

Introduction to the Pontryagin Maximum Principle for Quantum Optimal Control

U. Boscain,^{1,*} M. Sigalotti^{1,†} and D. Sugny^{2,‡}

¹*Laboratoire Jacques-Louis Lions, CNRS, Inria, Sorbonne Université, Université de Paris, France*

²*Laboratoire Interdisciplinaire Carnot de Bourgogne (ICB), UMR 6303 CNRS-Université de Bourgogne-Franche Comté, 9 Av. A. Savary, BP 47870, DIJON Cedex 21078, France*



(Received 10 December 2020; revised 21 May 2021; published 14 September 2021)

Optimal control theory is a powerful mathematical tool, which has known a rapid development since the 1950s, mainly for engineering applications. More recently, it has become a widely used method to improve process performance in quantum technologies by means of highly efficient control of quantum dynamics. This tutorial aims at providing an introduction to key concepts of optimal control theory that is accessible to physicists and engineers working in quantum control or in related fields. The different mathematical results are introduced intuitively, before being rigorously stated. This tutorial describes modern aspects of optimal control theory, with a particular focus on the Pontryagin maximum principle, which is the main tool for determining open-loop control laws without experimental feedback. The different steps to solve an optimal control problem are discussed, before moving on to more advanced topics such as the existence of optimal solutions or the definition of the different types of extremals, namely normal, abnormal, and singular. The tutorial covers various quantum control issues and describes their mathematical formulation suitable for optimal control. The connection between the Pontryagin maximum principle and gradient-based optimization algorithms used for high-dimensional quantum systems is described. The optimal solution of different low-dimensional quantum systems is presented in detail, illustrating how the mathematical tools are applied in a practical way.

DOI: 10.1103/PRXQuantum.2.030203

CONTENTS

I. INTRODUCTION	2	A. Why Lagrange multipliers appear in constrained optimization problems	9
A. Structure of the paper	2	B. Statement of the Pontryagin maximum principle	10
II. INTRODUCTION TO THE OPTIMAL CONTROL CONCEPTS: THE CASE OF A TWO-LEVEL QUANTUM SYSTEM	3	C. Use of the PMP	12
III. FORMULATION OF THE CONTROL PROBLEM	5	VII. GRADIENT-BASED OPTIMIZATION ALGORITHM	13
IV. THE DIFFERENT STEPS TO SOLVE AN OPTIMAL CONTROL PROBLEM	7	A. The weak Pontryagin maximum principle	13
V. EXISTENCE OF SOLUTIONS FOR THE OPTIMAL CONTROL PROBLEM: THE FILIPPOV TEST	8	B. Gradient-based optimization algorithm	14
VI. FIRST-ORDER CONDITIONS	9	VIII. EXAMPLE 1: A THREE-LEVEL QUANTUM SYSTEM WITH COMPLEX CONTROLS	15
		A. Formulation of the quantum control problem	15
		B. Existence	16
		C. Application of the PMP	17
		IX. EXAMPLE 2: A MINIMUM TIME TWO-LEVEL QUANTUM SYSTEM WITH A REAL CONTROL	20
		A. Formulation of the control problem	20
		B. Application of the PMP	21
		X. APPLICATIONS OF QUANTUM OPTIMAL CONTROL THEORY	23
		XI. CONCLUSION	24
		XII. LIST OF NOTATION	25
		ACKNOWLEDGMENTS	25

*ugo.boscain@sorbonne-universite.fr

†mario.sigalotti@inria.fr

‡dominique.sugny@u-bourgogne.fr

APPENDIX A: TEST OF CONTROLLABILITY	25
APPENDIX B: FILIPPOV'S THEOREM	26
REFERENCES	27

I. INTRODUCTION

Quantum technology aims at developing practical applications based on properties of quantum mechanics [1]. This objective requires precise manipulation of quantum objects by means of external electromagnetic fields. Quantum control encompasses a set of techniques to find the time evolution of control parameters that perform specific tasks in quantum physics [2–19]. In recent years, it has naturally become a key tool in the emergent field of quantum technologies [1,2,20], with applications ranging from quantum computing [2,21,22] to quantum sensing [23] and quantum simulation [24,25].

In the majority of quantum control protocols, the control law is computed in an open-loop configuration without experimental feedback. In this context, a powerful tool is optimal control theory (OCT) [2] that allows a given process to be carried out, while minimizing a cost (e.g., the control time). This approach has key advantages. Its flexibility makes it possible to adapt to experimental constraints or limitations and its optimal character leads to the physical limits of the driven dynamics. OCT can be viewed as a generalization of the classical calculus of variations for problems with dynamical constraints [26]. Its modern version was born with the Pontryagin maximum principle (PMP) in the late 1950s [27]. Since the pioneering study of Pontryagin and coworkers, OCT has undergone rapid development and is nowadays a recognized field of mathematical research. Recent tools from differential geometry have been applied to control theory, making these methods very effective in dealing with problems of growing complexity. Many reference textbooks have been published on the subject, both on mathematical results and engineering applications [26,28–36]. Originally inspired by problems of space dynamics, OCT was then applied in a wide spectrum of applications such as robotics or economics. OCT was first used for quantum processes [37,38] in the context of physical chemistry, the goals being to steer chemical reactions [4,39–42] or to control spin dynamics in nuclear magnetic resonance [43–46]. A lot of results have recently been established for quantum technologies, such as, for example, the minimum duration to generate high-fidelity quantum gates [2].

Two types of approach based on the PMP have been used to solve optimal control problems in low- and high-dimensional systems, respectively. In the first situation, called geometric optimal control theory, the equations for optimality are solved by using geometric and analytical tools. The results can be determined analytically or at least with a very high numerical precision. The PMP allows

one to deduce the structure of the optimal solutions and, in some cases, a proof of their global optimality can be established. In this context, a series of low-dimensional quantum control problems has been rigorously solved in recent years for both closed [47–56] and open quantum systems [57–64]. Specific numerical optimization algorithms have been developed and applied to design control fields in larger quantum systems [25,45,65–67]. Because of the complexity of the control landscape, only local optimal solutions are found with this numerical optimal control approach.

However, despite the recent success of quantum optimal control theory, the situation is still not completely satisfactory. The difficulty of the concepts used in this field does not allow a nonexpert to easily understand and apply these techniques. The mathematical textbooks use a specialized and sophisticated language, which makes these works difficult to access. Very few basic papers for physicists are available in the literature, while having a minimum grasp of these tools will be an important skill in the future of quantum technologies. The purpose of this tutorial is to provide an introduction to the core mathematical concepts and tools of OCT in a rigorous but understandable way by physicists and engineers working in quantum control and in related fields. A deep analogy can be carried out between OCT and finding the minima of a real function of several variables. This parallel is used throughout the text to qualitatively describe the key aspects of the PMP. The tutorial is based on an advanced course for PhD students in physics taught at Saarland University in Spring of 2019. It assumes a basic knowledge of standard topics in quantum physics, but also of mathematical techniques such as linear algebra or differential calculus and geometry. Finally, we hope that this paper will give the reader the prerequisites to access a more specialized literature and to apply optimal control techniques to their own control problems.

A. Structure of the paper

A tutorial about optimal control is a difficult task because a large number of mathematical results have been obtained and many techniques have been developed over the years for specific applications. Among others, we can distinguish the following problem classes: finite- or infinite-dimensional systems, open- or closed-loop control, linear or nonlinear dynamical systems, geometric or numerical optimal control, the PMP or Hamilton-Jacobi-Bellmann (HJB) approach, etc. We briefly recall that the HJB method, which is the result of the dynamic programming theory, leads to necessary and sufficient conditions for optimality in which the optimal cost is the solution of a nonlinear partial differential equation [26]. Unfortunately, this equation is generally very difficult to solve numerically.

This means making choices about which topics to include in this paper. We have deliberately selected specific aspects of OCT that are treated rigorously, while others are only briefly mentioned. The choice fell on basic mathematical concepts that are the most useful in quantum control. We limit our focus to the optimal control of an open-loop finite-dimensional system by using the PMP. In particular, we consider only analytical and geometric techniques to solve low-dimensional control problems. To ensure overall consistency and limit the length of the paper, we do not discuss in depth numerical optimization methods and the infinite-dimensional case [68,69], which are also key issues in quantum control. In order to connect this tutorial with the current applications of optimal control to high-dimensional quantum systems, we describe the link between the PMP and the most current implementation of the gradient-based optimization algorithm (the GRAPE algorithm [45]). Finally, we stress that a precise knowledge of the PMP is an essential skill for numerical optimization, and that the scope of the material of this paper is much broader than the examples presented.

The paper is built on three reading levels. A first level corresponds to the main text and explains the main concepts necessary to describe and apply the theory of optimal control. Some key ideas in optimal control are first introduced qualitatively for a simple quantum system in Sec. II. In addition to the two examples solved in Secs. VIII and IX, the different notions are described rigorously and then systematically illustrated by examples. This establishes a direct link between the mathematical concept and its practical application. A second reading level is given by footnotes, which recall a mathematical definition or correspond to a more specific comment that can be skipped on a first reading. A final reading level is available in the appendices. These different paragraphs explain in detail the mathematical origin of the theorems used for the controllability and the existence problem and some standard counterexamples or specificities the reader should have in mind. We point out that these sections are not mathematical proofs of theorems, but rather a description of the formalism introduced in a language accessible to a physicist. In order to facilitate the reading of the paper, a list of the main notation used is given in Sec. XII along with the place in the text where they are first introduced.

Although the paper is targeted to a physics audience and the mathematical details are kept as simple as possible, our objective is to stick to rigorous statements and claims, since this aspect becomes crucial while implementing optimal control ideas in numerical simulations or in experiments.

The paper is organized as follows. We first introduce the main ideas used in optimal control in the case of a simple quantum system in Sec. II. We then show how to formulate an optimal control problem from a mathematical

point of view in Sec. III. Closed and open quantum systems illustrate this discussion. The different steps to solve such a problem are presented in Sec. IV by using the analogy with finding a minimum of a function of several variables. The tutorial continues with a point that is crucial, but often overlooked in quantum control studies, namely the existence of optimal solutions. We present in Sec. V some results based on the Filippov test, which is one of the most important techniques to address this question. The first-order conditions are described in Sec. VI, with specific attention given to the different types of extremals and the statement of the PMP. The connection between the PMP and gradient-based optimization algorithms is described in Sec. VII. Sections VIII and IX are dedicated to the presentation of two examples in three- and two-level quantum systems, respectively. Recent advances in the application of OCT to quantum technologies are briefly described in Sec. X, where we mention some of the current directions that are being followed for the development of these techniques. A conclusion is given in Sec. XI. Mathematical details about the controllability and existence problems are respectively postponed to Appendices A and B.

II. INTRODUCTION TO THE OPTIMAL CONTROL CONCEPTS: THE CASE OF A TWO-LEVEL QUANTUM SYSTEM

In quantum control, a general problem is to prepare a given quantum state by means of a specific time-dependent electromagnetic pulse. This leads to some questions such as which states can be achieved or which shape of control is required to realize this objective. These aspects, which are addressed rigorously in the rest of the tutorial, are first introduced qualitatively in this section.

To this end, we consider the control driving a two-level quantum system from the ground to the excited state. The system is described by a wave function $\psi(t)$ whose dynamics is governed by the Schrödinger equation

$$i\dot{\psi} = \begin{pmatrix} E_0 & \Omega(t) \\ \Omega^*(t) & E_1 \end{pmatrix} \psi,$$

where units such that $\hbar = 1$ have been chosen. The parameters E_0 and E_1 respectively denote the energies of the ground and excited states, while $\Omega(t)$ corresponds (up to a multiplicative factor) to a complex external field whose real and imaginary parts are, e.g., the components of two orthogonal linearly polarized laser fields. We consider resonant fields for which the carrier frequency ω of the laser is equal to the energy difference $E_1 - E_0$, namely,

$$\Omega(t) = u(t)e^{i(E_1 - E_0)t},$$

where the amplitude $u(t)$ represents the control and is assumed to be real. We now apply a time-dependent

change of variables corresponding to the choice of a rotating frame. The time evolution of $\tilde{\psi} = \Upsilon^{-1}\psi$, with $\Upsilon = \text{diag}(e^{-iE_0 t}, e^{-iE_1 t})$, satisfies the differential equation

$$i\dot{\tilde{\psi}} = \begin{pmatrix} 0 & u(t) \\ u(t) & 0 \end{pmatrix} \tilde{\psi}.$$

We denote by $c_1 = x_1 + iy_1$ and $c_2 = x_2 + iy_2$ the two complex coordinates of $\tilde{\psi}$ in a basis of the Hilbert space \mathbb{C}^2 , where the indices 1 and 2 respectively correspond to the ground and excited states. Since ψ is a state of norm 1 and Υ a unitary operator, we deduce that $x_1^2 + y_1^2 + x_2^2 + y_2^2 = 1$. Starting from the state $x_1 = 1$, the goal of the control is to bring the system to a target for which $x_2^2 + y_2^2 = 1$. The Schrödinger equation is equivalent to the following set of equations for the coefficients x_k and y_k :

$$\begin{aligned} \dot{x}_1 &= uy_2, & \dot{y}_1 &= -ux_2, \\ \dot{x}_2 &= uy_1, & \dot{y}_2 &= -ux_1. \end{aligned}$$

Since u is a real control, we immediately see that the first and last equations are coupled to each other and decoupled from the two others. In other words, the initial state of the dynamics is only connected to states for which $y_1 = x_2 = 0$, i.e., such that $x_1^2 + y_2^2 = 1$. The system thus evolves on a circle. For our control objective, the only interesting states therefore correspond to $y_2 = \pm 1$. It is also straightforward to verify that such target states can be reached at least with a constant control u . In control theory, this formulation of the control problem and the analysis of the reachable set from the initial state constitute a basic prerequisite before deriving a specific control procedure. This step is detailed in Sec. III.

We now explore the optimal control of this system. We first use the circular geometry of the dynamics to simplify the corresponding equations. We introduce the angle θ such that $x_1 = \cos \theta$ and $y_2 = \sin \theta$, with $\theta(0) = 1$. We arrive at

$$\dot{\theta} = -u(t),$$

where the target state is here defined as $\theta_{\text{fi}} = \pm\pi/2$. By symmetry, we can fix, without loss of generality, $\theta_{\text{fi}} = -\pi/2$. Many control solutions u exist to reach this state and a specific protocol can be selected by minimizing at the same time a functional of the state of the system and of the control, called a cost. Here, an example is given by the control time. To summarize this example, the goal of the optimal control procedure is then to find the control u steering the system to the target state in minimum time. Consider first constant controls $u(t) = u_0$ with $u_0 \in \mathbb{R}$. The duration of the process is thus $T = \pi/2u_0$. This solution reveals a key problem in optimal control that corresponds to the existence of a minimum. In this example, arbitrarily fast controls can be achieved by considering larger and

larger amplitudes u_0 and an optimal trajectory minimizing the transfer time does not exist. The analysis of the existence of optimal solutions that is a building block of any rigorous description of an optimal control problem is discussed in Sec. V. It can be shown with the results presented in Sec. V that an optimal solution exists if the set of available controls is restricted to a bounded interval, e.g., $u(t) \in [-u_m, u_m]$, where u_m is the maximum amplitude. In this case, the optimal pulse is the control of maximum amplitude, the minimum time being equal to $\pi/2u_m$.

In order to illustrate the method of solving an optimal control problem, we consider the same transfer but in a fixed time T , the goal being to minimize the energy associated with the control, i.e., the term $\int_0^T u(t)^2 dt/2$. There is no additional constraint on the control and we have $u(t) \in \mathbb{R}$. The target state is reached if $\int_0^T u(t) dt = \pi/2$. Introducing the Lagrange multiplier $\lambda \in \mathbb{R}$, this constrained optimization problem can be transformed into the minimization of the functional

$$J = \int_0^T \left(\frac{u(t)^2}{2} + \lambda u(t) \right) dt - \lambda \frac{\pi}{2}.$$

If we denote by H the function $H = -u^2/2 - \lambda u$, the Euler-Lagrange principle leads to $\partial H / \partial u = 0$, i.e., $u(t) = -\lambda$. Using the constraint on the dynamics, we finally arrive at the optimal control $u(t) = \pi/2T$.

In this simple example, the optimal solution can be derived without the complete machinery of the Pontryagin maximum principle presented below. However, in the example we have introduced the main tools used in the PMP, such as the Lagrange multiplier, the Pontryagin Hamiltonian H , and the maximization condition $\partial H / \partial u = 0$. A few comments are in order here. The dynamical constraint is quite simple since the dimension of the state space is the same as the number of controls, the dynamics can be exactly integrated, and the set of controls satisfying the constraint $\int_0^T u(t) dt = \pi/2$ is regular. This is not the case for a general nonlinear control system for which (1) the Lagrange multiplier (which is not usually constant but a function of time) is not easily found, and (2) abnormal Lagrange multipliers appear if the set of controls satisfying the constraint is not regular. We observe that H can be rewritten as

$$H = \lambda \dot{\theta} - \frac{u^2}{2},$$

which corresponds to the general formulation of the Pontryagin Hamiltonian in the normal case. The maximization condition remains the same in a general setting if there is no constraint on the available control. These aspects are discussed in detail in Sec. VI.

III. FORMULATION OF THE CONTROL PROBLEM

The dynamics. A finite-dimensional control system is a dynamical system governed by an equation of the form

$$\dot{q}(t) = f[q(t), u(t)], \quad (1)$$

where $q : I \rightarrow M$ represents the state of the system, I is an interval in \mathbb{R} , and M is a smooth manifold whose dimension is denoted by n [70]. We recall that a manifold is a space that locally (but possibly not globally) looks like \mathbb{R}^n . Manifolds appear naturally in quantum control to describe, for instance, the $(2N - 1)$ -dimensional sphere S^{2N-1} , which is the set of wave functions of a N -level quantum system. The control law is $u : I \rightarrow U \subset \mathbb{R}^m$ and f is a smooth function such that $f(\cdot, \bar{u})$ is a vector field on M for every $\bar{u} \in U$. An example of a set U of possible values of $u(t)$ is given by $U = [-1, 1]^m$, meaning that the size of each of the coordinates of u is at most one. The set U can be the entire \mathbb{R}^m if there is no control constraint.

To be sure that Eq. (1) is well posed from a mathematical viewpoint, we consider the case in which $I = [0, T]$ for some $T > 0$ and u belongs to a space of regular enough functions \mathcal{U} called the class of *admissible controls* (see Ref. [71] for a precise definition). Piecewise continuous controls form a subset of admissible controls, and in experimental implementations in quantum control they are the only control laws that can be reasonably applied. However, the class of piecewise constant controls is not suited to prove existence of optimal controls [72].

Given an admissible control $u(\cdot)$ and an initial condition $q(0) = q_{\text{in}} \in M$, there exists a unique solution $q(\cdot)$ of Eq. (1), defined at least for small times [73]. A continuous curve $q(\cdot)$ for which there exists an admissible control $u(\cdot)$ such that Eq. (1) is satisfied is said to be an *admissible trajectory*.

Let us present some typical situations encountered in quantum control.

Consider the time evolution of the wave function of a closed N -level quantum system. In this case, under the dipolar approximation [10,74,75], the dynamics is governed by the Schrödinger equation (in units where $\hbar = 1$)

$$i\dot{\psi}(t) = \left(H_0 + \sum_{j=1}^m u_j(t) H_j \right) \psi(t), \quad (2)$$

where ψ , the *wave function*, belongs to the unit sphere in \mathbb{C}^N and H_0, \dots, H_m are $N \times N$ Hermitian matrices. The control parameters $u_j(t) \in \mathbb{R}$ are the components of the control $u(\cdot)$. This control problem has the form (1) with $n = 2N - 1$, $M = S^{2N-1} \subset \mathbb{C}^N$, $q = \psi$, and $f(\psi, u) = -i(H_0 + \sum_j^m u_j H_j)\psi$. Note that the uncontrolled part corresponding to the H_0 term is called the drift. The solution

of the Schrödinger equation can also be expressed in terms of the unitary operator $\mathbf{U}(t, t_0)$, which connects the wave function at time t_0 to its value at t : $\psi(t) = \mathbf{U}(t, t_0)\psi(t_0)$. The *propagator* $\mathbf{U}(t, t_0)$ also satisfies the Schrödinger equation

$$i\dot{\mathbf{U}}(t, t_0) = \left(H_0 + \sum_{j=1}^m u_j(t) H_j \right) \mathbf{U}(t, t_0) \quad (3)$$

with initial condition $\mathbf{U}(t_0, t_0) = \mathbb{I}_N$. In quantum computing, the control problem is generally defined with respect to the propagator \mathbf{U} . Equation (3) has the form (1) with $M = U(N) \subset \mathbb{C}^{N \times N}$ and $q = \mathbf{U}$.

The wave function formalism is well adapted to describe pure states of isolated quantum systems, but when one lacks information about the system, the correct formalism is that of mixed-state quantum systems. The state of the system is then described by a density operator ρ , which is a $N \times N$ positive semidefinite Hermitian matrix of unit trace. For a closed quantum system, the density operator is a solution of the von Neumann equation

$$\dot{\rho}(t) = -i[H, \rho(t)]$$

with $H = H_0 + \sum_{j=1}^m u_j(t) H_j$. For an open N -level quantum system interacting with its environment, the dynamics of ρ is governed in some cases [76,77] by the following first-order differential equation, called the Kossakowski-Lindblad equation [78,79]:

$$\dot{\rho}(t) = -i[H, \rho(t)] + \mathcal{L}[\rho(t)]. \quad (4)$$

This equation differs from the von Neumann one in that a dissipation operator \mathcal{L} acting on the set of density operators has been added. This linear operator that describes the interaction with the environment cannot be chosen arbitrarily. Its expression can be derived from physical arguments based on a Markovian regime and a small coupling with the environment [77,80]. From a mathematical point of view, the problem of finding dynamical generators for open systems that ensure complete positivity of the dynamical evolution was solved in finite- and infinite-dimensional Hilbert spaces [78,79]. The operator \mathcal{L} is a linear operator acting on the space of density matrices that can be expressed for a N -level quantum system as

$$\mathcal{L}[\rho(t)] = \frac{1}{2} \sum_{k,k'=1}^{N^2-1} a_{kk'} ([V_k \rho(t), V_{k'}^\dagger] + [V_k, \rho(t) V_{k'}^\dagger]),$$

where the matrices V_1, \dots, V_{N^2-1} are trace zero and orthonormal. The linear mapping \mathcal{L} is completely positive if and only if the matrix $a = (a_{kk'})_{k,k'=1}^{N^2-1}$ is positive [81,82]. The density operator ρ can be represented as a vector $\vec{\rho}$ by

stacking its columns. The corresponding time evolution is generated by superoperators in the Schrödinger-like form

$$i\dot{\vec{\rho}} = \mathbf{H}\vec{\rho}. \quad (5)$$

Equation (5) has the form (1) with $M = \mathbb{B}^{N^2-1} \subset \mathbb{R}^{N^2-1}$, $q = \vec{\rho}$, and $f : \vec{\rho} \mapsto \mathbf{H}\vec{\rho}$. Here \mathbb{B}^{N^2-1} denotes the ball of radius 1 in \mathbb{R}^{N^2-1} [83].

The initial and final states. When considering a quantum control problem, the goal in most situations is not to bring the system from an initial state q_{in} to a final state q_{fi} , but rather to reach at time T a smooth *submanifold* \mathcal{T} of M (see Ref. [84] for a precise definition), called the target:

$$q(0) = q_{\text{in}}, \quad q(T) \in \mathcal{T}. \quad (6)$$

This issue arises, for instance, in the population transfer from a state ψ_{in} to an eigenstate ψ_{fi} of the field-free Hamiltonian H_0 . In this case, since the phase of the final state is not physically relevant, \mathcal{T} is characterized by $\{e^{i\theta}\psi_{\text{fi}} \mid \theta \in [0, 2\pi]\}$. It can also happen that the initial condition $q(0) = q_{\text{in}}$ is generalized to $q(0) \in \mathcal{S}$, where \mathcal{S} is a smooth submanifold of M . However, for the sake of presentation, we will not treat this case here, the changes to be made to the method being straightforward. Finally, note that the time T can be fixed or free, such as, for instance, in a time-minimum control problem.

The optimal control problem. Two different optimal control approaches can be used to steer the system from q_{in} to a target \mathcal{T} .

- Approach A: prove that the target \mathcal{T} is reachable from q_{in} (in time T if the final time is fixed or in any time otherwise) and then find the best possible control realizing the transfer. This approach requires to solve the preliminary step of controllability. Essentially, we need to show that

$$\mathcal{T} \cap \mathcal{R}(q_{\text{in}}) \neq \emptyset \quad \text{if } T \text{ is free,}$$

where $\mathcal{R}(q_{\text{in}}) := \{\bar{q} \in M \mid \exists T \text{ and an admissible trajectory } q : [0, T] \rightarrow M \text{ such that } q(0) = q_{\text{in}}, q(T) = \bar{q}\}$

or that

$$\mathcal{T} \cap \mathcal{R}^T(q_{\text{in}}) \neq \emptyset \quad \text{if } T \text{ is fixed,}$$

where $\mathcal{R}^T(q_{\text{in}}) := \{\bar{q} \in M \mid \exists \text{ an admissible trajectory } q : [0, T] \rightarrow M \text{ such that } q(0) = q_{\text{in}}, q(T) = \bar{q}\}$,

and then solve the minimization problem

$$\int_0^T f^0[q(t), u(t)] dt \rightarrow \min, \quad (7)$$

where $f^0 : M \times U \rightarrow \mathbb{R}$ is a smooth function, which in many quantum control applications depends only on the control. An example is given by the functional $\int_0^T u^2(t) dt$ that represents the energy used in the control process. The control time T is fixed or free. This integral is generally called the *cost functional*. A schematic illustration of the reachable set $\mathcal{R}^T(q_{\text{in}})$ and the target \mathcal{T} is given in Fig. 1.

The test of controllability is sometimes easy (as, for instance, for low-dimensional closed quantum systems [85,86]) and sometimes more difficult. We recall that a closed quantum system is controllable if the matrix Lie algebra generated by the matrices H_0, \dots, H_m is $\text{su}(N)$. For general systems, a useful sufficient condition for controllability is described in Appendix A. When the test of controllability can be performed, this approach is to be preferred since it permits reaching the final state exactly.

- Approach B: find a control that brings the system as close as possible to the target, while minimizing the cost. This approach is used for systems for which the controllability step cannot be easily verified. In this case, the initial point is fixed and the final point is free, but the cost contains a term [denoted $d(\cdot, \cdot)$ in the next formula] depending on the distance between the final state of the dynamics and the target:

$$\int_0^T f^0[q(t), u(t)] dt + d[\mathcal{T}, q(T)] \rightarrow \min \quad (8)$$

with T fixed or free.

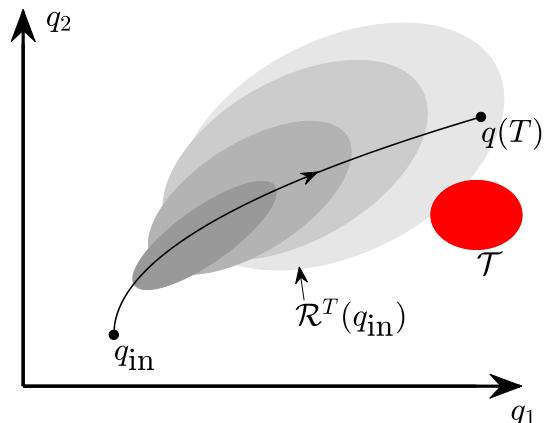


FIG. 1. The reachable set $\mathcal{R}^T(q_{\text{in}})$ from q_{in} at different times T (the shades of gray indicate an increase of the control time) and the target \mathcal{T} (red area). The manifold M is \mathbb{R}^2 and the coordinates of q are (q_1, q_2) . The intersection between the two sets $\mathcal{R}^T(q_{\text{in}})$ and \mathcal{T} is nonempty for a long enough time T . Note that the initial point of the dynamics does not belong to the reachable set at time T . This is due to a specific choice of the dynamical system with a drift.

TABLE I. Summary of the different optimal control approaches.

Approach A	$\dot{q}(t) = f[q(t), u(t)]$ $q(0) = q_{\text{in}}, q(T) \in \mathcal{T}$
When controllability can be verified, i.e., one can prove that $\mathcal{T} \cap \mathcal{R}(q_{\text{in}}) \neq \emptyset$ if T is free or $\mathcal{T} \cap \mathcal{R}^T(q_{\text{in}}) \neq \emptyset$ if T is fixed	$\int_0^T f^0[q(t), u(t)] dt \rightarrow \min$ T fixed or free
Approach B	$\dot{q}(t) = f[q(t), u(t)]$ $q(0) = q_{\text{in}}, q(T)$ free
When controllability cannot be verified	$\int_0^T f^0[q(t), u(t)] dt + d[\mathcal{T}, q(T)] \rightarrow \min$ T fixed or free

Example 1. An example is given by open quantum systems governed by the Kossakowski-Lindblad equation, for which the characterization of the reachable set is quite involved [87,88]. If we denote by ρ_{fi} the target state, a cost functional to minimize penalizing the energy of the control and the distance to the target can be

$$\int_0^T \sum_{j=1}^m \frac{u_j(t)^2}{2} dt + \|\rho(T) - \rho_{\text{fi}}\|^2,$$

where $\|\cdot\|$ is the norm corresponding to the scalar product of density matrices $\langle \rho_1 | \rho_2 \rangle = \text{Tr}[\rho_1^\dagger \rho_2]$.

The optimization problems in these two approaches should of course be considered together with the dynamics (1) and the initial and final conditions. They are summarized in Table I.

IV. THE DIFFERENT STEPS TO SOLVE AN OPTIMAL CONTROL PROBLEM

The steps to determine a solution to minimization problems (7) and (8) are similar to finding the minimum of a smooth function $f^0 : \mathbb{R} \rightarrow \mathbb{R}$.

- 1. Find conditions that guarantee the existence of solutions.** We recall that among smooth functions $f^0 : \mathbb{R} \rightarrow \mathbb{R}$, it is easy to find examples not admitting a minimum (e.g., the function $x \mapsto e^{-x}$ and the function $x \mapsto x$ do not have minima). This step is crucial. If it is skipped, first-order conditions may give a wrong candidate for optimality (see below for details) and numerical optimization schemes may either not converge or converge towards a

solution that is not a minimum. For optimal control problems, there exist several existence tests, but they are not always applicable or easy to use. In Sec. V, we present the Filippov test.

- 2. Apply first-order necessary conditions.** For a smooth function $f^0 : \mathbb{R} \rightarrow \mathbb{R}$, this means that if \bar{x} is a minimum then $df^0(\bar{x})/dx = 0$. This condition gives candidates for minima, i.e., identifies local minima, local maxima, and saddles. Note that if one does not verify *a priori* existence of minima, first-order conditions could give wrong candidates. Think, for instance, of the function $x \mapsto (x^2 + 1/2)e^{-x^2}$. This function has a single local minimum, obtained at $x = 0$, whose value is $1/2$, which is well identified by first-order conditions. However, its infimum is zero (for $x \rightarrow \pm\infty$, the function tends to zero). For optimal control problems, first-order necessary conditions should be given in an infinite-dimensional space (a space of curves) and they are expressed by the PMP, which is presented in Sec. VI B. In Approach A, note that the condition that the system exactly reaches the target is a *constraint* leading to the appearance of *Lagrange multipliers (normal and abnormal)*. This point is discussed in detail in Sec. VI A.

- 3. Apply second-order conditions.** For instance, for a smooth function $f^0 : \mathbb{R} \rightarrow \mathbb{R}$, among the points at which we have $df^0(\bar{x})/dx = 0$, a necessary condition to have a minimum is $d^2f^0(\bar{x})/dx^2 \geq 0$. This step is generally used to further reduce the candidates for optimality. For optimal control problems, there are several second-order conditions, such as higher-order Pontryagin maximum principles or Legendre-Clebsch conditions (see, for instance, Refs. [28,31,32]). In some cases, this step is difficult and it could be more convenient to go directly to the next one.

- 4. Selection of the best solution among all candidates.** Among the set of candidates for optimality identified in step 1 and (possibly) further reduced in step 2, one should select the best one. This step is often done by hand if the previous steps have identified a finite number of candidates for optimality. For optimal control problems, one often ends up with infinitely many candidates for optimality and this step is generally very difficult.

There are of course specific examples for which the solution is particularly simple. This is the case for convex problems for which only first-order conditions should be applied, since the existence step is automatic and first-order conditions are both necessary and sufficient for optimality. This situation is however rare in quantum control and we will not discuss it further.

V. EXISTENCE OF SOLUTIONS FOR THE OPTIMAL CONTROL PROBLEM: THE FILIPPOV TEST

The existence theory for optimal control is difficult and, unfortunately, there is no general procedure that can be applied in any situation. In this section, we present the most important technique, the Filippov test that allows one to tackle several types of problem. In order to keep this paragraph as accessible as possible, we present below only the main ideas and some propositions derived from the Filippov test. These results can be directly applied to quantum systems. A complete statement of the Filippov test is provided in Appendix B. We emphasize that it is fundamental to verify the existence of optimal controls before applying first-order conditions (i.e., the PMP). Otherwise, as discussed in the finite-dimensional case, it may occur that the PMP has solutions, but none of them are optimal.

Let us consider the problem in Approach A with T fixed.

Problem P1. *We have*

$$\begin{aligned} \dot{q}(t) &= f[q(t), u(t)], \\ q(0) &= q_{\text{in}}, \quad q(T) \in \mathcal{T}, \\ \int_0^T f^0[q(t), u(t)] dt &\rightarrow \min, \\ T > 0 &\text{ fixed.} \end{aligned}$$

Here $q : [0, T] \rightarrow M$, where M is a smooth n -dimensional manifold, f, f^0 are smooth functions of their arguments, $u \in \mathcal{U}$, $U \subset \mathbb{R}^m$, and \mathcal{T} is a smooth submanifold of M .

In order to tackle the existence problem, we define a new variable q^0 obtained as the value of the cost during the time evolution, that is,

$$q^0(t) = \int_0^t f^0[q(s), u(s)] ds,$$

and we let $\hat{q} = (q^0, q)$. The dynamics of the new state \hat{q} in $\mathbb{R} \times M$ is given by

$$\begin{aligned} \dot{\hat{q}}(t) &= \begin{pmatrix} \dot{q}^0(t) \\ \dot{q}(t) \end{pmatrix} = \begin{pmatrix} f^0[q(t), u(t)] \\ f[q(t), u(t)] \end{pmatrix} =: \hat{f}[q(t), u(t)], \\ \hat{q}(0) &= (0, q_{\text{in}}), \quad \hat{q}(T) \in \mathbb{R} \times \mathcal{T}. \end{aligned}$$

This control system is called the *augmented system*. The minimization problem in integral form, $\min \int_0^T f^0[q(t), u(t)] dt$, becomes a problem of minimization of one of the coordinates at the final time, i.e., $\min q^0(T)$.

We denote by $\hat{\mathcal{R}}^T(0, q_{\text{in}})$ the reachable set at time T starting from $(0, q_{\text{in}})$ for the augmented system. The key observation on which optimal control is based is expressed by the following proposition.

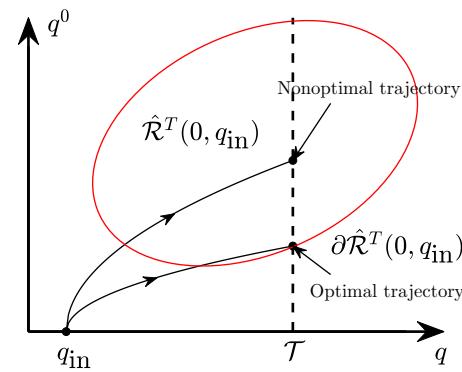


FIG. 2. The reachable set of the augmented system (area delimited by the red curve). The target T is represented by the vertical dashed line.

Proposition 1. *If $q(\cdot)$ is an optimal trajectory for Problem P1 then $\hat{q}(T) \in \partial \hat{\mathcal{R}}^T(0, q_{\text{in}})$.*

Proof. By contradiction, if $\hat{q}(T) = [q^0(T), q(T)] \in \text{int} \hat{\mathcal{R}}^T(0, q_{\text{in}})$ then there exists a trajectory reaching a point $[\alpha, q(T)]$ with $\alpha < q^0(T)$, i.e., arriving at the same point in M , but with a smaller cost. See Fig. 2. ■

When $\hat{\mathcal{R}}^T(0, q_{\text{in}}) \cap (\mathbb{R} \times \mathcal{T})$ is nonempty and compact, an optimal trajectory for Problem P1 exists [89]. We deduce the following property.

Proposition 2. *If $\hat{\mathcal{R}}^T(0, q_{\text{in}})$ is compact, T is closed, and $\hat{\mathcal{R}}^T(0, q_{\text{in}}) \cap (\mathbb{R} \times \mathcal{T})$ is nonempty, then there exists a solution to Problem P1.*

Hence, the compactness of $\hat{\mathcal{R}}^T(0, q_{\text{in}})$ is a key point. A similar reasoning allows one to relate the compactness of the reachable set within time T and the existence of solutions for the minimum time problem. A sufficient condition for compactness of the reachable set is given by Filippov's theorem, which is stated in Appendix B. A consequence of Filippov's theorem is the following (see also Propositions 11 and 12 in Appendix B).

Proposition 3. *Consider the Schrödinger equation*

$$i\dot{\psi}(t) = \left(H_0 + \sum_{j=1}^m u_j(t) H_j \right) \psi(t), \quad (9)$$

where $\psi(t)$ evolves in the unit sphere S^{2N-1} of \mathbb{C}^N , the matrices H_0, \dots, H_m are $N \times N$ Hermitian, and $u(t) \in U$. Let ψ_0 be the initial condition for $\psi(\cdot)$, and let the final target T be closed. Assume that the set U is convex and compact. Then each of the following two optimal control problems admit a solution:

1. T is fixed, $\mathcal{R}^T(q_{\text{in}})$ intersects \mathcal{T} , and the cost function $f^0 : S^{2N-1} \times U \rightarrow \mathbb{R}$ is convex;
2. minimum time problem when $\mathcal{R}(q_{\text{in}})$ intersects \mathcal{T} .

Example 2. The compactness of U is a key assumption to ensure the existence of an optimal solution. We come back to the example of Sec. II and we consider the control of a two-level quantum system whose dynamics is governed by the Hamiltonian $H = u \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ with u real valued [50]. The goal of the control is to steer the wave function from $\psi_{\text{in}} = (1, 0)$ to the target $\mathcal{T} = \{(0, e^{i\theta}) \mid \theta \in \mathbb{R}\}$ in minimum time. If $U = \mathbb{R}$ then there is no optimal solution since the target state can be reached in an arbitrary small time, while the control cannot be realized in a zero time. When $U = (-1, 1)$, the transfer can be achieved in a time larger than $\pi/2$, but not exactly in time $T = \pi/2$. It is only in the case where U is compact, e.g., $U = [-1, 1]$, that an optimal solution exists. For the constraint $|u(t)| \leq 1$, we find a pulse that allows us to bring the population from one level to another in a time $T = \pi/2$.

We show in Secs. VIII and IX how to use these results in two standard quantum control examples.

VI. FIRST-ORDER CONDITIONS

For a smooth real-valued function of one variable $f^0 : \mathbb{R} \rightarrow \mathbb{R}$, first-order optimality conditions are obtained from the observation that, at points where $df^0/dx \neq 0$, the function f^0 , which is well approximated by its first-order Taylor series, cannot be optimal since it behaves locally as a nonconstant affine function. In this way, one obtains the necessary condition: if \bar{x} is minimal for f^0 then $df^0(\bar{x})/dx = 0$. First-order conditions in optimal control are derived in the same way. We have to require that, for a small control variation, there is no cost variation at first order.

More precisely, if $J[u(\cdot)]$ is the value of the cost for a reference admissible control $u(\cdot)$ [for instance, $J[u(\cdot)] = \int_0^T f^0[q(t), u(t)]dt$ in Approach A or $J[u(\cdot)] = \int_0^T f^0[q(t), u(t)]dt + d[\mathcal{T}, q(T)]$ in Approach B], and $v(\cdot)$ is another admissible control, one would like to consider a condition of the form

$$\frac{\partial J[u(\cdot) + hv(\cdot)]}{\partial h} \Big|_{h=0} = 0. \quad (10)$$

But difficulties may arise for the following reasons.

We work in an infinite-dimensional space (the space of controls) and hence condition (10) should be required for infinitely many $v(\cdot)$. It may very well happen that if $u(\cdot)$ and $v(\cdot)$ are admissible controls then $u(\cdot) + hv(\cdot)$ is not admissible for every h close to 0. Think, for instance, of the case in which $m = 1$ and $U = [a, b]$. If $u(t) \equiv b$ is the reference control then $u(t) + hv(t)$ is not admissible for any nonzero perturbation $v(\cdot)$ when $hv(t)$ is strictly positive.

Hence, one should be very careful in choosing the admissible variations in order to fulfill the control restrictions. In Approach A, one should only restrict to control variations for which the corresponding trajectory reaches the target. More precisely, if $\tilde{q}(\cdot)$ is the trajectory corresponding to the control $\tilde{u}(\cdot) := u(\cdot) + hv(\cdot)$, one should add the condition

$$\tilde{q}(T) \in \mathcal{T} \quad (11)$$

with T either free or constrained to be the fixed final time depending on the problem under study. Condition (11) should be considered as a *constraint* for the minimization problem, which results in the use of *Lagrange multipliers (normal and abnormal)*.

The occurrence of Lagrange multipliers in optimal control is therefore not due to the fact that the optimization takes place in an infinite-dimensional space, but is rather a general feature of constrained minimization problems, as explained in Sec. VIA.

A. Why Lagrange multipliers appear in constrained optimization problems

We first recall how to find the minimum of a function of n variables $f^0(x)$, where $x = (x_1, \dots, x_n)$, under the constraint $f(x) = 0$, with the method of Lagrange multipliers. Here f^0 and f are two smooth functions $\mathbb{R}^n \rightarrow \mathbb{R}$. We have two cases.

Case 1. If \bar{x} is a point such that $f(\bar{x}) = 0$ with $\nabla f(\bar{x}) \neq 0$ then the implicit function theorem guarantees that $\{x \mid f(x) = 0\}$ is a smooth hypersurface in a neighborhood of \bar{x} . In this case, a necessary condition for f^0 to have a minimum at \bar{x} is that the level set of f^0 (i.e., the set on which f^0 takes a constant value) is not transversal to the set $\{x \mid f(x) = 0\}$ at \bar{x} . See Fig. 3.

More precisely, this means that

$$\exists \lambda \in \mathbb{R} \text{ such that } \nabla f^0(\bar{x}) = \lambda \nabla f(\bar{x}). \quad (12)$$

This statement can be proved by assuming, for instance, that $\partial_{x_n} f(\bar{x}) \neq 0$. The set $\{x \mid f(x) = 0\}$ can then be expressed locally around \bar{x} as $x_n = g(x_1, \dots, x_{n-1})$. The requirements that

$$\begin{aligned} \partial_{x_i} f[x_1, \dots, x_{n-1}, g(x_1, \dots, x_{n-1})] &\equiv 0, \\ i &= 1, \dots, n-1, \\ \partial_{x_i} f^0[x_1, \dots, x_{n-1}, g(x_1, \dots, x_{n-1})]|_{x=\bar{x}} &= 0, \\ i &= 1, \dots, n-1, \end{aligned}$$

immediately provide condition (12) with $\lambda = \partial_{x_n} f^0(\bar{x}) / \partial_{x_n} f(\bar{x})$.

Note that λ could be equal to zero. This case corresponds to the situation in which f^0 has a critical point at \bar{x} even in the absence of the constraint.

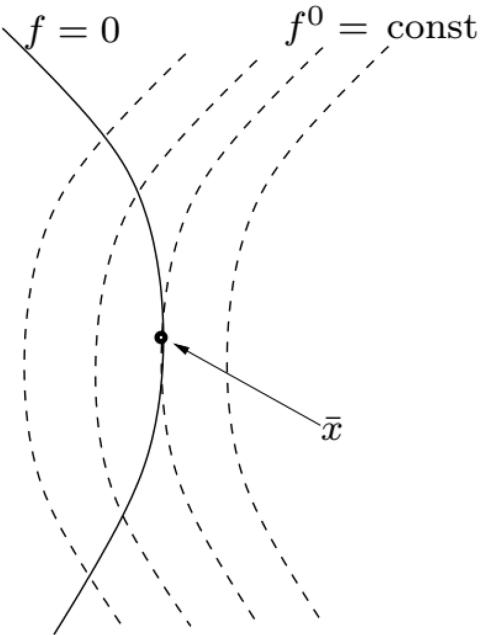


FIG. 3. Intersection of the set $f(x) = 0$ (solid line) with the level set of f^0 (dashed line). The two gradients $\nabla f(x)$ and $\nabla f^0(x)$ are parallel at $x = \bar{x}$.

Case 2. If \bar{x} is a point such that $f(\bar{x}) = 0$ with $\nabla f(\bar{x}) = 0$ then the set $\{x \mid f(x) = 0\}$ could be very complicated in a neighborhood of \bar{x} (typical examples are a single point, two crossing curves, . . . , but it could be any closed set). In general, the value of f^0 at these points cannot be compared with neighboring points by requiring that a certain derivative is zero [think, for instance, of the case in which $\{x \mid f(x) = 0\}$ is an isolated point]. However, they are candidates to optimality. As an illustrative example, consider the case where $n = 2$, $f^0(x_1, x_2) = x_1^2 + (x_2 - 1/4)^2$, and $f(x_1, x_2) = (x_1^2 + x_2^2)(x_1^2 + x_2^2 - 1)$. The point $\bar{x} = (0, 0)$ is an isolated point for which $f(\bar{x}) = 0$ and $\nabla f(\bar{x}) = 0$.

These results can be rewritten in the following form.

Theorem 4 (Lagrange multiplier rule in \mathbb{R}^n). *Let f^0 and f be two smooth functions from \mathbb{R}^n to \mathbb{R} . If f^0 has a minimum at \bar{x} on the set $\{x \mid f(x) = 0\}$ then there exists $(\bar{\lambda}, \bar{\lambda}_0) \in \mathbb{R}^2 \setminus \{(0, 0)\}$ such that, setting $\Lambda(x, \lambda, \lambda_0) = \lambda f(x) + \lambda_0 f^0(x)$, we have*

$$\nabla_x \Lambda(\bar{x}, \bar{\lambda}, \bar{\lambda}_0) = 0, \quad \nabla_\lambda \Lambda(\bar{x}, \bar{\lambda}, \bar{\lambda}_0) = 0. \quad (13)$$

To show that this statement is equivalent to what we just discussed, we observe that the second equality in Eq. (13) gives the constraint $f(\bar{x}) = 0$. For the first equation, we have two cases. If $\bar{\lambda}_0 \neq 0$ then we can normalize $\bar{\lambda}_0 = -1$ and we get $\bar{\lambda} \nabla_x f(\bar{x}) - \nabla_x f^0(\bar{x}) = 0$, i.e., Eq. (12) with the change of notation $\lambda \rightarrow \bar{\lambda}$. If $\bar{\lambda}_0 = 0$ then $\bar{\lambda} \neq 0$ and we get $\nabla_x f(\bar{x}) = 0$, that is, the second case studied above.

The quantities $\bar{\lambda}$ and $\bar{\lambda}_0$ are respectively called the *Lagrange multiplier* and *abnormal Lagrange multiplier*. If $(\bar{x}, \bar{\lambda}, \bar{\lambda}_0)$ is a solution of Eq. (13) with $\bar{\lambda}_0 \neq 0$ (respectively $\bar{\lambda}_0 = 0$) then \bar{x} is called a *normal extremal* (respectively *abnormal extremal*). An abnormal extremal is a candidate for optimality and occurs, in particular, when we cannot guarantee (at first order) that the set $\{x \mid f(x) = 0\}$ is a smooth curve. Abnormal extremals are candidates for optimality regardless of cost f^0 . Note that if \bar{x} is such that $\nabla_x f(\bar{x}) = 0$ and $\nabla_x f^0(\bar{x}) = 0$ then \bar{x} is both normal and abnormal. This is the case in which \bar{x} satisfies the first-order condition for optimality even without the constraint, but we cannot guarantee that the constraint is a smooth curve.

In the (infinite-dimensional) case of an optimal control problem, normal and abnormal Lagrange multipliers appear in a very similar way.

B. Statement of the Pontryagin maximum principle

In this section, we state the first-order necessary conditions for optimal control problems, namely the PMP. The basic idea is to define a new object [the pre-Hamiltonian; see Eq. (14) below] that allows us to formulate the Lagrange multiplier conditions in a simple and direct way.

The theorem is stated in a more general form that unifies and slightly generalizes the optimal control problems of Approaches A and B. In particular, we add to the cost $\int_0^T f^0[q(t), u(t)]dt$ a general terminal cost $\phi[q(T)]$. In Approach A, we have $\phi = 0$, while in Approach B, ϕ represents the distance from $q(T)$ to the target \mathcal{T} . We allow the target \mathcal{T} to coincide with M . This corresponds to leaving the final point $q(T)$ free in Approach B.

Theorem 5. *Consider the optimal control problem*

$$\begin{aligned} \dot{q}(t) &= f[q(t), u(t)], \\ q(0) &= q_{\text{in}}, \quad q(T) \in \mathcal{T}, \\ \int_0^T f^0[q(t), u(t)]dt + \phi[q(T)] &\rightarrow \min, \end{aligned}$$

where

- M is a smooth manifold of dimension n , $U \subset \mathbb{R}^m$,
- \mathcal{T} is a (nonempty) smooth submanifold of M ; it can be reduced to a point (fixed terminal point) or coincide with M (free terminal point),
- f, f^0 are smooth,
- $u \in \mathcal{U}$,
- $q : [0, T] \rightarrow M$ is a continuous curve [73].

Define the function (called the pre-Hamiltonian)

$$\mathcal{H}(q, p, u, p^0) = \langle p, f(q, u) \rangle + p^0 f^0(q, u) \quad (14)$$

with

$$(q, p, u, p^0) \in T^*M \times U \times \mathbb{R}$$

(see Ref. [70] for a precise definition of T^*M).

If the pair $(q, u) : [0, T] \rightarrow M \times U$ is optimal then there exists a never vanishing continuous pair $(p, p^0) : [0, T] \ni t \mapsto [p(t), p^0] \in T_{q(t)}^*M \times \mathbb{R}$, where $p^0 \leq 0$ is a constant, such that, for almost every (a.e.) $t \in [0, T]$, we have

- (i) q satisfies the Hamiltonian equation $\dot{q}(t) = \frac{\partial \mathcal{H}}{\partial p} [q(t), p(t), u(t), p^0]$;
- (ii) p satisfies the Hamiltonian equation $\dot{p}(t) = -\frac{\partial \mathcal{H}}{\partial q} [q(t), p(t), u(t), p^0]$;
- (iii) the quantity $\mathcal{H}_M[q(t), p(t), p^0] := \max_{v \in U} \mathcal{H}[q(t), p(t), v, p^0]$ is well defined and $\mathcal{H}[q(t), p(t), u(t), p^0] = \mathcal{H}_M[q(t), p(t), p^0]$,

which corresponds to the maximization condition.

Moreover,

- (iv) there exists a constant $c \geq 0$ such that $\mathcal{H}_M[q(t), p(t), p^0] = c$ on $[0, T]$, with $c = 0$ if the final time is free (value of the Hamiltonian);
- (v) for every $v \in T_{q(T)}\mathcal{T}$, we have $\langle p(T), v \rangle = p^0 \langle d\phi[q(T)], v \rangle$ (transversality condition), where $d\phi$ is the differential of the function ϕ .

Some comments are in order.

• A proof of the PMP can be found in, for instance, Refs. [27,28]. An intuitive derivation based on the Lagrange multiplier rule is presented in Sec. VII A for the case in which T is fixed, the final point is free, and $U = \mathbb{R}^m$.

• The covector p is called the *adjoint state* in the control theory literature [70], while p^0 is the *abnormal multiplier*. The quantities $p(\cdot)$ and p^0 play the roles of Lagrange multipliers for the constrained optimization problem. We point out the similarity between the expressions of \mathcal{H} and of Λ in Theorem 4 (with the change of notation $q \rightarrow x$ and $p \rightarrow \lambda$).

• A trajectory $q(\cdot)$ for which there exist $p(\cdot)$, $u(\cdot)$, and p^0 such that $[q(\cdot), p(\cdot), u(\cdot), p^0]$ satisfies all the conditions given by the PMP is called an *extremal trajectory* and the 4-uple $[q(\cdot), p(\cdot), u(\cdot), p^0]$ an *extremal* or, equivalently, an *extremal lift* of $q(\cdot)$. Such an extremal is called *normal* if $p^0 \neq 0$ and *abnormal* if $p^0 = 0$. It may happen that an extremal trajectory $q(\cdot)$ admits both a normal extremal lift $[q(\cdot), p_1(\cdot), u(\cdot), p^0]$ and an abnormal one $[q(\cdot), p_2(\cdot), u(\cdot), 0]$. In this case, we say that the extremal trajectory $q(\cdot)$ is a *nonstrict abnormal trajectory*. Note that (as in the finite-dimensional case) abnormal trajectories are candidates for optimality regardless of the cost. In the finite-dimensional case, they correspond to singularities of the constraint function, while here they correspond to singularities of the functional that maps the control $v(\cdot)$ to the

final point at time T of the solution of $\dot{q}(t) = f[q(t), v(t)]$, $q(0) = q_{\text{in}}$. It is worth noting that abnormal extremals do not only appear in pathological cases, but they are often present in real-world applications, such as, for instance, in the two examples presented in Secs. VIII and IX.

• The PMP is only a necessary condition for optimality. It may very well happen that an extremal trajectory is not optimal. The PMP can therefore provide several candidates for optimality, only some of which are optimal (or even none of them if the step of existence has not been verified; see Sec. V).

• Since the equation for $p(\cdot)$ in Theorem 5(ii) is linear, if $[q(\cdot), p(\cdot), u(\cdot), p^0]$ is an extremal then, for every $\alpha > 0$, $[\alpha q(\cdot), \alpha p(\cdot), u(\cdot), \alpha p^0]$ is an extremal as well. As a consequence, some useful normalizations are possible. A typical normalization for normal extremals is to require that $p^0 = -\frac{1}{2}$, but other choices are also possible.

• When there is no final cost ($\phi = 0$), the transversality condition (v) of Theorem 5 simplifies to

$$\langle p(T), T_{q(T)}\mathcal{T} \rangle = 0. \quad (15)$$

When the final point is fixed ($\mathcal{T} = \{q_{\text{fi}}\}$), $T_{q(T)}\mathcal{T}$ is a zero-dimensional manifold and, hence, condition (15) is empty. When the final point is free ($\mathcal{T} = M$), the transversality condition simplifies to $p(T) = p^0 d\phi[q(T)]$. In local coordinates, we recover that $p(T)$ is proportional to the gradient of ϕ evaluated at the point $q(T)$. Note that, since $[p(T), p^0] \neq 0$, in this case one necessarily has $p^0 \neq 0$.

Table II gives a list of the possible extremal solutions of the PMP.

Example 3. As a general example in quantum control, we consider a dynamical system governed by Eq. (2) where the goal is to minimize at the fixed final time T the cost $-|\langle \psi_{\text{fi}} | \psi(T) \rangle|^2 + \frac{1}{2} \int_0^T \sum_{j=1}^m u_j^2(t) dt$, where ψ_{fi} is a target state towards which we want to drive the system (up to a global phase) [90]. A direct application of the PMP shows

TABLE II. List of the possible extremal solutions of the PMP.

Name	Definition
Extremal	4-uple $[q(\cdot), p(\cdot), u(\cdot), p^0]$ solution of the PMP
Normal extremal	Extremal with $p^0 \neq 0$
Abnormal extremal	Extremal with $p^0 = 0$
Nonstrict abnormal trajectory	Trajectory which admits both abnormal and normal lifts
Regular control	When the maximization condition of the PMP gives $u = \omega(q, p)$
Singular control	When the control is not regular

that the pre-Hamiltonian \mathcal{H} can be expressed as

$$\mathcal{H}(\psi, \chi, u_j, p^0) = \text{Re}(\langle \chi | \dot{\psi} \rangle) + \frac{p^0}{2} \sum_j u_j^2,$$

where the adjoint state, denoted here by χ , is an abstract wave function that can be chosen so that it belongs to the unit sphere in \mathbb{C}^N , $\langle \chi | \chi \rangle = 1$. Since ψ and χ are complex-valued functions, the pre-Hamiltonian is defined through the real part of the scalar product between χ and $\dot{\psi}$. The standard definition used in Theorem 5 can be found by introducing the real and imaginary parts of the wave functions. Using Eq. (2), we deduce that

$$\begin{aligned} \mathcal{H}(\psi, \chi, u_j, p^0) &= \text{Im}\left(\langle \chi | H_0 + \sum_j u_j H_j | \psi \rangle\right) \\ &\quad + \frac{p^0}{2} \sum_j u_j^2, \end{aligned} \quad (16)$$

which leads to

$$i\dot{\chi}(t) = \left(H_0 + \sum_{j=1}^m u_j(t) H_j\right) \chi(t),$$

i.e., χ also satisfies the Schrödinger equation. We stress that this condition is only true in the bilinear case. A specific equation has to be computed for other dynamics, such as, e.g., the Gross-Pitaevskii equation [91]. The final condition $\chi(T)$ is given by the transversality condition (v) of the PMP:

$$\chi(T) = p^0 \langle \psi_{\text{fi}} | \psi(T) \rangle \psi_{\text{fi}}. \quad (17)$$

The maximization condition of the PMP leads to the constraints

$$\frac{\partial \mathcal{H}}{\partial u_j}(\psi, \chi, u_j, p^0) = 0$$

for $j = 1, \dots, m$. A direct computation from Eq. (16) gives

$$\text{Im}(\langle \chi | H_j | \psi \rangle) + p^0 u_j = 0.$$

For the normal extremal with $p^0 = -1$, we finally get

$$u_j = \text{Im}(\langle \chi | H_j | \psi \rangle). \quad (18)$$

C. Use of the PMP

The application of the PMP is not so straightforward. Indeed, there are many conditions to satisfy and all of them are coupled. This section is aimed at describing how to use it in practice.

The following points should be followed first for normal extremals ($p^0 < 0$) and then for abnormal extremals

($p^0 = 0$). In the first case, p^0 can be normalized to $-1/2$ since p^0 is defined up to a multiplicative positive factor. In the different steps, several difficulties (that are briefly mentioned) may arise. Most of them should be solved case by case, since they can be of different nature depending on the problem under study.

Step 1. Use maximization condition (iii) of Theorem 5 to express, when possible, the control as a function of the state and of the covector, i.e., $u = w(q, p)$. Note that if we have m controls (e.g., if \mathcal{T} is an open subset of \mathbb{R}^m) then the first-order maximality conditions give m equations for m unknowns. When the maximization condition permits expressing u as a function of q and p , we say that the control is *regular*; otherwise, the control is said to be *singular*. We may have regions where the control is regular and regions where it is singular. For singular controls, finer techniques have to be used to derive the expression of the control. These different points are discussed in the examples in Secs. VIII and IX.

Step 2. Insert the control found in the previous step into the Hamiltonian equations (i) and (ii) of Theorem 5:

$$\begin{aligned} \dot{q}(t) &= \frac{\partial \mathcal{H}}{\partial p}\{q(t), p(t), w[q(t), p(t)], p^0\}, \\ \dot{p}(t) &= -\frac{\partial \mathcal{H}}{\partial q}\{q(t), p(t), w[q(t), p(t)], p^0\}. \end{aligned} \quad (19)$$

In the case where the previous step provides a smooth $w(\cdot, \cdot)$, this is a well-defined set of $2n$ equations for $2n$ unknowns. Note, however, that the boundary conditions are given in a nonstandard form since we know $q(0)$ but not $p(0)$. Instead of $p(0)$, we have partial information on $q(T)$ and $p(T)$ depending on the dimension of \mathcal{T} [see the next step to understand how these final conditions are shared between $q(T)$ and $p(T)$]. We then solve Eq. (19) for *fixed* $q(0) = q_{\text{in}}$ and *any* $p(0) = p_{\text{in}} \in T_{q_{\text{in}}}^* M$. Let us denote the solution as

$$q(t; p_{\text{in}}, p^0), \quad p(t; p_{\text{in}}, p^0).$$

We stress that, when $w(\cdot, \cdot)$ is not regular enough, solutions to the Cauchy problem (19) with $q(0) = q_{\text{in}}$ and $p(0) = p_{\text{in}}$ may fail to exist or be unique.

Step 3. Find p_{in} such that

$$q(T; p_{\text{in}}, p^0) \in \mathcal{T}. \quad (20)$$

Note that if \mathcal{T} is reduced to a point and T is fixed, we get n equations for n unknowns (the components of p_{in}). If T is free then an additional equation is needed. This condition is given by relation (iv) in the PMP. If \mathcal{T} is a k -dimensional submanifold of M ($k \leq n$) then Eq. (20) provides only $n - k$ equations and the remaining ones correspond to the transversality condition (v) of the PMP.

Step 4. If Eq. (20) [together with the transversality condition and condition (iv) of the PMP if T is free] has a unique solution p_{in} and if we have verified *a priori* the existence step, then the optimal control problem is solved. Unfortunately, in general, there is no reason for Eq. (20) to provide a unique solution. Indeed, the PMP is only a necessary condition for optimality. If several solutions are found, one should choose the best among them by a direct comparison of the value of the cost. This is, in general, a nontrivial step, complicated by the difficulty of explicitly solving Eq. (20). For this reason, several techniques have been developed to select the extremals. Among others, we mention the sufficient conditions for optimality given by Hamilton-Jacobi-Bellman theory and synthesis theory. We refer the reader to Ref. [92] for a discussion. In Example 1 (Sec. VIII) we are able to select the optimal solution without the use of sufficient conditions for optimality, while this is not the case in Example 2 (Sec. IX).

Example 4. We come back to the case of Example 3. We have shown with Eq. (18) that the maximization condition allows us to express the m controls u_j as functions of ψ and χ in the normal case. This situation therefore corresponds to step 1 above where the control is regular. For abnormal extremals for which $p^0 = 0$, we get

$$\text{Im}(\langle \chi | H_j | \psi \rangle) = 0, \quad (21)$$

and the control is singular because this relation does not directly give the expression for u_j .

Applying step 2, we obtain in the regular situation the following coupled equations for ψ and χ :

$$\begin{aligned} i\dot{\psi} &= \left[H_0 + \sum_j \text{Im}(\langle \chi | H_j | \psi \rangle) H_j \right] \psi, \\ i\dot{\chi} &= \left[H_0 + \sum_j \text{Im}(\langle \chi | H_j | \psi \rangle) H_j \right] \chi, \end{aligned} \quad (22)$$

with the boundary conditions $\psi(0) = \psi_{\text{in}}$ and Eq. (17). In step 3, we then solve Eq. (22) to find the initial condition $\chi(0)$ such that the final state $\chi(T)$ satisfies Eq. (17) at time T .

The numerical procedures used to select the initial condition $\chi(0)$, called *shooting methods* in control theory, are based on suitable adaptations of the Newton algorithm [35,36]. Step 4 consists finally in comparing the cost of the different solutions found in step 3.

In the abnormal case, we use the fact that Eq. (21) is satisfied in a nonzero time interval, so the time derivatives of $\text{Im}(\langle \chi | H_j | \psi \rangle)$ are also zero. The first time derivative leads to the m relations

$$\sum_{j=1}^m u_j \text{Re}(\langle \chi | [H_k, H_j] | \psi \rangle) = \text{Re}(\langle \chi | [H_0, H_k] | \psi \rangle)$$

with $k = 1, \dots, m$. This linear system can be expressed in a more compact form as

$$Ru = s,$$

where R is a $m \times m$ matrix with elements $R_{kj} = \text{Re}(\langle \chi | [H_k, H_j] | \psi \rangle)$ and s a vector of coordinates $s_k = \text{Re}(\langle \chi | [H_0, H_k] | \psi \rangle)$. We deduce that the control u is given as a function of ψ and χ as $u = R^{-1}s$. If this system is singular then the second time derivative has to be used. This is the case, e.g., for $m = 1$, when a further constraint has to be fulfilled, namely, $\text{Re}(\langle \chi | [H_0, H_1] | \psi \rangle) = 0$. From the derivation of u , we then apply steps 2 and 3 to the abnormal extremals.

VII. GRADIENT-BASED OPTIMIZATION ALGORITHM

The aim of this section is to introduce a first-order gradient-based optimization algorithm based on the PMP. We first derive the necessary conditions of the PMP in the case of a fixed control time without any constraint on the final state and on the control. This construction is known in control theory as the weak PMP. Note that we consider only the case of regular control. Iterative algorithms can be deduced from these conditions. In a second step, we apply this idea to quantum control and we show how a gradient-based optimization algorithm, GRAPE [45], can be designed from this approach.

A. The weak Pontryagin maximum principle

We consider a control system whose dynamics is governed by Eq. (1), when the final state is free and the control is unconstrained. The objective is to solve a control problem in the Approach B, as defined in Sec. III with a fixed control time T . We recall that the cost functional to minimize can be expressed as

$$J[u(\cdot)] = \int_0^T f^0[q(t), u(t)] dt + d[T, q(T)].$$

Considering the evolution equation (1) as a dynamical constraint (in infinite dimension), in order to apply (formally) the Lagrange multiplier rule for normal extremals, we introduce the functional

$$\begin{aligned} \Lambda[p(\cdot), u(\cdot)] &= d[T, q(T)] + \int_0^T f^0[q(t), u(t)] dt \\ &\quad + \int_0^T \langle p(t), \dot{q}(t) - f[q(t), u(t)] \rangle dt. \end{aligned} \quad (23)$$

We stress that the Lagrange multiplier $p(\cdot)$ is here a function on $[0, T]$. Integrating Eq. (23) by parts, we obtain

$$\begin{aligned}\Lambda[p(\cdot), u(\cdot)] &= d[\mathcal{T}, q(T)] + \langle p(T), q(T) \rangle \\ &\quad - \langle p(0), q(0) \rangle \\ &\quad - \int_0^T \{H[q(t), p(t), u(t)] + \langle \dot{p}(t), q(t) \rangle\} dt\end{aligned}$$

with

$$H(q, p, u) = \langle p, f(q, u) \rangle - f^0(q, u). \quad (24)$$

Note that the scalar function H has the same expression (with $p^0 = -1$) as the pre-Hamiltonian \mathcal{H} introduced in Sec. VI B for the statement of the PMP. Since there is no constraint on the control law, i.e., $u(t) \in \mathbb{R}^m$ for any time t , we consider the variation $\delta\Lambda$ in Λ at first order due to the variation δu of u . This change of control induces a variation of the trajectory $\delta q(t)$ with $\delta q(0) = 0$, the initial point being fixed. Note that the adjoint state p is not modified. We arrive at

$$\begin{aligned}\delta\Lambda &= \left\langle \frac{\partial d(\mathcal{T}, q)}{\partial q} \Big|_{q=q(T)} + p(T), \delta q(T) \right\rangle \\ &\quad - \int_0^T \left[\left\langle \frac{\partial H}{\partial q} + \dot{p}, \delta q \right\rangle + \frac{\partial H}{\partial u} \delta u \right] dt. \quad (25)\end{aligned}$$

A necessary condition for Λ to be an extremum is $\delta\Lambda = 0$ for any variation δu . A solution is given by taking an adjoint state p satisfying

$$\dot{p} = -\frac{\partial H}{\partial q}, \quad (26)$$

the final boundary condition

$$p(T) = -\frac{\partial d(\mathcal{T}, q)}{\partial q} \Big|_{q=q(T)}, \quad (27)$$

and requiring that

$$\frac{\partial H}{\partial u} = 0 \quad \text{on } [0, T].$$

As could be expected, we find here a weak version of the equations of the PMP introduced in Sec. VI B where the maximization of the pre-Hamiltonian is replaced by an extremum condition given by the partial derivative with respect to u . We point out that this approach works if the set U is open.

B. Gradient-based optimization algorithm

The set of nonlinear coupled differential equations can be solved numerically by an iterative algorithm. The basic idea used in such algorithms can be formulated as follows. Assume that a control $u(\cdot)$ sufficiently close to the optimal solution is known. If $p(\cdot)$ satisfies Eqs. (26) and (27) then we deduce from Eq. (25) that

$$\delta\Lambda = - \int_0^T \frac{\partial H}{\partial u} \delta u dt.$$

This suggests that a better control can be achieved with the choice $\delta u = \epsilon \partial H / \partial u$, where ϵ is a small positive parameter. The iterative algorithm is then described by the following steps.

- (1) Choose a guess control $u(\cdot)$.
- (2) Propagate forward the state of the system q from $\dot{q} = f(q, u)$, with the initial condition $q(0) = q_0$.
- (3) Propagate backward the adjoint state of the system from $\dot{p} = -\partial H / \partial q$, with the final condition $p(T) = -\partial d(\mathcal{T}, q) / \partial q|_{q=q(T)}$.
- (4) Compute the correction δu to the control law, $\delta u(t) = \epsilon \partial H / \partial u$, where $\epsilon > 0$ is a small parameter.
- (5) Define the new control $u \mapsto u + \delta u$.
- (6) Go to step (2) and repeat until a given accuracy is reached.

This algorithm is an example of a first-order gradient-based optimization algorithm. By construction, it converges towards an extremal control of J that is not, in general, a global minimum solution of the control problem, only a local one. Note that several numerical details are hidden in the description of this method. Among others, we mention the choice of the guess control that allows one to reach a good local solution and the determination of the parameter ϵ . The latter must be sufficiently small to remain in the first-order approximation, but large enough to reduce the number of iterations and therefore the computational time. We refer the reader to standard numerical optimization textbooks to address these issues [36].

This approach can be directly applied to quantum systems. The bilinearity of quantum dynamics allows us to simplify the different terms used in the algorithm. We consider a quantum system whose dynamics is governed by the Schrödinger equation

$$i\dot{\psi}(t) = [H_0 + u(t)H_1]\psi(t).$$

The goal of the control process is to bring the system from ψ_{in} towards ψ_{fi} in a fixed time T . The control problem aims at minimizing the cost

$$J = \frac{1}{2} \int_0^T u(t)^2 dt - |\langle \psi_{\text{fi}} | \psi(T) \rangle|^2.$$

In the normal case, the pre-Hamiltonian H can be expressed as

$$H = \text{Re}[\langle \chi | \dot{\psi} \rangle] - \frac{u^2}{2} = \text{Im}[\langle \chi | H_0 + uH_1 | \psi \rangle] - \frac{u^2}{2},$$

where the wave function χ is the adjoint state of the system. We thus deduce that the gradient on which the iterative algorithm is based is given by

$$\frac{\partial H}{\partial u} = \text{Im}[\langle \chi | H_1 | \psi \rangle] - u(t). \quad (28)$$

We find with Eq. (28) the standard control correction used in the GRAPE algorithm in quantum control [45].

VIII. EXAMPLE 1: A THREE-LEVEL QUANTUM SYSTEM WITH COMPLEX CONTROLS

In this section, we mainly use the results of Ref. [50]; see also Refs. [49, 93, 94]. Note, however, that original results concerning the selection of the best extremal among all the possible solutions are presented.

A. Formulation of the quantum control problem

We consider a three-level quantum system whose dynamics is governed by the Schrödinger equation. The system is described by a pure state $\psi(t)$ belonging to a three-dimensional complex Hilbert space. The system is characterized, in the absence of external fields, by three energy levels, E_1 , E_2 , and E_3 , and is controlled by the pump and Stokes pulses that respectively couple states one and two and states two and three [95]. Note that there is no direct coupling between levels one and three. The time evolution of $\psi(t)$ is given by

$$i\dot{\psi}(t) = H(t)\psi(t),$$

where

$$H(t) = \begin{pmatrix} E_1 & \Omega_1(t) & 0 \\ \Omega_1^*(t) & E_2 & \Omega_2(t) \\ 0 & \Omega_2^*(t) & E_3 \end{pmatrix}.$$

Here $\Omega_1(t), \Omega_2(t) \in \mathbb{C}$ are the two time-dependent complex control parameters. We denote by $\psi_1(t)$, $\psi_2(t)$, and $\psi_3(t)$ the coordinates of $\psi(t)$, that is, $\psi(t) = [\psi_1(t), \psi_2(t), \psi_3(t)]$. They satisfy

$$|\psi_1(t)|^2 + |\psi_2(t)|^2 + |\psi_3(t)|^2 = 1,$$

leading to $M = S^5$, a manifold of real dimension 5. The goal of the control process is to transfer the population from the first eigenstate to the third one in a fixed time T . In other words, the aim is to find a trajectory in M going

from the submanifold $|\psi_1|^2 = 1$ to that with $|\psi_3|^2 = 1$. The system is completely controllable thanks to point 2 of Proposition 9 below and Approach A can be chosen. The optimal control problem is defined through the cost functional

$$C = \int_0^T [|\Omega_1(t)|^2 + |\Omega_2(t)|^2] dt$$

to be minimized. The cost C can be interpreted as the energy of the control laws used in the control process. We consider the specific case in which the control parameters are in resonance with the energy transition. More precisely, we assume that the pulses $\Omega_1(t)$ and $\Omega_2(t)$ can be expressed as

$$\begin{aligned} \Omega_1(t) &= u_1(t)e^{i(E_2-E_1)t}, \\ \Omega_2(t) &= u_2(t)e^{i(E_3-E_2)t}, \end{aligned}$$

with $u_1(t), u_2(t) \in \mathbb{R}$ to be optimized. Note that this assumption is not restrictive since it can be shown that the resonant case corresponds to the optimal solution [50, 93]. The uncontrolled part, called the drift, together with the imaginary unit in the Schrödinger equation, can be eliminated through a unitary transformation $Y(t)$ given by

$$Y(t) = \text{diag}(e^{-iE_1 t}, e^{-i(E_2 t + \pi/2)}, e^{-i(E_3 t + \pi)}).$$

Defining a new wave function x such that $\psi(t) = Y(t)x(t)$, we find that $x(t)$ solves the Schrödinger equation

$$i\dot{x}(t) = H'(t)x(t),$$

where $H' = Y^{-1}HY - iY^{-1}\dot{Y}$. Since Y only modifies the phases of the coordinates, $\psi(t)$ and $x(t)$ correspond to the same population distribution, i.e., setting $x = (x_1, x_2, x_3)$ we have $|x_j(t)|^2 = |\psi_j(t)|^2$, $j = 1, 2, 3$.

Computing explicitly H' we arrive at

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{pmatrix} = \begin{pmatrix} 0 & -u_1(t) & 0 \\ u_1(t) & 0 & -u_2(t) \\ 0 & u_2(t) & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}. \quad (29)$$

Without loss of generality, the optimal control problem can be restricted to the submanifold $S^2 \subset M$ defined by $\{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_1^2 + x_2^2 + x_3^2 = 1\}$. This statement is trivial if the initial condition belongs to S^2 [i.e., if $x(0)$ is real]. Otherwise, a straightforward change of phase of the coordinates ψ_k allows one to come back to this condition.

Equation (29) can be expressed in a more compact form as

$$\dot{x} = u_1(t)F_1(x) + u_2(t)F_2(x), \quad (30)$$

where

$$F_1(x) = \begin{pmatrix} -x_2 \\ x_1 \\ 0 \end{pmatrix}, \quad F_2(x) = \begin{pmatrix} 0 \\ -x_3 \\ x_2 \end{pmatrix}.$$

Note that F_1 and F_2 are two vector fields defined on the sphere S^2 respectively representing a rotation along the x_3 axis and along the x_1 axis. Since $\Omega_1(t)$ and $\Omega_2(t)$ differ from $u_1(t)$ and $u_2(t)$ only in phase factors, the minimization problem becomes

$$C = \int_0^T [u_1(t)^2 + u_2(t)^2] dt \rightarrow \min \quad (31)$$

with T fixed. Concerning initial and final conditions, since the goal is to go from the submanifold $|\psi_1|^2 = 1$ to the submanifold $|\psi_3|^2 = 1$, we can assume without loss of generality that $x(0) = (1, 0, 0)$ (again, a straightforward change of coordinates allows one to come back to this condition otherwise). Since we are now restricted to real variables, the target is $\mathcal{T} = \{(0, 0, +1), (0, 0, -1)\}$. Now, as the target is made of two points only, one should separately compute the optimal trajectories going from $(1, 0, 0)$ to $(0, 0, 1)$ and those going from $(1, 0, 0)$ to $(0, 0, -1)$. Finally, between all these trajectories, one should take those having the smallest cost. Because of the symmetries of the system, the two families of optimal trajectories have precisely the same cost (this will be clear in the explicit computations later on). As a consequence, without loss of generality, we can fix the final condition as $x(T) = (0, 0, 1)$.

Problem (30)–(31) with fixed initial and final conditions is actually a celebrated problem in OCT called the Grushin model on the sphere [49, 96, 97].

B. Existence

For convenience, let us rewrite the optimal control problem as follows.

Problem $P^{Grushin}(T)$. We have

$$\begin{aligned} \dot{x} &= u_1(t)F_1(x) + u_2(t)F_2(x), \\ \int_0^T [u_1(t)^2 + u_2(t)^2] dt &\rightarrow \min, \quad T \text{ fixed}, \\ x(0) &= (1, 0, 0), \quad x(T) = (0, 0, 1), \\ u_1, u_2 &\in \mathcal{U}, \quad U = \mathbb{R}. \end{aligned}$$

To prove the existence of solutions to $P^{Grushin}(T)$, one could be tempted to use Propositions 3 and 11. However, u_1 and u_2 take values in \mathbb{R} and, hence, the second hypothesis of the proposition is not verified.

Instead, we are going to use the following fact.

Proposition 6. If $u_1(t), u_2(t)$ are optimal controls for $P^{Grushin}(T)$ then $u_1(t)^2 + u_2(t)^2$ is almost everywhere constant and positive on $[0, T]$. Moreover, for every $\alpha > 0$, $\alpha u_1(\alpha t)$ and $\alpha u_2(\alpha t)$ are optimal controls for $P^{Grushin}(T/\alpha)$.

To prove this proposition, we first state a general lemma for driftless systems.

Lemma 7. Consider a control system of the form $\dot{x} = \sum_{j=1}^m u_j(t)F_j(x)$, where $x \in M$ and $u_j \in \mathcal{U}$, $j = 1, \dots, m$. Then any admissible trajectory $x(\cdot)$ defined on $[0, T]$ and corresponding to controls $u_j(\cdot)$, $j = 1, \dots, m$, is a reparameterization of an admissible trajectory $\bar{x}(\cdot)$ defined on the same time interval whose controls $\bar{u}_j(\cdot)$ satisfy a.e. $\sqrt{\sum_{j=1}^m \bar{u}_j(t)^2} = L/T$, where $L = \int_0^T \sqrt{\sum_{j=1}^m u_j(t)^2} dt$. Alternatively, $x(\cdot)$ is a reparameterization of a trajectory $\bar{x}(\cdot)$ defined on $[0, L]$ whose controls satisfy a.e. $\sqrt{\sum_{j=1}^m \bar{u}_j(t)^2} = 1$.

We recall that, given an admissible trajectory $x : [0, T] \rightarrow M$, $T > 0$, corresponding to controls $u_j(\cdot)$, $j = 1, \dots, m$, a reparameterization of $x(\cdot)$ is a trajectory $\bar{x}(\cdot) = x[\tau(\cdot)]$ with $\tau : [0, \bar{T}] \rightarrow [0, T]$ a function such that $d\tau(t)/dt > 0$ a.e. on $[0, \bar{T}]$. Such a trajectory is defined on $[0, \bar{T}]$. Lemma 7 is a consequence of the fact that, for a.e. $t \in [0, \bar{T}]$, we have

$$\begin{aligned} \dot{\bar{x}}(t) &= \frac{d}{dt} x[\tau(t)] \\ &= \dot{x}[\tau(t)]\dot{\tau}(t) \\ &= \left(\sum_{j=1}^m u_j[\tau(t)]F_j\{x[\tau(t)]\} \right) \dot{\tau}(t) \\ &= \sum_{j=1}^m \{u_j[\tau(t)]\dot{\tau}(t)\} F_j[\bar{x}(t)], \end{aligned}$$

from which it follows that $\bar{x}(\cdot)$ is admissible and corresponds to controls $u_j[\tau(\cdot)]\dot{\tau}(\cdot)$, $j = 1, \dots, m$.

In the following, for the optimal control problem under study, it is convenient to normalize T in such a way that $u_1(t)^2 + u_2(t)^2 = 1$. Usually, when one makes this choice, the trajectories are said to be *parameterized by arc length*. If the objective is to reach the target at time T' , it is sufficient to use the controls $[\alpha u_1(\alpha t), \alpha u_2(\alpha t)]$, where $\alpha = T/T'$.

When T is fixed in such a way that $u_1(t)^2 + u_2(t)^2 = 1$, we call problem $P^{Grushin}(T)$ simply $P^{Grushin}$.

Note that, as a consequence of Lemma 7, if T is not fixed then $P^{Grushin}(T)$ has no solution. Indeed, assume by

contradiction that $x(\cdot)$, defined on $[0, T_0]$ and corresponding to controls $u_j(\cdot)$, $j = 1, 2$, is a minimizer for T free. Let $c \in (0, 1)$. The trajectory corresponding to controls $\bar{u}_j(\cdot) = u_j(c \cdot) c$, $j = 1, 2$, is a reparameterization of $x(\cdot)$ reaching the final point at time $\bar{T} = T_0/c$ with a cost

$$\begin{aligned} \int_0^{T_0/c} \sum_{j=1}^2 u_j(ct)^2 c^2 dt &= c \int_0^{T_0} \sum_{j=1}^2 u_j(t)^2 dt, \\ &< \int_0^{T_0} \sum_{j=1}^2 u_j(t)^2 dt, \end{aligned}$$

which leads to a contradiction.

Proof of Proposition 6. Let us define

$$\begin{aligned} L[u(\cdot)] &= \int_0^T \sqrt{u_1(t)^2 + u_2(t)^2} dt, \\ E[u(\cdot)] &= \int_0^T [u_1(t)^2 + u_2(t)^2] dt. \end{aligned}$$

We are going to use the Cauchy-Schwarz inequality $\langle f, g \rangle^2 \leq \|f\|^2 \|g\|^2$ (with equality holding if and only if f and g are proportional), which holds in any Hilbert space. This inequality simply tells us that the scalar product of two vectors is less than or equal to the product of the norms of the two vectors (with equality holding if and only if the two vectors are collinear). In particular, this can be used in the space $L^2([0, T], \mathbb{R})$ of measurable functions $f : [0, T] \rightarrow \mathbb{R}$ with $\int_0^T f(t)^2 dt < \infty$. Namely,

$$\left(\int_0^T f(t)g(t) dt \right)^2 \leq \int_0^T f(t)^2 dt \int_0^T g(t)^2 dt$$

(with equality holding if and only if $f \propto g$, a.e.). Now, let $f(t) = \sqrt{u_1(t)^2 + u_2(t)^2}$ and $g(t) = 1$ for $t \in [0, T]$. Note that $f, g \in \mathcal{U} \subset L^2([0, T], \mathbb{R})$. We have

$$L[u(\cdot)]^2 \leq E[u(\cdot)]T$$

[with equality holding if and only if $u_1(t)^2 + u_2(t)^2 = \text{const a.e.}$]. Now, let $u(\cdot)$ be a minimizer of E defined on $[0, T]$. Assume by contradiction that $u_1(t)^2 + u_2(t)^2$ is not a.e. constant. Then $L[u(\cdot)]^2 < E[u(\cdot)]T$. Let $\bar{x}(\cdot)$ be an admissible trajectory defined on $[0, T]$ corresponding to controls satisfying $\sqrt{\bar{u}_1(t)^2 + \bar{u}_2(t)^2} = L[u(\cdot)]/T$ a.e., of which $x(\cdot)$ is a reparameterization as in Lemma 7. One immediately checks that $L[u(\cdot)] = L[\bar{u}(\cdot)]$. For this trajectory, we have

$$E[\bar{u}(\cdot)]T = L[\bar{u}(\cdot)]^2 = L[u(\cdot)]^2 < E[u(\cdot)]T,$$

contradicting the fact that $u(\cdot)$ is a minimizer of E . ■

Now we have the following result.

Proposition 8. Problem P^{Grushin} is equivalent to the problem of minimizing time with the constraint on the controls $u_1(t)^2 + u_2(t)^2 \leq 1$.

Proof. Since for P^{Grushin} we have $u_1(t)^2 + u_2(t)^2 = 1$ a.e., then

$$\int_0^T [u_1(t)^2 + u_2(t)^2] dt = T.$$

Hence, P^{Grushin} is equivalent to the problem of minimizing T with the constraint on the controls $u_1(t)^2 + u_2(t)^2 = 1$. To conclude the proof, let us show that a trajectory corresponding to controls for which the condition

$$u_1(t)^2 + u_2(t)^2 = 1 \quad \text{a.e.} \quad (32)$$

is not satisfied cannot be optimal for the time-optimal control problem mentioned in the statement. Actually, if Eq. (32) is not satisfied then $L = \int_0^T \sqrt{u_1(t)^2 + u_2(t)^2} < T$ and an arc length reparameterization of the trajectory reaches the target in exactly time L . ■

Let us now go back to the problem of existence of optimal trajectories for P^{Grushin} . Thanks to Proposition 8, P^{Grushin} can be equivalently recast as a time-optimal control problem with controls in the convex and compact set $U = \{(u_1, u_2) \in \mathbb{R}^2 \mid u_1^2 + u_2^2 \leq 1\}$. We can then apply Propositions 3 and 12 and deduce the existence of an optimal trajectory for P^{Grushin} [as well as $P^{\text{Grushin}}(T)$ for every $T > 0$].

C. Application of the PMP

Before applying the PMP, it is convenient to reformulate the problem in spherical coordinates. Indeed, one can prove the following statement.

Claim 1. Consider an optimal control problem as in the statement of the PMP (Theorem 5). If all admissible trajectories starting from q_{in} are contained in a submanifold of M of dimension strictly smaller than n , then each admissible trajectory has an abnormal extremal lift.

This property can be qualitatively justified as follows. We have already mentioned that abnormal trajectories correspond to singularities of the functional that maps a control law to the final point of the corresponding controlled trajectory. If all admissible trajectories starting from q_{in} are contained in a proper submanifold of M then the endpoint functional is everywhere singular, meaning that each admissible trajectory is abnormal.

As a consequence, since in our case all trajectories are contained in the sphere S^2 , if we apply the PMP in \mathbb{R}^3 , all optimal trajectories admit an abnormal extremal lift. This

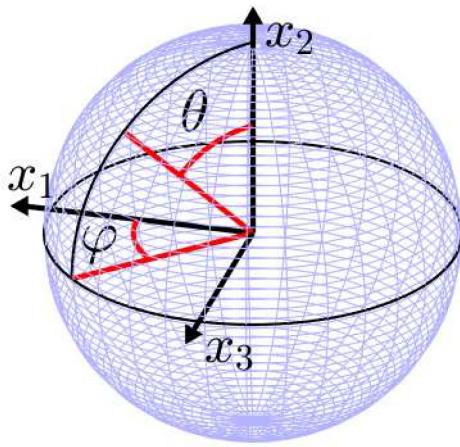


FIG. 4. Picture of the sphere with the spherical coordinates θ and φ .

creates additional difficulties that can be avoided working directly on S^2 in spherical coordinates.

Let us introduce the coordinates (θ, φ) as displayed in Fig. 4 such that

$$x_1 = \sin \theta \cos \varphi, \quad x_2 = \cos \theta, \quad x_3 = \sin \theta \sin \varphi.$$

In these coordinates, the starting point $x(0) = (1, 0, 0)$ and the final point $x(T) = (0, 0, 1)$ become $(\theta, \varphi)(0) = (\pi/2, 0)$, $(\theta, \varphi)(T) = (\pi/2, \pi/2)$. Note that these coordinates are singular for $\theta = 0$ and $\theta = \pi$, but such a singularity does not create any problem, as can be checked by using a second system of coordinates around the singularity. The control system then takes the form

$$\begin{aligned} \dot{\theta} &= -u_1(t) \cos \varphi + u_2(t) \sin \varphi, \\ \dot{\varphi} &= \cot(\theta)[u_1(t) \sin \varphi + u_2(t) \cos \varphi]. \end{aligned} \quad (33)$$

It can be simplified by using the controls v_1 and v_2 defined by

$$\begin{aligned} v_1 &= -u_1 \cos \varphi + u_2 \sin \varphi, \\ v_2 &= u_1 \sin \varphi + u_2 \cos \varphi, \end{aligned} \quad (34)$$

which do not modify the expression of the cost C since $u_1^2 + u_2^2 = v_1^2 + v_2^2$. The control system now becomes

$$\begin{pmatrix} \dot{\theta} \\ \dot{\varphi} \end{pmatrix} = v_1(t) X_1(\theta, \varphi) + v_2(t) X_2(\theta, \varphi),$$

where

$$X_1(\theta, \varphi) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad X_2(\theta, \varphi) = \begin{pmatrix} 0 \\ \cot(\theta) \end{pmatrix}.$$

Let us now apply the PMP. Set $q = (\theta, \varphi)$ and let $p = (p_\theta, p_\varphi)$. The pre-Hamiltonian (14) has the form

$$\begin{aligned} \mathcal{H}(q, p, v, p^0) &= v_1 \langle p, X_1(q) \rangle + v_2 \langle p, X_2(q) \rangle + p^0(v_1^2 + v_2^2) \\ &= v_1 p_\theta + v_2 p_\varphi \cot(\theta) + p^0(v_1^2 + v_2^2). \end{aligned}$$

We consider the steps of Sec. VI C first for abnormal ($p^0 = 0$) and then for normal ($p^0 = -\frac{1}{2}$) extremals.

Step 1. In this step, we have to apply the maximization condition to find the control as a function of q and p . Since the controls are unbounded and the Hamiltonian is concave, the maximization condition is equivalent to

$$\begin{aligned} \frac{\partial \mathcal{H}}{\partial v_1}[q(t), p(t), v(t), p^0] &\equiv 0, \\ \frac{\partial \mathcal{H}}{\partial v_2}[q(t), p(t), v(t), p^0] &\equiv 0. \end{aligned} \quad (35)$$

For abnormal extremals, we obtain

$$\begin{aligned} \langle p(t), X_1[q(t)] \rangle &= p_\theta(t) \equiv 0, \\ \langle p(t), X_2[q(t)] \rangle &= p_\varphi(t) \cot[\theta(t)] \equiv 0. \end{aligned}$$

These conditions do not permit one to obtain the control as a function of q and p . Hence, for this problem, abnormal extremals correspond to singular controls. Since p and p^0 cannot be simultaneously zero, the only possibility to have an abnormal extremal is that $\theta(t) \equiv \pi/2$ on $[0, T]$. In this case, from Eq. (33), we deduce that $\varphi(t)$ should be constant. As a consequence, an abnormal extremal trajectory starting from the initial condition $(\theta, \varphi)(0) = (\pi/2, 0)$ will never reach the final condition $(\theta, \varphi)(T) = (\pi/2, \pi/2)$ and we can disregard these trajectories.

For normal extremals, condition (35) gives

$$\begin{aligned} v_1(t) &= \langle p(t), X_1[q(t)] \rangle = p_\theta(t), \\ v_2(t) &= \langle p(t), X_2[q(t)] \rangle = p_\varphi(t) \cot[\theta(t)]. \end{aligned} \quad (36)$$

Hence, we obtain the controls as a function of q and p and we can conclude that normal extremals correspond to regular controls.

Step 2. Let us insert Eq. (36) into the Hamiltonian equations (i) and (ii) of Theorem 5. We have to consider the case $p^0 = -1/2$ only. We obtain

$$\begin{aligned} \dot{\theta}(t) &= \frac{\partial \mathcal{H}}{\partial p_\theta}[q(t), p(t), v(t), -1/2] \\ &= v_1(t) \\ &= p_\theta(t), \end{aligned} \quad (37)$$

$$\begin{aligned}\dot{p}_\theta(t) &= -\frac{\partial \mathcal{H}}{\partial \theta}[q(t), p(t), v(t), -1/2] \\ &= v_2(t)p_\varphi(t)\{1 + \cot[\theta(t)]^2\} \\ &= p_\varphi(t)^2 \cot[\theta(t)]\{1 + \cot^2[\theta(t)]\},\end{aligned}\quad (38)$$

$$\begin{aligned}\dot{\varphi}(t) &= \frac{\partial \mathcal{H}}{\partial p_\varphi}[q(t), p(t), v(t), -1/2] \\ &= v_2(t) \cot[\theta(t)] \\ &= p_\varphi(t) \cot^2[\theta(t)],\end{aligned}\quad (39)$$

$$\dot{p}_\varphi(t) = -\frac{\partial \mathcal{H}}{\partial \varphi}[q(t), p(t), v(t), -1/2] = 0.\quad (40)$$

Equation (40) tells us that p_φ is a constant of the motion, denoted by a in the following. We are then left with the differential equations

$$\dot{\theta} = p_\theta, \quad \dot{p}_\theta = a^2 \cot(\theta)[1 + \cot^2(\theta)].\quad (41)$$

Once these are solved, φ is obtained by integrating in time Eq. (39), which now has the form

$$\dot{\varphi} = a \cot^2(\theta).\quad (42)$$

Equations (41) and (42) should be solved for every value of $a \in \mathbb{R}$ with the initial conditions

$$\theta(0) = \pi/2, \quad \varphi(0) = 0, \quad p_\theta(0) = \pm 1.$$

The last condition comes from the property that the maximized Hamiltonian is now fixed to $\frac{1}{2}$ (corresponding to the choice of taking T in such a way that optimal trajectories are parameterized by arc length). More precisely,

$$\begin{aligned}\frac{1}{2} &= \mathcal{H}[q(t), p(t), v(t), -\frac{1}{2}] \\ &= v_1(t)p_\theta(t) + v_2(t)a \cot[\theta(t)] \\ &\quad - \frac{1}{2}[v_1(t)^2 + v_2(t)^2] \\ &= \frac{1}{2}\{p_\theta(t)^2 + a^2 \cot^2[\theta(t)]\}.\end{aligned}$$

Requiring this condition at $t = 0$, one gets $p_\theta(0) = \pm 1$.

The system of Eqs. (41) can be solved by again using the fact that the maximized Hamiltonian is equal to $\frac{1}{2}$, which

implies that

$$\dot{\theta}(t)^2 = 1 - a^2 \cot^2[\theta(t)].$$

Using a separation of variables, we arrive [with the initial condition $\theta(0) = \pi/2$] at

$$\theta(t) = \begin{cases} \arccos\left(\frac{\sin(\sqrt{1+a^2}t)}{\sqrt{1+a^2}}\right) & \text{if } p_\theta(0) = -1, \\ \pi - \arccos\left(\frac{\sin(\sqrt{1+a^2}t)}{\sqrt{1+a^2}}\right) & \text{if } p_\theta(0) = 1. \end{cases}$$

These expressions lead to a simple formula for $x_2(t)$, namely,

$$x_2(t) = \pm \left(\frac{\sin(\sqrt{1+a^2}t)}{\sqrt{1+a^2}} \right).\quad (43)$$

As already explained, the expression for φ can be obtained by integrating in time Eq. (42) using the expression of $\theta(t)$ with initial condition $\varphi(0) = 0$. The final result is more easily expressed in Cartesian coordinates:

$$\begin{aligned}x_1(t) &= \frac{a \sin(at) \sin(\sqrt{1+a^2}t)}{\sqrt{1+a^2}} \\ &\quad + \cos(at) \cos(\sqrt{1+a^2}t),\end{aligned}\quad (44)$$

$$\begin{aligned}x_3(t) &= \sin(at) \cos(\sqrt{1+a^2}t) \\ &\quad - \frac{a \sin(\sqrt{1+a^2}t) \cos(at)}{\sqrt{1+a^2}}.\end{aligned}\quad (45)$$

The corresponding controls can be obtained via Eqs. (37), (38), and (34), or from Eq. (29), providing

$$\begin{aligned}u_1(t) &= -\dot{x}_1(t)/x_2(t) = \pm \cos(at), \\ u_2(t) &= \dot{x}_3(t)/x_2(t) = \mp \sin(at).\end{aligned}$$

Step 3. In this step, we have to find the initial covector [i.e., $p_\theta(0) \in \{-1, +1\}$ and $a = p_\varphi \in \mathbb{R}$] whose corresponding trajectory arrives at the final target $(x_1, x_2, x_3)(T) = (0, 0, 1)$.

From Eq. (43), requiring that $x_2(T) = 0$ we get $\sin(\sqrt{1+a^2}T) = 0$. Then, from Eq. (44), requiring that $x_1(T) = 0$ we arrive at $\cos(aT) = 0$. We then have the conditions

$$\sqrt{1+a^2}T = n_1\pi, \quad aT = \frac{\pi}{2} + n_2\pi,$$

with $n_1, n_2 \in \mathbb{Z}$. Note that $n_1 > 0$ since $T > 0$, and, hence,

$$\frac{a}{\sqrt{1+a^2}} = \frac{n_2 + 1/2}{n_1},$$

from which we deduce that $|n_2 + \frac{1}{2}| < n_1$. It follows that

$$a = \frac{(n_2 + 1/2)/n_1}{\sqrt{1 - [(n_2 + 1/2)/n_1]^2}}. \quad (46)$$

The target is reached at time

$$T = \pi n_1 \sqrt{1 - \left(\frac{n_2 + 12}{n_1}\right)^2}. \quad (47)$$

Step 4. The previous step provided a discrete set of trajectories reaching the target. Those arriving in a shorter time correspond to $n_1 = 1$ and $n_2 = 0$ or $n_2 = -1$, for which $T = \pi\sqrt{3}/2$ and $a = \pm 1/\sqrt{3}$. The final expressions for the optimal trajectories and optimal controls are

$$\begin{aligned} x_1(t) &= \cos^3(t/\sqrt{3}), \\ x_2(t) &= \pm \frac{\sqrt{3}}{2} \sin\left(\frac{2t}{\sqrt{3}}\right), \\ x_3(t) &= -\varepsilon \sin^3(t/\sqrt{3}), \end{aligned}$$

and

$$\begin{aligned} u_1(t) &= \pm \cos(t/\sqrt{3}), \\ u_2(t) &= \mp \varepsilon \sin(t/\sqrt{3}), \end{aligned}$$

where $\varepsilon = \pm 1$ is the sign of a .

We point out that the two trajectories arriving at time T at the point $(0, 0, 1)$ are those corresponding to $n_2 = 0$ (i.e., to $a = 1/\sqrt{3}$) and to $p_\theta(0) = \mp 1$ [i.e., to the \pm sign in Eq. (43)], while the two trajectories arriving at time T at the point $(0, 0, -1)$ are associated with $n_2 = -1$ (i.e., to $a = -1/\sqrt{3}$). See Sec. VIII A. The four trajectories obtained in this way are optimal since we know that optimal trajectories exist and those that we have selected are the best among all trajectories satisfying a necessary condition for optimality. In this example, note that we can select the optimal trajectories by applying the PMP and finding by hand the best extremals, without using second-order conditions nor other sufficient conditions for optimality.

Figures 5 and 6 respectively display the two symmetric extremal trajectories reaching the target state $(0, 0, 1)$ at time T and the time evolution of the corresponding controls.

IX. EXAMPLE 2: A MINIMUM TIME TWO-LEVEL QUANTUM SYSTEM WITH A REAL CONTROL

A. Formulation of the control problem

We consider the control of a spin-1/2 particle whose dynamics is governed by the Bloch equation in a given

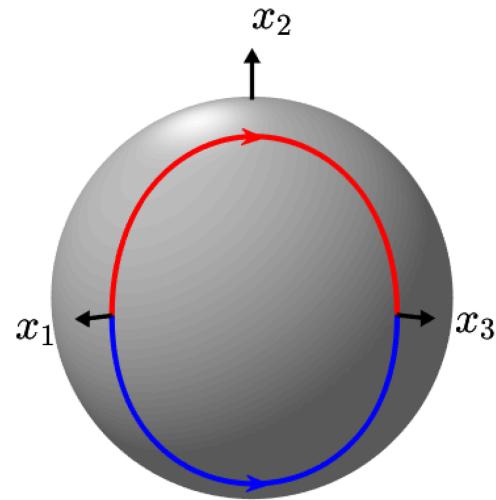


FIG. 5. Plot of the two extremal trajectories (in red and blue) on the sphere going from the point $(1, 0, 0)$ to the point $(0, 0, 1)$ and minimizing the cost functional C .

rotating frame [98,99]:

$$\begin{aligned} \dot{M}_x &= -\omega M_y, \\ \dot{M}_y &= \omega M_x - \omega_x(t) M_z, \\ \dot{M}_z &= \omega_x(t) M_y. \end{aligned}$$

Here $\mathbf{M} = (M_x, M_y, M_z)$ is the magnetization vector and ω is the offset term. The system is controlled through a single magnetic field along the x axis that satisfies the constraint $|\omega_x| \leq \omega_{\max}$. We introduce normalized coordinates $(x, y, z) = \mathbf{M}/M_0$, where M_0 is the thermal equilibrium magnetization, a normalized control $u = \omega_x/\omega_{\max}$ that satisfies the constraint $|u| \leq 1$, and a normalized time $\tau = \omega_{\max} t$ (denoted t below). Dividing the previous system by $\omega_{\max} M_0$, we find that the time evolution of the normalized

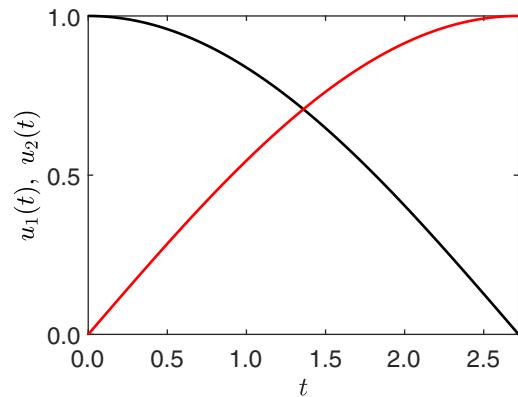


FIG. 6. Time evolution of the controls u_1 (in black) and u_2 (in red). The control time T and the parameter a are respectively set to $\pi\sqrt{3}/2$ and $-1/\sqrt{3}$.

coordinates is given by the equations

$$\begin{pmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \end{pmatrix} = \begin{pmatrix} -\Delta y \\ \Delta x \\ 0 \end{pmatrix} + u(t) \begin{pmatrix} 0 \\ -z \\ y \end{pmatrix},$$

where $\Delta = \omega/\omega_{\max}$ is the normalized offset. The trajectories of the system lie on the Bloch sphere defined by the equation $x^2 + y^2 + z^2 = 1$. The manifold M is here the sphere S^2 . The differential system can be written in a more compact form as

$$\dot{q} = (\mathcal{F} + u\mathcal{G})q, \quad (48)$$

where $q = (x, y, z)$ is the state of the system, $u(t) \in U = [-1, 1]$, and \mathcal{F} and \mathcal{G} are the skew-symmetric 3×3 matrices

$$\mathcal{F} = \begin{pmatrix} 0 & -\Delta & 0 \\ \Delta & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \mathcal{G} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}.$$

The vector fields $\mathcal{F}q$ and $\mathcal{G}q$ generate rotations respectively around the z and x axes.

Existence of time-optimal trajectories. By point 1 of Proposition 9, any initial point q_0 on the Bloch sphere M can be connected by an admissible trajectory of the control system to any other point q_1 on the Bloch sphere (see also Ref. [86]). In this case the existence of a time-optimal trajectory connecting q_0 to q_1 is a direct consequence of Proposition 12 below or Proposition 3. Indeed, $U = [-1, 1]$ is compact and the set $\{\mathcal{F}q + u\mathcal{G}q \mid u \in U\}$ is convex for any point q of the Bloch sphere.

B. Application of the PMP

The goal of the two control processes that we are considering is to steer the system in a minimum time from the north pole $(0, 0, 1)$ of the Bloch sphere to the south pole $(0, 0, -1)$ or the state $(1, 0, 0)$, respectively. We follow here the results established in Refs. [48, 100] (see also Ref. [101] for an experimental implementation).

The time-optimal control problem is solved by the application of the PMP. Here the cost C to minimize is

$$C = \int_0^T dt = T \rightarrow \min,$$

where T is free. The pre-Hamiltonian can be expressed as

$$\mathcal{H}(q, p, u, p^0) = p(\mathcal{F} + u\mathcal{G})q + p^0,$$

where $p = (p_x, p_y, p_z) \in \mathbb{R}^3$ is the covector and p^0 is a non-positive constant such that p and p^0 are not simultaneously equal to 0. (Here p is seen as a row vector and q as a column vector.) The value of \mathcal{H} is constantly equal to 0

since the final time is free. The PMP states that the optimal trajectories are solutions of the equations

$$\begin{aligned} \dot{q}(t) &= \frac{\partial \mathcal{H}}{\partial p}[q(t), p(t), u(t), p^0], \\ \dot{p}(t) &= -\frac{\partial \mathcal{H}}{\partial q}[q(t), p(t), u(t), p^0], \\ \mathcal{H}[q(t), p(t), u(t), p^0] &= \max_{|v| \leq 1} \mathcal{H}[q(t), p(t), v, p^0]. \end{aligned}$$

The dynamics of the adjoint state p is given by

$$\dot{p}(t) = -p(t)[\mathcal{F} + u(t)\mathcal{G}]. \quad (49)$$

Note that $d\|p(t)\|^2/dt = 0$, so that $\|p(t)\|$ does not depend on time. Its constant value is nonzero since $\mathcal{H} = 0$ and $(p, p_0) \neq 0$.

Steps 1 and 2. Since the only term of the pre-Hamiltonian depending on the control is $up\mathcal{G}q$, the maximization condition of the PMP leads to the introduction of the switching function

$$\Phi(t) = p(t)\mathcal{G}q(t).$$

In the regular case in which $\Phi(t) \neq 0$, we deduce from the maximization condition that the optimal control is given by the sign of Φ , $u(t) = \text{sign}[\Phi(t)]$. The corresponding trajectory is called a *bang trajectory*. If Φ has an isolated zero in a given time interval then the control function may *switch* from -1 to 1 or from 1 to -1 . A *bang-bang trajectory* is a trajectory obtained after a finite number of switches.

Using Eq. (49), we have

$$\dot{\Phi}(t) = p(t)[\mathcal{G}, \mathcal{F}]q(t), \quad (50)$$

where $[\cdot, \cdot]$ denotes the matrix commutator operator. In particular, Φ is a C^1 function and, for almost every t ,

$$\ddot{\Phi}(t) = p(t)[\mathcal{F}, [\mathcal{F}, \mathcal{G}]]q(t) + u(t)p(t)[\mathcal{G}, [\mathcal{F}, \mathcal{G}]]q(t).$$

Since $[\mathcal{F}, [\mathcal{F}, \mathcal{G}]] = -\Delta^2\mathcal{G}$ and $[\mathcal{G}, [\mathcal{F}, \mathcal{G}]] = \mathcal{F}$, we have, for a.e. t ,

$$\begin{aligned} \ddot{\Phi}(t) &= -\Delta^2\Phi(t) + u(t)p(t)\mathcal{F}q(t) \\ &= -\Omega^2\Phi(t) - p^0u(t), \end{aligned} \quad (51)$$

where $\Omega = \sqrt{1 + \Delta^2}$ and the second equality follows from the identity $\mathcal{H} = 0$.

Abnormal extremals. Abnormal extremals are characterized by the equality $p^0 = 0$, from which, together with $\mathcal{H} = 0$, we deduce that $\Phi(t) = -p(t)\mathcal{F}q(t)/u(t)$. If $\Phi(t) = 0$ at time t then $p(t)$ is orthogonal both to $\mathcal{F}q(t)$ and $\mathcal{G}q(t)$. If, moreover, t were not an isolated zero of Φ then $\dot{\Phi}(t) = 0$, since Φ is C^1 . It would follow from Eq.

(50) that $p(t)$ is also orthogonal to $[\mathcal{G}, \mathcal{F}]q(t)$. Since, for every $q \in M$, the vectors $\mathcal{F}q$, $\mathcal{G}q$, and $[\mathcal{G}, \mathcal{F}]q$ span the tangent plane $T_q M$ to the sphere M , we would deduce that $p(t) = 0$, contradicting the PMP. This means that abnormal extremals are necessarily bang-bang extremals. Moreover, we deduce from Eq. (51) that the switching times are the zeros of a nontrivial solution of the equation

$$\ddot{\Phi} + \Omega^2\Phi = 0.$$

The length of an arc between any two successive switching times is then equal to π/Ω .

Singular arcs. When the trajectory is normal, there might exist extremals for which Φ is zero on a nontrivial time interval. The control is singular on such an interval, since it cannot directly be obtained from the maximization condition. The restriction of the trajectory to an interval on which $\Phi \equiv 0$ is called a *singular arc*. Singular arcs are characterized by the fact that the time derivatives of Φ at all orders are zero. Since $p(t)$ is different from zero, the only possibility to have simultaneously $\Phi(t) = 0$ and $\dot{\Phi}(t) = 0$ is that the vectors $\mathcal{G}q(t)$ and $[\mathcal{G}, \mathcal{F}]q(t)$ are parallel. Since $\mathcal{G}q(t)$ and $[\mathcal{G}, \mathcal{F}]q(t)$ respectively generate the rotations around the x and the y axes, we deduce that singular arcs are contained in the equator $z = 0$ of the sphere. The singular control law u_s can be calculated from Eq. (51) by enforcing that Φ and its second time derivative are zero, yielding $u_s(t) = 0$. As could be expected, this control law generates a rotation along the equator. It is admissible because $|u_s| \leq 1$.

Normal bang-bang extremals. Consider a normal extremal and an interior bang arc of duration T between the switching times t_0 and $t_0 + T$ on which the control u is constantly equal to $+1$ or -1 . An interior arc is an arc that is neither at the beginning nor at the end of the extremal. We normalize p_0 to -1 . According to Eq. (51), the function $\Psi = \Phi - u/\Omega^2$ is a solution of $\ddot{\Psi} + \Omega^2\Psi = 0$. Moreover, since Φ is nonconstant, then Ψ is nontrivial. Hence, $\Psi(t) = v \cos(\Omega t + \theta_0)$ for some $v > 0$ and $\theta_0 \in \mathbb{R}$. Moreover, $v > 0$ is uniquely identified by u and $\dot{\Phi}(t_0)$ through the equalities $\Psi(t_0) = -u/\Omega^2$ and $\dot{\Psi}(t_0) = \dot{\Phi}(t_0)$. Switchings occur if $\Phi = \Psi + u/\Omega^2$ vanishes and changes sign. Since, moreover, $\text{sign}[u] = \text{sign}[\Phi]$, it follows that Ψ is larger than the negative value $-1/\Omega^2$ on $(t_0, t_0 + T)$ when $u = +1$ and smaller than the positive value $1/\Omega^2$ on $(t_0, t_0 + T)$ when $u = -1$. Hence, T is larger than π/Ω and

$$\dot{\Phi}(t_0 + T) = -\dot{\Phi}(t_0).$$

If $\dot{\Phi}(t_0) = 0$, we deduce that $T = 2\pi/\Omega$. Note that if u is constantly equal to $+1$ or -1 then the solutions of $\dot{q} = (\mathcal{F} + u\mathcal{G})q$ are $2\pi/\Omega$ periodic rotations around the axis spanned by $(u, 0, \Delta)$. Since a time-optimal trajectory cannot self-intersect, we conclude that $T < 2\pi/\Omega$ and $\dot{\Phi}(t_0) \neq 0$.

If $t_0 + T$ is the starting time of another internal bang arc then by the above considerations the duration of such an internal bang arc is also equal to T . Given a normal bang-bang trajectory, there then exists $T \in (\pi/\Omega, 2\pi/\Omega)$ such that the trajectory is the concatenation of bang arcs of duration T , except possibly for the first and last bang arcs, whose lengths can be smaller than T .

General extremals. As we have seen in the previous paragraphs, if a trajectory contains an internal bang arc then it is a bang-bang trajectory. Otherwise the set of zeros of Φ is connected, that is, either Φ has a single zero or it vanishes on a nontrivial singular arc and is different from zero out of it.

To summarize, extremal trajectories are of two types.

- Bang-bang trajectories whose internal bang arcs all have the same length $T \in [\pi/\Omega, 2\pi/\Omega]$ (the case $T = \pi/\Omega$ corresponding to abnormal extremals) and for which the first and last bang arcs have lengths at most T .
- Concatenations of a (possibly trivial) bang arc of length smaller than $2\pi/\Omega$, a singular arc on which $u_s = 0$, and another (possibly trivial) bang arc of length smaller than $2\pi/\Omega$.

Steps 3 and 4. We solve in this paragraph two time-optimal control problems. Starting from the north pole $(0, 0, 1)$, the goal is to reach in a minimum time the points $(0, 0, -1)$ [problem (P1)] and $(1, 0, 0)$ [problem (P2)]. To simplify the derivation of the optimal solutions, we assume that $|\Delta| \leq 1$ [48].

Before solving (P1) and (P2), we first derive analytical results describing the dynamics of the system. Consider a bang extremal trajectory starting from the north pole at time $t = 0$ with control $u(t) = \varepsilon = \pm 1$. The corresponding trajectory is given by

$$\begin{aligned} x(t) &= \frac{\varepsilon\Delta}{\Omega^2}[1 - \cos(\Omega t)], \\ y(t) &= -\frac{\varepsilon}{\Omega}\sin(\Omega t), \\ z(t) &= 1 + \frac{1}{\Omega^2}[\cos(\Omega t) - 1]. \end{aligned}$$

The first two times for which $z(t) = 0$ are $t_1 = (1/\Omega)[\pi - \arccos(\Delta^2)]$ and $t_2 = (1/\Omega)[\pi + \arccos(\Delta^2)]$. Note that all other times for which $z(t) = 0$ are larger than $2\pi/\Omega$ and cannot be the duration of a bang arc of an optimal trajectory.

The optimal solution of (P1) is a bang-bang trajectory with a first switch on the equator at $t = t_1$. The total duration of the process is $t_1 + t_2 = 2\pi/\Omega$. A symmetric configuration is possible with a first switch at $t = t_2$. The two trajectories are displayed in Fig. 7. Let us discuss how the optimality of such a trajectory can be asserted. The

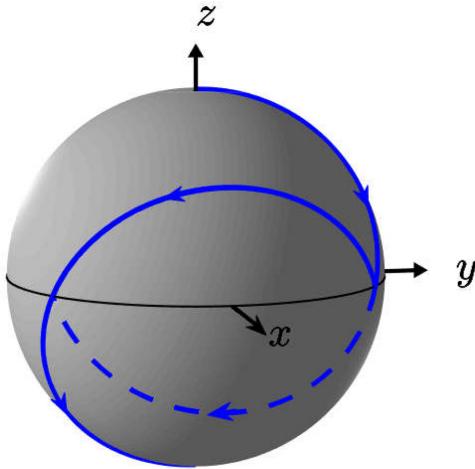


FIG. 7. Optimal trajectories (in blue) going from the north pole to the south pole of the Bloch sphere. The solid black line indicates the position of the equator. The parameter Δ is set to -0.5 . For one of the two optimal trajectories, a control $u = -1$ is first applied during a time t_1 , followed by a control $u = +1$ during a time t_2 , while for the other trajectory, $u = -1$ lasts for a time t_2 and $u = +1$ for a time t_1 .

proposed trajectory is clearly extremal and connects the chosen initial and final points. Since a bang-bang trajectory with at least two internal bang arcs has duration larger than $2\pi/\Omega$, it follows that any bang-bang trajectory with four or more bang arcs has a duration larger than the candidate optimal trajectory (that is, $2\pi/\Omega$). One is then left to compare $2\pi/\Omega$ with finitely many types of trajectories: bang-bang trajectories with two or three bang arcs, and trajectories obtained by concatenation of a bang arc, a singular arc, and a bang arc. By setting the initial and final points, this leaves few competitors to the optimal trajectory, which can be easily excluded by enumeration. As an example, we consider the concatenation of a bang arc with $u = -1$, a singular arc, and a new bang arc with $u = +1$. The two bang arcs last for a time t_1 , while the duration of the singular arc is $(2/\Delta)\arctan(\sqrt{1-\Delta^2}/\Delta)$. It is then straightforward to deduce that the duration of this extremal solution is larger than $t_1 + t_2 = 2\pi/\Omega$, except in the case $\Delta = 1$ for which the singular arc is of length zero.

Let us now discuss the solution of (P2). Using the results of Sec. IX B, we consider the concatenation of a bang extremal with $u = +1$ during the time t_1 and of a singular extremal with $u = 0$ during the time t_s . At time $t = t_1$, the trajectory reaches the point $(\Delta, -\sqrt{1-\Delta^2}, 0)$. We deduce that $t_s = (1/\Delta)\arctan(\sqrt{1-\Delta^2}/\Delta)$. The total duration of the control process is $t_1 + t_s$. The corresponding trajectory is represented in Fig. 8. One can check that all other candidates for optimality join $(1, 0, 0)$ in a longer time. The situation is more complicated than for (P1), since here $t_s \rightarrow \infty$ as $\Delta \rightarrow 0$, so the candidate trajectory should be compared with trajectories with more and

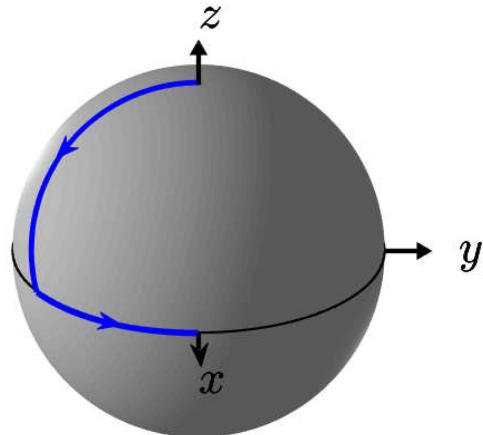


FIG. 8. Optimal trajectory (in blue) going from the north pole to the point $(1, 0, 0)$ of the x axis. The solid black line indicates the position of the equator. The parameter Δ is set to 0.5 .

more bangs as $\Delta \rightarrow 0$. A proof of the optimality of the trajectory described above can be obtained, for instance, using optimal synthesis theory, i.e., describing all the optimal trajectories starting from the north pole, as done in Ref. [48].

X. APPLICATIONS OF QUANTUM OPTIMAL CONTROL THEORY

We discuss in this section the link between the material and results presented in this tutorial and the current research objectives in this field, both from a theoretical and an experimental point of view. The ability to quickly and accurately perform operations in a quantum device is a key task in quantum technologies. Quantum optimal control theory addresses this challenge by combining analytical and numerical tools to design procedures adapted to the experimental setup under study. This approach has the key advantage of being based on a rigorous theoretical framework on which all developments have been based since the 1980s. In the case of a low-dimensional quantum system, as in the two examples presented in Secs. VIII and IX, the optimal control problem may be solved analytically or at least with a very high numerical accuracy. For high-dimensional systems, an efficient alternative is provided by numerical optimal control procedures, which include first- and second-order gradient ascent algorithms [45, 66, 102, 103] whose structure is based on the PMP. The connection between the PMP and gradient-based optimization algorithms is described in Sec. VII. Their flexibility allows them to adapt to many experimental situations for which precise modeling of the dynamics is known. As concrete and recent examples, we mention the experimental implementation of such techniques in Rydberg atoms [104, 105], in spin-wave states of atomic ensembles [106], in electron spin resonance [107], or in superconducting

cavity resonator [108]. The diversity of these examples shows the key role that optimal control now plays in quantum technologies. We describe below some examples of recent applications that are based on the mathematical formalism described in this tutorial. We also indicate some open issues in this field. We stress that our aim is not to provide an exhaustive list of all optimal control studies in this very active area (we refer the interested reader to recent reviews on this subject [2,4–6]), but rather to give an overview of the different aspects that can be analyzed.

The analytical techniques have been applied recently to a series of fundamental and practical issues in quantum technologies such as robustness [109–115] and selectivity [116,117] of the control process with respect to parameters of the system Hamiltonian. The robustness property is a key factor for the experimental implementation of open-loop optimal pulses. The basic idea consists in simultaneously controlling an ensemble of identical systems that differ only by the value of the unknown parameter. Robust or selective optimal pulses are respectively obtained in the case where the target states are the same or different according to the systems. In the same direction, numerical optimal control has also been employed to improve the estimation of Hamiltonian parameters [118,119]. Here, from a specific cost functional, the control is used to optimize the discrimination between the system dynamics, which leads to a precise characterization of the Hamiltonian [120]. Using the same type of approach, optimal control has proven to be very effective for quantum sensing [23,121,122].

A key aspect of quantum technologies is the minimum time to perform a given operation. Lower bounds on the time can be established in the framework of quantum speed limits [13,123] where the time is expressed as a ratio between the distance to the target state and the dynamical speed of evolution. Time-optimal trajectories correspond by definition to the speed limit of the processes. The solutions of the two examples solved in this tutorial therefore give the quantum speed limits of the problems under study. However, in a higher-dimensional quantum system, the many local minima of the control landscape make it very difficult to find a good approximation of the optimal trajectory and generally lead to an upper bound of the minimum time. Recent studies have explored the link between these two formalisms [54,124]. A major step forward would be to rigorously describe the concept of speed limit in terms of the Pontryagin maximum principle [125,126].

Quantum thermodynamics is nowadays a very useful tool to describe the dynamics of open quantum systems used in quantum technologies [127]. Optimal control and the PMP can be applied to analyze many issues in this field, such as the enhancement of quantum engine performances [128,129], work extraction [130,131], or qubit purification [132–134]. The quantification of the resources to control

a quantum system has also been discussed and could be optimized from the PMP [135].

An intense effort is being made to develop new quantum optimal control algorithms, which are well suited to experimental limitations and constraints [136–140] or which are particularly effective in designing optimal pulses [67,141,142]. In particular, hybrid optimal control using feedback from the experiment has been developed to avoid problems due to inaccuracies in the modeling of the dynamics. The structure of the control landscape for the preparation of quantum many-body systems has been identified and exhibits a spin-glass-like phase transition in the space of control protocols close to the optimal fidelity [143]. This structure explains the difficulty in numerically finding the optimal solution in this region. Combining optimal control and machine learning approaches as proposed in Ref. [144] could be a way to solve this problem and to improve the efficiency of optimal algorithms [145,146]. However, we stress that a rigorous mathematical description of these numerical methods and results will be necessary in the future to systematically apply these approaches to other quantum devices.

We conclude this section by pointing out that the PMP may also play a more unexpected role in quantum computing. It has recently been shown that Grover's quantum search problem can be mapped to a time-optimal control problem, and then described through the PMP [147]. A connection between optimal control theory and variational quantum algorithms [148] has been established in Ref. [149]. A general description of this intrinsic link is given in Ref. [150] where optimal control is used to precisely adjust the parameters of a quantum circuit. These results highlight that the PMP as a general optimization tool is not only interesting for the computation of time-dependent control pulses, but also in other optimization problems of interest in quantum technologies.

XI. CONCLUSION

In this tutorial, we have attempted to give the reader a minimal background on the mathematical techniques of OCT. In our opinion, this is a fundamental prerequisite to rigorously and correctly apply these tools in quantum control.

The objectives of the tutorial are twofold. First, we have highlighted the key concepts of the PMP using ideas based on the finding of extrema of functions of several variables. This analogy gives nonexperts an intuition of the tools of optimal control that might seem abstract on first reading. We have then stated the PMP and described in detail the different steps to be followed in order to solve an optimal control problem. Some are rarely discussed in quantum control, such as the existence of solutions or abnormal and singular extremals, while they play a crucial role in

some problems. The link between the PMP and gradient-based optimization algorithms has been explained, which also highlights the role of such mathematical tools in any numerical optimization process. Second, we have solved two basic control problems, namely the control of a three-level quantum system by means of two complex resonant fields and the control of a spin-1/2 particle through a real off-resonance driving. The low dimension of the two systems allows us to analytically express the optimal solutions and to give a complete geometric description of the control protocol. Such examples can be used as a starting point by the reader to apply the PMP to more complex quantum systems. A series of recent results using OCT has briefly been discussed to highlight current research directions in this area.

XII. LIST OF NOTATION

We provide in Table III a list of the main notation used in the paper as well as the first section in which they appear or are defined.

ACKNOWLEDGMENTS

The authors have contributed equally to this work. This research is partially supported by the ANR projects “SRGI” ANR-15-CE40-0018 and “QUACO” ANR-17-CE40-0007-01. This project has received funding from the European Union Horizon 2020 research and innovation program under the Marie Skłodowska-Curie Grant No. 765267 (QUSCO). We gratefully acknowledge useful discussions with Professor Christiane Koch, who suggested that we write this tutorial and provided advice throughout the project. We thank the PhD students J. Fischer,

L. Kerber, M. Krauss, J. Langbehn, A. Marshall, and F. Roy for their helpful comments.

APPENDIX A: TEST OF CONTROLLABILITY

We give in this section a general and useful sufficient condition for controllability. This section completes the discussion of Sec. III.

Proposition 9. *Let M be a smooth manifold, and let U be a subset of \mathbb{R}^m containing a neighborhood of the origin. Consider a control system of the form $\dot{q} = F_0(q) + \sum_{j=1}^m u_j(t)F_j(q)$, with F_0, \dots, F_m smooth vector fields on M and $u(t) = [u_1(t), \dots, u_m(t)] \in U$. Let \mathcal{L}_0 be the Lie algebra generated by the vector fields F_0, \dots, F_m and \mathcal{L}_1 be the Lie algebra generated by the vector fields F_1, \dots, F_m . The system is controllable if at least one of the following conditions is satisfied:*

1. *F_0 is a recurrent vector field and $\dim[\mathcal{L}_0(q)]$ is equal to the dimension of M at every $q \in M$;*
2. *$U = \mathbb{R}^m$ and $\dim[\mathcal{L}_1(q)]$ is equal to the dimension of M at every $q \in M$.*

For a definition of Lie algebra generated by a set of vector fields and for the notion of a recurrent vector field, we refer the reader to Refs. [28,151]. For our purposes, it is sufficient to recall that

- (a) if $F_0(q) = A_0q, \dots, F_m(q) = A_mq$ are linear vector fields then $G \in \mathcal{L}_0$ if and only if $G(q) = Bq$ with B in the matrix Lie algebra (for the commutator product) generated by A_0, \dots, A_m ;
- (b) if a vector field F is such that every solution of $\dot{q} = F(q)$ is periodic then F is recurrent.

TABLE III. Main notation used in the paper.

Notation	Definition	Section
$q(t)$	The state of the system at time t	III
f	A smooth vector-valued function of q and u that defines the control system	III
f_0	A smooth scalar-valued function of q and u that defines the cost functional	III
M	The manifold on which the state $q(t)$ evolves	III
T	The final control time	III
$u(t)$	The value of the control law at time t	III
U	The set of possible values of $u(t)$	III
\mathcal{U}	The set of admissible control laws	III
\mathcal{T}	The set of target states	III
$d[\mathcal{T}, q(T)]$	The distance between \mathcal{T} and the final state $q(T)$	III
$\mathcal{R}(q_{\text{in}})$	The set of reachable states from q_{in}	III
$\mathcal{R}^T(q_{\text{in}})$	The set of reachable states in time T from q_{in}	III
$\mathbf{F}(q)$	The set of directions $f(q, u)$ at q as u varies in U	V
$\phi[q(T)]$	The terminal cost that depends on the final state $q(T)$	V
$J[u(\cdot)]$	The value of the cost for a control $u(\cdot)$	VI
\mathcal{H}	The pre-Hamiltonian in the PMP	VI
p	The adjoint state in the PMP	VI
p_0	The abnormal multiplier	VI

With these sufficient conditions, we recover the standard controllability conditions introduced in quantum control for closed systems [85,86] and recalled in Sec. III.

APPENDIX B: FILIPPOV'S THEOREM

We state here Filippov's theorem, which gives a sufficient condition for the compactness of the reachable set [26]. This section completes the discussion of Sec. V.

Theorem 10 *Filippov. Fix $T > 0$. Consider the control system $\dot{q}(t) = f[q(t), u(t)]$, $q \in M$, $u \in \mathcal{U}$, where M is a n -dimensional manifold and $U \subset \mathbb{R}^m$. Fix an initial condition $q_{\text{in}} \in M$. Assume that the following conditions hold:*

- the set U is compact,
- the set $\mathbf{F}(q) = \{f(q, u) \mid u \in U\}$ is convex for every $q \in M$,
- for every $u \in \mathcal{U}$, the solution of $\dot{q}(t) = f[q(t), u(t)]$, $q(0) = q_{\text{in}}$, is defined on the whole interval $[0, T]$.

Then the sets $\mathcal{R}^T(q_{\text{in}})$ and $\mathcal{R}^{\leq T}(q_{\text{in}})$ are compact.

Here $\mathcal{R}^{\leq T}(q_{\text{in}})$ denotes the reachable set from q_{in} within time T , i.e., the set of points \bar{q} in M such that there exist $T' \in [0, T]$ and an admissible trajectory $q : [0, T'] \rightarrow M$ such that $q(0) = q_{\text{in}}$ and $q(T') = \bar{q}$.

Note that the third hypothesis of Theorem 10 is automatically satisfied when M is compact.

By applying Filippov's theorem to the augmented system for Problem P1 introduced in Sec. V, we obtain the following result.

Proposition 11. *Fix $T > 0$. Assume that*

- \mathcal{T} is closed and $\mathcal{R}^T(q_{\text{in}}) \cap \mathcal{T} \neq \emptyset$,
- the set U is compact,
- the set

$$\hat{\mathbf{F}}(q) = \left\{ \begin{pmatrix} f^0(q, u) \\ f(q, u) \end{pmatrix} \mid u \in U \right\}$$

is convex for every $q \in M$,

- for every $u \in \mathcal{U}$, the solution of $\dot{q}(t) = f[q(t), u(t)]$, $q(0) = q_{\text{in}}$, is defined on the whole interval $[0, T]$.

Then there exists a solution to problem P1.

Example 5. The conclusion of Proposition 11 may fail to hold if we drop the convexity assumption on the set $\mathbf{F}(q)$ of admissible velocities, as illustrated by the following example. Take $M = \mathbb{R}^2$, $U = \{-1, 1\}$, and $f(q, u) = A_u q$

with

$$A_1 = \begin{pmatrix} -1 & 1 \\ -1 & 0 \end{pmatrix} \quad \text{and} \quad A_{-1} = \begin{pmatrix} -1 & -1 \\ 1 & 0 \end{pmatrix}.$$

We point out that the corresponding control system satisfies all the assumptions of Theorem 10 except for the convexity of $\mathbf{F}(q)$, since the control can only take two discrete values -1 and 1 . Pick $q_{\text{in}} = (1, 0)$ and any $T > 0$. Then $q_T := e^{-T} q_{\text{in}}$ is in the closure of $\mathcal{R}^T(q_{\text{in}})$, since we can end up arbitrarily close to q_T at time T by applying a control that switches fast enough between -1 and 1 . On the other hand, if $\dot{q}(t) = f[q(t), u(t)]$ at time t and $q(t)$ has a nonzero vertical component, then $d\|q(t)\|/dt > -\|q(t)\|$. Since every trajectory of the control system starting at q_{in} necessarily leaves the horizontal axis, its final point at time T has norm larger than e^{-T} . Consequently, $q_T \notin \mathcal{R}^T(q_{\text{in}})$. This proves that $\mathcal{R}^T(q_{\text{in}})$ is not closed; hence, the conclusion of Theorem 10 does not hold. This result is illustrated in Fig. 9 that numerically shows that the distance to the target state tends to zero when the number of switches goes to $+\infty$. This example also highlights possible numerical problems when the existence of the optimal solution is not verified. Here, an optimization algorithm cannot converge towards a well-defined control law.

The idea of reducing the problem of existence of an optimal control to the compactness of the reachable set of the augmented system can be used for more general problems. For instance, if we add a terminal cost $\phi[q(T)]$ to the cost $\int_0^T f^0[q(t), u(t)] dt$, where ϕ is a smooth function (as, for instance, in Approach B, or in the general formulation given in Sec. VIB), we get a similar result adding to $f^0[q(t), u(t)]$ the directional derivative of ϕ along $f[q(t), u(t)]$, that is, replacing $f^0[q(t), u(t)]$ by $f^0[q(t), u(t)] + \langle d\phi[q(t)], f[q(t), u(t)] \rangle$. We recall that here $d\phi[q(t)]$ denotes the differential of ϕ evaluated at $q(t)$ and that $\langle \cdot, \cdot \rangle$ is the duality product between covectors of T^*M and vectors of TM .

When the final time is free, it is more difficult to get the existence of optimal trajectories. However, the compactness of $\mathcal{R}^{\leq T}(q_{\text{in}})$ in the Filippov theorem can be used to find conditions for the existence of optimal controls in minimum time. We state this result in the case where M is compact. Note that the problem of minimizing time can be written in the form of Problem P1 with T free and $f^0 = 1$.

Proposition 12. *Consider Problem P1 with T free, $f^0 = 1$, and M compact. Assume that*

- \mathcal{T} is closed and $\mathcal{R}(q_{\text{in}}) \cap \mathcal{T} \neq \emptyset$,
- the set U is compact,
- the set $\mathbf{F}(q) = \{f(q, u) \mid u \in U\}$ is convex for every $q \in M$.

Then there exists a solution to the problem.

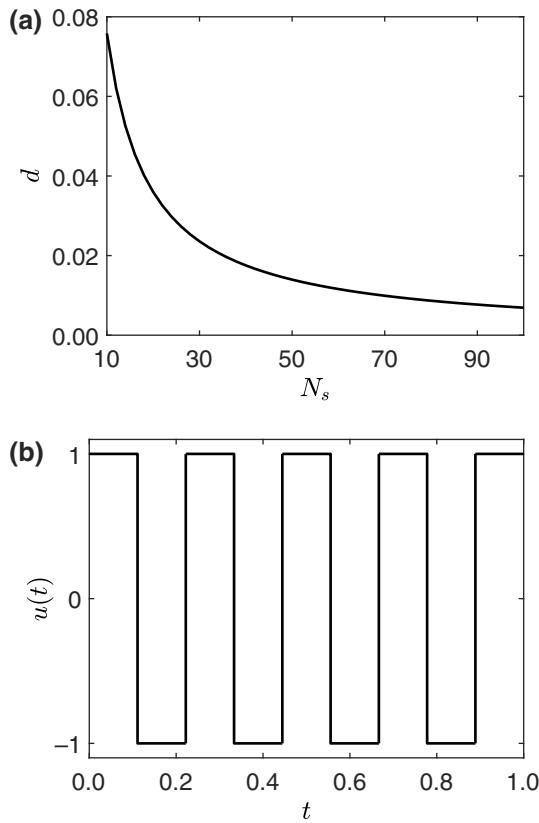


FIG. 9. Illustration of the existence problem. Panels (a) and (b) respectively depict the Euclidian distance d to the target as a function of the number N_s of switchings and an example of control law u with $N_s = 8$.

A straightforward application of Propositions 11 and 12 to closed quantum systems yields Proposition 3 of the main text.

- [1] A. Acin, 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] S. J. Glaser, U. Boscain, T. Calarco, C. P. Koch, W. K öckenberger, R. Kosloff, I. Kuprov, B. Luy, S. Schirmer, and T. Schulte-Herbrüggen *et al.*, Training Schrödinger's cat: Quantum optimal control, *European Physical Journal D* **69**, 279 (2015).
- [3] D. D'Alessandro, *Introduction to Quantum Control and Dynamics*, Applied Mathematics and Nonlinear Science Series (Chapman, Hall/CRC, Boca Raton, FL, 2008).
- [4] C. Brif, R. Chakrabarti, and H. Rabitz, Control of quantum phenomena: Past, present and future, *New J. Phys.* **12**, 075008 (2010).
- [5] C. Altafini and F. Ticozzi, Modeling and control of quantum systems: An Introduction, *IEEE Trans. Automat. Control* **57**, 1898 (2012).

- [6] D. Dong and I. R. Petersen, Quantum control theory and applications: A survey, *IET Control Theory & Applications* **4**, 2651 (2010).
- [7] C. P. Koch, Controlling open quantum systems: Tools, achievements, and limitations, *Journal of Physics: Condensed Matter* **28**, 213001 (2016).
- [8] 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).
- [9] N. V. Vitanov, A. A. Rangelov, B. W. Shore, and K. Bergmann, Stimulated Raman adiabatic passage in physics, chemistry, and beyond, *Rev. Mod. Phys.* **89**, 015006 (2017).
- [10] C. P. Koch, M. Lemeshko, and D. Sugny, Quantum control of molecular rotation, *Rev. Mod. Phys.* **91**, 035005 (2019).
- [11] D. Daems, A. Ruschhaupt, D. Sugny, and S. Guérin, Robust Quantum Control by a Single-Shot Shaped Pulse, *Phys. Rev. Lett.* **111**, 050404 (2013).
- [12] G. T. Genov, D. Schraft, T. Halfmann, and N. V. Vitanov, Correction of Arbitrary Field Errors in Population Inversion of Quantum Systems by Universal Composite Pulses, *Phys. Rev. Lett.* **113**, 043001 (2014).
- [13] S. Deffner and S. Campbell, Quantum speed limits: From Heisenberg's uncertainty principle to optimal quantum control, *Journal of Physics A: Mathematical and Theoretical* **50**, 453001 (2017).
- [14] H. Wakamura and T. Koike, A general formulation of time-optimal quantum control and optimality of singular protocols, *New. J. Phys.* **22**, 073010 (2020).
- [15] D. C. Brody and D. M. Meier, Solution to the Quantum Zermelo Navigation Problem, *Phys. Rev. Lett.* **114**, 100502 (2015).
- [16] A. Carlini, A. Hosoya, T. Koike, and Y. Okudaira, Time-Optimal Quantum Evolution, *Phys. Rev. Lett.* **96**, 060503 (2006).
- [17] X. Wang, M. Allegra, K. Jacobs, S. Lloyd, C. Lupo, and M. Mohseni, Quantum Brachistochrone Curves as Geodesics: Obtaining Accurate Minimum-Time Protocols for the Control of Quantum Systems, *Phys. Rev. Lett.* **114**, 170501 (2015).
- [18] A. Ruschhaupt, X. Chen, D. Alonso, and J. G. Muga, Optimally robust shortcuts to population inversion in two-level quantum systems, *New J. Phys.* **14**, 093040 (2012).
- [19] S. Lloyd and S. Montangero, Information Theoretical Analysis of Quantum Optimal Control, *Phys. Rev. Lett.* **113**, 010502 (2014).
- [20] F. K. Wilhelm, S. Kirchhoff, S. Machnes, N. Wittler, and D. Sugny, Lecture Notes for the 51st IFF Spring School (2020).
- [21] J. P. Palao and R. Kosloff, Quantum Computing by an Optimal Control Algorithm for Unitary Transformations, *Phys. Rev. Lett.* **89**, 188301 (2002).
- [22] J. Geng, Y. Wu, X. Wang, K. Xu, F. Shi, Y. Xie, X. Rong, and J. Du, Experimental Time-Optimal Universal Control of Spin Qubits in Solids, *Phys. Rev. Lett.* **117**, 170501 (2016).
- [23] P. Rembold, N. Oshnik, M. M. Müller, S. Montangero, T. Calarco, and E. Neu, Introduction to quantum optimal control for quantum sensing with nitrogen-vacancy

- centers in diamond, *AVS Quantum Science* **2**, 024701 (2020).
- [24] T. Schaetz, C. R. Monroe, and T. Esslinger, Focus on quantum simulation, *New J. Phys.* **15**, 085009 (2013).
- [25] P. Doria, T. Calarco, and S. Montangero, Optimal Control Technique for Many-Body Quantum Dynamics, *Phys. Rev. Lett.* **106**, 190501 (2011).
- [26] D. Liberzon, *Calculus of Variations and Optimal Control Theory* (Princeton University Press, Princeton, NJ, 2012).
- [27] L. S. Pontryagin, V. Boltianski, R. Gamkrelidze, and E. Mitchtchenko, *The Mathematical Theory of Optimal Processes* (John Wiley and Sons, New York, 1962).
- [28] A. A. Agrachev and Y. L. Sachkov, *Control Theory From the Geometric Viewpoint*, Encyclopaedia of Mathematical Sciences Vol. 87 (Springer-Verlag, Berlin, 2004).
- [29] A. Agrachev, D. Barilari, and U. Boscain, *A Comprehensive Introduction to Sub-Riemannian Geometry*, Cambridge Studies in Advanced Mathematics Vol. 181 (Cambridge University Press, Cambridge, 2020).
- [30] A. Bressan and B. Piccoli, *Introduction to the Mathematical Theory of Control*, AIMS Series on Applied Mathematics Vol. 2 (American Institute of Mathematical Sciences (AIMS), Springfield, MO, 2007).
- [31] H. Schättler and U. Ledzewicz, *Geometric Optimal Control*, Interdisciplinary Applied Mathematics Vol. 38 (Springer, New York, 2012).
- [32] U. Boscain and B. Piccoli, *Optimal Syntheses for Control Systems on 2-D Manifolds*, Mathématiques & Applications (Berlin) [Mathematics & Applications] Vol. 43 (Springer-Verlag, Berlin, 2004).
- [33] V. Jurdjevic, *Geometric Control Theory*, Cambridge Studies in Advanced Mathematics Vol. 52 (Cambridge University Press, Cambridge, 1997), ISBN 0-521-49502-4.
- [34] M. M. Lee and L. Markus, *Foundations of Optimal Control Theory* (John Wiley and Sons, New York, 1967).
- [35] B. Bonnard and D. Sugny, *Optimal Control in Space and Quantum Dynamics* (American Institute of Mathematical Sciences, Springfield, MO, 2012), Vol. 5.
- [36] A. E. Bryson and Y. C. Ho, *Applied Optimal Control: Optimization, Estimation, and Control* (Taylor and Francis, Philadelphia, 1975).
- [37] A. P. Peirce, M. A. Dahleh, and H. Rabitz, Optimal control of quantum-mechanical systems: Existence, numerical approximation, and applications, *Phys. Rev. A* **37**, 4950 (1988).
- [38] R. Kosloff, S. Rice, P. Gaspard, S. Tersigni, and D. Tannor, Wavepacket dancing: Achieving chemical selectivity by shaping light pulses, *Chem. Phys.* **139**, 201 (1989).
- [39] S. A. Rice and M. Zhao, *Optimal Control of Molecular Dynamics* (John Wiley and Sons, New York, 2000).
- [40] J. Somloj, V. A. Kazakov, and D. J. Tannor, Controlled dissociation of I₂ via optical transitions between the X and B electronic states, *Chem. Phys.* **172**, 85 (1993).
- [41] L. Levin, W. Skomorowski, L. Rybak, R. Kosloff, C. P. Koch, and Z. Amitay, Coherent Control of Bond Making, *Phys. Rev. Lett.* **114**, 233003 (2015).
- [42] D. Sugny, A. Keller, O. Atabek, D. Daems, C. M. Dion, S. Guérin, and H. Jauslin, Laser control for the optimal evolution of pure quantum states, *Phys. Rev. A* **71**, 063402 (2005).
- [43] S. Conolly, D. Nashimura, and A. Macovski, Optimal control solutions to the magnetic resonance selective excitation problem, *IEEE Trans. Med. Imag.* **5**, 106 (1986).
- [44] T. E. Skinner, T. O. Reiss, B. Luy, N. Khaneja, and S. J. Glaser, Application of optimal control theory to the design of broadband excitation pulses for high-resolution NMR, *Journal of Magnetic Resonance* **163**, 8 (2003).
- [45] N. Khaneja, T. Reiss, C. Kehlet, T. Schulte-Herbrueggen, and S. J. Glaser, Optimal control of coupled spin dynamics: Design of NMR pulse sequences by gradient ascent algorithms, *Journal of Magnetic Resonance* **172**, 296 (2005).
- [46] K. Kobzar, S. Ehni, T. E. Skinner, S. J. Glaser, and B. Luy, Exploring the limits of broadband 90° and 180° universal rotation pulses, *Journal of Magnetic Resonance* **225**, 142 (2012).
- [47] D. D'Alessandro and M. Dahleh, Optimal control of two-level quantum systems, *IEEE Trans. Automat. Contr.* **46**, 866 (2001).
- [48] U. Boscain and P. Mason, Time minimal trajectories for a spin 1–2 particle in a magnetic field, *J. Math. Phys.* **47**, 062101 (2006).
- [49] U. Boscain, T. Chambrion, and G. Charlot, Nonisotropic 3-level quantum systems: complete solutions for minimum time and minimal energy, *Discrete Contin. Dyn. Syst. Ser. B* **5**, 957 (2005).
- [50] U. Boscain, G. Charlot, J.-P. Gauthier, S. Guérin, and H.-R. Jauslin, Optimal control in laser-induced population transfer for two- and three-level quantum systems, *J. Math. Phys.* **43**, 2107 (2002).
- [51] A. Garon, S. J. Glaser, and D. Sugny, Time-optimal control of SU(2) quantum operations, *Phys. Rev. A* **88**, 043422 (2013).
- [52] N. Khaneja, R. Brockett, and S. J. Glaser, Time optimal control in spin systems, *Phys. Rev. A* **63**, 032308 (2001).
- [53] N. Khaneja, S. J. Glaser, and R. Brockett, Sub-Riemannian geometry and time optimal control of three spin systems: Quantum gates and coherence transfer, *Phys. Rev. A* **65**, 032301 (2002).
- [54] G. C. Hegerfeldt, Driving at the Quantum Speed Limit: Optimal Control of a Two-Level System, *Phys. Rev. Lett.* **111**, 260501 (2013).
- [55] D. Stefanatos, J. Ruths, and J.-S. Li, Frictionless atom cooling in harmonic traps: A time-optimal approach, *Phys. Rev. A* **82**, 063422 (2010).
- [56] X. Chen, Y. Ban, and G. C. Hegerfeldt, Time-optimal quantum control of nonlinear two-level systems, *Phys. Rev. A* **94**, 023624 (2016).
- [57] N. Khaneja, B. Luy, and S. J. Glaser, Boundary of quantum evolution under decoherence, *Proceedings of the National Academy of Sciences* **100**, 13162 (2003).
- [58] S. E. Sklarz, D. J. Tannor, and N. Khaneja, Optimal control of quantum dissipative dynamics: analytic solution for cooling the three-level Λ system, *Phys. Rev. A* **69**, 053408 (2004).
- [59] M. Lapert, Y. Zhang, M. Braun, S. J. Glaser, and D. Sugny, Singular Extremals for the Time-Optimal Control of Dissipative Spin 1/2 Particles, *Phys. Rev. Lett.* **104**, 083001 (2010).

- [60] B. Bonnard, M. Chyba, and D. Sugny, Time-minimal control of dissipative two-level quantum systems: the generic case, *IEEE. Trans. Automat. Contr.* **54**, 2598 (2009).
- [61] B. Bonnard, O. Cots, S. J. Glaser, M. Lapert, D. Sugny, and Y. Zhang, Geometric optimal control of the contrast imaging problem in nuclear magnetic Resonance, *IEEE. Trans. Automat. Contr.* **57**, 1957 (2012).
- [62] D. Stefanatos, N. Khaneja, and S. J. Glaser, Optimal control of coupled spins in the presence of longitudinal and transverse relaxation, *Phys. Rev. A* **69**, 022319 (2004).
- [63] M. Lapert, E. Assémat, S. J. Glaser, and D. Sugny, Understanding the global structure of two-level quantum systems with relaxation: Vector fields organized through the magic plane and the steady-state ellipsoid, *Phys. Rev. A* **88**, 033407 (2013).
- [64] V. Mukherjee, A. Carlini, A. Mari, T. Caneva, S. Montangero, T. Calarco, R. Fazio, and V. Giovannetti, Speeding up and slowing down the relaxation of a qubit by optimal control, *Phys. Rev. A* **88**, 062326 (2013).
- [65] D. M. Reich, M. Ndong, and C. P. Koch, Monotonically convergent optimization in quantum control using Krotov's method, *J. Chem. Phys.* **136**, 104103 (2012).
- [66] S. Machnes, U. Sander, S. J. Glaser, P. de Fouquières, A. Gruslys, S. Schirmer, and T. Schulte-Herbrüggen, Comparing, optimizing, and benchmarking quantum-control algorithms in a unifying programming framework, *Phys. Rev. A* **84**, 022305 (2011).
- [67] S. Machnes, E. Assémat, D. Tannor, and F. K. Wilhelm, Tunable, Flexible, and Efficient Optimization of Control Pulses for Practical Qubits, *Phys. Rev. Lett.* **120**, 150401 (2018).
- [68] A. Borzi, G. Ciaramella, and M. Sprengel, *Formulation and Numerical Solution of Quantum Control Problems*, Computational Science & Engineering Vol. 16 (Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2017).
- [69] U. Boscain, J.-P. Gauthier, F. Rossi, and M. Sigalotti, Approximate controllability, exact controllability, and conical eigenvalue intersections for quantum mechanical systems, *Comm. Math. Phys.* **333**, 1225 (2015).
- [70] When not specified differently, we consider in this paper manifolds without boundary. Recall that the *tangent space* $T_q M$ at a point q of a smooth manifold M is the vector space of all the vectors tangent to M at q . The *tangent bundle* TM is the set of all the tangent spaces at any $q \in M$. At every point $q \in M$ one can also attach the *cotangent space* $T_q^* M$, which is the dual space to $T_q M$. An element p of $T_q^* M$ is called a *covector* and is a scalar-valued linear function on $T_q M$, whose value at a vector $v \in T_q M$ is denoted by $\langle p, v \rangle$ and referred to as the *duality product* between p and v . The collection of the cotangent spaces at all points of M forms the *cotangent bundle* $T^* M$. If M has dimension n , $p \in T_q^* M$, and $v \in T_q M$, in local coordinates, one can think of p as a row vector and v as a column vector, so that the duality product $\langle p, v \rangle$ is equal to $p v$.
- [71] The class of *admissible controls* is the space $\mathcal{U} = L^\infty([0, T], U)$, i.e., the space of measurable and essentially bounded functions from $[0, T]$ to U . Recall that a function $u : [0, T] \rightarrow U$ is said to be *essentially bounded* if there exists a set $\omega \subset [0, T]$ of measure zero such that u is bounded on $[0, T] \setminus \omega$.
- [72] The reason is that a minimizing sequence of piecewise constant controls may fail to have a cluster point in the class (intervals on which the control is constant may have zero duration at the limit). For this reason, optimal control problems are set for a larger class of controls in which optimal solutions can be shown to exist (under rather general conditions that we discuss in the main text) and can be characterized through suitable necessary conditions. Such conditions often allow us to ensure that the optimal controls are indeed piecewise continuous. We should stress, however, that there exist some examples, such as the Fuller problem [152], for which the optimal control law has infinitely many switches between two values of U in finite time. Such a control law is not piecewise continuous, but belongs to $L^\infty(0, T, U)$.
- [73] The unique solution $q(\cdot)$ of Eq. (1) is locally Lipschitz continuous. A locally Lipschitz continuous curve $q(\cdot)$ for which there exists an admissible control $u(\cdot)$ such that Eq. (1) is satisfied is said to be an *admissible trajectory*. Recall that, given an interval $I \subset \mathbb{R}$, a function $q : I \rightarrow M$ is said to be *Lipschitz continuous* if there exists $C > 0$ such that, for every $t_1, t_2 \in I$, the distance in M between $q(t_1)$ and $q(t_2)$ is at most $C|t_1 - t_2|$, while $q(\cdot)$ is said to be *locally Lipschitz continuous* if, for every $t \in I$, there exists $\varepsilon > 0$ such that the restriction of q to $I \cap (t - \varepsilon, t + \varepsilon)$ is Lipschitz continuous.
- [74] M. Lapert, R. Tehini, G. Turinici, and D. Sugny, Monotonically convergent optimal control theory of quantum systems under a nonlinear interaction with the control field, *Phys. Rev. A* **78**, 023408 (2008).
- [75] Y. Ohtsuki and K. Nakagami, Monotonically convergent algorithms for solving quantum optimal control problems of a dynamical system nonlinearly interacting with a control, *Phys. Rev. A* **77**, 033414 (2008).
- [76] G. L. Giorgi, A. Saharyan, S. Guérin, D. Sugny, and B. Bellomo, Microscopic and phenomenological models of driven systems in structured reservoirs, *Phys. Rev. A* **101**, 012122 (2020).
- [77] H.-P. Breuer and F. Petruccione, *The Theory of Open Quantum Systems* (Oxford University, Oxford, 2002).
- [78] V. Gorini, A. Kossakowski, and E. C. G. Sudarshan, Completely positive dynamical semigroups of N-level systems, *J. Math. Phys.* **17**, 821 (1976).
- [79] G. Lindblad, On the generators of quantum dynamical semigroups, *Commun. Math. Phys.* **48**, 119 (1976).
- [80] R. Alicki and K. Lendi, *Quantum Dynamical Semigroups* (Springer, Berlin, 1987).
- [81] S. G. Schirmer and A. I. Solomon, Constraints on relaxation rates for N -level quantum systems, *Phys. Rev. A* **70**, 022107 (2004).
- [82] S. G. Schirmer, T. Zhang, and J. V. Leahy, Orbits of quantum states and geometry of Bloch vectors for N -level systems, *Journal of Physics A: Mathematical and General* **37**, 1389 (2004).
- [83] Note that even if M is here a manifold with a boundary ∂M , the forward time dynamics is well defined because the admissible velocities at any point of ∂M are either directed towards the interior of M or are tangent to ∂M [153].

- [84] Given a manifold M of dimension n , a *submanifold* of M of dimension $m \leq n$ is a subset of M that locally looks like \mathbb{R}^m .
- [85] F. Albertini and D. D'Alessandro, Notions of controllability for bilinear multilevel quantum systems, *IEEE Trans. Automat. Control* **48**, 1399 (2003).
- [86] S. G. Schirmer, H. Fu, and A. I. Solomon, Complete controllability of quantum systems, *Phys. Rev. A* **63**, 063410 (2001).
- [87] C. Altafini, Controllability properties for finite dimensional quantum Markovian master equations, *J. Math. Phys.* **44**, 2357 (2003).
- [88] G. Dirr, U. Helmke, I. Kurniawan, and T. Schulte-Herbrueggen, Lie-semigroup structures for reachability and control of open quantum systems: Kossakowski-Lindblad generators form Lie wedge to Markovian channels, *Rep. Math. Phys.* **64**, 93 (2009).
- [89] Note that this statement can be deduced at once by taking any converging subsequence of a minimizing sequence.
- [90] J. Werschnik and E. K. U. Gross, Quantum optimal control theory, *J. Phys. B* **40**, 175 (2007).
- [91] U. Hohenester, P. K. Rekdal, A. Borzì, and J. Schmiedmayer, Optimal quantum control of Bose-Einstein condensates in magnetic microtraps, *Phys. Rev. A* **75**, 023602 (2007).
- [92] U. Boscain and B. Piccoli, in *Encyclopedia of Systems and Control*, edited by J. Baillieul and T. Samad, Progr. Math. (Springer Publishing Company, Incorporated, London, 2015).
- [93] U. Boscain, T. Chambrion, and J.-P. Gauthier, On the $K + P$ problem for a three-level quantum system: optimality implies resonance, *J. Dyn. Control Syst.* **8**, 547 (2002).
- [94] D. Sugny and C. Kontz, Optimal control of a three-level quantum system by laser fields plus von Neumann measurements, *Phys. Rev. A* **77**, 063420 (2008).
- [95] N. V. Vitanov, T. Halfmann, B. W. Shore, and K. Bergmann, Laser-induced population transfer by adiabatic passage techniques, *Annu. Rev. Phys. Chem.* **52**, 763 (2001).
- [96] B. Bonnard and J.-B. Caillau, Metrics with equatorial singularities on the sphere, *Ann. Mat. Pura Appl. (4)* **193**, 1353 (2014).
- [97] V. V. Grusin, On a class of hypoelliptic operators, *Mat. Sb.* **83**, 456 (1970).
- [98] R. R. Ernst, G. Bodenhausen, and A. Wokaun, *Principles of Nuclear Magnetic Resonance in One and Two Dimensions* (Clarendon Press, Oxford, 1987).
- [99] M. H. Levitt, *Spin Dynamics: Basics of Nuclear Magnetic Resonance* (Wiley, New York, 2008).
- [100] U. Boscain and Y. Chitour, Time-optimal synthesis for left-invariant control systems on $SO(3)$, *SIAM J. Control Optim.* **44**, 111 (2005).
- [101] E. Assémat, M. Lapert, Y. Zhang, M. Braun, S. J. Glaser, and D. Sugny, Simultaneous time-optimal control of the inversion of two spin-1/2 particles, *Phys. Rev. A* **82**, 013415 (2010).
- [102] P. de Fouquieres, S. G. Schirmer, S. J. Glaser, and I. Kuprov, Second order gradient ascent pulse engineering, *J. Magn. Reson.* **212**, 412 (2011).
- [103] G. Ciaramella, A. Borzì, G. Dirr, and D. Wachsmuth, Newton methods for the optimal control of closed quantum spin systems, *SIAM J. Sci. Comput.* **37**, A319 (2015).
- [104] A. Omran, H. Levine, A. Keesling, G. Semeghini, T. T. Wang, S. Ebadi, H. Bernien, A. S. Zibrov, H. Pichler, S. Choi, J. Cui, M. Rossignolo, P. Rembold, S. Montangero, T. Calarco, M. Endres, M. Greiner, V. Vuletić, and M. D. Lukin, Generation and manipulation of Schrödinger cat states in Rydberg atom arrays, *Science* **365**, 570 (2019).
- [105] A. Larrouy, S. Patsch, R. Richaud, J.-M. Raimond, M. Brune, C. P. Koch, and S. Gleyzes, Fast Navigation in a Large Hilbert Space Using Quantum Optimal Control, *Phys. Rev. X* **10**, 021058 (2020).
- [106] Y. He, L. Ji, Y. Wang, L. Qiu, J. Zhao, Y. Ma, X. Huang, S. Wu, and D. E. Chang, Geometric Control of Collective Spontaneous Emission, *Phys. Rev. Lett.* **125**, 213602 (2020).
- [107] S. Probst, V. Ranjan, Q. Ansel, R. Heeres, B. Albanese, E. Albertinale, D. Vion, D. Esteve, S. J. Glaser, D. Sugny, and P. Bertrand, Shaped pulses for transient compensation in quantum-limited electron spin resonance spectroscopy, *J. Magn. Reson.* **303**, 42 (2019).
- [108] R. W. Heeres, P. Reinhold, N. Ofek, L. Frunzio, L. Jiang, M. H. Devoret, and R. J. Schoelkopf, Implementing a universal gate set on a logical qubit encoded in an oscillator, *Nature Commun.* **8**, 94 (2017).
- [109] G. Dridi, K. Liu, and S. Guérin, Optimal Robust Quantum Control by Inverse Geometric Optimization, *Phys. Rev. Lett.* **125**, 250403 (2020).
- [110] L. Van Damme, Q. Ansel, S. J. Glaser, and D. Sugny, Robust optimal control of two-level quantum systems, *Phys. Rev. A* **95**, 063403 (2017).
- [111] J. Zeng, X.-H. Deng, A. Russo, and E. Barnes, General solution to inhomogeneous dephasing and smooth pulse dynamical decoupling, *New J. Phys.* **20**, 033011 (2018).
- [112] J. Zeng and E. Barnes, Fastest pulses that implement dynamically corrected single-qubit phase gates, *Phys. Rev. A* **98**, 012301 (2018).
- [113] J. Zeng, C. H. Yang, A. S. Dzurak, and E. Barnes, Geometric formalism for constructing arbitrary single-qubit dynamically corrected gates, *Phys. Rev. A* **99**, 052321 (2019).
- [114] Jr.-S. Li, J. Ruths, and S. J. Glaser, Exact broadband excitation of two-level systems by mapping spins to springs, *Nature Comm.* **8**, 446 (2017).
- [115] V. Martikyan, A. Devra, D. Guéry-Odelin, S. J. Glaser, and D. Sugny, Robust control of an ensemble of springs: Application to ion cyclotron resonance and two-level quantum systems, *Phys. Rev. A* **102**, 053104 (2020).
- [116] L. Van Damme, Q. Ansel, S. J. Glaser, and D. Sugny, Time-optimal selective pulses of two uncoupled spin-1/2 particles, *Phys. Rev. A* **98**, 043421 (2018).
- [117] D. Basilewitsch, H. Yuan, and C. P. Koch, Optimally controlled quantum discrimination and estimation, *Phys. Rev. Res.* **2**, 033396 (2020).
- [118] H. Yuan and C.-H. Fred Fung, Optimal Feedback Scheme and Universal Time Scaling for Hamiltonian Parameter Estimation, *Phys. Rev. Lett.* **115**, 110401 (2015).
- [119] Q. Ansel, M. Tesch, S. J. Glaser, and D. Sugny, Optimizing fingerprinting experiments for parameter identification: Application to spin systems, *Phys. Rev. A* **96**, 053419 (2017).
- [120] Z. Hou, R.-J. Wang, J.-F. Tang, H. Yuan, G.-Y. Xiang, C.-F. Li, and G.-C. Guo, Control-Enhanced Sequential

- Scheme for General Quantum Parameter Estimation at the Heisenberg Limit, *Phys. Rev. Lett.* **123**, 040501 (2019).
- [121] F. Poggiali, P. Cappellaro, and N. Fabbri, Optimal Control for One-Qubit Quantum Sensing, *Phys. Rev. X* **8**, 021059 (2018).
- [122] P. Konzelmann, T. Rendler, V. Bergholm, A. Zappe, V. Pfannenstill, M. Garsi, F. Ziem, M. Niethammer, M. Widmann, S.-Y. Lee, P. Neumann, and J. Wrachtrup, Robust and efficient quantum optimal control of spin probes in a complex (biological) environment. Towards sensing of fast temperature fluctuations, *New J. Phys.* **20**, 123013 (2018).
- [123] M. R. Frey, Quantum speed limits—primer, perspectives, and potential future directions, *Quantum Inf. Process.* **15**, 3919 (2016).
- [124] T. Caneva, M. Murphy, T. Calarco, R. Fazio, S. Montangero, V. Giovannetti, and G. E. Santoro, Optimal Control at the Quantum Speed Limit, *Phys. Rev. Lett.* **103**, 240501 (2009).
- [125] M. Gajdacz, K. K. Das, J. Arlt, J. F. Sherson, and T. Opatrný, Time-limited optimal dynamics beyond the quantum speed limit, *Phys. Rev. A* **92**, 062106 (2015).
- [126] A. A. Díaz, V. Martikyan, S. J. Glaser, and D. Sugny, Purity speed limit of open quantum systems from magic subspaces, *Phys. Rev. A* **102**, 033104 (2020).
- [127] R. Kosloff, Quantum thermodynamics: A dynamical Viewpoint, *Entropy* **15**, 2100 (2013).
- [128] V. Cavina, A. Mari, and V. Giovannetti, Slow Dynamics and Thermodynamics of Open Quantum Systems, *Phys. Rev. Lett.* **119**, 050601 (2017).
- [129] V. Cavina, A. Mari, A. Carlini, and V. Giovannetti, Optimal thermodynamic control in open quantum systems, *Phys. Rev. A* **98**, 012139 (2018).
- [130] G. Manzano, F. Plastina, and R. Zambrini, Optimal Work Extraction and Thermodynamics of Quantum Measurements and Correlations, *Phys. Rev. Lett.* **121**, 120602 (2018).
- [131] H. J. D. Miller, M. Scandi, J. Anders, and M. Perarnau-Llobet, Work Fluctuations in Slow Processes: Quantum Signatures and Optimal Control, *Phys. Rev. Lett.* **123**, 230603 (2019).
- [132] J. Fischer, D. Basilewitsch, C. P. Koch, and D. Sugny, Time-optimal control of the purification of a qubit in contact with a structured environment, *Phys. Rev. A* **99**, 033410 (2019).
- [133] D. Basilewitsch, J. Fischer, D. M. Reich, D. Sugny, and C. P. Koch, Fundamental bounds on qubit reset, *Phys. Rev. Res.* **3**, 013110 (2021).
- [134] D. Basilewitsch, R. Schmidt, D. Sugny, S. Maniscalco, and C. P. Koch, Beating the limits with initial correlations, *New J. Phys.* **19**, 113042 (2017).
- [135] O. Abah, R. Puebla, A. Kiely, G. De chiara, M. Paternostro, and S. Campbell, Energetic cost of quantum control protocols, *New J. Phys.* **21**, 103048 (2019).
- [136] D. J. Egger and F. K. Wilhelm, Adaptive Hybrid Optimal Quantum Control for Imprecisely Characterized Systems, *Phys. Rev. Lett.* **112**, 240503 (2014).
- [137] N. Wittler, F. Roy, K. Pack, M. Werninghaus, A. S. Roy, D. J. Egger, S. Philipp, F. K. Wilhelm, and S. Machnes, Integrated Tool Set for Control, Calibration, and Characterization of Quantum Devices Applied to Superconducting Qubits, *Phys. Rev. Appl.* **15**, 034080 (2021).
- [138] F. Sauvage and F. Mintert, Optimal quantum control with poor statistics, *PRX Quantum* **1**, 020322 (2020).
- [139] R. Mukherjee, H. Xie, and F. Mintert, Bayesian Optimal Control of Greenberger-Horne-Zeilinger States in Rydberg Lattices, *Phys. Rev. Lett.* **125**, 203603 (2020).
- [140] G. Dridi, M. Lapert, J. Salomon, S. J. Glaser, and D. Sugny, Discrete-valued-pulse optimal control algorithms: Application to spin systems, *Phys. Rev. A* **92**, 043417 (2015).
- [141] I. Klickstein, A. Shirin, and F. Sorrentino, Locally Optimal Control of Complex Networks, *Phys. Rev. Lett.* **119**, 268301 (2017).
- [142] J. Li, X. Yang, X. Peng, and C.-P. Sun, Hybrid Quantum-Classical Approach to Quantum Optimal Control, *Phys. Rev. Lett.* **118**, 150503 (2017).
- [143] 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).
- [144] 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).
- [145] M. Y. Niu, S. Boixo, V. N. Smelyanskiy, and H. Neven, Universal quantum control through deep reinforcement learning, *npj Quantum Inf.* **5**, 33 (2019).
- [146] E. Perrier, C. Ferrie, and D. Tao, Quantum geometric machine learning for quantum circuits and control, *New J. Phys.* **22**, 103056 (2020).
- [147] C. Lin, Y. Wang, G. Kolesov, and U. Kalabic, Application of Pontryagin's minimum principle to Grover's quantum search problem, *Phys. Rev. A* **100**, 022327 (2019).
- [148] M. Cerezo, A. Arrasmith, R. Babbush, S. C. Benjamin, S. Endo, K. Fujii, J. R. McClean, K. Mitarai, X. Yuan, L. Cincio, and P. J. Coles, Variational Quantum Algorithms, arXiv:2012.09265 [quant-ph] (2020).
- [149] 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).
- [150] A. B. Magann, C. Arenz, M. D. Grace, T.-S. Ho, R. L. Kosut, J. R. McClean, H. A. Rabitz, and M. Sarovar, From pulses to circuits and back again: A quantum optimal control perspective on variational quantum algorithms, *PRX Quantum* **2**, 010101 (2021).
- [151] U. Boscain and M. Sigalotti, in *Contemporary Research in Elliptic PDEs and Related Topics*, Springer INdAM Ser. Vol. 33 (Springer, Cham, 2019), p. 203.
- [152] A. T. Fuller, Study of an optimum nonlinear system, *J. Electronics Control* **15**, 63 (1963).
- [153] C. Altafini, Coherent control of open quantum dynamical systems, *Phys. Rev. A* **70**, 062321 (2004).