

## Robust Quantum Optimal Control with Trajectory Optimization

Thomas Propson<sup>1,2,\*</sup>, Brian E. Jackson,<sup>3</sup> Jens Koch<sup>1,4,5</sup>, Zachary Manchester<sup>1,3</sup>, and David I. Schuster<sup>1,2,6</sup>

<sup>1</sup>James Franck Institute, University of Chicago, Chicago, Illinois 60637, USA

<sup>2</sup>Department of Physics, University of Chicago, Chicago, Illinois 60637, USA

<sup>3</sup>Robotics Institute, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, USA

<sup>4</sup>Department of Physics and Astronomy, Northwestern University, Evanston, Illinois 60208, USA

<sup>5</sup>Northwestern–Fermilab Center for Applied Physics and Superconducting Technologies, Northwestern University, Evanston, Illinois 60208, USA

<sup>6</sup>Pritzker School of Molecular Engineering, University of Chicago, Chicago, Illinois 60637, USA

(Received 15 April 2021; revised 1 November 2021; accepted 9 December 2021; published 27 January 2022)

The ability to engineer high-fidelity gates on quantum processors in the presence of systematic errors remains the primary barrier to achieving quantum advantage. Quantum optimal control methods have proven effective in experimentally realizing high-fidelity gates, but they require exquisite calibration to be performant. We apply robust trajectory optimization techniques to suppress gate errors arising from system parameter uncertainty. We propose a derivative-based approach that maintains computational efficiency by using forward-mode differentiation. Additionally, the effect of depolarization on a gate is typically modeled by integrating the Lindblad master equation, which is computationally expensive. We employ a computationally efficient model and utilize time-optimal control to achieve high-fidelity gates in the presence of depolarization. We apply these techniques to a fluxonium qubit and suppress simulated gate errors due to parameter uncertainty below  $10^{-7}$  for static parameter deviations of the order of 1%.

DOI: 10.1103/PhysRevApplied.17.014036

### I. INTRODUCTION

Quantum optimal control (QOC) is a class of optimization algorithms for accurately and efficiently manipulating quantum systems. Early techniques were proposed for nuclear magnetic resonance experiments [1–7], and applications now include superconducting circuits [8–25], neutral atoms and ions [26–37], vacancy centers in diamond [38–44], and Bose-Einstein condensates [45–48]. In the context of quantum computation, optimal control is employed to achieve high-fidelity gates while adhering to experimental constraints. Experimental errors such as parameter drift, noise, and finite control resolution cause the system to deviate from the model used in optimization, hampering experimental performance [10,15,21, 34,49]. Robust control improves upon standard optimal control by encoding model parameter uncertainties in optimization objectives, yielding performance guarantees over a range of parameter values [50–52]. We adapt robust control techniques from the robotics community to mitigate parameter-uncertainty errors for a superconducting fluxonium qubit.

Analytically derived control pulses that mitigate parameter-uncertainty errors include composite pulses

[53–56], pulses designed by considering dynamic and geometric phases [57–59], and pulses obtained with the derivative-removal-by-adiabatic-gate scheme [60]. As compared to analytical techniques, QOC is advantageous for designing pulses that consider all experimental constraints and performance trade-offs [18], and for constructing operations without a known analytic solution [10,15]. Accordingly, recent work has sought to achieve robustness in QOC frameworks using closed-loop methods [61–65] and open-loop methods [3,21,43,66–71].

In this work, we study three open-loop robust control techniques that make the quantum state trajectory less sensitive to the uncertainties of static and time-dependent parameters:

1. a sampling method, similar to the work in Refs. [3,21,43,68];
2. an unscented sampling method [72–74] adapted from the unscented transform [75,76] used in state estimation;
3. a derivative method, which penalizes the sensitivity of the quantum state trajectory to uncertain parameters.

We apply these techniques to the fluxonium qubit presented in Ref. [77]. We also show that QOC can solve important problems associated with fluxonium-based

\*tcpopson@uchicago.edu

qubits: exploiting the dependence of  $T_1$  on the controls to mitigate depolarization and synchronizing the phase of qubits with distinct frequencies. To ameliorate depolarization, we perform time-optimal control and employ an efficient depolarization model for which the computational cost is independent of the Hilbert space dimension. Leveraging recent advances in trajectory optimization within the field of robotics, we solve these optimization problems using the augmented Lagrangian trajectory optimizer (ALTRO) [78], which can enforce constraints on the control fields and the quantum state trajectory.

This paper is organized as follows. First, we describe ALTRO in the context of QOC in Sec. II. We outline realistic constraints for operating the fluxonium and define the associated QOC problem in Sec. III. Then, we formulate a method for suppressing depolarization in Sec. IV. Next, we describe three techniques for achieving robustness to static parameter uncertainties in Sec. V. We adapt the same techniques to mitigate  $1/f$  flux noise in Sec. VI.

## II. BACKGROUND

In this section, we review the QOC problem statement and describe the ALTRO solver [78]. QOC concerns a vector  $\mathbf{a}(t)$  of time-dependent control fields that steer the evolution of a quantum state  $|\psi(t)\rangle$ . The evolution of the state is governed by the time-dependent Schrödinger equation (TDSE)

$$i\hbar \frac{d}{dt} |\psi(t)\rangle = H[\mathbf{a}(t), t] |\psi(t)\rangle. \quad (1)$$

The Hamiltonian  $H[\mathbf{a}(t), t]$  is determined by the quantum system and the external control fields. The QOC problem is to find the controls that minimize a functional  $J[\mathbf{a}(t)]$ , which we call the objective. To make the problem numerically tractable, the quantum state and controls are discretized into  $N$  time steps,  $|\psi(t_k)\rangle \rightarrow |\psi_k\rangle$  and  $\mathbf{a}(t_k) \rightarrow \mathbf{a}_k$ , where  $t_k = t_{k-1} + \Delta t$  and  $k \in \{1, \dots, N\}$ . In the case of a single state-transfer problem, the objective is the infidelity of the time-evolved final state  $|\psi_N\rangle$  and the intended target state  $|\psi_T\rangle$ ,  $J(\mathbf{a}) = 1 - |\langle\psi_T|\psi_N(\mathbf{a})\rangle|^2$ . Standard QOC solvers compute derivatives of the objective  $\nabla J(\mathbf{a})$ , which can easily be used to implement first-order optimization methods [3,18,79,80].

Alternatively, the QOC problem can be formulated as a trajectory optimization problem and solved using specialized solvers developed by the robotics community [78,81–83]. The objective  $J(\mathbf{a}) = \sum_k \ell_k(\mathbf{x}_k, \mathbf{u}_k)$  is expressed in terms of the cost function at each time step  $\ell_k$ , where  $\mathbf{x}_k$  is the augmented state vector and  $\mathbf{u}_k$  is the augmented control vector. We use the term *augmented* because these vectors contain all of the relevant variables in the optimization problem, not just the quantum state and the control fields; for an example, see Sec. III. The augmented control contains all variables that the experimentalist may manipulate,

and the augmented state contains all variables that depend on those in the augmented control. The variables in the augmented states depend on those in the augmented controls as defined by the differential equations governing the physical system, which are encoded in the discrete relation  $\mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_k, \mathbf{u}_k)$ . For QOC,  $\mathbf{f}(\mathbf{x}_k, \mathbf{u}_k)$ —which we call the discrete dynamics function—propagates the quantum state by integrating the TDSE (1) using a Runge-Kutta method [84] or an exponential integrator [85–88].

We incorporate constraints on the augmented controls and states by formulating them as inequalities  $\mathbf{g}_k(\mathbf{x}_k, \mathbf{u}_k) \leq \mathbf{0}$  or equalities  $\mathbf{h}_k(\mathbf{x}_k, \mathbf{u}_k) = \mathbf{0}$ . The constraint functions  $\mathbf{g}_k$  and  $\mathbf{h}_k$  may be vector valued to encode multiple constraints, and equalities and inequalities are understood componentwise. To quantify constraint satisfaction, we define each constraint's *violation* as the magnitude of its deviation:  $\max(g(\cdot), 0)$  or  $|h(\cdot)|$ , where  $g$  and  $h$  are components of constraint functions  $\mathbf{g}_k$  and  $\mathbf{h}_k$ , respectively. Stated concisely, the trajectory optimization problem is

$$\underset{\substack{\mathbf{x}_1, \dots, \mathbf{x}_N \\ \mathbf{u}_1, \dots, \mathbf{u}_{N-1}}}{\text{minimize}} \quad \sum_{k=1}^N \ell_k(\mathbf{x}_k, \mathbf{u}_k) \quad (2a)$$

$$\text{subject to} \quad \mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_k, \mathbf{u}_k) \quad \text{for all } k, \quad (2b)$$

$$\mathbf{g}_k(\mathbf{x}_k, \mathbf{u}_k) \leq \mathbf{0} \quad \text{for all } k, \quad (2c)$$

$$\mathbf{h}_k(\mathbf{x}_k, \mathbf{u}_k) = \mathbf{0} \quad \text{for all } k. \quad (2d)$$

We have formulated the problem such that the cost and constraint functions at time step  $k$  may only depend on the augmented control and state at time step  $k$ . Although this structure may appear limiting, the problem can typically be reformulated to accommodate any cost or constraint function (for an example, see Sec. III), and the ALTRO solver, which we introduce in the following discussion, exploits this structure to efficiently solve the problem.

Standard techniques for solving Eqs. (2a)–(2d) typically fall into two categories: direct methods [89,90] and indirect methods [91]. For indirect methods, the augmented controls are the *decision variables*, i.e., the variables the optimizer adjusts to solve the problem. The augmented states are obtained from the augmented controls using the discrete dynamics function, and they are used to evaluate derivatives of the cost functions. Then, the derivative information is employed to update the augmented controls. This approach is taken by standard QOC solvers such as GOAT [79], GRAPE [3,18], and Krotov's method [80]. Conversely, direct methods treat both the augmented controls and states as decision variables. In addition to minimizing the cost functions, the optimizer uses derivative information for the discrete dynamics function to satisfy the dynamics constraint (2b) to a specified tolerance. In this sense, the TDSE (1) is a constraint that may be violated for intermediate steps of the optimization, where the quantum states need not be physical. The direct approach

lends itself to a nonlinear program formulation, for which a variety of general-purpose solvers exist [92,93].

Recent state-of-the-art solvers, such as ALTRO, combine the indirect and direct methods in a two-stage approach. First, ALTRO employs an indirect solving stage using the iterative linear-quadratic regulator (ILQR) algorithm [94] as the internal solver of an augmented Lagrangian method (ALM) [95–97]. In the second direct stage, ALTRO uses a projected Newton method [98,99]. Next, we provide a more detailed summary of these two stages.

ILQR is an indirect method for minimizing the objective subject to the dynamics constraint, i.e., solving Eqs. (2a) and (2b). First, ILQR uses an initial guess for the augmented controls to obtain the augmented states with the discrete dynamics function. ILQR then constructs quadratic models for each cost function using their zeroth-, first- and second-order derivatives in a Taylor expansion about the current augmented controls and states. These models are used with a recurrence relation between time steps to obtain the locally optimal update for the augmented controls. This recurrence relation is possible to derive in closed form because cost function contributions come only from the augmented control and state at a single time step [100]. Finally, a line search [101] is performed in the direction of the locally optimal update to ensure a decrease in the objective. This procedure is repeated until convergence is reached.

While indirect solvers like ILQR are computationally efficient and maintain high accuracy for the discrete dynamics throughout the optimization, they cannot handle nonlinear equality and inequality constraints (2c) and (2d). For QOC, a popular approach to handle such constraints is to add the constraint functions to the objective [15,18,21,70]. However, this strategy does not guarantee that the constraints are satisfied as the solver trades minimization of the cost functions and constraint functions against each other. ALM remedies this issue by adaptively adjusting a Lagrange multiplier estimate for each constraint function to ensure the constraints are satisfied. ALM adds terms that are linear and quadratic in the constraint functions to the objective. Then, the new objective is minimized with ILQR. If the solution obtained with ILQR does not satisfy the constraints, the prefactors for the constraint terms in the objective are increased intelligently and the procedure is repeated.

ALM converges superlinearly, but poor numerical conditioning may lead to small decreases in the constraint violations near the locally optimal solution [102]. To address this shortcoming, ALTRO projects the solution from the ALM stage onto the constraint manifold using a (direct) projected Newton method, achieving ultralow constraint violations on the order of  $10^{-8}$ . For more information on the details of the ALTRO solver, see Refs. [78,103].

As opposed to standard QOC solvers, ALTRO can satisfy constraints on both the control fields and quantum states to tight tolerances. This advantage is crucial for this work, where multiple medium-priority cost functions are minimized subject to many high-priority constraints.

### III. QOC FOR THE FLUXONIUM

In the following, we optimize quantum gates for the superconducting fluxonium qubit—a promising building block for quantum computers due to its high coherence times [77,104–108]. In this section, we use the trajectory optimization formalism (2a)–(2d) to define the optimization problem (6a)–(6h) below, which we extend in subsequent sections to account for experimental error channels. To high accuracy, we approximate the fluxonium Hamiltonian near the flux-frustration point as a two-level system:

$$\frac{H}{\hbar} = f_q \frac{\sigma_z}{2} + a(t) \frac{\sigma_x}{2}. \quad (3)$$

Here,  $f_q$  is the qubit frequency at the flux-frustration point,  $a(t)$  is the control governing the flux offset from the flux-frustration point,  $\hbar$  is Planck's constant, and  $\sigma_z, \sigma_x$  are Pauli matrices. Although the coherent dynamics can be described with this two-level system model, our noise model, experimental constraints, and system parameters consider the full system, and they are representative of the fluxonium presented in Ref. [77].

First, we introduce the augmented control and state for the fluxonium gate problem. Since the ALTRO implementation we use does not currently support complex numbers, we represent the quantum states in the isomorphism  $\mathcal{H}(\mathbb{C}^n) \cong \mathcal{H}(\mathbb{R}^{2n})$  given in Ref. [18],

$$H|\psi\rangle \cong \begin{pmatrix} H_{\text{re}} & -H_{\text{im}} \\ H_{\text{im}} & H_{\text{re}} \end{pmatrix} \begin{pmatrix} |\psi\rangle_{\text{re}} \\ |\psi\rangle_{\text{im}} \end{pmatrix}. \quad (4)$$

We use  $\psi$ —abandoning bra-ket notation—to denote the real representation of a state given by the right-hand side of Eq. (4). To refer to the discrete moments of the flux, we introduce the notation  $\int_t a_k \equiv \int_{t_1}^{t_k} a(t) dt$ ,  $a_k \equiv a(t_k)$ ,  $d_t^n a_k \equiv d^n a(t)/dt^n|_{t=t_k}$ . The augmented control and state are

$$\mathbf{u}_k = (d_t^2 a_k), \quad \mathbf{x}_k = \begin{pmatrix} \psi_k^0 \\ \psi_k^1 \\ \int_t a_k \\ a_k \\ d_t a_k \end{pmatrix}. \quad (5)$$

Here, the superscript on the quantum states  $i \in \{0, 1\}$  acts as a label. In standard QOC frameworks, the derivatives of the control fields are obtained with finite difference

methods, e.g.,  $d_t a_k \approx (a_{k+1} - a_k)/\Delta t$  [18]. Because ALTRO requires that cost functions do not use information from multiple time steps, we make  $d_t^2 a_k$  a decision variable and numerically integrate coupled ordinary differential equations (ODEs) to obtain  $d_t a_k$ ,  $a_k$ , and  $\int_t a_k$  so that we may penalize them in cost functions. Similarly, the quantum states are obtained by numerically integrating the TDSE (1) with the fluxonium Hamiltonian (3) and the given flux  $a_k$ . These numerical integration rules are implemented in the discrete dynamics function for the problem, and they give rise to the dynamics constraint (6b).

Next, we outline the constraints for this problem. Casting this problem in terms of a multistate transfer problem, we fix as the initial states  $|\psi_1^0\rangle = |0\rangle$ ,  $|\psi_1^1\rangle = |1\rangle$  (6c). The states at the final time step are constrained to be the target states  $|\psi_N^i\rangle = |\psi_T^i\rangle \equiv U|\psi_1^i\rangle$  for all  $i$  (6e), where  $U = X/2, Y/2, Z/2$  denotes the target gate. Furthermore, we impose the normalization constraint  $|\langle\psi_k^i|\psi_k^i\rangle|^2 = 1$  for all  $i, k$  (6g) to ensure the solver does not take advantage of discretization errors in numerical integration. For the flux, we have the initial condition  $\int_t a_1 = d_t a_1 = 0$  (6d). We also enforce the boundary condition  $a_1 = a_N = 0$  (6d), (6f) so the gates may be concatenated arbitrarily. We impose the zero net-flux constraint  $\int_t a_N = 0$  (6f) that mitigates the inductive drift ubiquitous in flux-bias lines [77,109,110]. Additionally, the flux is constrained by  $|a_k| \leq 0.5$  GHz for all  $k$  (6h) to ensure the two-level approximation remains valid (3). Above 0.5 GHz, the relationship between the energy levels and the flux becomes strongly nonlinear. All gates presented in this work satisfy these constraints to a maximum violation of approximately  $10^{-8}$ .

The cost function at each time step is  $\ell_k(\mathbf{x}_k, \mathbf{u}_k) = (\mathbf{x}_k - \mathbf{x}_T)^T Q_k (\mathbf{x}_k - \mathbf{x}_T) + \mathbf{u}_k^T R_k \mathbf{u}_k$ , where  $Q_k$  and  $R_k$  are diagonal matrices of hyperparameters that assign weights to cost function contributions. The  $Q_k$  term penalizes deviations from the target augmented state  $\mathbf{x}_T = (\psi_T^0, \psi_T^1, 0, 0, 0)^T$ , which is consistent with the constraints we have imposed on  $|\psi_N^i\rangle$ ,  $\int_t a_N$ , and  $a_N$ . Accordingly, this term penalizes the squared difference of  $\psi_k^i$  and  $\psi_T^i$  and penalizes the norm of  $\int_t a_k$ ,  $a_k$ , and  $d_t a_k$ . We penalize the squared difference of the final and target quantum states, rather than their infidelities, because the Hessian of the squared-difference cost function is diagonal—which makes matrix multiplications fast—and we wish to optimize  $Z/2$  gates, which requires a metric that is sensitive to global phases for the initial states  $|0\rangle$  and  $|1\rangle$ . Additionally, the  $R_k$  term penalizes the norm of  $d_t^2 a_k$ . Penalizing the norm of  $d_t^2 a_k$  and  $d_t a_k$  makes  $a_k$  smooth, which mitigates high-frequency arbitrary waveform generator transitions. Stated succinctly, the optimization problem takes the form

$$\underset{\substack{\mathbf{x}_1, \dots, \mathbf{x}_N \\ \mathbf{u}_1, \dots, \mathbf{u}_{N-1}}}{\text{minimize}} \quad \sum_{k=1}^N (\mathbf{x}_k - \mathbf{x}_T)^T Q_k (\mathbf{x}_k - \mathbf{x}_T) + \sum_{k=1}^{N-1} \mathbf{u}_k^T R_k \mathbf{u}_k \quad (6a)$$

$$\text{subject to } \mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_k, \mathbf{u}_k) \quad \text{for all } k, \quad (6b)$$

$$|\psi_1^0\rangle = |0\rangle, \quad |\psi_1^1\rangle = |1\rangle, \quad (6c)$$

$$\int_t a_1 = a_1 = d_t a_1 = 0, \quad (6d)$$

$$|\psi_N^i\rangle = |\psi_T^i\rangle \quad \text{for all } i, \quad (6e)$$

$$\int_t a_N = a_N = 0, \quad (6f)$$

$$|\langle\psi_k^i|\psi_k^i\rangle|^2 = 1 \quad \text{for all } i, k, \quad (6g)$$

$$|a_k| \leq 0.5 \text{ GHz} \quad \text{for all } k. \quad (6h)$$

Next, we remark on our problem formulation. We put a cost function at all time steps because it benefits the ILQR solving stage [103]; although this may incentivize early achievement of the desired gate, as in Ref. [18], we are primarily concerned with achieving the gate at the final time step, which the target-state constraint (6e) ensures. Additionally, the target-state constraint requires the final state to match the target state, including its global phase, up to our chosen maximum constraint violation of approximately  $10^{-8}$ . If we did not impose this constraint, the optimizer would be allowed to sacrifice the closed-system gate error to achieve better performance on the other cost functions, which is undesirable. To enforce a constraint in standard QOC frameworks, the prefactor for the constraint function is manually increased between separate optimization instances until the constraint is satisfied [15,18,21], which becomes infeasible for more than one constraint. ALM automates these prefactor updates to find a solution that satisfies all of the given constraints. Hence, ALTRO’s ability to handle multiple constraints makes it an attractive solver for QOC problems.

In extraordinarily difficult cases of QOC, it may be impossible to obey the physics of the system and achieve the desired gate [8], i.e., the dynamics constraint (6b) and the target-state constraint (6e) may be mutually unsatisfiable. In this case, the prefactors for the constraint function terms in the ALM objective will tend to infinity—leading to numerical instability—and the optimization will not converge. To maintain a constrained approach in this situation, the maximum constraint violation for the target-state constraint can be raised to a level commensurate with the minimum acceptable gate error.

Finally, for ALTRO's first indirect stage, the augmented states are obtained explicitly with the discrete dynamics function, so the dynamics constraint and initial conditions (6b)–(6d) are satisfied by construction. In this stage, the rest of the constraint functions (6e)–(6h) are added to the objective in their isomorphism-equivalent form (4). Conversely, for the second direct stage, all of the constraints (6b)–(6h) are used to define the projection onto the constraint manifold, and the objective is unmodified. Hence, the quantum states become free parameters that are adjusted to satisfy the TDSE. Although the final solution's deviation from the TDSE is never more than the maximum constraint violation, we explicitly integrate the TDSE when reporting gate errors to ensure accuracy. Exploring the benefit of direct optimization approaches for QOC is an interesting direction for future work.

#### IV. DEPOLARIZATION MITIGATION

In this section, we outline a method for optimizing the flux to mitigate depolarization. For many superconducting circuits, the depolarization time  $T_1$  is independent of the control parameters, so the fastest possible gate incurs the least depolarization error [111]. For the fluxonium, however,  $T_1$  is strongly dependent on the flux. We enable the optimizer to trade longer gate times for longer  $T_1$  times, or shorter  $T_1$  times for shorter gate times, by making the gate time a decision variable. Additionally, previous work has modeled the gate error due to depolarization by evolving density matrices under a master equation [43,111], or evolving a large number of states in a quantum trajectory approach [112]. We avoid the increase in computational complexity required for these techniques by penalizing the integrated depolarization rate in optimization.

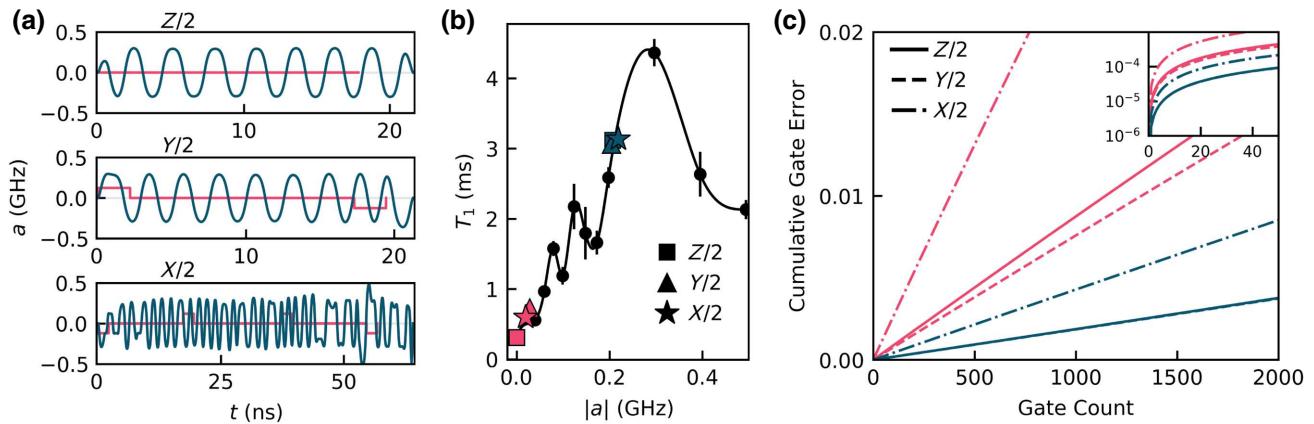


FIG. 1. (a) Flux pulses for the numerical gates (dark blue) and the analytic gates (light pink). (b) The  $T_1$  interpolation function used in optimization. Circle markers indicate measured  $T_1$  times. Noncircle markers are plotted at the time-averaged absolute flux and the time-averaged  $T_1$  time for each pulse. (c) Cumulative gate errors due to depolarization as a function of the number of gates applied. Cumulative gate errors for the numerical  $Z/2$  and  $Y/2$  gates are indistinguishable. Inset shows log-scaled cumulative gate errors for small gate counts.

The integrated depolarization rate is given by

$$D_1(t) = \int_0^t T_1^{-1}[a(t')]dt'. \quad (7)$$

For the gates we consider here, where the gate time is small compared to  $T_1$ , the integrated depolarization rate is proportional to the probability of a depolarization event. Additionally, the integrated depolarization rate is a reasonable proxy for the gate error incurred because depolarization errors are incoherent—they increase monotonically in time without interference. The integrated depolarization rate is appended to the augmented state (5) and its norm is penalized in the  $Q_k$  term of the objective by setting the corresponding element of the target augmented state to zero; see Eq. (6a). The time  $T_1$  as a function of the flux is obtained by evaluating a spline fit to experimental data; see Fig. 1(b).

Alternatively, modeling the depolarization with a master equation approach would require adding density matrices of size  $n \times n$  to the augmented state, and a quantum trajectory approach would require adding many states of size  $n$  to the augmented state, where  $n$  is the dimension of the Hilbert space. By contrast, the integrated depolarization rate is a single real number; thus, the computational complexity of evaluating this depolarization model does not scale with the dimension of the Hilbert space.

To perform time-optimal control, we make the duration between time steps a decision variable [78]. The square root of the duration  $\sqrt{\Delta t_k}$  is appended to the augmented control (5) and its square  $|\Delta t_k|$  is used for integration in the discrete dynamics function. Although we constrain the bounds of the duration between reasonable positive values to maintain numerical stability, the optimizer may assign

negative values to the duration for intermediate optimization iterations, so this squaring approach maintains positivity.

We analyze the effect of depolarization on the  $X/2$ ,  $Y/2$ , and  $Z/2$  gates obtained with our numerical method and the corresponding analytic gates presented in Ref. [77]. We use the Lindblad master equation to simulate  $T_1$  dissipation for successive gate applications, and compute the cumulative gate error after each application; see Appendix A. The gate error reported in this text is the infidelity of the evolved state and the target state averaged over 1000 pseudorandomly generated initial states.

The flux pulses for the numerical gates are approximately periodic with amplitudes of about 0.2 GHz; see Fig. 1(a). They are reminiscent of the analytically determined Floquet operations for a fluxonium described in Ref. [113] and realized in Ref. [114]. The numerical gate times are greater than the analytic gate times, but the numerical flux pulses spend more time at larger flux values, achieving higher  $T_1$  times on average; see Fig. 1(b). The single-gate errors for both the analytic and numerical gates are less than  $10^{-4}$ , which makes them sufficient for quantum error correction—a prerequisite for fault-tolerant quantum computing [115–117]. However, the numerical gates achieve single-gate errors about 5 times less than those for the analytic gates, which tracks closely with their relative improvement on the integrated depolarization rate metric; see Appendix A. This advantage in single-gate errors corresponds to a significant reduction in error correction resources [118,119]. Furthermore, for successive gate applications, the gate error due to depolarization is approximately linear in the gate count, which we expect for  $t \ll T_1$ ; see Fig. 1(c). The gate error reduction for large gate counts is important for noisy, intermediate-scale quantum (NISQ) applications. These improvements are significant given the constraints we have imposed on the gates, and do not represent a fundamental limit to the optimization methods we have employed.

## V. ROBUSTNESS TO STATIC PARAMETER UNCERTAINTY

We have formulated the QOC problem as an *open-loop* optimization problem, i.e., we do not incorporate feedback from the experiment into the optimization. However, the precise device parameters will deviate from the parameters we use in optimization, leading to poor experimental performance. We combat errors of this form using robust control techniques, making the state evolution insensitive to parameter uncertainty. As an example, we mitigate errors arising from the drift and finite measurement precision of the qubit frequency, which modifies the fluxonium Hamiltonian (3) by  $f_q \rightarrow f_q + \delta f_q$ . We consider three robust control techniques to accomplish this task: a sampling

method, an unscented sampling method, and a derivative method.

### A. Sampling method

The sampling method incentivizes the optimizer to ensure that multiple copies of a state, each evolving with a distinct value of the uncertain parameter, achieve the same target state. Variants of this technique have been proposed in the context of QOC [3,21,43,68]. For each initial state, we add two sample states  $|\psi^\pm\rangle$  to the augmented state (5). The discrete dynamics function is modified so the sample states evolve under the fluxonium Hamiltonian (3) with  $f_q \rightarrow f_q \pm \sigma_{f_q}$  for a fixed standard deviation  $\sigma_{f_q}$  of the qubit frequency, acting as a hyperparameter. We penalize the infidelities of the sample states with respect to the target state by adding a cost function to the objective of the form  $\sum_{k,\pm} b_k (1 - |\langle \psi_T | \psi_k^\pm \rangle|^2)$ , where  $b_k$  is a constant we supply. For this method, the standard orthonormal basis states are an insufficient choice for the initial states. As an example, a  $Z/2$  gate achieved by idling at the flux frustration point ( $a_k = 0$  for all  $k$ ) will be robust to qubit frequency detunings for the initial state  $|0\rangle$  or  $|1\rangle$  because the infidelity metric is insensitive to global phases, but this gate will not be robust for any other initial states. Therefore, we choose the four initial states  $\{|0\rangle, |1\rangle, (|0\rangle + i|1\rangle)/\sqrt{2}, (|0\rangle - |1\rangle)/\sqrt{2}\}$  [120], whose outer products span the operators on the Hilbert space, and we refer to them as the operator basis.

### B. Unscented sampling method

Whereas the sampling method penalizes the deviations of the sample states from the target state, the unscented sampling method penalizes the deviations of the sample states from the nominal state [72–74]. Accordingly, the cost function we add to the objective takes the form  $\sum_{k,j} c_k (\psi_k^j - \psi_k)^T (\psi_k^j - \psi_k)$ , where  $c_k$  is a constant we supply,  $\psi_k$  is the evolved initial state (nominal state), and  $\psi_k^j$  is a sample state that evolves under a modified Hamiltonian similar to that in the sampling method. The sample states are chosen to encode a unimodal distribution over the  $2n$  elements of the nominal state, modeling the uncertainty in the state as a result of the uncertainty in the parameter. We use the unscented transform [75,76] to accurately propagate the mean and covariance of this distribution between time steps, or, equivalently, through the transformation of the TDSE (1). Unlike the sampling method, the cost function for the unscented sampling method is sensitive to global phases. Accordingly, we do not observe a performance increase when using more than one initial state. A detailed procedure for the unscented transformation is given in Appendix B.

### C. Derivative method

The derivative method penalizes the sensitivity of the state to the uncertain parameter, which is encoded in the  $l$ th-order state derivative  $|\partial_{f_q}^l \psi\rangle \equiv \partial_{f_q}^l |\psi\rangle$ . In the  $m$ th-order derivative method, we append all state derivatives of order  $1, \dots, m$  to the augmented state (5) for each initial state. We obtain the state derivatives at each time step by performing forward-mode differentiation on the TDSE (1). For example, the dynamics for the first-order derivative method are

$$i\hbar \frac{d}{dt} |\psi\rangle = H|\psi\rangle, \quad (8)$$

$$i\hbar \frac{d}{dt} |\partial_{f_q} \psi\rangle = H|\partial_{f_q} \psi\rangle + (\partial_{f_q} H)|\psi\rangle. \quad (9)$$

We integrate the coupled ODEs with exponential integrators; see Appendix C. While the state  $|\psi\rangle$  has unit norm, the state derivatives  $|\partial_{f_q}^l \psi\rangle$  need not, as is evident from the nonunitary dynamics (9). We penalize the norms of the isomorphism-equivalent state derivatives in the  $Q_k$  term of the objective by setting the corresponding elements of the target augmented state to zero; see Eq. (6a). Intuitively, this corresponds to penalizing the sensitivity of each state element to the uncertain parameter. As was the case for the unscented sampling method, we do not observe a performance increase when using more than one initial state for the derivative method. We present the runtimes of our implementations of the three robust control methods in Appendix D.

### D. Comparison

We examine the gate errors due to a static qubit frequency detuning for the  $Z/2$  gates obtained with the robust control techniques and the analytic  $Z/2$  gate. To compute the gate error, an initial state is evolved under the fluxonium Hamiltonian (3) two separate times with the transformations  $f_q \rightarrow f_q \pm \delta f_q$  at the stated qubit frequency detuning  $\delta f_q$ . The reported gate error is the infidelity of the evolved state and the target state averaged over the two transformations for each of 1000 pseudorandomly generated initial states. We set  $\sigma_{f_q}/f_q = 1\%$  for the sampling and unscented sampling methods.

The analytic gate corresponds to idling at the flux frustration point  $a_k = 0$  for all  $k$ ; see Fig. 2(a). Its gate time  $1/4f_q \sim 18$  ns is the shortest possible for a  $Z/2$  gate on the device. The gate's erroneous rotation angle  $2\pi \delta f_q/4f_q$  is linear in the qubit frequency detuning, resulting in a gate error that is quadratic in the detuning. At a one-percent detuning  $|\delta f_q/f_q| = 1\%$ , the gate error is approximately  $4.7 \times 10^{-5}$ , which is sufficient for quantum error correction.

For the sampling method, the gate error at a one-percent qubit frequency detuning does not decrease substantially over the range of gate times, and begins to increase above  $5 \times 10^{-5}$  for gate times greater than approximately 50 ns; see Fig. 2(b). Optimization results for the sampling method reveal that it is typically able to achieve a high fidelity for one sample  $|\psi^\pm\rangle$ , but not the other  $|\psi^\mp\rangle$ , indicating that it is difficult for the optimizer to make progress on both objectives. For the unscented sampling method, the gate error at a one-percent detuning does not decrease substantially over the gate times, but it does reach a minimum

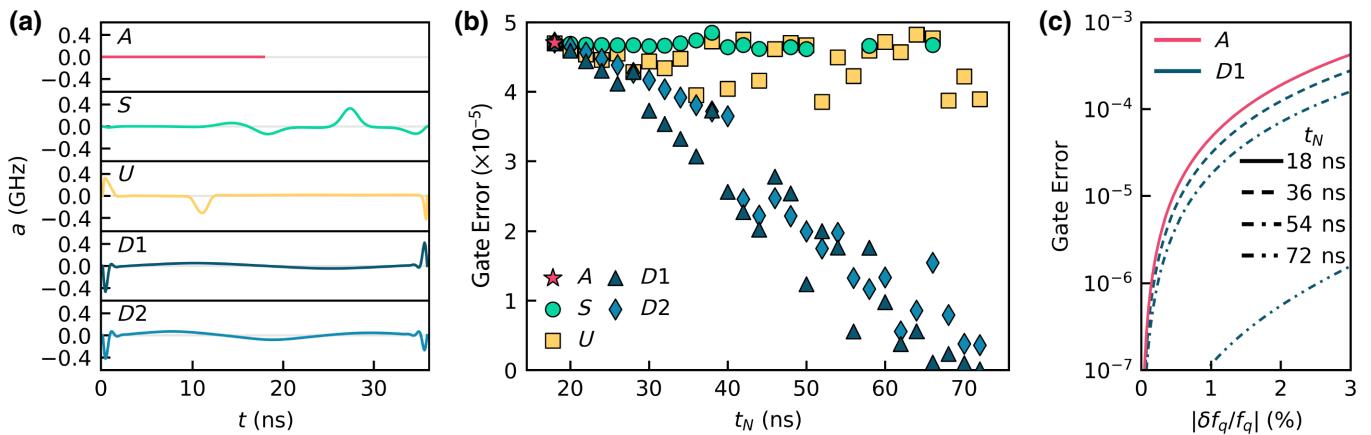


FIG. 2. (a) Flux pulses for  $Z/2$  gates robust to qubit frequency detunings constructed with the analytic ( $A$ ), sampling ( $S$ ), unscented sampling ( $U$ ), and the first- and second-order derivative methods ( $D1, D2$ ). The flux pulses shown for the sampling, unscented sampling, and derivative methods are optimized for twice the gate time of the analytic gate. (b) Single-gate error at a one-percent qubit frequency detuning as a function of the gate time. Missing data points represent gates with a gate error greater than  $5 \times 10^{-5}$ . (c) Single-gate error as a function of the qubit frequency detuning. The gate errors for the analytic and first-order derivative methods are shown for gate times that are multiples of  $1/4f_q \sim 18$  ns. The gate errors for the two methods are indistinguishable at the gate time 18 ns.

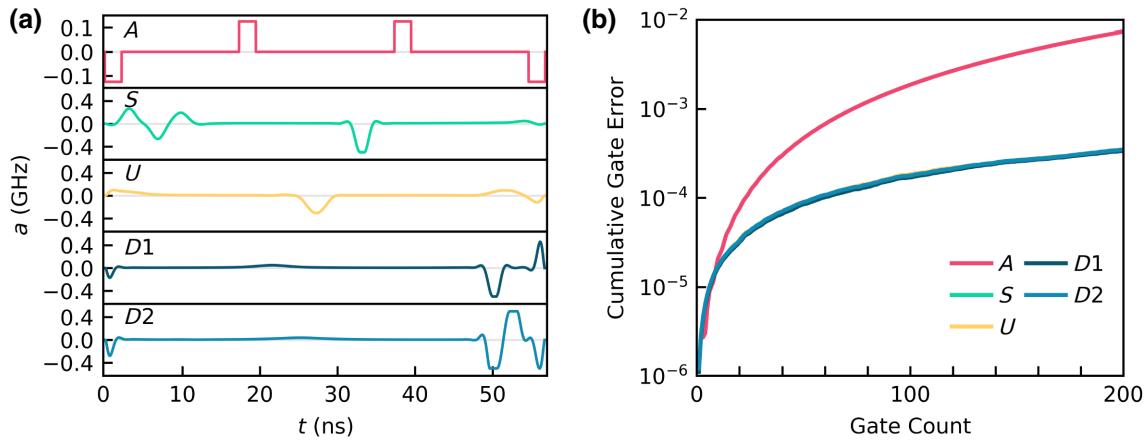


FIG. 3. (a) Flux pulses for  $X/2$  gates robust to flux noise constructed with the analytic ( $A$ ), sampling ( $S$ ), unscented sampling ( $U$ ), and the first- and second-order derivative methods ( $D1$ ,  $D2$ ). (b) Cumulative gate error due to  $1/f$  flux noise for successive gate applications. The cumulative gate errors for the sampling, unscented sampling, and the derivative methods are indistinguishable.

of approximately  $3.9 \times 10^{-5}$  near fractions of the Larmor period:  $2/4f_q \sim 36$  ns,  $3/4f_q \sim 54$  ns, and  $4/4f_q \sim 72$  ns.

The two derivative methods converge on qualitatively similar flux pulses that idle near the flux frustration point and use fast triangle movements at the boundaries, similar to the flux pulse produced by the unscented sampling method. For both derivative methods, the gate error at a one-percent qubit frequency detuning decreases superlinearly in the gate time. For the first-order method, the gate error at a one-percent detuning reaches  $10^{-7}$  at the Larmor period  $1/f_q \sim 72$  ns; see Fig. 2(c). This result mimics the ability of composite pulses to mitigate parameter uncertainty errors to arbitrary order with sufficiently many pulses [56]. It is difficult to choose an appropriate composite pulse for the problem studied here due to our Hamiltonian and experimental constraints. A comparison between composite pulses and numerical techniques could be an interesting topic for future study.

Furthermore, the ability to perform  $Z$ -type gates in any given time is critical for synchronizing phases in multiqubit experiments, where the qubits have distinct frequencies. Notably, the analytic gate studied here cannot be extended to gate times other than  $1/4f_q$ . We can find gates using the numerical methods at all gate times at and above  $1/4f_q$ ; see Fig. 2(b). These numerical methods offer an effective scheme for synchronizing multiqubit experiments.

## VI. ROBUSTNESS TO TIME-DEPENDENT PARAMETER UNCERTAINTY

An additional source of experimental error arises from time-dependent parameter uncertainty. For many flux-biased and inductively coupled superconducting circuit elements, magnetic flux noise is the dominant source of coherent errors, and therefore pure dephasing [121–124].

Flux noise modifies the fluxonium Hamiltonian (3) by  $a(t) \rightarrow a(t) + \delta a(t)$ , where  $\delta a(t)$  is the flux noise. The spectral density of flux noise is observed to follow a  $1/f$  distribution [77,121–126], so the noise is dominated by low-frequency components. The analytic gate considered here takes advantage of the low-frequency characteristic and treats the noise as quasistatic, performing a generalization of the spin-echo technique to compensate for erroneous drift [127,128].

We modify the robust control techniques presented in the previous section to combat  $1/f$  flux noise. The unscented sampling method is modified so that the sample states are subject to  $1/f$  flux noise. The noise is generated by filtering white noise sampled from a standard normal distribution with a finite impulse response filter [129]. The noise is then scaled by the flux noise amplitude of our device  $A_\Phi = 5.21\mu\Phi_0$ , which implies that  $\sigma_a = 2.5 \times 10^{-5}$  GHz. In principle, we could modify the sampling method similarly; however, we choose to subject the sample states to static noise  $a(t) \rightarrow a(t) \pm \sigma_a$  for comparison. The derivative methods require no algorithmic modification from the static case, but the TDSE is now differentiated with respect to  $a(t)$  instead of  $f_q$  as in Eq. (9).

We analyze the gate errors due to  $1/f$  flux noise for the  $X/2$  gates constructed with the robust control techniques and the analytic  $X/2$  gate. To compute the gate error, we evolve an initial state under the fluxonium Hamiltonian (3) where the optimized flux is modified,  $a(t) \rightarrow a(t) + \delta a(t)$ . We generate the flux noise as we described for the unscented sampling method. The reported gate error is the infidelity averaged over 1000 pseudorandomly generated initial states, each of which is subject to a distinct pseudorandomly generated flux noise instance. To observe the effect of interfering coherent errors, we simulate successive applications of the gate constructed by each method; we compute the cumulative gate error after each

application; see Fig. 3. Both the analytic and numerical gates yield single-gate errors sufficient for quantum error correction. Despite converging on qualitatively different solutions, the numerical gates perform similarly in the concatenated gate application comparison. Their gate errors after 200 gate applications—approximately  $11 \mu\text{s}$  in duration—are two orders of magnitude less than the gate error produced by the analytic gate. The  $1/f$  flux noise is a significant source of coherent errors in NISQ applications, and these numerical techniques offer effective avenues to mitigate it.

For qubit platforms where pure dephasing is caused by high-frequency noise, i.e., noise that has a non-negligible spectral weight at frequencies that are greater than the sampling rate of the control pulses, it is unclear that pulse shaping alone can ameliorate pure dephasing. In such cases, applying the time-optimal control scheme we have developed in Sec. IV to decrease the gate duration may be an effective strategy to mitigate pure dephasing.

## VII. CONCLUSION

We have introduced state-of-the-art trajectory optimization techniques in the context of quantum optimal control, enabling us to achieve tight tolerances for multiple constraints on the control fields and quantum states. Using these capabilities, we have mitigated decoherence and achieved robustness to parameter uncertainty errors on a superconducting fluxonium qubit. We have proposed a scheme for suppressing depolarization with time-optimal control and the integrated depolarization rate model. The computational complexity of evaluating this model is independent of the dimension of the Hilbert space, enabling inexpensive optimization on high-dimensional quantum systems. We have also proposed the derivative method for robust control that achieves superlinear gate error reductions in the gate time for the static parameter uncertainty problem we studied. We have shown that the derivative, sampling, and unscented sampling methods can mitigate  $1/f$  flux noise errors—which dominate coherent errors for flux controlled qubits. These robust control techniques can be applied to any Hamiltonian, allowing experimentalists in all domains to engineer robust operations on their quantum systems. Furthermore, they can be used to achieve the low gate errors required for fault-tolerant quantum computing applications. Our implementations of the techniques described in this work are available online [130].

## ACKNOWLEDGMENTS

We thank Helin Zhang for experimental assistance and Taylor Howell, Tanay Roy, Colm Ryan, and Daniel Weiss for useful discussions. This work was made possible by many open-source software projects, including but not limited to: DifferentialEquations.jl [131], Distributions.jl [132], ForwardDiff.jl [133], Matplotlib [134], NumPy

[135], TrajectoryOptimization.jl [78], and Zygote.jl [136]. This work is funded in part by EPiQC, an NSF Expedition in Computing, under Grant No. CCF-1730449. This work is supported by the Army Research Office under Grant No. W911NF1910016.

## APPENDIX A: DEPOLARIZATION

We comment on the depolarization metrics and then give our procedure for integrating the Lindblad master equation. The integrated depolarization rate and the gate error due to depolarization are compared in Table I for the numerical experiment described in Sec. IV. The ratio of the value obtained on the metric with the analytic technique to the value obtained with the numerical technique is similar across the two metrics.

We employ the Lindblad master equation to compute the gate error due to depolarization. This equation takes the form

$$\frac{d}{dt}\rho = -\frac{i}{\hbar}[H, \rho] + \sum_i \gamma_i \left( L_i \rho L_i^\dagger - \frac{1}{2} \{L_i^\dagger L_i, \rho\} \right). \quad (\text{A1})$$

For depolarization,  $\gamma_\pm = T_\pm^{-1}$ ,  $L_\pm = \sigma^\pm \equiv (\sigma_x \pm i\sigma_y)/2$ . Our device operates in the regime where  $hf \ll k_B T$  such that  $T_+ = T_- = 2T_1$ , where  $T_1$  is obtained at each time step from the spline shown in Fig. 1(b). We obtain the  $T_1$  values in this spline by driving the qubit at the desired flux bias and monitoring the resultant decay. For more details on these measurements, consult Ref. [77]. Because  $T_1$  depends on the flux, so do the decay rates  $\gamma_\pm$ . Integrating the master equation with time-dependent decay rates provides a heuristic for how gates might perform in the experiment. This procedure may not be strictly correct when decay rates change significantly on the time scale of the relaxation time, which is the regime we are operating in. Standard derivations of the Lindblad master equation do not account for time-dependent decay rates [137]. A more thorough treatment of this regime in future work would unlock insights for quantum computing platforms where decoherence is strongly dependent on the control parameters.

In order to use exponential integrators, we employ the vector (Choi-Jamiolkowski) isomorphism [138],

$$\frac{d}{dt}\text{vec}(\rho) = \hat{\mathcal{L}} \text{vec}(\rho), \quad (\text{A2})$$

$$\begin{aligned} \hat{\mathcal{L}} = & -i(\mathbb{1} \otimes H - H^T \otimes \mathbb{1}) + \sum_i \gamma_i \\ & \times \left( L_i^* \otimes L_i - \frac{1}{2}(\mathbb{1} \otimes L_i^\dagger L_i - L_i^T L_i^* \otimes \mathbb{1}) \right), \end{aligned} \quad (\text{A3})$$

where  $\rho = \sum_{ij} \alpha_{ij} |i\rangle \langle j|$  and  $\text{vec}(\rho) = \sum_{ij} \alpha_{ij} |i\rangle \otimes |j\rangle$ . Because the flux is constant between time steps due to our

TABLE I. Single-gate integrated depolarization rate ( $D_1$ ) and single-gate error (GE) due to depolarization. Values are reported for the analytic ( $A$ ) and numerical ( $N$ ) gates.

Gate	$D_{1A} (10^{-5})$	$D_{1N} (10^{-5})$	$D_{1A}/D_{1N}$	$GE_A (10^{-5})$	$GE_N (10^{-5})$	$GE_A/GE_N$
Z/2	5.745	1.149	5.000	0.888	0.185	4.791
Y/2	5.253	1.157	4.540	0.770	0.186	4.132
X/2	16.251	2.660	6.109	2.674	0.432	6.200

numerical discretization, the Hamiltonian and decay rates are also constant between time steps. Therefore, the exact solution to Eq. (A2) is

$$\text{vec}(\rho_{k+1}) = \exp(\Delta t_k \hat{\mathcal{L}}_k) \text{vec}(\rho_k). \quad (\text{A4})$$

The vector isomorphism transforms  $(n \times n) \times (n \times n)$  matrix-matrix multiplications to  $(n^2 \times n^2) \times n^2$  matrix-vector multiplications. For small  $n$ , we find that it is faster to use an exponential integrator on the vectorized equation than to perform Runge-Kutta on the unvectorized equation. The latter requires decreasing the interval  $\Delta t_k$  to maintain accuracy, resulting in more time steps.

## APPENDIX B: UNSCENTED SAMPLING METHOD

In this section, we outline the full unscented sampling procedure. We consider a state  $\psi \in \mathbb{R}^{2n}$ , an uncertain set of parameters  $\lambda \in \mathbb{R}^d$ , and discrete dynamics  $\psi_{k+1} = f(\psi_k, \lambda_k)$ . The nominal initial state is given by  $\bar{\psi}_1$  with an associated covariance matrix  $P_1 \in \mathbb{S}_{++}^{2n}$  that describes the uncertainty in the initial state. We use the notation  $\mathbb{S}_{++}^m$  to denote the set of real, symmetric, and positive-definite  $m \times m$  matrices. By the positive-definite requirement,  $P_1$  must be nonzero even if the state-preparation error is negligible. The uncertain parameter has zero mean and its distribution is given by the covariance matrix  $L_k \in \mathbb{S}_{++}^d$  at time step  $k$ . The zero-mean assumption is convenient for deriving the update procedure. A nonzero mean can be encoded in the discrete dynamics function  $f(\psi_k, \lambda_k)$ .

The initial  $4n + 2d$  sample states and initial  $4n + 2d$  uncertain parameters are sampled from the initial distributions,

$$\begin{pmatrix} \psi_1^j \\ \lambda_1^j \end{pmatrix} = \begin{pmatrix} \bar{\psi}_1 \\ 0 \end{pmatrix} \pm \beta \sqrt{\begin{pmatrix} P_1 & 0 \\ 0 & L_1 \end{pmatrix}}^j. \quad (\text{B1})$$

Here,  $\beta$  is a hyperparameter that controls the spacing of the covariance contour. The “ $\pm$ ” is understood to take “ $+$ ” for  $j \in \{1, \dots, 2n + d\}$  and “ $-$ ” for  $j \in \{2n + d + 1, \dots, 4n + 2d\}$ . We use the Cholesky factorization to compute the square root of the joint covariance matrix, though other methods such as the principal square root may be employed. The superscript on the matrix square root indicates the  $j$ th column (mod  $2n + d$ ) of the lower

triangular Cholesky factor. Then, the sample states are normalized,

$$\psi_1^j \rightarrow \frac{\psi_1^j}{\sqrt{\psi_1^{j^T} \psi_1^j}}. \quad (\text{B2})$$

The sample states are propagated to the next time step,

$$\psi_2^j = f(\psi_1^j, \lambda_1^j). \quad (\text{B3})$$

The mean and covariance of the sample states are computed,

$$\bar{\psi}_2 = \frac{1}{4n + 2d} \sum_{j=1}^{4n+2d} \psi_2^j, \quad (\text{B4})$$

$$P_2 = \frac{1}{2\beta^2} \sum_{j=1}^{4n+2d} (\psi_2^j - \bar{\psi}_2)(\psi_2^j - \bar{\psi}_2)^T. \quad (\text{B5})$$

The sample states are then resampled and propagated to the next time step using Eqs. (B1)–(B3). Our choice of sample states (sigma points) follows Eq. (11) of Ref. [75]. Prescriptions that require fewer sigma points exist [139].

## APPENDIX C: DERIVATIVE METHOD

Here, we outline how to efficiently integrate the dynamics for the derivative method using exponential integrators. General exponential integrators break the dynamics into a linear term and a nonlinear term. For example, the dynamics for the first state derivative are

$$\frac{d}{dt} |\partial_\lambda \psi\rangle = -\frac{i}{\hbar} H |\partial_\lambda \psi\rangle - \frac{i}{\hbar} (\partial_\lambda H) |\psi\rangle. \quad (\text{C1})$$

The linear term is  $L = -iH/\hbar$  and the nonlinear term is  $N = -(i/\hbar)(\partial_\lambda H)|\psi\rangle$ . With piecewise-constant controls, the exact solution to Eq. (C1) is

$$|\partial_\lambda \psi_{k+1}\rangle = \exp(\Delta t_k L_k) |\partial_\lambda \psi_k\rangle + \int_0^{\Delta t_k} \exp[(\Delta t_k - t') L_k] N(t_k + t') dt'. \quad (\text{C2})$$

General exponential integrators proceed by breaking the integral in Eq. (C2) into a discrete sum, similar to the

TABLE II. Average runtimes in seconds for  $Z/2$  optimizations using the base, depolarization, sampling ( $S$ ), unscented sampling ( $U$ ), and the first- and second-order derivative methods ( $D1, D2$ ).

Method	$t_N$ (ns)		
	18	36	72
Base	$0.155 \pm 0.008$	$7.0 \pm 0.4$	$15.9 \pm 0.8$
Depolarization	$1.69 \pm 0.08$	...	...
$S$	$1.77 \pm 0.09$	$48 \pm 2$	$280 \pm 10$
$U$	$75 \pm 4$	$340 \pm 20$	$400 \pm 20$
$D1$	$6.1 \pm 0.3$	$27 \pm 1$	$65 \pm 3$
$D2$	$15.7 \pm 0.8$	$17.3 \pm 0.9$	$54 \pm 3$

procedure for Runge-Kutta schemes. We use a simple approximation known as the Lawson-Euler method [86],

$$|\partial_\lambda \psi_{k+1}\rangle \approx \exp(\Delta t_k L_k) |\partial_\lambda \psi_k\rangle + \exp(\Delta t_k L_k) N_k \Delta t_k. \quad (C3)$$

This method provides a good trade-off between accuracy and efficiency, requiring one unique matrix exponential computation per stage. Integration accuracy for the state derivatives is not of the utmost importance because they are used in the robustness cost function—as opposed to the states themselves that are experimental parameters that must be realized with high accuracy.

## APPENDIX D: COMPUTATIONAL PERFORMANCE

In this section we provide runtimes for our optimizations. The runtimes for the base optimization in Sec. III, the depolarization optimization in Sec. IV, and the robust optimizations in Sec. V are presented in Table II for a  $Z/2$  gate at gate times that are multiples of  $1/4f_q \sim 18$  ns. We performed optimizations on a single core of an AMD Ryzen Threadripper 3970X 32-Core Processor @ 3.7 GHz. Future work will parallelize the robustness methods using GPUs [18], which will enable fast optimizations on high-dimensional Hilbert spaces.

experiments using monotonically convergent algorithms, *J. Chem. Phys.* **128**, 184505 (2008).

- [5] N. C. Nielsen, C. Kehlet, S. J. Glaser, and N. Khaneja, in *eMagRes* (American Cancer Society, Kennesaw, GA, USA, 2010).
- [6] 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, *J. Magn. Reson.* **163**, 8 (2003).
- [7] Z. Tošner, T. Vosegaard, C. Kehlet, N. Khaneja, S. J. Glaser, and N. C. Nielsen, Optimal control in NMR spectroscopy: Numerical implementation in SIMPSON, *J. Magn. Reson.* **197**, 120 (2009).
- [8] M. Abdelhafez, B. Baker, A. Gyenis, P. Mundada, A. A. Houck, D. Schuster, and J. Koch, Universal gates for protected superconducting qubits using optimal control, *Phys. Rev. A* **101**, 022321 (2020).
- [9] S. Boutin, C. K. Andersen, J. Venkatraman, A. J. Ferris, and A. Blais, Resonator reset in circuit qed by optimal control for large open quantum systems, *Phys. Rev. A* **96**, 042315 (2017).
- [10] S. Chakram, K. He, A. V. Dixit, A. E. Oriani, R. K. Naik, N. Leung, H. Kwon, W.-L. Ma, L. Jiang, and D. I. Schuster, Multimode photon blockade, *ArXiv:2010.15292* (2020).
- [11] D. J. Egger and F. K. Wilhelm, Optimized controlled-Z gates for two superconducting qubits coupled through a resonator, *Supercond. Sci. Technol.* **27**, 014001 (2013).
- [12] R. Fisher, F. Helmer, S. J. Glaser, F. Marquardt, and T. Schulte-Herbrüggen, Optimal control of circuit quantum electrodynamics in one and two dimensions, *Phys. Rev. B* **81**, 085328 (2010).
- [13] P. Gokhale, Y. Ding, T. Propson, C. Winkler, N. Leung, Y. Shi, D. I. Schuster, H. Hoffmann, and F. T. Chong, in *Proceedings of the 52nd Annual IEEE/ACM International Symposium on Microarchitecture* (Association for Computing Machinery, New York, NY, USA, 2019), p. 266.
- [14] S.-Y. Huang and H.-S. Goan, Optimal control for fast and high-fidelity quantum gates in coupled superconducting flux qubits, *Phys. Rev. A* **90**, 012318 (2014).
- [15] 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, *Nat. Commun.* **8**, 1 (2017).
- [16] J. Kelly, *et al.*, Optimal Quantum Control Using Randomized Benchmarking, *Phys. Rev. Lett.* **112**, 240504 (2014).
- [17] Z. Leng, P. Mundada, S. Ghadimi, and A. Houck, Robust and efficient algorithms for high-dimensional black-box quantum optimization, *ArXiv:1910.03591* (2019).
- [18] N. Leung, M. Abdelhafez, J. Koch, and D. Schuster, Speedup for quantum optimal control from automatic differentiation based on graphics processing units, *Phys. Rev. A* **95**, 042318 (2017).
- [19] S. Li, T. Chen, and Z.-Y. Xue, Fast holonomic quantum computation on superconducting circuits with optimal control, *Adv. Quantum Technol.* **3**, 2000001 (2020).
- [20] P. J. Liebermann and F. K. Wilhelm, Optimal Qubit Control Using Single-Flux Quantum Pulses, *Phys. Rev. Appl.* **6**, 024022 (2016).

[21] P. Reinhold, Ph.D. thesis, Yale University (2019).

[22] P. Rebentrost and F. K. Wilhelm, Optimal control of a leaking qubit, *Phys. Rev. B* **79**, 060507(R) (2009).

[23] P. Rebentrost, I. Serban, T. Schulte-Herbrüggen, and F. K. Wilhelm, Optimal Control of a Qubit Coupled to a Non-Markovian Environment, *Phys. Rev. Lett.* **102**, 090401 (2009).

[24] R. J. Spiteri, M. Schmidt, J. Ghosh, E. Zahedinejad, and B. C. Sanders, Quantum control for high-fidelity multi-qubit gates, *New J. Phys.* **20**, 113009 (2018).

[25] A. Spörl, T. Schulte-Herbrüggen, S. J. Glaser, V. Bergholm, M. J. Storcz, J. Ferber, and F. K. Wilhelm, Optimal control of coupled Josephson qubits, *Phys. Rev. A* **75**, 012302 (2007).

[26] I. Brouzos, A. I. Streltsov, A. Negretti, R. S. Said, T. Caneva, S. Montangero, and T. Calarco, Quantum speed limit and optimal control of many-boson dynamics, *Phys. Rev. A* **92**, 062110 (2015).

[27] G. De Chiara, T. Calarco, M. Anderlini, S. Montangero, P. J. Lee, B. L. Brown, W. D. Phillips, and J. V. Porto, Optimal control of atom transport for quantum gates in optical lattices, *Phys. Rev. A* **77**, 052333 (2008).

[28] M. Grace, C. Brif, H. Rabitz, I. A. Walmsley, R. L. Kosut, and D. A. Lidar, Optimal control of quantum gates and suppression of decoherence in a system of interacting two-level particles, *J. Phys. B: At. Mol. Opt. Phys.* **40**, S103 (2007).

[29] M. H. Goerz, T. Calarco, and C. P. Koch, The quantum speed limit of optimal controlled phasegates for trapped neutral atoms, *J. Phys. B: At. Mol. Opt. Phys.* **44**, 154011 (2011).

[30] J. Guo, X. Feng, P. Yang, Z. Yu, L. Q. Chen, C.-H. Yuan, and W. Zhang, High-performance Raman quantum memory with optimal control in room temperature atoms, *Nat. Commun.* **10**, 148 (2019).

[31] J. H. M. Jensen, J. J. Sørensen, K. Mølmer, and J. F. Sherson, Time-optimal control of collisional  $\sqrt{\text{swap}}$  gates in ultracold atomic systems, *Phys. Rev. A* **100**, 052314 (2019).

[32] 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).

[33] V. Nebendahl, H. Häffner, and C. F. Roos, Optimal control of entangling operations for trapped-ion quantum computing, *Phys. Rev. A* **79**, 012312 (2009).

[34] 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).

[35] S. Rosi, A. Bernard, N. Fabbri, L. Fallani, C. Fort, M. Inguscio, T. Calarco, and S. Montangero, Fast closed-loop optimal control of ultracold atoms in an optical lattice, *Phys. Rev. A* **88**, 021601(R) (2013).

[36] P. Treutlein, T. W. Hänsch, J. Reichel, A. Negretti, M. A. Cirone, and T. Calarco, Microwave potentials and optimal control for robust quantum gates on an atom chip, *Phys. Rev. A* **74**, 022312 (2006).

[37] S. van Frank, M. Bonneau, J. Schmiedmayer, S. Hild, C. Gross, M. Cheneau, I. Bloch, T. Pichler, A. Negretti, T. Calarco, *et al.*, Optimal control of complex atomic quantum systems, *Sci. Rep.* **6**, 34187 (2016).

[38] Y. Chou, S.-Y. Huang, and H.-S. Goan, Optimal control of fast and high-fidelity quantum gates with electron and nuclear spins of a nitrogen-vacancy center in diamond, *Phys. Rev. A* **91**, 052315 (2015).

[39] F. Dolde, V. Bergholm, Y. Wang, I. Jakobi, B. Naydenov, S. Pezzagna, J. Meijer, F. Jelezko, P. Neumann, T. Schulte-Herbrüggen, *et al.*, High-fidelity spin entanglement using optimal control, *Nat. Commun.* **5**, 1 (2014).

[40] 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).

[41] T. Nöbauer, A. Angerer, B. Bartels, M. Trupke, S. Rotter, J. Schmiedmayer, F. Mintert, and J. Majer, Smooth Optimal Quantum Control for Robust Solid-State Spin Magnetometry, *Phys. Rev. Lett.* **115**, 190801 (2015).

[42] F. Poggiali, P. Cappellaro, and N. Fabbri, Optimal Control for One-Qubit Quantum Sensing, *Phys. Rev. X* **8**, 021059 (2018).

[43] 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 Sci.* **2**, 024701 (2020).

[44] J. Tian, T. Du, Y. Liu, H. Liu, F. Jin, R. S. Said, and J. Cai, Optimal quantum optical control of spin in diamond, *Phys. Rev. A* **100**, 012110 (2019).

[45] S. Amri, R. Corgier, D. Sugny, E. M. Rasel, N. Gaaloul, and E. Charron, Optimal control of the transport of Bose-Einstein condensates with atom chips, *Sci. Rep.* **9**, 1 (2019).

[46] P. Doria, T. Calarco, and S. Montangero, Optimal Control Technique for Many-Body Quantum Dynamics, *Phys. Rev. Lett.* **106**, 190501 (2011).

[47] J. J. Sørensen, J. Jensen, T. Heinzel, and J. F. Sherson, QEngine: A C++ library for quantum optimal control of ultracold atoms, *Comput. Phys. Commun.* **243**, 135 (2019).

[48] J. J. W. H. Sørensen, M. O. Aranburu, T. Heinzel, and J. F. Sherson, Quantum optimal control in a chopped basis: Applications in control of Bose-Einstein condensates, *Phys. Rev. A* **98**, 022119 (2018).

[49] P. V. Klimov, J. Kelly, J. M. Martinis, and H. Neven, The snake optimizer for learning quantum processor control parameters, *ArXiv:2006.04594* (2020).

[50] K. Zhou and J. C. Doyle, *Essentials of Robust Control* (Pearson, Upper Saddle River, NJ, USA, 1998), 1st ed.

[51] J. Morimoto and C. Atkeson, Minimax differential dynamic programming: An application to robust biped walking, *Adv. Neural Inf. Process. Syst.* **15**, 1563 (2002).

[52] Z. Manchester and S. Kuindersma, Robust direct trajectory optimization using approximate invariant funnels, *Auton. Robot.* **43**, 375 (2018).

[53] H. K. Cummins and J. A. Jones, Use of composite rotations to correct systematic errors in NMR quantum computation, *New J. Phys.* **2**, 6 (2000).

[54] H. K. Cummins, G. Llewellyn, and J. A. Jones, Tackling systematic errors in quantum logic gates with composite rotations, *Phys. Rev. A* **67**, 042308 (2003).

[55] Å. Kupce and R. Freeman, Stretched adiabatic pulses for broadband spin inversion, *J. Magn. Reson. Ser. A* **117**, 246 (1995).

[56] J. T. Merrill and K. R. Brown, Progress in compensating pulse sequences for quantum computation, *Quantum Inf. Comput. Chem.* **241** (2014).

[57] Z. Han, Y. Dong, B. Liu, X. Yang, S. Song, L. Qiu, D. Li, J. Chu, W. Zheng, J. Xu, *et al.*, Experimental realization of universal time-optimal non-Abelian geometric gates, *ArXiv:2004.10364* (2020).

[58] J. Xu, S. Li, T. Chen, and Z.-Y. Xue, Nonadiabatic geometric quantum computation with optimal control on superconducting circuits, *ArXiv:2004.10199* (2020).

[59] G. Dridi, K. Liu, and S. Guérin, Optimal Robust Quantum Control by Inverse Geometric Optimization, *Phys. Rev. Lett.* **125**, 250403 (2020).

[60] F. Motzoi, J. M. Gambetta, P. Rebentrost, and F. K. Wilhelm, Simple Pulses for Elimination of Leakage in Weakly Nonlinear Qubits, *Phys. Rev. Lett.* **103**, 110501 (2009).

[61] D. J. Egger and F. K. Wilhelm, Adaptive Hybrid Optimal Quantum Control for Imprecisely Characterized Systems, *Phys. Rev. Lett.* **112**, 240503 (2014).

[62] G. Feng, F. H. Cho, H. Katiyar, J. Li, D. Lu, J. Baugh, and R. Laflamme, Gradient-based closed-loop quantum optimal control in a solid-state two-qubit system, *Phys. Rev. A* **98**, 052341 (2018).

[63] J. Li, X. Yang, X. Peng, and C.-P. Sun, Hybrid Quantum-Classical Approach to Quantum Optimal Control, *Phys. Rev. Lett.* **118**, 150503 (2017).

[64] N. Wittler, F. Roy, K. Pack, M. Werninghaus, A. S. Roy, D. J. Egger, S. Filipp, F. K. Wilhelm, and S. Machnes, An integrated tool-set for control, calibration and characterization of quantum devices applied to superconducting qubits, *ArXiv:2009.09866* (2020).

[65] A. Pechen and N. Il'in, Trap-free manipulation in the Landau-Zener system, *Phys. Rev. A* **86**, 052117 (2012).

[66] H. Ball, M. Biercuk, A. Carvalho, J. Chen, M. R. Hush, L. A. de Castro, L. Li, P. J. Liebermann, H. Slatyer, C. Edmunds, V. Frey, C. Hempel, and A. Milne, Software tools for quantum control: Improving quantum computer performance through noise and error suppression, *Quantum Sci. Technol.* **6**, 044011 (2021).

[67] A. R. Carvalho, H. Ball, M. J. Biercuk, M. R. Hush, and F. Thomsen, Error-robust quantum logic optimization using a cloud quantum computer interface, *ArXiv:2010.08057* (2020).

[68] J. Allen, Ph.D. thesis, University of Surrey (2019).

[69] R. L. Kosut, M. D. Grace, and C. Brif, Robust control of quantum gates via sequential convex programming, *Phys. Rev. A* **88**, 052326 (2013).

[70] M. Y. Niu, S. Boixo, V. N. Smelyanskiy, and H. Neven, Universal quantum control through deep reinforcement learning, *NPJ Quantum Inf.* **5**, 33 (2019).

[71] I. R. Petersen, in *Encyclopedia of Systems and Control*, edited by J. Baillieul and T. Samad (Springer London, London, 2013), p. 1.

[72] T. A. Howell, C. Fu, and Z. Manchester, Direct policy optimization using deterministic sampling and collocation, *ArXiv:2010.08506* (2020).

[73] A. Lee, Y. Duan, S. Patil, J. Schulman, Z. McCarthy, J. van den Berg, K. Goldberg, and P. Abbeel, in *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems* (IEEE, New York, NY, USA, 2013), p. 5660.

[74] S. Thangavel, R. Paulen, and S. Engell, Robust multi-stage nonlinear model predictive control using sigma points, *Processes* **8**, 851 (2020).

[75] S. J. Julier and J. K. Uhlmann, Unscented filtering and nonlinear estimation, *Proc. IEEE* **92**, 401 (2004).

[76] J. K. Uhlmann, Ph.D. thesis, University of Oxford Oxford (1995).

[77] H. Zhang, S. Chakram, T. Roy, N. Earnest, Y. Lu, Z. Huang, D. K. Weiss, J. Koch, and D. I. Schuster, Universal Fast-Flux Control of a Coherent, Low-Frequency Qubit, *Phys. Rev. X* **11**, 011010 (2021).

[78] T. A. Howell, B. E. Jackson, and Z. Manchester, in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (IEEE, New York, NY, USA, 2019), p. 7674.

[79] 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).

[80] M. H. Goerz, D. Basilewitsch, F. Gago-Encinas, M. G. Krauss, K. P. Horn, D. M. Reich, and C. P. Koch, Krotov: A Python implementation of Krotov's method for quantum optimal control, *SciPost Phys.* **7**, 080 (2019).

[81] J. Schulman, J. Ho, A. X. Lee, I. Awwal, H. Bradlow, and P. Abbeel, in *Robotics: Science and Systems* (Robotics: Science and Systems Foundation, New York, NY, USA, 2013), Vol. 9, p. 1.

[82] R. Tedrake, and the Drake Development Team, *Drake A planning, control, and analysis toolbox for nonlinear dynamical systems* (2016).

[83] A. Hereid and A. D. Ames, in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (IEEE, Vancouver, BC, Canada, 2017).

[84] L. Jørgensen, D. L. Cardozo, and E. Thibierge, *Numerical Resolution Of The Schrödinger Equation*, Tech. Rep. (École Normale Supérieure de Lyon, 2011).

[85] N. Auer, L. Einkemmer, P. Kandolf, and A. Ostermann, Magnus integrators on multicore CPUs and GPUs, *Comput. Phys. Commun.* **228**, 115 (2018).

[86] H. Berland and B. Skaflestad, Solving the nonlinear Schrödinger equation using exponential integrators, *Int. J. Model. Identif. Control.* **27**, 201 (2006).

[87] L. Einkemmer, M. Tokman, and J. Loffeld, On the performance of exponential integrators for problems in magnetohydrodynamics, *J. Comput. Phys.* **330**, 550 (2017).

[88] R. Shillito, J. A. Gross, A. D. Paolo, Élie Genois, and A. Blais, Fast and differentiable simulation of driven quantum systems, *ArXiv:2012.09282* (2020).

[89] C. R. Hargraves and S. W. Paris, Direct trajectory optimization using nonlinear programming and collocation, *J. Guidance* **10**, 338 (1987).

[90] M. Kelly, An introduction to trajectory optimization: How to do your own direct collocation, *SIAM Rev.* **59**, 849 (2017).

[91] J. T. Betts, Survey of numerical methods for trajectory optimization, *J. Guid. Control Dyn.* **21**, 193 (1998).

[92] P. E. Gill, W. Murray, and M. A. Saunders, SNOPT: An SQP algorithm for large-scale constrained optimization, *SIAM Rev.* **47**, 99 (2005).

[93] A. Wächter and L. T. Biegler, On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming, *Math. Program.* **106**, 25 (2006).

[94] W. Li and E. Todorov, in *Proceedings of the 1st International Conference on Informatics in Control, Automation and Robotics* (Setubal, Portugal, 2004).

[95] G. Lantoine and R. P. Russell, A hybrid differential dynamic programming algorithm for constrained optimal control problems. Part 1: Theory, *J. Optim. Theory Appl.* **154**, 382 (2012).

[96] B. Plancher, Z. Manchester, and S. Kuindersma, in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems* (IEEE, New York, NY, USA, 2017), p. 5674.

[97] J. Nocedal and S. Wright, *Numerical Optimization* (Springer Science & Business Media, New York, NY, USA, 2006).

[98] D. P. Bertsekas, Projected Newton methods for optimization problems with simple constraints, *SIAM J. Control Optim.* **20**, 221 (1982).

[99] C. V. Rao, S. J. Wright, and J. B. Rawlings, Application of interior-point methods to model predictive control, *J. Optim. Theory Appl.* **99**, 723 (1998).

[100] D. Q. Mayne, A second-order gradient method of optimizing non-linear discrete time systems, *Int. J. Control.* **3**, 8595 (1966).

[101] L. Zhang, W. Zhou, and D. Li, Global convergence of a modified Fletcher–Reeves conjugate gradient method with Armijo-type line search, *Numer. Math.* **104**, 561 (2006).

[102] D. P. Bertsekas, *Constrained Optimization and Lagrange Multiplier Methods* (Academic Press, New York, NY, USA, 2014).

[103] B. E. Jackson, T. Punnoose, D. Neamati, K. Tracy, R. Jitosh, and Z. Manchester, in *International Conference on Robotics and Automation ICRA* (2021), in Review.

[104] N. Earnest, S. Chakram, Y. Lu, N. Irons, R. K. Naik, N. Leung, L. Ocola, D. A. Czaplewski, B. Baker, J. Lawrence, J. Koch, and D. I. Schuster, Realization of a  $\Lambda$  System with Metastable States of a Capacitively Shunted Fluxonium, *Phys. Rev. Lett.* **120**, 150504 (2018).

[105] Y.-H. Lin, L. B. Nguyen, N. Grabon, J. San Miguel, N. Pankratova, and V. E. Manucharyan, Demonstration of Protection of a Superconducting Qubit from Energy Decay, *Phys. Rev. Lett.* **120**, 150503 (2018).

[106] V. E. Manucharyan, J. Koch, L. I. Glazman, and M. H. Devoret, Fluxonium: Single cooper-pair circuit free of charge offsets, *Science* **326**, 113 (2009).

[107] A. Somoroff, Q. Ficheux, R. A. Mencia, H. Xiong, R. V. Kuzmin, and V. E. Manucharyan, Millisecond coherence in a superconducting qubit, *ArXiv:2103.08578* (2021).

[108] L. B. Nguyen, Y.-H. Lin, A. Somoroff, R. Mencia, N. Grabon, and V. E. Manucharyan, High-Coherence Fluxonium Qubit, *Phys. Rev. X* **9**, 041041 (2019).

[109] M. A. Rol, F. Battistel, F. K. Malinowski, C. C. Bultink, B. M. Tarasinski, R. Vollmer, N. Haider, N. Muthusubramanian, A. Bruno, B. M. Terhal, and L. DiCarlo, Fast, High-Fidelity Conditional-Phase Gate Exploiting Leakage Interference in Weakly Anharmonic Superconducting Qubits, *Phys. Rev. Lett.* **123**, 120502 (2019).

[110] P. Krantz, M. Kjaergaard, F. Yan, T. P. Orlando, S. Gustavsson, and W. D. Oliver, A quantum engineer’s guide to superconducting qubits, *Appl. Phys. Rev.* **6**, 021318 (2019).

[111] T. Schulte-Herbrüggen, A. Spörl, N. Khaneja, and S. J. Glaser, Optimal control for generating quantum gates in open dissipative systems, *J. Phys. B: At. Mol. Opt. Phys.* **44**, 154013 (2011).

[112] M. Abdelhafez, D. I. Schuster, and J. Koch, Gradient-based optimal control of open quantum systems using quantum trajectories and automatic differentiation, *Phys. Rev. A* **99**, 052327 (2019).

[113] Z. Huang, P. S. Mundada, A. Gyenis, D. I. Schuster, A. A. Houck, and J. Koch, Engineering Dynamical Sweet Spots to Protect Qubits from  $1/f$  Noise, *Phys. Rev. Appl.* **15**, 034065 (2021).

[114] P. S. Mundada, A. Gyenis, Z. Huang, J. Koch, and A. A. Houck, Floquet-engineered enhancement of coherence times in a driven fluxonium qubit, *ArXiv:2007.13756* (2020).

[115] D. Aharonov and M. Ben-Or, Fault-tolerant quantum computation with constant error rate, *SIAM J. Comput.* **38**, 1207 (2008).

[116] E. Knill, Quantum computing with realistically noisy devices, *Nature* **434**, 39 (2005).

[117] D. Gottesman, Stabilizer codes and quantum error correction, *ArXiv:quant-ph/9705052* (1997).

[118] A. Paetznick, Ph.D. thesis, University of Waterloo (2014).

[119] M. Suchara, A. Faruque, C.-Y. Lai, G. Paz, F. T. Chong, and J. Kubiatowicz, Comparing the overhead of topological and concatenated quantum error correction, *ArXiv:1312.2316* (2013).

[120] J. M. Chow, J. M. Gambetta, L. Tornberg, J. Koch, L. S. Bishop, A. A. Houck, B. R. Johnson, L. Frunzio, S. M. Girvin, and R. J. Schoelkopf, Randomized Benchmarking and Process Tomography for Gate Errors in a Solid-State Qubit, *Phys. Rev. Lett.* **102**, 090502 (2009).

[121] R. C. Bialczak, R. McDermott, M. Ansmann, M. Hofheinz, N. Katz, E. Lucero, M. Neeley, A. D. O’Connell, H. Wang, A. N. Cleland, and J. M. Martinis,  $1/f$  Flux Noise in Josephson Phase Qubits, *Phys. Rev. Lett.* **99**, 187006 (2007).

[122] K. Kakuyanagi, T. Meno, S. Saito, H. Nakano, K. Semba, H. Takayanagi, F. Deppe, and A. Shnirman, Dephasing of a Superconducting Flux Qubit, *Phys. Rev. Lett.* **98**, 047004 (2007).

[123] P. Kumar, S. Sendelbach, M. A. Beck, J. W. Freeland, Z. Wang, H. Wang, C. C. Yu, R. Q. Wu, D. P. Pappas, and R. McDermott, Origin and Reduction of  $1/f$  Magnetic Flux Noise in Superconducting Devices, *Phys. Rev. Appl.* **6**, 041001(R) (2016).

[124] F. Yoshihara, K. Harrabi, A. O. Niskanen, Y. Nakamura, and J. S. Tsai, Decoherence of Flux Qubits due to  $1/f$  Flux Noise, *Phys. Rev. Lett.* **97**, 167001 (2006).

[125] R. H. Koch, D. P. DiVincenzo, and J. Clarke, Model for 1/ $f$  Flux Noise in SQUIDs and Qubits, *Phys. Rev. Lett.* **98**, 267003 (2007).

[126] F. Yoshihara, Y. Nakamura, and J. S. Tsai, Correlated flux noise and decoherence in two inductively coupled flux qubits, *Phys. Rev. B* **81**, 132502 (2010).

[127] E. L. Hahn and D. E. Maxwell, Spin echo measurements of nuclear spin coupling in molecules, *Phys. Rev.* **88**, 1070 (1952).

[128] S. Meiboom and D. Gill, Modified spin-echo method for measuring nuclear relaxation times, *Rev. Sci. Instrum.* **29**, 688 (1958).

[129] J. O. Smith, *Spectral Audio Signal Processing* (2020), online book, 2011 ed.

[130] <https://github.com/SchusterLab/rbqoc>.

[131] C. Rackauckas and Q. Nie, DifferentialEquations.jl—A performant and feature-rich ecosystem for solving differential equations in Julia, *J. Open Res. Softw.* **5**, 15 (2017).

[132] M. Besançon, D. Anthoff, A. Arslan, S. Byrne, D. Lin, T. Papamarkou, and J. Pearson, Distributions.jl: Definition and modeling of probability distributions in the JuliaStats ecosystem, [ArXiv:1907.08611](https://arxiv.org/abs/1907.08611) (2019).

[133] J. Revels, M. Lubin, and T. Papamarkou, Forward-mode automatic differentiation in Julia, [ArXiv:1607.07892](https://arxiv.org/abs/1607.07892) (2016).

[134] J. D. Hunter, Matplotlib: A 2D graphics environment, *Comput. Sci. Eng.* **9**, 90 (2007).

[135] C. R. Harris, *et al.*, Array programming with NumPy, *Nature* **585**, 357 (2020).

[136] M. Innes, Don't unroll adjoint: differentiating SSA-form programs, [ArXiv:1810.07951](https://arxiv.org/abs/1810.07951) (2018).

[137] D. Manzano, A short introduction to the Lindblad master equation, *AIP Adv.* **10**, 025106 (2020).

[138] G. T. Landi, *Quantum information and quantum noise* (Instituto de Física da Universidade de São Paulo, 2018).

[139] S. J. Julier and J. K. Uhlmann, in *Proceedings of the 2002 American Control Conference (IEEE Cat. No. CH37301)* (IEEE, 2002), Vol. 2, p. 887.