

Counterdiabaticity and the quantum approximate optimization algorithm

Jonathan Wurtz and Peter Love

Department of Physics and Astronomy, Tufts University, Medford, Massachusetts 02155, USA

The quantum approximate optimization algorithm (QAOA) is a near-term hybrid algorithm intended to solve combinatorial optimization problems, such as MaxCut. QAOA can be made to mimic an adiabatic schedule, and in the $p \rightarrow \infty$ limit the final state is an exact maximal eigenstate in accordance with the adiabatic theorem. In this work, the connection between QAOA and adiabaticity is made explicit by inspecting the regime of p large but finite. By connecting QAOA to counterdiabatic (CD) evolution, we construct CD-QAOA angles which mimic a counterdiabatic schedule by matching Trotter “error” terms to approximate adiabatic gauge potentials which suppress diabatic excitations arising from finite ramp speed. In our construction, these “error” terms are helpful, not detrimental, to QAOA. Using this matching to link QAOA with quantum adiabatic algorithms (QAA), we show that the approximation ratio converges to one at least as $1 - C(p) \sim 1/p^\mu$. We show that transfer of parameters between graphs, and interpolating angles for $p+1$ given p are both natural byproducts of CD-QAOA matching. Optimization of CD-QAOA angles is equivalent to optimizing a continuous adiabatic schedule. Finally, we show that, using a property of variational adiabatic gauge potentials, QAOA is at least counterdiabatic, not just adiabatic, and has better performance than finite time adiabatic evolution. We demonstrate the method on three examples: a 2 level system, an Ising chain, and the MaxCut problem.

1 Introduction

Recently, the quantum approximate optimization algorithm (QAOA) [1] has grown in interest as a possible algorithm to apply on noisy intermediate scale quantum (NISQ) devices [2, 3]. Given a parameterized wavefunction, a classical optimizer maximizes the expectation value of some objective function, which encodes approximate solutions to NP-HARD combinatorial optimization problems such as MaxCut. The ansatz alternates between unitaries generated by a “simple” H_S and “target” H_T Hamiltonian, which act p times on the maximal eigenstate of the simple Hamiltonian. The ansatz wavefunction, parameterized by $2p$ variational angles, γ, β is

$$|\gamma, \beta\rangle = \exp(i\beta_p H_S) \exp(i\gamma_p H_T) (\dots) \exp(i\beta_1 H_S) \exp(i\gamma_1 H_T) |0_S\rangle. \quad (1)$$

In the original introduction of QAOA, Farhi et. al. [1] show that in the limit $p \rightarrow \infty$, the QAOA is exact, in that the ansatz state can be in the subspace of the maximal eigenstates of the objective function. Their proof leverages the adiabatic theorem: Given a very slowly

Jonathan Wurtz: Corresponding author: jonathan.wurtz@gmail.com

time varying Hamiltonian, an initial eigenstate will evolve as an instantaneous eigenstate of the time varying Hamiltonian. Under adiabatic evolution, a simple-to-prepare ground state of an initial Hamiltonian is “dragged along” to prepare a ground state of a final Hamiltonian that encodes the solution to a combinatorial optimization problem [4, 5]. The adiabatic condition requires that the duration of the evolution be (roughly) bounded by the inverse gap squared [6]. In the $p \rightarrow \infty$ limit, a set of perturbatively small QAOA angles can satisfy the adiabatic condition by reproducing a Trotterized adiabatic evolution. A naïve choice of angles [7] are angles incrementally increasing and decreasing along each step q as $\gamma_q = \gamma_0(q/p)$ and $\beta_q = \beta_0(1 - q/p)$. If the angles are perturbatively small $\gamma_0, \beta_0 \ll 1$, the Trotter error is suppressed and the q th QAOA step reproduces a continuous time unitary for a time step $\tau = |\gamma_0| + |\beta_0|$ along an effective QAOA Hamiltonian

$$H_q = (q/p)H_T + (1 - q/p)H_S - \frac{i(q/p)(1 - q/p)\gamma_0\beta_0}{2(\gamma_0 + \beta_0)}[H_T, H_S] + \dots \quad (2)$$

where ellipsis denote higher order terms of a Baker-Campbell-Hausdorff (BCH) expansion. In the limit $\gamma_0, \beta_0 \rightarrow 0$, the higher-order BCH terms are zero, and the effective Hamiltonian reproduces a linear annealing schedule between simple and target Hamiltonian. In the limit $p \rightarrow \infty$, the total time $T \rightarrow \infty$, which satisfies the adiabatic condition and so the final QAOA ansatz state is an exact maximal state.

The $p \rightarrow \infty$ limit is excessive. It is the limit in which the angles for each step are perturbatively small to minimize the Trotter error, while the sum of all angles (eg. the total time) is large to enforce the adiabatic condition. To satisfy both, p must be very large. What happens when p is large but finite? This paper focuses on relaxing the $p \rightarrow \infty$ condition, computing and simulating QAOA in the large but fixed- p regime. It will be shown that QAOA angles at finite p can reproduce, and perform better than, a finite time adiabatic evolution. The main result of this work is that the Trotter “error” from alternating between terms in the generator can be a boon, not a bane, for a finite- p QAOA that mimics adiabatic protocols, by suppressing diabatic excitations. This leads to the fact that QAOA is at least *counterdiabatic*, not just *adiabatic*.

The Trotter error, to lowest order, gives a contribution to the effective QAOA Hamiltonian which is the commutator of the simple and target Hamiltonian $[H_T, H_S]$, where the magnitude of the angles γ, β give the relative strength of the Trotter error term. There is a trade off in the magnitude of these angles, as sketched in Fig. 1 for fixed p . If the angles are too small (Fig. 1 left), the adiabatic timescale is too short, and the performance is lessened by finite-speed diabatic excitations. If the angles are too large (Fig. 1 right), the Trotter error from combining digital terms into a combined unitary is large, accumulating error and reducing the performance. Somewhere in the middle (Fig. 1 center), these two error sources accumulate error minimally, resulting in a maximum performance.

This maximum can be found analytically by matching the Trotter “error” term to a low-order approximation of the adiabatic gauge potential (AGP), and thus show that the first-order Trotter error can act as a diabatic counter term to reduce excitations. The AGP and the broader concept of “Shortcuts to adiabaticity” (STA) [8] covered in section 3 aims to reduce finite-speed diabatic excitations [9, 10] by adding counter terms to the effective Hamiltonian. In doing so, the magnitude of QAOA angles is fixed by an analytically computed coefficient related to the variational AGP, and the total set of angles can be analytically computed. A counterdiabatic (CD) or STA Hamiltonian has better finite-time performance over analogous adiabatic-only protocols, and so these analytically-derived QAOA angles have better performance than an analogous adiabatic-only protocol.

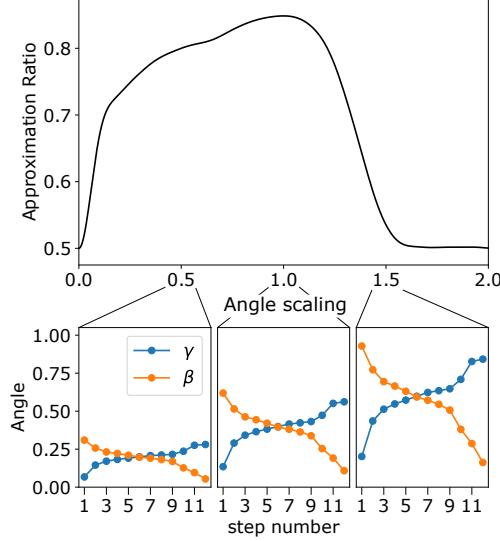


Figure 1: Optimizing CD-QAOA angles of fixed depth p . Naively, a smooth set of angles that mimic an adiabatic process leaves the relative magnitude of each angle undefined. If the angles are too small (left), the adiabatic timescale is too short and the state accumulates an excess of finite speed diabatic excitations, reducing the performance. If the angles are too large (right), the Trotter error of alternating between Hamiltonian terms accumulates, reducing the performance. Somewhere in between, these two error sources accumulate minimally, resulting in an optimum. We seek this optimum by matching the Trotter error with low order counterdiabatic terms. These data are variationally optimized $p = 12$ angles for a MaxCut graph, as further described in Fig. 6.

Crucially, we show that the second order BCH term can always be matched to a low order variational adiabatic gauge potential, as both the sign of the BCH term and the AGP coefficient are negative. The implication of this matching is that for counterdiabatic annealing protocols of total time T , there exists some finite $p(T)$ which is optimal. The optimal $p(T)$ connects performance scaling of adiabatic processes with QAOA scaling with p .

To demonstrate using the Trotter error as a counterdiabatic term, we construct an explicit procedure, which maps between continuous time counterdiabatic and STA protocols and discrete QAOA angles. The procedure is completely analytic, numerically implementable, and does not require quantum simulation. The forward procedure, given a counterdiabatic Hamiltonian $H_{CD}(t)$, is outlined as follows. Each QAOA step is expanded into a single unitary using a perturbative Baker-Campbell-Hausdorff (BCH) expansion, which is reasonably numerically tractable for perturbation order $\lesssim 5$. The generator of the step is matched with the effective counterdiabatic time evolution generator over some finite step in time, computed using a low order ($\lesssim 5$) Magnus expansion. The effective generators along each step q of QAOA are matched with the time evolution generators along some time step by classically variationally optimizing QAOA angles γ_q, β_q as well as time step τ_q to minimize some trace error. If the matching is close for every step, the QAOA ansatz state is close to the time evolved counterdiabatic state. If the first-order counterdiabatic term is exact, as is the case for the two level system, then an exact $p = 1$ QAOA can be derived using these methods, as shown in section 5.

Matching QAOA angles to finite-speed adiabatic schedules has several interesting corollaries. One useful implication is deriving the large- p behavior as the approximation ratio converges to 1. Kibble Zurek (KZ) theory [11] relates energy density to finite-speed an-

nealing through quantum critical points [12]. Using KZ scaling we show that the large- p QAOA behavior of the approximation ratio approaches unity at least polynomially

$$C(p) \sim 1 - p^{-\mu}. \quad (3)$$

Curiously, such scaling may be slower than optimal QAOA angles for a given p . We show in section 6 that $C(p) \sim 1 - p^{-1/2}$ for the Ising chain and finite-speed protocols. For optimal protocols [13], we show that $C(p) \sim 1 - p^{-1} \log(p)$. However, the optimal QAOA angles [1] go as $C(p) = 1 - 1/2p$. This scaling suggests that QAOA angles which mimic counterdiabatic evolution may not be optimal in all systems, and globally optimal protocols may be hard to find.

The map between continuous time counterdiabatic protocols and QAOA angles can be done in reverse. Given a set of smooth QAOA angles, an approximately equivalent continuous time counterdiabatic protocol with varying speeds $\lambda(t)$ and auxiliary fields $s(t)$ can be derived by variationally optimizing or matching each QAOA step with an equivalent continuous-time counterdiabatic Hamiltonian. This procedure will be shown in section 4 and demonstrated in section 7. Consequently, variational optimization of smooth QAOA angles in a large p limit is equivalent to variational optimization of continuous time counterdiabatic protocols [14]. The phenomena of inducing larger- p protocols [15] is also naturally explained by this mapping. Given an induced counterdiabatic protocol for some p , one can compute new angles for a different p' , which will have a similar “shape”, and similar performance.

Crucial to the matching procedure is the addition of some ancillary field s , which adds extra terms to the effective QAOA Hamiltonian

$$H(t) \mapsto H(t) + s(t)\tilde{H}(t). \quad (4)$$

Here, $\tilde{H}(t)$ slowly changes and includes terms from the BCH expansion, and $s(t) \rightarrow 0$ at the beginning and end of the protocol, so that in the $T \rightarrow \infty$ the final state is the exact ground state of the target Hamiltonian. This implies that the QAOA state, instead of following the ground state of the original Hamiltonian $\lambda H_S + (1 - \lambda)H_T$, follows the ground state of some other Hamiltonian which follows a more complicated path in a multi-dimensional parameter space. In this sense, matching QAOA with a simple Trotterized adiabatic path is inadequate to describe the effective evolution, and these extra terms are required.

Additionally, if two graphs have a similar low energy eigenspectrum, it is reasonable to expect that the induced optimal counterdiabatic annealing protocol for the first graph \mathcal{G} will have similar performance for the other graph \mathcal{G}' . Consequentially, the induced angles for a graph \mathcal{G}' given the counterdiabatic protocol induced from variationally optimized angles from graph \mathcal{G} will have similar performance. This counterdiabatic insight may explain the phenomena of transfer of parameters [16], where optimal angles for one graph are close to optimal for another. A more in-depth discussion of transfer of parameters in the counterdiabatic context is shown in sections 4 and 7.

Finally, we show that, in a certain sense, QAOA does better than an analogous continuous time adiabatic evolution. Given a total “angle budget” $T = \sum_q |\gamma_q| + |\beta_q|$, a set of QAOA angles which mimic an adiabatic schedule does better than a finite-time smooth evolution of total time T . This fact is a consequence of Pontryagin’s minimum principle [17, 18]. We show that more specifically, this is due to the fact that the sign of the variational adiabatic gauge potential is always the same as the first-order Trotter “error” term. This makes the effective QAOA evolution at least counterdiabatic, not just

adiabatic, which suppresses excitations and improves performance. This fact is discussed further in Sec. 8.

We comment that while this mapping is perturbatively exact in the large p limit, the method nonetheless produces protocols with good performance even for small $p \sim 1$. For instance, the method finds the exact $p = 1$ angles for a 2 level system, as shown in section 5, and comes very close to the exact angles for the Ising chain, as shown in section 6. These observations indicate that analytically computed CD-QAOA angles may be pre-computed without any variational feedback step [19] and used directly, even for the low depths $p \sim 5$ of tomorrow’s NISQ hardware.

2 Outline of the paper

The rest of the paper is as follows. Section 3 begins by introducing the adiabatic theorem, adiabatic protocols, and the concept of shortcuts to adiabaticity (STA). This section will also introduce the variational adiabatic gauge potential, which will be matched with the Trotter “error” terms. Section 4 introduces the method of matching continuous time counterdiabatic evolution with QAOA angles, as well as justifies the perturbative expansions of the matching procedure in the large p limit. Then, sections 5-7 demonstrates the methods on three examples. Section 5 outlines the method on a simple two level system, and show that the derived QAOA angles are exact for all $p \geq 1$ up to perturbative error, as expected. Section 6 outlines the method on the 1d Ising chain, where the variational AGP is no longer exact but the model is exactly solvable. Section 7 outlines the method on 3 regular graphs, where the goal is to find approximate solutions to MaxCut. This section will also illustrate reversing the CD-QAOA procedure to derive approximately optimal counterdiabatic protocols. Finally, sections 8-9 concludes with some discussion of the implications of CD-QAOA, and some interesting next steps for the method.

3 Adiabatic protocols and shortcuts to adiabaticity

An adiabatic protocol is generally a process which leave certain dynamical properties invariant [20]. For example, under time evolution which slowly changes the underlying Hamiltonian and eigensystem, an adiabatic process could leave the relative probability distribution of eigenstates of the time-evolved wavefunction unchanged. For this reason, a common implementation of adiabatic protocols is state preparation and quantum adiabatic algorithms (QAA) [21, 5]. Given some simple to prepare state $|0_S\rangle$ which is an eigenstate (usually ground or maximal) of a Hamiltonian H_S , parameters are slowly tuned to turn off H_S and turn on some target Hamiltonian H_T , where the (ground or maximal) eigenstate is the target state $|0_T\rangle$. A particular time dependent protocol which achieves this simply interpolates between the two with some time dependent fields $\lambda(t)$ and $s(t)$, such that $\lambda(0) = 0$, $\lambda(T) = 1$, and $s(0) = s(T) = 0$

$$H(t) = \lambda(t)H_T + (1 - \lambda(t))H_S + s(t)\tilde{H}(t). \quad (5)$$

Here, we also include some slowly changing auxiliary field $\tilde{H}(t)$ which may bypass gap closings to speed adiabatic annealing. If the rate of change of the Hamiltonian as a function of time is small in comparison to the energy gap between the eigenstate and neighboring states, the initial eigenstate is “dragged along” and persists as an eigenstate. The speed is characterized by the adiabatic condition $1/\dot{\lambda} \gg |\langle 0|\partial_t H|n\rangle|/\Delta_n^2$, where n

indexes any other instantaneous eigenstate, and Δ_n is the energy difference [6]. If the condition is violated, the eigenstate accumulates diabatic excitations which, undesirably, makes the state no longer an eigenstate. This condition can occur especially around phase transitions, where the gap becomes very small and diabatic excitations are unavoidable for large system sizes through processes such as the Kibble-Zurek mechanism (KZ) [11].

The field is usually taken to be a linear ramp $\lambda(t) = t/T$ for a total time T . Given a finite time to anneal between simple and target Hamiltonians a more general nonlinear field may be more advantageous, which is the goal of optimal control theory [22, 23, 18, 24]. Without constraint on the smoothness of the protocol, optimal protocols are “Bang-Bang” QAOA protocols in accordance with Pontryagin’s minimum principle [17] as long as the controls are nonsingular [25]. Similar to QAOA, good protocols may include a smooth ramp, plus a rapidly oscillatory term [26, 27]. Depending on the structure of the gap, an optimal smooth protocol may go fast where the gap is large, go slowly around phase transitions, where the gap is small [13, 28]. In this work we focus on second order phase transitions, characterized by smooth changes in the ground state. This is opposed to first order phase transitions, which are characterized by sharp changes in order parameter and exponentially small level crossings. However, it may be possible to change from a first to second order phase transition by eg. adding non-stoquastic terms to a stoquastic Hamiltonian [29].

Quantum computers are usually limited in the total time for annealing [30] and so it is advantageous to find fast protocols which add extra terms to the time dependence to avoid diabatic excitations or generally get a high-quality final state $|0_T\rangle$. This is the goal of Shortcuts to adiabaticity (STA) [8]: given control parameters $\lambda(t)$ and extra auxiliary fields $s(t)$, choose specific time dependent protocols which optimize the final state of the process, which may include non-adiabatic intermediary processes.

We will use two specific concepts from STA to construct QAOA angles. The first concept is path dependent protocols [23, 31] which choose nonzero values of auxiliary fields to find optimal paths between initial and final Hamiltonians in the multidimensional parameter space of control fields, potentially avoiding regions where diabatic excitations are likely. The second concept is counterdiabatic protocols [32], which add counter terms to the Hamiltonian to suppress diabatic excitations which occur from the finite speed evolution. Formally, the exact counter term is the adiabatic gauge potential (AGP) [33], which serves as the generator of parameter evolution of instantaneous eigenstates. For typical Hamiltonians, the AGP is exponentially large in norm and highly nonlocal, rendering any experimental implementation impossible [34]. For this reason, implementable counterdiabatic protocols must use approximate AGPs which are local. One choice are variational AGPs [35], which use some local parameterized ansatz $A(\alpha)$ for an approximation of the AGP and best match it with the exact AGP. This is done through a trace minimization of an action

$$\text{MIN: } S(\alpha) = \left\| \partial_\alpha H - i[A(\alpha), H] \right\|_{\mathcal{P}}, \quad (6)$$

which suppresses the diabatic excitations of each eigenstate, averaged over the positive semidefinite projector $\mathcal{P} \succcurlyeq 0$ [33]. The norm is defined as $\|Q\|_{\mathcal{P}} = \text{Tr}[\mathcal{P}Q^2]$. Note that if $\mathcal{P} = \mathbb{1}$ this trace minimization is analytically tractable for any number of qubits, using trace identities to avoid any computation in Hilbert space. Recently, there has been a proposal for an intuition-agnostic ansatz for the AGP as an expansion of commutators [27]

$$A(\alpha) = i\alpha_0[H, \partial_\lambda H] + i\alpha_1[H, [H, [H, \partial_\lambda H]]] + i\alpha_n[H, [\dots, [H, \partial_\lambda H] \dots]] + \dots \quad (7)$$

which is an expansion of Hasting's formulation of the adiabatic gauge potential for adiabatic continuation [36]. Using the Hamiltonian of Eq. (5), the lowest-order AGP from Eq. (7) is simply the commutator of the simple and target Hamiltonian $A = i\alpha_0[H_S, H_T]$. A critical observation that will be used repeatedly though this work is the fact that the lowest order AGP ansatz is equivalent to the lowest order of BCH expansion used to compute the effective generator of each QAOA step. The coefficient α of Eq. (6) can be optimized by setting $\partial_\alpha S(\alpha) = 0$. Assuming that the projector \mathcal{P} is diagonal in the instantaneous eigenbasis of H , using some trace identities the coefficient of the variational AGP is

$$A(\lambda) = i\alpha(\lambda)[H_T, H_S] \quad \text{with} \quad \alpha(\lambda) = \frac{\|[H, \partial_\lambda H]\|_{\mathcal{P}}}{\|[H, \partial_\lambda H], H\|_{\mathcal{P}}}. \quad (8)$$

Observe that the sign of α is always negative over all projectors $\mathcal{P} \succcurlyeq 0$. The numerator is the norm of an antihermitian operator, which is negative; the denominator is the norm of a Hermitian operator, which is positive. The sign of $\alpha \leq 0$ even if $\mathcal{P} = |0\rangle\langle 0|$, the instantaneous ground state, or any low-energy subspace. The norm only includes at most the sixth power of the Hamiltonian and so the AGP is only aware of local structure of the system, albeit over a larger neighborhood than the Hamiltonian terms.

The approximate AGP can be added to the Hamiltonian as a counter term with a strength proportional to the speed $\dot{\lambda}$ [33]. Additionally, we add an auxiliary field s that is also the same commutator term. The effective time evolution operator (eg, counterdiabatic Hamiltonian) becomes

$$H_{\text{CD}}(\lambda, \dot{\lambda}, s) = \lambda H_T + (1 - \lambda) H_S + i(s + \dot{\lambda}\alpha(\lambda))[H_T, H_S]. \quad (9)$$

In the limit $T \rightarrow 0$, the counterdiabatic Hamiltonian is equivalent to evolving in parameter space $\lambda \in 0 \rightarrow 1$ with respect to the variational AGP, and in the limit $T \rightarrow \infty$ and $s = 0$ Eq. (9) recovers the original adiabatic schedule of Eq. (5). It is a physically reasonable conjecture to expect that adding an approximate adiabatic gauge potential to the Hamiltonian as a counter term will decrease the amount of diabatic excitations to extremal states. Even though the variational parameter is optimized to minimize excitations over all eigenstates (Eq. (6)), it has been found that the AGP still works in the $T \rightarrow 0$ limit for extremal eigenstates of interacting models [37]. Similarly, it is reasonable to expect that a local approximation of the AGP will suppress transitions far apart in energy [36], and that a local AGP will generally mimic the action of the full AGP away from phase transitions. Reducing diabatic excitations will minimize energy variance and thus increase the value of the objective function (for optimization) or decrease energy density (for ground states), resulting in better performance of the algorithm.

Note that this reasonable conjecture may be false in certain cases. If the AGP is optimized for all eigenstates, and it may be the case that optimizing for low energy or ground states only may result in a smaller value of α . Given an extreme enough difference, a counterdiabatic evolution may “overcorrect” and result in negative performance for the ground state, while having on average positive performance over all eigenstates. A characterization of such systems is beyond the scope of this paper. Nonetheless, the value of $\alpha \leq 0$ even for the ground state, so an infinitesimal value of α will always improve performance.

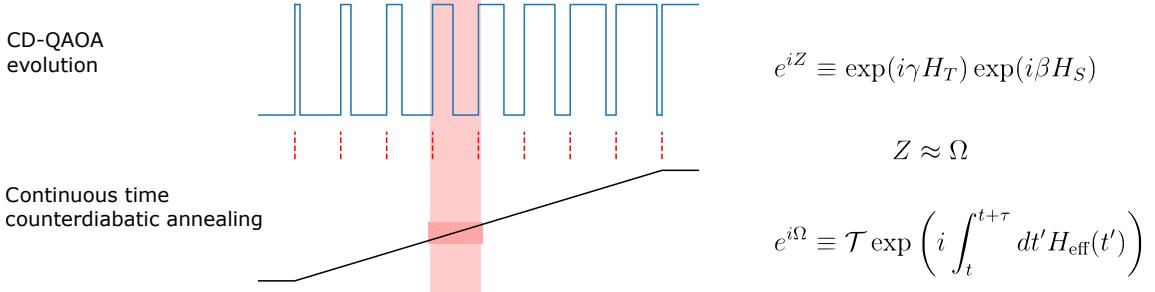


Figure 2: Schematic diagram of the CD-QAOA angle matching procedure. As input, the procedure receives the counterdiabatic annealing protocol (Black line), effective Hamiltonian H_{eff} , and number of QAOA steps p . For each step of the CD-QAOA evolution (pink), the effective unitaries over the interval are matched by minimizing the difference of their generators (right). This procedure sets the value of each angle γ_q, β_q in the QAOA (blue). By best matching along each step, the procedure fixes the angles $\{\gamma, \beta\}$ and total adiabatic annealing time T .

4 Counterdiabatic QAOA

We will now derive the mapping between continuous time counterdiabatic annealing, and QAOA angles. The QAOA ansatz of Eq. (1) is a set of unitaries that alternates between the simple Hamiltonian and the target Hamiltonian, which act on the maximal eigenstate of the simple Hamiltonian. In general, the $2p$ optimal angles $\{\gamma, \beta\}$ may be arbitrary, with the difference of angles between steps being large. A subset of QAOA angles are those which are smooth, where the difference of angles between steps is small. Such QAOA angles mimic an adiabatic process by smoothly interpolating between β large γ small at the beginning, and β small γ large at the end. Smooth QAOA angles are herein defined as a counterdiabatic QAOA (CD-QAOA). A set of CD-QAOA angles is identified by best matching the unitary generated by one step of QAOA, with the unitary generated by time evolution with the counterdiabatic Hamiltonian between some time interval $[t, t + \tau]$, for each QAOA step q

$$\exp(i\gamma H_T) \exp(i\beta H_S) \equiv e^{iZ} \approx e^{i\Omega} \equiv \mathcal{T} \exp\left(i \int_t^{t+\tau} dt' H_{\text{eff}}(t')\right). \quad (10)$$

A schematic of this matching is shown in Fig. 2. Z and Ω are the generators of each unitary, which may be defined perturbatively in a Baker-Campbell-Hausdorff (BCH) expansion [38] and Magnus expansion [39], respectively.

To lowest order of small time steps $\tau \ll 1$, the values of γ and β can be computed by de-Trotterizing the QAOA step by computing Z to the lowest order in BCH expansion, and computing Ω to the lowest order in the Magnus expansion. The lowest order BCH is simply a sum of the terms, and the lowest order Magnus expansion is simply the time average of the time dependent adiabatic-only Hamiltonian. For some small time step $\tau \ll 1$ the generators are matched by fixing angles

$$Z \approx \Omega, \quad (11)$$

$$\gamma H_T + \beta H_S \approx \tau \lambda(t) H_T + \tau (1 - \lambda(t)) H_S; \quad (12)$$

$$\Rightarrow \gamma = \tau \lambda(t). \quad (13)$$

$$\Rightarrow \beta = \tau (1 - \lambda(t)). \quad (14)$$

The Magnus and BCH expansion, as well as matching generators instead of unitaries,

works in the perturbative limit $\tau \rightarrow 0$. For sufficiently large T to suppress diabatic excitations from finite-time adiabatic evolution, and sufficiently small τ to suppress Trotterization errors, one can construct an adiabatic QAOA protocol with $p = T/\tau$ steps which asymptotically returns the ground state of H_T . This is the argument of Farhi et al. in the original introduction of QAOA [1]. However, this protocol leaves the relative magnitude of angles $\{\gamma, \beta\}$ free, and requires the limit $\tau \rightarrow 0$ to suppress higher order BCH error.

The angles can be fixed, and BCH error can be used advantageously, by using the counterdiabatic insights developed in section 3. Given some fixed finite p , one can still construct approximately adiabatic QAOA parameters as follows. First, instead of using the adiabatic-only Hamiltonian of Eq. (5), one can use the counterdiabatic Hamiltonian of Eq. (9) as the generator of time evolution. To better match generators, one can extend the effective QAOA evolution generator to the next order in the BCH expansion. Angles can then be fixed by matching the effective generators; observe that, crucially, the second order BCH term is the same as the approximate AGP and auxiliary term

$$Z \approx \Omega, \quad (15)$$

$$\gamma H_T + \beta H_S - \frac{i\gamma\beta}{2} [H_T, H_S] \approx \tau \bar{\lambda} H_T + \tau(1 - \bar{\lambda}) H_S + i\tau(\bar{s} + \dot{\lambda}\bar{\alpha}) [H_T, H_S]; \quad (16)$$

$$\Rightarrow \gamma = \tau \bar{\lambda}, \quad (17)$$

$$\Rightarrow \beta = \tau(1 - \bar{\lambda}), \quad (18)$$

$$\Rightarrow \frac{-\gamma\beta}{2} = \tau \left(\bar{s} + \dot{\lambda}\bar{\alpha} \right). \quad (19)$$

Matching the second order BCH expansion to the CD Hamiltonian sets not only the angles to fixed values, but also the time step

$$\tau = \frac{-2(\bar{s} + \dot{\lambda}\bar{\alpha})}{\bar{\lambda}(1 - \bar{\lambda})} ; \quad \gamma = \frac{-2(\bar{s} + \dot{\lambda}\bar{\alpha})}{1 - \bar{\lambda}} ; \quad \beta = \frac{-2(\bar{s} + \dot{\lambda}\bar{\alpha})}{\bar{\lambda}}. \quad (20)$$

Note that $\alpha < 0$ by Eq. (8), and so this solution always exists, as long as \bar{s} is not too large. In order for the expansions to be controlled, the next order terms must be much smaller. For the BCH expansion the perturbative criterion is that the third order term (see Eq. (48)) is much smaller than the second

$$\left| \frac{\gamma^2\beta}{12} \right| + \left| \frac{\beta^2\gamma}{12} \right| \ll \left| \frac{\gamma\beta}{2} \right| \Rightarrow \left| \frac{s + \dot{\lambda}\alpha}{3(1 - \lambda)} \right| + \left| \frac{s + \dot{\lambda}\alpha}{3\lambda} \right| \ll 1. \quad (21)$$

For the Magnus expansion the perturbative criterion is that the second order term (See Eq. (50)) is much smaller than the first

$$\frac{\dot{\lambda}\tau^3}{12} \|[H_T, H_S]\| \ll \tau(\|H_T\| + \|H_S\|) \Rightarrow \frac{\dot{\lambda}(s + \dot{\lambda}\alpha)^2}{3\lambda^2(1 - \lambda)^2} \ll \frac{\|H_T\| + \|H_S\|}{\|[H_T, H_S]\|}. \quad (22)$$

For $T \rightarrow \infty$, $s \ll 1$, and away from the edges of the protocol, both criterion are satisfied and the expansions are valid. This link allows relating the total time to p . From Eq. (20), the time for each QAOA step τ goes as $\dot{\lambda}\alpha$ for $s = 0$. Given $\dot{\lambda} \sim 1/T$, the total time $T = \sum \tau$ is quadratic in p . For $s < 0$ and $T \rightarrow \infty$, the time for each QAOA step τ goes as s , as $s \gg \dot{\lambda}\alpha$ and the total time $T = \sum \tau$ is linear in p

$$-\alpha T^2 \sim p \quad ; \quad -sT \sim p. \quad (23)$$

This scaling is easily verified with dimensional analysis: α has units of inverse energy squared, and so ($\hbar = 1$) time must be quadratic in p . s has units of inverse energy, and so time must be linear in p . Crucially, the sign of α or s must be negative in order to have a positive value of p .

4.1 General Counterdiabatic QAOA

The matching procedure of Eq. (16) can be made more general by including arbitrarily high orders of BCH and Magnus expansions. For the q th step in the counterdiabatic annealing schedule between times $[t_q, t_q + \tau_q]$, compute the effective generators $Z(\gamma_q, \beta_q)$ and $\Omega(t_q, \tau_q)$ through the BCH and Magnus expansions to appropriately high order. Then, make the equality of Eq. (10) as close as possible by minimizing the size of the difference $Z - \Omega$.

$$\exp(iZ(\gamma_q, \beta_q)) \approx \exp(i\Omega(t_q, \tau_q)) \Rightarrow \exp(iZ - i\Omega) \approx 1. \quad (24)$$

For each step, there is some error e_i which adds to the effective counterdiabatic time evolution

$$e_q(\gamma_q, \beta_q, \tau_q) = \frac{\sqrt{\text{Tr}[(Z - \Omega)^2]}}{\tau_i \sqrt{N}}. \quad (25)$$

The effective time evolution Hamiltonian which best mimics the QAOA angles is

$$H_{\text{eff}}(t) = \lambda(t)H_T + (1 - \lambda(t))H_S + i(s(t) + \dot{\lambda}\alpha(\lambda, s))[H_T, H_S] + e(t)H_E(t), \quad (26)$$

where $e(t) \approx e_q$ for $t \in [t_q, t_q + \tau_q]$ and $H_E \propto Z - \Omega$, with norm $\|H_s\| = N$ is the extra “error” term. If $e_q \approx 0$ for each QAOA step then the unitaries mimic the continuous time version, and final state of the CD-QAOA is that of the counterdiabatic time evolution. To make the effective evolution close to the counterdiabatic evolution, one can minimize the total error

$$\text{MIN: } \sum_q e_q(\gamma_q, \beta_q, \tau_q) \quad (27)$$

by variationally minimizing $3p$ parameters $\gamma_q, \beta_q, \tau_q$. This minimization completely fixes the CD-QAOA angles $\{\gamma, \beta\}$, as well as optimizes the best total evolution time $T = \sum_q \tau_q$ of the analogous CD annealing for given p . The trace can be computed efficiently using Pauli identities eg $\text{Tr}[(\sigma_x)^2] = 1$ and $\text{Tr}[\sigma_x] = 0$. The computation of $Z(\gamma, \beta)$ and $\Omega(t, \tau)$ may be unwieldy for large orders of expansion: the number of terms generically grow exponentially with the expansion order. However, such expansions are reasonably tractable with computational tools using Pauli identities, for low order ~ 5 . Eq. (27) reduces to Eq. (20) for a 2nd order BCH and first order Magnus expansion. The optimization step can also be done with reasonable efficiency using a two-step procedure. For a fixed total time T , each term e_q is optimized in a serial manner for each step q to find optimal values of time τ_q , and angles γ_q, β_q . These optimal values will depend on the choice of total time T , due to the difference in speed and thus magnitude of the counterdiabatic term. This procedure leaves the total number of steps p undefined, and will generally “overshoot” the total fixed time T . As such, in an outer loop the total time T can be optimized through

a binary search to find the optimal T for a particular choice of p . As τ changes with T , and one should choose a fixed p , this optimization is not as immediate as simply setting $T = \sum_q \tau_q$ after one step. In principle, Eq. 27 can be optimized simultaneously, but this is usually much slower as a general optimizer is not aware of the specific structure of the function.

Thus, using minimization, CD-QAOA angles can be derived from some counterdiabatic continuous time protocol to arbitrary orders in perturbation theory. For some fixed p , given some counterdiabatic schedule from Eq. (9) and counterdiabatic coefficient $\alpha(\lambda)$ from Eq. (8), a set of $2p$ QAOA angles $\{\gamma, \beta\}$ and final counterdiabatic protocol time T can be found by trace minimization.

4.2 Reverse Counterdiabatic QAOA

The matching procedure can be done in reverse. Given some set of smooth CD-QAOA angles $\{\gamma, \beta\}$, one can minimize Eq. (27), leaving γ and β fixed but variationally optimizing the time interval τ , as well as the protocol $\lambda(t)$ and $s(t)$ to some smooth function

$$\underset{\lambda(t), s(t), \tau_q}{\text{MIN:}} \quad \sum_q e_q(\gamma_q, \beta_q, \tau_q). \quad (28)$$

The minimization optimizes both the total counterdiabatic annealing time $T = \sum \tau_q$ for given angles $\{\gamma, \beta\}$, as well as generally yields nonlinear counterdiabatic annealing schedules. As will be discussed further in the example of section 7, optimizing a set of smooth CD-QAOA angles is analogous to optimizing a nonlinear annealing schedule $\lambda(t), s(t)$. To lowest order, the protocol can be derived piecewise by matching a first order Magnus expansion to a second order BCH expansion along each step

$$\begin{aligned} \lambda(t) &\Leftrightarrow \frac{1}{\tau_p} \int_{t_p}^{t_p + \tau_p} \lambda(t) dt = \frac{\gamma_p}{\gamma_p + \beta_p} \quad ; \quad \tau_p = \gamma_p + \beta_p \quad ; \quad t_p = \sum_{q < p} \tau_p, \quad (29) \\ s(t) &\Leftrightarrow \frac{1}{\tau_p} \int_{t_p}^{t_p + \tau_p} s(t) dt = \frac{-\gamma_p \beta_p}{2\tau_p} - \left(\frac{\lambda_p - \lambda_{p-1}}{\tau_p} \right) \alpha(\lambda_p), \end{aligned}$$

using a low-order polynomial interpolation. The term $s(t)$ includes removing the counterdiabatic contribution $\alpha(\lambda)$. Not subtracting out the counterdiabatic term will result in an incorrect value of $s(t)$, as protocols with different p and thus total annealing times will result in different magnitudes of $\dot{\lambda}$ and thus different prefactors of the equivalent 1st order BCH expansion term. If the magnitude $\dot{\lambda} \ll 1$ is small for large p and T , the counterdiabatic contribution is small. For intermediate p , especially near the edges, excluding the counterdiabatic term will result in incorrect $s(t)$.

4.3 Conclusion: Counterdiabatic QAOA

This section has outlined a procedure to map back and forth between continuous-time counterdiabatic evolution, and discrete CD-QAOA angles. Given some counterdiabatic Hamiltonian H_{CD} , counterdiabatic protocol $\lambda(t)$, $s(t)$, and number of QAOA steps p , one can variationally minimize the trace norm of the difference between continuous time counterdiabatic evolution and effective QAOA evolution to find fixed angles $\{\gamma, \beta\}$, as well as analogous total annealing time T

$$(H_{CD}(\lambda, \dot{\lambda}, s), \lambda(t), s(t), p) \mapsto (\{\gamma, \beta\}, T).$$

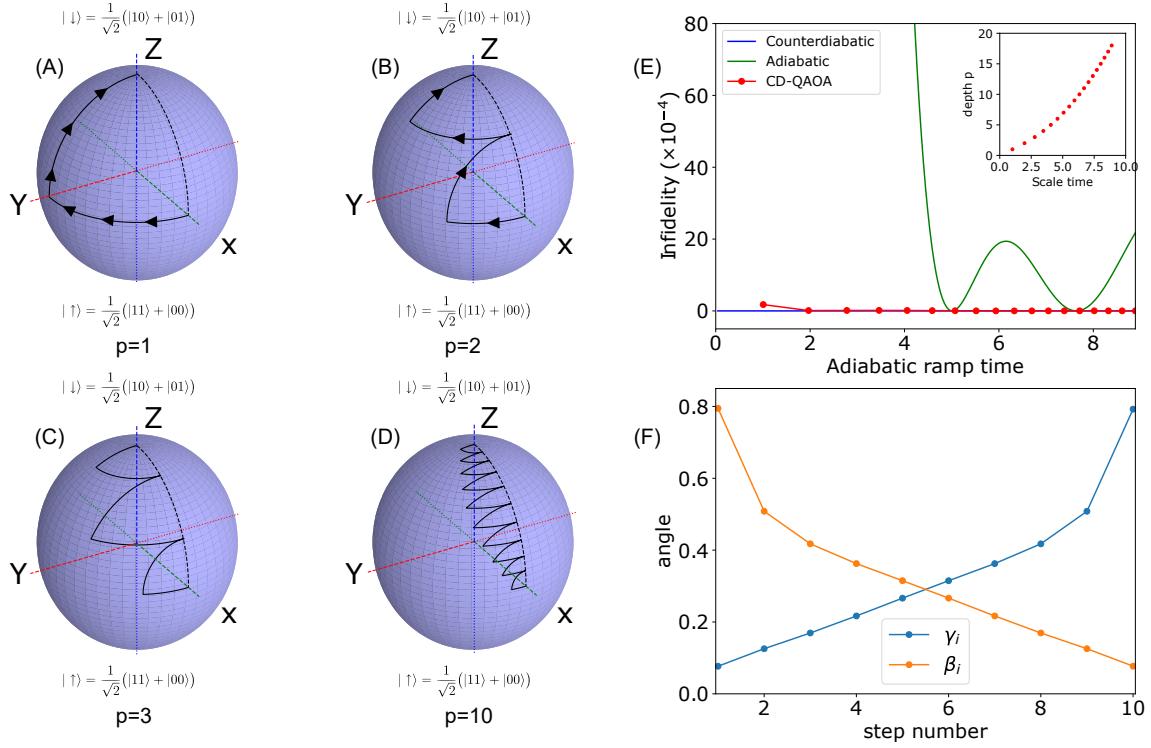


Figure 3: Counterdiabatic QAOA evolution on the Bloch sphere of the two level system. **(A-D)** CD-QAOA evolution along each step as a black line evolving continuously along each QAOA unitary. **(A)** A $p = 1$ QAOA may be exact by realizing a rotation of $\pi/2$ around the “Z” axis, then a rotation of $\pi/2$ around the “X” axis. **(B-D)** Higher order CD-QAOA derived from Eq. (27) can mimic the counterdiabatic evolution (black dashed) evolving along the “XZ” plane by matching the time evolution unitary with the unitary of one step of CD-QAOA using the trace minimization procedure of Eq. (27). Such a matching is exact up to perturbative error from the finite order of forth-order BCH and third-order Magnus expansions (Red, **(E)**) and outperforms an equivalent adiabatic-only evolution (Green, **(E)**). The large- p behavior scales as $T \sim \sqrt{p}$, as expected (inset, **(E)**). **(F)** The $p = 10$ CD-QAOA angles for a linear schedule $\lambda(t) = t/T$.

Such a mapping always exists for large finite p and T because the sign of the AGP term $\alpha(\lambda)$ is the same as the sign of the second order BCH term. In reverse, given smooth CD-QAOA angles $\{\gamma, \beta\}$, one can variationally minimize the smooth protocol λ, s and total annealing time T to best match the effective QAOA evolution to the continuous time counterdiabatic evolution

$$(H_{CD}(\lambda, \dot{\lambda}, s), \{\gamma, \beta\}) \mapsto (\lambda(t), s(t), T).$$

The next sections will outline three examples of increasing complexity demonstrating this CD-QAOA map. Section 5 will start with an example where the first order AGP is exact, and so $p = 1$ CD-QAOA is also exact. Section 6 will relax the exactness of the AGP with an Ising chain, to demonstrate the varying performance from choosing different counterdiabatic protocols $\lambda(t), s(t)$. Finally, Section 7 will demonstrate the reverse procedure and derive optimal continuous time counterdiabatic protocols from variationally optimized smooth QAOA angles.

5 Example 1: 2-level system

Let us begin with the simplest problem: the two level system. We will show that for any $p \geq 1$, we can derive perturbatively exact CD-QAOA angles which reproduce a counterdiabatic evolution. For the two level system, the first order commutator term is the exact adiabatic gauge potential [40]. Due to this fact, a counterdiabatic Hamiltonian of the form of Eq. (9) can generate exact eigenstates for any protocol speed. In the spirit of choosing Ising-like Hamiltonians for all examples, we choose the two site Ising Hamiltonian

$$H_{\text{CD}}(\lambda, \dot{\lambda}) = -\frac{\lambda}{2}\sigma_z^0\sigma_z^1 + (1-\lambda)(\sigma_x^0 + \sigma_x^1) + \frac{i\dot{\lambda}\alpha}{2}[\sigma_z^0\sigma_z^1, \sigma_x^0 + \sigma_x^1]. \quad (30)$$

$$H_S = \sigma_x^0 + \sigma_x^1 \quad ; \quad H_T = \frac{-1}{2}\sigma_z^0\sigma_z^1. \quad (31)$$

The Hamiltonian has a \mathbb{Z}_2 symmetry which separates the rank 4 Hilbert space into 2×2 level Hilbert spaces. The maximal state of H_S is the uniform superposition state $|+\rangle$ and is in the $+Z_2$ sector spanned by the basis vectors $\{ (|00\rangle + |11\rangle)/\sqrt{2}, (|01\rangle + |10\rangle)/\sqrt{2} \}$. By symmetry, the two site Ising Hamiltonian with this initial state is a two level system. The variational parameter α can be computed with Eq. (8) to find

$$\alpha(\lambda) = \frac{-1}{16(1-\lambda)^2 + \lambda^2}. \quad (32)$$

Evolving H_{CD} for any speed results in an exact final eigenstate, as the counterdiabatic term is exact. Similarly, $p = 1$ QAOA can generate the exact eigenstate of $\sigma_z^0\sigma_z^1$. The second order BCH term is the adiabatic gauge potential, and higher order terms must be constructed only of terms in Eq. (30), as the operator space is spanned by three operators. This fact can be seen visually by considering the evolution of the two level system on the Bloch sphere in Fig. 3A-D. The initial wavefunction starts in the “X” direction as an X eigenstate. First, the state is rotated in the “XY” plane by the $\sigma_z^0\sigma_z^1$ generator by $\pi/2$, eg. $\gamma = \pi/4$, then rotated in the “YZ” plane by the σ_x generator by $\pi/2$, eg. $\beta = \pi/4$. The final wavefunction is an exact eigenstate of the $\sigma_z^0\sigma_z^1$ operator, and thus $p = 1$ QAOA is exact.

We may verify this insight by applying the algorithm of Eq. (27) to the two level system for a linear schedule $\lambda(t) = t/T$ and $p = 1$ to derive the two CD-QAOA angles. To second order, the Magnus term is

$$-i \log \left(\mathcal{T} \exp \left(i \int_0^T dt H_{\text{CD}}(t) \right) \right) \approx \left(\frac{\pi T}{22} + \frac{T}{2} \right) \frac{\sigma_x^0 + \sigma_x^1}{2} + \left(\frac{\pi T}{22} + \frac{T}{2} \right) \sigma_z^0\sigma_z^1 + \left(\frac{\pi}{4} - \frac{T^2}{6} \right) \frac{\sigma_y^0\sigma_z^1 + \sigma_z^0\sigma_y^1}{2}. \quad (33)$$

To third order, the BCH term is

$$-i \log \left(\exp(i\gamma\sigma_z^0\sigma_z^1) \exp(i\gamma(\sigma_x^0 + \sigma_x^1)/2) \right) \approx \left(\beta - \frac{\gamma^2\beta}{3} \right) \frac{\sigma_x^0 + \sigma_x^1}{2} + \left(\gamma + \frac{\gamma\beta^2}{3} \right) \sigma_z^0\sigma_z^1 + \left(\gamma\beta \right) \frac{\sigma_y^0\sigma_z^1 + \sigma_z^0\sigma_y^1}{2}. \quad (34)$$

The difference between the Magnus generator and BCH generator may be made exactly zero by minimizing $||\Omega(T) - Z(\gamma, \beta)||$, yielding

$$T = 0.9974 \quad ; \quad \gamma = 0.2506\pi \quad ; \quad \beta = 0.2506\pi. \quad (35)$$

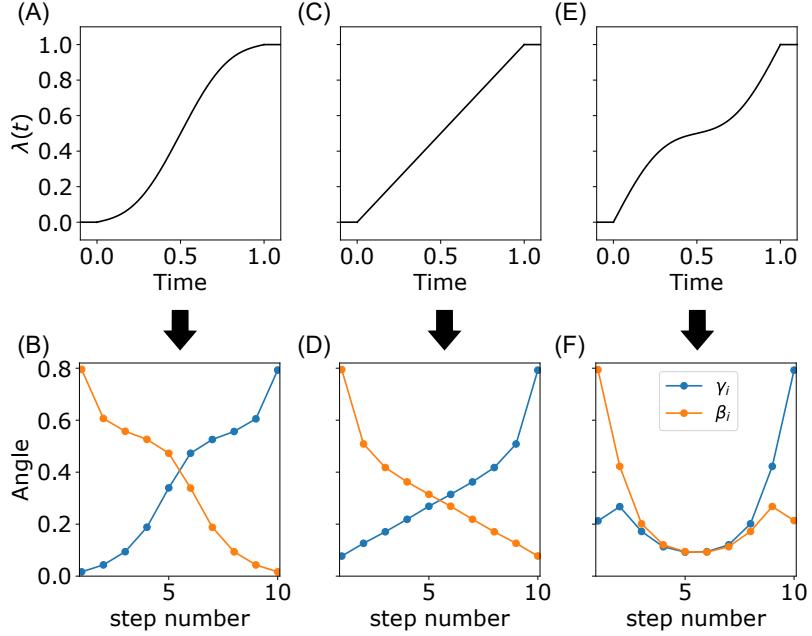


Figure 4: Induced CD-QAOA angles for a two level system and $p = 10$. Depending on the adiabatic schedule $\lambda(t)$ (Top), the induced CD-QAOA angles (bottom) also change. For the two level system, these protocols are all exact up to perturbative corrections. For situations where the adiabatic gauge potential is not exact, different protocols may result in different counterdiabatic performance, and so optimization of smooth QAOA angles is analogous to optimizing an adiabatic schedule $\lambda(t)$.

Including higher orders of BCH and Magnus expansion converge the values γ, β to $\pi/4$. These values show that the perturbative matching of CD-QAOA to a counterdiabatic time evolution can find CD-QAOA angles which are very close to the exact angles, even for small p .

An exact higher order $p > 1$ set of CD-QAOA angles which mimic a counterdiabatic schedule can also be generated. The exact eigenstate rotates in the “XZ” plane of the Bloch sphere (black dashed of Fig. 3), while a higher p CD-QAOA would, through repeated rotations around the “X” and “Z” axes, set the wavefunction at the end of each step as the exact instantaneous eigenstate along the “XZ” plane (Fig. 3A-D). Such protocols are not unique, as there is choice of where each step puts the wavefunction in the “XZ” plane. These non-unique protocols are analogously the choice of adiabatic protocol $\lambda(t)$: different protocols will generate different CD-QAOA angles. Some example adiabatic protocols $\lambda(t)$ and induced CD-QAOA angles for the two level system are shown in Fig. 4. Using a fourth order BCH expansion and third order Magnus expansion to compute the generators Z and Ω , Equation (27) is minimized for various $\lambda(t)$. It is important to reemphasize that such a minimization does not require any quantum simulation, only traces of sums of Pauli operators and computation of Magnus and BCH expansions to constant order.

6 Example 2: Ising chain

A more involved example is the transverse Ising chain, which is a 1d chain of qubits interacting with a nearest neighbor Ising term ($\sigma_z \sigma_z$) and on-site transverse field (σ_x). We will show that finite p CD-QAOA angles naturally converge to an equivalent finite time counterdiabatic annealing. The ZZ term of the Ising chain is a simple example of a MaxCut objective function on the ring of disagrees if the boundary conditions are periodic.

$$H_{CD}(\lambda, \dot{\lambda}, s) = \frac{-\lambda}{2} \sum_i \sigma_z^i \sigma_z^{i+1} + (1 - \lambda) \sum_i \sigma_x^i - 2(s + \dot{\lambda}\alpha) \sum_i \sigma_z^i \sigma_y^{i+1} + \sigma_y^i \sigma_z^{i+1}. \quad (36)$$

$$H_S = \sum_i \sigma_x^i \quad ; \quad H_T = \frac{-1}{2} \sum_i \sigma_z^i \sigma_z^{i+1}. \quad (37)$$

The Hamiltonian includes an additional term $\sigma_z \sigma_y$ which is the commutator of the Ising and transverse terms, and acts as the lowest order approximation to the adiabatic gauge potential. The transverse Ising model is exactly solvable (free integrable) using a Jordan Wigner mapping to free fermions [41], and so the eigensystem and nonequilibrium dynamics can be efficiently computed for any system size. There is a critical point at $\lambda = 1/2$, $s = 0$, with a phase transition between an ordered and disordered phase [28], which generates diabatic excitations for any finite speed adiabatic protocol. The system is free integrable, so the adiabatic gauge potential is simple to construct. For $s = 0$, the exact AGP is

$$A(\lambda) = \alpha_0(\lambda) \left(\sum_i \sigma_y^i \sigma_z^{i+1} + \sigma_z^i \sigma_y^{i+1} \right) + \alpha_1(\lambda) \left(\sum_i \sigma_y^i \sigma_x^{i+1} \sigma_z^{i+2} + \sigma_z^i \sigma_x^{i+1} \sigma_y^{i+2} \right) + \alpha_2(\lambda) \left(\sum_i \sigma_y^i \sigma_x^{i+1} \sigma_x^{i+2} \sigma_z^{i+3} + \sigma_z^i \sigma_x^{i+1} \sigma_x^{i+2} \sigma_y^{i+3} \right) + \dots \quad (38)$$

where the ellipsis denote larger Pauli string terms $YXX \cdots XXZ$, for some functions $\alpha(\lambda)$ [33]. For $\lambda = 0$ or 1 , $\alpha_0 = -1/4$ and $\alpha_{i>0} = 0$, meaning that the first order counterdiabatic term of H_{CD} is the exact adiabatic gauge potential only at the beginning and end of the annealing process, and a counterdiabatic evolution will not be exact, unlike the two level system of example 1. Different choices of protocol $\lambda(t)$ will generate different final counterdiabatic annealing wavefunctions $|\psi\rangle$ with different amounts of diabatic excitations, and there is some optimal smooth protocol which will minimize this error. The coefficient of the approximate adiabatic gauge potential of the Hamiltonian of Eq. (36) for $s = 0$ may be computed using Eq. (8) as

$$\alpha(\lambda) = \frac{-(1 - \lambda)^2 - \lambda^2}{8((1 - \lambda)^2 + \lambda^2)^2 + 8\lambda^2(1 - \lambda)^2}. \quad (39)$$

Given some protocol $\lambda(t)$ and $s(t)$ and depth p , one may induce the angle parameters γ and β

$$(H_{CD}(\lambda, \dot{\lambda}, s), \lambda(t), s(t), p) \mapsto (\{\gamma, \beta\}, T)$$

using the trace minimization procedure of Eq. (27). For instance, an adiabatic schedule $\lambda(t) = t/T$ and $s(t) = 0.05 \sin(\pi t/T)$ (Fig. 5A) and $p = 20$ derives CD-QAOA angles of Fig. 5B with an equivalent continuous annealing timescale of $T = 4.588$. We use a 4th order BCH expansion and 3rd order Magnus expansion to compute the effective generators along each step.

Because the CD-QAOA protocols are directly mapped from continuous adiabatic schedules, it makes sense to compare the CD-QAOA approximation ratio to the continuous time performance via the energy density. Such a comparison is shown in Fig. 5D-F for a range of p and values $s(t) = s_0 \sin(\pi t/T)$. In the perturbative limit of $p \rightarrow \infty$ and s small, the

angles γ and β are also small and so the unitary matching can be made close to exact. As can be seen (eg. for Fig. 5E for $s = 0$) as p is large the wavefunction of the QAOA protocol matches that of the analogous continuous time counterdiabatic protocol, indicated by the matching approximation ratio. As s becomes large, higher-order BCH terms accumulate error, diverging the QAOA protocol from its analogous continuous time counterdiabatic protocol.

For $p = 1$, it is notable that the approximation ratio is very close to the optimal value of $C_1 = 0.7500$. For instance, the $s = 0$ $p = 1$ induced angles have an approximation ratio of $C_1 = 0.7368$. Without any query to a quantum simulator, this method is able to compute very close to optimal angles.

As p gets large, the CD-QAOA approaches the continuous counterdiabatic value. It is a natural question to ask how the approximation ratio converges to one as a function of p . From Sec. 4, for $s = 0$ the time of an adiabatic evolution scales as $p \sim \alpha T^2$, where the prefactor is proportional to the average AGP coefficient. For $s \neq 0$, the time of an adiabatic evolution scales as $p \sim sT$, where the prefactor is proportional to the average value of s . For continuous adiabatic schedules, the scaling of energy density given a finite-speed passage of a second order phase transition is known from the Kibble-Zurek mechanism [42]. Any smooth finite-time protocol will cross a symmetry breaking phase transition with some velocity $\dot{\lambda}_c \sim 1/T$. Around the critical point, the process becomes nonadiabatic and excitations are “frozen out”. The density of these excitations, or “defects”, is related to the correlation length at that freeze-out point. The freeze out time is (roughly) related to the speed of quasiparticle propagation; when the correlation length grows faster than the propagation speed of the quasiparticles, the system is frozen. These excitations, or “defects”, persist through the rest of the adiabatic process on the other side of the critical point. The density of defects means that the final state has a finite energy density. Inverting ground states to maximal states of the objective function, the energy density of defects becomes the approximation ratio as

$$C(T) \sim 1 - T^{-d\nu/(1+z\nu)}, \quad (40)$$

where ν is the critical exponent, and z is the dynamic exponent. The critical and dynamic exponents for the 1d Transverse Ising model are well known [43] to be $\nu = 1$, $z = 1$. Given $T \sim p$ for $s \neq 0$, the large p scaling for a CD-QAOA protocol for the ring of disagrees is

$$C(p) \sim 1 - p^{-1/2}. \quad (41)$$

The prefactor will depend on the fine details of the protocol, including average value of s and the phase transition crossing speed. If the protocol moves through the critical point in some optimal fashion as a power law [13] as $|\lambda - \lambda_c| \sim |t/T|^r$, the approximation ratio instead goes as

$$C(T) \sim 1 - (\log(T)/T)^{d/z} \Rightarrow C(p) \sim 1 - p^{-1} \log(p) \quad (42)$$

The extra logarithmic dependence of the approximation ratio is different from the numerically observed optimal values of $C(p) = (2p+1)/(2p+2)$, which has a scaling exponent of 1 in the large p limit [1] for the ring of disagrees. Any smooth set of QAOA angles may be induced from some continuous-time adiabatic protocol, and so they must obey the counterdiabatic QAOA scaling of $p^{-1/2}$ or $\log(p)/p$. By exclusion, optimal protocols, which scale as p^{-1} , must be non-smooth. Along each step of such a non-smooth protocol, the intermediate wavefunction would be far from any semblance of an instantaneous

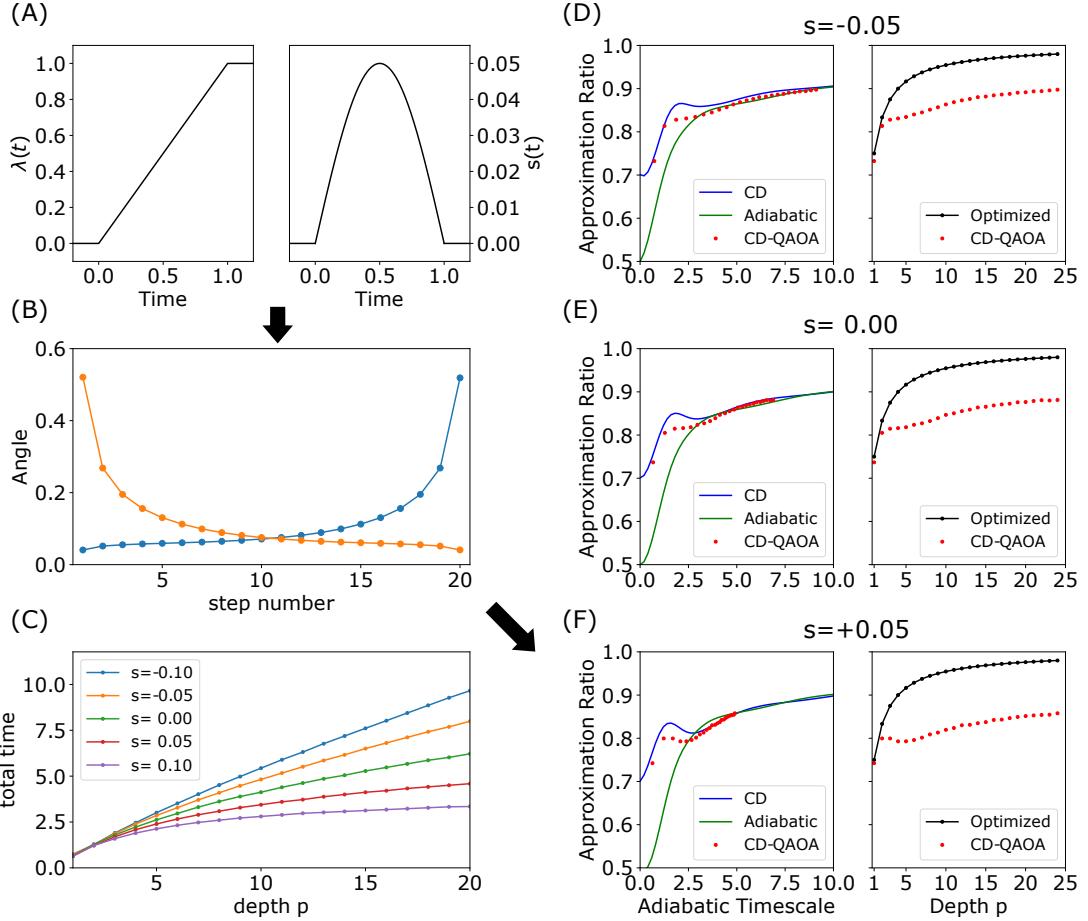


Figure 5: Counterdiabatic QAOA protocols for the 1d transverse Ising chain. Given a continuous time adiabatic schedule $\lambda(t)$, $s(t)$ (A) and QAOA depth p , analogous CD-QAOA angles are derived (B). (D-F) plots the approximation ratio of the derived CD-QAOA angles (red) for various p and strengths of drive s . (D-F) left plots CD-QAOA performance at their analogous adiabatic timescale (red) vs. continuous time counterdiabatic (blue) and adiabatic (green) protocols. In the perturbative limit of p large and s small, the CD-QAOA angles closely follows its analogous counterdiabatic protocol. (D-F) right plots CD-QAOA performance as a function of p . Black plots the conjectured maximum of $(2p+1)/(2p+2)$ [1], which approach unity as $1/2p$, while CD protocols approach unity as $1/\sqrt{p}$ for $s \neq 0$, suggesting optimal protocols for the Ising chain are non-adiabatic. (C) plots the matched time for a given value of s . For $s = 0$, the time goes as \sqrt{p} , while for $s < 0$, time is linear in p , in accordance with Eq. (23). For $s > 0$, there is some maximum total time which cannot be matched for any p .

ground state, and only recovers as an approximate eigenstate at the beginning and end of the QAOA protocol.

The difference between counterdiabatic and optimal QAOA angles for the chain also raises a point about optimizing these angles. For a smooth protocol, one can warm start the optimizer [44] with some adiabatic-like values or derived angles from Fig. 5. However, for a non-smooth optimal protocol, it is not immediately clear how an optimizer can find optimal angles, which may be hidden within the high-dimensional parameter space. In general, it may be easy to find smooth protocols with adequate QAOA performance derived from adiabatic limits, but it may be hard to find globally optimal protocols with the best performance for fixed p [45].

Curiously, a $\mu = 1$ scaling with no logarithmic correction that is an exactly optimal smooth protocol for the Ising chain has been found in [46]. These findings suggest that optimal QAOA protocols derived from optimal passage through critical points may have more subtlety than expected. For example, the counterdiabatic term may suppress just enough excitations to remove any logarithmic correction. Alternatively, the effective auxiliary term s ultimately coherently suppresses excitations of low momentum modes due to the strict locality of QAOA. The presence of such terms may change the low energy behavior of the original Hamiltonian enough to remove the logarithmic correction. Needless to say, a study of such behavior is beyond the scope of this work.

It may be interesting to generalize the exact protocol for the two level system and $p = 1$ of Section 5 to the more general Ising chain and larger p . For instance, a 2 step BCH expansion for $p = 2$ can generate terms which are the next-largest terms in the exact AGP of Eq. (38), which may enable one to compute exact $p = 2$ QAOA angles for a 3 site ring of disagrees. Similarly, going to larger p and using operator Krylov expansions [47] or high order Suzuki-Trotter decomposition [38] to compute the effective QAOA generator, one may be able to compute optimal protocols by matching the generator to the exact counterdiabatic Hamiltonian. In this way, one might be able to analytically derive exact QAOA angles which saturate the conjectured bound of $C(p) = (2p+1)/(2p+2)$ [1]. Such an investigation is beyond the scope of this work.

7 Example 3: regular MaxCut graphs

As a final example, let us consider QAOA solving MaxCut. We will demonstrate the two-way mapping between counterdiabatic protocols and CD-QAOA angles, and use it to show that transfer of parameters between graphs [16] and interpolating for larger p [15] are natural consequences. Let us consider a more general Ising Hamiltonian on a graph \mathcal{G} of the form

$$H_{\text{CD}}(\lambda, \dot{\lambda}, s) = \lambda \left[\frac{1}{2} \sum_{\langle ij \rangle \in \mathcal{G}} (1 - \sigma_z^i \sigma_z^j) \right] + (1 - \lambda) \left[\sum_i \sigma_x \right] - 2(s + \dot{\lambda} \alpha(\lambda)) \left[\sum_{\langle ij \rangle \in \mathcal{G}} (\sigma_y^i \sigma_z^j + \sigma_z^i \sigma_y^j) \right], \quad (43)$$

$$H_S = \sum_i \sigma_x \quad ; \quad H_T = \frac{1}{2} \sum_{\langle ij \rangle \in \mathcal{G}} (1 - \sigma_z^i \sigma_z^j). \quad (44)$$

Maximal energy states of the target Hamiltonian encode MaxCut solutions for the graph \mathcal{G} , and the maximal state of the simple Hamiltonian is a uniform superposition state $|+\rangle$. The third term is the commutator of the first two terms, and serves as the lowest order approximate adiabatic gauge potential, as well as auxiliary field s . The

Hamiltonian can include example 2 as a special case if the graph \mathcal{G} is an infinite chain of singly connected qubits. This example will focus on 3 regular graphs only.

Similar to the first two examples, given an counterdiabatic annealing schedule $\lambda(t)$ and $s(t)$, one can compute a set of CD-QAOA angles which mimics the continuous time evolution. For this example, we show the opposite approach. A set of smooth, optimized CD-QAOA angles can induce the continuous time counterdiabatic annealing schedule which best mimics the CD-QAOA angles, using the methods of Sec. 4.2. This procedure is shown in Fig. 6.

First, a classical optimizer maximizes the objective function (Fig. 6A) for a specific graph \mathcal{G} and p , and yield a set of angles $\{\gamma, \beta\}$ (Fig. 6B). If reasonably smooth, the angles can be used to induce the continuous time counterdiabatic protocol using Eq. (29) (Fig. 6C). It is important to subtract the counterdiabatic term. In the very large p limit, the total time is large and thus the counterdiabatic term is small and thus has negligible change for the form of $s(t)$. However, for intermediate p , different values of p would spuriously change the form of $s(t)$ if the AGP value was not excluded due to a small adiabatic timescale and thus large $\lambda\alpha$. For a ν -regular graph with no cycles of size 3, the variational parameter can be computed using Eq. (8) to be

$$\alpha(\lambda) = \frac{-32(1-\lambda)^2 - 8(3\nu-2)\lambda^2}{256((1-\lambda)^2 + 4(3\nu-2)\lambda^2)^2 + 256\lambda^2(1-\lambda)^2(\nu-1) + 96(\nu-1)(\nu-2)\lambda^4}. \quad (45)$$

The annealing schedule $\lambda(t), s(t)$ can then be induced using Eq. (29) using a polynomial interpolation (Fig. 6C). Then, given the fixed smooth protocol, the process can be done in reverse to derive new counterdiabatic QAOA angles (Fig. 6D) for a different value of p .

Optimizing a counterdiabatic QAOA schedule is equivalent to optimizing a smooth protocol (λ, s) . Different continuous time protocols have better or worse performance, and for some constraint of total time T and smooth ansatz $\lambda(t), s(t)$, there must be some smooth continuous time protocol that is optimal. Note that without the restriction of smoothness, by Pontryagin's minimum principle optimal protocols are bang-bang, eg. QAOA [17, 18]. In the large p limit, we know from section 4 that a counterdiabatic QAOA closely reproduces a continuous adiabatic schedule and, from Eq. (29), one can be mapped to the other. In this way, the classical optimizer is effectively searching for the best smooth adiabatic schedule [14] by optimizing angles which best mimic that schedule.

We comment that the best continuous schedule may differ slightly from that derived from CD-QAOA angles, due to a slight bias. For a continuous protocol, the total time T is fixed, while for a CD-QAOA protocol the number of steps p is fixed, and the time depends on the average value of s , as $T \sim sp$. By the adiabatic theorem, larger total evolution times have less excitations, and thus the optimizer is biased to a larger average value of s .

We show an example of this implementation in Fig. 6. Here, variational parameters for a $p = 12$ QAOA are initialized in an adiabatic inspired manner for a small 3 regular graph with 14 vertices (Fig. 6A). These initial values are optimized for the particular graph using a numerical gradient ascent procedure, which returns only after settling into a local minimum where the gradient is zero. The resultant optimal angles (Fig. 6B) are smooth and mimic a continuous time counterdiabatic schedule, as expected. The approximation ratio is $C_{12} = 0.99004$, very close to the exact result of unity. Using the procedure of Eq. (29) and a polynomial interpolation, the analogous continuous counterdiabatic protocol is derived, with an equivalent adiabatic time of $T \approx 9.3$ (Fig. 6C). With this continuous protocol, an equivalent CD-QAOA protocol may be derived for any value of

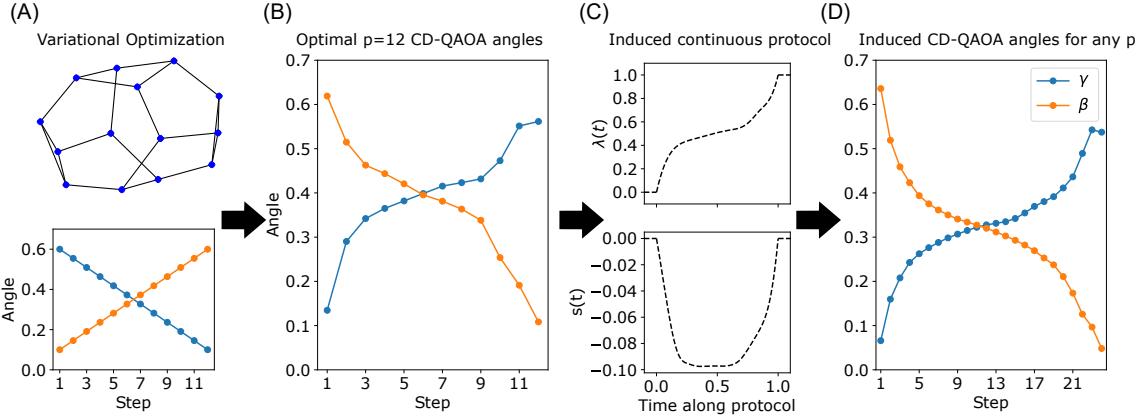


Figure 6: Inducing continuous adiabatic protocols and CD-QAOA angles from variational optimization. Given some small graph instance and adiabatic-inspired initial point **(A)**, a classical optimizer maximizes the objective function and yields an optimal set of smooth CD-QAOA angles **(B)**. Given these angles and a variationally derived prefactor for the adiabatic gauge potential, a continuous protocol for $\lambda(t)$ and $s(t)$ can be induced **(C)**. With the continuous protocol, a set of CD-QAOA angles for any p is induced **(D)**.

p using the variational minimization procedure of Eq. (27). For example, this protocol derives $p = 24$ angles (Fig. 6D) with an approximation ratio $C_{24} = 0.99994$ and an equivalent adiabatic time of $T \approx 15.4$.

The procedure of inducing larger p angles from some fixed p optimization formalizes the original observations of [15], which interpolate smooth QAOA angles for a certain p to induce different angles p' with good performance. These observations can now be justified under these counterdiabatic procedures: both p and p' share the same underlying approximately optimal counterdiabatic protocol. However, this procedure also allows for computing CD-QAOA angles for small $p \sim 1$, where the effective annealing timescale is small and the contribution from the counterdiabatic term $\dot{\lambda}\alpha(t)$ may be large. Near the edges of an evolution where the rate of change is large, the counterdiabatic term $\dot{\lambda}\alpha(\lambda)$ may be large, even for p large, a phenomenon that a simple interpolation may fail to capture.

It has been numerically observed that optimal QAOA angles of one graph can be transferred to other graphs of the same class [16]. It is reasonable to expect that the optimal continuous protocol derived from CD-QAOA angles of one graph should be transferable. If two graphs share the same low energy eigenspectrum, an optimal counterdiabatic protocol for one graph should be similar to an optimal protocol $\lambda(t)$, $s(t)$ for the other.

By extension, it seems reasonable to expect that optimized CD-QAOA angles for one graph should have a similar performance to the other. The counterdiabatic protocol induced from the optimized for the first graph should be close to optimal for the other graph. Similarly, the CD-QAOA angles induced from that protocol for the new graph should also be similar to those for the old graph. In lieu of this procedure, which ultimately results in very similar CD-QAOA angles, one can instead directly transfer angles from one graph to another. This indicates that transfer of parameters, at least for smooth sets of angles, is a natural consequence of counterdiabaticity.

8 QAOA is better than continuous-time adiabatic protocols

A natural question to ask is how continuous time adiabatic annealing protocols perform in comparison to QAOA. Here we show that under specific restrictions QAOA is better, by comparing finite time adiabatic annealing with finite time counterdiabatic annealing via a chain of inequalities relating approximation ratios:

$$\begin{bmatrix} \text{Optimal QAOA} \\ \text{with fixed } T \\ (1) \end{bmatrix} \geq \begin{bmatrix} \text{Optimal CD-QAOA} \\ \text{with fixed } T \\ (2) \end{bmatrix} \geq \begin{bmatrix} \text{CD-QAOA} \\ \text{with fixed } T, p \\ (3) \end{bmatrix} \stackrel{p \text{ large}}{=} \begin{bmatrix} \text{Counterdiabatic} \\ \text{with fixed } T \\ (4) \end{bmatrix} \geq \begin{bmatrix} \text{Adiabatic} \\ \text{with fixed } T \\ (5) \end{bmatrix}.$$

Suppose for some optimization problem, the $2p$ optimal QAOA angles (1) with a total magnitude T , which we call an “angle budget”

$$T = \sum_{q=1}^p |\gamma_q| + |\beta_q|. \quad (46)$$

Now, suppose this angle budget T acts as an optimization constraint. While the constraint is not as useful for digital computing, it has a natural interpretation as the total adiabatic evolution time, which is a common constraint for QAA. By the nature of global optima under constraints, these optimal angles (which may be non-smooth) have a greater than or equal approximation ratio than any CD-QAOA angles (2) with fixed angle budget T and variable p . This includes optimal CD-QAOA angles which optimize the effective counterdiabatic schedule $\lambda(t)$ and some $p(T)$ to best match counterdiabatic terms. Because $\alpha(\lambda)$ has the same sign as the second order BCH term, $p(T)$ will always be finite.

The value p can always be made large and fixed (3), which again by nature of optima under constraints must have a less than or equal approximation ratio to the optimal CD-QAOA angles. Increasing p reduces the strength of the effective counterdiabatic term; If large enough, the higher order BCH error can be made small enough to mimic the effective Hamiltonian (3) by suppressing higher-order error terms

$$H_{\text{eff}} = \frac{\gamma(t)}{\gamma + \beta} H_T + \frac{\beta(t)}{\gamma + \beta} H_S - \frac{i\gamma\beta}{\gamma + \beta} [H_T, H_S] \quad (47)$$

with some (less than optimal) effective counterdiabatic term $-\gamma_q\beta_q/(\gamma_q + \beta_q)$. Finally, due to the sign of the effective CD term (See Sec. 3), the finite time counterdiabatic annealing (4) has a larger approximation ratio than an evolution of the adiabatic-only annealing (5). By transitivity, any optimal QAOA with angles $\{\gamma, \beta\}$ has a better approximation ratio than an adiabatic-only finite time annealing evolution with total time $T = \sum_q |\gamma_q| + |\beta_q|$. In this sense, QAOA can always outperform an adiabatic evolution by at least reproducing a counterdiabatic evolution.

This observation is an instance of Pontryagin’s minimum principle [17, 18]: given some fixed total time for adiabatic evolution, optimal protocols are bang-bang, as long as the controls are nonsingular [25]. However, the critical intermediate step of CD-QAOA protocols requires observation of the *sign* of the counterdiabatic term. For the lowest order approximation of the adiabatic gauge potential, $\alpha(\lambda) \leq 0$ always by Eq. (8). Similarly, the second order BCH term is also < 0 always. The matching of the effective CD-QAOA Hamiltonian to the counterdiabatic Hamiltonian with finite p and fixed T is only possible with this choice of signs: for some small but negative α , the effective Hamiltonian of eq. (47) will have better adiabatic performance on the ground or maximal state of the

model. If $\alpha > 0$ the variational matching procedure would choose $p \rightarrow \infty$ to minimize the weight of the commutator term to be optimal. Due to this, the last inequalities (3) \geq (5) hold and QAOA does better than adiabatic, as QAOA at least reproduces a counterdiabatic evolution instead of an adiabatic one.

Although we show that smooth CD-QAOA can have good finite- p performance, it is not necessarily clear that smooth angles are always optimal. By Pontryagin’s minimum principle, optimal protocols are QAOA-like, but there is no necessary requirement on smoothness. An optimal protocol for a given p may be highly non-adiabatic, with the intermediate state being far from an instantaneous ground state and angles being highly nontrivial. It may be an interesting future study to find when optimal QAOA angles are smooth, or non-smooth, which may motivate an optimizer which biases its search to “non-smooth” regions of parameter space.

9 Conclusion

In this work, we have made the connection between QAA and QAOA explicit by inspecting the regime of p large but finite. The results are centered around one observation: the second order BCH term, which determines the effective QAOA evolution generator, is the same as a low order approximation of the adiabatic gauge potential, which suppresses diabatic excitations. By identifying and matching counterdiabatic annealing with QAOA evolution, we analytically construct sets of CD-QAOA angles which best mimic the continuous time evolution. Importantly, this process is completely tractable for any system size. Computation of coefficients requires only well-behaved variational minimization of traces of Pauli operators, which can be done efficiently on a classical computer without needing quantum simulations. While these results are perturbatively exact in the large- p limit, the CD-QAOA angle derivation is well behaved and has good performance even for small to intermediate $p \lesssim 10$. This suggests that these analytically derived angles can be used directly [19] bypassing the optimization step, or as variational warm starts on tomorrow’s NISQ devices.

While this method was applied to the MaxCut problem, these counterdiabatic insights may be applied to a wide range of objective Hamiltonians. For example, these methods may be applied to the variational quantum eigensolver to find low energy states of quantum chemistry Hamiltonians, which have off-diagonal matrix elements and are thus more general than binary optimization problems. Given some general setting of simple and target Hamiltonians H_S and H_T , a QAOA-like protocol can generally be derived using these methods to prepare low energy states in a counterdiabatic manner. Such a study is beyond the scope of this paper, although these results should transfer naturally to off-diagonal Hamiltonians.

Critical to the matching procedure is the fact that the sign of the variational adiabatic gauge potential coefficient α is the same sign as the second order BCH coefficient. This implies that for large T , there is some large but finite $p(T)$ which is optimal given that counterdiabatic evolution has better performance than adiabatic-only annealing. The existence of such finite- p protocols also implies that, in some sense, QAOA does better than finite time adiabatic-only annealing. Given some time T or equivalently angle budget $T = \sum_q |\gamma_q| + |\beta_q|$, the effective QAOA evolution can include terms which reduce diabatic excitations, or in accordance with Pontryagin’s minimum principle [17] take nonadiabatic paths with better use of the angle budget.

Matching of p with T , as well as improving performance over adiabatic behavior, also allows for insight of lower bounds on how the approximation ratio must converge with

p . Using Kibble Zurek scaling, we show that the approximation ratio converges to one at least polynomially $1 - C(p) \sim p^{-\mu}$, where exponent μ is derived from the critical exponents of the model. While this bound exists for smooth CD-QAOA angles, optimal angles may converge faster than the adiabatic limit. For instance, we find exponent $\mu = 1/2$ for the transverse Ising model, slower than the conjectured optimal behavior of $\mu = 1$. This suggests that optimal QAOA angles may be non-smooth and potentially hard to find.

Another important part of the construction is the addition of the auxiliary term s , which adds extra terms to the effective QAOA Hamiltonian. Instead of following the ground state of the simple Hamiltonian $\lambda H_S + (1 - \lambda) H_T$, the QAOA wavefunction may follow the ground state of a Hamiltonian with extra terms $H(t) + s(t) \tilde{H}(t)$. We numerically observe that near the edges of the optimized protocols of example 3, the protocol is “fast”, and thus counterdiabatic, whereas near the middle, the protocol is “slow” and thus adiabatic only, except with a large term s which allows for a slow effective annealing time $T \sim p$. The necessity of the auxiliary term s to match optimized QAOA angles with an effective QAOA Hamiltonian suggests that further study is required on the subject beyond the scope of this work.

Note that optimal QAOA angles may be degenerate (see for example Table 1 of [48]). For example, the approximation ratio is unchanged under $\{\gamma, \beta\} \rightarrow \{-\gamma, -\beta\}$ under time reversal symmetry. Additionally, a $U(2)$ symmetry allows addition of factors of π to angles. More non-trivial symmetries may exist, which relate points in parameter space under symmetry transformations. A set of smooth CD-QAOA angles could be “scrambled” under these symmetry transformations into a class of non-smooth QAOA angles that are nonetheless counterdiabatic. Alternatively, a non-smooth set of QAOA angles may be in fact non-counterdiabatic if they are not part of a class that cannot be made smooth under symmetry transformations. It may be interesting to distinguish between optimal angles which are counterdiabatic under transformations, and those which are not; however, this is beyond the scope of this work.

We also show that two heuristically observed behaviors of QAOA are natural consequences of CD-QAOA. It has been observed [16] that optimal parameters for one MaxCut graph “transfer”, or have similar performance, on other graphs with a similar low-energy eigenspectrum. We posit that this is a consequence of the fact that smooth CD-QAOA angles can derive continuous-time counterdiabatic annealing schedules $\lambda(t)$, and vice versa. Similarly, the interpolation procedure of [15], which induces parameters for level p given optimal parameters of $p - 1$, can be justified in the same way. We add the extra improvement of including the counterdiabatic term to adjust to different annealing times from different p .

In conclusion, this paper studies the subtle interplay between QAOA and the quantum adiabatic algorithm (QAA). Ultimately, QAOA is always at least *counterdiabatic*, not just *adiabatic*. It is an interesting future direction to further explore the extra advantage of QAOA and other digital algorithms over their analog counterparts. Here, we show that QAOA can at least beat quantum annealing, though give no bounds on improvement. Given an analogous continuous counterpart, it will be interesting to see where and how there is an explicit separation in performance given analogous quantum resources.

Acknowledgements

This work was supported by NSF STAQ project (PHY-1818914) and the Defense Advanced Research Projects Agency (DARPA) under Contract No. HR001120C0068.

A Appendix: BCH and Magnus Expansion terms

For completeness, we include the Baker-Campbell-Hausdorff and Magnus expansions here. The BCH expansion is a sum of terms $Z = Z_1 + Z_2 + Z_3 + Z_4 + \dots$ where

$$\begin{aligned} Z_1 &= \gamma H_T + \beta H_S & (48) \\ Z_2 &= -\frac{i\gamma\beta}{2} [H_T, H_S] \\ Z_3 &= \frac{\gamma\beta^2}{12} [H_S, [H_T, H_S]] - \frac{\gamma^2\beta}{12} [H_T, [H_T, H_S]] \\ Z_4 &= \frac{-i\gamma^2\beta^2}{24} [H_T, [H_S, [H_T, H_S]]] \\ Z_5 &= \dots \end{aligned}$$

Higher order terms can be found constructively, e.g. [38]. The Magnus expansion is sum of terms $\Omega = \Omega_1 + \Omega_2 + \Omega_3 + \dots$ where, for a smooth protocol around t_0 is $\lambda(t - t_0) \approx \lambda_0 + \dot{\lambda}t + \ddot{\lambda}t^2/2$,

$$\begin{aligned} \Omega_1 &= \int_{t_0}^{t_0+\tau} dt \lambda(t) H_T + (1 - \lambda(t)) H_S + i(s(t) + \dot{\lambda}(t)\alpha(\lambda(t))) [H_T, H_S] \\ &= \tau \bar{\lambda} H_T + \tau(1 - \bar{\lambda}) H_S + i\tau(\bar{s} + \dot{\lambda}\bar{\alpha}) [H_T, H_S], \end{aligned} \quad (49)$$

where $\bar{\lambda} = \frac{1}{\tau} \int_{t_0}^{t_0+\tau} dt \lambda(t)$ and $\bar{\alpha} = \int_{\lambda(t_0)}^{\lambda(t_0+\tau)} d\lambda \alpha(\lambda)$ are the average coefficients over the interval. For smooth $\alpha(\lambda)$ and adiabatic schedule $\lambda(t)$, $\bar{\alpha} = \dot{\lambda}\tau\bar{\alpha}$, with $\bar{\alpha} = \frac{1}{\tau} \int_{t_0}^{t_0+\tau} dt \alpha(t)$ the average coefficient over the interval. The second order term is

$$\Omega_2 = -\frac{1}{2} \int_{t_0}^{t_0+\tau} dt_1 \int_{t_0}^{t_1} dt_2 [H_{CD}(t_1), H_{CD}(t_2)] \quad (50)$$

$$= \frac{1}{2} \int_{t_0}^{t_0+\tau} dt_1 \int_{t_0}^{t_1} dt_2 \left(\lambda(t_1)(1 - \lambda(t_2)) - (1 - \lambda(t_1))\lambda(t_2) \right) [H_T, H_S] \quad (51)$$

$$+ \left(\lambda(t_1)\dot{\lambda}(t_2)\alpha(t_2) - \lambda(t_2)\dot{\lambda}(t_1)\alpha(t_1) \right) [H_T, [H_T, H_S]]$$

$$+ \left((1 - \lambda(t_1))\dot{\lambda}(t_2)\alpha(t_2) - (1 - \lambda(t_2))\dot{\lambda}(t_1)\alpha(t_1) \right) [H_S, [H_T, H_S]]$$

$$\approx \left(\frac{-\dot{\lambda}\tau^3}{12} + \frac{-\ddot{\lambda}\tau^4}{24} \right) [H_T, H_S] + \mathcal{O}(\dot{\lambda}^2\tau^3, \ddot{\lambda}\tau^3, \dot{\lambda}\ddot{\lambda}\tau^4, \dots) \quad (52)$$

where the higher order contributions are from the doubly nested commutators arising from the counterdiabatic term. The third order term Ω_3 will be constructed of at least doubly nested commutators such as $[H_T, [H_T, H_S]]$ at lowest order, and have prefactors such as $\dot{\lambda}\tau^4$, $\ddot{\lambda}\tau^5$, $\dot{\lambda}^2\tau^5$. The terms of third order and higher can be found in [39].

B Derivation of Eq. 32

In this section, we derive Eq. 32 in more detail. This has been done in several places elsewhere, e.g. ref. [33] but is repeated for completeness here. The AGP term is defined by Eq. 8

$$A(\lambda) = i\alpha(\lambda)[H_T, H_S] \quad \text{with} \quad \alpha(\lambda) = \frac{\|[H, \partial_\lambda H]\|_{\mathcal{P}}}{\|[[H, \partial_\lambda H], H]\|_{\mathcal{P}}}. \quad (53)$$

and the Hamiltonian terms for the 2 site Ising model are defined by Eq. 31

$$H_S = \sigma_x^0 + \sigma_x^1 \quad ; \quad H_T = \frac{-1}{2} \sigma_z^0 \sigma_z^1. \quad (54)$$

Plugging in the derivative of $\partial_\lambda H = -\sigma_z^0 \sigma_z^1/2 - \sigma_x^0 - \sigma_z^1$, the commutators are

$$[H, \partial_\lambda H] = \left[-\frac{\lambda}{2} \sigma_z^0 \sigma_z^1/2 + (1 - \lambda)(\sigma_x^0 + \sigma_x^1), -\sigma_z^0 \sigma_z^1 - \sigma_x^0 - \sigma_z^1 \right], \quad (55)$$

$$= (+i\lambda + i(1 - \lambda))(\sigma_y^0 \sigma_z^1 + \sigma_z^0 \sigma_y^1) = i(\sigma_y^0 \sigma_z^1 + \sigma_z^0 \sigma_y^1), \quad (56)$$

$$[[H, \partial_\lambda H], H] = i[\sigma_y^0 \sigma_z^1 + \sigma_z^0 \sigma_y^1, -\frac{\lambda}{2} \sigma_z^0 \sigma_z^1/2 + (1 - \lambda)(\sigma_x^0 + \sigma_x^1)], \quad (57)$$

$$= \lambda(\sigma_x^0 + \sigma_x^1) + 4(1 - \lambda)(\sigma_z^0 \sigma_z^1 - \sigma_y^0 \sigma_y^1). \quad (58)$$

The norm is defined as $\|A\|_{\mathcal{P}} = \text{Tr}[\mathcal{P} A^2]$. The projective subspace are the \mathbb{Z}_2 symmetric states $\{(|11\rangle + |00\rangle)/\sqrt{2}, (|01\rangle + |10\rangle)/\sqrt{2}\}$. The square of the terms are

$$([H, \partial_\lambda H])^2 = -2(\mathbb{1} + \sigma_x^0 \sigma_x^1), \quad (59)$$

$$([[H, \partial_\lambda H], H])^2 = 2(\mathbb{1} + \sigma_x^0 \sigma_z^1)(\lambda^2 + 16(1 - \lambda)^2). \quad (60)$$

Observe that the operators select the +1 sector of \mathbb{Z}_2 , as $\sigma_x^0 \sigma_x^1$ is the generator of \mathbb{Z}_2 symmetry, so the trace over the subspace and operators is 4 for both numerator and denominator. Eq. (32) can then be found by dividing the numerator by denominator above. A similar procedure can be done for the following examples, keeping care to factors of -1 , 2 , and i .

References

- [1] E. Farhi, J. Goldstone, and S. Gutmann, (2014), [arXiv:1411.4028 \[quant-ph\]](https://arxiv.org/abs/1411.4028) .
- [2] J. Preskill, [Quantum **2**, 79 \(2018\)](https://quantum-journal.org/papers/29/).
- [3] K. Bharti, A. Cervera-Lierta, T. H. Kyaw, T. Haug, S. Alperin-Lea, A. Anand, M. De-groote, H. Heimonen, J. S. Kottmann, T. Menke, W.-K. Mok, S. Sim, L.-C. Kwek, and A. Aspuru-Guzik, (2021), [arXiv:2101.08448 \[quant-ph\]](https://arxiv.org/abs/2101.08448) .
- [4] A. Lucas, [Frontiers in Physics **2** \(2014\), 10.3389/fphy.2014.00005](https://frontiersin.org/articles/10.3389/fphy.2014.00005).
- [5] T. Albash and D. A. Lidar, [Rev. Mod. Phys. **90**, 015002 \(2018\)](https://doi.org/10.1103/RevModPhys.90.015002).
- [6] M. H. S. Amin, [Phys. Rev. Lett. **102**, 220401 \(2009\)](https://doi.org/10.1103/PhysRevLett.102.220401).
- [7] S. H. Sack and M. Serbyn, (2021), [arXiv:2101.05742 \[quant-ph\]](https://arxiv.org/abs/2101.05742) .
- [8] D. Guéry-Odelin, A. Ruschhaupt, A. Kiely, E. Torrontegui, S. Martínez-Garaot, and J. G. Muga, [Rev. Mod. Phys. **91**, 045001 \(2019\)](https://doi.org/10.1103/RevModPhys.91.045001).
- [9] G. Rigolin, G. Ortiz, and V. H. Ponce, [Phys. Rev. A **78**, 052508 \(2008\)](https://doi.org/10.1103/PhysRevA.78.052508).
- [10] S. Bachmann, W. De Roeck, and M. Fraas, [Phys. Rev. Lett. **119**, 060201 \(2017\)](https://doi.org/10.1103/PhysRevLett.119.060201).
- [11] A. del Campo and W. H. Zurek, [International Journal of Modern Physics A **29**, 1430018 \(2014\)](https://doi.org/10.1142/S0217751X1430018).
- [12] J. I. Latorre and R. Orús, [Phys. Rev. A **69**, 062302 \(2004\)](https://doi.org/10.1103/PhysRevA.69.062302).

[13] R. Barankov and A. Polkovnikov, *Physical Review Letters* **101** (2008), 10.1103/physrevlett.101.076801.

[14] B. F. Schiffer, J. Tura, and J. I. Cirac, (2021), arXiv:2103.01226 [quant-ph] .

[15] L. Zhou, S.-T. Wang, S. Choi, H. Pichler, and M. D. Lukin, *Physical Review X* **10** (2020), 10.1103/physrevx.10.021067.

[16] F. G. S. L. Brandao, M. Broughton, E. Farhi, S. Gutmann, and H. Neven, (2018), arXiv:1812.04170 [quant-ph] .

[17] Z.-C. Yang, A. Rahmani, A. Shabani, H. Neven, and C. Chamon, *Physical Review X* **7** (2017), 10.1103/physrevx.7.021027.

[18] L. T. Brady, C. L. Baldwin, A. Bapat, Y. Kharkov, and A. V. Gorshkov, *Phys. Rev. Lett.* **126**, 070505 (2021).

[19] M. Streif and M. Leib, *Quantum Science and Technology* **5**, 034008 (2020).

[20] M. Born and V. Fock, *Zeitschrift für Physik* **51**, 165 (1928).

[21] E. Farhi, J. Goldstone, S. Gutmann, and M. Sipser, (2000), arXiv:quant-ph/0001106 [quant-ph] .

[22] A. T. Rezakhani, W.-J. Kuo, A. Hamma, D. A. Lidar, and P. Zanardi, *Phys. Rev. Lett.* **103**, 080502 (2009).

[23] C. Brif, M. D. Grace, M. Sarovar, and K. C. Young, *New Journal of Physics* **16**, 065013 (2014).

[24] L. T. Brady, L. Kocia, P. Bienias, A. Bapat, Y. Kharkov, and A. V. Gorshkov, (2021), arXiv:2107.01218 [quant-ph] .

[25] S. Bao, S. Kleer, R. Wang, and A. Rahmani, *Phys. Rev. A* **97**, 062343 (2018).

[26] T. Chasseur, L. S. Theis, Y. R. Sanders, D. J. Egger, and F. K. Wilhelm, *Phys. Rev. A* **91**, 043421 (2015).

[27] P. W. Claeys, M. Pandey, D. Sels, and A. Polkovnikov, *Phys. Rev. Lett.* **123**, 090602 (2019).

[28] S. Sachdev, *Quantum Phase Transitions*, 2nd ed. (Cambridge University Press, 2011).

[29] H. Nishimori and K. Takada, *Frontiers in ICT* **4**, 2 (2017).

[30] M. W. Johnson, M. H. S. Amin, S. Gildert, T. Lanting, F. Hamze, N. Dickson, R. Harris, A. J. Berkley, J. Johansson, P. Bunyk, E. M. Chapple, C. Enderud, J. P. Hilton, K. Karimi, E. Ladizinsky, N. Ladizinsky, T. Oh, I. Perminov, C. Rich, M. C. Thom, E. Tolkacheva, C. J. S. Truncik, S. Uchaikin, J. Wang, B. Wilson, and G. Rose, *Nature* **473**, 194 (2011).

[31] S. Sugiura, P. W. Claeys, A. Dymarsky, and A. Polkovnikov, *Phys. Rev. Research* **3**, 013102 (2021).

[32] M. V. Berry, *Journal of Physics A: Mathematical and Theoretical* **42**, 365303 (2009).

[33] M. Kolodrubetz, D. Sels, P. Mehta, and A. Polkovnikov, *Physics Reports* **697**, 1 (2017).

[34] M. Pandey, P. W. Claeys, D. K. Campbell, A. Polkovnikov, and D. Sels, *Phys. Rev. X* **10**, 041017 (2020).

[35] D. Sels and A. Polkovnikov, *Proceedings of the National Academy of Sciences* **114**, E3909 (2017).

[36] M. B. Hastings and X.-G. Wen, *Phys. Rev. B* **72**, 045141 (2005).

[37] J. Wurtz and A. Polkovnikov, *Phys. Rev. B* **101**, 195138 (2020).

[38] N. Hatano and M. Suzuki, *Lecture Notes in Physics* , 37–68 (2005).

[39] S. Blanes, F. Casas, J. Oteo, and J. Ros, *Physics Reports* **470**, 151–238 (2009).

[40] X. Chen, I. Lizuain, A. Ruschhaupt, D. Guéry-Odelin, and J. G. Muga, *Phys. Rev. Lett.* **105**, 123003 (2010).

[41] J. Dziarmaga, *Phys. Rev. Lett.* **95**, 245701 (2005).

- [42] M. Kolodrubetz, B. K. Clark, and D. A. Huse, *Physical Review Letters* **109** (2012), [10.1103/physrevlett.109.015701](https://doi.org/10.1103/physrevlett.109.015701).
- [43] A. Dutta, G. Aeppli, B. K. Chakrabarti, U. Divakaran, T. F. Rosenbaum, and D. Sen, (2015), [arXiv:1012.0653 \[cond-mat.stat-mech\]](https://arxiv.org/abs/1012.0653) .
- [44] D. J. Egger, J. Mareček, and S. Woerner, *Quantum* **5**, 479 (2021).
- [45] A. G. R. Day, M. Bukov, P. Weinberg, P. Mehta, and D. Sels, *Phys. Rev. Lett.* **122**, 020601 (2019).
- [46] G. B. Mbeng, R. Fazio, and G. Santoro, (2019), [arXiv:1906.08948 \[quant-ph\]](https://arxiv.org/abs/1906.08948) .
- [47] V. Viswanath and G. Müller, *The Recursion Method* (Springer US, 2008).
- [48] J. Wurtz and P. Love, *Phys. Rev. A* **103**, 042612 (2021).