Counterdiabatic Optimized Local Driving

Ieva Čepaitė,[1,*] Anatoli Polkovnikov,[2] Andrew J. Daley,[1] and Callum W. Duncan[1]

[1] *Department of Physics, SUPA and University of Strathclyde, Glasgow G4 0NG, United Kingdom*

[2] *Department of Physics, Boston University, Boston, Massachusetts 02215, USA*

(Received 10 March 2022; revised 20 October 2022; accepted 1 December 2022; published 30 January 2023)

Adiabatic protocols are employed across a variety of quantum technologies, from implementing state preparation and individual operations that are building blocks of larger devices, to higher-level protocols in quantum annealing and adiabatic quantum computation. The problem of speeding up these processes has garnered a large amount of interest, resulting in a menagerie of approaches, most notably quantum optimal control and shortcuts to adiabaticity. The two approaches are complementary: optimal control manipulates control fields to steer the dynamics in the minimum allowed time, while shortcuts to adiabaticity aims to retain the adiabatic condition upon speed-up. We outline a new method that combines the two methodologies and takes advantage of the strengths of each. The new technique improves upon approximate local counterdiabatic driving with the addition of time-dependent control fields. We refer to this new method as counterdiabatic optimized local driving (COLD) and we show that it can result in a substantial improvement when applied to annealing protocols, state preparation schemes, entanglement generation, and population transfer on a lattice. We also demonstrate a new approach to the optimization of control fields that does not require access to the wave function or the computation of system dynamics. COLD can be enhanced with existing advanced optimal control methods and we explore this using the chopped randomized basis method and gradient ascent pulse engineering.

DOI: 10.1103/PRXQuantum.4.010312

I. INTRODUCTION

Time-dependent manipulation of few and many-particle quantum systems is important across all implementations of quantum computing and simulation. In such processes, decoherence and undesired transitions reducing the state fidelity are relatively ubiquitous. One important example is given by the undesired transitions that can occur between instantaneous eigenstates of the dynamical Hamiltonian upon the application of an external drive. This is why many driving protocols rely on adiabatic dynamics, where the system follows the instantaneous eigenstates and transitions are naturally suppressed. Ideal adiabatic processes are reversible, making them—in principle—robust. However, to approach ideal adiabatic processes, the dynamics must always be very slow, requiring compromises on the timescales of competing heating and decoherence processes.

Speeding up adiabatic protocols to enable their completion within the system's coherence time is important for the development of any quantum technologies relying on such protocols [1]. One approach to do this is the implementation of optimal driving protocols, which aim to end up with the system in a desired final state. For example, numerically optimized paths can be employed to avoid points where gaps in the spectrum of the system become small, or additional control fields can be tuned to increase the size of these gaps [2–4]. In broad terms, this is the goal of protocols collectively referred to as quantum optimal control. Another option is to design techniques that speed up the adiabatic dynamics, often termed shortcuts to adiabaticity (STA). The primary aim of STA is to entirely remove or suppress diabatic transitions between instantaneous eigenstates of the dynamical Hamiltonian [5,6]. One particularly successful technique is counterdiabatic driving (CD), which was utilized in physical chemistry by Demirplak and Rice [7,8], and was independently introduced by Berry [9] under the name "transitionless driving." CD aims to suppress losses that arise due to fast deformations of the system far from the adiabatic limit by analytically compensating for them in the Hamiltonian. In general, to suppress diabatic losses exactly, the full analytical or numerical solutions

*ieva.cepaite@strath.ac.uk

of the Schrödinger equation are required. This makes the implementation of CD in complex systems—e.g., for many-body dynamics—difficult and requires the need for new techniques to be introduced.

Links between optimal control and STA have existed throughout the development of both approaches [10–12]. This has included the emulation of CD through fast oscillations of the original Hamiltonian [13,14] as well as, more recently, a fusion of machine learning (ML) methods and STA, demonstrating significant improvements for optimizing quantum protocols through ML with the inclusion of concepts from CD [15–17]. In addition, there is work on the application of STA techniques to improve aspects of quantum machine learning protocols [18]. Optimal control is of crucial importance when it comes to the implementation of STA techniques, which are generally constrained by what is practically possible. Examples of this include the use of a variational ansatz for the construction of the CD term that does not require knowledge of the full instantaneous spectrum of the Hamiltonian [19–21] and the recent method of extended STA that analytically improves the control protocol when standard STA methods are intractable [22]. Finally, we note a recent emergence of methods at the intersection of optimal control and CD for digital quantum computation, with examples including the implementation of the quantum approximate optimization algorithm [23,24], the implementation of variational algorithms [25], and insights into optimizing continuous CD protocols [23].

A key ingredient in the development of COLD is a recent approach designed for implementing CD in the setting of larger, more complex systems: local counterdiabatic driving (LCD) [21,26,27]. LCD offers a method to derive *approximate* CD protocols, with the aim of suppressing undesired transitions instead of fully eliminating them. This allows it to account for some physical constraints of the system, e.g., locality conditions. However, the approximate nature of the LCD protocol can lead to poor performance, necessitating the introduction of additional nonlocal, long-range corrections [21,28,29]. If all possible corrections are added then LCD is equivalent to the normal analytical approaches of CD, but the additional terms are generally difficult to control experimentally. COLD offers an alternative approach, with additional control fields that allow for an optimization of the dynamical Hamiltonian for a given form of LCD. The impact of more complex corrections can then be radically reduced, giving a corresponding improvement in the desired protocol. Note that the term local counterdiabatic driving was separately previously utilized to refer to the use of unitary transformations to absorb the CD term into the original bare Hamiltonian [30], which is similar in motivation but distinct from the LCD defined above.

An important consequence of optimizing for a given local order of the LCD is the possibility of bypassing the need to have access to the wave function, dynamics, or experimental data of the given system in order to perform the optimization. LCD is an analytic method and can be calculated using only the coefficients of the Hamiltonian. We find that it is possible to perform numerical optimization of the path of the system by simply minimizing higher-order LCD integrals and/or amplitudes, a method that not only bypasses the need to compute system dynamics but is also independent of system size. This makes it an exceptionally useful tool in practical settings.

The structure of this paper is as follows. First, we give a detailed description of the new method, COLD, with a focus on the elements of quantum optimal control and CD required for its implementation. In Sec. III we explore a two-spin annealing protocol that showcases the strengths of COLD. Section IV describes and analyses the improvements gained with COLD and its combination with other optimal control techniques in the case of state preparation in the Ising model as well as the potential computational advantage of optimizing higher-order integrals of LCD instead of the final state fidelity. In Sec. V we show the improvement that COLD can achieve on the recently realized example of LCD for state transfer on a synthetic lattice in ultracold atoms and in Sec. VI we demonstrate that, when implemented for second-order LCD, COLD can be used to quickly and effectively prepare highly entangled multipartite Greenberger-Horne-Zeilinger (GHZ) states. Finally, in Sec. VII we explore the possibility of minimizing both the power and amplitude of higher-order LCD drives as a means to efficiently optimize COLD parameters, bypassing the requirement of computing system dynamics.

II. AN INTRODUCTION TO COUNTERDIABATIC OPTIMIZED LOCAL DRIVING

A. Quantum optimal control

In the context we consider, we employ quantum optimal control to optimize the function $f(\psi, \theta)$ in the Schrödinger equation

$$\dot{\psi} = f(\psi, \theta), \tag{1}$$

where ψ is the quantum wave function and θ is the set of optimizable control parameters. Optimization of Eq. (1) in most cases means taking the system from an initial state $|\psi_0\rangle$ to a final target state $|\psi_T\rangle$ by finding the optimal values of θ with respect to some target metric (e.g., the time taken to evolve the system from $|\psi_0\rangle$ to $|\psi_T\rangle$). There is a large variety of techniques available to achieve this goal [3,31].

The success metric needs to be defined prior to the optimization of θ. Often this is done by constructing a *cost function*, which in turn defines the optimization landscape. In general, we can optimize for any desired property of the

final state of the system, with some examples being the entropy, energy, energy fluctuations, or some other observable. A commonly used cost function in state preparation is related to the fidelity of the final, postevolution state $|\psi_f(\boldsymbol{\theta})\rangle$ with respect to the target state:

$$C(\boldsymbol{\theta}) = 1 - |\langle\psi|\psi_f(\boldsymbol{\theta})\rangle|^2. \tag{2}$$

In performing such a numerical optimization, it is common to take the target state to be parameterized via a Hamiltonian split into two parts. The first is the so-called *bare* Hamiltonian $H_0(t)$, which can be time dependent and describes the dynamics of the quantum system in question. The second part is then an additional driving term that includes a function f parameterized by the control parameters $\boldsymbol{\theta}$, as well as operators O_{opt} that provide additional degrees of freedom in the dynamics of the system. The full Hamiltonian of the control system is then

$$H_{\boldsymbol{\theta}}(t,\boldsymbol{\theta}) = H_0(t) + f(t,\boldsymbol{\theta})O_{\text{opt}}. \tag{3}$$

The parameters $\boldsymbol{\theta}$ can then be optimized for the optimal dynamics with respect to the metric defined by the cost function.

In this work, we generally use the Powell minimization [32] and dual annealing [33] approaches for the numerical optimization as implemented in PYTHON's *SciPy* [34]. When performing this optimization without any CD terms in the Hamiltonian, we refer to them as bare Powell optimization (BPO) and bare dual annealing (BDA), respectively, with bare referring to the lack of CD. Furthermore, we implement the chopped randomized basis (CRAB) approach [35,36] and combine its methodology with that of COLD, to obtain the method of COLDCRAB. CRAB expands the size of the parameter landscape by employing randomization, usually in the optimized pulse driving the system. The approach was developed for quantum many-body problems whose simulation requires the time-dependent density matrix renormalization group, despite the fact that these were thought to not be tractable in the quantum control setting [36,37]. CRAB has benefits in that it can avoid traps in the control landscape [38], and has built-in flexibility for open-loop or closed-loop optimization [36,39], although these advantages come at a higher computational cost due to requiring a far larger search space for the optimization.

B. Counterdiabatic driving

An important class of optimization problems deals with the case where the initial and final states are ground states of a Hamiltonian $H_0(t)$ at some initial $t = t_i$ and final $t = t_f$ time. The adiabatic theorem guarantees that for an infinitesimally slow transformation of the system $t_f - t_i \to \infty$, it should follow the instantaneous (nondegenerate) ground state of $H_0(t)$ and hence reach the target state with unit fidelity. This process is generally known as quantum annealing.

In large, complex systems with many degrees of freedom, quantum annealing tends to require prohibitively long protocol times due to vanishingly small gaps typically present in such systems. This often makes annealing protocols impractical [23,40]. It has been found that this problem can be formally overcome by using CD, where velocity-dependent terms are added to the Hamiltonian analytically enforcing the adiabatic wave function to be the solution of the time-dependent Schrödinger equation [7–9]. In this case, the dynamical state will follow the instantaneous eigenstate with no transitions regardless of the driving time. The form of the dynamical Hamiltonian enforcing this is [9]:

$$H_{\text{CD}}(t) = H_0(t) + i\sum_n \left(|\partial_t n\rangle\langle n| - \langle n|\partial_t n\rangle\,|n\rangle\langle n|\right) \tag{4}$$

with $|n\rangle \equiv |n(t)\rangle$ the nth eigenstate of the instantaneous Hamiltonian $H_0(t)$. The last term enforces the phases $(\langle n|\partial_t n\rangle)$ on the instantaneous eigenstates, which are arbitrary and thus will be omitted. In general, knowledge of the CD Hamiltonian of Eq. (4) requires knowledge of the full spectrum of $H_0(t)$ at each instantaneous moment in time.

C. Counterdiabatic optimized local driving

We now introduce the main idea of COLD. The principle is to take the same approach as Sec. II B but with the original Hamiltonian given by $H_{\boldsymbol{\theta}}(t,\boldsymbol{\theta})$; see Eq. (3). Quantum annealing then applies to the whole family of control Hamiltonians $H_{\boldsymbol{\theta}}(t,\boldsymbol{\theta})$ as long as the additional control function $f(t,\boldsymbol{\theta})$ vanishes at the protocol boundaries: $f(t_i,\boldsymbol{\theta}) = f(t_f,\boldsymbol{\theta}) = 0$. This flexibility was explored in finding the optimal adiabatic path characterized by, e.g., the shortest distance between the initial and the final states, i.e., a geodesic [41]. A similar geodesic approach was developed in the context of dissipative systems to minimize energy losses [42]. During the protocol, a dynamical Hamiltonian $H_{\boldsymbol{\theta}}(t,\boldsymbol{\theta})$ generally induces transitions between the quantum states that it drives and the question about what is the optimal path remains open.

Hamiltonian $H_{\boldsymbol{\theta}}(t,\boldsymbol{\theta})$ and its eigenstates depend on time only through the driving parameters, which include $\boldsymbol{\theta}$ and any additional control terms in the particular protocol. This makes it convenient to introduce a path in the coupling space parametrized by a dimensionless parameter $\lambda \in [0,1]$ such that both H_0 and f are functions of λ satisfying $H_{\boldsymbol{\theta}}(\lambda = 0) = H_0(t_i)$ and $H_{\boldsymbol{\theta}}(\lambda = 1) = H_0(t_f)$, i.e., being equal to the initial and the final Hamiltonians at the protocol boundaries. By construction, this implies that any additional fields introduced to the bare Hamiltonian H_0 must go to zero at the boundaries. In this way, any

protocol can be uniquely characterized by first specifying the path $f(\lambda, \dot{\lambda})$ in the coupling space manifold and then the time dependence $\lambda(t)$ along it. The path determines the sequence of couplings of the Hamiltonian during time evolution and hence the sequence of ground-state wave functions followed by the driven state. Furthermore, the time dependence encodes the speed of traversing this path. We can then introduce a Hermitian operator called the (path-dependent) adiabatic gauge potential [21]: $A_\lambda = i\sum_n |\partial_\lambda n\rangle\langle n|$, which satisfies a closed-form equation

$$[G_\lambda, H_{\dot{\lambda}}] = 0, \tag{5}$$

where

$$G_\lambda = \partial_\lambda H_{\dot{\lambda}} + \frac{i}{\sim}[A_\lambda, H_{\dot{\lambda}}] \tag{6}$$

with both $H_{\dot{\lambda}}$ and A_λ having a dependence on λ and $\dot{\lambda}(\lambda)$. We note that in the case of a nonlinear Schrödinger equation where the dynamics are described by classical Hamiltonian equations of motion, the commutators need only be replaced with Poisson brackets and the same idea applies [27].

Thus, the CD Hamiltonian reads

$$H_{\mathrm{CD}}(\lambda, \dot{\lambda}) = H_{\dot{\lambda}}(\lambda, \dot{\lambda}) + \dot{\lambda} A_\lambda(\lambda, \dot{\lambda}), \tag{7}$$

and is equivalent to the CD Hamiltonian of Eq. (4) given knowledge of the exact adiabatic gauge potential. However, generally, the adiabatic gauge potential is a very nonlocal object and solutions of Eq. (5) are unstable to small perturbations containing exponentially many terms in the number of degrees of freedom.

LCD aims to find approximate gauge potentials that satisfy particular requirements like robustness and locality, thus circumventing many of the difficulties in determining the second components in Eqs. (4) and (7) exactly. The goal, in essence, is to suppress the most relevant diabatic effects rather than completely eliminate them. This method has recently been experimentally implemented to speed up state transfer for synthetic lattices in ultracold atoms [43], for preparing states in nuclear-magnetic-resonance systems [44], and annealing protocols on an IBM quantum computer [45].

Following the methods of Ref. [21], the problem of finding the optimal adiabatic gauge potential can be cast as the minimization of the Hilbert-Schmidt norm of G_λ, which is equivalent to minimization of the action

$$S(A_\lambda) = \mathrm{Tr}[G_\lambda (A_\lambda)^2] \tag{8}$$

with respect to A_λ. In most cases, this is achieved by first choosing an operator ansatz—i.e., a set of linearly independent operators $\{O_{\mathrm{LCD}}\}$—and then using this set as an operator basis for the adiabatic gauge potential $A_\lambda = \sum_j \alpha_j O^{(j)}_{\mathrm{LCD}}$. The action can then be minimized with respect to the set of coefficients, α. The choice of operators $\{O_{\mathrm{LCD}}\}$ can be made easier when noting that A_λ acts as a generator of motion in the parameter space. This implies that, for, real Hamiltonians like those we will be exploring in the following sections, a good ansatz for the adiabatic gauge potential is one that is *imaginary*. In the example of an Ising spin chain we may take $A_\lambda = \sum_j^N \alpha_j \sigma_j^y$, where j labels the N chain sites, and $\{O_{\mathrm{LCD}}\}$ is a set the y-Pauli matrices.

Without any additional control fields $f(\lambda, \dot{\lambda})$, LCD is essentially an informed choice of the operator set $\{O_{\mathrm{LCD}}\}$ in a way that the resulting control protocol from the minimization of Eq. (8) is optimal for a given $H_0(\lambda)$. In this case we explore the family of Hamiltonians

$$H_{\mathrm{LCD}}(\lambda) = H_0(\lambda) + \sum_j \alpha_j(\lambda) O^{(j)}_{\mathrm{LCD}}. \tag{9}$$

The performance of such LCD protocols is determined by how accurately the variational manifold spanned by the set $\{O_{\mathrm{LCD}}\}$ can approximate an exact A_λ such that Eq. (5) holds.

In the case of the new protocol COLD, we allow for extra exploration of the family of Hamiltonians due to the additional control fields as in Eq. (3). This expands the family of Hamiltonians to

$$H_{\mathrm{COLD}}(\lambda, \dot{\lambda}) = H_0(\lambda) + \alpha(\lambda, \dot{\lambda}) O_{\mathrm{LCD}} + f(\lambda, \dot{\lambda}) O_{\mathrm{opt}}. \tag{10}$$

Note that the coefficients of the optimal control field change the form of the LCD driving coefficients, i.e., $\alpha(\lambda) \to \alpha(\lambda, \dot{\lambda})$. The aim of COLD is then to optimize the coefficients $\dot{\lambda}$ in such a way that the LCD term in the above equation allows for the greatest suppression of nonadiabatic effects for the dynamical Hamiltonian $H_0(\lambda) + f(\lambda, \dot{\lambda}) O_{\mathrm{opt}}$. One can picture it as changing the path that the system takes between its initial and final states, with the express goal of picking a path that maximizes the effects of the approximate counterdiabatic drive given by the second term in the equation. This path will depend on the form of the optimal pulse function, the operators O_{opt}, and on the values of $\dot{\lambda}$. We focus on the optimization of the control parameters $\dot{\lambda}$ for a given choice of $f(\lambda, \dot{\lambda})$ and O_{opt}, although the choice of operators O_{opt} as well as the function of the control pulse $f(\lambda, \dot{\lambda})$ can also be optimized over as an extension.

With COLD, we have two methods to improve on the existing LCD protocol. As previously shown in Refs. [46,47], there is a possibility to add more terms to the LCD making it less local, e.g., through long-range interactions. In the spin chain case, we could take the aforementioned sum over σ^y terms to be the *first-order* ansatz for the

LCD, where higher-order ansatzes would contain sets of operators $\{O_{\mathrm{LCD}}\}$ with terms odd in σ^y such as $\sigma_j^y \sigma_{j+1}^{(z,x)}$. This procedure generally improves the performance of CD protocols at the expense of adding more complex operators that may be experimentally impractical depending on the scenario. Alternately, with COLD and the introduction of additional local control fields to the Hamiltonian, we can improve the performance of LCD at a fixed complexity of the CD term by significantly modifying the adiabatic landscape at intermediate couplings to enhance the performance of the given order of LCD.

In this work we pursue two directions of optimizing the local control fields: numerical optimization of the dynamics and minimization of the higher-order LCD terms. For the most part, we focus on numerical optimization of the dynamics directly, as these will reach optimal values for specific protocols when implemented. COLD opens the possibility of minimizing the higher-order LCD terms instead, which benefits from not requiring calculation of the systems dynamics. This approach, as discussed in Sec. VII, allows for optimal control procedures using COLD to be implemented for large systems that would be cumbersome for procedures that require the numerical optimization of the dynamics.

We also note that, while it may seem prudent to treat the LCD coefficients $\alpha(t)$ as control pulses that may be parameterized and optimized in the same way that f is, we find that this method fails to perform well compared to using the variational form of the LCD as we have done. This is likely due to a difficulty in choosing an accurate form of the drive as well as parameterizing it. On top of that, the loss function space in this case may become intractable when compared to that of COLD as we have presented it in this section.

Furthermore, we compare COLD to the use of CRAB, as discussed in Sec. II A. An advantage of COLD is that it can be combined with many advanced optimal control procedures, owing to the standard way additional control fields are introduced to the Hamiltonian. In this work we find the combination of COLD and CRAB particularly useful and we refer to this as COLDCRAB.

III. TWO-SPIN QUANTUM ANNEALING

To showcase and explore the use of COLD in a relatively simple setting, we consider a two-spin quantum annealing problem with bare Hamiltonian

$$H_0(t) = -2J\sigma_1^z\sigma_2^z - h(\sigma_1^z + \sigma_2^z) + 2h\lambda(t)(\sigma_1^x + \sigma_2^x), \tag{11}$$

where σ_j^a, $a \in \{x,y,z\}$, are the Pauli matrices applied to spins indexed by j. For the scaling function $\lambda(t)$, we pick the form

$$\lambda(t) = \sin^2\left(\frac{\pi}{2}\sin^2\left(\frac{\pi t}{2\tau}\right)\right) \tag{12}$$

such that $\lambda(0) = 0$ and $\lambda(\tau) = 1$. We consider the case of $J/h = 0.5$, which means that the spins start in the initial state of $|\uparrow\uparrow\rangle$ and finish in a superposition of all of the symmetric states.

As discussed in Ref. [21], since H_0 has a standard Ising spin chain form, the first-order LCD terms are given by the following ansatz for the adiabatic gauge potential:

$$A_\lambda = \alpha \sum_{i=1}^{N^2} \sigma_i^y \tag{13}$$

with the sum being over the full length of the N spin chain. Minimizing Eq. (8) for this A_λ with respect to the coefficient α gives

$$\alpha = -\frac{h^2}{4(h\lambda)^2 + h^2 + 4J^2}. \tag{14}$$

To further improve on the first-order LCD, we can implement COLD, as we will discuss shortly, or we can introduce higher-order terms to the ansatz for A_λ. This second method serves as a good benchmark against COLD, since it offers an improvement to first-order LCD in the same way as COLD does, but requires more complicated interactions between the two spins increasing the implementation overhead. The second-order LCD can be found by taking an ansatz for the adiabatic gauge potential:

$$A_\lambda^{(2)} = \alpha \sum_j \sigma_j^y + \gamma(\sigma_1^x\sigma_2^y + \sigma_1^y\sigma_2^x) + \zeta(\sigma_1^z\sigma_2^y + \sigma_1^y\sigma_2^z). \tag{15}$$

Here to solve for α, γ, and ζ we once again minimize the action given by Eq. (8) and obtain three coupled equations that can be solved numerically (see Appendix A for a detailed derivation).

We now consider three distinct cases in this two-spin quantum annealing example: no LCD, first-order LCD, and second-order LCD. The fidelity of the final state for each case over a wide range of driving times is shown in Fig. 1(a), with an easily distinguishable advantage in the case of LCD. The final fidelity where no LCD is implemented decreases rapidly as the ramp times are made short, with the system getting stuck in its initial state. On the contrary, first-order LCD retains good final state fidelities into short times, as the driving Hamiltonian becomes that of only the LCD term. The second-order LCD then gives unit fidelity, in agreement with previous observations [46], as

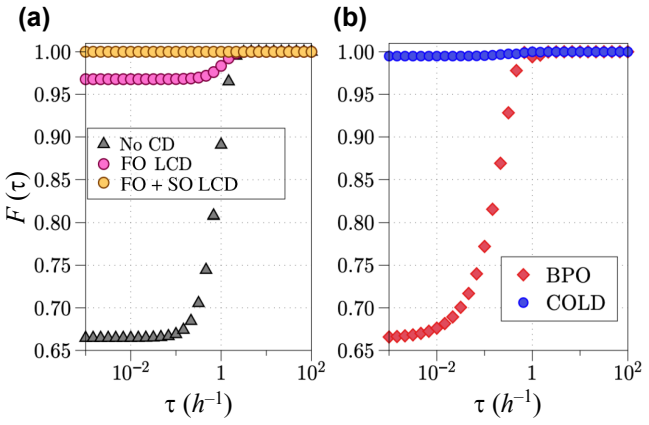

FIG. 1. Optimization of the annealing protocol for the two-spin Hamiltonian given by Eq. (11) for $h/J = 2$. (a) Final fidelities of the annealing protocol with triangles (black) representing the case where no CD is applied and circles showing the case of first-order (FO) LCD (pink) as well as the combination of first- and second-order (SO) LCD (orange). (b) Final fidelities achieved when using the optimal control method BPO (red diamonds) and the new approach of COLD (blue circles), both with $N_k = 1$.

for a two-spin Hamiltonian the highest-order corrections are those including two spin terms.

We now add an optimizable term, as described in Sec. II A, so that the new Hamiltonian reads

$$H_\beta(t) = H_0(t) + \sum_{k=1}^{N_k} \beta^k \sin(\pi k t/\tau) \sum_i \sigma_i^z \tag{16}$$

with N_k the number of optimization coefficients β, and $\beta^k \in \beta$ the coefficient of the kth frequency of the control function. Note that we consider

$$f(t,\beta) = \sum_{k=1}^{N_k} \beta^k \sin(\pi k t/\tau) = \sum_{k=1}^{N_k} \beta^k \sin(\pi k g(\lambda)) \tag{17}$$

with

$$g(\lambda) = \frac{2}{\pi} \arcsin \sqrt{\frac{2}{\pi} \arcsin(\sqrt{\lambda})}. \tag{18}$$

The form of the additional control field fulfils the requirement that the boundary conditions are $H(0) = H_0(0)$ and $H(\tau) = H_0(\tau)$. Note that numerically optimizing the β^k for the best final state fidelity *without* adding LCD terms results in the BPO method introduced in Sec. II A. We show the results of BPO in Fig. 1(b), where it is observed that BPO gives better results than the case of no LCD in Fig. 1(a). However, for short times, the BPO approach still results in the system getting stuck in the initial state.

Finally we present and compare the results of the new method, COLD. In this case the Hamiltonian before adding LCD terms is given by Eq. (16) and the coefficient of the first-order LCD is

$$\alpha = -\frac{h(h + f(\lambda,\beta)) + h\lambda \dot{f}(\lambda,\beta)/\dot{\lambda}}{4(h\lambda)^2 + (h + f(\lambda,\beta))^2 + 4J^2}. \tag{19}$$

Note that the optimization of the additional control field also feeds into the coefficient of the adiabatic gauge potential during the dynamics as discussed in Sec. II C. The results of the COLD approach for this two-spin annealing protocol are shown in Fig. 1(b), where we observe an improvement of the final state fidelity beyond what is possible with first-order LCD alone in Fig. 1(a). In this example, LCD alone reaches a final state fidelity of $1 - F = 3\%$ at short times; however, COLD improves this error in the final state to $1 - F = 0.005\%$. This is due to the extended family of dynamical Hamiltonians in Eq. (10) owing to the addition of an optimizable control field. This result shows that COLD can provide an advantageous alternative to the addition of higher-order LCD that may be experimentally impractical.

We have found that COLD performs better than LCD of the same order or BPO when the system dynamics are calculated numerically. This does not, however, imply anything about the performance of COLD in more complex scenarios, like in the case of an unknown target ground state. In that case the fidelity is a poor optimization metric. There is, however, a way to come to the same conclusions as those presented in Fig. 1 without the need to compute the dynamics exactly. We can do this by first using a guess for the COLD protocol to find the approximate adiabatic gauge potential and then minimizing the integral of the norm of the second-order correction to the adiabatic gauge potential along the path. Note that the ground state can in turn be obtained through first-order COLD, so there is no need to diagonalize the Hamiltonian. This integral should be small if COLD has implemented a dynamical Hamiltonian that makes the first-order adiabatic gauge potential the leading term. It is effectively a measure of the error of COLD and can be given by

$$I_1(0) = \int_0^\tau dt' [\langle \psi_g(t')| O^2(t') |\psi_g(t')\rangle - \langle \psi_g(t')| O(t') |\psi_g(t')\rangle^2]^{1/2} \tag{20}$$

with $|\psi_g(t)\rangle$ the instantaneous ground state along the path and

$$O(t) = \gamma(t)(\sigma_1^y \sigma_2^x + \sigma_1^x \sigma_2^y), \tag{21}$$

one of the second-order correction terms. In order to confirm this is the case, we compare the different paths—COLD and LCD only—in the two-spin example in order to determine if I_1 is small for COLD. If I_1 is small when compared to the same measure for lower-order LCD

as $\tau \to 0$ then we know that COLD is enforcing a better dynamical Hamiltonian. In the case of the two-spin annealing protocol we find that, as $\tau \to 0$, $I_1 \to 0.04$ for COLD and $I_1 \to 0.2$ for LCD, showing that COLD is minimizing the second-order correction along the path. A simpler integral

$$I_2(\gamma) = \int_0^\tau dt' |\gamma(t')| \tag{22}$$

also reflects this correction in this two-spin example, with $I_2 \to 0.03$ for COLD and $I_2 \to 0.1$ for LCD as $\tau \to 0$. This is particularly useful in more complex scenarios as I_2 is relatively straightforward to calculate, as we demonstrate in the next section. We also observe the reduction of the corresponding integrals of the $(\sigma_1^y\sigma_2^z + \sigma_1^z\sigma_2^y)$ term of the second-order LCD. By minimizing these integrals, it is possible to extend the COLD approach to more complex scenarios, including where the exact calculation of the dynamics is not possible.

IV. ONE-DIMENSIONAL ISING MODEL

In this section we apply COLD for state preparation on a one-dimensional Ising spin chain in the presence of a transverse and longitudinal field. We consider an annealing protocol where the aim is to prepare the ground state across the Ising phase transition. The annealing Hamiltonian is given by

$$H_0(t) = -J\sum_j^{N-1} \sigma_j^z\sigma_{j+1}^z + Z_0\sum_j^N \sigma_j^z + \lambda(t)X_f\sum_j^N \sigma_j^x, \tag{23}$$

where Z_0 is a small offset parameter to break ground-state degeneracies and X_f is the final x-field strength. Note that the breaking of the ground-state degeneracies is not a requirement but allows for easier consideration of the adiabatic path. As before, $\lambda(t)$ is a scaling function that has the boundary conditions $\lambda(0) = 0$ and $\lambda(\tau) = 1$, with τ the driving time. This means that we start from the ground state of all spins up and drive across the quantum phase transition to the ground state that is a superposition of all basis states. We again take the scaling function to be given by Eq. (12). In this example, we use $X_f = 10J$ and $Z_0 = 0.02J$.

For the Hamiltonian of Eq. (23), the LCD to first and second orders is well known, as the wave functions are entirely real. We take the first-order adiabatic gauge potential to be given by

$$A_\lambda = \alpha\sum_j^N \sigma_j^y, \tag{24}$$

where the coefficients for the general periodic spin chain of Eq. (23) are

$$\alpha(\lambda) = \frac{1}{2}\frac{Z_0X_f}{Z_0^2 + \lambda^2X_f^2 + 2J^2}. \tag{25}$$

Note that the quoted α above is technically for a periodic or infinite size system, with $J^2 \to J^2(1 - 1/N)$ for a finite system. However, we find that the inclusion of the factor for the finite system sizes we consider only changes the final achieved converged fidelities at short times by about $10^{-6}\%$. The second-order adiabatic gauge potential is of the form

$$A_\lambda^{(2)} = \alpha\sum_j \sigma_j^y + \gamma\sum_j (\sigma_j^x\sigma_{j+1}^y + \sigma_j^y\sigma_{j+1}^x) + \zeta\sum_j (\sigma_j^z\sigma_{j+1}^y + \sigma_j^y\sigma_{j+1}^z) \tag{26}$$

with the coefficients α, γ, and ζ again obtained by minimizing the action given by Eq. (8) and solving the coupled set of equations numerically (see Appendix A for a detailed derivation).

In this example, optimal control is implemented by introducing an additional driving field so that the dynamical Hamiltonian is given by

$$H_\beta(t,\beta) = H_0(t) + \sum_j f_j(t,\beta)\sigma_j^z \tag{27}$$

with β being the terms to optimize over. We take our additional terms to again respect the boundary conditions $f_j(t = 0,\beta) = 0$ and $f_j(t = \tau,\beta) = 0$, meaning that a natural choice is

$$f_j(t,\beta) = \sum_k^{N_k} \beta_k \sin(\omega_k t/\tau) = \sum_k^{N_k} \beta_k \sin(\omega_k g(\lambda)) \tag{28}$$

with $\omega_k = 2\pi k$ the kth principal frequency and $g(\lambda)$ given by Eq. (18). To implement the CRAB algorithm discussed in Sec. II A, we use $k \to k(1 + r_k)$ instead with r_k drawn from a uniform random distribution $r_k \in [-0.5, 0.5]$. Note that there is a strong distinction between CRAB, which is an established optimal control method in its own right and our own version COLDCRAB, which includes a LCD term along with the optimal control field in the Hamiltonian. To be more precise, the COLDCRAB Hamiltonian can be expressed as

$$H_{CC}(t,\beta,\mathbf{r}) = H_0(t) + \alpha(t,\beta,\mathbf{r})\sum_j^N \sigma_j^y + \sum_j f_j(t,\beta,\mathbf{r})\sigma_j^z, \tag{29}$$

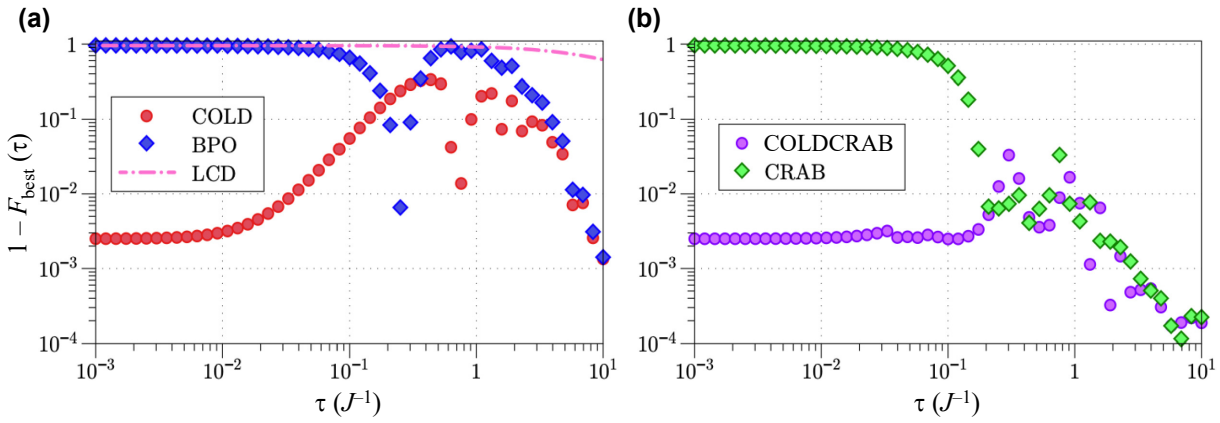

FIG. 2. Optimization of the annealing protocol for the Ising model given by Eq. (23) for $N = 5$ spins. (a) A comparison of final state fidelities (plots show $1 - F$) for different driving times using the optimal control technique BPO (blue diamonds), first-order LCD (pink dash-dot line), and COLD (red circles). The same is shown in (b) for CRAB (green diamonds) and COLDCRAB (purple circles). CD-enhanced techniques (COLD and COLDCRAB) introduced in this work show a clear convergence to good fidelities at short driving times. All results are for the best (lowest) fidelity achieved over 500 optimizations.

where, for each optimizable parameter θ^k associated with a kth principal frequency, we also assign a random value $r_k \in \mathbf{r}$ as described earlier. Note that the dependence on $\mathbf{r}$ is inherited by the LCD drive term α, since it is a function of $f(t, \theta, \mathbf{r})$.

As before, we choose the first-order adiabatic gauge potential given by Eq. (24) and find that the coefficients are

$$\alpha(\lambda, \theta) = \frac{X_f}{2} \frac{(Z_0 + f(\lambda, \theta)) - \lambda \dot{f}(\lambda, \theta)/\dot{\lambda}}{(Z_0 + f(\lambda, \theta))^2 + \lambda^2 X_f^2 + 2J^2}. \tag{30}$$

Note that, with the introduction of the additional control fields f, it is possible for α to be nonzero at the start or end of the protocol, as $\dot{f}$ is not fixed to be zero. However, this can be enforced by a suitable choice of the additional control field; we consider replacing $\alpha \to S(\lambda)\alpha$, where $S(\lambda)$ is a scaling function that tends to zero as $\lambda \to 0$ and $\lambda \to 1$. We find that the scaling function only has a minimal effect on the final fidelities observed. This issue could also be resolved by a suitable choice of f, with our example drive being an extreme case as $\dot{f}$ is maximal at the boundaries of the protocol. Note that this issue is present in LCD as much as in COLD and we have chosen to highlight it here as it may become a concern in an experimental setting where a discontinuous drive is simply impossible at the beginning and end of a protocol. The suitable choice of the form of f in a given example is a problem we leave for future work, with our focus being on the introduction of the COLD protocol.

We first compare the final state fidelity when using COLD versus BPO, as shown in Fig. 2(a) for different driving times in a system of $N = 5$ spins and a single $N_k = 1$ optimization coefficient. As expected, at long timescales the two methods agree as we approach the adiabatic limit of the dynamics. However, at shorter time scales the difference in behavior is dramatic. We observe that the BPO approach fails in the case of very fast driving as the state gets stuck in the initial state but the COLD approach converges to $1 - F \approx 10^{-3}$. We note that the advantage achieved by COLD is not due to the introduction of first-order LCD terms alone, as in Fig. 2(a) we see that this will result in $F = 0.0440$ for $\tau = 10^{-3} J^{-1}$. COLD is instead achieving this by making the LCD term dominant for the dynamical Hamiltonian through the additional control fields.

To confirm this, we plot the maximum amplitudes of both the first- and second-order adiabatic gauge potentials in Fig. 3, where Fig. 3(a) shows the case of no optimization and Fig. 3(b) the case of applying COLD. We can see that without COLD the second-order $(\sigma_j^x \sigma_{j+1}^y + \sigma_j^y \sigma_{j+1}^x)$ corrections to the LCD are far larger than the first order, resulting in the small final state fidelities when only first-order LCD is implemented. In the case of COLD, this relationship reverses and the first-order LCD terms dominate the dynamics. This gives us an indication that one way to optimize the control pulse may be a minimization of higher-order LCD terms, which we explore further in Sec. VII.

We find that the results of BPO and COLD at short driving times are stable against increasing system size, as shown for $\tau = 10^{-2} J^{-1}$ in Fig. 4(a), with only a small decrease in final state fidelity for larger systems with COLD. Similarly, increasing the number of optimization coefficients N_k results in little improvement in the values obtained at short times for this example, as shown in Fig. 4(b). It is possible that in more complex systems, more optimization coefficients will be needed to gain a larger advantage. We also note that, by increasing the number of coefficients, we are increasing the complexity of the cost

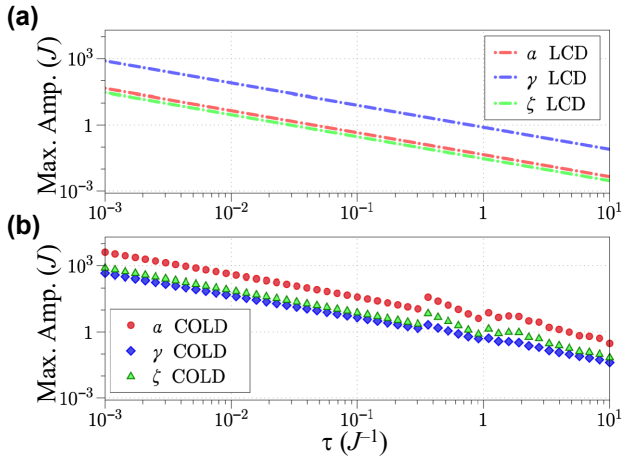

FIG. 3. Maximum amplitudes of CD terms in the Ising model annealing protocol for (a) first- and second-order LCD only with no additional optimal control fields and (b) the COLD approach optimized for the best final state fidelity implementing first-order LCD as shown in Fig. 2(a). The plot shows the maximum amplitude at each driving time for the first-order α (red circles) and the two second-order terms γ (blue diamonds) and ζ (green triangles) as given in Eq. (26) (although the second-order LCD is not actually implemented in COLD). An inversion in the strength of the second-order and first-order LCD terms for (a) no additional optimal control fields and (b) the addition of optimal control fields shows that COLD implements a dynamical Hamiltonian that is favorable for the applied order of LCD (first order in this case).

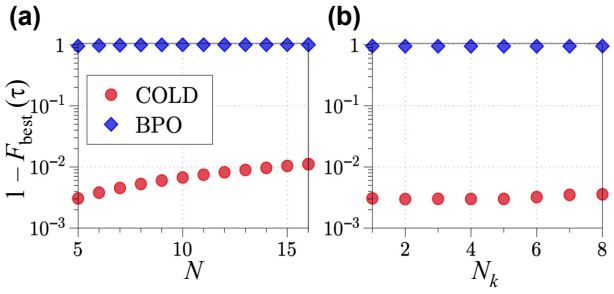

FIG. 4. Scaling of fidelities (plots show $1 - F$) in the annealing protocol for the Ising model with (a) system size N and (b) optimization parameters N_k at driving time $\tau = 10^{-2}J^{-1}$. Plots show a comparison between BPO (blue diamonds) and COLD (red circles). In (a) we see that the COLD fidelity decreases as a function of N but remains quite high when compared to BPO, while (b) shows the nonexistent improvement for both BPO and COLD with an increasing number of parameters in the $N = 5$ spin case. Once again, plotted best fidelities are obtained across 500 optimizations.

function landscape to be explored by the minimization procedure, hence leading to slightly worse final fidelities. This can mean that alternative approaches than the Powell minimization used so far, e.g., that of CRAB, could be better suited for probing the cost function for high N_k. We also note that this lack of improvement in the results is likely the consequence of the form of the control field given by Eq. (28) rather than due to a failure of the optimizer in the face of a complex parameter space. We find that the parameter space is relatively smooth in the case of $N_k = 1, 2, 3$ and a better solution for this form of control field does not exist.

We now consider the combined method of COLDCRAB for this annealing example as shown in Fig. 2(b). We point out that with our application of CRAB in this scenario we are not enforcing β to be zero at the start and end of the dynamics, allowing for their to be a tuning of the z-field offset. This is consistent between CRAB and COLDCRAB and therefore does not influence our comparison of the two. First, it is important to note that CRAB alone results in a overall speedup of the dynamics for a high final state fidelity $1 - F \approx 10^{-3}$ at long timescales. However, CRAB still suffers from getting stuck in the initial state at fast driving times and the final state fidelity again tends to zero. This is not the case for COLDCRAB, which converges to large final state fidelities $1 - F \approx 10^{-3}$ at short driving times $\tau \leq 10^{-1}J^{-1}$. Note that the difference between the convergence to final state fidelities are only marginally different between COLD and COLDCRAB at longer times, but at short times COLDCRAB performs a lot better. Further improvement could be gained by combining COLD with more advanced versions of CRAB or other optimal control methods.

As shown in Fig. 3, the amplitude of the driving required to achieve the fidelities discussed so far scales with the driving time. Practical scenarios will necessarily place limits both on achievable driving times and the maximum amplitude of any term that is being driven. However, the scaling of the drivings shown do not mean that everything diverges in the limit of $\tau \to 0$. To see this, we can first write the Scrödinger equation for COLD as

$$i\partial_t |\psi\rangle = (H_\beta + \dot{\lambda} A_\lambda) |\psi\rangle, \tag{31}$$

we then divide through by $\dot{\lambda}$ to get

$$i\partial_\lambda |\psi\rangle = \left(\frac{H_\beta}{\dot{\lambda}} + A_\lambda\right) |\psi\rangle, \tag{32}$$

in the limit of $\tau \to 0$ then $\dot{\lambda} \to \infty$ to result in the Hamiltonian term disappearing, or in other words, we turn off the Hamiltonian. We then only drive the system in the $\tau \to 0$ limit with the COLD or LCD driving term:

$$i\partial_\lambda |\psi\rangle = A_\lambda |\psi\rangle. \tag{33}$$

In this limit then λ plays the role of time, and this could then be implemented in a practical scenario in finite time as it corresponds to some manipulation of the couplings in the

system. This renormalized time cannot then be infinitesimally short if the couplings are bounded, but we have shown that the protocol does not diverge as $\tau \to 0$. In the case of a spin chain, evolving under Eq. (33) is effectively to first order in LCD implementing independent single spin rotations along the chain, and COLD can be easily applied [48,49].

If it is not possible to switch off the Hamiltonian as discussed above then as an alternative we can implement COLD with experimental constraints accounted for directly in the optimal control minimization. We consider an extreme example of constraints to show that even in this scenario COLD can provide an advantage and corresponding speed-up. In the constrained case the annealing protocol remains that of Hamiltonian (23) but we choose to introduce a bound of X_f on the maximum amplitude of all drivings. This makes it so that no optimal control or LCD term can go beyond the original amplitude of the x-field drive. We show the final state fidelities achieved for the constrained example in Fig. 5. As can be seen in Fig. 5(a), COLD provides a substantial improvement beyond what is achievable with BPO. BPO manages $F < 0.5$ for $\tau < 1J^{-1}$, but COLD can reach final state fidelities $F \gtrsim 0.9$ for $\tau < 1J^{-1}$. The real improvement, however, comes with the application of CRAB and COLDCRAB. CRAB already improves the fidelities substantially, and would allow for a speed up in the annealing protocol, but with COLDCRAB, the final state fidelities are even better, with $F \gtrsim 0.99$ achievable when approaching $\tau \gtrsim 0.1J^{-1}$. Signs are seen of the onset of the convergence to small values for COLDCRAB in Fig. 5(b) before the maximum amplitude required becomes too large and the short time results tend towards zero fidelity and the state being stuck again. With this example and the discussion on implementation via turning off the Hamiltonian, we have shown that COLD is capable of delivering improvements beyond other schemes even for practical problems with strict and rather extreme constraints.

V. TRANSPORT IN A SYNTHETIC LATTICE

The efficient transfer of states between opposite ends of a lattice could have future applications in the settings of quantum computation and simulation due to its promise of efficient transport of information [50]. This objective is often tackled in the setting of ultracold atoms in optical lattices. While the problem can be tuned to be a single-particle system and the analytical solutions of the corresponding instantaneous Schrödinger equation are known [51,52] even for a finite system [53], efficient evolution for state transfer is not straightforward. This is due to the fact that the majority of the states are delocalized across the lattice, meaning that the $|\psi\rangle\langle\psi|$ terms of the CD Hamiltonian of Eq. (4) are global in reach. It is normal to consider this system in the tight-binding limit where the implementation of global terms is not straightforward. Such terms can be generated via the interactions of the atoms with cavity modes [54,55] or from dipolar interactions [56–58]. However, it would be ambitious to expect this control to be general enough to implement the CD Hamiltonian of the exact solutions. This is one of the reasons that LCD has been pursued in this setting.

Recently, LCD has been successfully applied to improve an adiabatic rapid passage (ARP) protocol for population transfer across a synthetic lattice [43]. In this realization, population transfer was achieved in a synthetic tight-binding lattice of laser coupled atomic momentum states. We consider the same problem as in Ref. [43] but with the improvement that can be gained by COLD. This system is described by the Hamiltonian

$$H_0(t) = -\sum_n J_n(t)(c_n^\dagger c_{n+1} + \text{H.c.}) + \sum_n V_n(t) c_n c_n^\dagger, \tag{34}$$

where $J_n(t)$ is the time-dependent tunneling that describes the nearest-neighbor coupling, $V_n(t)$ is the on-site energy offset with respect to neighboring sites, and $c_n^\dagger$ (c_n) is the creation (annihilation) operator on a given synthetic lattice site. In the ARP protocol, the population gets moved from one end of the lattice to the other by linearly ramping the lattice from a positive tilt to a negative tilt via

$$J_n(t) = J_0(1.1 - \lambda) = J_0 0.1 + \frac{t}{\tau}, \tag{35}$$

$$V_n(t) = nV_0 2(\lambda - 1/2) = nV_0 1 - \frac{2t}{\tau}, \tag{36}$$

where $V_0 = 4J_0$ is the initial site energy slope and J_0 is the characteristic tunneling scale of the lattice. The scaling function in this case is given by

$$\lambda(t) = 1 - \frac{t}{\tau}. \tag{37}$$

In order to implement LCD as shown in Ref. [43], the first-order LCD can be accounted for by taking

$$J_n(t) \to J_{n,\mathrm{CD}}(t) e^{-i\varphi_{n,\mathrm{CD}}(t)}, \tag{38}$$

where

$$J_{n,\mathrm{CD}}(t) = \sqrt{J_n(t)^2 + (\alpha_n(t)/\tau)^2}, \tag{39}$$

$$\varphi_{n,\mathrm{CD}}(t) = \arctan\left(-\frac{J_n(t)\tau}{\alpha_n(t)}\right), \tag{40}$$

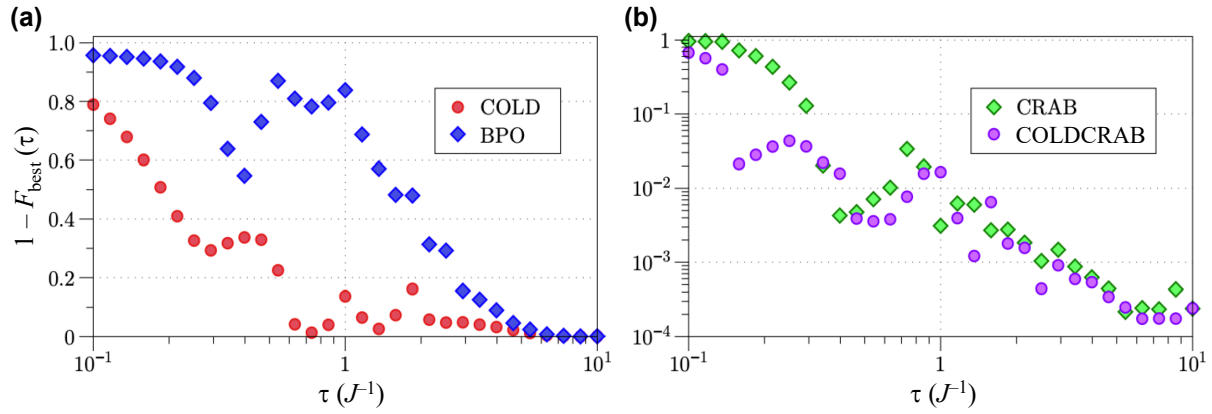

FIG. 5. Optimization of the *constrained* annealing protocol for the Ising model for $N = 5$ spins with a maximum amplitude limit on each term in the Hamiltonian of Eq. (23) of $10J$. (a) A comparison between BPO (blue diamonds) and COLD (red circles) that both give much lower fidelities (plots show $1 - F$) than in the unconstrained case in Fig. 2, although COLD persists in giving better results. In (b) the comparison is between CRAB (green diamonds) and COLDCRAB (purple circles) that show orders of magnitude better fidelities than those in (a), with COLDCRAB eking out higher fidelities at short driving times. The plotted best results are obtained from 200 optimizations for each method.

and the $\alpha_n(t)$ are the CD terms that can be found by solving a set of linear equations

$$\begin{aligned} &-3(J_nJ_{n+1})\alpha_{n+1} + (J_{n-1}^2 + 4J_n^2 + J_{n+1}^2)\alpha_j \\ &-3(J_nJ_{n-1})\alpha_{n-1} + (V_{n+1} - V_n)^2\alpha_n \\ &\quad = -\partial_\lambda J_n(V_{n+1} - V_n). \end{aligned} \tag{41}$$

In order to implement COLD, we include additional terms to the tunneling of the lattice

$$J_n(t) \to J_n(t, \beta) = J_n(t) + f(t, \beta), \tag{42}$$

which can then be incorporated into the forms of both $J_{n,\text{CD}}(t)$ and $\varphi_{n,\text{CD}}(t)$. We again want the additional control terms to go to zero around the problem boundaries and a natural choice is the same as in the Ising spin chain example in Eq. (28). The parameters β are optimized as before by minimizing with respect to the fidelity of the final state, where the population has been fully transferred to the opposite lattice site.

We first consider a system size of $N = 7$ sites that was successfully experimentally probed in Ref. [43], where final state fidelities of 0.75 were achieved for $\tau = 1$ ms with a final tunneling strength of $J/\sim = 1/2\pi$ kHz (equivalent to $\tau \approx 1J^{-1}$ in our units). We initially confirm the breakdown of ARP in this setting for fast times, and the success of the LCD protocol at short times, as shown in Fig. 6(a) and found in Ref. [43]. Implementing BPO on its own manages to enhance the achievable fidelities at intermediate times of $\tau > 0.03J^{-1}$. However, eventually, as observed in all scenarios in this work, BPO becomes stuck in the initial state at fast times, and the fidelity goes to zero. Implementing the newly introduced COLD protocol achieves an order-of-magnitude improvement in the fidelity over LCD. This is also plotted in Fig. 6(a) alongside previous results of ARP and first-order LCD.

One concern could be that COLD is achieving this improvement by simply pumping power into the tunneling term, but as we can see in Fig. 6(c), the maximum amplitude of the tunneling term tracks that of LCD. A key issue for experiments is the maximum amplitude achievable by a driving term and with this result we can stipulate that COLD is likely to be feasible in the same regimes as LCD in this synthetic lattice system. There is single outlier at intermediate times as indicated by the single point peaking in maximum amplitude in Fig. 6(c); this is the exception to the rule, where the optimization has found a marginally higher fidelity [see the offset point in Fig. 6(a)] by pumping in power.

A large concern for state transfer techniques is the robustness of a protocol with respect to an increas-ing system size. We show the best achievable fidelities with increasing system size for both BPO and COLD in Fig. 6(c). While both protocols show a decreasing fidelity with system size as is to be expected, once again COLD does not suffer from getting stuck in the initial state. This is shown by the BPO fidelities going to unity for large systems in Fig. 6(c), and is the same mechanism for this as for the short driving times in Fig. 6(a).

Another concern could be that BPO will beat COLD if enough parameters are allowed for the optimization, i.e., if we increase N_k enough. We observed no evidence of this for the Ising model example and we again do not observe this in this synthetic lattice example, as is shown in Fig. 6(d). Small improvements are made in the fidelities achieved with BPO and COLD for larger N_k, but this is not substantial.

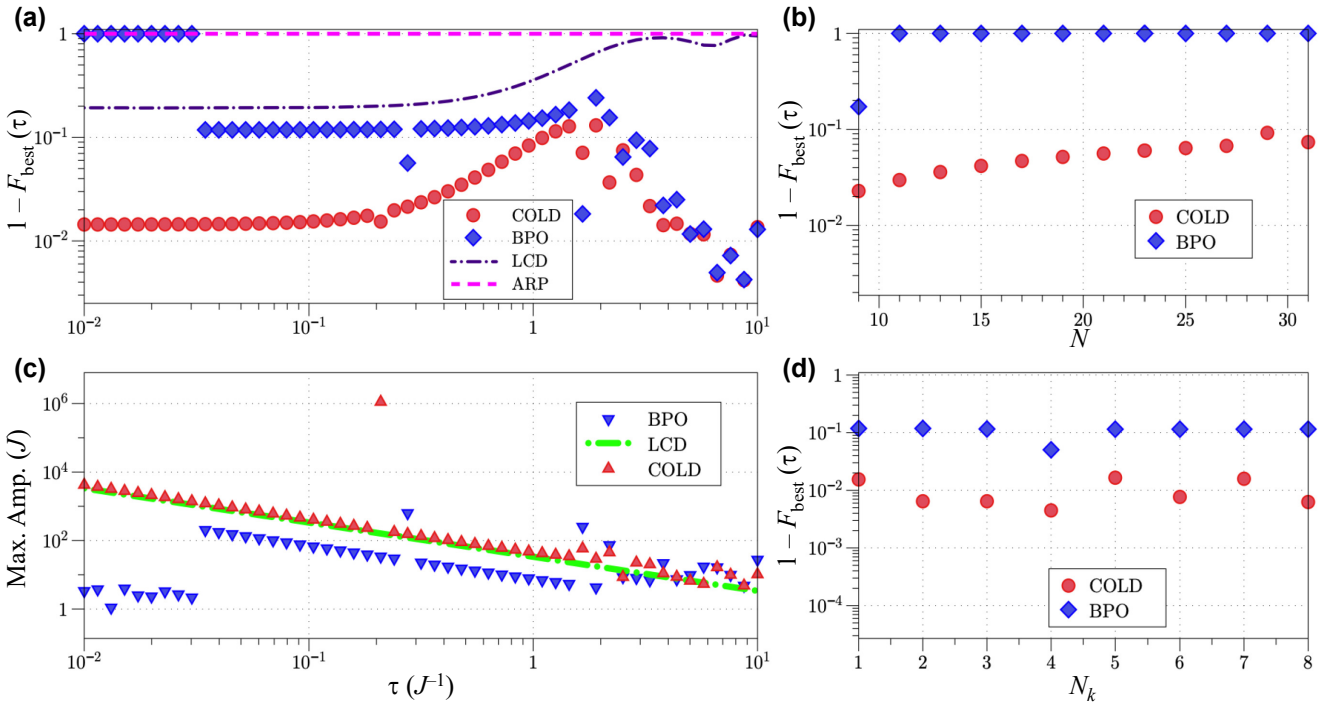

FIG. 6. Optimization of state transfer in a synthetic lattice. In (a) we compare the fidelities (plots show $1 - F$) obtained via the bare ARP protocol (pink dashed line) and first-order LCD previously implemented in Ref. [43] (purple dash-dot line) to BPO (blue diamonds) and the COLD method (red circles). (c) Maximum amplitude of the tunneling term at each driving time for LCD (green diamonds) as given by Eq. (38) as well as COLD (red triangles) that includes additional control parameters as shown in Eq. (42) and BPO (blue triangles) that omits the modifications due to CD but retains the control terms θ. In both (a) and (c) we simulate $N = 7$ lattice sites and use the $N_k = 1$ parameter for optimization of BPO and COLD. (b) Scaling of fidelities with an increasing number of lattice sites (where $N_k = 1$) for both COLD (red circles) and BPO (blue diamonds), noting that the latter performs very poorly for $N > 9$. Panel (d) shows the same for the number of parameters while keeping $N = 7$, with the trend indicating that increasing N_k does not lead to better fidelities in either the BPO or COLD case. Note that both (b) and (d) are simulated for driving time $\tau = 0.5J^{-1}$ and the best fidelities are obtained across 500 optimizations.

VI. GHZ STATE PREPARATION

As a final example, we focus on the preparation of multipartite GHZ [59] states

$$|\psi_{\mathrm{GHZ}}\rangle = \frac{1}{\sqrt{2}}(|0\rangle^{\otimes N} + |1\rangle^{\otimes N}) \tag{43}$$

in a system of frustrated spins [see Fig. 7(a)]. We start out with a system of all spins pointing down and drive a bare Hamiltonian of the form

$$H_0(t) = -J\sum_{j}^{N-1}\sigma_j^x\sigma_{j+1}^x + \sum_{j}^{N-2}\sigma_j^x\sigma_{j+2}^x - h(1-\lambda(t))\sum_{j}^{N}(\sigma_j^x + \sigma_j^z), \tag{44}$$

where $J = 1$ and $h = 10J$ with the same $\lambda(t)$ as used previously, given by Eq. (12). The form of the LCD to first and second orders is the same as in the case of the Ising spin chain [see Eqs. (24) and (26)] with the couplings in the case of the second order now including the additional terms between spins j and $j + 2$.

In order to explore the versatility of combining optimal control with LCD, for this example, we design our optimal control drive according to the gradient ascent pulse engineering (GRAPE) method [61]. Our control coefficients θ^k are now discretized on a finite grid of N_m time intervals t_m with uniform step δt to obtain control sequences in which individual elements $\theta^{k,m}(t_m)$ are treated as continuous parameters

$$f(t,\theta) \to [[f(\theta^{1,1},t_1),\ldots,f(\theta^{1,N_m},t_{N_m})],\ldots,[f(\theta^{N_k,1},t_1),\ldots,f(\theta^{N_k,N_m},t_{N_m})]], \tag{45}$$

where the total driving time $\tau = N_m\delta t$ and k is used to denote a localized drive for a subset of spins where N_k is the total number of control pulses. As in the Ising model case, we take our optimal control Hamiltonian to be of the form in Eq. (27) with each kth drive acting on the specified subset of spins with local σ^z operators. At each time

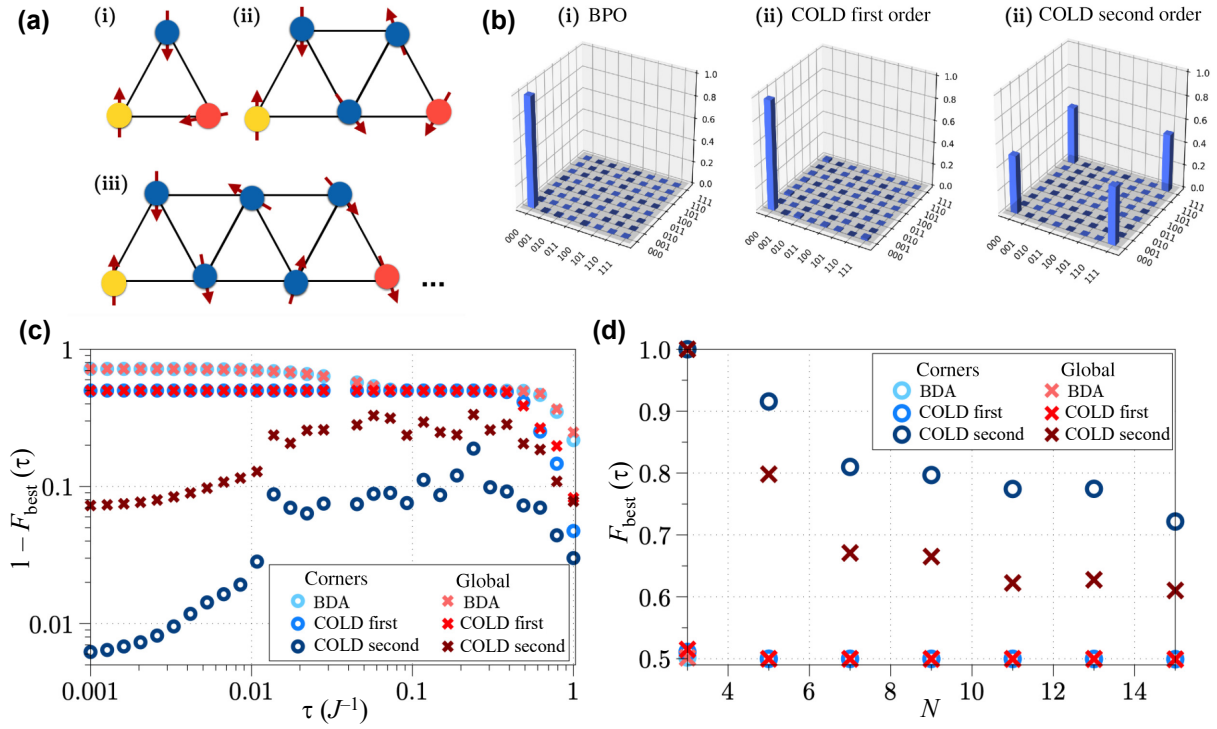

FIG. 7. GHZ state preparation in systems of frustrated spins. Spins are arranged in triangular formations as depicted in (a) for (i) three, (ii) five, and (iii) seven spins, with spins on the vertices and edges representing couplings. In the case of corner optimization, three separate optimizable drives are applied: one for the yellow corner spin, one for the red corner spin, and then a third drive for all of the blue spins in between. (b) Density matrix plots of the final state of a three-spin triangle after an evolution time $\tau = 0.1J^{-1}$ when optimized using (i) BDA, (ii) first-order COLD, and (iii) second-order COLD and corner optimization. (c) Final fidelities (plots show $1 - F$) of the GHZ state for the five-spin configuration depicted in (a)(ii) for an optimized global drive (red crosses) and locally driven corner spins (blue rings). (d) Final fidelities at driving time $\tau = 0.1J^{-1}$ for different system sizes N. In the global case we use ten total optimizable parameters with $N_k = 1$ drive and $N_m = 10$ time intervals, while in the corners case there are 30 total parameters as we increase to $N_k = 3$ separate drives. The plotted fidelities are the best results of five optimizations for each data point.

interval t_m the kth control drive strength is calculated as

$$f(\mathcal{B}^{k,m}, t_m) = \mathcal{B}^{k,m} \tanh(\kappa\vartheta(t_m)) \tanh(-\kappa\vartheta(t_m - \tau)) \quad (46)$$

with $\vartheta(t) = \sin \pi t/2\tau$ and $\kappa = 30$ an offset parameter used to control the shape of the drive. We use spline interpolation to calculate the derivatives of the control drive when they are required to obtain the LCD. The resulting function requires more parameters than the Fourier basis we chose to use in previous examples; however, it also allows for more flexibility in the final shape of the drive. Furthermore, due the increased number of parameters and search space, instead of Powell optimization as in previous examples we choose to instead implement dual annealing, which is a global optimizer and, while computationally more costly, is far better in the case of a complex parameter space with multiple minima.

Since such a preparation of GHZ states involves the generation of entanglement in a system that initially contains none, we expect that the first-order COLD may not contain the leading order of the counterdiabatic drive and thus may not be as effective. For this reason, we include second-order COLD terms as given in Eq. (26). We also explore the idea of using multiple control drives and localizing them to parts of the system. Thus we implement both a global drive that is uniform across all spins as well as a "corner" evolution, in which three different optimizable drives are used: one each for the first and last spins in the lattice as well as one for all of the remaining spins. This is depicted in Fig. 7(a), where different vertex (spin) colors represent different control pulses.

In Fig. 7(c) we plot $(1 - F)$, with F the fidelity of state preparation for a five-spin system with control drives consisting of $N_m = 10$ time intervals. There we observe that first-order COLD is indeed not particularly effective at short driving times and does not move the system out of its initial state [see the density matrix plots in (b)], regardless of whether or not separate control is applied to the corner spins. This is very likely due to the fact that the local σ^y terms are only a small contribution to the full counterdiabatic drive and thus we need to look to higher-order LCD to see any improvements. This is exactly what the

results indicate, as second-order COLD shows a fivefold improvement over the first order when a global optimizable drive is applied and up to 2 orders-of-magnitude improvement when the corner spins are driven separately at short times ($\tau = 0.001J^{-1}$). We then run the optimizations for larger systems at time $\tau = 0.1J^{-1}$ and find that this advantage is retained even with increasing system size.

This is a big improvement over recent results in digitized adiabatic evolution with LCD [25], where optimization was used to determine optimal coefficients for second-order LCD in order to prepare a GHZ state on an Ising spin chain. At ten spins the final state fidelity for $\tau = 1J^{-1}$ obtained in Ref. [25] was 0.18, while COLD can reach a fidelity of 0.72 for 15 spins when using corner optimization at $\tau = 0.1J^{-1}$.

This example shows that COLD can be used to speed up protocols that generate entanglement and is further evidence for the benefits of combining COLD different optimal control methods such as GRAPE as well as optimization algorithms like dual annealing.

VII. MINIMIZATION OF HIGHER-ORDER LCD TERMS

As alluded to in Sec. IV, the results plotted in Fig. 3 indicate that in optimizing the control pulse through the parameters β we maximize the largest amplitude of the first-order LCD and simultaneously reduce the second-order LCD. In the Ising spin chain case this corresponds to increasing the largest amplitude of $\alpha(\lambda, \beta)$ in Eq. (26) throughout the evolution while reducing the maximum amplitude of both $\gamma(\lambda, \beta)$ and $\zeta(\lambda, \beta)$. These results are a further indication that the implementation of COLD through the minimization of the second-order corrections discussed in Sec. III may be fruitful in more complex and/or larger systems, where the dynamics cannot be calculated.

We thus investigate replacing the original cost function of Eq. (2) with one that depends (a) explicitly on the maximum amplitude of the second-order drives $\gamma(\lambda, \beta)$ and $\zeta(\lambda, \beta)$ and (b) one that depends on the total power for either drive. Given that the LCD are functions of β, one can imagine that if there is indeed a relationship between minimizing a higher-order LCD and how effective the lower-order LCD is in producing the target state as a result, then we can determine parameters of the control drive that lead to a better final state fidelity.

We take the Ising Hamiltonian from Eq. (23) and supplement it again with the parameterized control pulse from Eq. (28). We once again take our first-order LCD to be of the form $\alpha(\lambda, \beta) \sum_j \sigma_j^y$ and the second-order drives to be $\gamma(\lambda, \beta) \sum_j (\sigma_j^x \sigma_{j+1}^y + \sigma_j^y \sigma_{j+1}^x)$ and $\zeta(\lambda, \beta) \sum_j (\sigma_j^z \sigma_{j+1}^y + \sigma_j^y \sigma_{j+1}^z)$. In Fig. 8(a) we show the results when the cost function used to optimize the parameters β is the integral from Eq. (22) that captures the total power of the drive:

$$C(\beta) = \int_0^\tau dt' |\zeta(\lambda(t'), \beta)| = I_2(\zeta(\lambda, \beta)). \tag{47}$$

In Fig. 8(c) we instead choose to minimize the largest amplitude of the drive reached throughout the evolution:

$$C(\beta) = \max_{t' \in [0,\tau]_\beta} (|\zeta(\lambda(t'), \beta)|). \tag{48}$$

In both cases we plot the resulting final state fidelities for different evolution times τ and compare them to those obtained earlier in Fig. 2(a) for five spins. The results are surprising in that while optimizing for fidelity, as was done previously, outperforms second-order minimization in both the integral and amplitude cases at most times, there is a stretch of driving times around $\tau \in [0.05, 0.5]$ where second-order minimization does better. This can be attributed to the fact that the parameter landscape for the new cost functions is completely different and allows for a more optimal value of β to be reached without being lost in a suboptimal minimum during the optimization.

In Fig. 8(b) we plot the final state fidelities at evolution time $\tau = 0.1J^{-1}$ for up to 50 spins in order to check how this type of optimization scales with system size and to compare the performance of both cost functions. We find that minimizing one of the two second-order LCDs while driving with the other still leads to impressive fidelities, but not as good as those where only first-order COLD is used. Indeed, we do not have any reason to expect an absolute optimum fidelity when using this method; however, the results in Fig. 8 are very encouraging.

While the new cost functions in Eqs. (47) and (48) may seem like a roundabout way to get to the same result—a better final state fidelity in shorter time—they have several particularly important advantages over the cost function given by Eq. (2). First and foremost, this approach does not require access to the wave function or experimental data at any point of the optimization process. In optimizing for final state fidelity directly we must compute the evolution of the system many times over in order to extract the fidelity at each iteration, but computing the drive integrals or their amplitudes is completely independent of the state of the system. This allows us to determine an optimal set of parameters β for an arbitrary system size extremely efficiently when compared to methods that require access to $|\psi_f\rangle$. A single optimization in their case, depending on the method used and the desired quality of the final outcome, may take hours or even days for larger system sizes. The new method allows us to perform an optimization with good results within minutes regardless of the number of spins, only requiring the wave function in order to

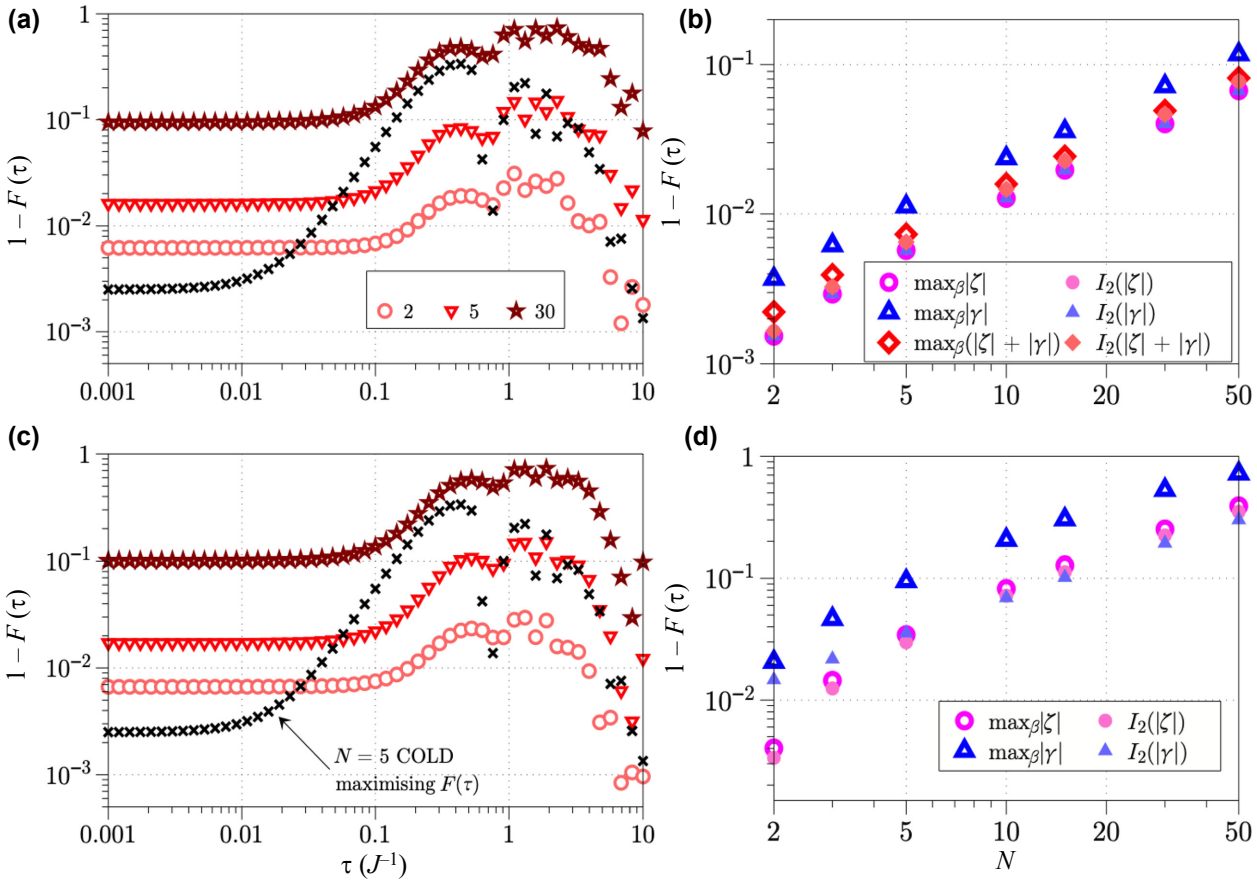

FIG. 8. Optimization of β via minimizing second-order LCD terms for the Ising model (all plots show $1-F$). In (a) we plot the final state fidelities after minimizing the integral from Eq. (22) of the $|\zeta(t)|$ drive $[I_2(|\zeta(t)|)]$ and then apply the result at each driving time to different system sizes N. The results of minimizing the maximum amplitude $\max_\beta|\zeta(t)|$ instead are plotted in (c). The black crosses in both (a) and (c) are the results of optimizing β by maximizing the final state fidelity $F(\tau)$ and are the same as the red circle plot in Fig. 2(a). Plotted in (b) and (d) are final state fidelities for $\tau = 0.1J^{-1}$ for different system sizes N when optimizing β either by minimizing the maximum amplitude of a drive $\max_\beta(\cdot)$ or its integral $I_2(\cdot)$. In (b) only first-order COLD is applied postoptimization, while in (d) one of the second-order drives is also applied after minimizing the other, e.g., if $\max_\beta|\zeta|$ is minimized to determine the optimal β then both the first-order drive α and the other second-order drive γ are applied. For system sizes above $N = 10$, we use ITensor [60] matrix product state calculations that are converged with a truncation level of 10^{-14} per time step and at each site reaching a maximum bond dimension of 4. In all cases, a single optimizable parameter is used ($N_k = 1$).

check the resulting fidelity after the optimization is finished. This is a very useful tool given that most optimal control methods demand access to the wave function while sacrificing efficiency.

It is not obvious that such a relationship between lower- and higher-order COLD as well as the fidelity of the final state must exist. In fact, this may be a fruitful new research direction to explore, combining the results obtained in this work along with, e.g., the methods in Ref. [46], where an approximate gauge potential can be systematically built up as a series of nested commutators. This might be a way to determine which operator ansatz O_{LCD} has a maximal amplitude for each driven Hamiltonian and lead to a systematic optimization of control pulses without ever having to simulate the system evolution. There is clearly a lot of new territory to explore both in terms of optimal control and in understanding the adiabatic gauge potential a little better.

VIII. DISCUSSION AND OUTLOOK

We have introduced a new hybrid approach combining quantum optimal control and shortcuts to adiabaticity: COLD. Inspired by the successes of LCD, where diabatic transitions are suppressed and locality conditions can be met, COLD improves on its methodology by combining it with quantum optimal control. The natural way to enhance the performance of LCD is by introducing higher-order CD terms, but these are often nonlocal and difficult to engineer in experiments. COLD circumvents

this by allowing for additional control fields that extend the family of dynamical Hamiltonians that can be explored. In this way, our method may find the best possible path where the effect of lower-order LCD is most relevant and higher-order corrections are suppressed.

COLD has a clear potential in efficiently speeding up adiabatic evolution in various settings. We demonstrate this numerically via several example protocols that indicate improvements beyond a classical optimization approach as well as LCD of different orders. Our work shows that COLD reduces the strength of higher-order LCD corrections, and that it performs well for increasing system sizes. We have shown that COLD can be implemented in the limit of fast driving by a "switching off" of the original dynamical Hamiltonian. For scenarios where removing the Hamiltonian is not possible, we have shown that an alternative way to implement COLD is to use a bounded optimization where amplitudes are restricted. We find that both the COLD and COLDCRAB protocols perform extremely well in this setting.

COLD will be most beneficial when the LCD is only realizable to a certain order but the higher-order corrections are large. This means that the diabatic transitions are not being sufficiently suppressed by the choice of LCD and COLD can be used to find the dynamical Hamiltonian for which the required order of the LCD term dominates. Note that this goes the other way too, with COLD not providing substantial improvements when the chosen lower-order LCD is small across the path. This can be thought of as being the case in two limits. First is the adiabatic limit, for which any CD correction is small and COLD will tend towards the adiabatic result. Second, the low-order LCD terms can be small compared to the driving as the exact CD would be correcting transitions due to interactions at longer ranges. In this scenario, the order of LCD being implemented with COLD needs to be increased, so that the CD term is accounting for the longer range terms. We show this in Sec. VI where it is clear that the generation of correlations or entanglement requires the suppression of diabatic terms that are nonlocal and thus first-order COLD cannot achieve a notable speed-up. In this case, higher-order corrections would need to be implemented with COLD, and finding methods for executing these nonlocal terms will be beneficial in these scenarios.

A further option is to combine COLD with one of a large variety of numerical optimal control methods, as we have done for the example of CRAB and GRAPE. We have shown a substantial improvement for state preparation in the Ising model that can be obtained from the COLDCRAB combination—particularly in the constrained case. Fusions of COLD with advanced optimal control methods for complex systems could prove even more fruitful with further study.

Another finding of our work is that COLD can be applied to more complex systems where exact dynamics are not possible, e.g., due to an excessively large Hilbert space. This may be achieved by variationally minimizing the integrals and maximum amplitude of the driving coefficients for the higher-order corrections to the LCD. This opens up a brand new research direction as it allows for the possibility to optimize the system's path without requiring access to the system's wave function or any sort of experimental resource. Note that this finding is more general than COLD itself, as it can even be used to optimize protocols that do not implement LCD terms, i.e., the menagerie of control procedures currently in use, providing a cost function that does not scale with the system size. This would be implemented by minimizing the highest orders of LCD in order to find a path that allows for the least diabatic losses. The data for this manuscript is available in open access at [62].

ACKNOWLEDGMENTS

Work at the University of Strathclyde was supported by the EPSRC Quantum Technologies Hub for Quantum Computing and Simulation (EP/T001062/1), and the European Union's Horizon 2020 research and innovation program under Grant Agreement No. 817482 PASQuanS. A.P. acknowledges support from the NSF under Grant No. DMR-2103658 and the AFOSR under Grants No. FA9550-16-1-0334 and No. FA9550-21-1-0342.

APPENDIX A: DERIVATION OF LOCAL COUNTERDIABATIC DRIVING TERMS FOR THE ISING MODEL

We consider here the derivation of the coupled set of equations to be solved for the second-order LCD of the Ising model; from this, it is possible to reach all terms quoted in the main text for the examples considered. We consider a finite size chain of size N. We take the Hamiltonian to be of the general form

$$H = -J\sum_{j=1}^{N-1}\sigma_j^z\sigma_{j+1}^z + Z\sum_{j=1}^{N}\sigma_j^z + X\sum_{j=1}^{N}\sigma_j^x, \qquad \text{(A1)}$$

where we consider each coefficient to be homogeneous across the chain and dependent upon the scaling factor of λ that is itself time dependent as noted in the main text. We take the second-order ansatz of LCD to be that given by Eq. (26). We then want to obtain G_λ as given by Eq. (6), which requires utilization of standard commutation rules and the commutation relations of the Pauli matrices. Following several pages of working, the following form of G_λ can be obtained:

$$
\begin{aligned}
G_\lambda = &-(\dot{J} + 4X\zeta)\sum_{j=1}^{N-1}\sigma_j^z\sigma_{j+1}^z + (\dot{Z} + 2X\alpha)\sum_{j=1}^{N}\sigma_j^z + (\dot{X} - 2\alpha Z + 4J\zeta)\sum_{j=1}^{N}\sigma_j^x \\
&+ 4J\zeta\sum_{j=1}^{N-2}\sigma_j^z\sigma_{j+1}^x\sigma_{j+2}^z + (2J\alpha + 2X\gamma - 2Z\zeta)\sum_{j=1}^{N-1}(\sigma_j^x\sigma_{j+1}^z + \sigma_j^z\sigma_{j+1}^x) \\
&+ 4(Z\gamma - X\zeta)\sum_{j=1}^{N-1}\sigma_j^y\sigma_{j+1}^y - 4Z\gamma\sum_{j=1}^{N-1}\sigma_j^x\sigma_{j+1}^x \\
&+ 2J\gamma\sum_{j=1}^{N-2}(\sigma_j^x\sigma_{j+1}^z\sigma_{j+2}^z + \sigma_j^z\sigma_{j+1}^z\sigma_{j+2}^x + \sigma_j^z\sigma_{j+1}^y\sigma_{j+2}^y + \sigma_j^y\sigma_{j+1}^y\sigma_{j+2}^z).
\end{aligned}
\tag{A2}
$$

Note that the three spin terms would trivially go to zero for the two-spin example considered in the main text. As Pauli operators are traceless, we can easily compute the Hilbert-Schmidt norm of G_λ and we simply need to keep track of factors from the finite size of the lattice to get

$$
\begin{aligned}
2^{-N}\mathrm{Tr}(G_\lambda^2) = &\,(N-1)(\dot{J} + 4X\zeta)^2 + N(\dot{Z} + 2X\alpha)^2 + N(\dot{X} - 2\alpha Z + 4J\zeta) \\
&+ 2(N-1)(2J\alpha + 2X\gamma - 2Z\zeta)^2 + 16(N-2)J^2\gamma^2 \\
&+ 16(N-1)(Z\gamma - X\zeta)^2 + 16(N-1)Z^2\gamma^2 + 16(N-2)J^2\zeta^2,
\end{aligned}
\tag{A3}
$$

where the factor on the left-hand side comes from the size of the Hilbert space. To find the system of equations to be solved, we need to minimize $\mathrm{Tr}(G_\lambda^2)$ with respect to α, γ, and ζ to obtain

$$
\begin{bmatrix}
2(X^2 + Z^2 + 2(1-1/N)J^2) & -4(1-1/N) & 8(1-1/N) \\
-JX & X^2 + 2(1-1/(N-1))J^2 + 4Z^2 & -3ZX \\
4JZ & -6XZ & 2(4X^2 + (4-3/(N-1))J^2 + Z^2)
\end{bmatrix}
\begin{bmatrix}\alpha\\ \gamma\\ \zeta\end{bmatrix}
= \begin{bmatrix} ZX - XZ \\ 0 \\ \dot{J}X - X\dot{J} \end{bmatrix}.
\tag{A4}
$$

If only the first-order correction of α is needed then this can be obtained by taking the first equation and setting γ and ζ to zero. From this, the two-spin and Ising model first-order corrections can be obtained. Note that in the limit of periodic boundary conditions or an infinite system we can take $N \to \infty$ to obtain the correct coefficients. We find that the coefficients that are proportional to system size only have a significant impact when the system is very small, e.g., in the two-spin case, and, therefore, have little impact on the results of the Ising model with $N \geq 5$.

APPENDIX B: OPTIMIZATION DISTRIBUTIONS

The results presented in Figs. 2, 5, and 6 of the main text contain plots of the best (highest) fidelities from a number of optimizations in each instance. Multiple optimization runs with different initial guesses for the optimizable parameters are included to avoid pitfalls such as local minima in the parameter landscape.

In the context of a physical implementation of one of these protocols, the optimal set of parameter values (ones that return the highest fidelity with respect to the target state) matter more than the average. However, in practice, these optimizations can be very computationally costly, in particular for larger system sizes and higher numbers of parameters. This means that we need to understand the behavior of the average and the worst case as they relate to the computational resources required.

If the parameter landscape is smooth and few local minima exist then only a few optimizations are needed to determine the best values of the optimizable parameters. However, this is never a guarantee and particularly in the

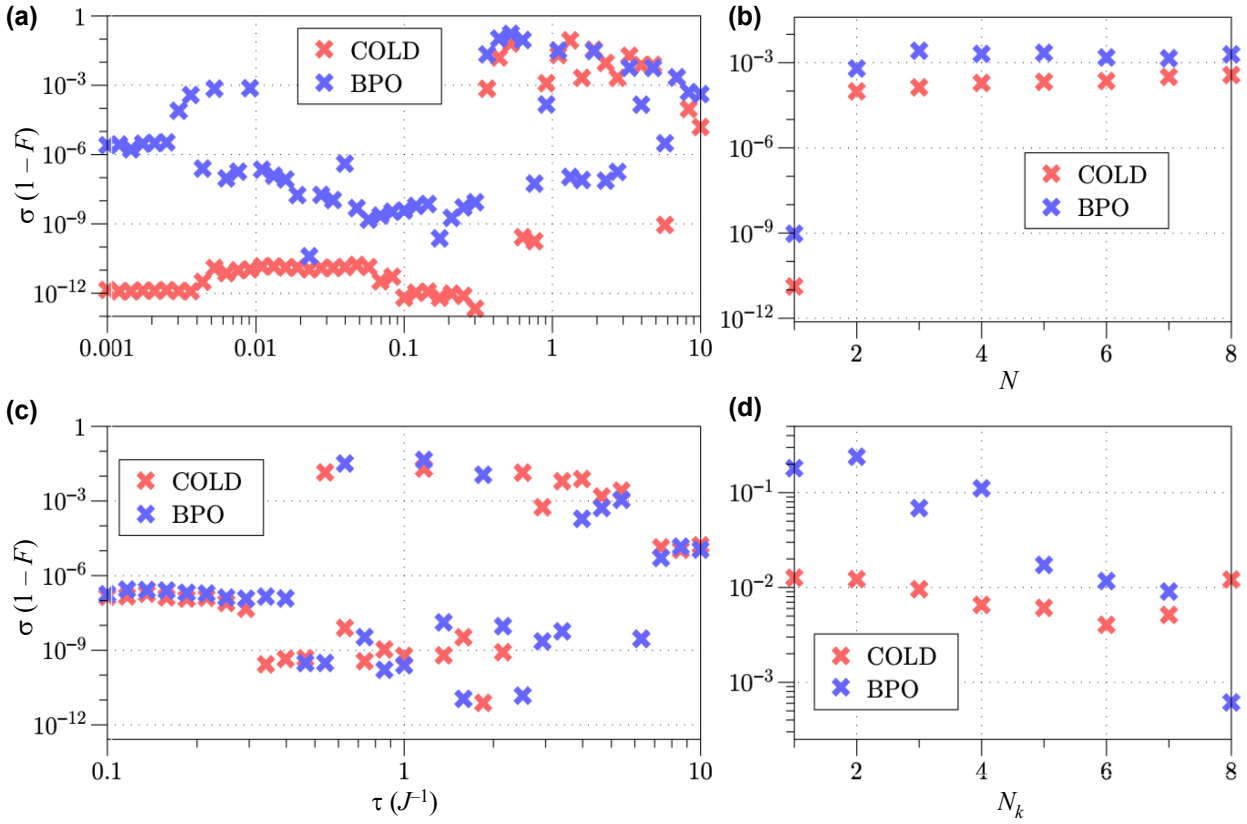

FIG. 9. Plot of the standard deviations from the mean for fidelities after 500 optimizations in the case of COLD and BPO for the Ising spin chain as discussed in Sec. IV. Panel (a) depicts the unconstrained case while (b) shows the constrained case. These correspond to the best result plots in Figs. 2(a) and 5(a), respectively. The plot in (c) gives the standard deviation for increasing lengths of the chain for spin number N while (d) shows the same for an increasing number of parameters N_k. As in Fig. 4, both (b) and (c) are plotted for driving time $\tau = 10^{-2}J^{-1}$. In all plots, results for COLD are depicted with red crosses while those for BPO are depicted with blue crosses.

case of the CRAB protocol (along with COLDCRAB), the behavior of the optimization is suboptimal when it comes to the number of optimizations needed to determine the parameter values that return the best fidelity of the target state. This is due to the fact that we *modify* the parame-ter landscape for every optimization by randomly changing the frequency components in the control field. While this allows each optimization to access a new solution space and thus increases the chances of converging to a more optimal form of the control field, it also increases variance in optimization outcomes. Since we cannot know which frequency gives the best results *a priori*, the only way to really reap the benefits of CRAB and COLDCRAB is to perform as many optimizations as possible.

This can be readily seen when we look at the standard deviation in the final fidelities over all optimizations. Figure 9 depicts these for the Ising spin chain of Sec. IV, both in the unconstrained and constrained cases as well as for varying the number of spins N and parameters N_k. We can see that in most cases for COLD the standard deviation of the fidelities stays below 10^{-3} barring longer driving times in the unconstrained case in Fig. 9(a) as well as some in Fig. 9(b) for the constrained case. BPO generally displays slightly higher standard deviations, but neither shows very significant variations in the results postoptimization. Note that the small variation in fidelity for an increasing number of parameters in Fig. 9(d) gives further evidence for the fact that additional parameters do not improve the results of COLD or BPO in the case of the Ising chain.

When it comes to CRAB and COLDCRAB, however, the picture is quite different. We find that the resulting fidelities are a lot more varied across optimizations, as would be expected given the additional component of randomness. Figure 10 shows not only the mean fidelities across optimizations but also the interquartile range of the data and the maximum and minimum values for each driving time. We find that across optimizations we are just as likely—and in some cases far more likely—to get a much worse final fidelity as we are to get a better one. This is reflected in the large range between the maximum (worst)

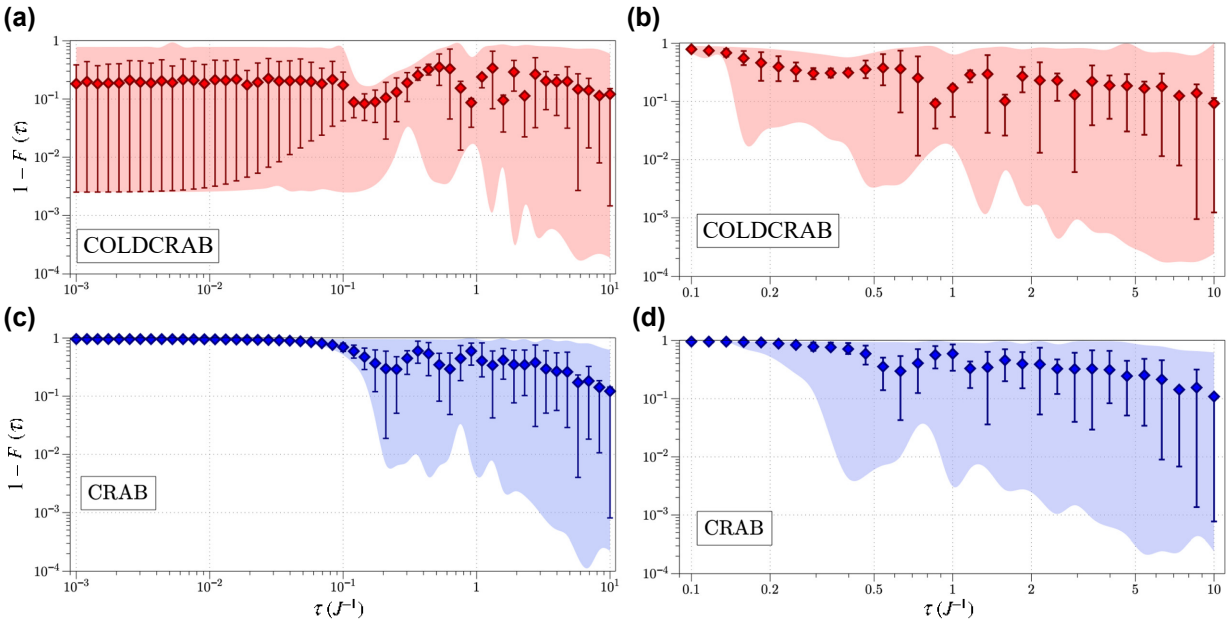

FIG. 10. Plots of the mean fidelities (diamonds) obtained over 500 optimization runs for the Ising spin chain as discussed in the main text. The error bars represent the interquartile range of the data while the shaded region encompasses the minimum and maximum fidelities obtained at each driving time. Panel (a) shows the case of COLDCRAB for the constrained instance, (b) plots the same for CRAB with (c) and (d) showing results for the unconstrained Ising chain case.

and minimum (best) fidelities for both methods as well as the interquartile range, which shows that the mean fidelity is a result of a large variation between large and small fidelities rather than a convergence to some in-between value.

These results are useful in an assessment of computational resources for such optimizations as well as giving an insight into the range of possible outcomes, particularly when implementing more unpredictable optimal control methods like CRAB.

APPENDIX C: CHOICE OF LCD ANSATZ

In determining the optimal choice of operator basis O_{LCD}, we turn back to Eq. (5) and note that it gives us some clues about the form of the LCD. Firstly, we note that if we know nothing about the system other than, that it is a spin chain described by Pauli matrices, then we take the first-order LCD to be all one-body terms while the second order can be two-body terms and so on. This is a natural choice due to the locality of the terms but also with respect to their practical implementation in an experiment. Given these considerations, it makes sense that, for a system of spins, the first-order LCD is a set of local σ^y terms.

To illustrate, in the case of the Ising spin chain case, we know that all wave functions have real coefficients, so we know that the exact CD is given by entirely imaginary terms. We can confirm this by attempting to use local σ^x or σ^z terms as our ansatz for the operator basis O_{LCD} and find that their coefficients are equal to 0 throughout the driving time. In the case of ansatz $O_{\mathrm{LCD}} = \alpha_x \sum_j \sigma_j^x$ we find that

$$G_{\lambda,\alpha_x} = \dot{J}\sum_j^{N-1}\sigma_j^z\sigma_{j+1}^z + \dot{X}\sum_j^N\sigma_j^x + \dot{Z}\sum_j^N\sigma_j^z + 2\alpha_x J\sum_j^{N-1}\sigma_j^y\sigma_{j+1}^z + 2\alpha_x J\sum_j^{N-1}\sigma_j^z\sigma_{j+1}^y + 2\alpha_x Z\sum_j^N\sigma_j^y, \tag{C1}$$

according to Eq. (6), Then the action, as in Eq. (8), is

$$\begin{aligned} S(A_\lambda) &= 2^{-N}\mathrm{Tr}[G_{\lambda,\alpha_x}(A_\lambda)^2] \\ &= \left(1-\tfrac{1}{N}\right)\dot{J}^2 + \dot{X}^2 + \dot{Z}^2 + \left(1-\tfrac{1}{N}\right)8\alpha_x^2 J^2 + 4\alpha_x^2 Z^2, \end{aligned} \tag{C2}$$

which, when minimized with respect to α_x, gives $\alpha_x = 0$.

The same procedure can be done for $O_{\mathrm{LCD}} = \alpha_z \sum_j \sigma_j^z$, i.e.,

$$G_{\lambda,\alpha_z} = \dot{J}\sum_j^{N-1}\sigma_j^z\sigma_{j+1}^z + \dot{X}\sum_j^N\sigma_j^x + \dot{Z}\sum_j^N\sigma_j^z - 2\alpha_z X\sum_j^N\sigma_j^y, \tag{C3}$$

where again we take the action

$$S(A_\lambda) = 2^{-N}\,\mathrm{Tr}[G_{\lambda,\alpha_x}(A_\lambda)^2] = 1 - \frac{1}{N}\,\dot{J}^2 + \dot{X}^2 + \dot{Z}^2 + 4\alpha_z^2 X^2, \qquad \text{(C4)}$$

which, when minimized with respect to α_z, once again gives $\alpha_z = 0$.

[1] A. Acín, I. Bloch, H. Buhrman, T. Calarco, C. Eichler, J. Eisert, D. Esteve, N. Gisin, S. J. Glaser, and F. Jelezko, *et al.*, The quantum technologies roadmap: A European community view, New J. Phys. **20**, 080201 (2018).

[2] D. E. Kirk, *Optimal Control Theory: An Introduction* (Dover Publications, New York, 2004), p. 3.

[3] S. J. Glaser, U. Boscain, T. Calarco, C. P. Koch, W. Köckenberger, R. Kosloff, I. Kuprov, B. Luy, S. Schirmer, and T. Schulte-Herbrüggen, *et al.*, Training Schrödinger's cat: Quantum optimal control, Eur. Phys. J. D **69**, 1 (2015).

[4] D. d'Alessandro, *Introduction to Quantum Control and Dynamics* (CRC Press, Boca Raton, 2007).

[5] E. Torrontegui, S. Ibáñez, S. Martínez-Garaot, M. Modugno, A. del Campo, D. Guéry-Odelin, A. Ruschhaupt, X. Chen, and J. G. Muga, in *Advances in Atomic, Molecular, and Optical Physics*, Adv. Atom. Mol. Opt. Phy., Vol. 62, edited by E. Arimondo, P. R. Berman, and C. C. Lin (Academic Press, 2013), p. 117.

[6] D. Guéry-Odelin, A. Ruschhaupt, A. Kiely, E. Torrontegui, S. Martínez-Garaot, and J. G. Muga, Shortcuts to adiabaticity: Concepts, methods, and applications, Rev. Mod. Phys. **91**, 045001 (2019).

[7] M. Demirplak and S. A. Rice, Adiabatic population transfer with control fields, J. Phys. Chem. A **107**, 9937 (2003).

[8] M. Demirplak and S. A. Rice, Assisted adiabatic passage revisited, J. Phys. Chem. B **109**, 6838 (2005).

[9] M. V. Berry, Transitionless quantum driving, J. Phys. A: Math. Theor. **42**, 365303 (2009).

[10] D. Stefanatos, J. Ruths, and J.-S. Li, Frictionless atom cooling in harmonic traps: A time-optimal approach, Phys. Rev. A **82**, 063422 (2010).

[11] D. Stefanatos and E. Paspalakis, A shortcut tour of quantum control methods for modern quantum technologies, EPL **132**, 60001 (2021).

[12] Q. Zhang, X. Chen, and D. Guéry-Odelin, Connection between inverse engineering and optimal control in shortcuts to adiabaticity, Entropy **23**, 84 (2021).

[13] F. Petiziol, B. Dive, F. Mintert, and S. Wimberger, Fast adiabatic evolution by oscillating initial Hamiltonians, Phys. Rev. A **98**, 043436 (2018).

[14] F. Petiziol, B. Dive, S. Carretta, R. Mannella, F. Mintert, and S. Wimberger, Accelerating adiabatic protocols for entangling two qubits in circuit QED, Phys. Rev. A **99**, 042315 (2019).

[15] M. Bukov, A. G. R. Day, D. Sels, P. Weinberg, A. Polkovnikov, and P. Mehta, Reinforcement Learning in Different Phases of Quantum Control, Phys. Rev. X **8**, 031086 (2018).

[16] J. Yao, L. Lin, and M. Bukov, Reinforcement Learning for Many-Body Ground-State Preparation Inspired by Counterdiabatic Driving, Phys. Rev. X **11**, 031070 (2021).

[17] I. Khait, J. Carrasquilla, and D. Segal, Optimal control of quantum thermal machines using machine learning, Phys. Rev. Res. **4**, L012029 (2022).

[18] Y. Ban, X. Chen, E. Torrontegui, E. Solano, and J. Casanova, Speeding up quantum perceptron via shortcuts to adiabaticity, Sci. Rep. **11**, 1 (2021).

[19] H. Saberi, T. c. v. Opatrný, K. Mølmer, and A. del Campo, Adiabatic tracking of quantum many-body dynamics, Phys. Rev. A **90**, 060301 (2014).

[20] S. Campbell, G. De Chiara, M. Paternostro, G. M. Palma, and R. Fazio, Shortcut to Adiabaticity in the Lipkin-Meshkov-Glick Model, Phys. Rev. Lett. **114**, 177206 (2015).

[21] D. Sels and A. Polkovnikov, Minimizing irreversible losses in quantum systems by local counterdiabatic driving, PNAS **114**, E3909 (2017).

[22] C. Whitty, A. Kiely, and A. Ruschhaupt, Quantum control via enhanced shortcuts to adiabaticity, Phys. Rev. Res. **2**, 023360 (2020).

[23] J. Wurtz and P. J. Love, Counterdiabaticity and the quantum approximate optimization algorithm, Quantum **6**, 635 (2022).

[24] P. Chandarana, N. N. Hegade, K. Paul, F. Albarrán-Arriagada, E. Solano, A. del Campo, and X. Chen, Digitized-counterdiabatic quantum approximate optimization algorithm, Phys. Rev. Res. **4**, 013141 (2022).

[25] D. Sun, P. Chandarana, Z.-H. Xin, and X. Chen, Optimizing counterdiabaticity by variational quantum circuits, Phil. Trans. R. Soc. A **380**, 20210282 (2022).

[26] M. Kolodrubetz, D. Sels, P. Mehta, and A. Polkovnikov, Geometry and non-adiabatic response in quantum and classical systems, Phys. Rep. **697**, 1 (2017).

[27] N. O. Gjonbalaj, D. K. Campbell, and A. Polkovnikov, Counter-diabatic driving in the classical β-Fermi-Pasta-Ulam-Tsingou chain, (2021), ArXiv:2112.02422.

[28] T. H. Kyaw and L.-C. Kwek, Cluster state generation in one-dimensional Kitaev honeycomb model via shortcut to adiabaticity, New J. Phys. **20**, 045007 (2018).

[29] A. del Campo, M. M. Rams, and W. H. Zurek, Assisted Finite-Rate Adiabatic Passage Across a Quantum Critical Point: Exact Solution for the Quantum Ising Model, Phys. Rev. Lett. **109**, 115703 (2012).

[30] A. del Campo, Shortcuts to Adiabaticity by Counterdiabatic Driving, Phys. Rev. Lett. **111**, 100502 (2013).

[31] C. P. Koch, Controlling open quantum systems: Tools, achievements, and limitations, J. Phys.-Condens. Mat. **28**, 213001 (2016).

[32] M. J. Powell, An efficient method for finding the minimum of a function of several variables without calculating derivatives, Comp. J. **7**, 155 (1964).

[33] Y. Xiang, D. Sun, W. Fan, and X. Gong, Generalized simulated annealing algorithm and its application to the Thomson model, Phys. Lett. A **233**, 216 (1997).

[34] P. Virtanen, *et al.*, SciPy 1.0 Contributors, SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python, Nat. Meth. **17**, 261 (2020).

[35] T. Caneva, T. Calarco, and S. Montangero, Chopped random-basis quantum optimization, Phys. Rev. A **84**, 022326 (2011).
[36] M. M. Müller, R. S. Said, F. Jelezko, T. Calarco, and S. Montangero, One decade of quantum optimal control in the chopped random basis, (2021), ArXiv:2104.07687.
[37] C. Brif, R. Chakrabarti, and H. Rabitz, Control of quantum phenomena: Past, present and future, New J. Phys. **12**, 075008 (2010).
[38] N. Rach, M. M. Müller, T. Calarco, and S. Montangero, Dressing the chopped-random-basis optimization: A bandwidth-limited access to the trap-free landscape, Phys. Rev. A **92**, 062343 (2015).
[39] R. Heck, O. Vuculescu, J. J. Sørensen, J. Zoller, M. G. Andreasen, M. G. Bason, P. Ejlertsen, O. Elíasson, P. Haikka, and J. S. Laustsen, *et al.*, Remote optimization of an ultracold atoms experiment by experts and citizen scientists, PNAS **115**, E11231 (2018).
[40] E. Farhi, J. Goldstone, S. Gutmann, and D. Nagaj, How to make the quantum adiabatic algorithm fail, Int. J. Quantum Inf. **06**, 503 (2008).
[41] M. Tomka, T. Souza, S. Rosenberg, and A. Polkovnikov, Geodesic paths for quantum many-body systems, (2016), ArXiv:1606.05890.
[42] D. A. Sivak and G. E. Crooks, Thermodynamic Metrics and Optimal Paths, Phys. Rev. Lett. **108**, 190602 (2012).
[43] E. J. Meier, K. Ngan, D. Sels, and B. Gadway, Counterdiabatic control of transport in a synthetic tight-binding lattice, Phys. Rev. Res. **2**, 043201 (2020).
[44] H. Zhou, Y. Ji, X. Nie, X. Yang, X. Chen, J. Bian, and X. Peng, Experimental Realization of Shortcuts to Adiabaticity in a Nonintegrable Spin Chain by Local Counterdiabatic Driving, Phys. Rev. Appl. **13**, 044059 (2020).
[45] N. N. Hegade, K. Paul, Y. Ding, M. Sanz, F. Albarrán-Arriagada, E. Solano, and X. Chen, Shortcuts to Adiabaticity in Digitized Adiabatic Quantum Computing, Phys. Rev. Appl. **15**, 024038 (2021).
[46] P. W. Claeys, M. Pandey, D. Sels, and A. Polkovnikov, Floquet-Engineering Counterdiabatic Protocols in Quantum Many-Body Systems, Phys. Rev. Lett. **123**, 090602 (2019).
[47] L. Prielinger, A. Hartmann, Y. Yamashiro, K. Nishimura, W. Lechner, and H. Nishimori, Diabatic quantum annealing by counter-diabatic driving, (2020), ArXiv:2011.02691.
[48] T. Caneva, T. Calarco, R. Fazio, G. E. Santoro, and S. Montangero, Speeding up critical system dynamics through optimized evolution, Phys. Rev. A **84**, 012312 (2011).
[49] M. Murphy, S. Montangero, V. Giovannetti, and T. Calarco, Communication at the quantum speed limit along a spin chain, Phys. Rev. A **82**, 022318 (2010).
[50] N. Lang and H. P. Büchler, Topological networks for quantum communication between distant qubits, npj Quantum Inf. **3**, 1 (2017).
[51] Y. Hatsugai, Chern Number and Edge States in the Integer Quantum Hall Effect, Phys. Rev. Lett. **71**, 3697 (1993).
[52] D. Hügel and B. Paredes, Chiral ladders and the edges of quantum Hall insulators, Phys. Rev. A **89**, 023619 (2014).
[53] C. W. Duncan, P. Öhberg, and M. Valiente, Exact edge, bulk, and bound states of finite topological systems, Phys. Rev. B **97**, 195439 (2018).
[54] R. Landig, L. Hruby, N. Dogra, M. Landini, R. Mottl, T. Donner, and T. Esslinger, Quantum phases from competing short-and long-range interactions in an optical lattice, Nature **532**, 476 (2016).
[55] T. Keller, S. B. Jäger, and G. Morigi, Phases of cold atoms interacting via photon-mediated long-range forces, J. Stat. Mech.-Theory E **2017**, 064002 (2017).
[56] M. Baranov, Ł. Dobrek, K. Góral, L. Santos, and M. Lewenstein, Ultracold dipolar gases–A challenge for experiments and theory, Phys. Scr. **2002**, 74 (2002).
[57] C. Menotti, M. Lewenstein, T. Lahaye, and T. Pfau, in *AIP Conference Proceedings*, Vol. 970 (American Institute of Physics, 2008), p. 332.
[58] C. Trefzger, C. Menotti, B. Capogrosso-Sansone, and M. Lewenstein, Ultracold dipolar gases in optical lattices, J. Phys. B - At. Mol. Opt. **44**, 193001 (2011).
[59] D. M. Greenberger, M. A. Horne, A. Shimony, and A. Zeilinger, Bell's theorem without inequalities, Am. J. Phys. **58**, 1131 (1990).
[60] M. Fishman, S. R. White, and E. M. Stoudenmire, The ITensor software library for tensor network calculations (2020).
[61] N. Khaneja, T. Reiss, C. Kehlet, T. Schulte-Herbrüggen, and S. J. Glaser, Optimal control of coupled spin dynamics: Design of NMR pulse sequences by gradient ascent algorithms, J. Magn. Reson. **172**, 296 (2005).
[62] https://doi.org/10.15129/d060ba54-88b8-4fa2-8214-41ae83b60b9f.