

Optimizing Variational Quantum Algorithms Using Pontryagin's Minimum Principle

Zhi-Cheng Yang,¹ Armin Rahmani,^{2,3} Alireza Shabani,⁴ Hartmut Neven,⁴ and Claudio Chamon¹

¹*Physics Department, Boston University, Boston, Massachusetts 02215, USA*

²*Department of Physics and Astronomy and Quantum Matter Institute, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z4*

³*Department of Physics and Astronomy, Western Washington University, 516 High Street, Bellingham, Washington 98225, USA*

⁴*Google Inc., Venice, California 90291, USA*

(Received 28 July 2016; revised manuscript received 18 March 2017; published 18 May 2017)

We use Pontryagin's minimum principle to optimize variational quantum algorithms. We show that for a fixed computation time, the optimal evolution has a bang-bang (square pulse) form, both for closed and open quantum systems with Markovian decoherence. Our findings support the choice of evolution ansatz in the recently proposed quantum approximate optimization algorithm. Focusing on the Sherrington-Kirkpatrick spin glass as an example, we find a system-size independent distribution of the duration of pulses, with characteristic time scale set by the inverse of the coupling constants in the Hamiltonian. The optimality of the bang-bang protocols and the characteristic time scale of the pulses provide an efficient parametrization of the protocol and inform the search for effective hybrid (classical and quantum) schemes for tackling combinatorial optimization problems. Furthermore, we find that the success rates of our optimal bang-bang protocols remain high even in the presence of weak external noise and coupling to a thermal bath.

DOI: [10.1103/PhysRevX.7.021027](https://doi.org/10.1103/PhysRevX.7.021027)

Subject Areas: Quantum Physics, Quantum Information

I. INTRODUCTION

Quantum annealing (QA) aims to solve computational problems by using a guided quantum drive. The dynamics is generated by a time-dependent Hamiltonian along a trajectory that ends at a final target Hamiltonian whose ground state contains the solution of the problem [1–3]. QA is based on the adiabatic theorem, which guarantees that if the Hamiltonian is changed sufficiently slowly, transitions to excited states are suppressed during the adiabatic evolution, thus preparing states that are close to the target ground state. Unfortunately, the adiabatic condition that ensures that the system remains in the instantaneous ground state leads to long time scales for the solution of the hard computational problem. Within the framework of adiabatic computation, there have been several theoretical proposals on the optimizations of the quantum adiabatic algorithm (QAA), such as heuristic guesses for the initial state [4], increasing the minimum gap [5,6], and the quantum adiabatic brachistochrone formulation [7].

The adiabatic trajectory is not the only path for reaching the ground state of a final Hamiltonian that encodes the solution of the computational problem. More generally, one could imagine many other paths, including those where the

Hamiltonian is varied rapidly, that land at the desired state or, of practical interest, reach low energy states. In fact, it has already been found that for certain hard instances of problems, fast nonadiabatic paths can sometimes prevent the system from getting stuck at local minima, thus improving the search results [8–10]. The variational quantum algorithm (VQA) is an example where one searches for such possible paths, using optimization of the outcome via the variation of a fixed number of parameters in the protocol. A hybrid machine, combining classical optimization and quantum evolution, optimizes the variational parameters. Such hybrid variational approaches have proved useful in the context of quantum state preparation [11–14]. Recently, Ref. [15] introduced a variational quantum eigensolver (VQE) for applications in quantum chemistry. This idea was further explored in Refs. [16–20] and experimentally tested in Refs. [21–23]. In a related approach [24–27], Farhi *et al.* introduced a quantum approximate optimization algorithm (QAOA) for combinatorial optimization problems based on a parametrized square-pulse ansatz for dynamical evolution of the solver.

In this paper, we make a connection between VQA and optimal control theory [28–31]. VQA is essentially an adaptive feedback control [32,33] of a quantum system with the objective function encoding the solution of a computational problem; see Fig. 1(a). It utilizes a hybrid system composed of a classical computer that searches for the optimal variational protocol using measurements done on a quantum machine that generates the final states

Published by the American Physical Society under the terms of the Creative Commons Attribution 3.0 License. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

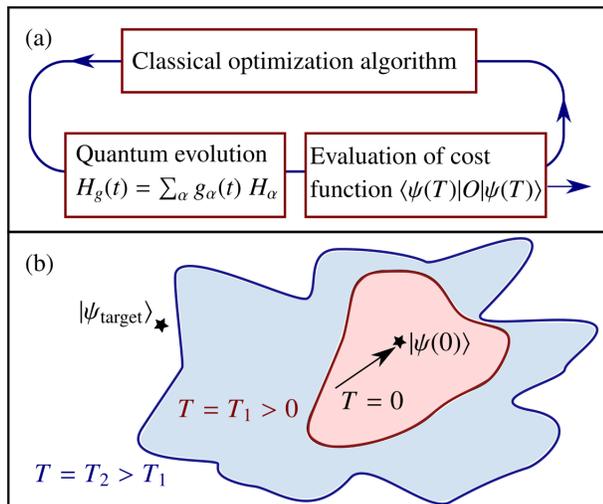


FIG. 1. (a) Variational quantum algorithm as a closed-loop learning control problem. (b) Increasing the total time expands the set of final states that one can reach with the variational protocols. The optimal protocol for a given time generates the closest state to a low energy target state within this reachable set.

corresponding to different variational protocols, via a closed-loop learning method [34].

Using Pontryagin’s minimum principle of optimal control, we show that the optimal protocol for VQA has a “bang-bang” form. Our results put the bang-bang ansatz of QAOA on a rigorous ground in contrast to VQA with a continuous-time evolution. A comparison of the performance of the optimal nonadiabatic bang-bang protocol with conventional (linear ramp) QAA demonstrates that the former significantly reduces error in the final state in the absence of noise or decoherence. The advantage over the linear ramp QAA survives weak dephasing white noise as well as weak coupling to a thermal bath. Furthermore, we perform a quantitative analysis of the characteristics of these optimal protocols. We numerically find a system-size independent distribution function for the duration of individual pulses, which may facilitate the development of effective algorithms for the classical optimizer through an efficient representation of the protocol with few variational parameters. Interestingly, each of the pulses in our bang-bang protocols contains commuting (either one-qubit or two-qubit) terms. Thus, our protocol can be implemented by applying a sequence of one-qubit gates [generated by the initial Hamiltonian, $g = 0$ in Eq. (9)] and two-qubit gates generated by the problem Hamiltonian ($g = 1$).

II. VARIATIONAL QUANTUM ALGORITHM

Consider a computational optimization problem such as finding a sequence of N bits that minimizes a certain function of all of the bits. To solve this problem with VQA,

we consider a system of N qubits with a parametrized Hamiltonian

$$H_g(t) = \sum_{\alpha} g_{\alpha}(t) H_{\alpha}. \quad (1)$$

Generically, we can cast the problem into generating a state $|\psi\rangle$ that minimizes a certain cost function, such as the expectation value of an operator O with respect to $|\psi\rangle$. A common example is finding the ground state of a disorderd classical Ising Hamiltonian [35], where the operator O is a Hamiltonian diagonal in the computational basis. In the context of quantum chemistry, VQE considers the operator O to be the Hamiltonian of a molecule [16].

The essence of VQA, as depicted in Fig. 1, is finding the time-dependent parameters $g_{\alpha}(t)$ over a time period T such that

$$|\psi(T)\rangle = \mathcal{T} e^{-i \int_0^T H_g(t) dt} |\psi(0)\rangle \quad (2)$$

minimizes a cost function $\langle \psi(T) | O | \psi(T) \rangle$. Generically, the controls $g_{\alpha}(t)$ belong to a permissible set determined by the experimental setup. A common such set is given by simple bounds, as shown in Eq. (3). The ideal solution could be a unique state $|\psi_{\text{target}}\rangle$ [as depicted in Fig. 1(b)] that is the ground state of the target Hamiltonian or, more generally, a set of states in the Hilbert space with an optimal figure of merit. One can either fix the initial state $|\psi(0)\rangle$ or add it to the list of the variational parameters (here, we fix it, motivated by experimental constraints). Generally, the longer the total time T , the closer we can get to an ideal solution.

One way to view this is to consider the reachable set, i.e., the set of all the final states one can reach by using one of the infinite number of permissible controls. The reachable set, naturally, grows with T (in fact, if $g_{\alpha} = 0$ is allowed, the reachable set for $T = T_1$ is strictly a subset of the reachable set for $T = T_2 > T_1$). As shown in Fig. 1(b), there could be a critical time beyond which the reachable set includes the target state and an exact solution is possible. There is no advantage in increasing T beyond this critical time. Generically, for smaller T , where the reachable set does not include the target state, the optimal protocols are highly constrained as they should prepare the closest point(s) of the reachable set to the target. For times longer than the critical time mentioned above, we expect an infinite number of protocols to produce the target as the evolution has extra time to meander in the Hilbert space. Our strategy is to fix T and find the best variational protocol $g_{\alpha}(t)$. If the solution is not acceptable, we can increase T . Next, we discuss how Pontryagin’s minimum principle from optimal control theory determines the form of optimal $g_{\alpha}(t)$ functions.

III. PONTRYAGIN'S MINIMUM PRINCIPLE APPLIED TO VQA

A. Bang-bang optimal protocols

The parameters in Hamiltonian Eq. (1) are generically constrained by their range

$$g_\alpha^{\min} \leq g_\alpha(t) \leq g_\alpha^{\max} \quad (3)$$

during the evolution $0 < t < T$. Equation (3) implies that, by assumption, each g_α can be tuned in the above range independently of the values of the other control parameters. For fixed initial state $|\psi(0)\rangle$, the coupling constants $g_\alpha(t)$ uniquely determine the final wave function. Consequently, the cost function, which we take as an arbitrary function of the final state, is a functional of $g_\alpha(t)$:

$$F[\{g_\alpha(t)\}] = \mathcal{F}(|\psi(T)\rangle). \quad (4)$$

Pontryagin's minimum principle [29] is directly applicable here. Briefly, this theorem states that, for a set of dynamical variables \mathbf{x} evolving from given initial values $\mathbf{x}(0)$ with the equations of motions $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{g})$, where \mathbf{g} are a set of control functions, the control functions \mathbf{g}^* that minimize an arbitrary function $\mathcal{F}[\mathbf{x}(T)]$ of the final values of the dynamical variable satisfy

$$\mathcal{H}(\mathbf{x}^*, \mathbf{p}^*, \mathbf{g}^*) = \min_{\mathbf{g}} \mathcal{H}(\mathbf{x}^*, \mathbf{p}^*, \mathbf{g}) \quad (5)$$

at any point in time and for each of the control functions. The optimal-control Hamiltonian is defined as $\mathcal{H}(\mathbf{x}, \mathbf{p}, \mathbf{g}) \equiv \mathbf{f}(\mathbf{x}, \mathbf{g}) \cdot \mathbf{p}$ for conjugate variables \mathbf{p} that evolve as $\dot{\mathbf{p}} = -\partial_{\mathbf{x}} \mathcal{H}$ with boundary conditions $\mathbf{p}(T) = \partial_{\mathbf{x}} \mathcal{F}[\mathbf{x}(T)]$. Here, the “*” superscript indicates the optimal solution corresponding to \mathbf{g}^* .

An important consequence of Eq. (5) is that if the equations of motion for \mathbf{x} , and consequently the optimal-control Hamiltonian \mathcal{H} , are linear in \mathbf{g} , generically, the optimal protocol is bang-bang; i.e., at any time during the evolution, we have $g_\alpha^*(t) = g_\alpha^{\min}$ or $g_\alpha^*(t) = g_\alpha^{\max}$. This follows from the fact that at any point in time we need to choose g_α to minimize $\mathcal{H}(\mathbf{x}^*, \mathbf{p}^*, \mathbf{g})$. If the sign of the coefficient of g_α in the optimal-control Hamiltonian is positive (negative), we should then choose the smallest (largest) g_α from the permissible range Eq. (3). In other words, the optimal protocol for each control function involves a sequence of sudden jumps between its minimum and maximum permissible values. The only caveat for the above argument is the possibility that the coefficient of a particular g_α in $\mathcal{H}(\mathbf{x}^*, \mathbf{p}^*, \mathbf{g})$ vanishes over a finite interval (since the sign of this coefficient determines whether we should choose the minimum or maximum value). We expect this special scenario to be nongeneric particularly for the disordered systems considered in the present paper.

In the quantum mechanical context, if the physical Hamiltonian is linear in the controls, the equations of motion and consequently the optimal-control Hamiltonian will also be linear, giving rise to bang-bang protocols, as verified in several recent studies on optimal topological quantum computing [36,37]. To find the protocol \mathbf{g} that minimizes the cost function in our case, we expand the wave function in a complete orthonormal basis, e.g., the computational basis $|z\rangle$ as $|\psi(t)\rangle = \sum_z A_z(t)|z\rangle$, and treat the real and imaginary parts of the amplitudes $A_z(t)$ as dynamical variables, which evolve according to the Schrödinger equation

$$\dot{A}_z^R = \frac{1}{2} \sum_{\alpha, z'} g_\alpha [(H_\alpha^{zz'} + H_\alpha^{z'z}) A_{z'}^I - i(H_\alpha^{zz'} - H_\alpha^{z'z}) A_{z'}^R], \quad (6)$$

$$\dot{A}_z^I = \frac{1}{2} \sum_{\alpha, z'} g_\alpha [-(H_\alpha^{zz'} + H_\alpha^{z'z}) A_{z'}^R - i(H_\alpha^{zz'} - H_\alpha^{z'z}) A_{z'}^I], \quad (7)$$

where $H_\alpha^{zz'} \equiv \langle z | H_\alpha | z' \rangle$ and $A_z^{R,I} \equiv \text{Re}, \text{Im}(A_z)$. Clearly, these equations of motion are linear in the control functions g_α , and the cost function Eq. (4) is a function of only the final values of the dynamical variables. Thus, the argument above holds and the optimal protocol is generically bang-bang regardless of the number of variational parameters. We remark that our optimal bang-bang protocol is non-adiabatic by construction, and we put no constraint on maximizing the degree of adiabaticity. The value of this result hinges upon the time scale over which a coupling constant is held fixed. The longer this time scale, the fewer parameters (switching times) are needed to represent the protocol. In fact, in the limit where this time scale goes to zero, any protocol can be approximated by a sequence of square pulses through Trotterization. In this paper, we find that the time scale above is indeed finite and is set by the energy scale of the Hamiltonian for the Sherrington-Kirkpatrick (SK) Ising spin-glass model [see Eq. (9)].

B. Presence of decoherence

From a practical point of view, it is important to assess the validity of the closed system findings in the presence of decoherence. Again, a straightforward application of the Pontryagin principle extends the above results for a closed system evolution to an open quantum system with Markovian dynamics described by a Lindblad equation,

$$\frac{d\rho}{dt} = -i \left[\sum_{\alpha} g_\alpha(t) H_{\alpha, \rho} \right] + \sum_{\beta} f_{\beta}(t) (2F_{\beta} \rho F_{\beta}^{\dagger} - \{F_{\beta}^{\dagger} F_{\beta}, \rho\}), \quad (8)$$

where the optimal protocol $\{g_\alpha(t), f_{\beta}(t)\}$, if controllable, is of type bang-bang. This is due to the linearity of the dynamical equation (8). A decoherence operator F_{β} can represent either noise in the Hamiltonian parameters (in

which case, F_β is Hermitian [38,39]) or an engineered bath [40]. In the former case, f_β 's are constant rates of noise processes, and in the latter, $f_\beta(t)$'s are control knobs that Pontryagin's minimum principle says should vary in bang-bang form for an optimal protocol. In the rest of the paper, we focus only on closed system Schrödinger dynamics when finding the optimal protocol. We do, however, discuss the effects of noise and open-system dynamics on our optimal protocols.

IV. VQA FOR THE SK SPIN-GLASS MODEL

We now focus on a canonical problem in combinatorial optimization, namely, the SK Ising spin glass [41] with the energy function

$$C = \frac{1}{\sqrt{n}} \sum_{i,j=1}^n J_{ij} \sigma_i^z \sigma_j^z + \sum_{i=1}^n h_i \sigma_i^z. \quad (9)$$

where J_{ij} and h_i are independent Gaussian random variables with zero mean and variance $J^2 = h^2 = 1$, and each σ^z spin can take the values ± 1 . The goal is to minimize C over all the 2^n spin configurations. A multitude of practical combinatorial optimization problems map to this model. The computational cost of finding the minimum with classical algorithms is exponential in n .

In analogy with the simple instances of quantum annealing, we focus on the case with only one control function $g(t)$ and use the following parametrized Hamiltonian:

$$H_g(t) = g(t)C + [1 - g(t)]B, \quad (10)$$

with the operator $B \equiv -\sum_{i=1}^n \sigma_i^x$ representing a transverse field, which generates quantum fluctuations.

For the initial state, we choose the ground states of B . It is easy to prepare the product state $|\psi(0)\rangle = \prod_i [(|\uparrow\rangle_i + |\downarrow\rangle_i)/\sqrt{2}]$ commonly used in other schemes, such as the QAA. Here, $\sigma_i^z |\uparrow\rangle_i = |\uparrow\rangle_i$ and $\sigma_i^z |\downarrow\rangle_i = -|\downarrow\rangle_i$. We would like to minimize the cost function $\langle \psi(T) | C | \psi(T) \rangle$. In the adiabatic scheme, a smooth ramp, such as $g(t) = (t/T)$, is applied for $0 < t < T$, and we can generate large overlap with the ground state of C in the limit of large T . Here, we allow for arbitrary time dependence of the control function in the fixed range $0 \leq g(t) \leq 1$. According to the general argument of Sec. III A, the optimal solution is bang-bang.

As we discuss in the Introduction, in VQA a classical optimization algorithm commands a quantum system to find the optimal protocol variationally from measurement of the cost function for many protocols. This requires many repetitions and it is to our advantage to use the shortest possible time T for which the final state has an acceptable overlap with the ground state of C (projective measurement is ultimately used in generating the ground state). In the

adiabatic scheme, we need only one shot, but there are important restrictions from the small energy gaps along the adiabatic trajectory, which can lead to exceedingly long time scales, over which quantum coherence cannot be even approximately sustained. Furthermore, the presence of noise or modulation in the control fields places important limitations on adiabatic schemes due to the emergence of the recently proposed noise-induced antiadiabaticity in the long-time limit [42]. Unlike QAA, which relies on the adiabatic theorem, VQA has no known connection to instantaneous ground states and the minimum gap to excitations as transitions to excited states during the time evolution are allowed as long as the system eventually lands at the ground state of the final Hamiltonian.

Given the limitations of the adiabatic scheme, a quantum approximate algorithm has been introduced for solving combinatorial optimization problems [24–27] in the spirit of VQA. The algorithm of Ref. [24] uses an ansatz

$$|\gamma, \beta\rangle = U(B, \beta_p) U(C, \gamma_p) \dots U(B, \beta_1) U(C, \gamma_1) |\psi(0)\rangle, \quad (11)$$

where the evolution operators are given by $U(C, \gamma) = e^{-i\gamma C}$ and $U(B, \beta) = e^{-i\beta B}$. The integer p is a parameter characterizing a variational ansatz.

For a given p , the goal of the algorithm is to find a set of variational parameters that minimizes the expectation value of $F_p(\gamma, \beta) = \langle \gamma, \beta | C | \gamma, \beta \rangle$, which ensures that the state $|\gamma, \beta\rangle$ approaches the ground state of C . Physically, the ansatz describes time evolution for a total time $T = \sum_{i=1}^p (\gamma_i + \beta_i)$, and a sequence of sudden switching between the Hamiltonians B and C . While this ansatz with a finite p is an intelligent guess, the result that we derive using Pontryagin's minimum principle implies that, given bounded independent control of Hamiltonian terms, the ansatz Eq. (11) is the optimal choice for a VQA approach to optimization. We reiterate that B and C are each a sum of commuting one- and/or two-qubit terms. Therefore, our protocol can be interpreted as a sequence of simple gates. Estimating the required p requires an analysis of the characteristic time scales of the pulses, which we carry out in this paper.

V. NUMERICAL STUDIES

We start by verifying for small system sizes and short annealing times that the optimal annealing protocol is indeed bang-bang, by using a Metropolis Monte Carlo (MC) algorithm, which makes no assumptions about the nature of the protocol. We divide the total time T into S slices of duration $\delta t = T/S$ and use a piecewise constant protocol. The method approaches an unbiased optimization, i.e., it explores all permissible controls and chooses the optimal one, if the protocols obtained converge upon increasing S . We then proceed by carrying out a MC

simulation starting from random initial protocols, without any assumption regarding the bang-bang nature of the protocol. In each step, we slightly change the control parameter g in a randomly chosen discretized time interval. If the cost function gets smaller, we accept the attempt; if the cost function gets larger, we accept the attempt with probability $e^{-\Delta E/T_{\text{MC}}}$, where T_{MC} is a fictitious temperature that is gradually reduced to zero.

In Fig. 2(a), we show the optimal protocol obtained from such MC simulation for a fixed instance of Hamiltonian Eq. (9) with $n = 5$ spins and total time $T = 0.8$.

Indeed, the MC simulation converge to a bang-bang protocol for different initial protocols in agreement with Pontryagin's minimum principle. Despite the convergence for short total time, the MC simulations often fail to converge for longer times and larger systems, signaling the difficulty of implementing VQA without any *a priori* knowledge about the form of the optimal protocol. However, based on the mathematical proof of the bang-bang nature of the optimal protocols, we can parametrize the protocol similar to QAOA [24] and use the durations of the pulses as variational parameters to be optimized with the interior-point minimization method (IPMM), increasing p to achieve convergence.

We check that IPMM results are indeed in agreement with MC results, e.g., Fig. 2(a) (it also runs much faster). In Fig. 2(b), we show a typical optimized protocol obtained with IPMM for a certain instance of Hamiltonian Eq. (9) with $n = 5$ spins and $T = 2$. Guided by MC results, we choose $\sim 20T$ variational parameters, which proves to be adequate (we converge to a smaller number of bangs than we allow in the ansatz).

We now turn to the critical question of the time scales of the pulses. We observe numerically and then argue analytically that the typical time scale of each bang is independent of the system size, and is only determined by some characteristic energy scale of Hamiltonian Eq. (9). Therefore, from a complexity theory perspective, this result implies that the hardness of the optimization problem

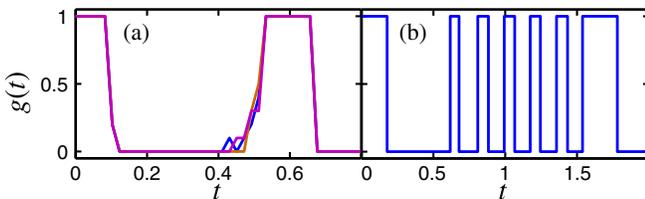


FIG. 2. (a) The optimal protocol obtained from MC simulations for a fixed instance of Hamiltonian Eq. (9) with $n = 5$ spins and total annealing time $T = 0.8$. Different colors represent different initial protocols. The plots are for $S = 40$, but the optimal protocol does not change upon increasing S . (b) A typical protocol obtained for a given instance of Hamiltonian Eq. (9) with $n = 5$ spins and $T = 2$, using a classical optimization solver. We start from a uniform initial protocol with S slices such that $\delta t = T/S = 0.1$.

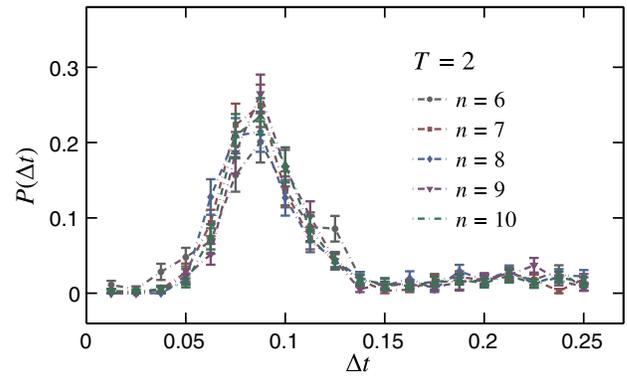


FIG. 3. The average probability distribution of the time scales of bangs for different system sizes: $n = 6, 7, 8, 9$, and 10 . The total annealing time is fixed to be $T = 2$, leading to an average success rate around 0.33 – 0.47 , depending on the system sizes. Each curve is averaged over 50 instances of Hamiltonian Eq. (9).

should translate into the number of pulses and/or the hardness of the protocol optimization. In Fig. 3, we plot the distribution of the time scales of each bang Δt for system sizes $n = 6, 7, 8, 9$, and 10 . For each system size, we fix the total annealing time to be $T = 2$, and average over 50 instances of the Hamiltonian Eq. (9).

We find that the distributions for the bang time collapse for different system sizes, and peak at almost the same value. This observation suggests a universal average distribution of the bang times for the near-optimal protocols, and a typical time scale (peak or average value) that is independent of the system size. Although we consider only a few system sizes, the dependence on n is extremely weak and we expect our results to extrapolate to large n .

Finally, we comment on the performance of our protocols. The cost function we minimize is the expectation value $\langle \psi(T) | C | \psi(T) \rangle$. Minimizing the energy expectation value results in larger overlap with the ground state of C . As expected, the time scales for our protocols are significantly shorter than those of the adiabatic algorithm with similar success rate. A comparison between the optimal bang-bang protocols and linear ramps $g(t) = t/T$ is shown in Fig. 4. The errors in the final wave functions $1 - |\langle \psi_{\text{GS}} | \psi(T) \rangle|^2$ and final energies $E - E_{\text{GS}} \equiv \langle \psi(T) | C | \psi(T) \rangle - E_{\text{GS}}$ are averaged over the 20 instances (out of 50 generated realizations) with the highest success rates for the optimal bang-bang protocol and linear QAA ramp, respectively. We find that, in the system sizes we consider in this work, the nonadiabatic bang-bang protocol with the same total time performs better than the linear ramp (commonly used in QAA) in the ideal case, where the thermal environment and external noise are neglected. Of course, in practice, one needs to include the overhead of searching for the optimal solution and understand how it scales as a function of system size. In particular, while implementing the bang-bang protocol consisting of square pulses on a quantum annealer is feasible [43], finding the optimal protocol with a

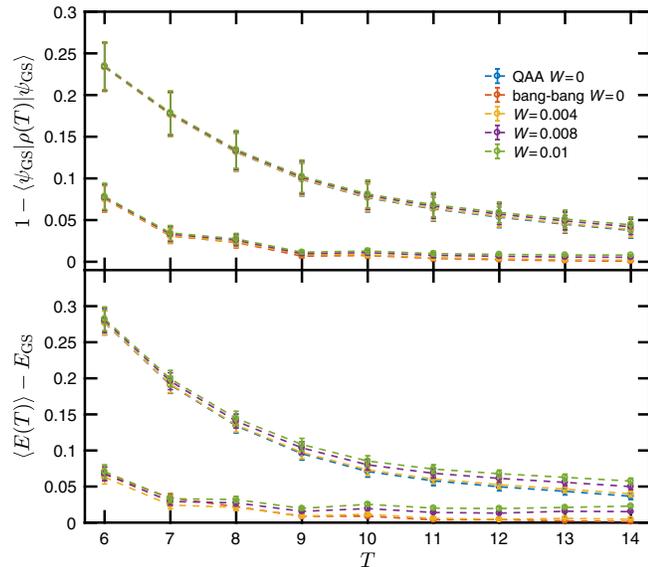


FIG. 4. Errors in the fidelity (upper panel) and final energies (lower panel) evolved with the bang-bang and QAA protocols in the presence of noise with different strengths for $n = 5$.

classical optimizer could be difficult for certain hard instances of problems.

VI. EFFECTS OF DISSIPATION AND DEPHASING

Real-world implementations of the bang-bang and QAA protocols are inevitably subject to noise either in the external controls or due to coupling to the thermal environment. Therefore, it is important to examine the effects of these perturbations on our optimal protocols for practical applications. Here, we consider two noise models in order to evaluate the robustness of our bang-bang protocol, at the same time comparing it with the performance of QAA.

A. Random dephasing noise

Here, we consider pure dephasing noise, where we introduce random fields in the x and z directions. This type of noise can capture noise induced by hardware electronics. Our error model can be viewed as the continuous time analog of the depolarizing channel commonly used for simulating noise in quantum circuits [44]. Since in a bang-bang protocol we have either $g(t) = 1$ or $g(t) = 0$ at any given time, we can write the stochastic Hamiltonian as

$$H(t) = C + \sum_{i=1}^n \delta h_i(t) \sigma_i^z + \sum_{i=1}^n \delta b_i(t) \sigma_i^x, \quad g=1, \quad (12)$$

$$H(t) = B + \sum_{i=1}^n \delta h_i(t) \sigma_i^z + \sum_{i=1}^n \delta b_i(t) \sigma_i^x, \quad g=0, \quad (13)$$

where $\delta h_i(t)$ and $\delta b_i(t)$ are noise in the z and x directions, respectively, with strengths independent of the value of the

coupling constants (additive noise). Assuming independent white noise for different terms with zero mean and second moments,

$$\overline{\delta h_i(t) \delta h_{i'}(t')} = W_h^2 \delta_{ii'} \delta(t - t'), \quad (14)$$

$$\overline{\delta b_i(t) \delta b_{i'}(t')} = W_b^2 \delta_{ii'} \delta(t - t'), \quad (15)$$

the noise-averaged density matrix evolves with the following master equation [39]:

$$\begin{aligned} \frac{d\rho(t)}{dt} = & -i[H, \rho(t)] - \frac{1}{2} W_h^2 \sum_{i=1}^n [[\rho(t), \sigma_i^z], \sigma_i^z] \\ & - \frac{1}{2} W_b^2 \sum_{i=1}^n [[\rho(t), \sigma_i^x], \sigma_i^x], \end{aligned} \quad (16)$$

where we take $W_b = W_h = W$ for simplicity. In the bang-bang case, the Hamiltonian H takes two different values, $H = C$ ($H = B$) for $g = 1$ ($g = 0$), while in the QAA case, H has the explicit time dependence of Eq. (10) with $g(t) = t/T$.

In Fig. 4, we show the errors in the fidelity $1 - \langle \psi_{GS} | \rho(T) | \psi_{GS} \rangle$ and final energy $\text{Tr}[\rho(T)C] - E_{GS}$ for different strengths of noise. We find that in the small W regime, the noise only slightly decreases the fidelity, acting like a perturbation without inducing any instability. The effects of the noise on the linear QAA ramp are similar both qualitatively and quantitatively, changing the $W = 0$ error by an amount of the same order of magnitude. For the strongest strength of noise that we study ($W = 0.01$), the fidelity of the optimal bang-bang protocol remains higher than that of the linear ramp protocol.

A comment is in order regarding the dimension of W and the range used. As $\delta(t - t')$ has a dimension of time (inverse energy), W^2 has a dimension of energy. Strictly speaking, the δ function introduces infinitely large (albeit completely uncorrelated) random fields. This is unrealistic. In real experiments there is a characteristic high frequency, introducing a characteristic short time scale $\Delta\tau$, over which noise is correlated. This frequency scale is typically several orders of magnitude larger than the characteristic energy of the Hamiltonian (it diverges for the δ function). Therefore, Eqs. (14) and (15) imply that $\delta h, \delta b \sim W/\sqrt{\Delta\tau}$, which means that for moderate noise in the random fields δh and δb , the corresponding values of W are suppressed by $\sqrt{\Delta\tau}$.

B. Weak thermal bath

Here, we consider coupling the system to a weak thermal bath at temperature $1/\beta$. In this regime, the dynamics of the open system can be approximately described by the Redfield master equation, which is commonly used to model noisy QAA for an actual annealing hardware

[45–47]. Here, we use the formulation in Ref. [45] and apply it to both QAA and bang-bang protocols.

The system of many qubits is coupled to the thermal bath via the Hamiltonian $\sum_i^n \sigma_i^z Q_i^z$, where Q_i^z are bath operators. We assume an Ohmic bosonic bath in thermal equilibrium, with the spectral density function given by

$$S_i^z(\omega) = \int_{-\infty}^{\infty} dt e^{i\omega t} \langle Q_i^z(t) Q_i^z(0) \rangle = \eta \omega \frac{1}{1 - e^{-\beta\omega}}, \quad (17)$$

where η is a dimensionless coefficient describing the strength of the coupling to the environment. We take the cutoff frequency of the bath to be infinite, so as to guarantee the Markovian assumption of dynamics. We employ Eqs. (4)–(9) in Ref. [45] to simulate the dynamics of open systems based on the Redfield master equation.

Figure 5 shows the errors in the fidelity for different strengths of coupling to the bath, for both QAA and bang-bang protocols. Similar to the case of the closed system in the presence of white noise, we find that the errors corresponding to both protocols change in an analogous manner due to weak coupling to the environment in both the short- and long-time regimes. There is an intermediate time regime $8.5 \lesssim T \lesssim 11.5$, where QAA exhibits remarkable robustness and a much smaller change in $\eta = 0$ error. However, the errors of the VQA and QAA get closer as we increase T . Once again, the fidelity of the optimal bang-bang protocol remains higher than the QAA even for open system dynamics.

VII. PULSE DURATION FROM PONTRYAGIN'S MINIMUM PRINCIPLE

Here, we provide more details on how Pontryagin's minimum principle can not only tell about the form of optimal solution for VQA but can also shed light on when the pulses should be switched on and off, in the context of the SK model.

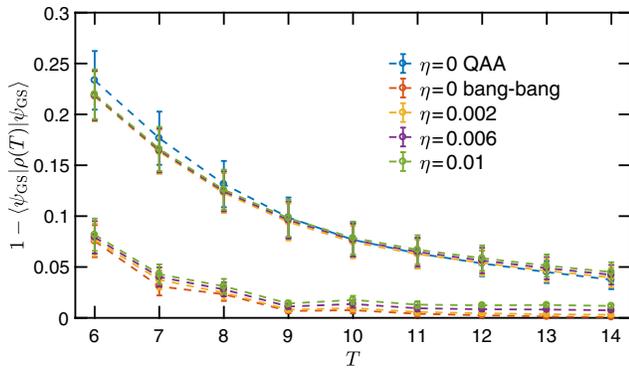


FIG. 5. Errors in the fidelity of final states evolved with the bang-bang and QAA protocols in the presence of different strengths of coupling to the environment for $n = 5$. The inverse temperature is chosen to be $\beta = 2/J$.

Using the computational basis $z = z_1, \dots, z_n$, we represent the wave function as $|\psi(t)\rangle = \sum_z A_z(t) |z\rangle$. The initial state with all the spins in the x direction corresponds to $A_z(0) = 1/\sqrt{2^n}$, and the Schrödinger equation reads

$$i\partial_t A_z(t) = g(t) C_z A_z(t) + [1 - g(t)] \sum_{k=1}^n A_{\bar{z}(k)}(t), \quad (18)$$

with $\bar{z}(k) = z_1, \dots, \bar{z}_k, \dots, z_n$, where \bar{z}_k represents a flipped spin with respect to z_k and C_z is the energy function we would like to minimize. In terms of the real and imaginary parts of $A_z(t) = R_z(t) + iI_z(t)$, we can then write

$$\partial_t R_z(t) = g(t) C_z I_z(t) + [1 - g(t)] \sum_{k=1}^n I_{\bar{z}(k)}(t), \quad (19)$$

$$\partial_t I_z(t) = -g(t) C_z R_z(t) - [1 - g(t)] \sum_{k=1}^n R_{\bar{z}(k)}(t). \quad (20)$$

Introducing conjugate momenta $P_z(t)$ and $Q_z(t)$, respectively, for the real and imaginary parts of $A_z(t)$, the explicit form of the optimal-control Hamiltonian is given by

$$\mathcal{H} = \sum_z \{g(t) C_z [P_z(t) I_z(t) - Q_z(t) R_z(t)] + [1 - g(t)] \sum_{k=1}^n [P_z(t) I_{\bar{z}(k)}(t) - Q_z(t) R_{\bar{z}(k)}(t)]\}. \quad (21)$$

The equations of motion for the conjugate momenta are $\partial_t P_z(t) = -\{\partial \mathcal{H} / [\partial R_z(t)]\}$ and $\partial_t Q_z(t) = -\{\partial \mathcal{H} / [\partial I_z(t)]\}$, which can be written explicitly as

$$\partial_t P_z(t) = g(t) C_z Q_z(t) + [1 - g(t)] \sum_{k=1}^n Q_{\bar{z}(k)}(t), \quad (22)$$

$$\partial_t Q_z(t) = -g(t) C_z P_z(t) - [1 - g(t)] \sum_{k=1}^n P_{\bar{z}(k)}(t), \quad (23)$$

where we use the relationships $\sum_{z,k} Q_z(t) R_{\bar{z}(k)}(t) = \sum_{z,k} Q_{\bar{z}(k)}(t) R_z(t)$ and $\sum_{z,k} P_z(t) I_{\bar{z}(k)}(t) = \sum_{z,k} P_{\bar{z}(k)}(t) I_z(t)$.

The cost function

$$F[g(t)] = \sum_z |A_z(T)|^2 C_z \quad (24)$$

leads to the following boundary conditions at $t = T$ for the conjugate momenta:

$$P_z(T) = 2R_z(T)C_z, \quad Q_z(T) = 2I_z(T)C_z. \quad (25)$$

Note that $g(t)$ uniquely determines $A_z(t)$. From $A_z(T)$ and the expression above, we can find $P_z(T)$ and $Q_z(T)$, solve the equations of motion backward in time, and determine the conjugate momenta as a function of time. Therefore, $g(t)$ also uniquely determines $P_z(t)$ and $Q_z(t)$. Potryagin's minimum principle states that the optimal protocol $g^*(t)$ satisfies

$$\mathcal{H}(g^*, \mathbf{R}^*, \mathbf{I}^*, \mathbf{P}^*, \mathbf{Q}^*) = \min_g \mathcal{H}(g, \mathbf{R}^*, \mathbf{I}^*, \mathbf{P}^*, \mathbf{Q}^*), \quad (26)$$

where \mathbf{R}^* , \mathbf{I}^* , \mathbf{P}^* , \mathbf{Q}^* are the corresponding optimal solution. As we argue in Sec. III, $g^*(t)$ is bang-bang and can take only two values: 0 and 1. Which value will depend on the sign of $[(\partial\mathcal{H})/\partial g]$ given by the expression

$$\frac{\partial\mathcal{H}}{\partial g} = \sum_z \left\{ C_z [P_z^*(t)I_z^*(t) - Q_z^*(t)R_z^*(t)] - \sum_{k=1}^n [P_z^*(t)I_{\bar{z}(k)}^*(t) - Q_z^*(t)R_{\bar{z}(k)}^*(t)] \right\}. \quad (27)$$

The sudden quenches at which $g(t)$ switches from 0 to 1 or vice versa correspond to the zeros of $[(\partial\mathcal{H})/\partial g]$ above.

Let us first combine P_z and Q_z into one complex momentum:

$$\Pi_z(t) = P_z(t) + iQ_z(t). \quad (28)$$

As we argue above, the optimal protocol is bang-bang with $g(t) = 0, 1$. In any interval with $g(t) = 1$, we can write

$$A_z(t) = e^{-iC_z(t-t_0)}A(t_0), \quad (29)$$

$$\Pi_z(t) = e^{-iC_z(t-t_0)}\Pi(t_0), \quad (30)$$

where t_0 is the beginning of the current bang $g(t_0) = 1$. We first note that the term $P_z(t)I_z(t) - Q_z(t)R_z(t) = \text{Im}[A_z(t)\Pi_z^*(t)]$ in Eq. (27) and does not change in intervals with $g(t) = 1$. Moreover, the terms

$$\begin{aligned} & P_z(t)I_{\bar{z}(k)}(t) - Q_z(t)R_{\bar{z}(k)}(t) \\ &= \text{Im}[e^{-i(C_z - C_{\bar{z}(k)})(t-t_0)}A_z(t_0)\Pi_{\bar{z}(k)}^*(t_0)]. \end{aligned} \quad (31)$$

The above equation allows for an estimation of the typical time scale of the interval $g(t) = 1$. Suppose at some t_0 , $g(t)$ switches from 0 to 1; i.e., a bang starts. From the discussion above, we know that at $t = t_0$ we must have $[(\partial\mathcal{H})/(\partial g)](t_0) = 0$. The time when the bang stops corresponds to the next $t > t_0$ when $[(\partial\mathcal{H})/(\partial g)](t) = 0$. More explicitly, in the interval $g(t) = 1$, Eq. (27) can be written as

$$w(t) = \sum_z \sum_{k=1}^n \text{Im}[(e^{-i(C_z - C_{\bar{z}(k)})(t-t_0)} - 1)A_z(t_0)\Pi_{\bar{z}(k)}^*(t_0)]. \quad (32)$$

We have $w(t_0) = 0$. The first root of the equation $w(t) = 0$ with $t > t_0$ then determines the duration of a pulse.

While we cannot derive analytically the average first root for $t > t_0$ from Eq. (32), one can see that the only time dependence in Eq. (32) within the current interval is in $e^{-i(C_z - C_{\bar{z}(k)})t}$. Thus, the energy difference $\Delta C_{z,k} \equiv C_z - C_{\bar{z}(k)}$, which has zero mean (as both C_z and $C_{\bar{z}(k)}$ have zero mean) and variance $\overline{\Delta C_{z,k}^2} = 4(J^2 + h^2)$, sets the characteristic time scale proportional to $1/\sqrt{J^2 + h^2}$ observed in Fig. 3. Importantly, this time scale is finite and system-size independent, distinguishing our bang-bang-type optimal protocol from the Trotterization of generic protocols, where the duration of individual pulses must be taken to zero. We believe the numerically obtained system-size independent distribution of Fig. 3 follows from Eq. (32), whose root determines one set of the switching times for quenching $g(t)$ from 1 to 0, giving the duration of a pulse with $g(t) = 1$ [the distributions of pulse durations with $g(t) = 0$ and $g(t) = 1$ are found numerically to be almost identical]. However, an analytical derivation of the distribution in Fig. 3 has remained elusive.

VIII. SUMMARY AND OUTLOOK

We show that the optimal VQA with bounded linear control parameters has a protocol of the bang-bang form. We verify this prediction by finding numerically the optimal protocol that minimizes the energy of a SK spin glass. The optimal nonadiabatic bang-bang protocols significantly reduce the error when compared to QAA within the same running time, and, at least for our system sizes, the advantage remains in the presence of weak additive white noise in the control parameters as well as weak coupling to a thermal environment.

Importantly, we show that the characteristic time scale between bangs is fixed by the energy scales in the problem and is independent of system size, which we confirm numerically. This finding significantly reduces the number of variational parameters in VQA, potentially decreasing the computational cost of the VQA outer-loop classical optimization algorithm to a great extent.

Our results, that the bang-bang protocols are optimal and the duration of each square pulse is size independent, inform the search for effective hybrid classical-quantum schemes for solving combinatorial optimization problems. Further progress relies on the development of efficient outer-loop algorithms as well as hardware development for quantum enhanced optimization. Ultimately, the power of our results lies in their application to larger systems, for which solving the time-dependent Schrödinger equation is

impossible on a classical computer. Rapid developments in quantum technologies [48], together with the relative robustness of our protocols to specific models of external noise and thermal environment, support the promise of such applications.

ACKNOWLEDGMENTS

We are grateful to Ryan Babbush, Pedram Roushan, Eduardo Mucciolo, Dries Sels, and Anatoli Polkovnikov for illuminating discussions. Z.-C. Y. would like to thank Quntao Zhuang for many early discussions on QAA. This work was supported by NSERC (A. R.), Max Planck-UBC Centre for Quantum Materials (A. R.), and DOE Grant No. DE-FG02-06ER46316 (C. C. and Z.-C. Y.).

-
- [1] T. Kadowaki and H. Nishimori, *Quantum Annealing in the Transverse Ising Model*, *Phys. Rev. E* **58**, 5355 (1998).
- [2] E. Farhi, J. Goldstone, S. Gutmann, and M. Sipser, *Quantum Computation by Adiabatic Evolution*, [arXiv:quant-ph/0001106](https://arxiv.org/abs/quant-ph/0001106).
- [3] E. Farhi, J. Goldstone, S. Gutmann, J. Laplan, A. Lundgren, and D. Preda, *A Quantum Adiabatic Evolution Algorithm Applied to Random Instances of an NP-Complete Problem*, *Science* **292**, 472 (2001).
- [4] A. Perdomo-Ortiz, S. E. Venegas-Andraca, and A. Aspuru-Guzik, *A Study of Heuristic Guesses for Adiabatic Quantum Computation*, *Quantum Inf. Process.* **10**, 33 (2011).
- [5] L. Zeng, J. Zhang, and M. Sarovar, *Schedule Path Optimization for Adiabatic Quantum Computing and Optimization*, *J. Phys. A* **49**, 165304 (2016).
- [6] Q. Zhuang, *Increase of Degeneracy Improves the Performance of the Quantum Adiabatic Algorithm*, *Phys. Rev. A* **90**, 052317 (2014).
- [7] A. T. Rezakhani, W.-J. Kuo, A. Hamma, D. A. Lidar, and P. Zanardi, *Quantum Adiabatic Brachistochrone*, *Phys. Rev. Lett.* **103**, 080502 (2009).
- [8] D. S. Steiger, T. F. Rønnow, and M. Troyer, *Heavy Tails in the Distribution of Time to Solution for Classical and Quantum Annealing*, *Phys. Rev. Lett.* **115**, 230501 (2015).
- [9] B. Heim, T. F. Rønnow, S. V. Isakov, and M. Troyer, *Quantum versus Classical Annealing of Ising Spin Glasses*, *Science* **348**, 215 (2015).
- [10] E. Crosson, E. Farhi, C. Y. Lin, H. H. Lin, and P. Shor, *Different Strategies for Optimization Using the Quantum Adiabatic Algorithm*, [arXiv:1401.7320](https://arxiv.org/abs/1401.7320).
- [11] W. Rohringer, R. Bücker, S. Manz, T. Betz, C. Koller, M. Gbel, A. Perrin, J. Schmiedmayer, and T. Schumm, *Stochastic Optimization of a Cold Atom Experiment Using a Genetic Algorithm*, *Appl. Phys. Lett.* **93**, 264101 (2008).
- [12] S. Rosi, A. Bernard, N. Fabbri, L. Fallani, C. Fort, M. Inguscio, T. Calarco, and S. Montangero, *Fast Closed-Loop Optimal Control of Ultracold Atoms in an Optical Lattice*, *Phys. Rev. A* **88**, 021601 (2013).
- [13] A. Rahmani, T. Kitagawa, E. Demler, and C. Chamon, *Cooling through Optimal Control of Quantum Evolution*, *Phys. Rev. A* **87**, 043607 (2013).
- [14] A. Rahmani, *Quantum Dynamics with an Ensemble of Hamiltonians*, *Mod. Phys. Lett. B* **27**, 1330019 (2013).
- [15] A. Peruzzo, J. McClean, P. Shadbolt, M. H. Yung, X. Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien, *A Variational Eigenvalue Solver on a Photonic Quantum Processor*, *Nat. Commun.* **5**, 4213 (2014).
- [16] M. Yung, J. Casanova, A. Mezzacapo, J. McClean, L. Lamata, A. Aspuru-Guzik, and E. Solano, *From Transistor to Trapped-Ion Computers for Quantum Chemistry*, *Sci. Rep.* **4**, 3589 (2014).
- [17] D. Wecker, M. B. Hastings, and M. Troyer, *Progress towards Practical Quantum Variational Algorithms*, *Phys. Rev. A* **92**, 042303 (2015).
- [18] D. Wecker, M. B. Hastings, and M. Troyer, *Training a Quantum Optimizer*, *Phys. Rev. A* **94**, 022309 (2016).
- [19] J. R. McClean, J. Romero, R. Babbush, and A. Aspuru-Guzik, *The Theory of Variational Hybrid Quantum-Classical Algorithms*, *New J. Phys.* **18**, 023023 (2016).
- [20] J. R. McClean, M. E. Schwartz, J. Carter, and W. A. de Jong, *Hybrid Quantum-Classical Hierarchy for Mitigation of Decoherence and Determination of Excited States*, *Phys. Rev. A* **95**, 042308 (2017).
- [21] Y. Shen, X. Zhang, S. Zhang, J. N. Zhang, M. H. Yung, and K. Kim, *Quantum Implementation of the Unitary Coupled Cluster for Simulating Molecular Electronic Structure*, *Phys. Rev. A* **95**, 020501(R) (2017).
- [22] C. Eichler, J. Mlynek, J. Butscher, P. Kurpiers, K. Hammerer, T. J. Osborne, and A. Wallraff, *Exploring Interacting Quantum Many-Body Systems by Experimentally Creating Continuous Matrix Product States in Superconducting Circuits*, *Phys. Rev. X* **5**, 041044 (2015).
- [23] P. O'Malley *et al.*, *Scalable Quantum Simulation of Molecular Energies*, *Phys. Rev. X* **6**, 031007 (2016).
- [24] E. Farhi, J. Goldstone, and S. Gutmann, *A Quantum Approximate Optimization Algorithm*, [arXiv:1411.4028](https://arxiv.org/abs/1411.4028).
- [25] E. Farhi, J. Goldstone, and S. Gutmann, *A Quantum Approximate Optimization Algorithm Applied to a Bounded Occurrence Constraint Problem*, [arXiv:1412.6062](https://arxiv.org/abs/1412.6062).
- [26] E. Farhi and A. W. Harrow, *Quantum Supremacy through the Quantum Approximate Optimization Algorithm*, [arXiv:1602.07674](https://arxiv.org/abs/1602.07674).
- [27] C. Y.-Y. Lin and Y. Zhu, *Performance of QAOA on Typical Instances of Constraint Satisfaction Problems with Bounded Degree*, [arXiv:1601.01744](https://arxiv.org/abs/1601.01744).
- [28] A. E. Bryson and Y.-C. Ho, *Applied Optimal Control: Optimization, Estimation and Control* (CRC Press, Boca Raton, FL, 1975).
- [29] L. S. Pontryagin, *Mathematical Theory of Optimal Processes* (CRC Press, Boca Raton, FL, 1987).
- [30] R. F. Stengel, *Optimal Control and Estimation* (Dover Publications, Mineola, NY, 1994).
- [31] C. Brif, M. D. Grace, M. Sarovar, and K. C. Young, *Exploring Adiabatic Quantum Trajectories via Optimal Control*, *New J. Phys.* **16**, 065013 (2014).
- [32] K. J. Astrom, *Adaptive Feedback Control*, *Proc. IEEE* **75**, 185 (1987).
- [33] C. Brif, R. Chakrabarti, and H. Rabitz, *Control of Quantum Phenomena: Past, Present and Future*, *New J. Phys.* **12**, 075008 (2010).

- [34] D. Dong and I. R. Petersen, *Quantum Control Theory and Applications: A Survey*, *IET Control Theory A* **12**, 2651 (2010).
- [35] A. Lucas, *Ising Formulations of Many NP Problems*, *Front. Phys.* **2**, 5 (2014).
- [36] T. Karzig, A. Rahmani, F. von Oppen, and G. Refael, *Optimal Control of Majorana Zero Modes*, *Phys. Rev. B* **91**, 201404 (2015).
- [37] A. Rahmani, B. Seradjeh, and M. Franz, *Topologically Protected Non-Abelian Braiding through Optimal Diabatic Dynamics*, arXiv:1605.03611.
- [38] H. Pichler, J. Schachenmayer, A. J. Daley, and P. Zoller, *Heating Dynamics of Bosonic Atoms in a Noisy Optical Lattice*, *Phys. Rev. A* **87**, 033606 (2013).
- [39] A. Rahmani, *Dynamics of Noisy Quantum Systems in the Heisenberg Picture: Application to the Stability of Fractional Charge*, *Phys. Rev. A* **92**, 042110 (2015).
- [40] H. P. Breuer and F. Petruccione, *The Theory of Open Quantum Systems* (Oxford University Press, New York, 2002).
- [41] D. Sherrington and S. Kirkpatrick, *Solvable Model of a Spin-Glass*, *Phys. Rev. Lett.* **35**, 1792 (1975).
- [42] A. Dutta, A. Rahmani, and A. del Campo, *Anti-Kibble-Zurek Behavior in Crossing the Quantum Critical Point of a Thermally Isolated System Driven by a Noisy Control Field*, *Phys. Rev. Lett.* **117**, 080402 (2016).
- [43] R. Barends *et al.*, *Digitized Adiabatic Quantum Computing with a Superconducting Circuit*, *Nature (London)* **534**, 222 (2016).
- [44] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, England, 2010).
- [45] M. H. S. Amin, C. J. S. Truncik, and D. V. Averin, *Role of Single-Qubit Decoherence Time in Adiabatic Quantum Computation*, *Phys. Rev. A* **80**, 022303 (2009).
- [46] T. Albash, S. Boixo, D. A. Lidar, and P. Zanardi, *Quantum Adiabatic Markovian Master Equations*, *New J. Phys.* **14**, 123016 (2012).
- [47] S. Boixo, V. N. Smelyanskiy, A. Shabani, S. V. Isakov, M. Dykman, V. S. Denchev, M. H. Amin, A. Y. Smirnov, M. Mohseni, and H. Neven, *Computational Multiqubit Tunneling in Programmable Quantum Annealers*, *Nat. Commun.* **7**, 10327 (2016).
- [48] M. Mohseni, P. Read, H. Neven, S. Boixo, V. Denchev, R. Babbush, A. Fowler, V. Smelyanskiy, and J. Martinis, *Commercialize Quantum Technologies in Five Years*, *Nature (London)* **543**, 171 (2017).