

**Optimal control for quantum metrology via Pontryagin's principle**Chungwei Lin<sup>1,\*</sup>, Yanting Ma,<sup>1</sup> and Dries Sels<sup>2,3,4</sup><sup>1</sup>*Mitsubishi Electric Research Laboratories, 201 Broadway, Cambridge, Massachusetts 02139, USA*<sup>2</sup>*Department of Physics, Harvard University, Cambridge, Massachusetts 02138, USA*<sup>3</sup>*Department of Physics, New York University, New York, New York 10003, USA*<sup>4</sup>*Center for Computational Quantum Physics, Flatiron Institute, 162 5th Ave., New York, New York 10010, USA*

(Received 12 October 2020; accepted 30 April 2021; published 14 May 2021)

Quantum metrology comprises a set of techniques and protocols that utilize quantum features for parameter estimation which can in principle outperform any procedure based on classical physics. We formulate the quantum metrology in terms of an optimal control problem and apply Pontryagin's maximum principle to determine the optimal protocol that maximizes the quantum Fisher information for a given evolution time. As the quantum Fisher information involves a derivative with respect to the parameter which one wants to estimate, we devise an augmented dynamical system that explicitly includes gradients of the quantum Fisher information. The necessary conditions derived from Pontryagin's maximum principle are used to quantify the quality of the numerical solution. The proposed formalism is generalized to problems with control constraints, and can also be used to maximize the classical Fisher information for a chosen measurement.

DOI: [10.1103/PhysRevA.103.052607](https://doi.org/10.1103/PhysRevA.103.052607)**I. INTRODUCTION**

Modern quantum technology [1–6] requires manipulating the wave function to achieve performance beyond the scope of classical physics. A typical quantum task starts from an easily prepared initial state, undergoes a designed control protocol, and hopefully ends up with a state sufficiently close to the target state. When the closeness to the target state can be quantified by a scalar metric (a terminal cost function), the quantum task can be formulated as an optimal control problem—one tries to find the best control protocol that maximizes the performance metric for given resources. Many quantum applications (or at least an intermediate step of the application) fit this description. Important examples include quantum state preparation [7–13], where the cost function is the overlap to the known target state, the “continuous-time” variation-principle based quantum computation [14–18], where the cost function is the ground-state energy, and quantum parameter estimation (quantum metrology) [19–33], where the cost function is the Fisher information. Maximal Fisher information has been used for optimal estimation of Hamiltonian parameters [34–37]. Numerically, the Fisher information can be optimized by, e.g., GRAPE (GRADient Ascent Pulse Engineering [38]) both for single and multiple parameter estimations in the presence of noise [39–42]. The Fisher information has also been used to quantify the precision to which certain parameters of external signals (external to the sensing qubit) can be measured [43–45]. For quantum metrology application, optimal control has been applied to the preparation of entangled superposition states that are required for optimal

measurement, e.g., squeezed spin states [13,46] or Ramsey interferometry with BEC on atom chips [47,48].

Pontryagin's maximum principle (PMP) [49–52] is a powerful tool in classical control theory, and it has been applied to quantum state preparation [7,53] and nonadiabatic quantum computation [54,55]. In essence, PMP adopts the variational principle to derive a set of necessary conditions for the optimal control. In particular, it provides an efficient way to compute the gradient of the cost function with respect to the control field as well as the evolution time by introducing an auxiliary system (described by costate variables) that follows the dynamics similar to the original problem. When the system degrees of freedom are small (such as a single qubit), these necessary conditions are very restrictive and analytical solutions can sometimes be constructed [7,53,55]. For systems of higher dimensions, these necessary conditions become less informative but the efficient procedure of computing gradient is still useful for numerical solutions. Moreover, PMP optimality conditions are valuable in quantifying the quality of a numerical solution and can be done with almost no extra computational overhead. In this work, we extend PMP to quantum metrology applications where the natural choice of the terminal cost function is the quantum or classical Fisher information (QFI or CFI). The fact that QFI and CFI involve a derivative with respect to the external parameter causes some nontrivial complications. To properly use PMP, we devise an augmented dynamical system that involves the variables appearing in QFI, based on which switching functions can be stably and efficiently obtained. With the provided formalism, we are able to numerically demonstrate that the optimal control indeed satisfies all the necessary conditions of PMP.

The rest of the paper is organized as follows. First we describe the concrete problem and review the necessary

\*clin@merl.com

background of PMP. We then introduce the augmented dynamics that is designed for QFI and CFI. The formalism will be applied to a few problems, including maximizing QFI within a given control constraint and maximizing CFI for a given measurement basis. A short conclusion is provided in the end.

## II. PROBLEM AND AUGMENTED DYNAMICS FOR PMP

The concrete problem we consider is the “twist and turn” Hamiltonian [27,36,56]

$$H(t) = \chi \hat{J}_z^2 + \omega \hat{J}_z + \Omega(t) \hat{J}_x, \quad (1)$$

with  $[\hat{J}_i, \hat{J}_j] = i \epsilon_{ijk} \hat{J}_k$  ( $i = x, y, z$ ) and the initial state the nonentangled maximum-eigenvalue state of  $\hat{J}_x$ , denoted as  $|\Psi_{\text{coh-x}}\rangle$ . The potential physical realizations include interacting (generalized) spins [28,35], the two-arm interferometer [57,58], and superradiance [59,60]. The goal of the control is to efficiently estimate the parameter  $\omega$  (around zero) in Eq. (1), i.e., to produce a final state  $|\psi(T)\rangle$  over the total evolution time  $T$  that is as sensitive as possible to the change of the parameter  $\omega$  around zero. The quantitative metric is QFI:

$$\mathcal{F}_Q(|\psi(T)\rangle) = 4[\langle \partial_\omega \psi(T) | \partial_\omega \psi(T) \rangle - |\langle \psi(T) | \partial_\omega \psi(T) \rangle|^2]. \quad (2)$$

In Eq. (1),  $\hat{J}_z^2$  is the source of entanglement and referred to as a “twist” term;  $\hat{J}_x$  is the external control and referred to as a “turn” term. For eigenstates of  $\hat{J}_z$ , denoted as  $|m\rangle_z$ ,  $\hat{J}_z^2$  determines their relative phases but not amplitudes, whereas  $\hat{J}_x$  determines their relative amplitudes but not phases. The optimal control problem is to find an  $\Omega(t)$  that steers  $|\psi(0)\rangle = |\Psi_{\text{coh-x}}\rangle$  to a final state  $|\psi(T)\rangle$  that maximizes QFI at a given terminal time  $T$ . Using the terminology of control theory, Eq. (1) is control affine, as it depends linearly on the control  $\Omega(t)$ , and is time invariant, as the time dependence of  $H(t)$  is exclusively through the control  $\Omega(t)$ .

Hamiltonian (1) represents a set of  $N$  all-to-all interacting spins where  $\hat{J}_i = \sum_{n=1}^N \frac{\sigma_i}{2}$  ( $i = x, y, z$  and  $\sigma$ 's are Pauli matrices). For a system composed of  $N$  spins,  $\text{QFI}(t) = Nt^2$  is referred to as the “shot-noise” limit (SNL), which can be achieved without any quantum entanglement, and  $\text{QFI}(t) = N^2t^2$  as the Heisenberg’s limit (HL), which is the upper bound of QFI and is achieved by preparing the initial state as  $|\Psi_{\text{HL}}\rangle = (|M\rangle_z + | -M\rangle_z) / \sqrt{2}$  with  $|\pm M\rangle_z$  the largest- or smallest-eigenvalue eigenstate of  $\hat{J}_z$  (the maximum eigenvalue  $M = N/2$ ) [21]. A system displays quantum enhancement when QFI is larger than SNL. One of the key insights from Haine and Hope in Ref. [36] is that, for a limited evolution time  $T$ , the process of state preparation (i.e., to produce an entangled state) should also be regarded as a degree of freedom to maximize QFI. This becomes essential when  $T$  is too short (small  $N\chi T$ ) to produce a highly entangled state.

To compute QFI we need  $|\partial_\omega \psi(T)\rangle$ , which can be obtained by evolving  $|\partial_\omega \psi(t)\rangle$  via the differential equation  $\partial_t |\partial_\omega \psi(t)\rangle = -i \hat{J}_z |\psi(t)\rangle - i H(\omega, t) |\partial_\omega \psi(t)\rangle$  and the initial condition  $|\partial_\omega \psi(t=0)\rangle = 0$ . To apply PMP, we regard  $|\psi\rangle$  and  $|\partial_\omega \psi\rangle$  as *independent* dynamical variables. Denoting  $|\psi\rangle$

as  $|\psi_0\rangle$  and  $|\partial_\omega \psi\rangle$  as  $|\psi_1\rangle$ , the *augmented* dynamics satisfies

$$\partial_t \begin{bmatrix} |\psi_0\rangle \\ |\psi_1\rangle \end{bmatrix} = \begin{bmatrix} -iH(\omega) & 0 \\ -i\hat{J}_z & -iH(\omega) \end{bmatrix} \begin{bmatrix} |\psi_0\rangle \\ |\psi_1\rangle \end{bmatrix}. \quad (3)$$

The initial augmented state is  $(|\psi_0\rangle, |\psi_1\rangle) = (|\Psi_{\text{coh-x}}\rangle, 0)$ . The terminal cost function (to minimize) is

$$\mathcal{C}_Q = -[\langle \psi_1(T) | \psi_1(T) \rangle - |\langle \psi_1(T) | \psi_0(T) \rangle|^2], \quad (4)$$

which, up to a positive factor, is the *negative* QFI. The subscript “Q” indicates the quantum case.

Given a dynamical system Eq. (3), PMP introduces a set of auxiliary costate variables based on which the switching function and control Hamiltonian (c-Hamiltonian) are defined. Following the standard procedure [7,51,53,55], we denote  $|\pi_0\rangle$  and  $|\pi_1\rangle$  as the costate variables (in the form of wave function) of  $|\psi_0\rangle$  and  $|\psi_1\rangle$ , and derive their dynamics as

$$\partial_t \begin{bmatrix} |\pi_0\rangle \\ |\pi_1\rangle \end{bmatrix} = \begin{bmatrix} -iH(\omega) & -i\hat{J}_z \\ 0 & -iH(\omega) \end{bmatrix} \begin{bmatrix} |\pi_0\rangle \\ |\pi_1\rangle \end{bmatrix}, \quad (5)$$

with the costate boundary conditions

$$\begin{aligned} |\pi_0(T)\rangle &= \frac{\partial \mathcal{C}_Q}{\partial \langle \psi_0(T) |} = +|\psi_1(T)\rangle \langle \psi_1(T) | \psi_0(T) \rangle, \\ |\pi_1(T)\rangle &= \frac{\partial \mathcal{C}_Q}{\partial \langle \psi_1(T) |} = -|\psi_1(T)\rangle + |\psi_0(T)\rangle \langle \psi_0(T) | \psi_1(T) \rangle. \end{aligned} \quad (6)$$

Note that  $\langle \psi_0(t) | \psi_1(t) \rangle = 0$  when  $\omega = 0$ . The switching function  $\Phi(t)$  and c-Hamiltonian  $\mathcal{H}_c(t)$  are

$$\begin{aligned} \Phi(t) &= \text{Im} \left\{ \langle \pi_0 | \langle \pi_1 | \begin{bmatrix} \hat{J}_x & 0 \\ 0 & \hat{J}_x \end{bmatrix} \begin{bmatrix} |\psi_0\rangle \\ |\psi_1\rangle \end{bmatrix} \right\}, \\ \mathcal{H}_c(t) &= \text{Im} \left\{ \langle \pi_0 | \langle \pi_1 | \begin{bmatrix} H(\omega) & 0 \\ \hat{J}_z & H(\omega) \end{bmatrix} \begin{bmatrix} |\psi_0\rangle \\ |\psi_1\rangle \end{bmatrix} \right\}. \end{aligned} \quad (7)$$

According to PMP,  $\Phi(t) \sim \frac{\delta \mathcal{C}}{\delta \Omega(t)}$  and  $\mathcal{H}_c \sim \frac{\partial \mathcal{C}}{\partial T}$ . The necessary conditions for an optimal control  $\Omega(t)$  are (i)  $\Phi(t) = 0$  and (ii)  $\mathcal{H}_c(t)$  is a constant over the entire evolution time  $T$  [51,52]. Condition (i) is general (optimal solution requires a zero gradient with respect to the cost function), whereas condition (ii) is specific to time-invariant control problems; both can be served to quantify the control quality. Practically,  $\Phi(t)$  can be used in the gradient-based optimization algorithm [i.e.,  $\Omega(t) \rightarrow \Omega(t) - \gamma \Phi(t)$  with a learning rate  $\gamma$ ] for numerical solutions. The sign of  $\mathcal{H}_c(t)$  tells if increasing the evolution time  $T$  reduces [ $\mathcal{H}_c(t) < 0$ ] or increases [ $\mathcal{H}_c(t) > 0$ ] the terminal cost function [53]. In all our simulations,  $\mathcal{H}_c(t) < 0$ , meaning increasing the evolution time increases QFI. This holds for unitary dynamics but is not expected to be the case in the presence of quantum decoherence.

Three general remarks are pointed out. First, as the dynamics based on the Schrödinger equation [Eq. (1)] is typically control affine, the optimal control is expected to contain some “bang” sector(s) [7,53–55]. Based on the control theory, this expectation requires a terminal cost function that is also linear in  $|\psi(T)\rangle$ , which is true when using the fidelity as the terminal cost for a known target state [7,53,55]. As QFI is quadratic in the final state, the optimal control is not expected to be bang-bang in general. Second, the augmented dynamics (3) is nonunitary. This is not essential for the formalism but imposes

TABLE I. Optimal (maximum) QFI for different number of controls. The subscript denotes the number of controls  $N_t$ ;  $\Phi_{sd}$  is defined in Eq. (8).  $(N, \chi) = (10, 4), (20, 1), (20, 2), (20, 4), (30, 1)$  are considered. For a given  $(N, \chi)$ , a control resulting in a smaller standard deviation  $\Phi_{sd}$  has a larger QFI.

$(N, \chi)$	QFI <sub>8</sub> ( $\Phi_{sd}$ )	QFI <sub>16</sub> ( $\Phi_{sd}$ )	QFI <sub>32</sub> ( $\Phi_{sd}$ )	QFI <sub>64</sub> ( $\Phi_{sd}$ )
(10,4)	80.16 ( $8.96 \times 10^{-1}$ )	87.96 ( $4.72 \times 10^{-2}$ )	88.15 ( $3.00 \times 10^{-3}$ )	88.15 ( $2.10 \times 10^{-3}$ )
(20,1)	270.13 ( $6.62 \times 10^{-1}$ )	273.19 ( $4.34 \times 10^{-2}$ )	273.28 ( $6.10 \times 10^{-3}$ )	273.28 ( $5.55 \times 10^{-3}$ )
(20,2)	320.38 (1.72)	330.26 ( $3.00 \times 10^{-1}$ )	331.86 ( $1.81 \times 10^{-2}$ )	331.88 ( $1.00 \times 10^{-2}$ )
(20,4)	223.31 (7.95)	341.35 (1.11)	356.37 ( $3.2 \times 10^{-1}$ )	364.60 ( $1.08 \times 10^{-2}$ )
(30,1)	648.24 (4.12)	659.27 ( $4.14 \times 10^{-1}$ )	661.74 ( $4.04 \times 10^{-2}$ )	661.78 ( $2.86 \times 10^{-2}$ )

demands on the numerical ODE (ordinary differential equation) solver. In the implementation we express the dynamics using real-valued variables and use the explicit Runge-Kutta method of order 5 as the ODE solver. Finally, the proposed formalism regards  $|\psi\rangle, |\partial_\omega\psi\rangle$  as *independent* dynamical variables and introduces  $|\lambda_0\rangle, |\lambda_1\rangle$  as their corresponding costate variables. Compared to the GRAPE algorithm where computing the gradient  $\frac{\delta\mathcal{C}}{\delta\Omega(t)}$  at each  $t$  requires an integration over time (see the Appendix of Ref. [39]), in the proposed formalism the gradient  $\frac{\delta\mathcal{C}}{\delta\Omega(t)} \sim \Phi(t)$  is *local* in time [Eq. (7)], greatly reducing the computation complexity. We notice that the forward augmented dynamics alone [Eq. (3)] can be used to compute the gradients with respect to multiple control parameters [61] and has been applied to construct the optimal gate operations [61,62]. Before moving to concrete examples we point out that the model considered here [Eq. (1)] contains three parameters: the number of spins  $N$ , the twist strength  $\chi$ , and the total evolution time  $T$ . In the following discussions  $T = 1$  unless assigned specifically. We now apply the formalism to analyze a few interesting cases.

### III. APPLICATIONS

#### A. Convergence of optimal control

The control function  $\Omega(t)$  is typically approximated by a piecewise constant function, i.e.,  $\Omega(t) = \Omega_i$  for  $t \in [t_i, t_{i+1})$ , with the evolution time  $T$  divided into  $N_t$  equal time intervals [38]. As the first application, we investigate how the optimal QFI converges upon increasing  $N_t$  to approximate  $\Omega(t)$ . The motivation is to quantify the solution quality from the smallness of the switching function, which can be characterized by a mean and a standard deviation:

$$\Phi_m \equiv \frac{1}{T} \int_0^T dt \Phi(t),$$

$$\Phi_{sd} \equiv \frac{1}{\sqrt{T}} \left[ \int_0^T dt \Phi^2(t) \right]^{1/2}. \quad (8)$$

The normalizations are chosen such that  $\Phi_m$  and  $\Phi_{sd}$  have no dependence on  $T$ . For an optimal control,  $\Phi(t) = 0$ , so both  $\Phi_m$  and  $\Phi_{sd}$  vanish. When the piecewise constant function is

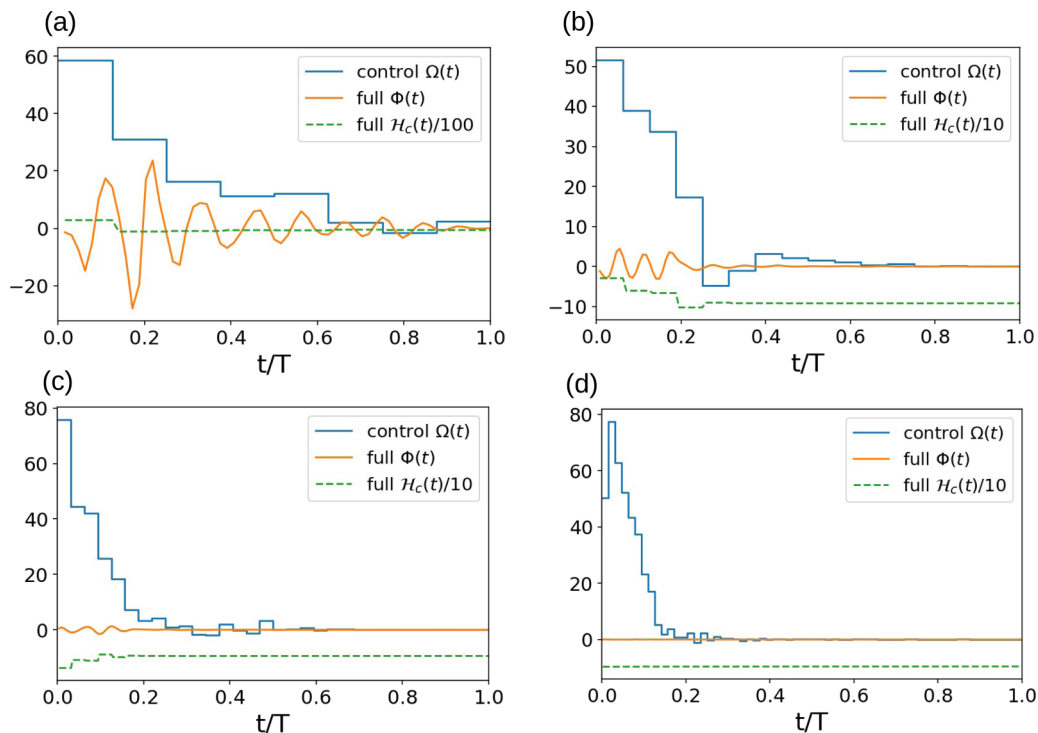


FIG. 1. Optimal controls using 8 (a), 16 (b), 32 (c), and 64 (d) controls (time intervals) with  $(N, \chi) = (20, 4)$ . Increasing the number of controls reduces the magnitudes of the switching function and results in a flatter negative c-Hamiltonian.

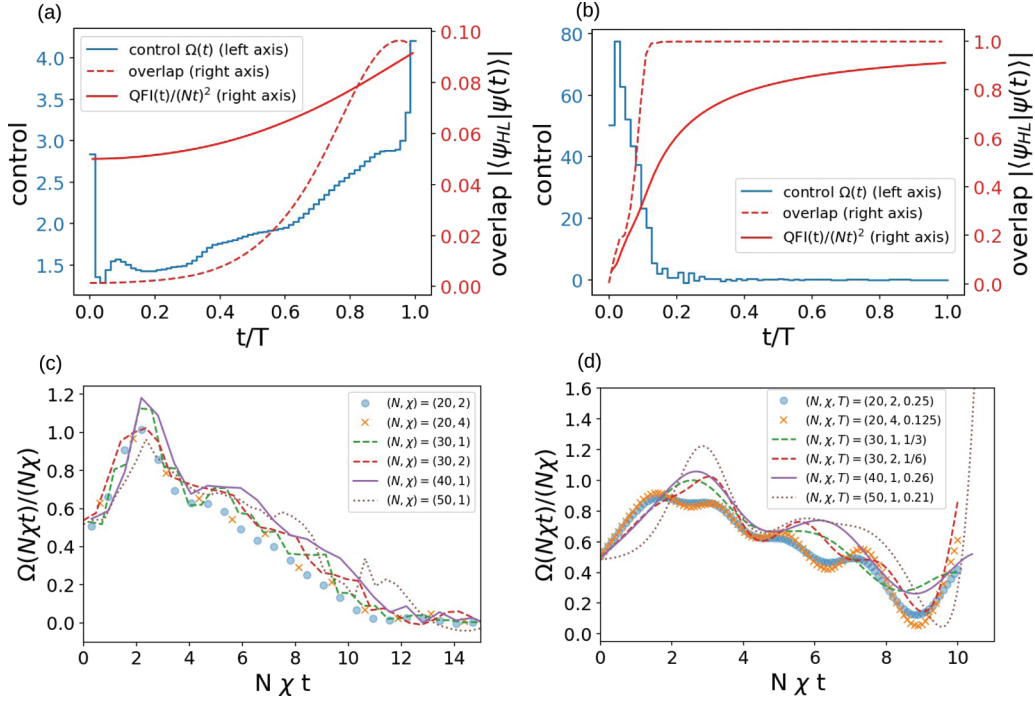


FIG. 2. Optimal controls (blue, left axis), the overlap  $|\langle \psi_{\text{HL}} | \psi(t) \rangle|$  (red dashed, right axis), and  $\text{QFI}/(Nt)^2$  (red solid, right axis) for  $(N, \chi) = (20, 0.1)$  (a) and  $(20, 4)$  (b). In (a),  $\Omega(t)$  is nonzero during the whole evolution, and the overlap to  $|\psi_{\text{HL}}\rangle$  increases to  $\lesssim 0.1$  around  $t = 0.96$ . In contrast in (b), the optimal control  $\Omega(t)$  vanishes after  $t \gtrsim 0.18$  around which  $|\langle \Psi_{\text{HL}} | \psi(t) \rangle|$  is approaching one. 64 controls are used in these simulations. (c) The dimensionless controls [Eq. (10)] for  $(N, \chi) = (20, 2), (20, 4), (30, 1), (30, 2), (40, 1),$  and  $(50, 1)$ : they almost collapse to a single curve. (d) The dimensionless controls using  $|\Psi_{\text{HL}}\rangle$  as the target state for  $(N, \chi, T) = (20, 2, 1/4), (20, 4, 1/8), (30, 1, 1/3), (30, 2, 1/6), (40, 1, 0.26),$  and  $(50, 1, 0.21)$ . All overlaps  $|\langle \Psi_{\text{HL}} | \psi(T) \rangle|$  are larger than 0.985.

approaching the optimal solution,  $\Phi_m$  is also close to zero and the value of  $\Phi_{\text{sd}} (> 0)$  can be used to characterize how good a solution is.

Table I summarizes the optimal QFI's for  $(N, \chi) = (10, 4), (20, 1), (20, 2), (20, 4),$  and  $(30, 1)$  using different numbers of controls and their corresponding  $\Phi_{\text{sd}}$ 's. As expected, the control that results in a smaller  $\Phi_{\text{sd}}$  gives a larger QFI. Figure 1 plots the optimal  $\Omega(t)$  and the corresponding  $\Phi(t)$  and  $\mathcal{H}_c(t)$  for  $(N, \chi) = (20, 4)$  using 8, 16, 32, and 64 controls. The optimal control using more time intervals gives a smaller switching function and a flatter (negative) c-Hamiltonian.

### B. Strong twist limit

When  $N\chi$  is large, the optimal control appears to be strong during early evolution and vanish after a certain amount of time (the same observation is also pointed out in Ref. [36]). This behavior can be understood by invoking the state that achieves HL. If preparing  $|\psi_{\text{HL}}\rangle$  takes only a small fraction of the total evolution time  $T$ , one way to maximize QFI is to first produce  $|\Psi_{\text{HL}}\rangle$  and then let the system interact freely with the environment. The resulting QFI is roughly  $N^2(T - t_{\text{prep}})^2$ , which approaches the HL  $N^2T^2$  when  $T \gg t_{\text{prep}}$  ( $t_{\text{prep}}$  is time to produce  $|\psi_{\text{HL}}\rangle$ ). To see what the optimal control does, Figs. 2(a) and 2(b) contrast the optimal controls for  $\chi = 0.1$  and 4 using  $N = 20$ . For  $\chi = 0.1$ ,  $\Omega(t)$  is nonzero over the entire  $T$ ; for  $\chi = 4$ ,  $\Omega(t)$  vanishes around  $t = 0.18$ . The overlap  $|\langle \psi_{\text{HL}} | \psi(t) \rangle|$  and the normalized  $\text{QFI}/(Nt)^2$  are also provided. As  $\hat{J}_x$  is the only term in Eq. (1) capable of changing

the  $|m\rangle_z$  population,  $\Omega(t)$  has to be nonzero to change the overlap; once  $\Omega(t)$  is zero the value of  $|\langle \psi_{\text{HL}} | \psi(t) \rangle|$  is fixed. For  $\chi = 0.1$ , where the entanglement source is too weak to bring the state close to  $|\psi_{\text{HL}}\rangle$ , the control is always nonzero and  $|\langle \psi_{\text{HL}} | \psi(t) \rangle|$  gradually increases to  $\lesssim 0.1$ . For  $\chi = 4$ , where the entanglement source is strong, the control steers the state close to  $|\psi_{\text{HL}}\rangle$  during  $t \leq 0.18$  and then is turned off; QFI is maximized via steering the state to  $|\psi_{\text{HL}}\rangle$  fast. This behavior appears to be general once  $N\chi$  is sufficiently large [Fig. 2(c)].

One can further analyze the optimal control by expressing Eq. (1) at  $\omega = 0$  as

$$\begin{aligned} i\partial_t |\psi\rangle &= N\chi \left[ \frac{\hat{J}_z^2}{N} + \frac{\Omega(t)}{N\chi} \hat{J}_x \right] |\psi\rangle \\ \Rightarrow i \frac{\partial |\psi\rangle}{\partial(N\chi t)} &= \left[ \frac{\hat{J}_z^2}{N} + \left( \frac{\Omega(t)}{N\chi} \right) \hat{J}_x \right] |\psi\rangle, \end{aligned} \quad (9)$$

with the given initial state  $|\Psi_{\text{coh-x}}\rangle$ . The additional information in the strong  $N\chi$  limit is that the target state is also known, at least approximately, to be  $|\Psi_{\text{HL}}\rangle$ . The second expression of Eq. (9) is scaled such that the spectral ranges of  $\hat{J}_z^2/N$  and  $\hat{J}_x$  are comparable for *all*  $N$  and therefore  $\Omega(t)/(N\chi)$  represents the  $N$ -independent strength ratio between the twist and turn. Equation (9) also introduces a dimensionless time  $N\chi t$ . Denoting  $\Omega_{N\chi}(t)$  to be the optimal control for a given  $(N, \chi, T = 1)$ , we define the corresponding dimensionless

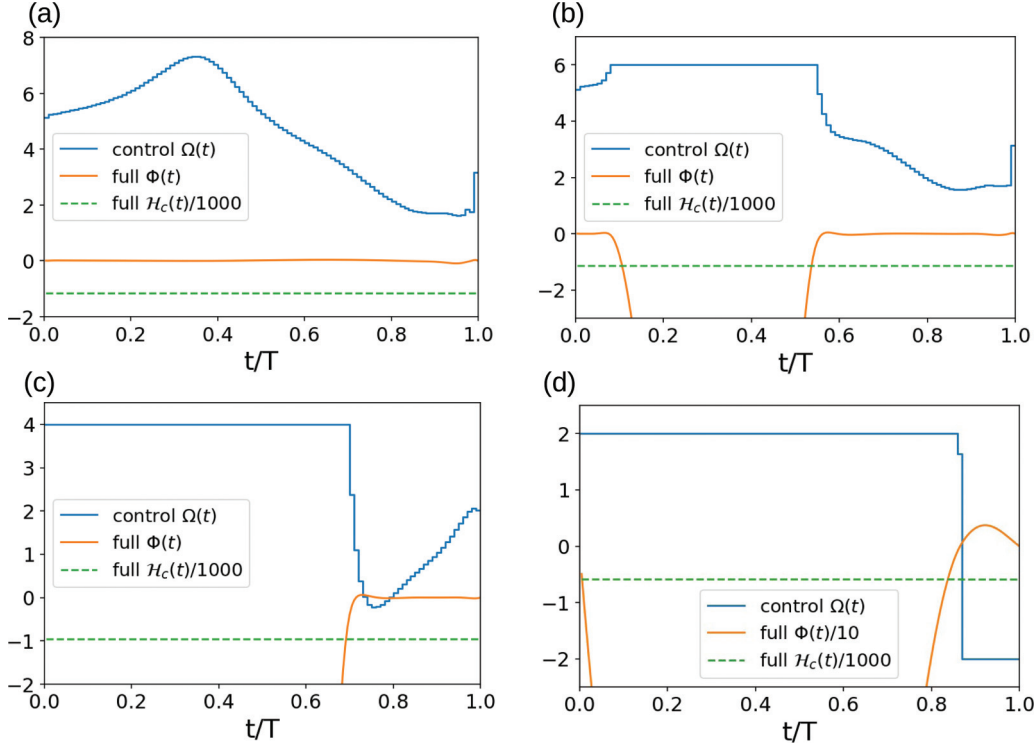


FIG. 3. Optimal control of  $N = 100$ ,  $\chi = 0.1$  using 100 time intervals. (a) No constraint on the control. (b)–(d) Optimal controls with  $\Omega(t) < u_{\max}$ : (b)  $u_{\max} = 6$ ; (c)  $u_{\max} = 4$ ; (d)  $u_{\max} = 2$ . When the optimal control takes one of the extreme values,  $\Phi(t)$  has an opposite sign.

control as

$$\bar{\Omega}_{N\chi}(t) \equiv \frac{\Omega_{N\chi}(N\chi t)}{N\chi}. \quad (10)$$

Because of similar structures of the initial and target states for all  $N$  (i.e.,  $|\Psi_{\text{coh-x}}\rangle$  is peaked at  $|m=0\rangle_z$  and is monotonously decreased as  $|m|$  increases;  $|\Psi_{\text{HL}}\rangle$  is only nonzero at  $|m=\pm N/2\rangle_z$ ), the dimensionless control  $\bar{\Omega}_{N\chi}(t)$  is expected to be only weakly dependent on  $N\chi$ . Figure 2(c) gives  $\bar{\Omega}_{N\chi}(t)$  for  $(N, \chi) = (20, 2), (20, 4), (30, 1), (30, 2), (40, 1),$  and  $(50, 1)$ : their optimal dimensionless controls [Eq. (10)] to a good approximation collapse to a single curve. A direct consequence is that the total input energy to maximize QFI, defined by  $\int_0^T dt \Omega(t)$ , is roughly a constant.

To examine this picture more carefully, Fig. 2(d) shows the dimensionless controls that maximize  $|\langle \Psi_{\text{HL}} | \psi(T) \rangle|^2$  [53,55] for  $(N, \chi, T) = (20, 2, 1/4), (20, 4, 1/8), (30, 1, 1/3), (30, 2, 1/6), (40, 1, 0.26),$  and  $(50, 1, 0.21)$ . The evolution time  $T$  is chosen to be close to and smaller than the optimal time (i.e., a  $T$  such that  $\mathcal{H}_c$  being small and negative) and all overlaps  $|\langle \Psi_{\text{HL}} | \psi(T) \rangle|$  are larger than 0.985; note the product  $N\chi T \sim 10$ . The optimal controls based on maximizing  $|\langle \Psi_{\text{HL}} | \psi(T) \rangle|^2$  [Fig. 2(d)] share the following features: (i)  $\Omega(t)$  has a peak around  $N\chi$  at  $t \sim T/3$ ; (ii)  $\Omega(t)$  increases drastically from  $\Omega \sim 0$  around  $t \lesssim T$ . Feature (i) is captured by the optimal controls that maximize QFI [Fig. 2(c)] but (ii) is not because maximizing QFI requires turning off  $\Omega$  once the state reaches  $|\Psi_{\text{HL}}\rangle$ . Feature (ii) thus highlights the difference between maximizing  $|\langle \Psi_{\text{HL}} | \psi(T) \rangle|^2$  (the traditional sensing intuition that separates the state preparation and the state evolution [36]) and maximizing QFI in the strong  $N\chi$  limit.

### C. System with constrained control amplitude

As a second application, we consider  $N = 100$ ,  $\chi = 0.1$ . These parameters are used as an example in Ref. [36]. With the ability to compute the gradient efficiently, we use 100 time intervals to approximate  $\Omega(t)$  and the obtained optimal control is given in Fig. 3(a). We see that the necessary conditions are to a good approximation satisfied; specifically  $|\Phi_m| \lesssim 10^{-3}$  and  $\Phi_{\text{sd}} \approx 0.006$ .

In practice the control amplitude is bounded, i.e.,  $|\Omega(t)| \leq u_{\max}$ , and to obtain the optimal control with amplitude constraint requires an additional step during the iteration:  $\Omega(t)$  is taken to be the closest extreme (bang) value when  $|\Omega(t)| > u_{\max}$ . The necessary condition is modified: when  $\Omega(t)$  takes the extreme value, the sign of the switching function  $\Phi(t)$  is opposite to that of  $\Omega(t)$ ; otherwise,  $\Phi(t) = 0$ . Figures 3(b)–3(d) show the results of  $u_{\max} = 6, 4, 2$ . We see that the necessary conditions are well satisfied. Imposing the maximum  $|\Omega(t)|$  reduces the optimal QFI from 2895.0 (no constraint), 2869.9 ( $u_{\max} = 6$ ), and 2431.1 ( $u_{\max} = 4$ ), to 1347.5 ( $u_{\max} = 2$ ). Consistent with the intuition, the bang control appears when  $|\Omega^*(t)| > u_{\max}$  with  $\Omega^*(t)$  the optimal control without constraints.

### D. Classical Fisher Information

As a final application, we use PMP to maximize CFI defined as

$$\mathcal{F}_C = \sum_m \frac{(\partial_\omega P_m)^2}{P_m}. \quad (11)$$

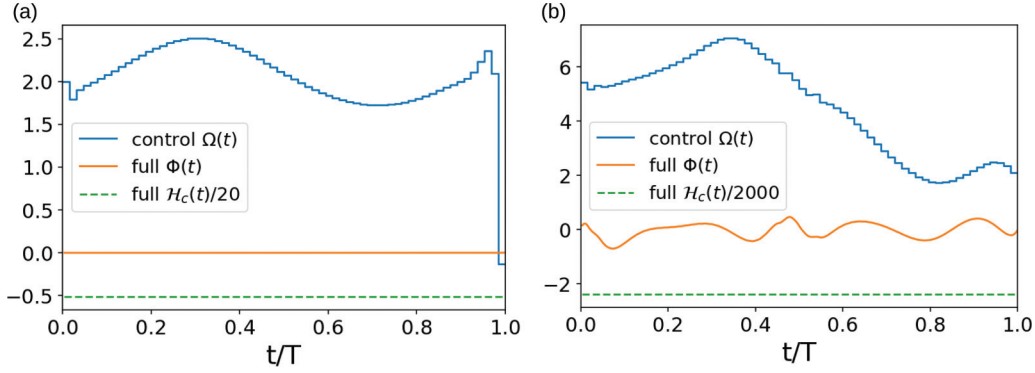


FIG. 4. (a) Optimal control for  $N = 4$ ,  $\chi = 1$  using 64 controls to maximize CFI. The optimal CFI is about 8.19 with  $\phi = \pi/2$ . (b) Optimal control for  $N = 100$ ,  $\chi = 0.1$  using 64 controls to maximize CFI. The optimal CFI is about 2867.5 with  $\phi = 0$ .

$P_m = \langle m | e^{i\phi \hat{J}_z} |\psi(T)\rangle \langle \psi(T) | m \rangle$  is the probability distribution of  $\hat{J}_x$  measurement ( $|m\rangle_{x,z}$ 's are eigenstates of  $\hat{J}_{x,z}$ ). Following Refs. [36,63] an additional phase offset  $\phi$  is introduced. The terminal cost function (to minimize) is chosen to be  $\mathcal{C}_C = -\mathcal{F}_C$  (the subscript ‘‘C’’ indicates ‘‘classical’’), and the most crucial step is to compute  $\frac{\partial \mathcal{C}_C}{\partial (\psi_{0,1}(T))}$  to get the boundary condition for the costates  $|\pi_{0,1}(T)\rangle$ .

Denoting the solution of Eq. (3) at  $t = T$  to be  $|\psi_0(T)\rangle = \sum_m \bar{\alpha}_m |m\rangle_z$  and  $|\psi_1(T)\rangle = \sum_m \bar{\beta}_m |m\rangle_z$ , and applying  $e^{i\phi \hat{J}_z}$  to the terminal state leads to

$$\begin{aligned} e^{i\phi \hat{J}_z} |\psi_0(T)\rangle &= \sum_m e^{i\phi m} \bar{\alpha}_m |m\rangle_z = \sum_m \alpha_m |m\rangle_x \\ &= \sum_m \left[ \sum_n U_{mn} \bar{\alpha}_n e^{in\phi} \right] |m\rangle_x, \\ e^{i\phi \hat{J}_z} |\psi_1(T)\rangle &= \sum_m e^{i\phi m} \bar{\beta}_m |m\rangle_z = \sum_m \beta_m |m\rangle_x \\ &= \sum_m \left[ \sum_n U_{mn} \bar{\beta}_n e^{in\phi} \right] |m\rangle_x. \end{aligned} \quad (12)$$

What we have directly from Eq. (3) are  $\bar{\alpha}_n$ ,  $\bar{\beta}_n$ , and  $U_{mn}$  (where each row vector of  $U$  is an eigenvector of  $\hat{J}_x$  and  $U$  is real valued in  $z$  basis), based on which we get  $\alpha_m = \sum_{n'} U_{mn'} \bar{\alpha}_{n'} e^{+in'\phi}$ ,  $\beta_m = \sum_{n'} U_{mn'} \bar{\beta}_{n'} e^{+in'\phi}$ ,  $P_m = \alpha_m^* \alpha_m$ ,  $\partial_\omega P_m = \beta_m^* \alpha_m + \alpha_m^* \beta_m$ , and

$$\mathcal{F}_C = \sum_m \frac{(\beta_m^* \alpha_m + \alpha_m^* \beta_m)^2}{\alpha_m^* \alpha_m}. \quad (13)$$

Straightforward derivatives give

$$\begin{aligned} \frac{\partial \mathcal{F}_C}{\partial \bar{\alpha}_n^*} &= \sum_m \frac{-(\partial_\omega P_m)^2}{P_m^2} \frac{\partial P_m}{\partial \bar{\alpha}_n^*} + \sum_m \frac{2(\partial_\omega P_m)}{P_m} \frac{\partial (\partial_\omega P_m)}{\partial \bar{\alpha}_n^*}, \\ \frac{\partial \mathcal{F}_C}{\partial \bar{\beta}_n^*} &= \sum_m \frac{2(\partial_\omega P_m)}{P_m} \frac{\partial (\partial_\omega P_m)}{\partial \bar{\beta}_n^*}, \\ \frac{\partial \mathcal{F}_C}{\partial \phi} &= \sum_m \frac{-(\partial_\omega P_m)^2}{P_m^2} \frac{\partial P_m}{\partial \phi} + \sum_m \frac{2(\partial_\omega P_m)}{P_m} \frac{\partial (\partial_\omega P_m)}{\partial \phi}, \end{aligned} \quad (14)$$

where  $\frac{\partial P_m}{\partial \bar{\alpha}_n^*} = [U_{mn} e^{-in\phi}] \alpha_m$ ,  $\frac{\partial (\partial_\omega P_m)}{\partial \bar{\alpha}_n^*} = [U_{mn} e^{-in\phi}] \beta_m$ ,  $\frac{\partial (\partial_\omega P_m)}{\partial \bar{\beta}_n^*} = [U_{mn} e^{-in\phi}] \alpha_m$ , and

$$\begin{aligned} \frac{\partial P_m}{\partial \phi} &= \left[ \sum_n U_{mn} (-i \cdot n) e^{-in\phi} \bar{\alpha}_n^* \right] \alpha_m + \text{c.c.}, \\ \frac{\partial (\partial_\omega P_m)}{\partial \phi} &= \left[ \sum_n U_{mn} (-i \cdot n) e^{-in\phi} \bar{\beta}_n^* \right] \alpha_m \\ &+ \beta_m^* \left[ \sum_n U_{mn} (+i \cdot n) e^{+in\phi} \bar{\alpha}_n \right] + \text{c.c.} \end{aligned} \quad (15)$$

The negative of Eq. (14) is used as the terminal boundary condition of the costate variables  $|\pi_0(t)\rangle$ ,  $|\pi_1(t)\rangle$ , i.e.,

$$\begin{aligned} |\pi_0(T)\rangle &= \sum_n \left[ -\frac{\partial \mathcal{F}_C}{\partial \bar{\alpha}_n^*} \right] |n\rangle_z, \\ |\pi_1(T)\rangle &= \sum_n \left[ -\frac{\partial \mathcal{F}_C}{\partial \bar{\beta}_n^*} \right] |n\rangle_z. \end{aligned} \quad (16)$$

The phase  $\phi$  is updated by  $\phi \rightarrow \phi + \gamma \frac{\partial \mathcal{F}_C}{\partial \phi}$ .

Results of  $N = 4$ ,  $\chi = 1$  and  $N = 100$ ,  $\chi = 0.1$  are presented in Figs. 4(a) and 4(b). 64 time intervals are used to approximate  $\Omega(t)$ . For  $N = 4$ ,  $\chi = 1$  [Fig. 4(a)], both the mean and standard deviation are smaller than  $10^{-3}$ . For  $N = 100$ ,  $\chi = 0.1$  [Fig. 4(b)],  $\Phi(t)$  does not converge to zero but its mean is close to zero. The mean and standard deviation of  $\Phi$  are respectively  $\sim 0.06$  and  $0.37$ . Overall, all necessary conditions are approximately satisfied for CFI optimization. As discussed in Refs. [36,63], the measurement uncertainty can be taken into account by replacing  $P_m$  by  $\hat{P}_m = \sum_{m'} \Gamma_{m,m'} P_{m'}$  in Eq. (11) (with  $\sum_m \Gamma_{m,m'} = 1$  for all  $m'$ ). The proposed method can apply to this problem as well (not shown).

#### IV. CONCLUSION

To conclude, we apply Pontryagin’s maximum principle to the quantum parameter estimation in the context of the ‘‘twist and turn’’ Hamiltonian. What PMP provides are (i) a formalism to efficiently evaluate the gradient with respect to a terminal cost function (the switching function) and (ii) a set of necessary conditions that can be used to quantify the quality of

an approximate solution. For the quantum metrology application, the performance is characterized by a single scalar—the quantum or classical Fisher information, and the optimal control finds the control protocol that maximizes QFI or CFI for a given evolution time. One nontrivial complication pertaining to quantum metrology is that the cost function involves derivatives with respect to the external parameter which one wants to estimate, and we overcome this obstacle by designing an augmented dynamical system where the wave function and its derivative to the external parameter  $|\psi\rangle$  and  $|\partial_\omega\psi\rangle$  are regarded as independent dynamical variables. By introducing the corresponding costate variables, all PMP quantities, particularly the switching function, can be stably obtained. The ability to efficiently compute the gradient greatly accelerates the optimization process and significantly expands the scope of problems one can solve. With the developed formalism, we are able to maximize QFI or CFI with more than 100 control variables. Moreover, the quality of an obtained control can be quantified by how well the PMP necessary conditions are satisfied (this applies to any approximate optimal controls). As a concrete example, we show how an optimal solution converges upon increasing the number of controls by corre-

lating the QFI and the smallness of the switching function. Specific to the twist and turn problem, we explicitly confirm the “traditional sensing intuition” in the strong twist limit: the main function of the optimal control is to steer the state to be close to  $|\psi_{\text{HL}}\rangle$  (the state that maximizes QFI) quickly and then let the system freely interact with the environment. An important and natural question is the effect of quantum decoherence, and a quantitative answer requires calculations using density matrix as dynamical variables with dynamics involving dissipation channel(s). We expect the maximum QFI to occur at a finite evolution time (as a compromise between  $\text{QFI} \sim t^2$  and decoherence), but the actual behavior should depend critically on the dissipation channel, especially when there is decoherence-free subspace [53].

### ACKNOWLEDGMENTS

We thank Y. Wang (Mitsubishi Electric Research Laboratories) for very helpful discussions. D.S. acknowledges support from the FWO as a postdoctoral fellow of the Research Foundation—Flanders.

- 
- [1] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, UK, 2011).
- [2] P. Kaye, R. Laflamme, and M. Mosca, *An Introduction to Quantum Computing* (Oxford University Press, New York, 2007).
- [3] P. W. Shor, *SIAM J. Comput.* **26**, 1484 (1997).
- [4] C. M. Caves and P. D. Drummond, *Rev. Mod. Phys.* **66**, 481 (1994).
- [5] S. L. Braunstein and P. van Loock, *Rev. Mod. Phys.* **77**, 513 (2005).
- [6] T. L. S. Collaboration, *Nat. Phys.* **7**, 962 (2011).
- [7] S. Bao, S. Kleer, R. Wang, and A. Rahmani, *Phys. Rev. A* **97**, 062343 (2018).
- [8] A. Omran, H. Levine, A. Keesling, G. Semeghini, T. T. Wang, S. Ebadi, H. Bernien, A. S. Zibrov, H. Pichler, S. Choi *et al.*, *Science* **365**, 570 (2019).
- [9] H. Jirari and W. Pötz, *Phys. Rev. A* **74**, 022306 (2006).
- [10] N. Friis, O. Marty, C. Maier, C. Hempel, M. Holzäpfel, P. Jurcevic, M. B. Plenio, M. Huber, C. Roos, R. Blatt *et al.*, *Phys. Rev. X* **8**, 021012 (2018).
- [11] M. W. Doherty, V. V. Struzhkin, D. A. Simpson, L. P. McGuinness, Y. Meng, A. Stacey, T. J. Karle, R. J. Hemley, N. B. Manson, L. C. L. Hollenberg *et al.*, *Phys. Rev. Lett.* **112**, 047601 (2014).
- [12] L. Van Damme, R. Zeier, S. J. Glaser, and D. Sugny, *Phys. Rev. A* **90**, 013409 (2014).
- [13] T. Pichler, T. Caneva, S. Montangero, M. D. Lukin, and T. Calarco, *Phys. Rev. A* **93**, 013851 (2016).
- [14] E. Farhi, J. Goldstone, S. Gurmman, and M. Sipser, [arXiv:quant-ph/0001106](https://arxiv.org/abs/quant-ph/0001106).
- [15] A. T. Rezakhani, W.-J. Kuo, A. Hamma, D. A. Lidar, and P. Zanardi, *Phys. Rev. Lett.* **103**, 080502 (2009).
- [16] Q. Zhuang, *Phys. Rev. A* **90**, 052317 (2014).
- [17] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O’Brien, *Nat. Commun.* **5**, 4213 (2014).
- [18] P. J. J. O’Malley, R. Babbush, I. D. Kivlichan, J. Romero, J. R. McClean, R. Barends, J. Kelly, P. Roushan, A. Tranter, N. Ding *et al.*, *Phys. Rev. X* **6**, 031007 (2016).
- [19] C. W. Helstrom, *Quantum Detection and Estimation Theory*, Mathematics in Science and Engineering Vol. 123 (Elsevier, Academic Press, New York, 1976).
- [20] A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory*, 1st ed. (Edizioni della Normale, Pisa, 2011).
- [21] V. Giovannetti, S. Lloyd, and L. Maccone, *Phys. Rev. Lett.* **96**, 010401 (2006).
- [22] C. Brif, R. Chakrabarti, and H. Rabitz, *New J. Phys.* **12**, 075008 (2010).
- [23] S. J. Glaser, U. Boscain, T. Calarco, C. P. Koch, W. Kockenberger, R. Kosloff, I. Kuprov, B. Luy, S. Schirmer, T. Schulte-Herbruggen *et al.*, *Eur. Phys. J. D* **69**, 279 (2015).
- [24] V. Giovannetti, S. Lloyd, and L. Maccone, *Nat. Photon.* **5**, 222 (2011).
- [25] M. Tsang, R. Nair, and X.-M. Lu, *Phys. Rev. X* **6**, 031033 (2016).
- [26] Q. Zhuang, Z. Zhang, and J. H. Shapiro, *Phys. Rev. A* **96**, 040304(R) (2017).
- [27] L. Pezzè, A. Smerzi, M. K. Oberthaler, R. Schmied, and P. Treutlein, *Rev. Mod. Phys.* **90**, 035005 (2018).
- [28] H. Zhou, J. Choi, S. Choi, R. Landig, A. M. Douglas, J. Isoya, F. Jelezko, S. Onoda, H. Sumiya, P. Cappellaro *et al.*, *Phys. Rev. X* **10**, 031003 (2020).
- [29] S. A. Haine, *Phys. Rev. Lett.* **110**, 053002 (2013).
- [30] P. Sekatski, M. Skotiniotis, J. Kołodyński, and W. Dür, *Quantum* **1**, 27 (2017).

- [31] T. Gefen, F. Jelezko, and A. Retzker, *Phys. Rev. A* **96**, 032310 (2017).
- [32] P. Rembold, N. Oshnik, M. M. Müller, S. Montangero, T. Calarco, and E. Neu, *AVS Quantum Sci.* **2**, 024701 (2020).
- [33] G. Tóth, T. Vértesi, P. Horodecki, and R. Horodecki, *Phys. Rev. Lett.* **125**, 020402 (2020).
- [34] S. Pang and A. N. Jordan, *Nat. Commun.* **8**, 14695 (2017).
- [35] R. Kaubruegger, P. Silvi, C. Kokail, R. van Bijnen, A. M. Rey, J. Ye, A. M. Kaufman, and P. Zoller, *Phys. Rev. Lett.* **123**, 260505 (2019).
- [36] S. A. Haine and J. J. Hope, *Phys. Rev. Lett.* **124**, 060402 (2020).
- [37] B. Koczor, S. Endo, T. Jones, Y. Matsuzaki, and S. C. Benjamin, *New J. Phys.* **22**, 083038 (2020).
- [38] N. Khaneja, T. Reiss, C. Kehlet, T. Schulte-Herbrüggen, and S. J. Glaser, *J. Magn. Reson.* **172**, 296 (2005).
- [39] J. Liu and H. Yuan, *Phys. Rev. A* **96**, 012117 (2017).
- [40] J. Liu and H. Yuan, *Phys. Rev. A* **96**, 042114 (2017).
- [41] D. Basilewitsch, H. Yuan, and C. P. Koch, *Phys. Rev. Res.* **2**, 033396 (2020).
- [42] J. Liu, H. Yuan, X.-M. Lu, and X. Wang, *J. Phys. A: Math. Theor.* **53**, 023001 (2019).
- [43] P. Hyllus, O. Gühne, and A. Smerzi, *Phys. Rev. A* **82**, 012337 (2010).
- [44] F. Poggiali, P. Cappellaro, and N. Fabbri, *Phys. Rev. X* **8**, 021059 (2018).
- [45] M. M. Müller, S. Gherardini, and F. Caruso, *Sci. Rep.* **8**, 14278 (2018).
- [46] Y. P. Huang and M. G. Moore, *Phys. Rev. Lett.* **100**, 250406 (2008).
- [47] C. Lovecchio, F. Schäfer, S. Cherukattil, M. Ali Khan, I. Herrera, F. S. Cataliotti, T. Calarco, S. Montangero, and F. Caruso, *Phys. Rev. A* **93**, 010304(R) (2016).
- [48] S. van Frank, A. Negretti, T. Berrada, R. Bucker, S. Montangero, J.-F. Schaff, T. Schumm, T. Calarco, and J. Schmiedmayer, *Nat. Commun.* **5**, 4009 (2014).
- [49] L. Pontryagin, *Mathematical Theory of Optimal Processes* (CRC Press, Boca Raton, FL 1987).
- [50] H. J. Sussmann, *SIAM J.* **25**, 433 (1987).
- [51] D. G. Luenberger, *Introduction to Dynamic Systems: Theory, Models, and Applications* (Wiley, New York, 1979).
- [52] U. L. Heinz Schattler, *Geometric Optimal Control: Theory, Methods and Examples*, 1st ed., Interdisciplinary Applied Mathematics Vol. 38 (Springer-Verlag New York, 2012).
- [53] C. Lin, D. Sels, and Y. Wang, *Phys. Rev. A* **101**, 022320 (2020).
- [54] Z.-C. Yang, A. Rahmani, A. Shabani, H. Neven, and C. Chamon, *Phys. Rev. X* **7**, 021027 (2017).
- [55] C. Lin, Y. Wang, G. Kolesov, and U. Kalabic, *Phys. Rev. A* **100**, 022327 (2019).
- [56] A. J. Hayes, S. Dooley, W. J. Munro, K. Nemoto, and J. Dunningham, *Quantum Sci. Technol.* **3**, 035007 (2018).
- [57] G. Tóth, *Phys. Rev. A* **85**, 022322 (2012).
- [58] B. Yurke, S. L. McCall, and J. R. Klauder, *Phys. Rev. A* **33**, 4033 (1986).
- [59] R. H. Dicke, *Phys. Rev.* **93**, 99 (1954).
- [60] M. Gross and S. Haroche, *Phys. Rep.* **93**, 301 (1982).
- [61] I. Kuprov and C. T. Rodgers, *J. Chem. Phys.* **131**, 234108 (2009).
- [62] S. Machnes, E. Assémat, D. Tannor, and F. K. Wilhelm, *Phys. Rev. Lett.* **120**, 150401 (2018).
- [63] S. P. Nolan, S. S. Szigeti, and S. A. Haine, *Phys. Rev. Lett.* **119**, 193601 (2017).