

Application of Pontryagin's minimum principle to Grover's quantum search problem

Chungwei Lin,^{1,*} Yebin Wang,¹ Grigory Kolesov,^{1,2} and Uroš Kalabić¹

¹*Mitsubishi Electric Research Laboratories, 201 Broadway, Cambridge, Massachusetts 02139, USA*

²*John A. Paulson School of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts 02138, USA*



(Received 22 February 2019; published 21 August 2019)

Grover's algorithm is one of the most famous algorithms which explicitly demonstrates how the quantum nature can be utilized to accelerate the searching process. In this work, Grover's quantum search problem is mapped to a time-optimal control problem. Resorting to Pontryagin's minimum principle, we find that the time-optimal solution has the bang-singular-bang structure. This structure can be derived naturally, without integrating the differential equations, using the geometric control technique where Hamiltonians in Schrödinger's equation are represented as vector fields. In view of optimal control, Grover's algorithm uses the bang-bang protocol to approximate the optimal protocol with a minimized number of bang-to-bang switchings to reduce the query complexity. Our work provides a concrete example of how Pontryagin's minimum principle is connected to quantum computation, and offers insight into how a quantum algorithm can be designed.

DOI: [10.1103/PhysRevA.100.022327](https://doi.org/10.1103/PhysRevA.100.022327)

I. INTRODUCTION

Quantum computation deliberately uses quantum-mechanical phenomena, such as superposition and entanglement, to reduce the computation time or the number of queries to accomplish certain tasks [1,2]. Well known quantum algorithms include Shor's algorithm for factoring [3] and Grover's algorithm for searching an unstructured database or an unordered list [4,5]. While the standard paradigm for quantum computation involves a *discrete* sequence of unitary logic gates [3–9], there exists another paradigm, pioneered by Farhi and Gutmann [10], where the quantum register evolves under some designed Hamiltonian which can vary *continuously* in time [10–13]. The concept of “continuous-time” quantum computation explicitly allows established physics principles such as the adiabatic theorem [14–16] and the Trotter-Suzuki decomposition [17,18] to guide how quantum algorithms can be designed. The adiabatic theorem is, for example, the foundation of the quantum annealing technique [11,19–22], and fast, nonadiabatic evolution is found to be helpful for other problems [23,24]. More recently, there are quantum algorithms based on the variational principle, notably the variational quantum eigensolver (VQE) [25–27] and the quantum approximate optimization algorithm (QAOA) [28–30]. They are more fault tolerant than quantum algorithms of the standard paradigm and are promising for noisy intermediate-scale quantum (NISQ) technology [31].

Generally, when applying quantum algorithms to solve a classical NP (nondeterministic polynomial-time) problem, we are given a quantum “problem Hamiltonian” (oracle) whose ground state is the solution of the original classical problem [21,32–34]. Designing a quantum algorithm is equivalent to find a “driving Hamiltonian” and an initial state, both easily

implemented, that can steer the initial state to the target state (e.g., ground state of the problem Hamiltonian) within the shortest time. From this point of view, time-optimal control [35–37] appears to be fundamentally connected to quantum computation, as both address (i) if the target state can be reached and (ii) how to reach the target state in the shortest time. Recently, Pontryagin's minimum principle (PMP) [38] has been applied to quantum state preparation [39] and nonadiabatic quantum computation [40]. Due to the linearity of Schrödinger's equation, time-optimal control generally has the bang-bang form, i.e., the control takes its extreme values. Indeed, Ref. [39] shows that the bang-bang protocol takes the minimum time to drive a two-qubit product state to the fully entangled state; Ref. [40] establishes the connection between the bang-bang protocol and the QAOA algorithm by solving spin-glass problems; Ref. [41] demonstrates that pulses of bang-bang type take the minimum time to cancel the second-order noise. In this paper, we formulate Grover's problem as a time-optimal control problem and apply PMP to solve it. By recasting the problem to involve only observable dynamical variables, we are able to show that the time-optimal control is of bang-singular-bang type without explicitly solving the necessary optimality conditions derived from PMP. In the language of control theory, Grover's algorithm can be regarded as a protocol that approximates the optimal control by the bang-bang control with a minimum number of switchings. Our analysis directly applies to a system effectively involving a single qubit [39,42,43], and may provide insight into the problems of higher dimensions [44,45].

The rest of the paper is organized as follows. In Sec. II we formulate the problem and state the necessary conditions for a time-optimal solution derived from PMP. In Sec. III we show that the bang-singular-bang structure satisfies the necessary conditions, and we compare this solution to other protocols including the pure singular control and the pure bang-bang control which turns out to be Grover's algorithm. Advantages of numerically applying PMP to problems of

*clin@merl.com

higher dimensions are discussed. In Sec. IV we reformulate the problem in terms of geometric control, and show that the time-optimal bang-singular-bang control can be naturally derived in this formalism. The relation to quantum state preparation is pointed out. Section V is the conclusion. In the Appendix, a few detailed steps regarding the geometric control technique are provided.

II. PROBLEM STATEMENT AND PONTRYAGIN'S MINIMUM PRINCIPLE

In this section we formulate Grover's problem in terms of two Hamiltonians and summarize the relevant results from PMP. Because both quantum mechanics and PMP use the term "Hamiltonian," we shall use "Hamiltonian" (symbol H) in the quantum-mechanical sense, i.e., it is a matrix that governs the dynamics of the wave function; we use "c-Hamiltonian" (symbol \mathcal{H}) to represent the control Hamiltonian, a scalar function defined in control theory.

A. Problem statement

Following Refs. [10,13], Grover's problem can be formulated using Hamiltonians. Given the "problem Hamiltonian" $H_w = |w\rangle\langle w|$ and the "driving Hamiltonian" $H_s = |s\rangle\langle s|$, where $\langle s|w\rangle \equiv x < 1$ [x and $|s\rangle$ are assumed to be real for now, and $|s\rangle$ for Grover's problem is given in Eq. (4)], we want to find the time-optimal protocol $u(t)$ that brings the initial state $|\psi_i\rangle = |s\rangle$ to the target state $|\psi_{\text{target}}\rangle = |w\rangle$ for the system evolved under a time-dependent Hamiltonian

$$\begin{aligned} H(t; u) &= \frac{1}{2}(H_w + H_s) + u(t) \frac{1}{2}(H_w - H_s) \\ &\equiv H_0 + u(t)H_d, \quad \text{with } |u(t)| \leq 1. \end{aligned} \quad (1)$$

The control $u(t)$ of the problem is assumed to be bounded by $|u| \leq 1$ [46]. In this paper, the terms "control," "protocol," and "algorithm" are used interchangeably; a given control/protocol/algorithm corresponds to a specific $u(t)$. Because all nontarget states are *degenerate* in energy, we can express both Hamiltonians using two orthogonal states $|w\rangle, |\bar{w}\rangle = [|s\rangle - x|w\rangle]/\sqrt{1-x^2}$. The initial state is given by $|\psi_i\rangle = |s\rangle = x|w\rangle + \sqrt{1-x^2}|\bar{w}\rangle$. In terms of Pauli matrices defined as

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix},$$

the Hamiltonians in Eq. (1) are

$$H_w = \frac{1}{2}(e + \sigma_z), \quad (2a)$$

$$H_s = \frac{1}{2}e + \frac{1}{2}[(2x\sqrt{1-x^2})\sigma_x + (2x^2 - 1)\sigma_z] \quad (2b)$$

$$H_0 = \frac{1}{2}e + \frac{x}{2}[\sqrt{1-x^2}\sigma_x + x\sigma_z], \quad (2c)$$

$$H_d = -\frac{x}{2}\sqrt{1-x^2}\sigma_x + \frac{1}{2}(1-x^2)\sigma_z, \quad (2d)$$

$$|\psi_i\rangle = \begin{bmatrix} x \\ \sqrt{1-x^2} \end{bmatrix}, \quad |\psi_{\text{target}}\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}. \quad (2e)$$

Note that the target state is allowed to have an arbitrary phase. For later discussions, we further define

$$\begin{aligned} H_X &= H_0 - H_d = H_s, \\ H_Y &= H_0 + H_d = H_w \end{aligned} \quad (3)$$

i.e., H_X corresponds to $u = -1$ whereas H_Y corresponds to $u = +1$. The initial state is chosen to be

$$|s\rangle = \frac{1}{\sqrt{N}} \sum_k |k\rangle, \quad (4)$$

with N the dimension of the search space (Hilbert space) spanned by $\{|k\rangle\}$. As Hamiltonians given in Eqs. (2) are just 2×2 matrices, the problem dimension N is encoded in the overlap $x = 1/\sqrt{N}$. Throughout this paper, x (different from \mathbf{x}) exclusively represents the overlap between the initial and the target state $\langle s|w\rangle$. The two-dimensional wave function will be referred to as a "qubit" state.

B. Time-optimal control problem and Pontryagin's minimum principle

The necessary conditions for a time-optimal solution derived from PMP are discussed in this subsection. We focus on the control-affine control system, where the dynamics of its "state variables" \mathbf{x} are described by

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}) + u(t)\mathbf{g}(\mathbf{x}), \quad \mathbf{x} \in R^n, u \in R. \quad (5)$$

To differentiate from the "quantum state," \mathbf{x} will be referred to as "dynamical variables"; \mathbf{f} and \mathbf{g} are smooth vector fields which are functions of \mathbf{x} . A control-affine system is directly related to Schrödinger's equation. In the language of differential geometry, \mathbf{x} defines a manifold; \mathbf{f} and \mathbf{g} belong to the tangent space of that manifold. We assume the admissible range of u is bounded by $|u| \leq 1$. The time-optimal control problem is defined as follows: given Eq. (5), find the optimal control $u(t)$ to minimize the cost function

$$J = \lambda_0 \int_0^{t_f} dt + \mathcal{C}(\mathbf{x}(t_f)), \quad (6)$$

where t_f is the final time, λ_0 is a positive constant, and $\mathcal{C}(\mathbf{x}(t_f))$ is a terminal cost function depending only on the values of the dynamical variables at t_f . To make sense of a "time-optimal" solution, λ_0 has to be positive. We only consider the time-invariant problem where \mathbf{f} , \mathbf{g} , and \mathcal{C} do not depend explicitly on time t .

Following PMP [35], a control Hamiltonian is defined as

$$\begin{aligned} \tilde{H}_c(t) &= \lambda_0 + \langle \boldsymbol{\lambda}(t), \mathbf{f}(\mathbf{x}) \rangle + u(t)\langle \boldsymbol{\lambda}(t), \mathbf{g}(\mathbf{x}) \rangle \\ &\equiv \lambda_0 + \langle \boldsymbol{\lambda}(t), \mathbf{f}(\mathbf{x}) \rangle + u(t)\Phi(t) \\ &\equiv \lambda_0 + \mathcal{H}_c(t). \end{aligned} \quad (7)$$

$\boldsymbol{\lambda}$ is referred to as a set of "costate" variables (or the conjugate momentum), which has the same dimension as \mathbf{x} . $\langle \cdot, \cdot \rangle$ is the inner product introduced for two real-valued vectors. A switching function $\Phi(t)$ is defined as

$$\Phi(t) = \langle \boldsymbol{\lambda}(t), \mathbf{g}(\mathbf{x}) \rangle, \quad (8)$$

which plays the most important role in determining the structure of optimal control.

Given an optimal solution $(\mathbf{x}^*, \boldsymbol{\lambda}^*; u^*)$ to the time-optimal control problem, it has to satisfy the following necessary conditions:

$$\dot{\mathbf{x}}^*(t) = +(\nabla_{\mathbf{x}} \mathcal{H}_c), \quad \mathbf{x}^*(0) \text{ is given}, \quad (9a)$$

$$\dot{\boldsymbol{\lambda}}^*(t) = -(\nabla_{\mathbf{x}} \mathcal{H}_c)^T, \quad \boldsymbol{\lambda}^*(t_f) = \nabla_{\mathbf{x}} \mathcal{C}|_{\mathbf{x}^*(t_f)}, \quad (9b)$$

$$\bar{\mathcal{H}}_c = \lambda_0 + \mathcal{H}_c = \text{const}, \quad (9c)$$

$$u^*(t) = \begin{cases} +1 & \text{if } \Phi(t) < 0, \\ -1 & \text{if } \Phi(t) > 0, \\ \text{undetermined} & \text{if } \Phi(t) = 0. \end{cases} \quad (9d)$$

Let us elaborate on these necessary conditions. Equation (9a) is identical to the dynamics defined in Eq. (5). Equation (9b) defines the dynamics of costate variables, whose boundary condition is fixed at the final time t_f . Equation (9c) holds for the time-invariant problem. If the final time t_f is not fixed (i.e., t_f is allowed to vary to minimize \mathcal{C}), then $\bar{\mathcal{H}}_c = 0$. Because λ_0 is a positive constant, we conclude that $\mathcal{H}_c(t)$ has to be a negative constant for a time-optimal solution. In the following analysis, we shall focus on \mathcal{H}_c (instead of $\bar{\mathcal{H}}_c$), and the condition (9c) is replaced by

$$\mathcal{H}_c = \text{const} \leq 0. \quad (10)$$

Equation (9d) means that the optimal control takes the extreme values (± 1 in this case) when the switching function is nonzero. This structure is referred to as the ‘‘bang-bang’’ protocol. If $\Phi(t) = 0$ over a finite interval of time, then u is undetermined from Eq. (9d) and more analysis is needed. When $\Phi = 0$, the time-optimal control u^* may not take its extreme values and is referred to as a singular control. The behavior of $\Phi = 0$ will be analyzed in Sec. IV B to establish the structure of the time-optimal solution.

C. Application to Schrödinger's equation

In Grover's problem, the dynamics of the wave function obey Schrödinger's equation

$$i \frac{d}{dt} |\Psi(t)\rangle = [H_0 + u(t)H_d] |\Psi(t)\rangle,$$

$$\text{with } |\Psi(t)\rangle = \begin{bmatrix} \Psi_0 \\ \Psi_1 \end{bmatrix} = \begin{bmatrix} \Psi_{0,R} + i\Psi_{0,I} \\ \Psi_{1,R} + i\Psi_{1,I} \end{bmatrix}. \quad (11)$$

The initial and final states are given in Eq. (2e). To make the final state as close to $[1, 0]^T$ as possible, the terminal cost function can be chosen as

$$\mathcal{C}(\Psi(t_f)) = -\frac{1}{2} |\Psi_0(t_f)|^2. \quad (12)$$

As explicitly shown in Eq. (11), $|\Psi\rangle$ contains four real variables. As a result costate variables $\Pi_{0,R}, \Pi_{0,I}, \Pi_{1,R}, \Pi_{1,I}$ are needed. It is convenient to express the four costate variables as a two-dimensional complex vector

$$|\Pi(t)\rangle = \begin{bmatrix} \Pi_{0,R} + i\Pi_{0,I} \\ \Pi_{1,R} + i\Pi_{1,I} \end{bmatrix}. \quad (13)$$

Using the property that H_0 and H_d are real-valued, we can express the c-Hamiltonian and switching function as

$$\begin{aligned} \mathcal{H}_{Q,c} &= \text{Im}\langle \Pi(t) | [H_0 + u(t)H_d] | \Psi(t) \rangle, \\ \Phi_Q(t) &= \text{Im}\langle \Pi(t) | H_d | \Psi(t) \rangle. \end{aligned} \quad (14)$$

Applying Eq. (9b), one can derive that the dynamics of $|\Pi(t)\rangle$ are governed by the same Schrödinger's equation, with the boundary condition given at t_f [39,40]:

$$i \frac{d}{dt} |\Pi(t)\rangle = [H_0 + u(t)H_d] |\Pi(t)\rangle,$$

$$\text{with } |\Pi(t_f)\rangle = -\begin{bmatrix} \Psi_0(t_f) \\ 0 \end{bmatrix}. \quad (15)$$

The condition (9d) still holds, with the switching function computed using Eq. (14). The subscript Q in Eqs. (14) indicates ‘‘quantum’’ and will be dropped from now on.

Equations (11) and (12) allow us to do brute-force optimization to determine the optimal $u(t)$. Equations (14) and (15) allow us to check if a control $u(t)$ satisfies the necessary conditions. One may expect that the optimal control is mostly of bang-bang type, as $\Phi(t) = 0$ can only occupy a negligible region in the manifold. In Sec. III we numerically show that the optimal protocol actually has the bang-singular-bang structure.

III. OPTIMAL PROTOCOL

In this section we show that the time-optimal control has the bang-singular-bang structure. In particular, we numerically check that a solution bearing this structure satisfies all necessary conditions imposed by PMP, and explicitly show that the time-optimal solution does take a shorter time to reach the target state when compared to other known protocols.

A. Two existing protocols

We begin the discussion by introducing two existing protocols which will be compared to the optimal protocol described shortly. The first one is provided by Farhi and Gutmann [10]. By choosing $u(t) = 0$ in Eq. (1), the wave function is

$$\begin{aligned} |\Psi(t)\rangle &= e^{-i\frac{1}{2}(H_s+H_w)t} |s\rangle \\ &= e^{-i\frac{t}{2}} \begin{bmatrix} x \cos \frac{xt}{2} - i \sin \frac{xt}{2} \\ \sqrt{1-x^2} \cos \frac{xt}{2} \end{bmatrix} \equiv \begin{bmatrix} \Psi_0(t) \\ \Psi_1(t) \end{bmatrix}. \end{aligned} \quad (16)$$

As $|\Psi_0(t)|^2 = x^2 \cos^2 \frac{xt}{2} + \sin^2 \frac{xt}{2}$, we find that for total computation time $t_f = \frac{\pi}{x}$, $|\Psi_0(t_f)|^2 = 1$. This protocol will be referred as the ‘‘singular’’ control as u does not take either of its extreme values. The second protocol is the celebrated Grover's algorithm [5], where the initial state $|s\rangle$ is evolved by alternating $U_w = 2|w\rangle\langle w| - I$ and $U_s = I - 2|s\rangle\langle s|$ about $N = \frac{\pi}{4x}$ times, i.e.,

$$|\psi_f\rangle = (U_s U_w)^N |s\rangle. \quad (17)$$

In terms of Eq. (1), U_w is achieved by $e^{-i\pi H_w}$ and U_s by $e^{-i\pi H_s}$ up to an irrelevant sign. The total computation time for Grover's algorithm is therefore $(\pi + \pi)N \sim \frac{\pi^2}{2x}$. Grover's algorithm and the singular protocol have the same asymptotic x^{-1} behavior, but the latter has a smaller prefactor. The annealing algorithm proposed by Roland and Cerf [13] is slower but more robust, as it applies the adiabatic theorem concerning the (instant) gap between the ground and the first excited state, but does not explicitly use the structure that all nontarget states are degenerate. The annealing algorithm will not be discussed

here. However, we note that the optimal $u(t)$ obtained using PMP is nonadiabatic, and the procedure does not know if there is a gap at any instant of time, which is essential for adiabatic quantum computing.

B. Optimal control and analytical analysis

We claim that the time-optimal control has the bang-singular-bang structure, which will be proved using geometric control in Sec. IV. In particular, the order is $u = +1$ (bang), $u = 0$ (singular), and $u = -1$ (bang). The time periods for the initial and the final bang control are identical because the problem is unchanged if we swap the initial and target states. Let us first investigate the consequences of applying this protocol.

Assume that the time intervals of two bang controls are both t_1 , and that of the singular control is t_2 . The wave function at $t_f = 2t_1 + t_2$ is given by

$$|\Psi(t_1; t_2)\rangle = e^{-it_1 H_s} e^{-i\frac{t_2}{2}(H_s + H_w)} e^{-it_1 H_w} |s\rangle \equiv \begin{bmatrix} \Psi_0(t_1, t_2) \\ \Psi_1(t_1, t_2) \end{bmatrix}. \quad (18)$$

A straightforward calculation gives (neglecting the phase factor)

$$\begin{aligned} \Psi_1(t_1, t_2) &= \sqrt{1-x^2} \left[\cos \frac{xt_2}{2} \left(1 - 4x^2 \sin^2 \frac{t_1}{2} \right) - 2x \sin \frac{xt_2}{2} \sin t_1 \right]. \end{aligned} \quad (19)$$

Note that maximizing $|\Psi_0(t_1, t_2)|^2$ is equivalent to minimizing $|\Psi_1(t_1, t_2)|^2$. If the target state can be exactly reached, $\Psi_1(t_1, t_2) = 0$, which allows us to represent t_2 as a function of t_1 :

$$\tan \frac{xt_2}{2} = \frac{1 - 4x^2 \sin^2 \frac{t_1}{2}}{2x \sin t_1} = \frac{1 - 2x^2 + 2x^2 \cos t_1}{2x \sin t_1}. \quad (20)$$

As a result the total computation time can be written as a function of t_1 :

$$t_f(t_1) = 2t_1 + t_2(t_1). \quad (21)$$

A time-optimal solution requires $\frac{dt_f}{dt_1} = 2 + \frac{dt_2}{dt_1} = 0$, from which we get $\frac{dt_2}{dt_1} = -2$. To get the optimal t_1 , we take $\frac{d}{dt_1}$ on both sides of Eq. (20):

$$\begin{aligned} \sec^2 \left(\frac{xt_2}{2} \right) \frac{x}{2} \frac{dt_2}{dt_1} &= \frac{1}{\cos^2 \left(\frac{xt_2}{2} \right)} \frac{x}{2} \frac{dt_2}{dt_1} \\ &\Rightarrow [4x^2(x^2 - 1)] \cos^2(t_1) + [-2(2x^2 - 1)^2] \\ &\quad \times \cos(t_1) + (2x^2 - 1)^2 = 0. \end{aligned} \quad (22)$$

Equation (22) gives the optimal t_1 . Solving Eq. (22) gives $\cos(t_1) = \frac{-1+2x^2}{2x^2}$ or $\frac{-1+2x^2}{2(-1+x^2)}$. We take the positive solution which corresponds to a smaller t_1 . The time-optimal solution (t_1^*, t_2^*) is therefore

$$\cos(t_1^*) = \frac{-1+2x^2}{2(-1+x^2)}, \quad \tan \left(\frac{xt_2^*}{2} \right) = \frac{1 - 4x^2 \sin^2 \frac{t_1^*}{2}}{2x \sin t_1^*}. \quad (23)$$

Taking $x = \frac{1}{2}$, we have $\cos(t_1^*) = \frac{1}{3}$ and thus $t_1^* \approx 0.392\pi$, $t_2^* \approx 0.784\pi$, $t_f^* = 2t_1^* + t_2^* \approx 1.5673\pi$. For $x = \frac{1}{\sqrt{32}}$, $t_1^* \approx 0.339\pi$ and $t_f^* \approx 5.221\pi$. These values are consistent with our numerical simulations.

When x is small, we get $\cos t_1^* \approx \frac{1}{2} - \frac{x^2}{2}$ from which we identify $t_1^* = \frac{\pi}{3} + \delta$. Matching the lowest nonvanishing order in x , we have

$$\begin{aligned} \cos(t_1^*) &= \cos \left(\frac{\pi}{3} + \delta \right) \approx \frac{1}{2} - \delta \sin \frac{\pi}{3} = \frac{1}{2} - \frac{x^2}{2} \\ &\Rightarrow \delta = \frac{1}{\sqrt{3}}, \quad t_1^* = \frac{\pi}{3} + \frac{x^2}{\sqrt{3}}. \end{aligned} \quad (24)$$

To get the small- x expansion of t_2^* , we use $\sin \theta = \frac{2 \tan(\theta/2)}{1 + \tan^2(\theta/2)}$ to write the second equation in Eq. (23) as

$$\begin{aligned} \sin(xt_2^*) &= 2 \frac{\left(1 - 4x^2 \sin^2 \frac{t_1^*}{2} \right) (2x \sin t_1^*)}{\left(1 - 4x^2 \sin^2 \frac{t_1^*}{2} \right)^2 + (2x \sin t_1^*)^2} \\ &\approx 4x \sin t_1^* = 2x\sqrt{3}. \end{aligned} \quad (25)$$

We expand $xt_2^* \approx \pi + \delta$ to get

$$\begin{aligned} \sin(xt_2^*) &= \sin(\pi + \delta) \approx -\delta = 2x\sqrt{3} \Rightarrow \delta = -2x\sqrt{3} \\ &\Rightarrow t_2^* = \frac{\pi + \delta}{x} = \frac{\pi}{x} - 2\sqrt{3}. \end{aligned} \quad (26)$$

The total time in the small x limit is

$$2t_1^* + t_2^* = \frac{\pi}{x} + \frac{2\pi}{3} - 2\sqrt{3} \approx \frac{\pi}{x} - 0.436\pi. \quad (27)$$

Note π/x is the total time needed in the singular protocol [10]. Figure 1(a) shows the total computation time for optimal bang-singular-bang control [by evaluating Eqs. (23)], the singular control, and Grover's algorithm. For n number of qubits, the overlap $x = 2^{-n/2}$. Their asymptotic behaviors are the same; the acceleration is only noticeable when the system is small. The time difference between the singular and the optimal control, given in Fig. 1(b), is indeed consistent with Eq. (27).

C. Example of $x = 1/2$

More details for the $x = 1/2$ case are provided. In particular, we are interested in the behavior for a fixed total time t_f where the total cost function defined in Eq. (6) only has the terminal cost function. The case we will consider is $t_f = 1.3\pi < t_f^*(x = 1/2) \sim 1.57\pi$, i.e., it is a representative example where the computation time is too short to reach the target state. The optimal bang-singular-bang protocol is

$$u_{\text{opt}}(t; t_1) = \begin{cases} 1 & \text{for } 0 < t < t_1, \\ 0 & \text{for } t_1 < t < t_f - t_1, \\ -1 & \text{for } t_f - t_1 < t < t_f. \end{cases} \quad (28)$$

For a given t_f , control (28) only has one parameter t_1 which will be determined by numerically minimizing the cost

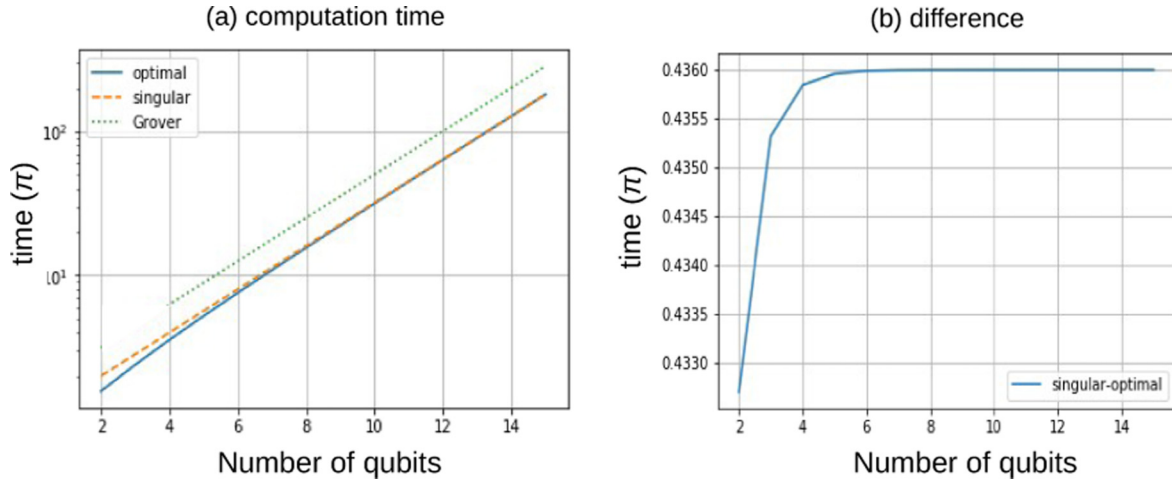


FIG. 1. (a) The total computation time for the optimal control, the singular control, and Grover’s algorithm in logarithmic scale. For n number of qubits, the overlap $x = 2^{-n/2}$. For the smaller system, the optimal control has a noticeable speedup. When the system becomes large, these two methods are almost equivalent. Grover’s algorithm has the same asymptotic behavior with a larger prefactor. (b) the difference of computation time between the optimal and the singular control in linear scale. The optimal control takes less time; the difference approaches 0.436π at large n .

function. We also consider the multiple bang-bang protocol

$$u_m(t; t_1, N) = \begin{cases} 1 & \text{for } 0 < t < t_1, \\ -1/1 & \text{alternate } N \text{ times for } t_1 < t < t_f - t_1, \\ -1 & \text{for } t_f - t_1 < t < t_f. \end{cases} \quad (29)$$

Between $t_1 < t < t_f - t_1$, the control u is alternating between -1 and 1 exactly N times, and each lasts $(t_f - 2t_1)/(2N)$. For a given t_f and N , t_1 of control (29) will also be determined by numerically minimizing the cost function.

Figure 2(a) shows the switching function Φ and the c-Hamiltonian \mathcal{H}_c [defined in Eq. (14)] using the multiple bang-bang control (29) with $N = 2$ (a total of 5 switchings). The numerical optimization gives $t_1 = 0.4446\pi$. We see that for t between about 0.3π and 0.9π , the control $u(t)$ and the

switching function $\Phi(t)$ have the same sign, which violates the necessary condition (9d). Also, the c-Hamiltonian \mathcal{H}_c is not a constant over $[0, t_f]$, which violates the necessary condition (10). Figure 2(b) shows the switching function and the c-Hamiltonian using the optimal bang-singular-bang control (28). The numerical optimization gives $t_1 = 0.3918\pi$. We see that all necessary conditions are satisfied, i.e., the control $u(t)$ and the switching function $\Phi(t)$ either have opposite signs or are both zero and the c-Hamiltonian \mathcal{H}_c is a constant over the entire $[0, t_f]$.

D. Summary and discussion on problems of higher dimensions

In this section, we have combined the analytical analysis and numerical optimization to show that the control of

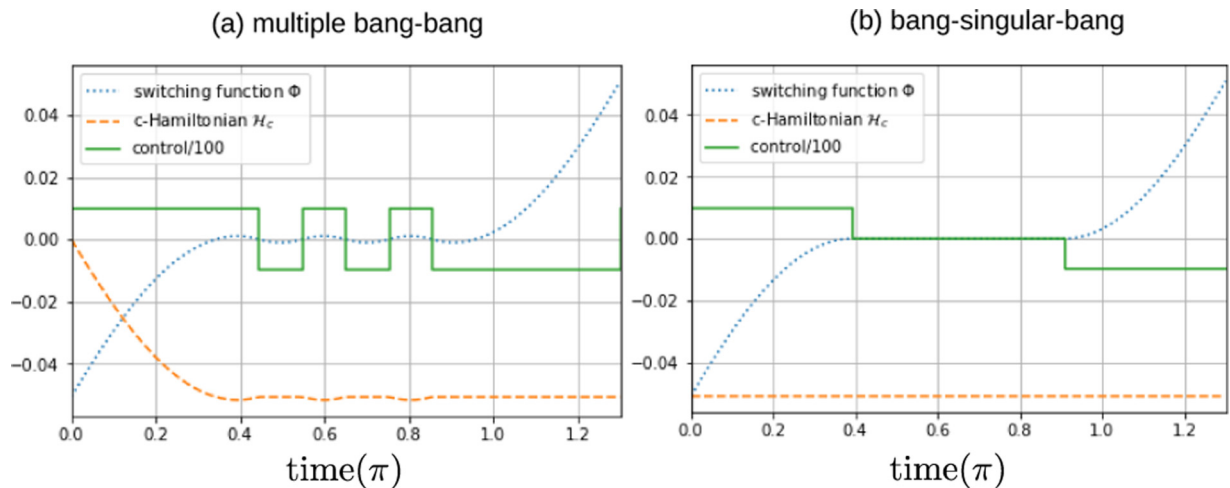


FIG. 2. The multiple bang-bang control (a) and the optimal control (b) for $x = \frac{1}{2}$ with the total computation time $t_f = 1.3\pi$, which is not long enough to reach the target state. In (a), the c-Hamiltonian (dashed curve) is not a constant over the time interval, and the switching function (dotted curve) and the control (solid curve) have the same sign, violating the necessary conditions derived from PMP. For the optimal bang-singular-bang control, all necessary conditions are satisfied.

bang-singular-bang type satisfies all necessary conditions for a time-optimal solution. The numerical optimization involves repeating the following three steps: (i) assuming a control $u(t)$; (ii) evolving the wave function from $|\Psi(t=0)\rangle = |s\rangle$ to $|\Psi(t=t_f)\rangle$ (by integrating a differential equation, a two-dimensional Schrödinger equation here); (iii) evaluating the cost function. To check the necessary conditions, we need to (iv) compute the conjugate momentum $|\Pi(t)\rangle$ backwards in time from t_f . If $u(t)$ contains many parameters, satisfying the necessary conditions during the entire computation time is practically impossible. Knowing the structure of the control is of great numerical value: it not only significantly decreases the optimization time by reducing the dimension of search domain but also finds a better solution (i.e., smaller cost function). In Sec. IV we show that the structure of optimal control can be determined even without integration for this problem.

For problems of higher dimensions, the structure of optimal control by itself can easily become prohibitively difficult to determine. Although the analytical analysis may not be possible, PMP can still be a valuable numerical tool. Specifically, by solving the Schrödinger equation two times for a given $u(t)$ —one for the wave function $|\Psi(t)\rangle$ [step (ii)] and one for the conjugate momentum $|\Pi(t)\rangle$ [step (iv)]—we are able to compute the switching function that reflects the gradient of the cost function with respect to the given $u(t)$. This is known as the “adjoint state method” [35]. The gradient obtained using the adjoint state method is exact (up to a numerical error), time efficient, and can be used in an iterative optimization algorithm to update $u(t)$.

IV. GEOMETRIC CONTROL

The Schrödinger equation (11) involves four (dependent) real variables, and directly applying PMP to a dynamical system of four variables is not immediately informative. In this section, we reduce the number of variables to two, and apply geometric control to derive the structure of the time-optimal solution. Grover’s algorithm will also be analyzed within this framework.

A. Dimension reduction and Bloch sphere representation

In view of control theory, a two-dimensional wave function has four (real-valued) dynamical variables. However, because the norm of the wave function is conserved and a global phase is not observable, only two dynamical variables are relevant. These two variables can be parametrized using two angles (θ, ϕ) by expressing a general qubit state as

$$|\Psi(\theta, \phi, \phi_0)\rangle = e^{i\phi_0} \begin{bmatrix} \cos \frac{\theta}{2} \\ \sin \frac{\theta}{2} e^{i\phi} \end{bmatrix} \rightarrow |\Psi(\theta, \phi)\rangle = \begin{bmatrix} \cos \frac{\theta}{2} \\ \sin \frac{\theta}{2} e^{i\phi} \end{bmatrix} \quad (30)$$

with $\theta \in [0, \pi]$ and $\phi \in [0, 2\pi]$. Neglecting the global phase ϕ_0 , a given qubit state corresponds to a point on a unit sphere S^2 :

$$(\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta) \Leftrightarrow |\Psi(\theta, \phi)\rangle = \begin{bmatrix} \cos \frac{\theta}{2} \\ \sin \frac{\theta}{2} e^{i\phi} \end{bmatrix}. \quad (31)$$

This is the “Bloch sphere” representation. In this convention, the first component $\cos \frac{\theta}{2}$ is always real and positive.

To investigate the dynamics in (θ, ϕ) manifold, we first need to translate the Hamiltonians in the Schrödinger equation to the corresponding vector fields. Since any 2×2 Hermitian matrix is a linear combination of three Pauli matrices and the identity matrix, we need to determine the vector fields corresponding to these matrices:

$$\begin{aligned} \sigma_z &\rightarrow V_z = 2\partial_\theta, \\ \sigma_x &\rightarrow V_x = -2\sin \phi \partial_\theta - 2\cos \phi \cot \theta \partial_\phi, \\ \sigma_y &\rightarrow V_y = 2\cos \phi \partial_\theta - 2\sin \phi \cot \theta \partial_\phi. \end{aligned} \quad (32)$$

$\{\partial_\theta, \partial_\phi\}$ form a complete basis on the tangent space of (θ, ϕ) manifold. Note that the following commutation relations hold:

$$[V_z, V_x] = -2V_y, \quad [V_y, V_z] = -2V_x, \quad [V_x, V_y] = -2V_z, \quad (33)$$

and the identity matrix that generates a global phase to the qubit wave function in the Schrödinger equation has no action on the dynamical variables (θ, ϕ) . The details of Eqs. (32) are provided in the Appendix. Using Eqs. (32), the vector fields in Eqs. (2) are

$$H_Y = H_w \rightarrow \mathbf{Y} = \frac{1}{2}V_z = \partial_\theta, \quad (34a)$$

$$\begin{aligned} H_X = H_s \rightarrow \mathbf{X} &= -2x\sqrt{1-x^2} \sin \phi \partial_\theta \\ &+ [(2x^2 - 1) - 2x\sqrt{1-x^2} \cos \phi \cot \theta] \partial_\phi, \end{aligned} \quad (34b)$$

$$\begin{aligned} H_0 = \frac{1}{2}(H_w + H_s) \rightarrow \mathbf{f} &= -x\sqrt{1-x^2} \sin \phi \partial_\theta \\ &+ [x^2 - (x\sqrt{1-x^2}) \cos \phi \cot \theta] \partial_\phi, \end{aligned} \quad (34c)$$

$$\begin{aligned} H_d = \frac{1}{2}(H_w - H_s) \rightarrow \mathbf{g} &= x\sqrt{1-x^2} \sin \phi \partial_\theta \\ &+ [(1-x^2) + x\sqrt{1-x^2} \cos \phi \cot \theta] \partial_\phi. \end{aligned} \quad (34d)$$

The control problem formulated in the Schrödinger equation [Eq. (11)] can now be recast in (θ, ϕ) :

$$\begin{aligned} \frac{d}{dt} \begin{bmatrix} \theta \\ \phi \end{bmatrix} &= \begin{bmatrix} -x\sqrt{1-x^2} \sin \phi \\ x^2 - x\sqrt{1-x^2} \cos \phi \cot \theta \end{bmatrix} \\ &+ u(t) \begin{bmatrix} x\sqrt{1-x^2} \sin \phi \\ (1-x^2) + x\sqrt{1-x^2} \cos \phi \cot \theta \end{bmatrix} \\ &\equiv \mathbf{f}(\theta, \phi) + u \cdot \mathbf{g}(\theta, \phi). \end{aligned} \quad (35)$$

The initial and target values of dynamical variables are

$$\begin{aligned} (\theta_i, \phi_i) &= \left(2 \arctan \frac{\sqrt{1-x^2}}{x}, 0 \right), \\ \theta_f &= 0, \quad \phi_f \text{ any value.} \end{aligned} \quad (36)$$

Although Eq. (35) becomes highly nonlinear [compared to Eq. (11)], it is an equation involving only two variables. A two-dimensional control-affine problem allows us to essentially enumerate all possibilities and thus determine the structure of optimal control.

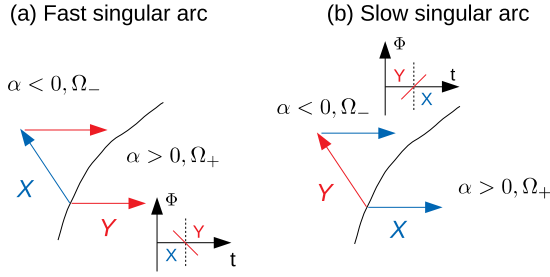


FIG. 3. The singular arc $\alpha = 0$ divides the manifold into $\alpha > 0, \Omega_+$ and $\alpha < 0, \Omega_-$ regions. (a) The fast singular arc where the singular arc represents the optimal-time trajectory. (b) The slow singular arc where the singular arc is not the optimal-time trajectory. X and Y represent the $u = -1$ and $u = +1$ bang controls respectively. The insets show the behavior of $\Phi(t) \sim 0$, based on which the allowed bang-bang controls in Ω_+ (YX) and Ω_- (XY) regions are indicated. The trajectory can be kept around $\alpha = 0$ by alternating the X and Y controls.

B. General structure of time-optimal solution for two-dimensional systems

In this subsection, we describe the general structure of the time-optimal solution in two dimension given by Sussmann [47–49]. Let us consider the following question: If $u(t)$ is an optimal solution, i.e., *all* conditions in (9) are satisfied but with a zero switching function at time t [i.e., $\Phi(\mathbf{x}(t)) = 0$], what are the possible values of the dynamical variables at a later time $t + dt$? The answer relies on $\dot{\Phi}(t)$. Assuming that \mathbf{f} and \mathbf{g} are linearly independent, their commutator can be expressed as a linear combination of \mathbf{f} and \mathbf{g} , i.e.,

$$[\mathbf{f}, \mathbf{g}] = \alpha(\mathbf{x})\mathbf{f} + \beta(\mathbf{x})\mathbf{g}, \quad (37)$$

where $\alpha(\mathbf{x})$ and $\beta(\mathbf{x})$ are smooth functions of \mathbf{x} . The time derivative of $\Phi(t)$, $\dot{\Phi}(t)$, can be computed as

$$\begin{aligned} \dot{\Phi} &= \langle \boldsymbol{\lambda}, [\mathbf{f}, \mathbf{g}](\mathbf{x}) \rangle \\ &= \alpha \langle \boldsymbol{\lambda}, \mathbf{f} \rangle + \beta \langle \boldsymbol{\lambda}, \mathbf{g} \rangle, \end{aligned} \quad (38)$$

where $\Phi = \langle \boldsymbol{\lambda}, \mathbf{g} \rangle$ is used. Because $\Phi(t) = 0$ and condition (10) indicates $\langle \boldsymbol{\lambda}, \mathbf{f} \rangle = -|C|$, we get

$$\dot{\Phi} = -\alpha|C|. \quad (39)$$

Using Eq. (9d), we conclude

$$\begin{aligned} \alpha > 0 &\Rightarrow \dot{\Phi} < 0 \quad \text{allow } u = -1 \text{ to } u = +1, XY \text{ control,} \\ \alpha < 0 &\Rightarrow \dot{\Phi} > 0 \quad \text{allow } u = +1 \text{ to } u = -1, YX \text{ control.} \end{aligned} \quad (40)$$

Here $u = -1$ is denoted as X control, $u = +1$ as Y control. When $\alpha \neq 0$, $\Phi(t) = 0$ can only be an isolated point in time and therefore the optimal control has to be of the bang-bang type. Equation (40) is illustrated in Fig. 3.

The next step is to investigate if the admissible control can keep the dynamical variables along $\alpha = 0$ that divides the manifold into $\alpha > 0, \Omega_+$ and $\alpha < 0, \Omega_-$ regions. If this happens, then $\alpha = 0$ defines a singular arc along which the optimal control can be singular [$\Phi(t) = 0$ over a finite amount of time]; otherwise, the optimal control is of the bang-bang type. To answer this question, we consider $\alpha(\mathbf{x}_0) = 0$ and compute the Lie derivative of α with respect to the vector

fields \mathbf{X} and \mathbf{Y} . Two scenarios arise:

(i) If $L_{\mathbf{X}}\alpha(\mathbf{x}_0)$ and $L_{\mathbf{Y}}\alpha(\mathbf{x}_0)$ have the same sign, then \mathbf{x}_0 will move out of $\alpha = 0$ for all admissible controls, so $\alpha = 0$ is not a singular arc.

(ii) If $L_{\mathbf{X}}\alpha(\mathbf{x}_0)$ and $L_{\mathbf{Y}}\alpha(\mathbf{x}_0)$ have opposite signs, then $\alpha = 0$ represents a singular arc. In this scenario, one can keep the dynamical variables around $\alpha = 0$ by alternating the X and Y controls.

The scenario (ii) has two possible cases. Case (a): if $L_{\mathbf{X}}\alpha(\mathbf{x}_0) < 0$, $L_{\mathbf{Y}}\alpha(\mathbf{x}_0) > 0$, then $\alpha = 0$ is the time-optimal trajectory; $\alpha = 0$ is referred to as a “fast” singular arc [see Fig. 3(a)]. In this case, the optimal control is obtained by $d\alpha = 0 = L_{\mathbf{f}}\alpha + uL_{\mathbf{g}}\alpha$:

$$u = \frac{L_{\mathbf{X}}\alpha(\mathbf{x}_0) + L_{\mathbf{Y}}\alpha(\mathbf{x}_0)}{L_{\mathbf{X}}\alpha(\mathbf{x}_0) - L_{\mathbf{Y}}\alpha(\mathbf{x}_0)}. \quad (41)$$

Case (b): if $L_{\mathbf{X}}\alpha(\mathbf{x}_0) > 0$, $L_{\mathbf{Y}}\alpha(\mathbf{x}_0) < 0$, then $\alpha = 0$ is not the time-optimal trajectory; $\alpha = 0$ is referred to as a “slow” singular arc [see Fig. 3(b)].

The reasoning follows. In case (a), \mathbf{X} drives \mathbf{x}_0 to the Ω_- region. To go back to the singular arc, we need to apply \mathbf{Y} but the XY control violates the condition (40). Therefore, to move along the singular arc, PMP implies that the singular control takes the minimum time. In case (b), \mathbf{X} takes \mathbf{x}_0 to the Ω_+ region. In Ω_+ , XY control is allowed by Eq. (40), which brings the trajectory back to the singular arc. Therefore, it is not obvious if the bang-bang control or the singular control takes the minimum time. It is shown that the bang-bang control [37] is the optimal solution. It may be a bit surprising that the optimal control can be determined without integrating the differential equation in time (only spatial derivatives are needed), and we pinpoint that the condition $\langle \boldsymbol{\lambda}, \mathbf{f} \rangle = -|C|$ leading to Eq. (39) is what makes this possible.

C. The structure of optimal solutions to Grover’s problem

The analysis provided in Sec. IV B is now applied to Eq. (35), where the two-dimensional dynamical variables are $\mathbf{x} = (\theta, \phi)$. The commutator $[\mathbf{f}, \mathbf{g}]$ is

$$\begin{aligned} [\mathbf{f}, \mathbf{g}] &= \left[-x\sqrt{1-x^2} \cot \theta \frac{1}{\sin \phi} - (1-x^2) \frac{\cos \phi}{\sin \phi} \right] \mathbf{f} \\ &\quad + \left[-x\sqrt{1-x^2} \cot \theta \frac{1}{\sin \phi} + x^2 \frac{\cos \phi}{\sin \phi} \right] \mathbf{g} \\ &\equiv \alpha(\theta, \phi)\mathbf{f} + \beta(\theta, \phi)\mathbf{g}. \end{aligned} \quad (42)$$

The curve defined by $\alpha(\theta, \phi) = -\sqrt{1-x^2} \left(x \frac{\cot \theta}{\sin \phi} + \sqrt{1-x^2} \frac{\cos \phi}{\sin \phi} \right) = 0$ is

$$-\frac{x}{\sqrt{1-x^2}} \cot \theta = \cos \phi. \quad (43)$$

This curve divides the region around $\alpha = 0$ into $\alpha > 0 (\Omega_+)$ and $\alpha < 0 (\Omega_-)$ regions. To identify these regions, we note that $\alpha(\frac{\pi}{2}, \frac{\pi}{2}) = 0$, and $\alpha(\frac{\pi}{2}, \frac{\pi}{2} + \epsilon) > 0$, $\alpha(\frac{\pi}{2}, \frac{\pi}{2} - \epsilon) < 0$ (with ϵ a small positive number).

To see if $\alpha = 0$ corresponds to a singular control, we compute the Lie derivative of $\alpha(\theta, \phi)$ with respect to \mathbf{Y} and \mathbf{X} . Using $\partial_{\theta}\alpha = \frac{x\sqrt{1-x^2}}{\sin \phi \sin^2 \theta}$ and $\cot \theta = -\frac{\sqrt{1-x^2}}{x} \cos \phi$ (the

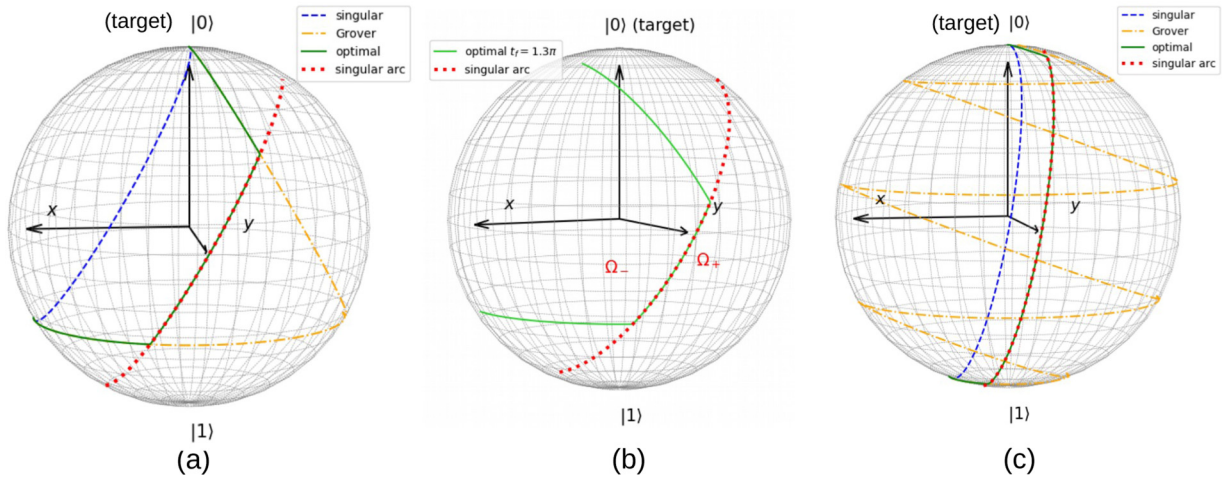


FIG. 4. Trajectories using the singular (dashed curves), the optimal (solid curves), and Grover's protocol (dash-dot curves) for (a) $x = \frac{1}{2}$ and $t_f = 1.567\pi$, (b) $x = \frac{1}{2}$ and $t_f = 1.3\pi$, and (c) $x = \frac{1}{4\sqrt{2}}$ and $t_f = 5.221\pi$. The parameters of (a) and (c) are computed from Eq. (23); those of (b) are given in Fig. 2(b). In (b), $t_f = 1.3\pi$ which is too short to reach the target state. The singular arc (dotted curves) determined by Eq. (43) is also plotted. The optimal trajectory and the singular arc coincide over a finite amount of time. The trajectory of Grover's algorithm is zigzagging through the optimal trajectory.

condition of $\alpha = 0$), we get

$$\begin{aligned} L_Y \alpha &= \partial_\phi \alpha = 1 - x^2 > 0, \\ L_X \alpha &= -(1 - x^2) = -L_Y \alpha < 0. \end{aligned} \quad (44)$$

Substituting Eq. (44) into Eq. (41), we see that $\alpha = 0$ corresponds to a singular control with $u = 0$. Figure 4 gives the trajectories using the singular, the optimal, and Grover's protocol for $x = 1/\sqrt{4}$ and $x = 1/\sqrt{32}$. Note that in Fig. 4(b) we purposely use $t_f = 1.3\pi$ which is too short to reach the target state [see Fig. 2(b)]. The singular arc determined by Eq. (43) is also plotted. The optimal trajectory does coincide with the singular arc over a finite interval of time.

D. Relation to the quantum state preparation

Although we only focus on Grover's problem, the same technique is applicable to problems which can be reduced to a single qubit, and some qualitative behaviors may provide insights into problems of higher dimensions. In this subsection we discuss two concrete examples related to the quantum state preparation.

For Grover's problem, the structure of the singular arc and the position of the initial state (see Fig. 4) naturally introduce two critical times T_1 and T_2 ($> T_1$). For the computation time $t_f < T_1$, the trajectory does not touch the singular arc and the optimal control is of bang-bang type (Y-X). T_2 is defined as the minimum time where the target state can be reached. For $T_1 < t_f < T_2$, the optimal control is of bang-singular-bang type (Y-0-X). For $t_f > T_2$, there is no unique trajectory to reach the target state. If only the bang-bang control is allowed, there will be many switchings along the singular arc. These behaviors are found in the quantum state preparation problem (for both single-qubit and multiqubit systems) studied in Refs. [44,45]. The fact that a singular control can be approximated by many bang-bang controls is reflected as the "glassy" phase (in the time domain) in Refs. [44,45].

For the second example, we consider the time-optimal control that steers a two-qubit product state to the fully entangled state, which is formulated and solved in Ref. [39]. The dimension of the two-qubit Hilbert space is 4. Using the symmetry of the Hamiltonian, the four-dimensional Hilbert space can be divided into two invariant subspaces and each subspace is effectively a single qubit. The formalism provided in Secs. II and IV A can then be applied to one of the subspaces. Without providing details, our analysis does show that the singular-bang control is the time-optimal solution, as found in Ref. [39]. One benefit of geometric control is that we are able to compute the value of the singular control [using Eq. (41)] without performing time integration.

E. Analysis of Grover's algorithm in the reduced dimension

We conclude this section by analyzing Grover's algorithm in the (θ, ϕ) manifold, where the problem is to steer the dynamical variables from $(\theta_i, \phi_i) = (2 \arctan \frac{\sqrt{1-x^2}}{x}, 0)$ to $\theta_f = 0$. We first note that, by substituting $\tan \theta_i = \frac{\sqrt{1-x^2}}{x}$ and $\phi_i = 0$ into Eq. (34a), one gets $\mathbf{X}(\theta_i, \phi_i) = 0$. This indicates that the X ($u = -1$) control cannot change the initial dynamical variables (θ_i, ϕ_i) and the bang control has to start with Y ($u = +1$). Next, we notice that only one of the bang controls, the X control, can change the value of θ . To drive θ from θ_i to 0 with a minimum number of bang-to-bang switchings, we need to maximize the reduction of θ during the X control. To obtain the maximum θ reduction, we consider the dynamics with the X control given in Eq. (34a):

$$\frac{d\theta}{dt} = -2x\sqrt{1-x^2} \sin \phi, \quad (45a)$$

$$\frac{d\phi}{dt} = (2x^2 - 1) - 2x\sqrt{1-x^2} \cos \phi \cot \theta. \quad (45b)$$

For small x , $x \cot \theta \leq 1/N$ holds for $\theta \in [Nx, \pi - Nx]$ (one can pick $N = 10$), indicating $\frac{d\phi}{dt} \approx -1$ [Eq. (45b)] over this

range. Integrating Eq. (45a) from ϕ_i to ϕ_f with $\frac{d\phi}{dt} \approx -1$, we get the change of θ as

$$\begin{aligned}\Delta\theta &= -2x\sqrt{1-x^2} \int_{\phi_i}^{\phi_f} \sin\phi \frac{dt}{d\phi} d\phi \\ &\approx 2x\sqrt{1-x^2} [\cos\phi_i - \cos\phi_f] \\ &\geq 2x\sqrt{1-x^2} [-2] \equiv -|\Delta\theta|_{\max}.\end{aligned}\quad (46)$$

The maximum θ reduction $|\Delta\theta|_{\max} = 4x\sqrt{1-x^2}$ can only be obtained when $\phi_i = \pi$ and $\phi_f = 0$. To achieve $|\Delta\theta|_{\max}$ from $\phi_i = 0$, one first (a) applies Y ($\mathbf{Y} = \partial_\phi$) control for time π that takes ϕ from 0 to π , and one then (b) applies X ($\mathbf{X} \approx -2x \sin\phi \partial_\theta - \partial_\phi$ for small x) control for time π that takes ϕ from π back to 0; it is (b) that reduces the θ by $|\Delta\theta|_{\max}$. (a) and (b) will be repeated N times to bring $\theta = \theta_i$ to $\theta = 0$. For small x , $\theta_i \approx \pi$ and N is estimated as

$$\frac{\theta_i}{|\Delta\theta|_{\max}} \approx \frac{\pi}{4x}.\quad (47)$$

This procedure is identical to Grover's algorithm described in Sec. III A. We conclude that, in terms of optimal control, Grover's algorithm uses the bang-bang protocol with a minimum number of switchings to approximate the optimal bang-singular-bang control. To visualize this point, in Fig. 4(c) we show that the trajectory of Grover's algorithm is zigzagging through the optimal trajectory. Cutting the number of switchings reduces the query complexity, which is how the quantum speedup is defined in quantum computation using a discrete sequence of gates.

V. CONCLUSION

Grover's quantum search problem is one of the most well known and important problems in quantum computation. In this paper, we formulate this problem as a time-optimal control problem, and apply Pontryagin's minimum principle to solve it. Designing a quantum algorithm is equivalent to solving a time-optimal control problem where the dynamical variables are the components of the wave function, the dynamics are the specified by Schrödinger's equation, and the target state is the ground state of the given "problem Hamiltonian." Although the dimension of the Hilbert space

of Grover's problem can be arbitrarily large, it can be reduced to a two-dimensional Schrödinger equation. We show that the optimal control for Grover's problem has the bang-singular-bang structure, and we explicitly demonstrate that the optimal protocol does take a shorter time to reach the target state when comparing to other known protocols. This bang-singular-bang can be derived without integrating the differential equations by using the geometric control technique. The key step is to derive the dynamical equations involving only two real-valued dynamical variables. Within this manifold, applying geometric control provides the optimal protocol at each point. The formalism provided in this paper can be beneficial to any quantum systems which can be effectively described by a single qubit. We are aware that the generalization to systems of higher dimensions may not be as informative, as the analysis relies on enumerating all possibilities, but it can be worth investigating. The key insight brought by time-optimal control theory is that in addition to the bang-bang protocol the singular control should be considered. However, PMP provides an efficient and exact formalism to compute the gradient of the cost function at a given control, which can be useful in numerical optimization for problems of higher dimensions. In view of optimal control, Grover's algorithm uses the bang-bang protocol, with a minimum number of bang-to-bang switchings to reduce the query complexity, to approximate the optimal protocol. Our work provides a concrete example how optimal control is connected to the quantum computation, and may shed some light on how a quantum algorithm can be designed.

ACKNOWLEDGMENTS

C.L. thanks Yanting Ma, Helena Zhang, and Andrew Millis for very helpful discussions. We thank Dries Sels for very valuable comments.

APPENDIX: OPERATOR AS A VECTOR FIELD IN (θ, ϕ)

In this Appendix, Eqs. (32) that relate the vector fields defined on the tangent space of (θ, ϕ) to the Pauli matrices are derived. Applying $H = \sigma_z$ on a state $|\Psi(\theta, \phi)\rangle$ over a time interval dt , we get

$$e^{-i\sigma_z dt} \begin{bmatrix} \cos \frac{\theta}{2} \\ \sin \frac{\theta}{2} e^{i\phi} \end{bmatrix} = \begin{bmatrix} \cos \frac{\theta}{2} e^{-i dt} \\ \sin \frac{\theta}{2} e^{i(\phi+dt)} \end{bmatrix} \rightarrow \begin{bmatrix} \cos \frac{\theta}{2} \\ \sin \frac{\theta}{2} e^{i(\phi+2 dt)} \end{bmatrix}.\quad (A1)$$

After forcing the first component to be real and positive, $e^{-i\sigma_z dt}$ only changes the ϕ component. We therefore conclude $\sigma_z \rightarrow 2\partial_\phi$.

Applying $H = \sigma_x$ on a state $|\Psi(\theta(0), \phi(0))\rangle$ over a time interval dt , we get

$$\begin{aligned}e^{-i\sigma_x dt} \begin{bmatrix} \cos \frac{\theta(0)}{2} \\ \sin \frac{\theta(0)}{2} e^{i\phi(0)} \end{bmatrix} &= \begin{bmatrix} 1 & -i dt \\ -i dt & 1 \end{bmatrix} \begin{bmatrix} \cos \frac{\theta(0)}{2} \\ \sin \frac{\theta(0)}{2} e^{i\phi(0)} \end{bmatrix} \\ &= \begin{bmatrix} \cos \frac{\theta(0)}{2} + dt \sin \frac{\theta(0)}{2} \sin \phi(0) - i dt \sin \frac{\theta(0)}{2} \cos \phi(0) \\ \sin \frac{\theta(0)}{2} e^{i\phi(0)} - i dt \cos \frac{\theta(0)}{2} \end{bmatrix} \rightarrow \begin{bmatrix} \cos \frac{\theta(dt)}{2} \\ \sin \frac{\theta(dt)}{2} e^{i\phi(dt)} \end{bmatrix}.\end{aligned}\quad (A2)$$

Applying σ_x changes both θ and ϕ components. We note that $A + i dt = A[1 + i(dt/A)] \sim Ae^{idt/A}$, so to the first order a small imaginary part does not change the amplitude. For the first component, we see that

$$\begin{aligned}\cos \frac{\theta(dt)}{2} &= \cos \frac{\theta(0)}{2} + dt \sin \frac{\theta(0)}{2} \sin \phi(0) \\ \Rightarrow \frac{d\theta}{dt} &= -2 \sin \phi.\end{aligned}\tag{A3}$$

The amplitude change of the first component determines the coefficient of ∂_θ . The first component also introduces a phase term $1 - i dt \tan \frac{\theta(0)}{2} \cos \phi(0)$ which has to be compensated. Therefore, the second component gives

$$\begin{aligned}\sin \frac{\theta(dt)}{2} e^{i\phi(dt)} &= \left[\sin \frac{\theta(0)}{2} e^{i\phi(0)} - i dt \cos \frac{\theta(0)}{2} \right] \left[1 + i dt \tan \frac{\theta(0)}{2} \cos \phi(0) \right] \\ \Rightarrow \sin \frac{\theta(dt)}{2} e^{i\phi(dt)} - \sin \frac{\theta(0)}{2} e^{i\phi(0)} &= i dt \left[\sin \frac{\theta(0)}{2} \tan \frac{\theta(0)}{2} \cos \phi(0) e^{i\phi(0)} - \cos \frac{\theta(0)}{2} \right] \\ \Rightarrow \frac{1}{2} \cos \frac{\theta}{2} \frac{d\theta}{dt} + \sin \frac{\theta}{2} i \frac{d\phi}{dt} &= i \left[\sin \frac{\theta}{2} \tan \frac{\theta}{2} \cos \phi - \cos \frac{\theta}{2} e^{-i\phi} \right].\end{aligned}\tag{A4}$$

Using $\frac{d\theta}{dt} = -2 \sin \phi$ and $e^{-i\phi} = \cos \phi - i \sin \phi$, we get

$$\begin{aligned}-\cos \frac{\theta}{2} \sin \phi + i \sin \frac{\theta}{2} \frac{d\phi}{dt} &= i \left[\sin \frac{\theta}{2} \tan \frac{\theta}{2} \cos \phi - \cos \frac{\theta}{2} (\cos \phi - i \sin \phi) \right] \\ \Rightarrow \frac{d\phi}{dt} &= \cos \phi \left[\tan \frac{\theta}{2} - \cot \frac{\theta}{2} \right] = -2 \cos \phi \cot \theta.\end{aligned}\tag{A5}$$

Combining Eq. (A3) and (A5), we get

$$\sigma_y \rightarrow -2 \sin \phi \partial_\theta - 2 \cos \phi \cot \theta \partial_\phi.\tag{A6}$$

The vector field corresponding to σ_y can be derived similarly. A straightforward calculation gives

$$n_x \sigma_x + n_z \sigma_z \rightarrow n_x V_x + n_z V_z = -2n_x \sin \phi \partial_\theta + 2[n_z - n_x \cos \phi \cot \theta] \partial_\phi,\tag{A7}$$

which is useful to map Hamiltonians appearing in Grover's problem to their corresponding vector fields.

-
- [1] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2011).
- [2] M. M. P. Kaye and R. Laflamme, *An Introduction to Quantum Computing* (Oxford University Press, New York, 2007).
- [3] P. W. Shor, Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer, *SIAM J. Comput.* **26**, 1484 (1997).
- [4] L. K. Grover, A fast quantum mechanical algorithm for database search, in *Proceedings of the Twenty-eighth Annual ACM Symposium on Theory of Computing, STOC '96, Philadelphia, May 1996* (ACM, New York, 1996), pp. 212–219.
- [5] L. K. Grover, Quantum Mechanics Helps in Searching for a Needle in a Haystack, *Phys. Rev. Lett.* **79**, 325 (1997).
- [6] D. Deutsch and R. Jozsa, Rapid solution of problems by quantum computation, *Proc. Math. Phys. Sci.* **439**, 553 (1992).
- [7] D. R. Simon, On the power of quantum computation, *SIAM J. Comput.* **26**, 1474 (1997).
- [8] C. Bennett, E. Bernstein, G. Brassard, and U. Vazirani, Strengths and weaknesses of quantum computing, *SIAM J. Comput.* **26**, 1510 (1997).
- [9] A. W. Harrow, A. Hassidim, and S. Lloyd, Quantum Algorithm for Linear Systems of Equations, *Phys. Rev. Lett.* **103**, 150502 (2009).
- [10] E. Farhi and S. Gutmann, Analog analogue of a digital quantum computation, *Phys. Rev. A* **57**, 2403 (1998).
- [11] T. Kadowaki and H. Nishimori, Quantum annealing in the transverse Ising model, *Phys. Rev. E* **58**, 5355 (1998).
- [12] E. Farhi, J. Goldstone, S. Gutmann, J. Lapan, A. Lundgren, and D. Preda, A quantum adiabatic evolution algorithm applied to random instances of an NP-complete problem, *Science* **292**, 472 (2001).
- [13] J. Roland and N. J. Cerf, Quantum search by local adiabatic evolution, *Phys. Rev. A* **65**, 042308 (2002).
- [14] E. Farhi, J. Goldstone, S. Gurmann, and M. Sipser, Quantum computation by adiabatic evolution, [arXiv:quant-ph/0001106](https://arxiv.org/abs/quant-ph/0001106).
- [15] A. T. Rezakhani, W.-J. Kuo, A. Hamma, D. A. Lidar, and P. Zanardi, Quantum Adiabatic Brachistochrone, *Phys. Rev. Lett.* **103**, 080502 (2009).
- [16] Q. Zhuang, Increase of degeneracy improves the performance of the quantum adiabatic algorithm, *Phys. Rev. A* **90**, 052317 (2014).
- [17] H. F. Trotter, On the product of semi-groups of operators, *Proc. Am. Math. Soc.* **10**, 545 (1959).
- [18] M. Suzuki, Generalized Trotter's formula and systematic approximants of exponential operators and inner derivations with applications to many-body problems, *Commun. Math. Phys.* **51**, 183 (1976).

- [19] J. Brooke, Quantum annealing of a disordered magnet, *Science* **284**, 779 (1999).
- [20] E. Santoro, G. E. Tosatti, Optimization using quantum mechanics: Quantum annealing through adiabatic evolution, *J. Phys. A* **39**, R393 (2006).
- [21] A. Das and B. K. Chakrabarti, *Colloquium*: Quantum annealing and analog quantum computation, *Rev. Mod. Phys.* **80**, 1061 (2008).
- [22] M. W. Johnson, P. Bunyk, F. Maibaum, E. Tolkacheva, A. J. Berkley, E. M. Chapple, R. Harris, J. Johansson, T. Lanting, I. Perminov, E. Ladizinsky, T. Oh, and G. Rose, A scalable control system for a superconducting adiabatic quantum optimization processor, *Supercond. Sci. Technol.* **23**, 065004 (2010).
- [23] D. S. Steiger, T. F. Rønnow, and M. Troyer, Heavy Tails in the Distribution of Time to Solution for Classical and Quantum Annealing, *Phys. Rev. Lett.* **115**, 230501 (2015).
- [24] B. Heim, T. F. Rønnow, S. V. Isakov, and M. Troyer, Quantum versus classical annealing of Ising spin glasses, *Science* **348**, 215 (2015).
- [25] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien, A variational eigenvalue solver on a photonic quantum processor, *Nat. Commun.* **5**, 4213 (2014).
- [26] P. J. J. O'Malley, R. Babbush, I. D. Kivlichan, J. Romero, J. R. McClean, R. Barends, J. Kelly, P. Roushan, A. Tranter, N. Ding, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, A. G. Fowler, E. Jeffrey, E. Lucero, A. Megrant, J. Y. Mutus, M. Neeley, C. Neill, C. Quintana, D. Sank, A. Vainsencher, J. Wenner, T. C. White, P. V. Coveney, P. J. Love, H. Neven, A. Aspuru-Guzik, and J. M. Martinis, Scalable Quantum Simulation of Molecular Energies, *Phys. Rev. X* **6**, 031007 (2016).
- [27] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, J. M. Gambetta, Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets, *Nature (London)* **549**, 242 (2017).
- [28] E. Farhi, J. Goldstone, and S. Gurmman, A quantum approximate optimization algorithm, [arXiv:1411.4028](https://arxiv.org/abs/1411.4028) (2014).
- [29] D. Wecker, M. B. Hastings, and M. Troyer, Progress towards practical quantum variational algorithms, *Phys. Rev. A* **92**, 042303 (2015).
- [30] C. Y.-Y. Lin and Y. Zhu, Performance of QAOA on typical instances of constraint satisfaction problems with bounded degree, [arXiv:1601.01744](https://arxiv.org/abs/1601.01744).
- [31] J. Preskill, Quantum Computing in the NISQ era and beyond, *Quantum* **2**, 79 (2018).
- [32] N. Xu, J. Zhu, D. Lu, X. Zhou, X. Peng, and J. Du, Quantum Factorization of 143 on a Dipolar-Coupling Nuclear Magnetic Resonance System, *Phys. Rev. Lett.* **108**, 130501 (2012).
- [33] E. Crosson, E. Farhi, C. Y.-Y. Lin, H.-H. Lin, and P. Shor, Different strategies for optimization using the quantum adiabatic algorithm, [arXiv:1401.7320](https://arxiv.org/abs/1401.7320).
- [34] D. Wecker, M. B. Hastings, and M. Troyer, Training a quantum optimizer, *Phys. Rev. A* **94**, 022309 (2016).
- [35] D. G. Luenberger, *Introduction to Dynamic Systems: Theory, Models, and Applications* (Wiley, New York, 1979).
- [36] D. Liberzon, *Calculus of Variations and Optimal Control Theory: A Concise Introduction* (Princeton University Press, Princeton, 2012).
- [37] H. Schättler and U. Ledzewicz, *Geometric Optimal Control: Theory, Methods and Examples*, Interdisciplinary Applied Mathematics, Vol. 38, 1st ed. (Springer-Verlag, New York, 2012).
- [38] L. S. Pontryagin, *Mathematical Theory of Optimal Processes* (CRC, Boca Raton, FL, 1987).
- [39] S. Bao, S. Kleer, R. Wang, and A. Rahmani, Optimal control of superconducting qubits using pontryagin's minimum principle: Preparing a maximally entangled state with singular bang-bang protocols, *Phys. Rev. A* **97**, 062343 (2018).
- [40] Z.-C. Yang, A. Rahmani, A. Shabani, H. Neven, and C. Chamon, Optimizing Variational Quantum Algorithms using Pontryagin's Minimum Principle, *Phys. Rev. X* **7**, 021027 (2017).
- [41] J. Zeng and E. Barnes, Fastest pulses that implement dynamically corrected single-qubit phase gates, *Phys. Rev. A* **98**, 012301 (2018).
- [42] G. H. Low, T. J. Yoder, and I. L. Chuang, Quantum Imaging by Coherent Enhancement, *Phys. Rev. Lett.* **114**, 100801 (2015).
- [43] G. H. Low and I. L. Chuang, Optimal Hamiltonian Simulation by Quantum Signal Processing, *Phys. Rev. Lett.* **118**, 010501 (2017).
- [44] 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).
- [45] A. G. R. Day, M. Bukov, P. Weinberg, P. Mehta, and D. Sels, Glassy Phase of Optimal Quantum Control, *Phys. Rev. Lett.* **122**, 020601 (2019).
- [46] Note that the conventional Hamiltonian is given by $H = u(t)H_w + [1 - u(t)]H_s$ with $0 < u < 1$. To be consistent with the convention in control theory, we define H_0 and H_d such that $|u| < 1$.
- [47] H. J. Sussmann, The structure of time-optimal trajectories for single-input systems in the plane: The c^∞ non-singular case, *SIAM J. Control Optim.* **25**, 433 (1987).
- [48] H. J. Sussmann, The structure of time-optimal trajectories for single-input systems in the plane: The general real analytic case, *SIAM J. Control Optim.* **25**, 868 (1987).
- [49] H. J. Sussmann, Regular synthesis for time-optimal control of single-input real analytic systems in the plane, *SIAM J. Control Optim.* **25**, 1145 (1987).