sec. VII.B) that gives a higher success probability for the same evolution time relative to the standard one-dimensional

---

[38]Here "gapped" means $O(1)$, whereas in AQC gapped usually means $O[1/\text{poly}(n)]$.

path for the Grover problem, in fact has *less* entanglement (negativity) (Rezakhani *et al.*, 2009).

The entanglement entropy in the adiabatic algorithm for the Exact Cover problem, where no speedup is known (recall Sec. VI.F.2), scales linearly with problem size for $n \leq 20$ (Latorre and Orus, 2004; Orús and Latorre, 2004).

Further studies have also shown this lack of correlation between performance and the amount of entanglement entropy. Hauke *et al.* (2015) performed simulations of adiabatic quantum optimization of a trapped ion Hamiltonian with $n = 16$ of the form

$$H_1 = J \sum_{i \neq j}^{n} \frac{\sigma_i^z \sigma_j^z}{|i-j|} + \sum_i h_i^z \sigma_i^z + V \sum_{i \neq j}^{n} \sigma_i^z \sigma_j^z, \quad (168)$$

with 100 disorder realizations of h_i^z. It was found that a large entanglement entropy has little significance for the success probability of the optimization task.

Overall, these results indicate that the connection between entanglement and algorithmic efficiency in AQC is currently wide open and deserves further study.

VII. CIRCUMVENTING SLOWDOWN MECHANISMS FOR AQC

In this section we collect several insights into mechanisms that explain slowdowns in the performance of adiabatic algorithms. We also discuss mechanisms for circumventing such slowdowns. Several important ideas will be reviewed: avoiding the use certain initial and final Hamiltonians, modifying the adiabatic schedule, avoiding quantum phase transitions, and avoiding perturbative energy level crossings.

A. Avoiding poor choices for the initial and final Hamiltonians

We first show that if one chooses the initial Hamiltonian to be the one-dimensional projector onto the uniform superposition state $|\phi\rangle$, and uses a linear interpolation, then an improvement beyond a Grover-like quadratic speedup is impossible as long as the final Hamiltonian H_1 is diagonal in the computational basis. Specifically, for an adiabatic algorithm of the form

$$H(t) = \left(1 - \frac{t}{t_f}\right) E(\mathbb{1} - |\phi\rangle\langle\phi|) + \frac{t}{t_f} H_1, \quad (169)$$

the run time t_f for measuring the ground state of H_1 with probability p is lower bounded by [Theorem 1 of Farhi *et al.* (2008); see also Žnidarič and Horvat (2006)]

$$t_f \geq \frac{2}{E}(1 - \sqrt{1-p})\sqrt{\frac{N}{k}} - 2\frac{\sqrt{p}}{E}, \quad (170)$$

where $N = 2^n$ and k is the degeneracy of the ground state of H_1. To see this, define an operator V_x for $x = 0, \ldots, N-1$ that is diagonal in the computational basis:

$$\langle z|V_x|z\rangle = e^{2\pi i z x/N}, \quad (171)$$

and let $|x\rangle = V_x|\phi\rangle = (1/\sqrt{N}) \sum_{z=0}^{N-1} e^{2\pi i z x/N}|z\rangle$. Now define the modified adiabatic algorithm

$$H_x(t) = \left(1 - \frac{t}{t_f}\right) E(\mathbb{1} - |x\rangle\langle x|) + \frac{t}{t_f} H_1. \quad (172)$$

Note that $|x=0\rangle = |\phi\rangle$ implies that $H_0(t) = H(t)$. For each x, the final state is given by $|\psi_x\rangle = U_x(t_f, 0)|x\rangle$, with success probability $p_x = \langle\psi_x|P|\psi_x\rangle$, where P is the projector onto the ground subspace of H_1. Using $H_x(t) = V_x H_0 V_x^\dagger$, we have $U_x(t, 0) = V_x U_0(t, 0) V_x^\dagger$, and hence $p_x = p$, $\forall x$ since V_x commutes with P. We should already see a potential problem for having t_f scale better than $\sqrt{N}$, since if we were to run the algorithm backward, we would find the state $|x\rangle$, which would be solving the Grover problem [note that the initial Hamiltonian (172) is the Grover Hamiltonian in a rotated basis].

Now define an evolution according to an x-independent Hamiltonian:

$$H_R(t) = \left(1 - \frac{t}{t_f}\right) E\mathbb{1} + \frac{t}{t_f} H_1, \quad (173)$$

and let $|g_x\rangle = (1/\sqrt{p})P|\psi_x\rangle$. Consider the difference in the reverse evolutions associated with $H_R(t)$ and $H_x(t)$ from $|g_x\rangle$:

$$S(t) = \sum_x \|[U_x^\dagger(t_f, t) - U_R^\dagger(t_f, t)]|g_x\rangle\|^2. \quad (174)$$

We can write $|g_x\rangle = \sqrt{p}|\psi_x\rangle + \sqrt{1-p}|\psi_x^\perp\rangle$, where $|\psi_x^\perp\rangle$ is orthogonal to $|\psi_x\rangle$. Using $U_x^\dagger(t_f, 0)|\psi_x\rangle = |x\rangle$ and defining $|R_x\rangle = U_R^\dagger(t_f, 0)|g_x\rangle$, we have

$$S(0) = \sum_x \|\sqrt{p}|x\rangle + \sqrt{1-p}|x^\perp\rangle - |R_x\rangle\|^2 \quad (175a)$$

$$= 2N - \sum_x [\sqrt{p}\langle x|R_x\rangle + \sqrt{1-p}\langle x^\perp|R_x\rangle + \text{c.c.}] \quad (175b)$$

$$\geq 2N - 2\sqrt{p}\sum_x |\langle x|R_x\rangle| - 2N\sqrt{1-p}. \quad (175c)$$

Since H_R commutes with H_1, the state $|R_x\rangle$ is an element of the k-dimensional ground subspace of H_1. Choosing a basis $\{|G_i\rangle\}_{i=1}^k$ for this subspace, and writing $|R_x\rangle = \sum_{i=1}^k \alpha_{x,i}|G_i\rangle$, we have

$$\sum_x |\langle x|R_x\rangle| \leq \sum_{x,i} |\alpha_{x,i}| \cdot |\langle x|G_i\rangle|$$

$$\leq \sqrt{\sum_{x,i} |\alpha_{x,i}| \sum_{x',i'} |\langle x'|G_{i'}\rangle|} = \sqrt{Nk}. \quad (176)$$

Therefore, we have

$$S(0) \geq 2N(1 - \sqrt{1-p}) - 2\sqrt{Nkp}. \quad (177)$$

In order to upper bound $S(0)$, we use $S(t_f) - S(0) \leq \int_0^{t_f} |(d/dt)S(t)| dt$ with $S(t_f) = 0$. The derivative can be computed using the Schrödinger equation

$$\frac{d}{dt}S(t) = -i\sum_x \langle g_x|U_x(t_f,t)[H_x(t) - H_R(t)]U_R^\dagger(t_f,t)|g_x\rangle$$
$$+ \text{c.c.}$$
$$= -2\Im\sum_x\left(1 - \frac{t}{t_f}\right)E\langle g_x|U_x(t_f,t)|x\rangle$$
$$\times \langle x|U_R^\dagger(t_f,t)|g_x\rangle. \tag{178}$$

Thus,

$$\left|\frac{d}{dt}S(t)\right| \le 2E\left(1 - \frac{t}{t_f}\right)\sum_x|\langle x|U_R^\dagger(t_f,t)|g_x\rangle| \tag{179a}$$

$$\le 2E\left(1 - \frac{t}{t_f}\right)\sqrt{Nk}, \tag{179b}$$

where in Eq. (179b) we used the same trick as in Eq. (176). Therefore, $\int_0^{t_f}|(d/dt)S(t)|dt \le Et_f\sqrt{Nk}$. Putting the upper and lower bounds for $S(0)$ together we have

$$Et_f\sqrt{Nk} \ge 2N(1 - \sqrt{1-p}) - 2\sqrt{Nkp}, \tag{180}$$

which yields Eq. (170).

As an example of the relevance of this result, consider the trivial case of n decoupled spins in a global magnetic field. For an initial Hamiltonian that reflects the bit structure of the problem, e.g., the standard $H_0 = -\sum_i \sigma_i^x$, the run time of the adiabatic algorithm scales as $\sqrt{n}$ (Muthukrishnan, Albash, and Lidar (2016); Brady and van Dam, 2017). If, however, we were to choose instead the projector initial Hamiltonian, the result shows that we would find a dramatically poor scaling despite the simplicity of the final Hamiltonian.

A similar result is found if all structure is removed from the final Hamiltonian. Namely, if $H_1 = \sum_z h(z)|z\rangle\langle z|$, we can define a permutation π over the N computational-basis states such that $h^{[\pi]}(z) = h(\pi^{-1}(z))$. Assume that the initial Hamiltonian is π independent and that $c(t)$ satisfies $|c(t)| \le 1$. Then, for the permuted Hamiltonian $H_{1,\pi} = \sum_z h(z)|\pi(z)\rangle\times\langle\pi(z)|$, one can show that if the adiabatic algorithm

$$H_\pi(t) = H_0 + c(t)H_{1,\pi} \tag{181}$$

succeeds with probability p for a set of $\epsilon N!$ permutations, then [Theorem 2 of Farhi et al. (2008)]

$$t_f \ge \frac{\epsilon^2 p}{16h^*}\sqrt{N-1} - \frac{\epsilon\sqrt{\epsilon/2}}{4h^*}, \tag{182}$$

where $h^* = \sqrt{\sum_z h(z)^2/N - 1}$. This result means that no algorithm of the form of Eq. (181) can find the minimum of $H_{1,\pi}$ with a constant probability for even a fraction of all permutations if t_f is $o(\sqrt{N})$.

The lesson from this analysis is *what not to do* when designing quantum adiabatic algorithms: avoid choosing the initial Hamiltonian to be the one-dimensional projector onto the uniform superposition state if a better-than-quadratic speedup is hoped for and avoid removing structure from the final Hamiltonian.

B. Quantum adiabatic brachistochrone

Modifying the adiabatic schedule adaptively so that it slows down as the gap decreases is an approach that is essential for obtaining a quadratic speedup using the adiabatic Grover algorithm (recall Sec. III.A). Here we discuss how such ideas, including the condition for the locally optimized schedule [Eq. (24)] can be understood as arising from a variational time-optimal strategy for determining the interpolating Hamiltonian between H_0 and H_1 (Rezakhani et al., 2009). By time optimal we mean a strategy that gives rise to the shortest total evolution time t_f while guaranteeing that the final evolved state $|\psi(t_f)\rangle$ is close to the desired final ground state $|\varepsilon_0(t_f)\rangle$. The success of the strategy is judged by the trade-off between t_f and the fidelity $F(t_f) = |\langle\psi(t_f)|\varepsilon_0(t_f)\rangle|^2$. We first discuss this method generally and then show how it applies to the adiabatic Grover case.

The interpolating Hamiltonian's time dependence comes from a set of control parameters $\vec{x}(t) = (x^1(t), \ldots, x^M(t))$, i.e., $H(t) = H[\vec{x}(t)]$. We can parametrize $\vec{x}(t)$ in terms of a dimensionless time parameter $s(t)$ with $s(0) = 0$ and $s(t_f) = 1$, where $v = ds/dt$ characterizes the speed of motion along the control trajectory $\vec{x}[s(t)]$. The total evolution time is then given by

$$t_f = \int_0^1 \frac{ds}{v(s)}. \tag{183}$$

Motivated by the form of the adiabatic condition, let us define the following Lagrangian:

$$\mathcal{L}[\vec{x}(s), \dot{\vec{x}}(s)] \equiv \frac{\|\partial_s H(s)\|_{\text{HS}}^2}{\Delta^p(s)}$$
$$= \sum_{i,j}\frac{\text{Tr}[\partial_{x^i}H(s)\partial_{x^j}H(s)]}{\Delta^p(s)}\partial_s x^i(s)\partial_s x^j(s) \tag{184}$$

($p > 0$), and adiabatic-time functional

$$\mathcal{T}[\vec{x}(s)] = \int_0^1 ds\,\mathcal{L}[\vec{x}(s), \dot{\vec{x}}(s)], \tag{185}$$

where $\|A\|_{\text{HS}} \equiv \sqrt{\text{Tr}(A^\dagger A)}$ is the Hilbert-Schmidt norm, chosen to ensure analyticity (this choice is not unique, but other choices may not induce a corresponding Riemannian geometry). The time-optimal curve $\vec{x}_{\text{QAB}}(s)$ is the quantum adiabatic brachistochrone (QAB) and is the solution of the variational equation $\delta\mathcal{T}[\vec{x}(s)]/\delta\vec{x}(s) = 0$.

Alternatively, the problem can be thought of in geometrical terms. The integral in Eq. (185) is of the form $\int ds \sum_{i,j} g_{ij}(\vec{x})\partial_s x^i\partial_s x^j$, which defines a reparametrization-invariant object. Therefore, using results from differential geometry, the Euler-Lagrange equations derived from extremizing Eq. (185) are simply the geodesic equations associated with the metric g_{ij} appearing in $\mathcal{L}[\vec{x}(s), \dot{\vec{x}}(s)] = g_{ij}(\vec{x})\dot{x}^i\dot{x}^j$ (Einstein summation convention):

$$\partial_s^2 x^k + \Gamma_{ij}^k \partial_s x^i \partial_s x^j = 0, \tag{186}$$

where

$$\Gamma_{ij}^k = \tfrac{1}{2} g^{kl} (\partial_j g_{li} + \partial_i g_{lj} - \partial_l g_{ij})$$

are the Christoffel symbols (connection coefficients) and

$$g_{ij}(\vec{x}) = \frac{\mathrm{Tr}[\partial_i H(\vec{x}) \partial_j H(\vec{x})]}{\Delta^p(\vec{x})}. \tag{187}$$

To find the variational time-optimal strategy associated with minimizing Eq. (185), the procedure is thus as follows: (a) solve Eq. (186) to find the optimal path $\vec{x}_{\mathrm{QAB}}(s)$, and (b) compute the adiabatic error using the Schrödinger equation along this optimal path (or multiparameter schedule). Note that to compute the metric requires knowledge of the gap or at least an estimate thereof.

The optimal path is a geodesic in the parameter manifold endowed with the Riemannian metric g. This metric gives rise to a curvature tensor $\mathbf{R}$, which can be computed from the metric tensor and the connection using standard methods (Nakahara, 1990). Namely, $\Gamma \sim g^{-1} \partial g \sim \Delta^{-1} \partial \Delta$, and $\mathbf{R} \sim \partial^2 g + g\Gamma^2 \sim \Delta^{-p-2}$. Thus, the smaller the gap, the higher the curvature.

Let us illustrate with a simple example. Consider the following Hamiltonian with a *single* control parameter $x^1(s)$:

$$H(s) = [1 - x^1(s)] P_a^\perp + x^1(s) P_b^\perp, \tag{188}$$

where we have defined the projector $P_a^\perp = \mathbb{1} - |a\rangle\langle a|$ and similarly for $P_b^\perp$. This includes the Grover problem as the special case where $|a\rangle$ is the uniform superposition and $|b\rangle$ is the marked state. We can always find a state $|a^\perp\rangle$ such that $|b\rangle = \alpha_0 |a\rangle + \alpha_1 |a^\perp\rangle$, where $\alpha_0 = \langle a|b\rangle$. Therefore, the evolution according to $H(s)$ occurs in a two-dimensional subspace spanned by $|a\rangle$ and $|a^\perp\rangle$, and

$$\partial_{x^1} H(s) = -P_a^\perp + P_b^\perp, \tag{189a}$$

$$\mathrm{Tr}[\partial_{x^1} H(s) \partial_{x^1} H(s)] = 2(1 - |\alpha_0|^2), \tag{189b}$$

$$\Delta(s) = \sqrt{1 - 4(1 - |\alpha_0|^2) x^1(s)[1 - x^1(s)]}, \tag{189c}$$

$$g_{11} = \frac{2(1 - |\alpha_0|^2)}{\Delta(s)^3}. \tag{189d}$$

The geodesic equation is then given by

$$\frac{d^2}{ds^2} x^1(s)$$
$$+ \frac{p[1 - 2x^1(s)](1 - |\alpha_0|^2)}{1 - 4x^1(s)[1 - x^1(s)](1 - |\alpha_0|^2)} \left(\frac{d}{ds} x^1(s)\right)^2 = 0. \tag{190}$$

In the case of $p = 4$, we can solve this equation analytically, and the solution with the boundary conditions $x^1(0) = 1 - x^1(1) = 0$ is given by

$$x^1(s) = \frac{1}{2} + \frac{|\alpha_0|}{2\sqrt{1 - |\alpha_0|^2}} \tan\left[\cos^{-1}(|\alpha_0|)(2s - 1)\right]. \tag{191}$$

Note that

$$\cos^{-1}(|\alpha_0|) = \tan^{-1}\left(\frac{\sqrt{1 - |\alpha_0|^2}}{|\alpha_0|}\right).$$

Remarkably, this is equivalent to the expression we found for the Grover problem [Eq. (29)] if we take $\alpha_0 = 1/\sqrt{N}$, despite the different choice of norm and value of p. This shows that the optimal schedule for the Grover problem has a deep differential geometric origin.

We can extend the analysis to *two* control parameters such that the time-dependent Hamiltonian is given by

$$H(s) = x^1(s) P_a^\perp + x^2(s) P_b^\perp. \tag{192}$$

The associated QAB (or geodesic) path can be found numerically, and it turns out that it is not of the form $x^2(s) = 1 - x^1(s)$, i.e., it is different from the (Roland and Cerf, 2002) path given by Eq. (191). The optimal two-parameter path reduces the adiabatic error relative to the latter [see Fig. 9(a)],
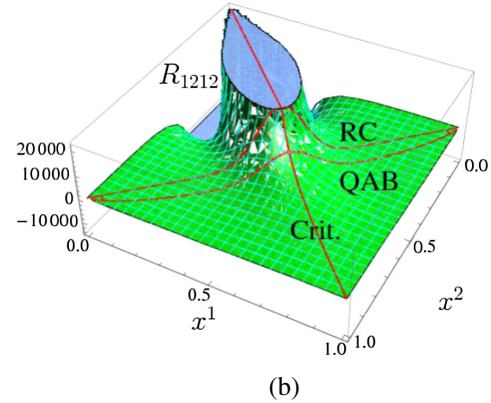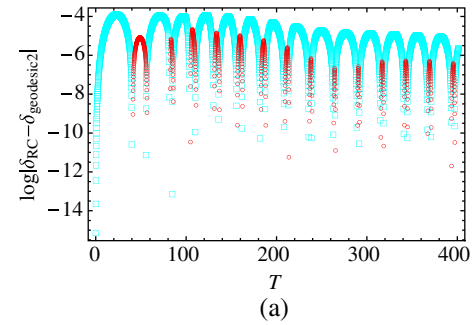

(a)

(b)

FIG. 9. (a) Final-time error $\delta(T) = \sqrt{1 - F(T)}$ ($T = T_f$ in our notation) for the single-control (denoted RC for Roland-Cerf) and two-parameter control (denoted geodesic2) geodesic paths for the Grover problem with $n = 6$. Squares (cyan) indicate where the two-parameter geodesic path outperforms (i.e., has a lower error than) the single-parameter path, and circles (red) correspond to the opposite case. (b) The curvature tensor component R_{1212} for $n = 3$. The curves on the curvature surface show the case of the standard linear interpolation $x_2 = 1 - x_1$ (denoted Crit.), the path followed by the one-parameter geodesic (denoted RC), and the path followed by the two-parameter geodesic (denoted QAB). From Rezakhani *et al.*, 2009.

but cannot, of course, reduce the (already optimal) $\sqrt{N}$ scaling. The two-parameter QAB also has lower curvature than the Roland-Cerf path [see Fig. 9(b)], which implies that it follows a path with a larger gap and less entanglement than the latter (Rezakhani *et al.*, 2009) as mentioned in Sec. VI.G.2.

The differential geometric approach to AQC was further explored by Rezakhani *et al.* (2010), where its connections to quantum phase transitions were elucidated within a unifying information-geometric framework; see also Zulkowski and DeWeese (2015).

C. Modifying the initial Hamiltonian

Rather than modifying the adiabatic interpolation, one may modify the initial Hamiltonian. Such a strategy was pursued by Farhi *et al.* (2011) and tried on a particular set of 3-SAT instances, where the clauses are picked randomly subject to satisfying two disparate planted solutions and then penalizing one of them with a single additional clause. This was done in order to generate instances with an avoided crossing at the final time $s = 1$, reproducing the type of obstacle to AQC envisioned by Altshuler, Krovi, and Roland (2010).

It was then shown that in this case, by picking a random initial Hamiltonian of the form

$$H_0 = \frac{1}{2}\sum_{i=1}^{n} c_i(\mathbb{1} - \sigma_i^x), \qquad (193)$$

where c_i is a random variable taking value $1/2$ or $3/2$ with equal probability, it is possible to remove the small gap encountered by the standard adiabatic algorithm with high probability. Since this strategy does not rely on information about the specific instance, it appears to be quite general. Therefore, if the algorithm is to be run on a single instance of some optimization problem, the adiabatic algorithm should be run repeatedly with different initial Hamiltonians (Farhi *et al.*, 2011).

An alternative approach based on modifying the initial Hamiltonian, with a different goal, was proposed by Perdomo-Ortiz, Venegas-Andraca, and Aspuru-Guzik (2011), whereby an initial guess for the solution (a computational-basis state) is used as the initial state of the adiabatic algorithm. An initial Hamiltonian is used with this state as its ground state. Evolution to the final Hamiltonian then proceeds according to a standard schedule. If the final state that is measured is not the ground state of the final Hamiltonian (due to diabatic transitions), the algorithm can be repeated with the measured state as the new initial state. Such "warm start" repetitions of the algorithm exhibited improved performance compared to the standard approach for 3-SAT problems, although the results were limited to relatively small system sizes of six and seven qubits.

D. Modifying the final Hamiltonian

The same problem can be specified by two or more different final Hamiltonians, as seen in the case of EC3, in terms of Eqs. (153) and (154). Altshuler, Krovi, and Roland (2010) claimed that adiabatic quantum optimization fails for random instances of EC3 because of Anderson localization. The claim,

discussed in more detail in Sec. VII.G, was based on the form given in Eq. (154). However, Choi (2011) argued that it is possible to reformulate the final Hamiltonian for EC3 such that the argument in Altshuler, Krovi, and Roland (2010) may not apply. Namely, for any pair of binary variables x_{C_i}, x_{C_j} in the same clause C, add a term $D_{ij}x_{C_i}x_{C_j}$ with $D_{ij} > 0$; this is permissible since in order for a clause to be satisfied, exactly one variable must take value 1, whereas the other two are 0. Numerical evidence for up to 15 bits suggests that the addition of the new set of arbitrary parameters D_{ij} may avoid the Anderson localization issue (Choi, 2011). This example illustrates a general principle that it can be incorrect to conclude from the failure of one specific choice of the final Hamiltonian that all quantum adiabatic algorithms fail for the same problem.

E. Adding a catalyst Hamiltonian

We define a "catalyst" as a term that (1) vanishes at the initial and final times, but is present at intermediate times, (2) is a sum of local terms with the same qubit-interaction graph as the final Hamiltonian H_1, and (3) does not use any other information specific to the particular instance.

Consider, e.g.,

$$H(s) = (1-s)H_0 + s(1-s)H_C + sH_1. \qquad (194)$$

The specific form of H_C is of course important, but even a randomly chosen catalyst can help (Farhi, Goldstone, and Gutmann, 2002b; Farhi *et al.*, 2011; Zeng, Zhang, and Sarovar, 2016). We illustrate how H_C can turn a slowdown (exponential run time) into a success (at worst polynomial run time) for a specific problem with a specific H_C that is analytically tractable. Consider a final Hamiltonian of the form

$$H_1 = \sum_z h(z)|z\rangle\langle z|, \qquad (195)$$

where z denotes an n-bit string, and $h(z) = \sum_{i<j<k} h_3(z_i, z_j, z_k)$ with

$$h_3(z_1, z_2, z_3) = \begin{cases} 0, & z_1 + z_2 + z_3 = 0, \\ 3, & z_1 + z_2 + z_3 = 1, \\ 1, & z_1 + z_2 + z_3 = 2, \\ 1, & z_1 + z_2 + z_3 = 3. \end{cases} \qquad (196)$$

The all-zero bit string minimizes the final Hamiltonian with energy 0.

The cost function $h(z)$ is bit-permutation symmetric and depends only on the Hamming weight $|z|$, which facilitates the analysis. Specifically (Farhi, Goldstone, and Gutmann, 2002a),

$$h(z) = \frac{3}{2}|z|(n-|z|)(n-|z|-1) + \frac{1}{2}|z|(|z|-1)(n-|z|)$$
$$+ \frac{1}{6}|z|(|z|-1)(|z|-2). \qquad (197)$$

The final Hamiltonian can then be written in terms of the total spin operators $S^\alpha = (1/2)\sum_{i=1}^n \sigma_i^\alpha$ by using the mapping $|z| \mapsto n/2 - S^z$. The initial Hamiltonian is taken to be

$$H_0 = \frac{(n-1)(n-2)}{2}\left(\frac{n}{2}\mathbb{1} - S^x\right) \qquad (198)$$

[the unconventional normalization is to ensure that both H_1 and H_0 scale similarly with n (Farhi *et al.*, 2000)]. H_C is taken to be identical for all combinations of 3 bits in order to preserve the permutation symmetry:

$$H_C = -2n(S^x S^z + S^z S^x). \qquad (199)$$

Note that this catalyst is nonstoquastic. A useful way to characterize the change due to the introduction of H_C is to study the semiclassical potential associated with the Hamiltonian:

$$V(s, \theta, \varphi) = \langle \theta, \varphi | H(s) | \theta, \varphi \rangle, \qquad (200)$$

where $|\theta, \varphi\rangle$ is the spin-coherent state defined in Eq. (161). In the large n limit we have (Farhi, Goldstone, and Gutmann, 2002b)

$$\lim_{n \to \infty} V/(2/n)^3 = 2(1-s)(1 - \sin\theta\cos\varphi)$$
$$+ \tfrac{1}{6}s(13 + 3\cos\theta - 9\cos^2\theta - 7\cos^3\theta)$$
$$- 8s(1-s)\cos\theta\sin\theta\cos\varphi, \qquad (201)$$

where the three terms arise from the initial, final, and catalyst Hamiltonians, respectively. We display the behavior of this potential in Fig. 10. In the absence of H_C, there is a value of s where the potential has degenerate minima, and the system must tunnel from the right well to the left well in order to follow the global minimum. This point is associated with an exponentially small gap (Farhi, Goldstone, and Gutmann, 2002a), i.e., the algorithm requires exponential time to follow the global minimum. However, in the presence of H_C the potential never exhibits such an obstacle; there is always a single global minimum that the system can follow from $s = 0$ to $s = 1$ with polynomial run time.

Using this method of introducing a catalyst Hamiltonian, improvements were generally observed on a large number of MAX 2-SAT instances of size $n = 20$ (by directly solving the Schrödinger equation) (Crosson *et al.*, 2014). Both stoquastic and nonstoquastic H_C were tried and improved the success rate, but the difference between stoquastic and nonstoquastic was not decisive.

A similar study was performed by Hormozi *et al.* (2017) on fully connected Ising instances $H_1 = \sum_{i=1}^n h_i \sigma_i^z + \sum_{i=j}^n J_{ij}\sigma_i^z\sigma_j^z$ of size $n \leq 17$, where the J_{ij}'s and h_i's were picked from a continuous Gaussian distribution with zero mean and unit variance. They observed that a stoquastic catalyst generally improves the performance of easy instances by boosting the minimum gap and reducing the number of anticrossings. The fraction of instances affected tends to grow with increasing problem size. This is in stark contrast to nonstoquastic catalysts that tend to improve the performance of the very hard instances, but the fraction of improved instances remains constant with increasing problem size.
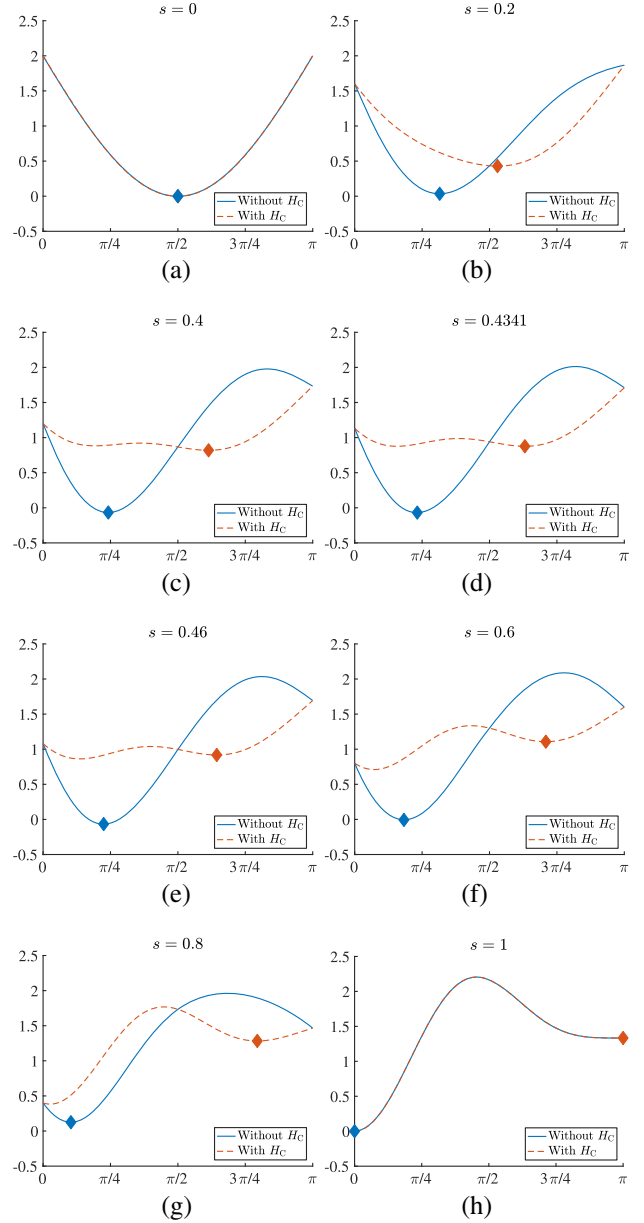

FIG. 10. The diamonds represent the minima followed by a polynomial run time. In the case with H_C, the potential can follow the global minimum polynomial time. In the case without H_C, there is an s value where the potential has a degenerate minimum, and the algorithm cannot tunnel to the new global minimum in polynomial time.

Furthermore, the gap does not generically increase with the addition of this catalyst, and the number of anticrossings grows. This latter feature leads to the increased success probability as population lost from the ground state at one anticrossing can be recovered at a later anticrossing.

F. Addition of nonstoquastic terms

The addition of nonstoquastic terms was already considered numerically in the previous section; here we focus on analytical results obtained for certain mean-field models.

Quantum statistical-mechanical techniques (Trotter-Suzuki decomposition, replica method under the replica-symmetric ansatz, and the static approximation) were used by Seki and Nishimori (2012, 2015), Seoane and Nishimori (2012), and Nishimori (2016) to analyze infinite-range Ising models with ferromagnetic as well as random interactions. These studies concluded that nonstoquastic terms can sometimes modify first-order quantum phase transitions (with an exponentially small gap) in the stoquastic Hamiltonian to second-order transitions (with a polynomially small gap) in the modified, nonstoquastic Hamiltonian.

The Hamiltonian is of the form

$$H(s, \lambda) = (1 - s)H_0 - s[\lambda H_{1,z}^{(p)} + (1 - \lambda)H_{1,x}^{(k)}], \quad (202)$$

where $H_0 = -\sum_{i=1}^{n} \sigma_i^x$ is a standard initial Hamiltonian, and

$$H_{1,\alpha}^{(q)} = n\left(\frac{1}{n}\sum_{i=1}^{n}\sigma_i^\alpha\right)^q, \quad \alpha \in \{x, z\}, \quad q \in \{p, k\}, \quad (203)$$

where $\lambda \in [0, 1]$ controls the strength of the nonstoquastic term $H_{1,x}^{(k)}$, and both p and k are integers ≥ 2 that determine the locality of the model. The parameter λ is increased to 1 along with s, so that the final Hamiltonian is the infinite-range p-body ferromagnetic Ising model $H(1, 1) = -H_{1,z}^{(p)}$. Also the r-pattern Hopfield model was studied, where $\lambda H_{1,z}^{(p)}$ is replaced by $-\sum_{1 \leq i_1 < \cdots < i_p \leq n} J_{i_1 \cdots i_p}\sigma_{i_1}^z \cdots \sigma_{i_p}^z$, where $J_{i_1 \cdots i_p}$ is given in Eq. (143), with ξ_{i_p} being ± 1 with equal probability.

In the ferromagnetic case Seki and Nishimori (2012) and Seoane and Nishimori (2012) showed that for $p \geq 4$ a two-local nonstoquastic XX term[39] changes the first-order phase transition to a second-order one, for an appropriately chosen path in the (λ, s) plane, starting from $(\lambda_0, 0)$ (with arbitrary λ_0) and ending at $(1,1)$. The situation in the Hopfield model case is identical to the ferromagnetic case, for an extensive number of patterns $r \propto n$. For a fixed number of patterns $p \geq 5$ is sufficient and $p > 3$ is necessary in order to avoid first-order phase transitions (Seki and Nishimori, 2015).

G. Avoiding perturbative crossings

An important slowdown mechanism we already alluded to in Sec. VII.C is due to anticrossings very close to the end of the evolution that can result in an extremely small minimum gap. These crossings are often referred to as perturbative, because a perturbative expansion back in time from the final Hamiltonian [e.g., perturbation theory in Γ for Eq. (142)] yields perturbed states that cross in energy very close to where the exact eigenstates anticross, with a gap that is exponentially small in the Hamming weight of the unperturbed crossing states (Amin and Choi, 2009) [shown there in the context of the weighted maximum independent set problem; see also Foini, Semerjian, and Zamponi (2010) and Farhi et al. (2011)]. This

---

[39]This addition does not result in a truly nonstoquastic Hamiltonian; there exists a local unitary transformation that makes the Hamiltonian stoquastic. Specifically, rotate σ^z to σ^x. In this new basis, the Hamiltonian is stoquastic.

problem of perturbative crossings was demonstrated for the NP-complete Exact Cover problem (recall Sec. VI.F.2) in Altshuler, Krovi, and Roland (2010), who related the mechanism of exponentially small spectral gaps to Anderson localization of the eigenfunctions of $H(s)$ in the space of the solutions. They showed that the Hamming weight between such states can be $\Theta(n)$, which is clearly problematic for the adiabatic algorithm. It was also claimed by Altshuler, Krovi, and Roland (2010) that these anticrossings appear with high probability as the transverse field goes to zero; however, the latter claim did not survive a more accurate analysis that took into account the extreme value statistics of the energy levels: the exponential degeneracy of the ground state, which is a distinguishing feature of random NP-complete problems with a discrete spectrum (such as Exact Cover), dooms the proposed mechanism (Knysh and Smelyanskiy, 2010; Knysh, 2016).

Nevertheless, this does not rule out the occurrence of exponentially small gaps close to the end of the evolution. Furthermore, it is plausible that the mechanism for avoided-level crossings presented by Altshuler, Krovi, and Roland (2010) may not necessarily be restricted to the end of the evolution, but may occur throughout a many-body-localized phase (Laumann et al., 2015). In light of this we now discuss a rather general way to circumvent such perturbative crossings that differs from the random initial Hamiltonian approach presented in Sec. VII.C.

Using the NP-hard maximum independent set problem, it was shown that this problem occurs only for one particular implementation of the adiabatic algorithm, and different choices can avoid the problem (Choi, 2010). In fact, Dickson and Amin (2011) showed that there is always some choice of the initial and final Hamiltonians that avoids such perturbative crossings (note that this does not include non-perturbative crossings). Furthermore, this choice can be made efficiently, i.e., in polynomial time, space, and energy (Dickson, 2011), as we now summarize.

The idea of Dickson (2011) was to cause the ground state to diverge from all other states by changing the degeneracy of the spectrum of the final Hamiltonian, such that the ground state is the most degenerate, the first excited state less degenerate, up to the highest excited state, which will be the least degenerate. Consider an n-qubit Ising Hamiltonian of the form

$$H_1 = \sum_{i \in M}h_i\sigma_i^z + \sum_{\{i,j\} \in M}J_{ij}\sigma_i^z\sigma_j^z, \quad (204)$$

where $h_i, J_{ij} \in \{0, \pm 1\}$, and M specifies the nonzero terms, of which there are m. In order to simplify the analysis, assume that there are no single bit-flip degeneracies, meaning that there are no degenerate states that are Hamming distance one from each other. For each nonzero h_i term that the ground state satisfies, i.e., $h_i\sigma_i^z = -1$, add $a \geq 1$ ancilla qubits with an interaction of the form

$$H_h = \sum_{k=1}^{a}b(h_i\sigma_i^z + 1)(\sigma_{i_k}^z + 1)/2, \quad (205)$$

where $\{i_1, \ldots, i_a\} \in M_h$. This term vanishes when the term h_i is satisfied, regardless of the orientation of the a ancillas, whereas otherwise it gives an energy bn_1, where n_1 is the

number of ancillas pointing up. Note that when the term is unsatisfied, when all ancilla spins point down the energy cost is zero. This is important because we do not want to change the energy of the ground state configuration.

Similarly, for each (nonzero) J_{ij}, also add a ancillas with the following interaction term:

$$H_J = \sum_{k=1}^{a} (J_{ij}\sigma_i^z\sigma_j^z + 1)(\sigma_{(ij)_k}^z + 1)/2, \quad (206)$$

where $\{ij_1, ..., ij_a\} \in M_J$. This introduces 3-local terms; it is possible to use 2-local terms to achieve the same result at the expense of introducing an additional ancilla for each term [see the Appendix of Dickson (2011) for details].

The spectrum of the new Hamiltonian (with ma additional ancilla qubits) is the original spectrum when all ancilla qubits point down, and all the new energy states correspond to flips of the ancilla qubits, with increased energy. Note that this means that no new local minima were introduced. Now consider the following adiabatic algorithm Hamiltonian:

$$H = \lambda H_0 + H_1', \qquad H_1' = H_1 + H_h + H_J, \quad (207)$$

with λ decreased from ∞ (proportional to abm suffices) to 0, and where the initial Hamiltonian includes transverse fields on the ancilla qubits:

$$H_0 = -\sum_{i \in M \cup M_h \cup M_J} \sigma_i^x. \quad (208)$$

Consider a nondegenerate classical state α with energy E_α under the action of H_1. It becomes degenerate under the action of H_1'. Let $|\alpha\rangle$ denote the uniform superposition over all these degenerate states with energy E_α. Introducing $\lambda > 0$ breaks the degeneracy, and from first-order degenerate perturbation theory (see Appendix E.1) the state $|\alpha\rangle$ is the new lowest energy eigenstate within the subspace spanned by the unperturbed degenerate states with energy E_α. The correction to its energy is $E_{|\alpha\rangle} = E_\alpha + \lambda E_{|\alpha\rangle}^{(1)} + \cdots$, where

$$E_\alpha = -(\text{No. of terms in } H_1' \text{ satisfied by } \alpha)$$
$$+ (\text{No. of terms in } H_1' \text{ unsatisfied by } \alpha)$$
$$= -2(\text{No. of terms in } H_1' \text{ satisfied by } \alpha) + m \quad (209)$$

(recall that $h_i, J_{ij} \in \{0, \pm 1\}$), and

$$E_{|\alpha\rangle}^{(1)} = \langle\alpha|H_0|\alpha\rangle$$
$$= -a(\text{No. of terms in } H_1' \text{ satisfied by } \alpha)$$
$$= \frac{a}{2}(E_\alpha - m). \quad (210)$$

Note that in E_α the contribution is entirely due to H_1, while in $E_{|\alpha\rangle}^{(1)}$ the contribution is entirely due to $H_h + H_J$.

Taking $a = b = n^2$, it can be shown that higher-order corrections do not depend on a, and hence the first-order correction dominates the behavior. Therefore, it is clear that a state $|\alpha\rangle$ with a lower (final) energy than a state $|\beta\rangle$ has a larger

negative slope (first-order perturbation energy correction). Therefore, the states $|\alpha\rangle$ and $|\beta\rangle$ grow farther apart for $\lambda > 0$ according to first-order perturbation theory. This means that the perturbative crossing is avoided.

This method works in general for the problem of finding the ground state of an arbitrarily connected Ising model with local fields and is fully stoquastic. Thus, all NP-complete problems can be attacked using StoqAQC without encountering perturbative crossings. Of course, this does not prove that StoqAQC can solve NP-complete problems in polynomial time. However, it does mean that proving otherwise requires identifying some effect other than perturbative crossings that unavoidably results in exponentially long adiabatic run times.

H. Evolving nonadiabatically

Our discussion so far has been restricted to adiabatic evolutions, where the minimum gap controls the efficiency of the quantum algorithm. However, as we have seen with the glued-trees problem in Sec. III.D, the quantum evolution can take advantage of the presence of two avoided-level crossings (and their associated exponentially small gaps) to leave and return to the ground state with high probability in polynomial time, whereas an adiabatic evolution would have required exponential time. Setting aside the field of an open-system AQC where relaxation can play a beneficial role in returning the computation to the ground state [the subject of a separate review (Albash and Lidar, 2018)], this is one among several cases where nonadiabatic, i.e., diabatic evolution enhances the performance of a quantum algorithm based on Hamiltonian computation. Another example was given by Crosson et al. (2014) [see also Hormozi et al. (2017)] where it was observed that evolving rapidly (as well as starting from excited states) increased the success probability on the hardest instances of randomly generated $n = 20$ MAX-2-SAT instances with a unique ground state. When evolving rapidly, population leaks into the first excited state before the avoided-level crossing and then returns to the ground state after the avoided-level crossing. An instance of this behavior is shown in Fig. 11.

A similar result was observed by Muthukrishnan, Albash, and Lidar (2016) for a large class of perturbed Hamming weight problems (recall Sec. VI.E.1), but with the difference that the rapid evolution diabatically pushes the population to higher excited states and then returns to the ground state through a series of avoided-level crossings, a phenomenon called "diabatic cascade."

These results raise the question of whether adiabatic evolution is in fact the most efficient choice for running a quantum adiabatic algorithm. After all the goal is to find the ground state once, with the highest probability and in the shortest amount of time. Therefore, rather than maximizing the probability by increasing the evolution time t_f, we can instead use many rapid repetitions of the algorithm to simultaneously shorten t_f and increase the success probability. Let $p_S(t_f)$ denote the single-run success probability of the algorithm with evolution time t_f. The probability of failing to find the ground state after R independent repetitions is $(1 - p_S)^R$, so the probability of succeeding at least once is $1 - (1 - p_S)^R$, which we set equal to the desired probability
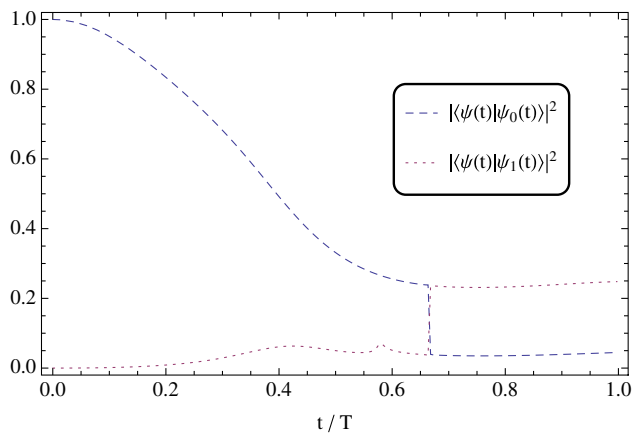
FIG. 11. Overlap squared of the evolved wave function $|\psi(t)\rangle$ and the instantaneous ground state $|\psi_0(t)\rangle$ and first excited state $|\psi_1(t)\rangle$ for an instance of MAX-2-SAT with $n = 20$ and with a total time $T = 10$. Because of the rapid evolution, population leaks out of the ground state and hence the decrease in the ground state population. There is an avoided-level crossing at approximately $t/T = 0.65$, where the population between the ground state and first excited state is effectively swapped. Therefore, if more substantial leaking into the first excited state occurs, this will lead to an increase in probability of finding the ground state at the end of the evolution. From Crosson $et~al.$, 2014.

p_d. The trade-off between success probability and run time is therefore well captured by the time-to-solution (TTS) metric, which measures the time required to find the ground state at least once with probability p_d (typically taken to be 99%):

$$\mathrm{TTS}(t_f) = t_f \frac{\ln(1 - p_d)}{\ln[1 - p_S(t_f)]}. \qquad (211)$$

Other metrics exist that quantify this trade-off, e.g., without insisting on finding the ground state (King $et~al.$, 2015), or that make use of optimal stopping theory and assign a cost to each run (Vinci and Lidar, 2016).

For $p_S \lesssim 1$ (close to the adiabatic limit), only a single (or few) repetitions of the algorithm are necessary and the TTS scales linearly with t_f. As t_f is lowered, the success probability typically decreases and more repetitions are necessary, but the TTS may in fact be lower because of the smaller t_f value. The optimal t_f for the algorithm minimizes the TTS and is defined as

$$\mathrm{TTS}_{\mathrm{opt}} = \min_{t_f > 0} \mathrm{TTS}(t_f). \qquad (212)$$

Benchmarking of algorithms then proceeds as follows. For a specific class of problem instances of varying sizes n, $\mathrm{TTS}_{\mathrm{opt}}$ is calculated for each size n. The scaling of the algorithm with n is then determined from the scaling with n of $\mathrm{TTS}_{\mathrm{opt}}$, as done by Boixo $et~al.$ (2014).

One benefit of this approach is in obtaining a quantum scaling advantage over specific classical algorithms. For example, the constant gap perturbed Hamming weight oracular problems (Reichardt, 2004) (see Sec. VI.D.1) and the spike problem of Farhi, Goldstone, and Gutmann (2002a) (see Sec. VI.E.1) with a polynomially closing quantum gap can be solved in $O(1)$ time using a classical algorithm. However, QA exhibits a scaling advantage over SA for these problems in the sense that QA offers a TTS that scales better than SA with single-spin updates (Muthukrishnan, Albash, and Lidar, 2016).

VIII. OUTLOOK AND CHALLENGES

Adiabatic quantum computing has blossomed from a speculative alternative approach for solving optimization problems to a formidable alternative to other universal models of quantum computing with deep connections to both classical and quantum complexity theory and condensed matter physics.

In this review we have given an account of most of the major theoretical developments in the field. Of course, some omissions were inevitable. For example, a potentially promising application of AQC is in quantum chemistry, where the calculation of molecular energies can be formulated in terms of a second-quantized fermionic Hamiltonian that is mapped, via a generalized Jordan-Wigner transformation (Jordan and Wigner, 1928; Bravyi and Kitaev, 2002), to a nonstoquastic qubit Hamiltonian (Aspuru-Guzik $et~al.$, 2005; Seeley, Richard, and Love, 2012). This mapping generates k-local interactions, but perturbative gadgets can be used to reduce the problem to only 2-local interactions (Babbush, Love, and Aspuru-Guzik, 2014). The ground state of the mapped Hamiltonian can then be prepared using adiabatic evolution followed by appropriate measurements to determine the energy spectrum. However, the scaling of the minimum gap for such a preparation procedure is not known, and hence this is an example of AQC with a nonstoquastic Hamiltonian for which it is unknown whether a quantum speedup is possible. A variety of other interesting AQC results with an unknown speedup, and which we did not have the space to review here in detail, can be found in Kurihara, Tanaka, and Miyashita (2014), Rajak and Chakrabarti (2014), Sato $et~al.$ (2014), Hashizume $et~al.$ (2015), Inack and Pilati (2015), O'Gorman $et~al.$ (2015), Behrman and Steck (2016), Cao $et~al.$ (2016), Chancellor (2016), Dulny and Kim (2016), Durkin (2016), Goto (2016), Karimi and Rosenberg (2016), Karimi and Ronagh (2016), Miyahara and Tsumura (2016), Raymond, Yarkoni, and Andriyash (2016), Rosenberg $et~al.$ (2016), and Santra, Shehab, and Balu (2016).

Moreover, to make the review comprehensive and detailed enough to be self-contained, we focused only on the closed-system setting, thus completely ignoring the important problem of AQC in open systems, with the associated questions of error correction and fault tolerance. We also left out the experimental work on AQC and quantum annealing. These important topics are the subject of a separate review (Albash and Lidar, 2018).

Because of the prominence of stoquastic Hamiltonians in the body of work on AQC, we coined a new term, StoqAQC, which is roughly what was meant when the term "quantum adiabatic algorithm" was first introduced. Correspondingly, we devoted a substantial part of this review to StoqAQC, despite the fact that there are indications that this model of computation may not be more powerful than classical computing. Its prominence is explained by the fact that it is easier to analyze than universal AQC, which requires nonstoquastic terms, and by the fact that it is easier to implement experimentally (Bunyk $et~al.$, 2014;

Weber *et al.*, 2017). The relatively short history of AQC has witnessed a battle between attempts to show that StoqAQC fails to deliver quantum speedups and corresponding refutations by clever tweaks. To put this and other results discussed in perspective, we conclude with a list of ten key theoretical challenges for the field of AQC:

(1) Prove or disprove that StoqAQC is classically efficiently simulatable.
(2) Find an NP-hard optimization for which AQC gives a quantum speedup in the worst case.
(3) Find a class of nonoracular, physically realizable optimization problems for which AQC gives a quantum speedup.
(4) Identify a subset of nonstoquastic Hamiltonians for which ground state preparation can be done efficiently using adiabatic evolution.
(5) Formulate every quantum algorithm that gives a speedup in the circuit model natively as an AQC algorithm (i.e., directly without using perturbative gadgets).
(6) Find a problem that can be solved with a quantum speedup using AQC that was not previously known from other models of quantum computing.
(7) Give a way to decide whether adiabatic evolution gives rise to a stronger or weaker speedup than nonadiabatic (diabatic) evolution for a given problem.
(8) Predict the optimal adiabatic schedule for a given problem without *a priori* knowledge of the size and/or position of its spectral gap.
(9) Prove or disprove that tunneling can generate a (scaling, not prefactor) quantum speedup in AQC.
(10) Establish the relation between entanglement and quantum speedups using AQC.

Solving these problems will likely keep researchers busy for years to come, require interdisciplinary collaborations, and will significantly advance our understanding of AQC. We hope that this review will catalyze new and productive approaches, enhancing our repertoire of algorithms that give rise to quantum speedups from the unique perspective of AQC.

ACKNOWLEDGMENTS

APPENDIX A: TECHNICAL CALCULATIONS

1. Upper bound on the adiabatic path length L

Right below Eq. (14) we claimed that an upper bound on L is $\max_s \|\dot{H}(s)\|/\Delta$. To see this differentiate the eigenstate equation $H|\varepsilon_a\rangle = \varepsilon_a|\varepsilon_a\rangle$ for the normalized instantaneous eigenstate $|\varepsilon_a\rangle$ and inner multiply by $\langle\varepsilon_b|$, with $b \neq a$, to get $(\varepsilon_a - \varepsilon_b)\langle\varepsilon_b|\dot{\varepsilon}_a\rangle = \langle\varepsilon_b|\dot{H}|\varepsilon_a\rangle$. Let $\Delta_{ba} = \varepsilon_b - \varepsilon_a$ and $\Delta_a = \min_b\min_s\Delta_{ba}(s)$. Using our phase choice

$$|\dot{\varepsilon}_a\rangle = \sum_b |\varepsilon_b\rangle\langle\varepsilon_b|\dot{\varepsilon}_a\rangle = \sum_{b\neq a} |\varepsilon_b\rangle\langle\varepsilon_b|\dot{\varepsilon}_a\rangle$$
$$= -\sum_{b\neq a} |\varepsilon_b\rangle\langle\varepsilon_b|\dot{H}|\varepsilon_a\rangle/\Delta_{ba}. \tag{A1}$$

Thus

$$\||\dot{\varepsilon}_a\rangle\| \leq \frac{1}{\Delta_a}\left\|\sum_{b\neq a}|\varepsilon_b\rangle\langle\varepsilon_b|\dot{H}|\varepsilon_a\rangle\right\|$$
$$\leq \frac{1}{\Delta_a}\left\|\sum_{b\neq a}|\varepsilon_b\rangle\langle\varepsilon_b|\right\|\|\dot{H}|\varepsilon_a\rangle\| \leq \frac{1}{\Delta_a}\|\dot{H}\|, \tag{A2}$$

where in the last equality we used the definition of the operator norm and the fact that $\sum_{b\neq a}|\varepsilon_b\rangle\langle\varepsilon_b|$ is a projector. Integration just multiplies by 1.

2. Proof of the inequality given in Eq. (25)

Note that

$$\|\partial_s H[A(s)]\| = |\partial_s A|\|H_1 - H_0\| \leq 2|\partial_s A|, \tag{A3a}$$

$$\|\partial_s^2 H[A(s)]\| \leq 2|\partial_s^2 A| \tag{A3b}$$

for the interpolating Hamiltonian (16). Also note that

$$\partial_s^2 A(s) = cp\Delta^{p-1}[A(s)]\frac{d\Delta}{dA}\partial_s A(s) \tag{A4a}$$

$$= c^2 p \frac{d\Delta}{dA}\Delta^{2p-1}[A(s)]. \tag{A4b}$$

Thus,

$$\int_0^1 \left(\frac{\|\partial_s^2 H[A(s)]\|}{\Delta^2[A(s)]} + \frac{\|\partial_s H[A(s)]\|^2}{\Delta^3[A(s)]}\right)ds$$
$$\leq 2\int_0^1 \left(\frac{|\partial_s^2 A|}{\Delta^2[A(s)]} + \frac{2|\partial_s A|^2}{\Delta^3[A(s)]}\right)ds \tag{A5a}$$
$$= \int_0^1 2c^2\left(p\frac{d\Delta}{dA}+2\right)\Delta^{2p-3}[A(s)]ds \tag{A5b}$$
$$= 2c\int_0^1 \left(p\frac{d\Delta}{du}+2\right)\Delta^{p-3}(u)du \tag{A5c}$$
$$= 4c\int_0^1 \Delta^{p-3}(u)du, \tag{A5d}$$

where in Eq. (A5c) we used the change of variables $u = A(s)$, so that $ds = du/\partial_s A = du/c\Delta^p(u)$, and in Eq. (A5d) we used $B(p) \equiv 2c \int_0^1 \Delta^{p-3} d\Delta = 0$, since $B(2) = 2c \ln[\Delta(1)/\Delta(0)] = 0$ and $B(p \neq 2) = [2c/(p-2)][\Delta^{p-2}(1) - \Delta^{p-2}(0)] = 0$, due to the boundary conditions $\Delta(0) = \Delta(1) = 1$ [Eq. (18a)].

APPENDIX B: A LOWER BOUND FOR THE ADIABATIC GROVER SEARCH PROBLEM

Here we show that there is no schedule that gives a better scaling for the adiabatic Grover problem than the one discussed in Sec. III.A.2, resulting in a quadratic quantum speedup. The argument is due to Roland and Cerf (2002), which in turn is based on the general Hamiltonian quantum computation argument by Farhi and Gutmann (1998).

To show this, consider two different searches, one for m and another for m'. We do not allow the schedule to depend on m, i.e., the same schedule must apply to all marked states. Let us denote the states for each at the end of the algorithm by $|\psi_m(t_f)\rangle$ and $|\psi_{m'}(t_f)\rangle$. In order to be able to distinguish if the search gave m or m', we must require that $|\psi_m(t_f)\rangle$ and $|\psi_{m'}(t_f)\rangle$ are sufficiently different. Let us define the distance (or infidelity)

$$D_{mm'}(t) \equiv 1 - |\langle \psi_m(t)|\psi_{m'}(t)\rangle|^2 \qquad (B1)$$

[note that $D_{mm'}(0) = D_{mm}(t) = 0$] and demand that

$$D_{mm'}(t_f) \geq \epsilon, \qquad m \neq m'. \qquad (B2)$$

First, we have a lower bound on the sum

$$\sum_{m,m'} D_{mm'}(t_f) = \sum_{m \neq m'} D_{mm'}(t_f)$$
$$\geq \sum_{m \neq m'} \epsilon = N(N-1)\epsilon. \qquad (B3)$$

Next, let us find an upper bound on the sum. We write the Hamiltonian (16) explicitly as $H(t) = \mathbb{1} - [1 - A(t)]|\phi\rangle \times \langle\phi| + H_{1m}(t)$, where $H_{1m}(t) = -A(t)|m\rangle\langle m|$.[40] Then

$$\frac{d}{dt} D_{mm'}(t) = 2\Im[\langle \psi_m|(H_{1m} - H_{1m'})|\psi_{m'}\rangle\langle\psi_{m'}|\psi_m\rangle]$$
$$\leq 2|\langle \psi_m|(H_{1m} - H_{1m'})|\psi_{m'}\rangle\langle\psi_{m'}|\psi_m\rangle|$$
$$\leq 2|\langle \psi_m|H_{1m}|\psi_{m'}\rangle| + 2|\langle\psi_m|H_{1m'}|\psi_{m'}\rangle|. \qquad (B4)$$

Let us now sum over all m and m':

---

[40]Optimality applies to arbitrary driving Hamiltonians. Hence the lower bound holds more generally and does in fact not require the initial Hamiltonian to be a projector onto the uniform superposition as we have done here for simplicity.

$$\sum_{m,m'} \frac{d}{dt} D_{mm'}(t) \leq 4\sum_{m,m'} |\langle \psi_m|H_{1m}|\psi_{m'}\rangle|$$
$$\leq 4\sum_{m,m'} \|H_{1m}|\psi_{m'}\rangle\|\,\||\psi_m\rangle\| = 4\sum_{m,m'} \|H_{1m}|\psi_{m'}\rangle\|, \qquad (B5)$$

where we first used the fact that under the sum the two terms in the last line of Eq. (B4) are identical, and then we used the Cauchy-Schwartz inequality ($|\langle x|y\rangle| \leq \|x\|\|y\|$). Now we note that

$$\sum_m \|H_{1m}|\psi_{m'}\rangle\|^2 = \sum_m \langle\psi_{m'}|H_{1m}H_{1m}|\psi_{m'}\rangle$$
$$= A^2(t) \sum_m \langle\psi_{m'}|m\rangle\langle m|\psi_{m'}\rangle = A^2(t), \quad (B6)$$

so that

$$\left(\sum_m \|H_{1m}|\psi_{m'}\rangle\|\right)^2 = (\vec{x}\cdot\vec{y})^2 \leq (\vec{x}\cdot\vec{x})(\vec{y}\cdot\vec{y}) = NA^2(t), \qquad (B7)$$

where $\vec{x} = (\|H_{11}|\psi_{m'}\rangle\|, \|H_{12}|\psi_{m'}\rangle\|, ..., \|H_{1N}|\psi_{m'}\rangle\|)$ and $\vec{y} = (1, 1, ..., 1)$. Therefore, we have

$$\sum_{m,m'} \frac{d}{dt} D_{mm'}(t) \leq 4\sum_{m,m'} \|H_{1m}|\psi_{m'}\rangle\|$$
$$\leq 4\sum_{m'} \sqrt{N} A(t) = 4N\sqrt{N}A(t). \qquad (B8)$$

If we integrate both sides, we have

$$\sum_{m,m'} D_{mm'}(t_f) \leq 4N\sqrt{N} \int_0^{t_f} A(t)dt \leq 4N\sqrt{N}t_f. \qquad (B9)$$

Combining this with Eq. (B3), we have $N(N-1)\epsilon \leq 4N\sqrt{N}t_f$, and hence

$$t_f \geq \frac{\epsilon}{4} \frac{N-1}{\sqrt{N}}, \qquad (B10)$$

so that the computation must last a minimum time of $O(\sqrt{N})$ if the schedule is to be agnostic about the identity of the marked state. Therefore, the solution using the locally optimized schedule is asymptotically optimal.

APPENDIX C: TECHNICAL DETAILS FOR THE PROOF OF THE UNIVERSALITY OF AQC USING THE HISTORY STATE CONSTRUCTION

We add details to the proof sketch given in Sec. IV.B.

1. $|\gamma(0)\rangle$ is the ground state of H_{init}

Let us first check that $|\gamma(0)\rangle$ is the ground state of H_{init} with eigenvalue 0. Note that H_{init} is a sum of projectors, so it is positive semidefinite. Therefore if we find a state with energy 0, then it is definitely a ground state. The all-zero clock state is annihilated by H_c, H_{input},

and $H_{\text{c-init}}$, so we have $H_{\text{init}}|\gamma(0)\rangle = 0$, i.e., $|\gamma(0)\rangle$ is a ground state of H_{input}. We show later that it is a unique ground state.

2. $|\eta\rangle$ is a ground state of H_{final}

Next we check that $|\eta\rangle$ is a ground state of H_{final} with eigenvalue 0. First we show that H_{final} is positive semidefinite. We already know that H_{input} and H_c are positive semidefinite, so we only need to show this to be the case for the H_ℓ's. This follows since it is easily checked that $H_\ell = H_\ell^\dagger = (1/2)H_\ell^2$, so that

$$\langle X|H_\ell|X\rangle = \tfrac{1}{2}\langle X|H_\ell^\dagger H_\ell|X\rangle = \tfrac{1}{2}\||H_\ell|X\rangle\|^2 \geq 0.$$

Thus $H(s)$ is positive semidefinite, since it is a sum of positive semidefinite terms.

Let us now check that H_{final} annihilates $|\eta\rangle$. First, because $|\eta\rangle$ involves only legal clock states, it is annihilated by H_c. Next,

$$H_{\text{input}}|\eta\rangle = H_{\text{input}} \frac{1}{\sqrt{L+1}}|\alpha(0)\rangle \otimes |0^L\rangle_c = 0. \quad (C1)$$

Finally, note that the only nonzero terms in $\sum_{\ell=0}^{L} H_\ell|\eta\rangle$ are of the form

$$H_\ell|\alpha(\ell-1)\rangle \otimes |1^{\ell-1}0^{L-\ell+1}\rangle_c$$
$$= |\alpha(\ell-1)\rangle \otimes |1^{\ell-1}0^{L-\ell+1}\rangle_c - |\alpha(\ell)\rangle \otimes |1^\ell 0^{L-\ell}\rangle_c,$$
$$(C2a)$$

$$H_\ell|\alpha(\ell)\rangle \otimes |1^\ell 0^{L-\ell}\rangle_c$$
$$= -|\alpha(\ell-1)\rangle \otimes |1^{\ell-1}0^{L-\ell+1}\rangle_c + |\alpha(\ell)\rangle \otimes |1^\ell 0^{L-\ell}\rangle_c,$$
$$(C2b)$$

which cancel. Therefore, $|\eta\rangle$ has eigenvalue 0 and is a ground state of H_{final}.

3. Gap bound in the space spanned by $\{|\gamma(\ell)\rangle\}_{\ell=0}^{L}$

Let $\mathcal{S}_0$ be the space spanned by $\{|\gamma(\ell)\rangle\}_{\ell=0}^{L}$. Let us show that $H(s)$ acting on any state in $\mathcal{S}_0$ keeps it in this subspace:

$$H_c|\gamma(\ell)\rangle = 0, \quad (C3a)$$

$$H_{\text{input}}|\gamma(\ell)\rangle = 0, \quad (C3b)$$

$$H_{\text{c-init}}|\gamma(\ell)\rangle = \begin{cases} 0, & \ell = 0, \\ |\gamma(\ell)\rangle, & \ell \neq 0, \end{cases} \quad (C3c)$$

$$H_\ell|\gamma(\ell')\rangle = \delta_{\ell',\ell}[|\gamma(\ell-1)\rangle - |\gamma(\ell)\rangle] + \delta_{\ell',\ell}[|\gamma(\ell)\rangle - |\gamma(\ell-1)\rangle]. \quad (C3d)$$

Since the initial state $|\gamma(0)\rangle \in \mathcal{S}_0$, the dynamics generated by $H(s)$ keep the state in $\mathcal{S}_0$. Thus, it is sufficient to bound the gap in this subspace. In the basis given by $\{|\gamma(\ell)\rangle\}_{\ell=0}^{L}$, we can write an $(L+1) \times (L+1)$ matrix representation of the Hamiltonian in the $\mathcal{S}_0$ subspace, which using Eq. (C3) is

$$H_{\mathcal{S}_0}(s) = (1-s) \begin{pmatrix} 0 & 0 & 0 & \cdots & & & \\ 0 & 1 & 0 & & & & \\ & & 1 & & & & \\ & & & \ddots & & & \\ & & & & 1 & 0 & \\ & & & & & 1 & \end{pmatrix}$$
$$+ s \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} & 0 & 0 & \cdots & 0 \\ -\frac{1}{2} & 1 & -\frac{1}{2} & & & \\ 0 & -\frac{1}{2} & 1 & -\frac{1}{2} & & \\ \vdots & & \ddots & \ddots & \ddots & \\ & & & -\frac{1}{2} & 1 & -\frac{1}{2} \\ 0 & & & & -\frac{1}{2} & \frac{1}{2} \end{pmatrix}. \quad (C4)$$

a. Bound for $s < 1/3$

Let us first bound the gap for $s < 1/3$. The Gerschgorin circle theorem states (Gershgorin, 1931):

Let A be any matrix with entries a_{ij}. Consider the disk D_i (for $1 \leq i \leq n$) in the complex plane defined as $D_i = \{z| |z - a_{ii}| \leq \sum_{j\neq i} |a_{ij}|\}$. Then the eigenvalues of A are contained in $\cup_i D_i$ and any connected component of $\cup_i D_i$ contains as many eigenvalues of A as the number of disks that form the component.

Consider the cases $i = 1$, $i = L+1$, and $i \neq 1, L+1$ separately. Note that when $s < 1/3$:

$$[H_{\mathcal{S}_0}(s)]_{11} = \frac{1}{2}s < \frac{1}{6}, \qquad \sum_{j\neq 1}|a_{1j}| = \frac{1}{2}s < \frac{1}{6}, \quad (C5a)$$

$$[H_{\mathcal{S}_0}(s)]_{L+1,L+1} = 1 - \frac{1}{2}s > 5/6$$
$$\sum_{j\neq L+1}|a_{L+1,j}| = s < \frac{1}{3}, \quad (C5b)$$

$$[H_{\mathcal{S}_0}(s)]_{ii} = 1, \qquad i \neq 1, L+1$$
$$\sum_{j\neq i}|a_{1j}| = \frac{1}{2}s < \frac{1}{3}. \quad (C5c)$$

Therefore, we can identify a disk D_1 centered at $z \leq 1/6$ on the real line with radius $\leq 1/6$. The closest possible disk to it which does not overlap it is the disk D_{L+1} centered at $z \geq 5/6$ with a radius $\leq 1/3$. Therefore, since no disks intersect D_1, it covers the smallest values on the real line, and it follows that the ground state for $s < 1/3$ is unique. Furthermore, we learned that the minimum gap is a constant of at least $1/6$, since that is the closest distance between D_1 and D_{L+1}. (This also proves that $|\gamma(0)\rangle$ is the unique ground state at $s = 0$.)

b. Bound for $s \geq 1/3$

Now let $s \geq 1/3$ and consider the matrix representation of $G(s) \equiv \mathbb{1} - H_{\mathcal{S}_0}(s)$ in the same basis:

$$G(s) = \begin{pmatrix} 1-\frac{1}{2}s & \frac{1}{2}s & & & & \\ \frac{1}{2}s & 0 & \frac{1}{2}s & & & \\ & \ddots & \ddots & \ddots & & \\ & & & 0 & \frac{1}{2}s & \\ & & & \frac{1}{2}s & \frac{1}{2}s \end{pmatrix}. \qquad (C6)$$

This matrix is Hermitian and has all non-negative real entries for $0 < s \leq 1$. Note that increasing powers of $G(s)$ fill more of the matrix, and $G(s)^{L+1}$ has all positive entries for $0 < s \leq 1$. We can thus invoke Perron's theorem:

Let G be a Hermitian matrix with real non-negative entries. If there exists a finite k such that all entries of G^k are positive, then G's largest eigenvalue is positive and all other eigenvalues are strictly smaller in absolute value. The eigenvector corresponding to the largest eigenvalue is unique, and all its entries are positive.

Therefore, by Perron's theorem, $G(s)$'s largest eigenvalue μ must be positive, and the associated unique eigenvector $\vec{\alpha} = (\alpha_1, ..., \alpha_{L+1})$ has $\alpha_i > 0$. Let us use this to define a matrix P with entries

$$P_{ij} = \frac{\alpha_j}{\mu \alpha_i} G_{ij} \geq 0,$$

such that

$$\sum_j P_{ij} = \frac{1}{\mu \alpha_i} \sum_j G_{ij} \alpha_j = 1, \qquad (C7)$$

where we used the fact that $\vec{\alpha}$ is an eigenvector of G with eigenvalue μ. Thus P is a stochastic matrix (it has only non-negative entries and its rows sum to 1). Now note that if $(\alpha_1 v_1, ..., \alpha_{L+1} v_{L+1})$ is a left eigenvector of P with eigenvalue ν/μ, then $(v_1, ..., v_{L+1})$ is an eigenvector of G with eigenvalue ν:

$$\frac{\nu}{\mu} \alpha_j v_j = \sum_i \alpha_i v_i P_{ij} = \sum_i v_i \frac{\alpha_j}{\mu} G_{ij} = \frac{\alpha_j}{\mu} \sum_i G_{ji} v_i$$

$$\Rightarrow \nu v_j = \sum_i G_{ji} v_i. \qquad (C8)$$

It is straightforward to check that the reverse also holds: if $(v_1, ..., v_{L+1})$ is an eigenvector of G with eigenvalue ν, then $(\alpha_1 v_1, ..., \alpha_{L+1} v_{L+1})$ is a left eigenvector of P with eigenvalue ν/μ. By taking $\vec{v} = \vec{\alpha}$, which corresponds to the largest eigenvalue $(\nu = \mu)$ of G, it then follows that $\vec{\alpha^2} = (\alpha_1^2, ..., \alpha_{L+1}^2)$ is a left eigenvector of P with the maximal eigenvalue 1. If we normalize $\vec{\alpha^2}$, i.e., define

$$\vec{\Pi} = \frac{1}{Z}(\alpha_1^2, ..., \alpha_{L+1}^2), \qquad Z = \sum_i \alpha_i^2, \qquad (C9)$$

then $\vec{\Pi}$ is the limiting distribution of P, i.e., $P\vec{\Pi} = \vec{\Pi}$. We can then relate the energy gap between the highest and second highest eigenvalue (let us denote it by δ/μ) of P to the energy gap between the ground state energy of H (given by $1 - \mu = \lambda$) and the first excited state (given by $1 - \delta$)

$$\Delta_{\text{largest}}(P) = 1 - \frac{\delta}{\mu} = \frac{\mu - \delta}{\mu} = \frac{\Delta(H_{\mathcal{S}_0})}{1 - \lambda}, \qquad (C10)$$

where "largest" denotes the gap from the largest eigenvalue of P to the next largest eigenvalue. We bound the gap of P and hence of $H(s)$. Let us define a nonempty set $\mathcal{B} \subseteq \{1, 2, ..., L + 1\}$ satisfying $\sum_{i \in \mathcal{B}} \Pi_i \leq 1/2$, where Π_i are the entries of $\vec{\Pi}$. Then the conductance $\varphi(P)$ is defined as

$$\varphi(P) = \min_{\mathcal{B}} \frac{F(\mathcal{B})}{\Pi(\mathcal{B})}, \qquad (C11)$$

where

$$F(\mathcal{B}) = \sum_{i \in \mathcal{B}} \sum_{j \notin \mathcal{B}} \Pi_i P_{ij}, \qquad (C12a)$$

$$\Pi(\mathcal{B}) = \sum_{i \in \mathcal{B}} \Pi_i. \qquad (C12b)$$

The conductance bound (Sinclair and Jerrum, 1989) then states that

$$\Delta_{\text{largest}}(P) \geq \frac{1}{2}\varphi(P)^2. \qquad (C13)$$

To use the result of the conductance bound, we show that the ground state of $H(s)$ [and hence the eigenstate associated with the largest eigenvalue of $G(s)$] is monotone, i.e., that $\alpha_1 \geq \alpha_2 \geq \cdots \geq \alpha_{L+1} \geq 0$. The case $s = 0$ is obvious, so consider $s > 0$. First note that $G(s)$ applied to a monotone vector generates a monotone vector, i.e., $G(s)$ preserves monotonicity. To see this consider $G(s)\vec{v} = \vec{w}$ with $\vec{v}$ monotone. The components of $\vec{w}$ are given by

$$w_1 = (1 - \frac{1}{2}s)v_1 + \frac{1}{2}sv_2, \qquad (C14a)$$

$$w_k = \frac{1}{2}sv_{k-1} + \frac{1}{2}sv_{k+1}, \qquad 2 \leq k \leq L, \qquad (C14b)$$

$$w_{L+1} = \frac{1}{2}sv_L + \frac{1}{2}sv_{L+1}. \qquad (C14c)$$

Therefore we have

$$w_1 - w_2 = (1 - s)v_1 + \frac{1}{2}s(v_2 - v_3), \qquad (C15a)$$

$$w_k - w_{k+1} = \frac{1}{2}s(v_1 - v_2 + v_3 - v_4), \qquad 2 \leq k \leq L - 1, \qquad (C15b)$$

$$w_L - w_{L+1} = \frac{1}{2}s(v_{L-1} - v_L), \qquad (C15c)$$

which clearly are all ≥ 0 by the monotonicity of $\vec{v}$ and $s \leq 1$. Therefore $\vec{w}$ is also monotone.

Recall that $G(s)$ is Hermitian, so it has an orthonormal set of eigenvectors $\{|v_i\rangle\}_{i=1}^{L+1}$ with eigenvalues μ_i, where $|v_1\rangle = \vec{\alpha}$ and $\mu_1 = \mu$.[41] Because these eigenvectors form a basis we can always find a set of coefficients $\{c_i\}_{i=1}^{L+1}$ such that

---

[41]Note that we abuse notation and mix kets with standard vectors here and also do not use transpose notation to distinguish column from row vectors.

$$\sum_i c_i|v_i\rangle = (1, ..., 1) = \vec{1}. \qquad (C16)$$

Then

$$\left(\frac{1}{\mu_1}G(s)\right)^k\sum_i c_i|v_i\rangle = \sum_i \left(\frac{\mu_i}{\mu_1}\right)^k c_i|v_i\rangle \qquad (C17a)$$

$$\Rightarrow \left(\frac{1}{\mu_1}G(s)\right)^k\vec{1}^T = \sum_i \left(\frac{\mu_i}{\mu_1}\right)^k\vec{1}^T. \qquad (C17b)$$

Using $|\mu_i| < \mu_1$ by Perron's theorem, we have from Eq. (C17a) that

$$\lim_{k\to\infty}\left(\frac{1}{\mu_1}G(s)\right)^k\sum_i c_i|v_i\rangle = c_1|v_1\rangle. \qquad (C18)$$

Since the quantities

$$\left(\frac{1}{\mu_1}G(s)\right)^k$$

(for $k \geq L+1$), $\sum_i c_i|v_i\rangle = \vec{1}$, and $|v_1\rangle = \vec{\alpha}$ are all positive, it follows that also $c_1 > 0$. Since $\vec{1}$ is monotone and $G(s)$ preserves monotonicity, we have finally established that $|v_1\rangle = \vec{\alpha}$ is monotone. This then implies that $\vec{\Pi}$ [Eq. (C9)] is monotone.

We are ready to calculate the conductance $\varphi(P)$. First consider the case where the first index (of $\vec{\Pi}$) is in the set $\mathcal{B}$, i.e., $1 \in \mathcal{B}$. Let k be the smallest index such that $k \in \mathcal{B}$ but $k+1 \notin \mathcal{B}$. (Note that from the form of P, only P_{11}, $P_{j,j+1}$, $P_{L+1,L+1}$ are nonzero.) Then we have the following for $F(\mathcal{B})$:

$$F(\mathcal{B}) = \sum_{i\in\mathcal{B},i\neq k}\sum_{j\notin\mathcal{B}}\Pi_i P_{ij} + \Pi_k P_{k,k+1} \geq \Pi_k P_{k,k+1}$$

$$= \Pi_k\frac{\sqrt{\Pi_{k+1}}}{\mu\sqrt{\Pi_k}}[G(s)]_{k,k+1} = \frac{\sqrt{\Pi_k\Pi_{k+1}}}{1-\lambda}[G(s)]_{k,k+1}$$

$$\geq \frac{\Pi_{k+1}}{1-\lambda}[G(s)]_{k,k+1}, \qquad (C19)$$

where the last inequality follows from the monotonicity of $\vec{\Pi}$. Because $0 < 1-\lambda \leq 1$, and $[G(s)]_{k,k+1} = (1/2)s \geq 1/6$ for $s \geq 1/3$, it follows that

$$F(\mathcal{B} = \{1, \text{others}, k\}) \geq \frac{\Pi_{k+1}}{6}. \qquad (C20)$$

Since by definition $\Pi(\mathcal{B}) \leq 1/2$, then $\Pi(\bar{\mathcal{B}}) \geq 1/2$ where $\bar{\mathcal{B}}$ is the complement of $\mathcal{B}$, but since the largest possible size of $\bar{\mathcal{B}}$ is L (recall that $1 \in \mathcal{B}$), it follows that $\Pi(\bar{\mathcal{B}}) \leq L\Pi_{k+1}$, so that $\Pi_{k+1} \geq 1/(2L)$, and hence

$$\frac{F(\mathcal{B} = \{1, \text{others}, k\})}{\Pi(\mathcal{B})} \geq \frac{1}{6L}. \qquad (C21)$$

Next consider the case where $1 \notin \mathcal{B}$. Now let k be the smallest index such that $k \notin \mathcal{B}$ but $k+1 \in \mathcal{B}$. Then

$$F(\mathcal{B}) = \sum_{i\in\mathcal{B},i\neq k+1}\sum_{j\notin\mathcal{B}}\Pi_i P_{ij} + \Pi_{k+1}P_{k+1,k} \qquad (C22a)$$

$$\geq \Pi_{k+1}P_{k+1,k} \geq \frac{\Pi_{k+1}}{6}. \qquad (C22b)$$

In this case, since the maximum size of $\mathcal{B}$ is L but it excludes the index 1, we have $\Pi(\mathcal{B}) \leq L\Pi_{k+1}$, so that $F(\mathcal{B}) \geq \Pi(\mathcal{B})/(6L)$. Therefore, we again find the condition (C21). Thus, by the conductance bound [Eq. (C13)]

$$\Delta(P) = \frac{\Delta(H_{\mathcal{S}_0})}{1-\lambda} \geq \frac{1}{2}\left(\frac{1}{6L}\right)^2. \qquad (C23)$$

Now since λ is the ground state of $H_{\mathcal{S}_0}$, for any state in $|v\rangle \in \mathcal{S}_0$, we must have $\langle v|H_{\mathcal{S}_0}|v\rangle \geq \lambda$. In particular,

$$\langle\gamma(0)|H_{\mathcal{S}_0}|\gamma(0)\rangle = \tfrac{1}{2}s \geq \lambda, \qquad (C24)$$

i.e., $\lambda \leq 1/2$. This finally yields Eq. (78).

4. Gap bound in the entire Hilbert space

Let us now go a step further and show how the global gap (i.e., not restricted to the $\mathcal{S}_0$ subspace) scales with L. Let $\mathcal{S}$ denote the subspace spanned by all legal clock states. The dimensions of this subspace are $\dim(\mathcal{S}) = (L+1)2^n$, since we have $L+1$ legal clock states and 2^n computational states. $H(s)$ acting on any state in $\mathcal{S}$ does not generate any illegal clock states, so for any $|v\rangle \in \mathcal{S}$ we have $H(s)|v\rangle \in \mathcal{S}$. Similarly, for any state in the orthogonal subspace $\mathcal{S}^\perp$, i.e., the subspace of illegal clock states, for any state $|v^\perp\rangle \in \mathcal{S}^\perp$, we have $H(s)|v^\perp\rangle \in \mathcal{S}^\perp$. Therefore, the eigenstates of $H(s)$ below either to $\mathcal{S}$ or to $\mathcal{S}^\perp$, and $H(s)$ is block diagonal with blocks $H_{\mathcal{S}}(s)$ and $H_{\mathcal{S}^\perp}(s)$ that can be diagonalized independently.

Let us first restrict to $H_{\mathcal{S}^\perp}(s)$. H_c penalizes all illegal clock states by at least one unit of energy and acts as the identity on the computational qubits. Therefore, it shifts the entire spectrum of $\mathcal{S}^\perp$ by at least one unit of energy. Since the remaining terms are positive semidefinite, they cannot lower the energy. Therefore, regardless of the form of the ground state in the subspace, it has an energy of at least one unit.

Let us now restrict to $H_{\mathcal{S}}(s)$ and define

$$|\gamma_j(\ell)\rangle = |\alpha_j(\ell)\rangle \otimes |1^\ell 0^{L-\ell}\rangle_c, \qquad (C25)$$

where $|\alpha_j(\ell)\rangle$ is the state of the circuit at time ℓ had the input state been given by the binary representation of j (e.g., if $j = 4$, the input configuration of the circuit would have been $|0^{n-3}1_30_20_1\rangle$). Note that $|\gamma_0(\ell)\rangle = |\gamma(\ell)\rangle$. Let $\mathcal{S}_j$ denote the space spanned by $\{|\gamma_j(\ell)\rangle\}_{\ell=0}^L$. Since $H_{\mathcal{S}}(s)$ cannot mix states with different j subindices (it can only propagate forward or backward in ℓ), $H_{\mathcal{S}}(s)$ is block diagonal in the subspaces $\mathcal{S}_j$. Therefore, we only need to find the minimum ground state energy of $H_{\mathcal{S}_{j>0}}(s)$ to determine the minimum gap from $H_{\mathcal{S}_0}(s)$.

In order to determine the ground state energy of $H_{\mathcal{S}_j}$, we can write

$$H_{\mathcal{S}_j}(s) = H_0(s) + H_{\mathcal{S}_j,\text{input}}, \qquad j > 0, \qquad \text{(C26)}$$

where $H_0(s)$ has exactly the same spectral properties as $H_{\mathcal{S}_0}$ except in the $\mathcal{S}_j$ subspace. The reason for this decomposition is because H_input is zero in $\mathcal{S}_0$ and hence is absent from $H_{\mathcal{S}_0}(s)$. Note that

$$H_{\mathcal{S}_j,\text{input}}|\gamma_j(\ell)\rangle = \begin{cases} k|\gamma_j(0)\rangle, & \ell = 0, \\ 0, & \ell > 0 \end{cases} \qquad \text{(C27)}$$

(recall that H_input projects onto the 0 clock state, which is why for $\ell > 0$ we have zero.) Therefore, in the basis $\{|\gamma_j(\ell)\rangle\}_{\ell=0}^L$, we can write the matrix representation of $H_{\mathcal{S}_j,\text{input}}$ as

$$H_{\mathcal{S}_j,\text{input}} = \begin{pmatrix} k & & & 0 \\ & 0 & & \\ & & \ddots & \\ 0 & & & 0 \end{pmatrix}, \qquad \text{(C28)}$$

where $k \geq 1$ denotes the number of 1's in the binary representation of $j > 0$. In particular, note that it is diagonal in this basis. We now use the geometrical lemma (Kitaev, Shen, and Vyalyi, 2000; Aharonov and Naveh, 2002):

Lemma 4: Geometrical lemma. Let H_1 and H_2 be two Hamiltonians with ground state energies g_1 and g_2, respectively. Both Hamiltonians have a ground state energy gap to the first excited state that is larger than Λ. Let the angle between the two ground subspaces be θ.[42] Then the ground state energy (g_0) of $H_0 = H_1 + H_2$ is at least $g_1 + g_2 + \Lambda(1 - \cos\theta)$.

Let $H_1 = H_0$ and $H_2 = H_{\mathcal{S}_1,\text{input}}$. We saw that the ground state gap of H_1 is $\Omega(1/L^2)$ and that of H_2 is 1, so we can take $\Lambda = \Omega(1/L^2)$. The ground state energy of H_2 is $g_2 = 0$. Therefore, using the geometrical lemma, we have $g_0 - g_1 \geq \Lambda(1 - \cos\theta)$. It remains to bound the angle between the two ground spaces. From Eq. (C28), it is clear that the (degenerate) ground state of H_2 can be written as a linear combination of $\{|\gamma_j(\ell)\rangle\}_{\ell=1}^L$, whereas the (unique) ground state of H_1 can be written as a monotone vector in $\{|\gamma_j(\ell)\rangle\}_{\ell=0}^L$. Therefore,

$$\cos\theta = \max_{\{c_{\ell'}\}} \left| \sum_{\ell=0}^L \alpha_\ell \langle\gamma_j(\ell)| \sum_{\ell'=1}^L c_{\ell'}|\gamma_j(\ell')\rangle \right|$$

$$= \max_{\{c_{\ell'}\}} \left| \sum_{\ell=1}^L \alpha_\ell c_\ell \langle\gamma_j(\ell)|\gamma_j(\ell)\rangle \right|$$

$$\leq \max_{\{c_{\ell'}\}} \sum_{\ell=1}^L \alpha_\ell |c_\ell| \leq \sqrt{\frac{L}{L+1}} \leq 1 - \frac{1}{2L}, \qquad \text{(C29)}$$

where we have used that $\vec{\alpha}$ is monotone so that $c_\ell = \alpha_\ell\sqrt{L+1}/L$ maximizes the sum. Therefore, the global gap can be bounded from below by $\Omega(1/L^3)$, which is Eq. (79).

---

[42]The angle θ is defined via $\cos\theta = \max_{v_1,v_2}|\langle v_1|v_2\rangle|$, where $|v_i\rangle$ belongs to space i.

APPENDIX D: PROOF OF THE AMPLIFICATION LEMMA (CLAIM 1)

To prove Claim 1, define a new verifier $V^*(\eta, X)$ which amounts to repeating $V(\eta, X)$ K times, where $K = \text{poly}(|\eta|)$ to keep the verifier efficient, and taking a majority vote on the output, i.e., $\Pr(V^*(\eta, X) = 1) = \Pr(\sum_{i=1}^K V_i > K/2)$, where $V_i \in \{0, 1\}$ is the random number associated with the ith run of $V(\eta, X)$.

Now recall the multiplicative Chernoff bound:

$$\Pr\left(\sum_{i=1}^K Y_i \leq (1-\beta)Kp\right) \leq e^{-\beta^2 Kp/2}, \qquad 0 < \beta < 1,$$

$$\Pr\left(\sum_{i=1}^K Y_i \geq (1+\beta)Kp\right) \leq e^{-\beta^2 Kp/(2+\beta)}, \qquad 0 < \beta, \quad \text{(D1)}$$

for $p = \mathbb{E}(Y)$ where $Y \in \{0, 1\}$ is a random variable. Consider first the case where $Q(\eta) = 1$. In that case, $p \geq 2/3$. If we now pick $\beta = 1 - 1/(2p)$ (i.e., $1/4 \leq \beta \leq 1/2$) in the Chernoff bound, then

$$\Pr\left(\sum_{i=1}^K V_i > \frac{K}{2}\right) = 1 - \Pr\left(\sum_{i=1}^K V_i \leq \frac{K}{2}\right)$$

$$\geq 1 - e^{-(p-1/2)^2 K/2p} \geq 1 - e^{-(2/3-1/2)^2 K/4/3}. \quad \text{(D2)}$$

For the case where $Q(\eta) = 0$, $p \leq 1/3$, take $\beta = 1/(2p) - 1 > 0$, so that

$$\Pr\left(\sum_{i=1}^K V_i > \frac{K}{2}\right)$$

$$\leq \Pr\left(\sum_{i=1}^K V_i \geq \frac{K}{2}\right)$$

$$\leq e^{-(p-1/2)^2 K/p(p+1/2)} \leq e^{-(1/3-1/2)^2 K/(1/3+1/2)/3}. \quad \text{(D3)}$$

This shows that $\text{MA}(2/3, 1/3) = \text{MA}(1 - e^{-|\eta|^g}, e^{-|\eta|^g})$, since $K = \text{poly}(|\eta|)$.

To show that $\text{MA}(c, c - 1/|\eta|^g) \subseteq \text{MA}(2/3, 1/3)$, it is sufficient to show that when $Q(\eta) = 0$ it is exponentially unlikely that Merlin is able to fool Arthur so that $Q(\eta) = 1$. Therefore consider the probability of fooling Arthur, i.e., $\Pr(V^*(\eta, X) = 1) > c$ when $Q(\eta) = 0$. Take $p = \Pr(V(\eta, X) = 1) = c - 1/|\eta|^g$. Then

$$\Pr(\text{Arthur fooled}) = \Pr\left(\sum_{i=1}^K V_i \geq Kc\right)$$

$$= \Pr\left(\frac{1}{K}\sum_{i=1}^K V_i \geq p + \epsilon\right), \quad \text{(D4)}$$

where we take $\epsilon = 1/|\eta|^g$. Recall the additive Chernoff bound

$$\Pr\left(\frac{1}{K}\sum_{i=1}^K Y_i \geq p + \epsilon\right) \leq e^{-KD(p+\epsilon\|p)}, \qquad \text{(D5)}$$

where

$$D(x\|y) = x\ln\frac{x}{y} + (1-x)\ln\frac{1-x}{1-y} \quad \text{(D6)}$$

is the Kullback-Leibler divergence. Expanding

$$D(p+\epsilon\|p) = \frac{\epsilon^2}{2p(1-p)} + O(\epsilon^3),$$

we see that if $K = \epsilon^{-2-\varepsilon}$, where $0 < \varepsilon \ll 1$, then we can exponentially suppress the probability that Arthur is fooled by Merlin while keeping $K = \text{poly}(|\eta|)$.

APPENDIX E: PERTURBATIVE GADGETS

In this Appendix we review the subject of perturbative gadgets, which have played an important role in the reduction of the locality of interactions in the proofs of QMA completeness and the universality of AQC. These tools are generally useful. Our discussion is based primarily on Jordan and Farhi (2008) [see also Bravyi, DiVencenzo, Loss, and Terhal (2008)]. To set up the appropriate tools we first briefly review degenerate perturbation theory.

1. Degenerate perturbation theory à la Bloch (1958)

Consider $H = H_0 + \lambda V$ where H_0 has a d-dimensional degenerate ground subspace $\mathcal{E}_0$ with energy 0. Let $|\psi_1\rangle, ..., |\psi_d\rangle$ be the lowest d energy eigenstates of H with energies $E_1, ..., E_d$, and let their span define the subspace $\mathcal{E}$. The goal is to define a perturbative expansion (in λ) for the effective Hamiltonian H_{eff} of H defined as

$$H_{\text{eff}}(H,d) = \sum_{j=1}^{d} E_j |\psi_j\rangle\langle\psi_j|. \quad \text{(E1)}$$

We show that this expansion converges provided λ satisfies

$$\|\lambda V\| < \gamma/4, \quad \text{(E2)}$$

where γ is the gap to the first excited state of H_0. We first show how to construct this effective Hamiltonian in terms of other, more convenient operators.

Let P_0 be the projection onto $\mathcal{E}_0$, and define

$$|\alpha_j\rangle = P_0|\psi_j\rangle, \quad j = 1, ..., d. \quad \text{(E3)}$$

For λ sufficiently small [this will amount to satisfying Eq. (E2)], the states $\{|\alpha_j\rangle\}_{j=1}^{d}$ are linearly independent since the states $\{|\psi_j\rangle\}_{j=1}^{d}$ are only slightly perturbed from the eigenstates of H_0. Note that the states $\{|\alpha_j\rangle\}_{j=1}^{d}$ are not necessarily orthogonal or normalized. There exists an operator $\mathcal{U}$ such that

$$\mathcal{U}|\alpha_j\rangle = |\psi_j\rangle, \quad j = 1, ..., d, \quad \text{(E4a)}$$

$$\mathcal{U}|\phi\rangle = 0, \quad \forall\, |\phi\rangle \in \mathcal{E}_0^{\perp}. \quad \text{(E4b)}$$

This means that

$$P_0\mathcal{U}|\alpha_j\rangle = P_0|\psi_j\rangle = |\alpha_j\rangle$$
$$\Rightarrow P_0^2\mathcal{U}|\alpha_j\rangle = P_0\mathcal{U}|\alpha_j\rangle = P_0|\alpha_j\rangle$$
$$\Rightarrow P_0\mathcal{U} = P_0. \quad \text{(E5)}$$

Let $\tilde{\mathcal{U}}$ be the operator satisfying

$$\tilde{\mathcal{U}}|\psi_j\rangle = |\alpha_j\rangle, \quad j = 1, ..., d, \quad \text{(E6a)}$$

$$\tilde{\mathcal{U}}|\phi\rangle = 0, \quad \forall\, |\phi\rangle \in \mathcal{E}^{\perp}. \quad \text{(E6b)}$$

Note that $\tilde{\mathcal{U}}$ is not the inverse of $\mathcal{U}$ because $\mathcal{U}$ is not invertible on the entire Hilbert space. Also, $\tilde{\mathcal{U}}$ is not P_0 because of Eq. (E6b) (it annihilates all states outside of $\mathcal{E}$). Note that

$$\mathcal{U}P_0\tilde{\mathcal{U}}|\psi_j\rangle = |\psi_j\rangle. \quad \text{(E7)}$$

Now define

$$\mathcal{A} = \lambda P_0 V \mathcal{U}. \quad \text{(E8)}$$

Note that the states $\{|\alpha_i\rangle\}_{i=1}^{d}$ are right eigenvectors of $\mathcal{A}$ with eigenvalues $E_1, ..., E_d$, respectively:

$$\mathcal{A}|\alpha_j\rangle = \lambda P_0 V|\psi_j\rangle = P_0(H_0 + \lambda V)|\psi_j\rangle = P_0 E_j|\psi_j\rangle$$
$$= E_j|\alpha_j\rangle, \quad \text{(E9)}$$

where we used the fact that $P_0 H_0 = 0$ because the eigenvalue of the ground subspace of H_0 is zero. The effective Hamiltonian associated with H can now be constructed using $\mathcal{U}, \tilde{\mathcal{U}}, \mathcal{A}$:

$$H_{\text{eff}}(H,d) = \mathcal{U}\mathcal{A}\tilde{\mathcal{U}}. \quad \text{(E10)}$$

To see this note that

$$\mathcal{U}\mathcal{A}\tilde{\mathcal{U}}|\phi\rangle = 0, \quad \forall\, |\phi\rangle \in \mathcal{E}^{\perp}, \quad \text{(E11a)}$$

$$\mathcal{U}\mathcal{A}\tilde{\mathcal{U}}|\psi_j\rangle = \mathcal{U}\mathcal{A}|\alpha_j\rangle = E_j\mathcal{U}|\alpha_j\rangle = E_j|\psi_j\rangle, \quad \text{(E11b)}$$

which is identical to the action of H_{eff} on a complete set of vectors. The strategy is now to find a perturbative expansion for $\mathcal{U}$ (we will not need the explicit expansion of $\tilde{\mathcal{U}}$, so we do not provide it here), construct $\mathcal{A}$ using Eq. (E8), find $\{|\alpha_j\rangle\}_{j=1}^{d}$ and $\{E_j\}_{j=1}^{d}$ as, respectively, the right eigenvectors and eigenvalues of $\mathcal{A}$, and apply $\mathcal{U}$ to $|\alpha_j\rangle$ to get a perturbative expansion for $|\psi_j\rangle$.

It can be shown that the desired perturbative expansion of $\mathcal{U}$ and $\mathcal{A}$ is given by

$$\mathcal{U} = P_0 + \sum_{m=1}^{\infty} \mathcal{U}_m, \quad \text{(E12a)}$$

$$\mathcal{A} = P_0 V \sum_{m=1}^{\infty} \mathcal{U}_m = \sum_{m=1}^{\infty} \mathcal{A}_m, \quad \text{(E12b)}$$

where

$$\mathcal{U}_m = \sum_{\substack{\ell_1 \geq 1, \ell_2 \geq 0, \ldots, \ell_m \geq 0 \\ \ell_1 + \cdots + \ell_m = m \\ \ell_1 + \cdots + \ell_p \geq p, 1 \leq p \leq m-1}} (S_{\ell_1} \lambda V)(S_{\ell_2} \lambda V) \cdots (S_{\ell_m} \lambda V) P_0, \quad \text{(E13a)}$$

$$S_\ell = \begin{cases} \frac{1}{(-H_0)^\ell}(\mathbb{1} - P_0), & \ell > 0, \\ -P_0, & \ell = 0. \end{cases} \quad \text{(E13b)}$$

The series in Eq. (E12) converges for $\|\lambda V\| < \gamma/4$. To see this note that

$$\|\mathcal{U}\| = \left\| \mathcal{U}_0 + \sum_{m=1}^\infty \mathcal{U}_m \right\| \leq \|\mathcal{U}_0\| + \sum_{m=1}^\infty \|\mathcal{U}_m\|$$

$$\leq 1 + \sum_{m=1}^\infty \lambda^m \sideset{}{'}\sum \|S_{\ell_1} V S_{\ell_2} \cdots S_{\ell_m} V P_0\|$$

$$\leq 1 + \sum_{m=1}^\infty \lambda^m \sideset{}{'}\sum \|S_{\ell_1}\| \cdots \|S_{\ell_m}\| \|V\|, \quad \text{(E14)}$$

where the sum $\sum'$ involves summing all the different ways to add up to m while satisfying convexity, i.e., $\ell_1 + \ell_2 + \cdots + \ell_p \geq p$. Because of the form of S_ℓ [Eq. (E13b)], we have

$$\|S_\ell\| = \left(\frac{1}{E_1^{(0)}}\right)^\ell = \frac{1}{\gamma^\ell}, \quad \text{(E15)}$$

where $E_1^{(0)}$ is the energy of the first excited state of H_0 (corresponds to the state that minimizes $H_0 Q_0$ to calculate the operator norm). Therefore, we have

$$\|\mathcal{U}\| \leq 1 + \sum_{m=1}^\infty \lambda^m \sideset{}{'}\sum \frac{\|V\|^m}{\gamma^m}. \quad \text{(E16)}$$

The sum $\sum'$ is less than the number of ways to add up to m using m non-negative integers, which is given by $\binom{2m-1}{m}$. However, since $\sum_{j=0}^{2m-1} \binom{2m-1}{j} = 2^{2m-1}$, it is clear that $\binom{2m-1}{m} \leq 2^{2m-1}$. Therefore, we can upper bound the sum with this value

$$\|\mathcal{U}\| \leq 1 + \sum_{m=1}^\infty 2^{2m-1} \frac{\|\lambda V\|^m}{\gamma^m}. \quad \text{(E17)}$$

This series converges if the condition for λ in Eq. (E2) is satisfied.

2. Perturbative gadgets

For a k-local target Hamiltonian H^{T}, the goal is to construct a 2-local "gadget" Hamiltonian H^{G}, whose low energy spectrum (captured by an effective Hamiltonian H_{eff}) approximates the spectrum of H^{T}. In order to do so, we use the expression in Eq. (E10) for the effective Hamiltonian in terms of the operators $\mathcal{U}$ and $\mathcal{A}$ and use their perturbative expansion from the previous section. We show that for our gadget Hamiltonian, the perturbative expansion of the effective Hamiltonian matches that of the target Hamiltonian.

The perturbative gadget we review here uses a strongly bound set of ancillas, coupled to the target qubits via weaker interactions, where the latter are treated as a perturbation. H^{T} is then generated in low order perturbation theory of the combined system consisting of both ancilla and target qubits. Such gadgets first appeared in the proof of QMA completeness of the 2-local Hamiltonian problem via a reduction from the 3-local Hamiltonian, where they were used to construct effective three-body interactions from two-body ones (Kempe, Kitaev, and Regev, 2006).

Let H_s denote a k-local term. For the ith qubit in H_s, we associate an arbitrary direction in $\mathbb{R}^3$ denoted by $\hat{n}_{s,i}$. A general k-local target Hamiltonian acting on n qubits can then be expressed as

$$H^{\mathrm{T}} = \sum_{s=1}^r c_s H_s, \quad \text{(E18)}$$

with $H_s = \sigma_{s,1} \sigma_{s,2} \cdots \sigma_{s,k}$ where $\sigma_{s,j} = \hat{n}_{s,j} \cdot \vec{\sigma}_{s,j}$. The goal is to simulate H^{T} using only 2-local interactions. Toward this end, introduce k ancilla qubits for each H_s, for a total of rk ancilla qubits. Define

$$H^{\mathrm{G}} = H^{\mathrm{A}} + \lambda V = \sum_{s=1}^r H_s^{\mathrm{A}} + \lambda \sum_{s=1}^r V_s, \quad \text{(E19a)}$$

$$H_s^{\mathrm{A}} = \sum_{i<j}^k \frac{1}{2}(\mathbb{1} - Z_{s,i} Z_{s,j}), \quad \text{(E19b)}$$

$$V_s = \sum_{j=1}^k c_{s,j} \sigma_{s,j} \otimes X_{s,j}, \quad \text{(E19c)}$$

$$c_{s,j} = \begin{cases} c_s & j = 1, \\ 1 & j \neq 1, \end{cases} \quad \text{(E19d)}$$

where $X_{s,j}, Z_{s,j}$ are the Pauli (x, z) operators on the jth ancilla qubit of H_s. Note that the ground state of H_s^{A} is given by the span of $\{|0_1 \cdots 0_k\rangle_s^{\mathrm{A}}, |1_1 \cdots 1_k\rangle_s^{\mathrm{A}}\}$.

Consider the k-local ancilla operator $X_s \equiv X_{s,1} \otimes X_{s,2} \otimes \cdots \otimes X_{s,k}$. This operator clearly commutes with H_{G}. Therefore H_{G} and the set of operators $\{X_s\}_{s=1}^r$ share a set of eigenstates. The operator X_s has eigenvalues ± 1, each with degeneracy 2^{k-1} (to see this simply write X_s in the basis $\{|\pm\rangle_{s,1} \otimes |\pm\rangle_{s,2} \otimes \cdots \otimes |\pm\rangle_{s,k}\}$). Therefore, H^{G} can be block diagonalized into 2^r blocks, where each block corresponds to a fixed $X_s = \pm 1$ for $s = 1, \ldots, r$ with dimension $2^n 2^{r(k-1)}$. Let H_+^{G} denote the block with $X_s = 1, \forall s$.

Note that since H_+^{G} will be used to approximate H^{G}, the system will need to be initialized to have $X_s = 1, \forall s$. The eigenstate of X_s with eigenvalue 1 is given by

$$|+\rangle_s = \frac{1}{\sqrt{2}}(|0_1 \cdots 0_k\rangle_s + |1_1 \cdots 1_k\rangle_s), \quad \text{(E20)}$$

so the ancilla qubits must be initialized to be in the state $\otimes_{r=1}^s |+\rangle_s$.

We show that the low energy spectrum of H_+^{G} approximates the spectrum of H^{T}. Our task is to calculate $H_{\mathrm{eff}}(H_+^{\mathrm{G}}, 2^n)$ [in the notation of Eq. (E1)] perturbatively to kth order in λ. λV will perturb the ground subspace of H^{A} in two ways:

(1) It shifts the energy of the entire subspace.
(2) It splits the degeneracy of the ground subspace beginning at kth order in perturbation theory. It is this splitting that will allow us to mimic the spectrum of H_{T}.

We analyze the shift and splitting separately. To do this, define

$$\tilde{H}_{\mathrm{eff}}(H, d, \Delta) \equiv H_{\mathrm{eff}}(H, d) - \Delta\Pi, \qquad (E21)$$

where Π is the projection onto the space spanned by $\{|E_j\rangle\}_{j=1}^d$. Note that the eigenstates of $\tilde{H}_{\mathrm{eff}}$ and H_{eff} are identical, and the energy gaps between energy levels are identical too.

Let us start with the case where $r = 1$, i.e., $H^{\mathrm{T}} = \sigma_1\sigma_2\cdots\sigma_k$, so that $H^{\mathrm{A}} = \sum_{i=1}^k\sum_{j=i+1}^k\frac{1}{2}(1 - Z_iZ_j)$, and $V = \sum_{j=1}^k\sigma_j\otimes X_j$. We first construct $\mathcal{A}$ [Eq. (E8)] for H_+^{G}. Note that H^{A} has a ground state of zero energy (corresponds to all qubits with $Z_i = 1$ or all qubits with $Z_i = -1$), and the first excited state has energy $\gamma = k - 1$ (let $Z_1 = -1$, all the rest are $+1$). Furthermore,

$$\|V\| = \left\|\sum_{j=1}^k\sigma_j\otimes X_j\right\| \le \sum_{j=1}^k\|\sigma_j\otimes X_j\| = k. \qquad (E22)$$

Therefore, by Eq. (E2), the perturbative expansion will converge if $\lambda < (k-1)/4k$. Because of the form of $\mathcal{A}$ [Eq. (E12b)], all $\mathcal{A}_m$ terms are sandwiched between P_0 operators. Thus, all nonzero terms in $\mathcal{A}$ must take states in $\mathcal{E}_0$ and return them to states in $\mathcal{E}_0$. Since we have restricted to the $X = \otimes_{i=1}^k X_i = +1$ sector, $\mathcal{E}_0$ is restricted to have the ancilla qubits in the $|+\rangle$ state [Eq. (E20)]. Therefore, we can write

$$P_0 = \mathbb{1}\otimes P_+, \qquad (E23)$$

where P_+ is the projection onto the $|+\rangle$ ancilla state.

Each term in $V = \sum_{j=1}^k\sigma_j\otimes X_j$ flips only a single ancilla qubit. Therefore, in order for $\mathcal{A}$ take a state out of $\mathcal{E}_0$ and return it, the power of V must either flip all ancilla qubits or flip some and return them back to their original value. The former process (flipping all qubits) first happens at kth order in perturbation theory. The latter process (flipping and returning) can happen at lower orders than k, but $\mathcal{A}$ is then proportional to P_0 since the product of V's effectively cancel. To see how this works, consider $\mathcal{A}$ up to second order for $k > 2$. From the perturbation expansion [Eq. (E12b)] we have

$$\mathcal{A}_{\le 2} = \lambda P_0VP_0 + \lambda^2 P_0VS_1VP_0, \qquad (E24)$$

but $P_0VP_0 = 0$ since $V|+\rangle$ is orthogonal to $|+\rangle$. On the other hand, VP_0 takes the system to a state with energy $k - 1$ for H^{A}, so $S_1VP_0 = -VP_0/(k-1)$. Therefore,

$$\mathcal{A}_{\le 2} = -\frac{\lambda^2}{k-1}P_0V^2P_0. \qquad (E25)$$

Now note that

$$V^2 = \sum_i(\sigma_i\otimes X_i)^2 + \sum_{i\ne j}(\sigma_i\otimes X_i)(\sigma_j\otimes X_j). \qquad (E26)$$

The cross terms are annihilated by $P_0\cdot P_0$ since they take the state out of $\mathcal{E}_0$. The diagonal term is proportional to the identity on the ancilla qubits, so we have

$$\mathcal{A}_{\le 2} = -\frac{\lambda^2}{k-1}\Omega P_0, \qquad (E27)$$

where Ω is an operator that depends on the particular orientation of the σ_j's. This argument extends to order $k - 1$ so that

$$\mathcal{A}_{\le k-1} = \sum_{m\,\text{even}}\lambda^m\Omega_m P_0. \qquad (E28)$$

At order k, something new happens. There are now cross terms that involve all X_i's once, i.e.,

$$\lambda^k P_0(\sigma_1\otimes X_1)S_1(\sigma_2\otimes X_2)S_1\cdots S_1(\sigma_k\otimes X_k)P_0.$$

The S_1 operator measures the successive change in energy of the system to give an overall constant of

$$\left(-\frac{1}{k-1}\right)\left(-\frac{1}{2(k-2)}\right)\cdots\left(-\frac{1}{(k-1)1}\right)$$
$$= \prod_{j=1}^{k-1}\left(-\frac{1}{j(k-j)}\right) = \frac{(-1)^{k-1}}{(k-1)!^2}. \qquad (E29)$$

Therefore, the cross terms (of which there are $k!$, since the operators can be multiplied in any order) then take the form

$$-\frac{(-\lambda)^k k!}{(k-1)!^2}P_0(\sigma_1\cdots\sigma_k\otimes X)P_0. \qquad (E30)$$

Thus,

$$\mathcal{A}_{\le k} = f(\lambda)P_0 - \frac{k(-\lambda)^k}{(k-1)!}P_0(H^{\mathrm{T}}\otimes X)P_0, \qquad (E31)$$

where $f(\lambda)$ is some kth order polynomial in λ, with coefficients that depend on H^{T}. Using the form of H_{eff} in Eq. (E10), we have

$$H_{\mathrm{eff}}(H_-^{\mathrm{G}}, 2^n) = f(\lambda)\mathcal{U}P_0\tilde{\mathcal{U}}$$
$$- \mathcal{U}\left(\frac{k(-\lambda)^k}{(k-1)!}P_0(H^{\mathrm{T}}\otimes X)P_0 + O(\lambda^{k+1})\right)\tilde{\mathcal{U}}. \qquad (E32)$$

Recall that $\mathcal{U}P_0\tilde{\mathcal{U}}|\psi_j\rangle = |\psi_j\rangle$ [Eq. (E7)] so $\mathcal{U}P_0\tilde{\mathcal{U}}$ acts as the identity in $\mathcal{E}$ so that the first term in Eq. (E32) can be dropped. Furthermore, we can replace $\mathcal{U}$ and $\tilde{\mathcal{U}}$ in the second term by their λ^0 counterpart since we are keeping terms only to order k

and the term in the parentheses is already of order k. Therefore,

$$
\begin{aligned}
\tilde{H}_{\text{eff}}&(H_+^{\text{G}}, 2^n, f(\lambda)) \\
&= -\frac{k(-\lambda)^k}{(k-1)!} P_0(H^{\text{T}} \otimes X)P_0 + O(\lambda^{k+1}) \\
&= -\frac{k(-\lambda)^k}{(k-1)!} (H^{\text{T}} \otimes P_+) + O(\lambda^{k+1}). \quad \text{(E33)}
\end{aligned}
$$

This shows that the target Hamiltonian H^{T} appears as the leading order term in the effective Hamiltonian that describes the 2^n-Hilbert space of the n target qubits, albeit with a diminished magnitude of order $\lambda^k/(k-1)!$.

Let us now consider the general r case, i.e., the Hamiltonian in Eq. (E18). We note that just as in the $r = 1$ case, H^{A} again has an energy gap of $k - 1$. Generalizing from Eq. (E22), the perturbative expansion then converges for

$$
\lambda < \frac{k-1}{4\|V\|}. \quad \text{(E34)}
$$

In the sector where $X_s = +1$, H^{A} has the state $\otimes_{s=1}^r |+\rangle_s$ as a ground state. Since H^{A} acts as the identity on the computational qubits, the ground state is 2^n-fold degenerate.

In the perturbation expansion for $\mathcal{A}$, products of V again appear. Each V_s acts on a different ancilla register. Therefore, at order k, cross terms of different V_s's cannot flip all k ancilla qubits in a register, so they are annihilated by $P_0 \cdot P_0$. The only cross terms that contribute are k products of a given s where each ancilla qubit appears once. Therefore, the natural generalization of the previous result is recovered, namely, Eq. (E33) continues to hold with H^{T} replaced by the sum over r terms as in Eq. (E18), where again $f(\lambda)$ is some polynomial in λ of order k with coefficients that depend on $c_s H_s$, and where P_+ is the projector onto $\otimes_s |+\rangle_s$.

Note that the convergence condition (E34) requires the interaction term V to be stronger than the effective interaction it generates, which scales as λ^k [as can be seen from Eq. (E33)]. This may pose implementation difficulties, since a practical device is likely to have only a limited range of interaction strengths. Weaker gadgets can be implemented that circumvent this problem, albeit at the cost of a larger overhead of ancillary qubits (Cao and Nagaj, 2015). The idea is to replace strong interactions by repetition of interactions with "classical" ancillas. Additional gadget simplifications and resource reductions were proposed by Cao *et al.* (2015).

REFERENCES

Aaronson, S., 2010, arXiv:1009.5104.

Aaronson, S., 2015, Nat. Phys. **11**, 291.

Adachi, S. H., and M. P. Henderson, 2015, arXiv:1510.06356.

Aharonov, D., D. Gottesman, S. Irani, and J. Kempe, 2009, Commun. Math. Phys. **287**, 41.

Aharonov, D., and T. Naveh, 2002, arXiv:quant-ph/0210077.

Aharonov, D., and A. Ta-Shma, 2003, in *Proceedings of the Thirty-fifth Annual ACM Symposium on Theory of Computing, STOC '03* (ACM, New York, NY, USA), pp. 20–29.

Aharonov, D., W. van Dam, J. Kempe, Z. Landau, S. Lloyd, and O. Regev, 2007, SIAM J. Comput. **37**, 166.

Albash, T., and D. Lidar, 2018, unpublished.

Altshuler, B., H. Krovi, and J. Roland, 2010, Proc. Natl. Acad. Sci. U.S.A. **107**, 12446.

Amara, P., D. Hsu, and J. E. Straub, 1993, J. Phys. Chem. **97**, 6715.

Ambainis, A., and O. Regev, 2004, arXiv:quant-ph/0411152.

Amico, L., R. Fazio, A. Osterloh, and V. Vedral, 2008, Rev. Mod. Phys. **80**, 517.

Amin, M. H., E. Andriyash, J. Rolfe, B. Kulchytskyy, and R. Melko, 2016, arXiv:1601.02036.

Amin, M. H. S., 2009, Phys. Rev. Lett. **102**, 220401.

Amin, M. H. S., and V. Choi, 2009, Phys. Rev. A **80**, 062326.

Andriyash, E., and M. H. Amin, 2017, arXiv:1703.09277.

Apolloni, B., C. Carvalho, and D. de Falco, 1989, Stoch. Proc. Appl. **33**, 233.

Apolloni, B., N. Cesa-Bianchi, and D. de Falco, 1988, in *Proceedings of the Ascona/Locarno Conference* (World Scientific, Singapore), p. 97.

Arecchi, F. T., E. Courtens, R. Gilmore, and H. Thomas, 1972, Phys. Rev. A **6**, 2211.

Aspuru-Guzik, A., A. D. Dutoi, P. J. Love, and M. Head-Gordon, 2005, Science **309**, 1704.

Averin, D. V., 1998, Solid State Commun. **105**, 659.

Avron, J. E., and A. Elgart, 1999, Commun. Math. Phys. **203**, 445.

Awerbuch, B., 1985, J. ACM **32**, 804.

Azinović, M., D. Herr, B. Heim, E. Brown, and M. Troyer, 2016, arXiv:1607.03329.

Babai, L., 2015, arXiv:1512.03547.

Babbush, R., V. Denchev, N. Ding, S. Isakov, and H. Neven, 2014, arXiv:1406.4203.

Babbush, R., P. J. Love, and A. Aspuru-Guzik, 2014, Sci. Rep. **4**, 6603.

Bapst, V., and G. Semerjian, **2012**, J. Stat. Mech. 12, P06007.

Barabasi, A. L., and R. Albert, 1999, Science **286**, 509.

Batu, T., L. Fortnow, R. Rubinfeld, W. D. Smith, and P. White, 2000, in *Proceedings of the 41st Annual Symposium on Foundations of Computer Science* (IEEE, Washington, DC), pp. 259–269.

Bauer, B., L. Wang, I. Pizorn, and M. Troyer, 2015, arXiv:1501.06914.

Becker, A., J.-S. Coron, and A. Joux, 2011, "Improved generic algorithms for hard knapsacks," in *Advances in Cryptology—EUROCRYPT 2011: Proceedings of the 30th Annual International Conference on the Theory and Applications of Cryptographic Techniques*, edited by K. G. Paterson (Springer, Berlin, Heidelberg), pp. 364–385.

Behrman, E. C., and J. E. Steck, 2016, arXiv:1603.01752.

Benedetti, M., J. Realpe-Gómez, R. Biswas, and A. Perdomo-Ortiz, 2016, Phys. Rev. A **94**, 022308.

Benioff, P., 1980, J. Stat. Phys. **22**, 563.

Benioff, P., 1982, Phys. Rev. Lett. **48**, 1581.

Bennett, C., E. Bernstein, G. Brassard, and U. Vazirani, 1997, SIAM J. Comput. **26**, 1510.

Bernstein, E., and U. Vazirani, 1993, in *Proceedings of the Twenty-fifth Annual ACM Symposium on Theory of Computing, STOC '93* (ACM, New York, NY, USA), pp. 11–20.

Bernstein, E., and U. Vazirani, 1997, SIAM J. Comput. **26**, 1411.

Bhatia, R., 1997, *Matrix Analysis*, Graduate Texts in Mathematics No. 169 (Springer-Verlag, New York).

Biamonte, J. D., and P. J. Love, 2008, Phys. Rev. A **78**, 012352.

Biham, E., O. Biham, D. Biron, M. Grassl, and D. A. Lidar, 1999, Phys. Rev. A **60**, 2742.

Bisseling, R. H., 2004, *Parallel Scientific Computation* (Oxford University Press Oxford).

Bloch, C., 1958, Nucl. Phys. **6**, 329.

Boixo, S., E. Knill, and R. D. Somma, 2009, Quantum Inf. Comput. **9**, 833.

Boixo, S., E. Knill, and R. D. Somma, 2010, arXiv:1005.3034 (unpublished).

Boixo, S., T. F. Ronnow, S. V. Isakov, Z. Wang, D. Wecker, D. A. Lidar, J. M. Martinis, and M. Troyer, 2014, Nat. Phys. **10**, 218.

Boixo, S., V. N. Smelyanskiy, A. Shabani, S. V. Isakov, M. Dykman, V. S. Denchev, M. H. Amin, A. Y. Smirnov, M. Mohseni, and H. Neven, 2016, Nat. Commun. **7**, 10327.

Boixo, S., and R. D. Somma, 2010, Phys. Rev. A **81**, 032308.

Bollobás, B., O. Riordan, J. Spencer, and G. Tusnády, 2001, Random Struct. Algorithms **18**, 279.

Borgs, C., J. Chayes, and B. Pittel, 2001, Random Struct. Algorithms **19**, 247.

Born, M., and V. Fock, 1928, Z. Phys. **51**, 165.

Born, M., and R. Oppenheimer, 1927, Ann. Phys. (Berlin) **389**, 457.

Boyer, M., G. Brassard, P. Høyer, and A. Tapp, 1998, Fortschr. Phys. **46**, 493.

Brady, L. T., and W. v. Dam, 2016, arXiv:1609.09137.

Brady, L. T., and W. van Dam, 2016a, Phys. Rev. A **93**, 032304.

Brady, L. T., and W. van Dam, 2016b, Phys. Rev. A **94**, 032309.

Brady, L. T., and W. van Dam, 2017, Phys. Rev. A **95**, 032335.

Brandao, F. G. S. L., and M. Horodecki, 2013, Nat. Phys. **9**, 721.

Bravyi, S., A. J. Bessen, and B. M. Terhal, 2006, arXiv:quant-ph/0611021.

Bravyi, S., D. P. DiVincenzo, D. Loss, and B. M. Terhal, 2008, Phys. Rev. Lett. **101**, 070503.

Bravyi, S., D. P. DiVincenzo, R. I. Oliveira, and B. M. Terhal, 2008, Quantum Inf. Comput. **8**, 0361.

Bravyi, S., and D. Gosset, 2016, arXiv:1612.05602.

Bravyi, S., and M. Hastings, 2014, arXiv:1410.0703.

Bravyi, S., and M. Hastings, 2017, Commun. Math. Phys. **349**, 1.

Bravyi, S., and B. M. Terhal, 2009, SIAM J. Comput. **39**, 1462.

Bravyi, S. B., and A. Y. Kitaev, 2002, Ann. Phys. (Amsterdam) **298**, 210.

Breuckmann, N. P., and B. M. Terhal, 2014, J. Phys. A **47**, 195304.

Brin, S., and L. Page, 1998, Comput. Netw. ISDN Syst. **30**, 107.

Brooke, J., D. Bitko, T. F. Rosenbaum, and G. Aeppli, 1999, Science **284**, 779.

Brooke, J., T. F. Rosenbaum, and G. Aeppli, 2001, Nature (London) **413**, 610.

Buhrman, H., R. Cleve, J. Watrous, and R. de Wolf, 2001, Phys. Rev. Lett. **87**, 167902.

Bunyk, P. I., *et al.*, 2014, IEEE Trans. Appl. Supercond. **24**, 1.

Cabrera, G. G., and R. Jullien, 1987, Phys. Rev. B **35**, 7062.

Cao, Y., R. Babbush, J. Biamonte, and S. Kais, 2015, Phys. Rev. A **91**, 012315.

Cao, Y., S. Jiang, D. Perouli, and S. Kais, 2016, arXiv:1608.08547.

Cao, Y., and D. Nagaj, 2015, Quantum Inf. Comput. **15**, 1197.

Cao, Z., and A. Elgart, 2012, J. Math. Phys. (N.Y.) **53**, 032201.

Chancellor, N., 2016, arXiv:1606.06833.

Chayes, L., N. Crawford, D. Ioffe, and A. Levit, 2008, J. Stat. Phys. **133**, 131.

Cheung, D., P. Høyer, and N. Wiebe, 2011, J. Phys. A **44**, 415302.

Childs, A. M., R. Cleve, E. Deotto, E. Farhi, S. Gutmann, and D. A. Spielman, 2003, in *Proceedings of the Thirty-fifth Annual ACM Symposium on Theory of Computing, STOC '03* (ACM, New York, NY, USA), pp. 59–68.

Childs, A. M., E. Farhi, J. Goldstone, and S. Gutmann, 2002, Quantum Inf. Comput. **2**, 181.

Childs, A. M., D. Gosset, and Z. Webb, 2013, arXiv:1311.3297.

Choi, V., 2010, arXiv:1004.2226.

Choi, V., 2011, Proc. Natl. Acad. Sci. U.S.A. **108**, E19.

Collins, D., K. W. Kim, and W. C. Holton, 1998, Phys. Rev. A **58**, R1633.

Crosson, E., and M. Deng, 2014, arXiv:1410.8484.

Crosson, E., E. Farhi, C. Y. -Y. Lin, H.-H. Lin, and P. Shor, 2014, arXiv:1401.7320.

Crosson, E., and A. Harrow, 2016, arXiv:1601.03030.

Cubitt, T., and A. Montanaro, 2016, SIAM J. Comput. **45**, 268.

Das, A., and B. K. Chakrabarti, 2008, Rev. Mod. Phys. **80**, 1061.

Das, A., B. K. Chakrabarti, and R. B. Stinchcombe, 2005, Phys. Rev. E **72**, 026701.

Das, S., R. Kobes, and G. Kunstatter, 2002, Phys. Rev. A **65**, 062310.

Das Sarma, A., A. R. Molla, G. Pandurangan, and E. Upfal, 2015, Special Issue on Distributed Computing and Networking, Theor. Comput. Sci. **561**, Part B, 113.

Denchev, V. S., S. Boixo, S. V. Isakov, N. Ding, R. Babbush, V. Smelyanskiy, J. Martinis, and H. Neven, 2016, Phys. Rev. X **6**, 031015.

Denchev, V. S., N. Ding, S. V. N. Vishwanathan, and H. Neven, 2012, arXiv:1205.1148.

Deutsch, D., 1985, Proc. R. Soc. A **400**, 97.

Deutsch, D., 1989, Proc. R. Soc. A **425**, 73.

Deutsch, D., and R. Jozsa, 1992, Proc. R. Soc. A **439**, 553.

Dickson, N. G., 2011, New J. Phys. **13**, 073011.

Dickson, N. G., and M. H. S. Amin, 2011, Phys. Rev. Lett. **106**, 050502.

Du, J., L. Hu, Y. Wang, J. Wu, M. Zhao, and D. Suter, 2008, Phys. Rev. Lett. **101**, 060403.

Dulny, J., and M. Kim, 2016, arXiv:1603.07980.

Durkin, G. A., 2016, Phys. Rev. A **94**, 043821.

Dusuel, S., and J. Vidal, 2005, Phys. Rev. B **71**, 224420.

Earl, D. J., and M. W. Deem, 2005, Phys. Chem. Chem. Phys. **7**, 3910.

Ehrenfest, P., 1916, Verslagen Kon. Akad. Amsterdam **25**, 412.

Einstein, A., 1914, Verh. Dtsch. Phys. Ges. **16**, 826.

Eisert, J., M. Cramer, and M. B. Plenio, 2010, Rev. Mod. Phys. **82**, 277.

Elgart, A., and G. A. Hagedorn, 2012, J. Math. Phys. (N.Y.) **53**, 102202.

Farhi, E., J. Goldstone, D. Gosset, S. Gutmann, H. B. Meyer, and P. Shor, 2011, Quantum Inf. Comput. **11**, 181.

Farhi, E., J. Goldstone, and S. Gutmann, 2002a, arXiv:quant-ph/0201031.

Farhi, E., J. Goldstone, and S. Gutmann, 2002b, arXiv:quant-ph/0208135.

Farhi, E., J. Goldstone, S. Gutmann, J. Lapan, A. Lundgren, and D. Preda, 2001, Science **292**, 472.

Farhi, E., J. Goldstone, S. Gutmann, and D. Nagaj, 2008, Int. J. Quantum. Inform. **06**, 503.

Farhi, E., J. Goldstone, S. Gutmann, and M. Sipser, 2000, arXiv:quant-ph/0001106.

Farhi, E., D. Gosset, I. Hen, A. W. Sandvik, P. Shor, A. P. Young, and F. Zamponi, 2012, Phys. Rev. A **86**, 052334.

Farhi, E., and S. Gutmann, 1998, Phys. Rev. A **57**, 2403.

Farhi, E., and A. W. Harrow, 2016, arXiv:1602.07674.

Feynman, R. P., 1982, Int. J. Theor. Phys. **21**, 467.

Feynman, R. P., 1985, Opt. News **11**, 11.

Finnila, A. B., M. A. Gomez, C. Sebenik, C. Stenson, and J. D. Doll, 1994, Chem. Phys. Lett. **219**, 343.

Fisher, D. S., 1995, Phys. Rev. B **51**, 6411.

Foini, L., G. Semerjian, and F. Zamponi, 2010, Phys. Rev. Lett. **105**, 167204.

Franz, S., M. Mézard, F. Ricci-Tersenghi, M. Weigt, and R. Zecchina, 2001, Europhys. Lett. **55**, 465.

Frees, A., J. K. Gamble, K. Rudinger, E. Bach, M. Friesen, R. Joynt, and S. N. Coppersmith, 2013, Phys. Rev. A **88**, 032307.

Freund, Y., R. Schapire, and N. Abe, 1999, Journal of Japanese Society for Artificial Intelligence **14**, 771.

Gaitan, F., and L. Clark, 2012, Phys. Rev. Lett. **108**, 010501.

Gaitan, F., and L. Clark, 2014, Phys. Rev. A **89**, 022342.

Garey, M. R. and D. S. Johnson, 1979, *Computers and Intractability: A Guide to the Theory of NP-Completeness* (W.H. Freeman, New York).

Garnerone, S., P. Zanardi, and D. A. Lidar, 2012, Phys. Rev. Lett. **108**, 230506.

Garrido, L. M. and F. J. Sancho, 1962, Physica (Utrecht) **28**, 553.

Ge, Y., A. Molnár, and J. I. Cirac, 2015, arXiv:1508.00570.

Gershgorin, S., 1931, Bull. Acad. Sci. USSR Cl. Sci. Math. Nat. **6**, 749.

Gharibian, S., 2013, Approximation, Proof Systems, and Correlations in a Quantum World, Ph.D. thesis (University of Waterloo).

Gharibian, S., Y. Huang, Z. Landau, and S. W. Shin, 2015, Theor. Comput. Sci. **10**, 159.

Goldreich, O., 2008, *Computational Complexity: A Conceptual Perspective* (Cambridge University Press, Cambridge, England).

Gosset, D., B. M. Terhal, and A. Vershynina, 2015, Phys. Rev. Lett. **114**, 140501.

Goto, H., 2016, Sci. Rep. **6**, 21686.

Gottesman, D., and M. B. Hastings, 2010, New J. Phys. **12**, 025002.

Gottesman, D., and S. Irani, 2013, Theory Comput. **9**, 31.

Grover, L. K., 1997, Phys. Rev. Lett. **79**, 325.

Guidetti, M., and A. P. Young, 2011, Phys. Rev. E **84**, 011102.

Haanpaa, H., M. Jarvisalo, P. Kaski, and I. Niemela, 2006, Journal on Satisfiability, Boolean Modeling and Computation **2**, 27.

Hagedorn, G. A., and A. Joye, 2002, J. Math. Anal. Appl. **267**, 235.

Hallgren, S., D. Nagaj, and S. Narayanaswami, 2013, Quantum Inf. Comput. **13**, 0721.

Han, Y., L. A. Hemaspaandra, and T. Thierauf, 1993, "Threshold computation and cryptographic security," in *Algorithms and Computation: Proceedings of the 4th International Symposium, ISAAC '93*, edited by K. W. Ng, P. Raghavan, N. V. Balasubramanian, and F. Y. L. Chin (Springer, Berlin, Heidelberg), pp. 230–239.

Hashizume, Y., T. Koizumi, K. Akitaya, T. Nakajima, S. Okamura, and M. Suzuki, 2015, Phys. Rev. E **92**, 023302.

Hastings, M. B., **2007**, J. Stat. Mech. 07, P08024.

Hastings, M. B., 2009, Phys. Rev. Lett. **103**, 050502.

Hastings, M. B., and M. H. Freedman, 2013, Quantum Inf. Comput. **13**, 1038.

Hauke, P., L. Bonnes, M. Heyl, and W. Lechner, 2015, Front. Phys. **3**.

Hayden, P., D. W. Leung, and A. Winter, 2006, Commun. Math. Phys. **265**, 95.

Hayes, B., 2002, Am. Sci. **90**, 113.

Hen, I., 2014a, J. Phys. A **47**, 045305.

Hen, I., 2014b, Europhys. Lett. **105**, 50005.

Hen, I., 2017, Europhys. Lett. **118**, 30003.

Hen, I., and A. P. Young, 2011, Phys. Rev. E **84**, 061152.

Hen, I., and A. P. Young, 2012, Phys. Rev. A **86**, 042310.

Hinton, G. E., S. Osindero, and Y.-W. Teh, 2006, Neural Comput. **18**, 1527.

Hoeffding, W., 1963, J. Am. Stat. Assoc. **58**, 13.

Hormozi, L., E. W. Brown, G. Carleo, and M. Troyer, 2017, Phys. Rev. B **95**, 184416.

Horn, R. A., and C. R. Johnson, 2012, *Matrix Analysis* (Cambridge University Press, Cambridge, England).

Houdayer, J., 2001, Eur. Phys. J. B **22**, 479.

Huang, Y., 2014, arXiv:1406.6355.

Hubbard, J., 1959, Phys. Rev. Lett. **3**, 77.

Hukushima, K., and K. Nemoto, 1996, J. Phys. Soc. Jpn. **65**, 1604.

Impagliazzo, R., and R. Paturi, 2001, J. Comput. Syst. Sci. **62**, 367.

Inack, E. M., and S. Pilati, 2015, Phys. Rev. E **92**, 053304.

Isakov, S. V., G. Mazzola, V. N. Smelyanskiy, Z. Jiang, S. Boixo, H. Neven, and M. Troyer, 2016, Phys. Rev. Lett. **117**, 180402.

Ishii, H., and T. Yamamoto, 1985, J. Phys. C **18**, 6225.

Jansen, S., M.-B. Ruskai, and R. Seiler, 2007, J. Math. Phys. (N.Y.) **48**, 102111.

Janzing, D., 2007, Phys. Rev. A **75**, 012307.

Janzing, D., P. Wocjan, and S. Zhang, 2008, New J. Phys. **10**, 093004.

Jarret, M., S. P. Jordan, and B. Lackey, 2016, Phys. Rev. A **94**, 042318.

Jerrum, M., 1992, Random Structures & Algorithms **3**, 347.

Johnson, M. W., *et al.*, 2011, Nature (London) **473**, 194.

Jordan, P. and E. Wigner, 1928, Z. Phys. **47**, 631.

Jordan, S. P., 2016a, "AQC 2016—Adiabatic Quantum Computer vs. Diffusion Monte Carlo," website.

Jordan, S. P., 2016b, "Quantum algorithm zoo," website.

Jordan, S. P., and E. Farhi, 2008, Phys. Rev. A **77**, 062329.

Jordan, S. P., D. Gosset, and P. J. Love, 2010, Phys. Rev. A **81**, 032331.

Jörg, T., F. Krzakala, J. Kurchan, and A. C. Maggs, 2008, Phys. Rev. Lett. **101**, 147204.

Jörg, T., F. Krzakala, J. Kurchan, A. C. Maggs, and J. Pujos, 2010, Europhys. Lett. **89**, 40004.

Jörg, T., F. Krzakala, G. Semerjian, and F. Zamponi, 2010, Phys. Rev. Lett. **104**, 207206.

Joye, A., 1994, Asymptotic Analysis **9**, 209.

Jozsa, R., and N. Linden, 2003, Proc. R. Soc. A **459**, 2011.

Kadowaki, T., and H. Nishimori, 1998, Phys. Rev. E **58**, 5355.

Karimi, H., and G. Rosenberg, 2016, arXiv:1606.07797.

Karimi, S., and P. Ronagh, 2016, arXiv:1605.09462.

Karmarkar, N., and R. Karp, 1982, "The differencing method of set partitioning," Technical Report No. UCB/CSD 82/113 (University of California at Berkeley).

Karp, R., 1972, in *Complexity of Computer Computations*, The IBM Research Symposia Series, edited by R. E. Miller and J. W. Thatcher (Plenum, New York), Chap. 9, p. 85.

Kato, T., 1950, J. Phys. Soc. Jpn. **5**, 435.

Katzgraber, H. G., F. Hamze, Z. Zhu, A. J. Ochoa, and H. Munoz-Bauza, 2015, arXiv:1505.01545.

Katzgraber, H. G., S. Trebst, D. A. Huse, and M. Troyer, **2006**, J. Stat. Mech. 03, P03018.

Kay, A., 2008, Phys. Rev. A **78**, 012346.

Kempe, J., A. Kitaev, and O. Regev, 2006, SIAM J. Comput. **35**, 1070.

Kempe, J., and O. Regev, 2003, Quantum Inf. Comput. **3**, 258.

King, J., S. Yarkoni, M. M. Nevisi, J. P. Hilton, and C. C. McGeoch, 2015, arXiv:1508.05087.

Kirkpatrick, S., C. D. Gelatt, and M. P. Vecchi, 1983, Science **220**, 671.

Kitaev, A. Y., A. H. Shen, and M. N. Vyalyi, 2000, *Classical and Quantum Computation, Graduate Studies in Mathematics*, Vol. 47 (American Mathematical Society, Providence, RI).

Klarsfeld, S., and J. A. Oteo, 1989, J. Phys. A **22**, 4565.

Klauder, J. R., 1979, Phys. Rev. D **19**, 2349.

Kleinberg, J. M., R. Kumar, P. Raghavan, S. Rajagopalan, and A. S. Tomkins, 1999, Computing and Combinatorics, in *Proceedings of the 5th Annual International Conference, COCOON'99* (Springer, Berlin/Heidelberg), p. 1.

Knysh, S., 2016, Nat. Commun. **7**, 12370.

Knysh, S., and V. Smelyanskiy, 2010, arXiv:1005.3011.

Knysh, S., and V. N. Smelyanskiy, 2006, arXiv:cond-mat/0602257.

Kong, L., and E. Crosson, 2015, arXiv:1511.06991.

Krzakala, F., A. Rosso, G. Semerjian, and F. Zamponi, 2008, Phys. Rev. B **78**, 134428.

Kumar, V., A. Grama, A. Gupta, and G. Karypis, 2003, *Introduction to Parallel Computing* (Addison-Wesley, Reading, MA), 2nd ed.

Kurihara, K., S. Tanaka, and S. Miyashita, 2014, arXiv:1408.2035.

Landau, L. D., 1932, Phys. Z. Sowjetunion **2**, 46.

Landau, Z., U. Vazirani, and T. Vidick, 2015, Nat. Phys. **11**, 566.

Langville, A. N., and C. D. Meyer, 2006, *Google's PageRank and Beyond: The Science of Search Engine Rankings* (Princeton University Press, Princeton, NJ).

Latorre, J. I., and R. Orus, 2004, Phys. Rev. A **69**, 062302.

Laumann, C., A. Scardicchio, and S. L. Sondhi, 2008, Phys. Rev. B **78**, 134424.

Laumann, C. R., R. Moessner, A. Scardicchio, and S. L. Sondhi, 2012, Phys. Rev. Lett. **109**, 030502.

Laumann, C. R., R. Moessner, A. Scardicchio, and S. L. Sondhi, 2015, Eur. Phys. J. Spec. Top. **224**, 75.

Lidar, D. A., A. T. Rezakhani, and A. Hamma, 2009, J. Math. Phys. (N.Y.) **50**, 102106.

Liu, C.-W., A. Polkovnikov, and A. W. Sandvik, 2015, Phys. Rev. Lett. **114**, 147203.

Lloyd, S., M. Mohseni, and P. Rebentrost, 2014, Nat. Phys. **10**, 631.

Lloyd, S., and B. M. Terhal, 2016, New J. Phys. **18**, 023042.

Lucas, A., 2014, Front. Phys. **2**, 5.

Mandrà, S., Z. Zhu, and H. G. Katzgraber, 2017, Phys. Rev. Lett. **118**, 070502.

Mandrà, S., Z. Zhu, W. Wang, A. Perdomo-Ortiz, and H. G. Katzgraber, 2016, Phys. Rev. A **94**, 022337.

Mao, W., 2005a, Phys. Rev. A **71**, 060309.

Mao, W., 2005b, Phys. Rev. A **72**, 052316.

Margolus, N., 1990, in *Complexity, Entropy, and the Physics of Information*, SFI Studies in the Sciences of Complexity, Vol. VIII, edited by W. Zurek (Addison-Wesley, Reading, MA), pp. 273–287.

Marinari, E., and G. Parisi, 1992, Europhys. Lett. **19**, 451.

Marriott, C., and J. Watrous, 2005, J. Comput. Complex. **14**, 122.

Marzlin, K.-P., and B. C. Sanders, 2004, Phys. Rev. Lett. **93**, 160408.

Matsuda, Y., H. Nishimori, and H. G. Katzgraber, 2009, New J. Phys. **11**, 073021.

Meir, R., 2003, in *Advanced Lectures on Machine Learning*, Lecture Notes in Computer Science, Vol. 2600, edited by S. Mendelson and A. Smola (Springer, Berlin/Heidelberg), pp. 118–183.

Mertens, S., 1998, Phys. Rev. Lett. **81**, 4281.

Mertens, S., 2001, Theor. Comput. Sci. **265**, 79.

Mertens, S., 2003, arXiv:cond-mat/0310317.

Messiah, A., 1962, *Quantum Mechanics*, Vol. II (North-Holland Publishing Company, Amsterdam).

Mezard, M., G. Parisi, and M. A. Virasoro, 1987, *Spin Glass Theory and Beyond*, World Scientific Lecture Notes in Physics (World Scientific, Singapore).

Miller, J., and D. A. Huse, 1993, Phys. Rev. Lett. **70**, 3147.

Miyahara, H., and K. Tsumura, 2016, arXiv:1606.01484.

Mizel, A., 2004, Phys. Rev. A **70**, 012304.

Mizel, A., D. A. Lidar, and M. Mitchell, 2007, Phys. Rev. Lett. **99**, 070502.

Mizel, A., M. W. Mitchell, and M. L. Cohen, 2001, Phys. Rev. A **63**, 040302.

Mizel, A., M. W. Mitchell, and M. L. Cohen, 2002, Phys. Rev. A **65**, 022315.

Moussa, J. E., 2013, arXiv:1310.6676.

Movassagh, R., and P. W. Shor, 2016, Proc. Natl. Acad. Sci. U.S.A. **113**, 13278.

Muthukrishnan, S., T. Albash, and D. A. Lidar, 2016, Phys. Rev. X **6**, 031010.

Nagaj, D., 2008, arXiv:0808.2117.

Nagaj, D., and S. Mozes, 2007, J. Math. Phys. (N.Y.) **48**, 072104.

Nagaj, D., and P. Wocjan, 2008, Phys. Rev. A **78**, 032311.

Nagaj, D., P. Wocjan, and Y. Zhang, 2009, Quantum Inf. Comput. **9**, 1053.

Nakahara, M., 1990, *Geometry, topology and Physics* (Institute of Physics Publishing, London, UK).

Nayak, C., S. H. Simon, A. Stern, M. Freedman, and S. Das Sarma, 2008, Rev. Mod. Phys. **80**, 1083.

Nenciu, G., 1993, Commun. Math. Phys. **152**, 479.

Neven, H., V. S. Denchev, G. Rose, and W. G. Macready, 2008, arXiv:0811.0416.

Neven, H., V. S. Denchev, G. Rose, and W. G. Macready, 2009, arXiv:0912.0779.

Neven, H., G. Rose, and W. G. Macready, 2008, arXiv:0804.4457.

Nielsen, M., and I. Chuang, 2000, *Quantum Computation and Quantum Information*, Cambridge Series on Information and the Natural Sciences (Cambridge University Press, Cambridge, England).

Nishimori, H., 2001, *Statistical Physics of Spin Glasses and Information Processing: An Introduction* (Oxford University Press, Oxford, England).

Nishimori, H., 2016, arXiv:1609.03785.

Nussbaum, R., 2003, arXiv:math/0307056.

O'Gorman, B., R. Babbush, A. Perdomo-Ortiz, A. Aspuru-Guzik, and V. Smelyanskiy, 2015, Eur. Phys. J. Spec. Top. **224**, 163.

O'Hara, M. J., and D. P. O'Leary, 2008, Phys. Rev. A **77**, 042319.

Oliveira, R., and B. M. Terhal, 2008, Quantum Inf. Comput. **8**, 900.

Orús, R., and J. Latorre, 2004, Phys. Rev. A **69**, 052308.

Osborne, T. J., 2007, Phys. Rev. A **75**, 032321.

Osborne, T. J., and M. A. Nielsen, 2002, Phys. Rev. A **66**, 032110.

Osterloh, A., L. Amico, G. Falci, and R. Fazio, 2002, Nature (London) **416**, 608.

Östlund, S., and S. Rommer, 1995, Phys. Rev. Lett. **75**, 3537.

Papageorgiou, A., and J. F. Traub, 2013, Phys. Rev. A **88**, 022316.

Perdomo-Ortiz, A., S. E. Venegas-Andraca, and A. Aspuru-Guzik, 2011, Quantum Inf. Process. **10**, 33.

Piddock, S., and A. Montanaro, 2015, arXiv:1506.04014.

Powles, J. G., 1958, Proc. Phys. Soc. London **71**, 497.

Pudenz, K. L., and D. A. Lidar, 2013, Quantum Inf. Process. **12**, 2027.

Rajak, A., and B. K. Chakrabarti, 2014, Indian Journal of Physics **88**, 951.

Ramsey, F. P., 1930, Proc. London Math. Soc. **s2-30**, 264.

Ranjbar, M., W. G. Macready, L. Clark, and F. Gaitan, 2016, Quantum Inf. Process. **15**, 3519.

Rauber, T., and G. Rünger, 2010, *Parallel programming: For multicore and cluster systems* (Springer, New York).

Raussendorf, R., and H. J. Briegel, 2001, Phys. Rev. Lett. **86**, 5188.

Ray, P., B. K. Chakrabarti, and A. Chakrabarti, 1989, Phys. Rev. B **39**, 11828.

Raymond, J., S. Yarkoni, and E. Andriyash, 2016, arXiv:1606.00919.

Rebentrost, P., M. Mohseni, and S. Lloyd, 2014, Phys. Rev. Lett. **113**, 130503.

Reichardt, B. W., 2004, in *Proceedings of the Thirty-sixth Annual ACM Symposium on Theory of Computing, STOC '04* (ACM, New York, NY, USA), pp. 502–510, erratum: http://www-bcf.usc.edu/~breichar/Correction.txt.

Rezakhani, A. T., D. F. Abasto, D. A. Lidar, and P. Zanardi, 2010, Phys. Rev. A **82**, 012321.

Rezakhani, A. T., W. J. Kuo, A. Hamma, D. A. Lidar, and P. Zanardi, 2009, Phys. Rev. Lett. **103**, 080502.

Rezakhani, A. T., A. K. Pimachev, and D. A. Lidar, 2010, Phys. Rev. A **82**, 052305.

Ricci-Tersenghi, F., 2010, Science **330**, 1639.

Robson, J., 2001, "Finding a maximum independent set in time $O(2^{n/4})$," Technical Report (Laboratoire Bordelais de Recherche en Informatique).

Roland, J., and N. J. Cerf, 2002, Phys. Rev. A **65**, 042308.

Rønnow, T. F., Z. Wang, J. Job, S. Boixo, S. V. Isakov, D. Wecker, J. M. Martinis, D. A. Lidar, and M. Troyer, 2014, Science **345**, 420.

Rosenberg, G., P. Haghnegahdar, P. Goddard, P. Carr, K. Wu, and M. L. de Prado, 2016, IEEE J. Selected Topics in Signal Processing **10**, 1053.

Sachdev, S., 2001, *Quantum Phase Transitions* (Cambridge University Press, Cambridge, England).

Santra, S., O. Shehab, and R. Balu, 2016, arXiv:1602.08149.

Sarandy, M. S., and D. A. Lidar, 2005, Phys. Rev. Lett. **95**, 250503.

Sato, I., K. Kurihara, S. Tanaka, H. Nakagawa, and S. Miyashita, 2014, arXiv:1408.2037.

Schaller, G., 2008, Phys. Rev. A **78**, 032328.

Schaller, G., and R. Schützhold, 2010, Quantum Inf. Comput. **10**, 109.

Schollwöck, U., 2005, Rev. Mod. Phys. **77**, 259.

Schuch, N., and F. Verstraete, 2009, Nat. Phys. **5**, 732.

Seeley, J. T., M. J. Richard, and P. J. Love, 2012, J. Chem. Phys. **137**, 224109.

Seki, Y., and H. Nishimori, 2012, Phys. Rev. E **85**, 051112.

Seki, Y., and H. Nishimori, 2015, J. Phys. A **48**, 335301.

Seoane, B., and H. Nishimori, 2012, J. Phys. A **45**, 435301.

Shor, P. W., 1994, *Proceedings of the 35th Annual Symposium on Foundations of Computer Science* (IEEE, Washington, DC), p. 124.

Shor, P. W., 1997, SIAM J. Comput. **26**, 1484.

Simon, D. R., 1997, SIAM J. Comput. **26**, 1474.

Sinclair, A., and M. Jerrum, 1989, Inf. Comput. **82**, 93.

Siu, M. S., 2005, Phys. Rev. A **71**, 062314.

Smelyanskiy, V. N., S. Knysh, and R. D. Morris, 2004, Phys. Rev. E **70**, 036702.

Smelyanskiy, V. N., U. v. Toussaint, and D. A. Timucin, 2002, arXiv:quant-ph/0202155.

Smelyanskiy, V. N., U. V. Toussaint, and D. A. Timucin, 2001, arXiv:quant-ph/0112143.

Somma, R., and S. Boixo, 2013, SIAM J. Comput. **42**, 593.

Somma, R. D., D. Nagaj, and M. Kieferová, 2012, Phys. Rev. Lett. **109**, 050501.

Somorjai, R. L., 1991, J. Phys. Chem. **95**, 4141.

Stadler, P. F., W. Hordijk, and J. F. Fontanari, 2003, Phys. Rev. E **67**, 056701.

Suzuki, S., J. Inoue, and B. Chakrabarti, 2013, *Quantum Ising Phases and Transitions in Transverse Ising Models*, Lecture Notes in Physics (Springer-Verlag, Berlin), Vol. 862, 2nd ed.

Swendsen, R. H., and J.-S. Wang, 1986, Phys. Rev. Lett. **57**, 2607.

Swendsen, R. H., and J.-S. Wang, 1987, Phys. Rev. Lett. **58**, 86.

Teufel, S., 2003, *Adiabatic Perturbation Theory in Quantum Dynamics*, Lecture Notes in Mathematics (Springer-Verlag, Berlin), Vol. 1821.

Tong, D. M., K. Singh, L. C. Kwek, and C. H. Oh, 2005, Phys. Rev. Lett. **95**, 110407.

Tsuda, J., Y. Yamanaka, and H. Nishimori, 2013, J. Phys. Soc. Jpn. **82**, 114004.

Usadel, K., 1986, Solid State Commun. **58**, 629.

Valiant, L. G., 1990, Commun. ACM **33**, 103.

van Dam, W., M. Mosca, and U. Vazirani, 2001, *Proceedings of the 42nd IEEE Symposium on Foundations of Computer Science* (IEEE, New York), p. 279.

Van den Nest, M., 2013, Phys. Rev. Lett. **110**, 060504.

Venuti, L. C., T. Albash, D. A. Lidar, and P. Zanardi, 2016, Phys. Rev. A **93**, 032118.

Verstraete, F., and J. I. Cirac, 2004, arXiv:cond-mat/0407066.

Verstraete, F., V. Murg, and J. I. Cirac, 2008, Adv. Phys. **57**, 143.

Vidal, G., J. I. Latorre, E. Rico, and A. Kitaev, 2003, Phys. Rev. Lett. **90**, 227902.

Vinci, W., and D. A. Lidar, 2016, arXiv:1608.05764.

Weber, S. J., *et al.*, 2017, arXiv:1701.06544.

Wei, Z., and M. Ying, 2006, Phys. Lett. A **354**, 271.

Wen, J., Y. Huang, and D. Qiu, 2009, Int. J. Quantum. Inform. **07**, 1531.

Wen, J., and D. Qiu, 2008, Int. J. Quantum. Inform. **06**, 997.

White, S. R., 1992, Phys. Rev. Lett. **69**, 2863.

White, S. R., and R. M. Noack, 1992, Phys. Rev. Lett. **68**, 3487.

Wiebe, N., and N. S. Babcock, 2012, New J. Phys. **14**, 013024.

Wiebe, N., A. Kapoor, and K. M. Svore, 2014, arXiv:1412.3489.

Wolff, U., 1989, Phys. Rev. Lett. **62**, 361.

Wu, J.-d., M.-s. Zhao, J.-l. Chen, and Y.-d. Zhang, 2008, Phys. Rev. A **77**, 062114.

Wu, L. A., M. S. Sarandy, and D. A. Lidar, 2004, Phys. Rev. Lett. **93**, 250404.

Yao, A., 1993, in *Proceedings of the 34th IEEE Symposium on Foundations of Computer Science* (IEEE, New York), p. 352.

Ye, J., S. Sachdev, and N. Read, 1993, Phys. Rev. Lett. **70**, 4011.

Young, A. P., S. Knysh, and V. N. Smelyanskiy, 2008, Phys. Rev. Lett. **101**, 170503.

Young, A. P., S. Knysh, and V. N. Smelyanskiy, 2010, Phys. Rev. Lett. **104**, 020502.

Zalka, C., 1999, Phys. Rev. A **60**, 2746.

Zanardi, P., and M. Rasetti, 1999, Phys. Lett. A **264**, 94.

Zdeborová, L., and M. Mézard, **2008**a, J. Stat. Mech. 08, P12004.

Zdeborová, L., and M. Mézard, 2008b, Phys. Rev. Lett. **101**, 078702.

Zener, C., 1932, Proc. R. Soc. A **137**, 696.

Zeng, L., J. Zhang, and M. Sarovar, 2016, J. Phys. A **49**, 165305.

Zhang, B. H., G. Wagenbreth, V. Martin-Mayor, and I. Hen, 2017, Sci. Rep. **7**, 1044.

Zhu, Z., A. J. Ochoa, and H. G. Katzgraber, 2015, Phys. Rev. Lett. **115**, 077201.

Zintchenko, I., M. B. Hastings, and M. Troyer, 2015, Phys. Rev. B **91**, 024201.

Žnidarič, M., and M. Horvat, 2006, Phys. Rev. A **73**, 022329.

Zulkowski, P. R., and M. R. DeWeese, 2015, Phys. Rev. E **92**, 032113.

Correction: A subset of references was improperly tagged, causing their disappearance in the PDF, and has been fixed.