

Robust Quantum Optimal Control

Thomas Propson,^{1,2,*} Brian Jackson,³ Zachary Manchester,³ and David I. Schuster^{1,2,4}

¹*James Franck Institute, University of Chicago, Chicago, Illinois 60637, USA*

²*Department of Physics, University of Chicago, Chicago, Illinois 60637, USA*

³*Robotics Institute, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, USA*

⁴*Pritzker School of Molecular Engineering, University of Chicago, Chicago, Illinois 60637, USA*

(Dated: January 12, 2021)

The ability to engineer high-fidelity gates on quantum processors in the presence of systematic errors remains the primary challenge requisite 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 method that takes advantage of uncertain parameter derivative information while maintaining computational efficiency by transforming high-order differential equations to coupled, first-order ODEs. Additionally, the effect of depolarization on a gate is most accurately modeled by integrating the Lindblad master equation, which is computationally expensive. We propose 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 with realistic parameters and constraints, achieving orders of magnitude gate-error reductions from our baseline gate set in simulations.

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 [8–16], neutral atoms and ions [17–27], nitrogen-vacancy centers in diamond [28–34], and Bose-Einstein condensates [35, 36]. For quantum computation, optimal control is employed to achieve high-fidelity gates while adhering to experimental constraints. Experimental errors may cause the system to deviate from the model used in optimization, leading to poor experimental performance. Robust control improves upon standard optimal control by encoding model parameter uncertainties in optimization objectives, yielding performance guarantees over a range of parameter values [37–39]. We adapt robust control techniques from the robotics community to mitigate parameter uncertainty errors for a superconducting fluxonium qubit.

QOC has had tremendous success in engineering high-fidelity quantum gates [13, 28, 29, 40–51]. However, the system often deviates from the model used in optimization due to experimental errors such as parameter drift, noise, finite control resolution, and decoherence. Multiple techniques have been developed to address this shortcoming. Analytically-derived control pulses that mitigate parameter uncertainty errors include composite pulses [52–55], pulses designed by considering dynamic and geometric phases [16, 56], and pulses obtained with the DRAG scheme [57]. To suppress decoherence, Floquet techniques have been employed [12, 58]. Numerical schemes to mitigate parameter uncertainty errors

include closed-loop methods [41, 59, 60] and open-loop methods [33, 61–64]. To model decoherence, numerical techniques typically employ master equations [33, 65] or Monte Carlo-style quantum trajectories [66].

In this work, we study three robust control techniques that make the system’s quantum state trajectory less sensitive to static and time-dependent parameter uncertainty:

1. A sampling method, similar to the work of [3, 33, 61–63].
2. An unscented sampling method adapted from the unscented transform used in state estimation [67–70].
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 [71]. We also show that QOC can solve important problems associated with fluxonium-based qubits: mitigating depolarization by taking advantage of the T_1 -dependence of the controls and synchronizing qubits with distinct frequencies by performing phase gates in arbitrary times. To mitigate depolarization, we perform time-optimal control and employ an efficient depolarization model that is significantly less expensive to compute than integrating a master equation. Leveraging recent advances in trajectory optimization within the field of robotics, we solve these optimization problems using ALTRO (Augmented Lagrangian TRajjectory Optimizer) [72], which uses the iterative linear quadratic regulator (iLQR) method [73] within an augmented Lagrangian framework [74, 75] to handle nonlinear equality and inequality constraints at each time step.

This paper is organized as follows. First, we reintroduce ALTRO in the context of QOC in Section II. We

* tcropson@uchicago.edu

describe realistic constraints for operating the fluxonium and define the associated QOC problem in Section III. Then, we outline a method for making the optimization aware of depolarization in Section IV. Next, we outline three techniques for achieving robustness to static parameter uncertainties in Section V. We adapt the same techniques to mitigate $1/f$ flux noise in Section VI.

II. BACKGROUND

In this section, we review the QOC problem statement and reintroduce the ALTRO solver [72]. QOC concerns a vector of time-dependent control parameters $u(t)$ 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),

$$i\hbar \frac{d}{dt} |\psi(t)\rangle = H(u(t), t) |\psi(t)\rangle \quad (1)$$

The Hamiltonian $H(u(t), t)$ is determined by the quantum system. The QOC problem is to find the controls that minimize a functional $J(u(t))$. To make the problem numerically tractable, the controls and state are discretized into N knot points (time steps). In the case of a single state transfer problem, the functional is the infidelity of the desired final state and the initial state evolved to the final knot point, $J(u) = 1 - |\langle \psi_f | \psi_N(u) \rangle|^2$. In general, $J(u)$ is a linear combination of cost functions on the state as well as cost functions on the controls [14]. Standard QOC solvers compute derivatives of the functional $\nabla_u J(u)$, which are used to implement first-order optimization methods [3, 14, 76, 77].

REM: Alternatively, the QOC problem can be formulated as a trajectory optimization problem and solved using a variety of specialized solvers developed by the robotics community [72, 78–80]. Trajectory optimization problems are typically of the following form:

$$\underset{x_{1:N}, u_{1:N-1}}{\text{minimize}} \quad \ell_f(x_N) + \sum_{k=1}^{N-1} \ell(x_k, u_k) \quad (2a)$$

$$\text{subject to} \quad x_{k+1} = f(x_k, u_k), \quad (2b)$$

$$g_k(x_k, u_k) \leq 0, \quad (2c)$$

$$h_k(x_k, u_k) = 0 \quad (2d)$$

where ℓ_f and ℓ are the final and stage cost functions, $x_k \in \mathbb{R}^l$ and $u_k \in \mathbb{R}^m$ are the state and control variables, $f(x_k, u_k)$ is the discrete dynamics function, and $g_k(x_k, u_k)$ and $h_k(x_k, u_k)$ are the inequality and equality constraints, potentially including initial and final conditions, at time step k .

ADD: Alternatively, the QOC problem can be formulated as a trajectory optimization problem and solved using a variety of specialized solvers developed by the robotics community [72, 78–80]. The functional $J(u)$ is divided into its constituent cost functions at each knot

point $\ell_k(x_k, u_k)$, where $k \in \{1, \dots, N\}$ denotes the knot point, x_k is the augmented state vector, and u_k is the augmented control vector. The augmented state includes all variables of interest that are not decision variables. In the case of QOC, the state is a subset of the augmented state $|\psi_k\rangle \subseteq x_k$. The solver we use does not currently support complex numbers, so we employ the isomorphism $\mathcal{H}(\mathbb{C}^n) \cong \mathcal{H}(\mathbb{R}^{2n})$ given in [14],

$$H |\psi\rangle \cong \begin{bmatrix} H_{\text{re}} & -H_{\text{im}} \\ H_{\text{im}} & H_{\text{re}} \end{bmatrix} \begin{bmatrix} |\psi\rangle_{\text{re}} \\ |\psi\rangle_{\text{im}} \end{bmatrix} \quad (3)$$

The augmented state trajectory obeys the physics of the system if the dynamics constraint is satisfied $x_{k+1} = f(x_k, u_k)$. For QOC, the discrete dynamics function $f(x_k, u_k)$ propagates the state by integrating the TDSE 1 with Runge-Kutta methods [81] or exponential integrators [82–84].

ADD: Additional constraints on the augmented states and augmented controls are encoded in constraint functions. The constraint functions are put into a form such that, when the constraint is satisfied, inequality constraint functions obey $g_k(x_k, u_k) \leq 0$ and equality constraint functions obey $h_k(x_k, u_k) = 0$. The constraint functions may be vector-valued to encode multiple constraints of each type. If a constraint function takes a positive value, the value is called the constraint’s violation and the constraint is said to be active. The trajectory optimization problem can be stated concisely,

$$\underset{x_{1:N}, u_{1:N-1}}{\text{minimize}} \quad \ell_N(x_N) + \sum_{k=1}^{N-1} \ell_k(x_k, u_k) \quad (4a)$$

$$\text{subject to} \quad x_{k+1} = f(x_k, u_k), \quad (4b)$$

$$g_k(x_k, u_k) \leq 0, \quad (4c)$$

$$h_k(x_k, u_k) = 0 \quad (4d)$$

REM: Many techniques have been proposed for solving (4a)–(4d). Standard methods include direction collocation [85] and differential-dynamic programming (DDP) [86]. Recent state-of-the-art solvers, such as ALTRO [72], have combined principles from both of these approaches.

ADD: Many techniques have been proposed for solving (4a)–(4d). Standard methods include direction collocation [85] and differential-dynamic programming (DDP) [86]. Recent state-of-the-art solvers, such as ALTRO, have combined principles from both of these approaches. ALTRO uses iLQR [73] as the internal solver of an augmented Lagrangian method (ALM) [74, 75] and employs a projected Newton method [87, 88] in its final solving stage.

REM: ALTRO uses iterative LQR (iLQR) [73] as the internal solver of an augmented Lagrangian method (ALM). iLQR solves an unconstrained trajectory optimization problem using a backward Riccati recursion to derive a closed-loop linear feedback law about the current trajectory. By simulating the system forward with

the feedback law, the trajectory is brought closer to the (locally) optimal trajectory. DDP-based solvers such as iLQR are popular since they are very computationally efficient, are always dynamically feasible, and provide a locally stabilizing closed-loop control policy about the optimal trajectory. However, standard implementations have no ability to deal with nonlinear equality and inequality constraints. ALM handles constraints by successively solving unconstrained minimization problems of the form:

$$\underset{z}{\text{minimize}} \quad f(z) + \lambda^T c(z) + \frac{1}{2} c(z)^T I_\mu c(z) \quad (5)$$

where $f(z)$ is the objective function, $c(z) : \mathbb{R}^{l+m} \mapsto \mathbb{R}^p$ is the constraint function, $\lambda \in \mathbb{R}^p$ is a Lagrange multiplier estimate, and I_μ is a diagonal matrix of penalty weights, μ , whose magnitudes depend on whether the corresponding constraint is active or inactive. For ALTRO, $f(z)$ is (4a), $c(z)$ is the concatenation of g_k and h_k , and z is the concatenation of the states and controls across all time steps. After minimizing (7) using iLQR, the Lagrange multiplier estimates are updated according to,

$$\lambda \leftarrow \lambda + \mu c(z) \quad (6)$$

the penalty terms are updated, and the process repeats until convergence.

ADD: iLQR simulates the augmented state forward using the discrete dynamics function and computes derivatives of the cost functions to update the augmented controls. This procedure is known as indirect shooting, and iLQR shares this approach with GOAT [76], GRAPE [3, 14], and Krotov's method [77]. iLQR differs from these solvers in its update procedure. iLQR computes the locally optimal update for the augmented controls using a dynamic programming algorithm, and a line search is performed in the direction of this update to ensure a decrease in the sum of the cost functions.

ADD: Indirect shooting solvers such as iLQR are popular for solving unconstrained trajectory optimization problems (4a)-(4b), but standard implementations have no ability to handle nonlinear equality and inequality constraints (4c)-(4d). One way to handle constraints is to employ a projected gradient method [89, 90]. However, within the indirect shooting framework, projected gradient methods can only handle constraints on the augmented controls, not the augmented states. Another technique is to promote the constraint functions to cost functions [14]. A simple implementation of this approach does not guarantee that the constraints are satisfied, as the solver trades minimization of the cost functions and constraint functions against each other. ALM promotes the constraint functions to cost functions, and adaptively adjusts a Lagrange multiplier estimate for each constraint function to ensure the constraints are satisfied. ALM considers a decision variable $z \in \mathbb{R}^s$ with s given, a cost function $f(z) : \mathbb{R}^s \mapsto \mathbb{R}$, and a constraint function $c(z) : \mathbb{R}^s \mapsto \mathbb{R}^p$ encoding p constraints. For ALTRO, z

is the vector concatenation of the augmented states and augmented controls across all knot points, $f(z)$ is the sum of the trajectory optimization cost functions (4a), and $c(z)$ is the vector concatenation of the trajectory optimization constraint functions (4c)-(4d). ALM successively solves unconstrained minimization problems of the form:

$$\underset{z}{\text{minimize}} \quad f(z) + \lambda^T c(z) + \frac{1}{2} c(z)^T I_\mu c(z) \quad (7)$$

$\lambda \in \mathbb{R}^p$ is a Lagrange multiplier estimate, and I_μ is a diagonal matrix of penalty weights, $\mu \in \mathbb{R}^p$, whose magnitudes depend on whether the corresponding constraint is active or inactive. After minimizing (7) using iLQR, the Lagrange multiplier estimates are updated according to,

$$\lambda \leftarrow \lambda + I_\mu c(z) \quad (8)$$

the penalty weight is updated, and the process repeats until convergence.

REM: ALM converges superlinearly but tends to exhibit slow constraint convergence near the optimal solution due to poor numerical conditioning. To address this shortcoming, ALTRO provides a solution-polishing phase that takes 1-2 Newton steps on the active constraint set to provide machine-precision constraint satisfaction. For more information on the details of the ALTRO solver, see [72, 91].

ADD: ALM converges superlinearly, but poor numerical conditioning may lead to small decreases in the constraint violations near the locally optimal solution [92]. To address this shortcoming, ALTRO projects the solution from the ALM solving stage onto the constraint manifold using a projected Newton method, achieving ultra-low constraint violations $\sim 1 \cdot 10^{-8}$. For more information on the details of the ALTRO solver, see [72, 91].

REM: As opposed to standard QOC solvers, the ALTRO solver provides a unified framework for imposing arbitrary, simultaneous constraints on both the states and the controls that converge superlinearly to tight tolerances. First-order solvers such as GRAPE use unconstrained optimization to satisfy desiderata on the states, such as achieving a desired gate fidelity, forbidding occupations of particular quantum states, and smoothing the controls below a threshold. The unconstrained approach relies on a careful choice of competing penalty weights, which is cumbersome and often intractable. To impose constraints on the controls, such as amplitude restrictions, first-order solvers employ projected gradient descent methods which restrict optimization to the constraint manifold and may hinder convergence to the optimal solution [89, 90]. Conversely, the ALTRO solver delays projecting onto the constraint manifold defined by (4c) and (4d) until the ALM solve is already at a coarse tolerance, achieving machine-precision satisfaction for constraints on the controls and the states. These advantages are critical for this work, where multiple con-

straints must be satisfied in addition to minimizing auxiliary metrics.

ADD: As opposed to standard QOC solvers, ALTRO is able to satisfy constraints on both the augmented states and the augmented controls 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 (unitary transformations) for the fluxonium qubit. The fluxonium is a promising building block for quantum computers, and the accurate two-level approximation of its Hamiltonian makes QOC on a classical computer inexpensive. In the two-level approximation, the Hamiltonian takes the form:

$$H/h = f_q \frac{\sigma_z}{2} + a(t) \frac{\sigma_x}{2} \quad (9)$$

where f_q is the qubit frequency at the flux frustration point, $a(t)$ is the flux-drive amplitude, h is Planck's constant, and σ_z, σ_x are Pauli matrices. We optimize $X/2$, $Y/2$, and $Z/2$ gates for the fluxonium presented in [71], and compare them to the analytically constructed gates for that device.

REM: The optimization problem takes the form:

$$\text{minimize}_{x_{1:N}, u_{1:N-1}} \sum_{k=1}^N \|x_k - x_f\|_{Q_k} + \sum_{k=1}^{N-1} \|u_k\|_{R_k} \quad (10a)$$

$$\text{subject to } x_{k+1} = f(x_k, u_k), \quad (10b)$$

$$|\psi_1^0\rangle = |0\rangle, |\psi_1^1\rangle = |1\rangle, \quad (10c)$$

$$|\psi_N^i\rangle = |\psi_f^i\rangle \quad \forall i \in \{0, 1\}, \quad (10d)$$

$$|\langle \psi_k^i | \psi_k^i \rangle|^2 = 1 \quad \forall i \in \{0, 1\}, \quad (10e)$$

$$\int_0^{t_1} a \, dt = \partial a_1 / \partial t = 0, \quad (10f)$$

$$\int_0^{t_N} a \, dt = 0, \quad (10g)$$

$$a_1 = a_N = 0, \quad (10h)$$

$$|a_k| \leq 0.5 \text{ GHz} \quad (10i)$$

The augmented state and augmented controls are:

$$x = \begin{bmatrix} \psi^0 \\ \psi^1 \\ \int_0^t a \, dt \\ a \\ \partial a / \partial t \end{bmatrix} \quad u = [\partial^2 a / \partial t^2] \quad (11)$$

The matrices Q_k and R_k define the penalty metric. Including objectives at each knot point smoothens the optimization landscape, though the importance of the final objective is encoded in the relative weight $Q_N \sim N \cdot Q_k$. We choose Q to be diagonal because it is computationally

efficient. This metric penalizes phase differences between the states, although phase-insensitive metrics such as infidelity may be employed.

REM: We impose constraints that reflect the physical limitations of the apparatus and improve the experimental realization of the control pulse. The discrete dynamics (14b) integrates the states according to the TDSE (1) and the fluxonium Hamiltonian (9) and integrates the moments of the flux amplitude. Exposing lower order moments of the flux amplitude allows us to penalize their norms in (14a), smoothing the flux amplitude and mitigating AWG ringing due to high frequency transitions. We ensure the desired gate is achieved by imposing the constraints (14c) and (14d). The initial states in (14c) are chosen to span the Hilbert space. The final states in (14d) are the image of the initial states under the desired gate. The ALTRO implementation we use does not currently support complex numbers so we compute in the isomorphism $\mathcal{H}(\mathbb{C}^n) \cong \mathcal{H}(\mathbb{R}^{2n})$ given in [14],

$$H |\psi\rangle \cong \begin{bmatrix} H_{\text{re}} & -H_{\text{im}} \\ H_{\text{im}} & H_{\text{re}} \end{bmatrix} \begin{bmatrix} \psi_{\text{re}} \\ \psi_{\text{im}} \end{bmatrix} \quad (12)$$

We ensure the solver does not take advantage of discretization error by imposing the normalization constraint (14e). The constraint (10f) is an initial condition. We ensure the pulse has zero net flux by imposing (14f), which mitigates the hysteresis ubiquitous in flux bias lines [71, 93, 94]. We ensure gates may be concatenated arbitrarily without inducing AWG ringing due to high-frequency transitions by imposing the boundary condition (10h). Finally, we ensure the two-level approximation (9) remains valid by imposing the amplitude constraint (14g). All of the gates presented in this work achieve a maximum individual constraint violation of 10^{-8} for (14b)–(14g).

ADD: First, we outline the constraints for the fluxonium gate problem. We formulate this problem as a multi-state-transfer problem. The initial conditions on the states are $|\psi_1^0\rangle = |0\rangle$, $|\psi_1^1\rangle = |1\rangle$ (14c) where the superscript is an index $i \in \{0, 1\}$, and the subscript indicates the first knot point $k = 1$. The states at the final knot point are constrained to be the image of the initial states under the desired gate U , $|\psi_N^i\rangle = U |\psi_1^i\rangle \quad \forall i$ (14d). Furthermore, we impose the normalization constraint $|\langle \psi_k^i | \psi_k^i \rangle|^2 = 1 \quad \forall i, k$ (14e) to ensure the solver does not take advantage of discretization errors in numerical integration. To refer to the discrete moments the flux amplitude, we introduce the notation $\int_{t_1}^{t_k} a(t) \, dt \equiv \int_{t_1}^{t_k} a_k$, $a(t)|_{t=t_k} \equiv a_k$, $\partial^n a(t) / \partial t^n |_{t=t_k} \equiv \partial_t^n a_k$. We impose the zero net flux constraint $\int_t a_N = 0$ (14f) which mitigates the hysteresis ubiquitous in flux-bias lines [71, 93, 94]. The flux amplitude is constrained by $|a_k| \leq 0.5 \text{ GHz} \quad \forall k$ (14g) to ensure the two-level approximation (9) remains valid. We also enforce the boundary condition $a_1 = a_N = 0$ (14h) so the gates may be concatenated arbitrarily. Additionally, we have the initial condition $\int_t a_1 = \partial_t a_1 = 0$ (14i).

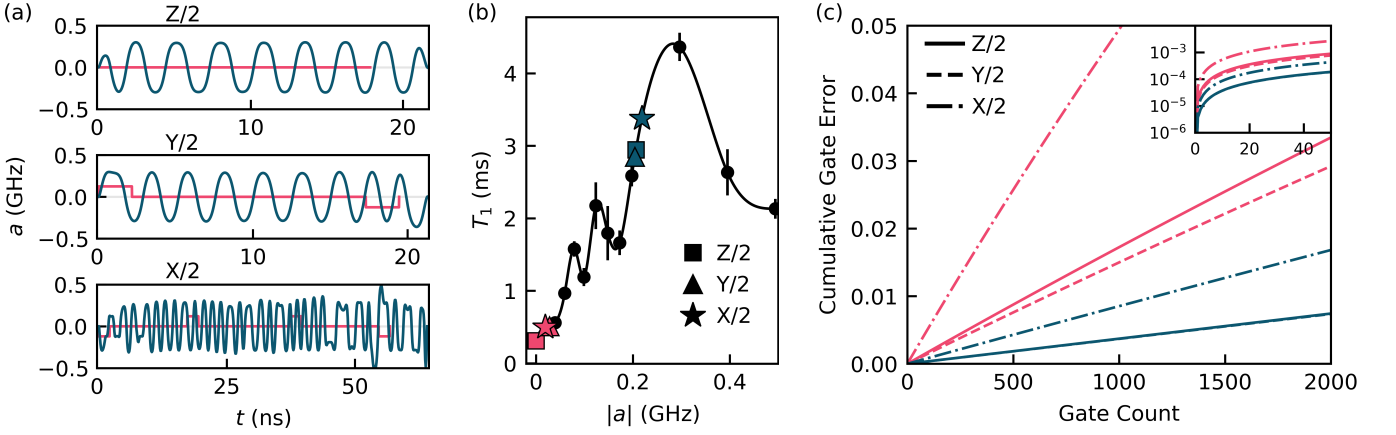


Figure 1: (a) Numerically optimized gates (dark blue) and analytically optimized gates (light pink). (b) T_1 interpolation function used in optimization. Markers denote the time-averaged, absolute amplitude of each gate. (c) Cumulative gate errors due to depolarization as a function of the number of gates applied. Gate errors for the numerically optimized Z/2 and Y/2 gates are indistinguishable. Inset shows the cumulative gate errors plotted on a log scale for small gate counts.

ADD: Next, we outline the cost functions for the fluxonium gate problem. The augmented state and augmented control are:

$$x_k = \begin{bmatrix} |\psi_k^0\rangle \\ |\psi_k^1\rangle \\ \int_t a_k \\ a_k \\ \partial_t a_k \end{bmatrix} \quad u_k = \begin{bmatrix} \partial_t^2 a_k \end{bmatrix} \quad (13)$$

The cost function at each knot point is $\ell_k(x_k, u_k) = (x_k - x_f)^T Q_k (x_k - x_f) + u_k^T R_k u_k$ where Q_k and R_k are matrices we supply, and x_f is given by the constraints we have imposed on $|\psi_N^i\rangle$, $\int_t a_N$, and a_N in addition to $\partial_t a_f = 0$. We choose Q_k to be diagonal so the cost function incentivizes the achievement of the constraints on $|\psi_N^i\rangle$, $\int_t a_N$, and a_N , and penalizes the norm of $\partial_t a_k$ and $\partial_t^2 a_k$, mitigating AWG ringing due to high-frequency transitions. In the discrete dynamics function for this problem (14b), we integrate the states according to the TDSE (1) and the fluxonium Hamiltonian (9) and integrate the second-derivative of the flux amplitude to obtain the first-derivative, proportional, and integral moments. Stated succinctly, the optimization problem takes

the form:

$$\underset{x_{1:N}, u_{1:N-1}}{\text{minimize}} \quad \sum_{k=1}^N (x_k - x_f)^T Q_k (x_k - x_f) + \sum_{k=1}^{N-1} u_k^T R_k u_k \quad (14a)$$

$$\text{subject to} \quad x_{k+1} = f(x_k, u_k), \quad (14b)$$

$$|\psi_1^0\rangle = |0\rangle, |\psi_1^1\rangle = |1\rangle, \quad (14c)$$

$$|\psi_N^i\rangle = U |\psi_1^i\rangle \quad \forall i, \quad (14d)$$

$$|\langle \psi_k^i | \psi_k^i \rangle|^2 = 1 \quad \forall i, k, \quad (14e)$$

$$\int_t a_N = 0, \quad (14f)$$

$$|a_k| \leq 0.5 \text{ GHz} \quad \forall k, \quad (14g)$$

$$a_1 = a_N = 0, \quad (14h)$$

$$\int_t a_1 = \partial_t a_1 = 0 \quad (14i)$$

IV. DEPOLARIZATION AWARENESS

The strength of depolarization varies with the control parameters in a range of superconducting circuit platforms, including the fluxonium. In this section, we outline a method for tuning the flux amplitude to mitigate depolarization. Previous work has modeled the gate error due to depolarization by evolving density matrices under a master equation [33, 65] or evolving a large number of states in a quantum trajectory approach [66]. We avoid the increase in computational complexity required for these techniques by penalizing the integrated rate (probability) of depolarization in optimization. Using this probability as a proxy for the gate error incurred is reasonable because losses due to depolarization increase

monotonically in time. Additionally, for a constant T_1 time—the $1/e$ depolarization time for the qubit—a shorter gate duration would favor a lower depolarization probability. We allow the optimizer to tune the gate time to minimize the depolarization probability.

The depolarization probability is given by,

$$P_1(t) = \int_0^t T_1^{-1}(a(t')) dt' \quad (15)$$

This value is appended to the augmented state (13) and its norm is penalized in the objective (14a) by setting the corresponding element of the final augmented state x_f to 0. $T_1(a_k)$ is obtained at each knot point by evaluating a spline fit to experimental data of the form $\{(a, T_1)\}$, see Figure 1b. It is also possible to fit a spline to theoretically obtained data. However, T_1 values are known to fluctuate greatly with laboratory temperatures [95]. Interpolating T_1 from experimental data increases the fridge truth of the simulation.

We allow the optimizer to tune the gate time by making the time step between each knot point Δt_k a decision variable. Promoting Δt_k to a decision variable, rather than the number of knot points N , preserves the Markovianity of the trajectory optimization problem. The square root of the time step $\sqrt{\Delta t_k}$ is appended to the augmented control (13) and the squared root of the time step $|\Delta t_k|$ is used for numerical integration in the discrete dynamics function (14b). To ensure numerical integration accuracy is maintained, we constrain the bounds of the time step at each knot point.

We compare the performance of the $X/2$, $Y/2$, and $Z/2$ gates obtained with our numerical method to the analytic gates. 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 flux amplitudes for the numerical gates are similar; the waveforms are approximately periodic with amplitudes $\sim 0.2\text{GHz}$, see Figure 1a. **TODO: Floquet.** Their gate times are greater than their analytic counterparts, but they reach higher amplitudes and therefore higher T_1 times. Although the analytic and numerical gates attain single gate errors sufficient for quantum error correction ($< 10^{-4}$), which are required for fault-tolerant quantum computing [96–98], the single gate errors for the numerical gates are a factor of 5 less than those for the analytic gates, commensurate to the depolarization probability reductions, see Appendix A. The gate error reported in this text is the infidelity of the evolved state and the target state averaged over 1000 pseudo-randomly generated initial states. The numerical $Z/2$ and $Y/2$ gates produce similar results in the concatenated gate application comparison, yielding cumulative gate errors below $8 \cdot 10^{-3}$ over 2000 gate applications $\sim 40\mu\text{s}$, see Figure 1c. The cumulative gate error for the numerical $X/2$ gate is $1.7 \cdot 10^{-2}$ over 2000 gate applications $\sim 124\mu\text{s}$. The low cumulative gate errors for high gate counts produced by the numerical gates

are critical for noisy, intermediate-scale quantum (NISQ) applications. These improvements are significant for the realistic 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, equivalently, we do not incorporate feedback from the experiment in optimization. However, the device typically deviates from the Hamiltonian 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 Hamiltonian 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 (9) by $f_q \leftarrow f_q + \delta f_q$. We consider three robust control techniques to accomplish this task: the sampling method, the unscented sampling method, and the derivative method.

The sampling method incentivizes the optimizer to ensure multiple copies of a state, each of which evolves with a distinct value of the uncertain parameter, achieve the same state transfer. For each initial state, we add two sample states $|\psi^\pm\rangle$ to the augmented state (13), and the discrete dynamics function (14b) is modified so the sample states evolve under the fluxonium Hamiltonian (9) with $f_q \leftarrow f_q \pm \sigma_{f_q}$ for a fixed hyperparameter σ_{f_q} which is the standard deviation of the qubit frequency. 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 $a_k = 0 \forall k$ will be robust to qubit frequency detunings for the initial states $|0\rangle$ or $|1\rangle$ because the fidelity metric is insensitive to global phases, but this gate will not be robust for any other initial states. Thus, we choose four initial states so that their outer products span the operators on the Hilbert space $\{|0\rangle, |1\rangle, (|0\rangle + i|1\rangle)/\sqrt{2}, (|0\rangle - |1\rangle)/\sqrt{2}\}$ [99]. The infidelities between the sample states and the images of their initial states under the desired gate are penalized by adding a cost function $\sum_{k,\pm} q_k (1 - |\langle \psi_k^\pm | U \psi_1^\pm \rangle|^2)$ to the objective (14a) for a constant q_k we supply.

For the unscented sampling method, we append $2(2n + d)$ sample states to the augmented state vector (13) for each initial state used in the sampling method. $2n = 4$ is twice the dimension of the Hilbert space, resulting from the isomorphism (12), and $d = 1$ is the number of uncertain parameters. The sample states encode a unimodal distribution over the $2n$ elements of the evolving initial state, modeling the uncertainty in the state due to the uncertainty in the parameter. The discrete dynamics function (14b) is modified so each sample state evolves under the fluxonium Hamiltonian (9) where $f_q \leftarrow f_q + \delta f_q$ and

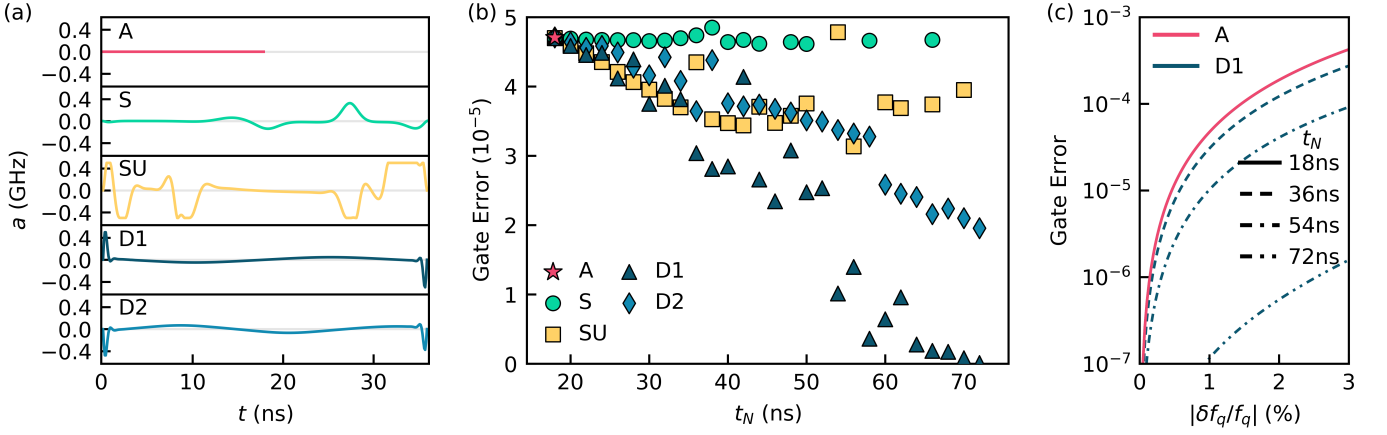


Figure 2: (a) $Z/2$ gates robust to qubit frequency detunings constructed with the analytic (A), sampling (S), unscented sampling (SU), and the 1st- and 2nd-order derivative methods (D1, D2). The numerical gates shown are optimized for twice the gate time of the analytic gate. (b) Single gate error at a one-percent qubit frequency detuning as a function of the gate time. Missing data points represent solutions with a gate error greater than $5 \cdot 10^{-5}$. (c) Single gate error as a function of the qubit frequency detuning. The gate errors for the analytic and 1st-order derivative methods are shown for gate times which are multiples of $1/4f_q \sim 18$ ns. The performance of the two methods is indistinguishable at the gate time 18ns.

δf_q is sampled at each knot point from a distribution with fixed standard deviation σ_{f_q} . After propagating the sample states to each knot point, we apply the unscented transformation to the states, which preserves the first and second moments of the distribution the states encode. This procedure makes the unscented sampling method more resistant to sample states collapsing towards the mean than in the sampling method. We add a contribution to the objective (14a) for the infidelity of the sample states and their target states as in the sampling method. A detailed procedure for the unscented transformation is given in Appendix B.

The derivative method penalizes the sensitivity of the state to the uncertain parameter, which is encoded in the l^{th} -order state derivative $\partial_{f_q}^l |\psi\rangle$. In the m^{th} -order derivative method, we append all state derivatives of order $1, \dots, m$ to the augmented state vector (13) for each initial state used in the sampling method. We penalize the norms of the state derivatives in the objective (14a) by setting the corresponding elements of the final augmented state x_f to 0. We could obtain the state derivatives at each knot point with backward-mode differentiation. In a naive automatic differentiation scheme, the discrete dynamics function at knot points $1, \dots, k-1$ would be differentiated to obtain the state derivative at knot point k , requiring $O(N^2)$ matrix multiplications. Instead, we employ forward-mode differentiation on the TDSE (1) to obtain coupled, first-order ODEs which require $O(N)$ matrix multiplications to integrate. For example, the dynamics for the 1st-order derivative method

are:

$$i\hbar \frac{d}{dt} |\psi\rangle = H |\psi\rangle \quad (16)$$

$$i\hbar \frac{d}{dt} |\partial_{f_q} \psi\rangle = H |\partial_{f_q} \psi\rangle + (\partial_{f_q} H) |\psi\rangle \quad (17)$$

We integrate the coupled ODEs with exponential integrators, see Appendix C, in the discrete dynamics function (14b). For runtimes and complexity analyses for the three robust control techniques, consult Appendix D.

We compare the gate errors due to a static qubit frequency detuning for $Z/2$ gates obtained with the robust control techniques and the analytic method. To compute the gate error, an initial state is evolved under the fluxonium Hamiltonian (9) two separate times with the transformations $f_q \rightarrow f_q \pm \delta f_q$ at the stated qubit frequency detuning. The reported gate error is the infidelity of the evolved state and the target state averaged over the two transformations and 1000 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 $a_k = 0 \forall k$, see Figure 2a. 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 $\pi \delta f_q / 2f_q$ is linear in the qubit frequency detuning, resulting in a gate error that is quadratic in the detuning. At a one-percent qubit frequency detuning ($|\delta f_q|/f_q = 1\%$), the gate error is $\sim 4.5 \cdot 10^{-5}$, which is sufficient for quantum error correction. Although the gate performs well, the method used to derive it can not be employed to achieve a $Z/2$ gate at gate times other than $1/4f_q$. The ability to perform phase gates in any given time is critical for

multi-qubit experiments, where the qubits operate at different frequencies $f_{q,i} \neq f_{q,j}$. We can find solutions using the numerical methods at all gate times above 18ns, see Figure 2b. These numerical methods offer an effective scheme for synchronizing multi-qubit experiments.

The solution for the unscented sampling method combines idling periods with fast ramps to the maximum amplitude, whereas that for the sampling method does not reach high amplitudes. The gate error at a one-percent qubit frequency detuning for the sampling method does not improve substantially over the range of gate times. Conversely, the gate error at a one-percent detuning for the unscented sampling method decreases linearly in the gate time until half the Larmor period $1/2f_q$ after which it has a consistent gate error $\sim 3.5 \cdot 10^{-5}$.

The derivative methods converge on qualitatively similar solutions that use fast triangle pulses at the boundaries and idle at low amplitudes. The gate errors at a one-percent qubit frequency detuning for both methods decrease super-linearly in their gate times. The 1st-order derivative method consistently achieves lower gate errors than the 2nd-order derivative method, which we believe is related to the greater importance of minimizing the 1st-order state derivative norm than the 2nd-order state derivative norm for this problem, see Appendix C. The gate error at a one-percent detuning for the 1st-order derivative method approaches zero at the Larmor period $1/f_q$, see Figure 2c. This result mimics the ability of composite pulses to mitigate parameter uncertainty errors to arbitrary order with sufficiently many pulses [55]. It is difficult to choose an appropriate composite pulse for the problem studied here due to our Hamiltonian and experimental constraints. We propose comparisons between composite pulses and numerical techniques for future work.

VI. ROBUSTNESS TO TIME-DEPENDENT PARAMETER UNCERTAINTY

An additional source of experimental error arises from time-dependent Hamiltonian parameter uncertainty. For many flux-biased and inductively-coupled superconducting circuit elements, magnetic flux noise is a significant source of coherent errors [TODO: coherent errors?](#). Flux noise modifies the fluxonium Hamiltonian (9) by $a(t) \rightarrow a(t) + \delta a(t)$. The spectral density of flux noise is known to follow a $1/f$ distribution [100–103], so the noise is dominated by low-frequency components. The analytic gate considered here takes advantage of the low-frequency characteristic and treat the noise as quasi-static, performing a generalization of the spin-echo technique to compensate for erroneous drift.

Additionally, 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 [104]. The noise is then scaled by the flux noise amplitude of our device $A_\Phi = 5.21\mu\Phi_0 \implies \sigma_a = 2.5 \cdot 10^{-5}\text{GHz}$. In principle, we could modify the sampling method similarly; however, we choose to sample statically $a(t) \rightarrow a(t) + \sigma_a$ 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 (17).

We compare the gate errors due to $1/f$ flux noise for $X/2$ gates constructed with the robust control techniques and the analytic method. To compute the gate error, an initial state is evolved under the fluxonium Hamiltonian (9) where the optimized flux amplitude is modified $a(t) \rightarrow a(t) + \delta a(t)$. The flux noise $\delta a(t)$ is generated as we described for the unscented sampling method. Each of the 1000 averages we perform uses a different random seed for generating the flux noise. We simulate successive applications of the gate constructed by each method and compute the cumulative gate error after each application, see Figure 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 $\sim 11\mu\text{s}$ are two orders of magnitude less than the gate error produced by the analytic gate. [TODO: motional narrowing.](#) $1/f$ flux noise is a significant source of coherent errors in NISQ applications and these numerical techniques offer effective avenues to mitigate it.

VII. CONCLUSION

In conclusion, we have applied state-of-the-art trajectory optimization techniques to mitigate decoherence and achieve robustness to parameter uncertainty errors on a quantum system. We have proposed a scheme for suppressing depolarization with time-optimal control and the depolarization probability model. The computational cost of this depolarization 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 which achieves super-linear 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. [REM: The derivative and sampling methods scale cubically with the dimension of the Hilbert space, allowing for applications to larger quantum systems. The derivative, sampling, and unscented sampling methods will benefit from interleaving optimization with experimental characterization, while the sampling and unscented sampling methods can also be applied to error channels](#)

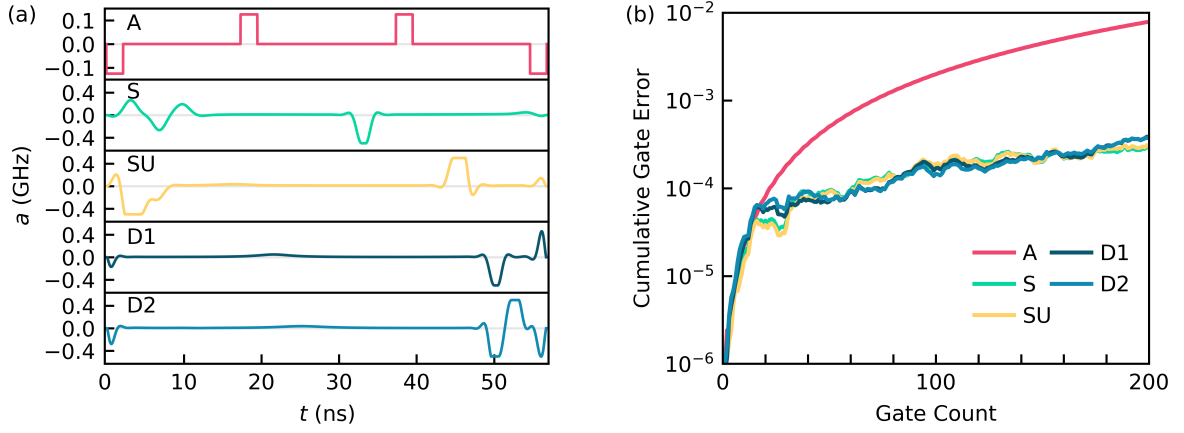


Figure 3: (a) $X/2$ gates robust to flux amplitude offsets constructed with the analytic (A), sampling (S), unscented sampling (SU), and the 1st- and 2nd-order derivative methods (D1, D2). (b) Cumulative gate error due to $1/f$ flux noise for successive gate applications.

without models. **ADD:** The robust control techniques can be applied to any Hamiltonian, allowing experimentalists in all domains to engineer robust operations on their quantum systems. These methods will be used to achieve the low gate errors required for fault-tolerant quantum computing applications. Our implementation of the techniques described in this work is available at <https://github.com/SchusterLab/rbqoc>.

ACKNOWLEDGMENTS

The authors would like to thank Helin Zhang for experimental assistance and Taylor Howell, Jens Koch, Tanay Roy, Colm Ryan, and Daniel Weiss for useful discussions. **TODO:** cite software packages. **TODO:** funding.

Appendix A: Depolarization

We comment on the depolarization metrics and then give our procedure for integrating the Lindblad master equation. The depolarization probability and the gate error due to depolarization metrics are compared in Table I for the numerical experiment described in Section IV. The relative performance of the analytic and numerical techniques is similar across the two metrics.

To compute the gate error due to depolarization we employ the Lindblad master equation. This equation takes the form:

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

where $\rho = |\psi\rangle\langle\psi|$ is the density matrix, n is the dimension of the Hilbert space, $[\cdot, \cdot]$ is the algebraic commutator, and $\{\cdot, \cdot\}$ is the algebraic anti-commutator. For

Gate	P_{1A} (10^{-5})	P_{1N} (10^{-5})	P_{1A}/P_{1N}	GE_A (10^{-5})	GE_N (10^{-5})	GE_A/GE_N
Z/2	5.745	1.149	5.000	1.776	0.371	4.787
Y/2	5.253	1.157	4.540	1.539	0.370	4.159
X/2	16.251	2.660	6.109	5.347	0.863	6.196

Table I: Single gate depolarization probability (P) ratios and single gate error due to depolarization (GE) ratios. Values are reported for the analytic (A) and numerical (N) techniques.

depolarization $\gamma_\uparrow = T_{1,\uparrow}^{-1}$, $\gamma_\downarrow = T_{1,\downarrow}^{-1}$, $L_\uparrow = \sigma^+/2$, and $L_\downarrow = \sigma^-/2$ where $\sigma^\pm = \sigma_x \pm i\sigma_y$. Both $T_{1,\uparrow}$ and $T_{1,\downarrow}$ are obtained from the spline shown in Figure 1b. We obtain the T_1 values in this spline by driving the qubit at the desired flux amplitude and monitoring the resultant decay. For more details consult [71].

Exponential integrators can be employed to integrate the Lindblad master equation using the Vectorization/Choi-Jamiolkowski isomorphism [105],

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

$$\begin{aligned} \hat{\mathcal{L}} = & -i(I \otimes H - H^T \otimes I) \\ & + \sum_{i=1}^{n^2-1} \gamma_i (L_i^* \otimes L_i - \frac{1}{2}(I \otimes L_i^\dagger L_i - L_i^T L_i^* \otimes I)) \end{aligned} \quad (\text{A3})$$

where $\rho = \sum_{i,j} \alpha_{i,j} |i\rangle\langle j|$ and $\text{vec}(\rho) = \sum_{i,j} \alpha_{i,j} |i\rangle \otimes |j\rangle$. We use zero-order hold on the controls so the exact solution is $\text{vec}(\rho_{k+1}) = \exp(\Delta t_k \hat{\mathcal{L}}_k) \text{vec}(\rho_k)$. This isomorphism transforms $(n^2 \times n^2) \times (n^2 \times n^2)$ matrix-matrix multiplications to $(n^4 \times n^4) \times n^4$ matrix-vector multiplications. For small n and zero-order hold on the controls,

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 integration time step to maintain accuracy, resulting in more knot points.

Appendix B: Unscented Transformation

In this section, we outline the full unscented sampling procedure. We consider a state $\psi \in \mathbb{R}^{2n}$ with a deviant parameter $\lambda \in \mathbb{R}^d$ and discrete dynamics $\psi_{k+1} = f(\psi_k, \lambda_k)$. The nominal initial state is given by ψ_0 with an associated positive-definite covariance matrix $P_0 \in \mathbb{R}_{++}^{2n \times 2n}$ which describes the uncertainty in the initial state. P_0 is typically non-zero even if the state preparation error is negligible. The deviant parameter has zero-mean and its distribution is given by the covariance matrix $L_k \in \mathbb{R}_{++}^{d \times d}$ at knot point k . The zero mean assumption is convenient for deriving the update procedure. A non-zero mean can be encoded in the discrete dynamics.

The initial $4n + 2d$ sigma points and initial $4n + 2d$ deviant parameters are sampled from the initial distributions,

$$\begin{bmatrix} \Psi_0^i \\ \Lambda_0^i \end{bmatrix} = \begin{bmatrix} \bar{\Psi}_0 \\ 0 \end{bmatrix} \pm \beta \sqrt{\begin{bmatrix} P_0 & 0 \\ 0 & L_0 \end{bmatrix}}^i \quad (\text{B1})$$

We have written $\bar{\Psi}_0 = \psi_0$. β is a hyperparameter which controls the spacing of the covariance contour. The (\pm) is understood to take $(+)$ for $i \in \{1, \dots, 2n + d\}$ and $(-)$ for $i \in \{2n + d + 1, \dots, 4n + 2d\}$. 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 i^{th} column (mod $2n + d$) of the lower triangular Cholesky factor. Then, the sigma points are normalized,

$$\Psi_0^i \leftarrow \frac{\Psi_0^i}{\sqrt{\Psi_0^{iT} \Psi_0^i}} \quad (\text{B2})$$

The sigma points are propagated to the next knot point,

$$\Psi_1 = f(\Psi_0, \Lambda_0) \quad (\text{B3})$$

The mean and covariance of the sigma points are computed,

$$\bar{\Psi}_1 = \frac{1}{4n + 2d} \sum_{i=1}^{4n+2d} \Psi_1^i \quad (\text{B4})$$

$$P_1 = \frac{1}{2\beta^2} \sum_{i=1}^{4n+2d} (\Psi_1^i - \bar{\Psi}_1)^T (\Psi_1^i - \bar{\Psi}_1) \quad (\text{B5})$$

The sigma points are then resampled and propagated to the next knot point using (B1), (B2), and (B3). Our choice of sigma points follows the prescription in equation 11 of [68]. Prescriptions that require fewer sigma points exist [106].

Appendix C: Derivative Method

We comment on the optimization metrics of the derivative methods and then outline how to efficiently integrate their dynamics. The 1st-order derivative method tends to outperform the 2nd-order derivative method in the small, static detuning regime, see Figure 2b. The norms of the state derivatives for the 1st- and 2nd-order methods are provided in Table II for the solutions to the numerical experiment detailed in Section V at $t_N = 60\text{ns}$. The 2nd-order method is able to decrease the 2nd-order state derivative norm relative to the 1st-order method at the expense of increasing its 1st-order derivative norm. In the problem we have studied it is likely the 2nd-order derivative norm has a smaller contribution to the gate error than the 1st-order derivative norm. A careful analysis of contributions at each order could be used to predict the efficacy of the derivative method for future problems.

Method	$\ \partial_{f_q} \psi_N\ _2 (10^3)$	$\ \partial_{f_q}^2 \psi_N\ _2 (10^6)$
D1	0.436	57.817
D2	1.702	9.030

Table II: Norm of state derivatives with respect to the qubit frequency for $Z/2$ gates optimized using the derivative methods. The norms are computed at the end of the gate $t_N = 60\text{ns}$ and are averaged over the four state derivatives.

The dynamics for the derivative methods can be integrated efficiently using exponential integrators. General exponential integrators break the dynamics into a linear term and a non-linear term. For example, the dynamics of the first state derivative in units of $i\hbar = 1$ are $\frac{d}{dt} |\partial_\lambda \psi\rangle = H |\partial_\lambda \psi\rangle + (\partial_\lambda H) |\psi\rangle$. The linear term is $L = H$ and the non-linear term is $N = (\partial_\lambda H) |\psi\rangle$. With zero-order hold on the controls the exact solution is:

$$\begin{aligned} |\partial_\lambda \psi_{k+1}\rangle &= \exp(\Delta t_k L_k) |\partial_\lambda \psi_k\rangle \\ &+ \int_0^{\Delta t_k} \exp((\Delta t_k - t') L_k) N(t_k + t') dt' \end{aligned} \quad (\text{C1})$$

General exponential integrators proceed by breaking the integral in (C1) into a discrete sum, similar to the procedure for Runge-Kutta schemes. We use a simple approximation known as the Lawson-Euler method [83],

$$\begin{aligned} |\partial_\lambda \psi_{k+1}\rangle &\approx \exp(\Delta t_k L_k) |\partial_\lambda \psi_k\rangle \\ &+ \exp(\Delta t_k L_k) N_k \end{aligned} \quad (\text{C2})$$

This method provides a good tradeoff between accuracy and efficiency, requiring one unique matrix exponential computation per stage. Integration accuracy is not of the utmost importance because the state derivatives guide the optimization, and do not correspond to experimental parameters which must be realized with high accuracy.

Appendix D: Computational Performance

We provide runtimes for our optimizations and comment on the scaling of the robustness methods. The runtimes for the base optimization in Section III, the longitudinally aware optimization in Section IV, and the robust optimizations in Section V are presented in Table III for a $Z/2$ gate at multiples of the analytic $Z/2$ gate time. Compiler optimizations for statically-sized arrays were utilized for all methods except for the unscented sampling method. The unscented sampling method's augmented state vector size was too large to take advantage of these optimizations, which adversely affected its run time. The runtimes of the robust methods scale super-linearly with the gate time. For large gate times, we expect the runtimes to scale linearly with the gate time because the number of knot points scales linearly with the gate time. We observe this trend for the base method, which has a smaller augmented state vector size than the robust methods. We performed optimizations on a single CPU thread in this work as a proof of concept. Future work will parallelize the robustness methods across the initial

states in the operator basis and utilize GPUs [14], which will enable fast optimizations on large Hilbert spaces.

t_N (ns)	Wall Time (± 0.001 s)					
	Base	Long.	S	SU	D1	D2
18	0.155	1.688	1.773	210.573	16.713	48.398
36	7.014	-	48.213	4566.236	67.838	81.030
72	15.906	-	281.372	16575.308	266.997	332.182

Table III: **TODO: check accuracy** Runtimes for $Z/2$ optimizations. Programs were executed on a single thread of an AMD Ryzen Threadripper 3970X 32-Core Processor.

Now we present the problem size complexities for the robustness methods. For the sampling method, the size of the augmented state vector is $O(dn^3)$, where d is the number of deviant parameters and n is the dimension of the Hilbert space. There are n^2 initial states in the operator basis, $2d$ sample states per initial state, and each state has $2n$ real numbers. For the unscented sampling method, the size of the augmented state vector is $O(dn^3 + n^4)$. There are n^2 initial states in the operator basis, $2(2n + d)$ sample states per initial state, and each state has $2n$ real numbers. For the m^{th} -order derivative method, the size of the augmented state vector is $O(dmn^3)$. There are n^2 initial states in the operator basis, dm state derivatives per initial state, and each state has $2n$ real numbers.

-
- [1] L. M. K. Vandersypen and I. L. Chuang, Nmr techniques for quantum control and computation, *Rev. Mod. Phys.* **76**, 1037 (2005).
 - [2] C. T. Kehlet, A. C. Sivertsen, M. Bjerring, T. O. Reiss, N. Khaneja, S. J. Glaser, and N. C. Nielsen, Improving solid-state nmr dipolar recoupling by optimal control, *Journal of the American Chemical Society* **126**, 10202 (2004), pMID: 15315406, <https://doi.org/10.1021/ja048786e>.
 - [3] N. Khaneja, T. Reiss, C. Kehlet, T. Schulte-Herbrüggen, and S. J. Glaser, Optimal control of coupled spin dynamics: design of nmr pulse sequences by gradient ascent algorithms, *Journal of magnetic resonance* **172**, 296 (2005).
 - [4] I. I. Maximov, Z. ToÅpner, and N. C. Nielsen, Optimal control design of nmr and dynamic nuclear polarization experiments using monotonically convergent algorithms, *The Journal of Chemical Physics* **128**, 184505 (2008), <https://doi.org/10.1063/1.2903458>.
 - [5] N. C. Nielsen, C. Kehlet, S. J. Glaser, and N. Khaneja, Optimal control methods in nmr spectroscopy, in *eMagRes* (American Cancer Society, 2010) <https://onlinelibrary.wiley.com/doi/pdf/10.1002/9780470034590.ch51>.
 - [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, *Journal of Magnetic Resonance* **163**, 8 (2003).
 - [7] Z. ToÅpner, T. Vosegaard, C. Kehlet, N. Khaneja, S. J. Glaser, and N. C. Nielsen, Optimal control in nmr spectroscopy: Numerical implementation in simpson, *Journal of Magnetic Resonance* **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. 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 preprint arXiv:2010.15292* (2020).
 - [10] 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).
 - [11] P. Gokhale, Y. Ding, T. Propson, C. Winkler, N. Leung, Y. Shi, D. I. Schuster, H. Hoffmann, and F. T. Chong, Partial compilation of variational algorithms for noisy intermediate-scale quantum machines, in *Proceedings of the 52nd Annual IEEE/ACM International Symposium on Microarchitecture* (2019) pp. 266–278.
 - [12] 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, arXiv preprint arXiv:2004.12458 (2020).
- [13] Z. Leng, P. Mundada, S. Ghadimi, and A. Houck, Robust and efficient algorithms for high-dimensional black-box quantum optimization, arXiv preprint arXiv:1910.03591 (2019).
- [14] N. Leung, M. Abdelhafez, J. Koch, and D. Schuster, Speedup for quantum optimal control from automatic differentiation based on graphics processing units, *Physical Review A* **95**, 042318 (2017).
- [15] S. Li, T. Chen, and Z.-Y. Xue, Fast holonomic quantum computation on superconducting circuits with optimal control, *Advanced Quantum Technologies* **3**, 2000001 (2020), <https://onlinelibrary.wiley.com/doi/pdf/10.1002/qute.202000001>.
- [16] J. Xu, S. Li, T. Chen, and Z.-Y. Xue, Nonadiabatic geometric quantum computation with optimal control on superconducting circuits, arXiv preprint arXiv:2004.10199 (2020).
- [17] 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, *Physical Review A* **92**, 062110 (2015).
- [18] G. De Chiara, T. Calarco, M. Anderlini, S. Montangero, P. Lee, B. Brown, W. Phillips, and J. Porto, Optimal control of atom transport for quantum gates in optical lattices, *Physical Review A* **77**, 052333 (2008).
- [19] M. H. Goerz, T. Calarco, and C. P. Koch, The quantum speed limit of optimal controlled phasegates for trapped neutral atoms, *Journal of Physics B: Atomic, Molecular and Optical Physics* **44**, 154011 (2011).
- [20] J. Guo, X. Feng, P. Yang, Z. Yu, L. Chen, C.-H. Yuan, and W. Zhang, High-performance raman quantum memory with optimal control in room temperature atoms, *Nature communications* **10**, 1 (2019).
- [21] J. H. M. Jensen, J. J. Sørensen, K. Mølmer, and J. F. Sherson, Time-optimal control of collisional swap gates in ultracold atomic systems, *Physical Review A* **100**, 052314 (2019).
- [22] 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, *Physical Review X* **10**, 021058 (2020).
- [23] A. Omran, H. Levine, A. Keesling, G. Semeghini, T. T. Wang, S. Ebadi, H. Bernien, A. S. Zibrov, H. Pichler, S. Choi, *et al.*, Generation and manipulation of schrödinger cat states in rydberg atom arrays, *Science* **365**, 570 (2019).
- [24] 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, *Physical Review A* **88**, 021601 (2013).
- [25] J. J. Sørensen, J. Jensen, T. Heinzl, and J. F. Sherson, Qengine: A c++ library for quantum optimal control of ultracold atoms, *Computer Physics Communications* **243**, 135 (2019).
- [26] 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, *Physical Review A* **74**, 022312 (2006).
- [27] 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, *Scientific reports* **6**, 34187 (2016).
- [28] 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).
- [29] 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, *Nature communications* **5**, 1 (2014).
- [30] 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).
- [31] 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).
- [32] F. Poggiali, P. Cappellaro, and N. Fabbri, Optimal control for one-qubit quantum sensing, *Phys. Rev. X* **8**, 021059 (2018).
- [33] 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, arXiv preprint arXiv:2004.12119 (2020).
- [34] 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).
- [35] 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, *Scientific reports* **9**, 1 (2019).
- [36] J. Sørensen, M. Aranburu, T. Heinzl, and J. Sherson, Quantum optimal control in a chopped basis: Applications in control of bose-einstein condensates, *Physical Review A* **98**, 022119 (2018).
- [37] K. Zhou, *Essentials of Robust Control*, 1st ed. (Pearson, 1997).
- [38] J. Morimoto and C. Atkeson, Minimax differential dynamic programming: An application to robust biped walking, *Advances in neural information processing systems* **15**, 1563 (2002).
- [39] Z. Manchester and S. Kuindersma, Robust Direct Trajectory Optimization Using Approximate Invariant Funnel, *Autonomous Robots* 10.1007/s10514-018-9779-5 (2018).
- [40] D. J. Egger and F. K. Wilhelm, Optimized controlled-z gates for two superconducting qubits coupled through a resonator, *Superconductor Science and Technology* **27**, 014001 (2013).
- [41] D. J. Egger and F. K. Wilhelm, Adaptive hybrid optimal quantum control for imprecisely characterized systems, *Phys. Rev. Lett.* **112**, 240503 (2014).
- [42] 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, *Journal of Physics B: Atomic, Molecular and Optical Physics* **40**, S103 (2007).
- [43] R. W. Heeres, P. Reinhold, N. Ofek, L. Frunzio, L. Jiang, M. H. Devoret, and R. J. Schoelkopf, Implementing a universal gate set on a logical qubit encoded in an oscillator, *Nature communications* **8**, 1 (2017).
- [44] S.-Y. Huang and H.-S. Goan, Optimal control for fast

- and high-fidelity quantum gates in coupled superconducting flux qubits, *Physical Review A* **90**, 012318 (2014).
- [45] J. Kelly, R. Barends, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, A. G. Fowler, I.-C. Hoi, E. Jeffrey, A. Megrant, J. Mutus, C. Neill, P. J. J. O'Malley, C. Quintana, P. Roushan, D. Sank, A. Vainsencher, J. Wenner, T. C. White, A. N. Cleland, and J. M. Martinis, Optimal quantum control using randomized benchmarking, *Phys. Rev. Lett.* **112**, 240504 (2014).
 - [46] P. J. Lieberman and F. K. Wilhelm, Optimal qubit control using single-flux quantum pulses, *Phys. Rev. Applied* **6**, 024022 (2016).
 - [47] V. Nebendahl, H. Häffner, and C. F. Roos, Optimal control of entangling operations for trapped-ion quantum computing, *Phys. Rev. A* **79**, 012312 (2009).
 - [48] P. Rebentrost and F. K. Wilhelm, Optimal control of a leaking qubit, *Phys. Rev. B* **79**, 060507 (2009).
 - [49] 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).
 - [50] R. J. Spiteri, M. Schmidt, J. Ghosh, E. Zahedinejad, and B. C. Sanders, Quantum control for high-fidelity multi-qubit gates, *New Journal of Physics* **20**, 113009 (2018).
 - [51] A. Spörl, T. Schulte-Herbrüggen, S. J. Glaser, V. Bergholm, M. J. Storz, J. Ferber, and F. K. Wilhelm, Optimal control of coupled josephson qubits, *Phys. Rev. A* **75**, 012302 (2007).
 - [52] H. K. Cummins and J. A. Jones, Use of composite rotations to correct systematic errors in NMR quantum computation, *New Journal of Physics* **2**, 6 (2000).
 - [53] 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).
 - [54] Ä. Kupce and R. Freeman, Stretched adiabatic pulses for broadband spin inversion, *Journal of Magnetic Resonance, Series A* **117**, 246 (1995).
 - [55] J. T. Merrill and K. R. Brown, Progress in compensating pulse sequences for quantum computation, *Quantum Information and Computation for Chemistry*, 241 (2014).
 - [56] 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 preprint arXiv:2004.10364* (2020).
 - [57] 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).
 - [58] 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 preprint arXiv:2007.13756* (2020).
 - [59] 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, *Physical Review A* **98**, 052341 (2018).
 - [60] 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 (2020), *arXiv:2009.09866* [quant-ph].
 - [61] J. Allen, *Robust Optimal Control of the Cross-Resonance Gate in Superconducting Qubits*, Ph.D. thesis, University of Surrey (2019).
 - [62] 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 preprint arXiv:2010.08057* (2020).
 - [63] P. Reinhold, *Controlling Error-Correctable Bosonic Qubits*, Ph.D. thesis, Yale University (2019).
 - [64] R. L. Kosut, M. D. Grace, and C. Brif, Robust control of quantum gates via sequential convex programming, *Phys. Rev. A* **88**, 052326 (2013).
 - [65] T. Schulte-Herbrüggen, A. Spörl, N. Khaneja, and S. J. Glaser, Optimal control for generating quantum gates in open dissipative systems, *Journal of Physics B: Atomic, Molecular and Optical Physics* **44**, 154013 (2011).
 - [66] M. Abdelhafez, D. I. Schuster, and J. Koch, Gradient-based optimal control of open quantum systems using quantum trajectories and automatic differentiation, *Physical Review A* **99**, 052327 (2019).
 - [67] T. A. Howell, C. Fu, and Z. Manchester, Direct policy optimization using deterministic sampling and collocation, *arXiv preprint arXiv:2010.08506* (2020).
 - [68] S. J. Julier and J. K. Uhlmann, Unscented filtering and nonlinear estimation, *Proceedings of the IEEE* **92**, 401 (2004).
 - [69] A. Lee, Y. Duan, S. Patil, J. Schulman, Z. McCarthy, J. Van Den Berg, K. Goldberg, and P. Abbeel, Sigma hulls for gaussian belief space planning for imprecise articulated robots amid obstacles, in *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems* (IEEE, 2013) pp. 5660–5667.
 - [70] Z. Manchester and S. Kuindersma, Derivative-free trajectory optimization with unscented dynamic programming, in *2016 IEEE 55th Conference on Decision and Control (CDC)* (IEEE, 2016) pp. 3642–3647.
 - [71] H. Zhang, S. Chakram, T. Roy, N. Earnest, Y. Lu, Z. Huang, D. Weiss, J. Koch, and D. I. Schuster, Universal fast flux control of a coherent, low-frequency qubit, *arXiv preprint arXiv:2002.10653* (2020).
 - [72] T. A. Howell, B. E. Jackson, and Z. Manchester, Altro: A fast solver for constrained trajectory optimization, in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (IEEE, 2019) pp. 7674–7679.
 - [73] W. Li and E. Todorov, Iterative Linear Quadratic Regulator Design for Nonlinear Biological Movement Systems, in *Proceedings of the 1st International Conference on Informatics in Control, Automation and Robotics* (Setubal, Portugal, 2004).
 - [74] G. Lantoiné and R. P. Russell, A hybrid differential dynamic programming algorithm for constrained optimal control problems. part 1: Theory, *Journal of Optimization Theory and Applications* **154**, 382 (2012).
 - [75] B. Plancher, Z. Manchester, and S. Kuindersma, Constrained unscented dynamic programming, in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (IEEE, 2017) pp. 5674–5680.
 - [76] 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).
- [77] 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 physics* **7** (2019).
- [78] J. Schulman, J. Ho, A. X. Lee, I. Awwal, H. Bradlow, and P. Abbeel, Finding locally optimal, collision-free trajectories with sequential convex optimization., in *Robotics: science and systems*, Vol. 9 (Citeseer, 2013) pp. 1–10.
- [79] R. Tedrake and the Drake Development Team, Drake: A planning, control, and analysis toolbox for nonlinear dynamical systems (2016).
- [80] A. Hereid and A. D. Ames, Frost: Fast robot optimization and simulation toolkit, in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (IEEE/RSJ, Vancouver, BC, Canada, 2017).
- [81] L. Jørgensen, D. L. Cardozo, and E. Thibierge, *Numerical Resolution Of The Schrödinger Equation*, Tech. Rep. (École Normale Supérieure de Lyon, 2011).
- [82] N. Auer, L. Einkemmer, P. Kandolf, and A. Ostermann, Magnus integrators on multicore cpus and gpus, *Computer Physics Communications* **228**, 115 (2018).
- [83] H. Berland and B. Skaflestad, *Solving the nonlinear Schrödinger equation using exponential integrators*, Tech. Rep. (Norwegian University of Science and Technology, 2005).
- [84] L. Einkemmer, M. Tokman, and J. Loffeld, On the performance of exponential integrators for problems in magnetohydrodynamics, *Journal of Computational Physics* **330**, 550 (2017).
- [85] C. R. Hargraves and S. W. Paris, Direct Trajectory Optimization Using Nonlinear Programming and Collocation, *J. Guidance* **10**, 338 (1987).
- [86] D. Q. Mayne, A second-order gradient method of optimizing non-linear discrete time systems, *Int J Control* **3**, 8595 (1966).
- [87] D. P. Bertsekas, Projected newton methods for optimization problems with simple constraints, *SIAM Journal on control and Optimization* **20**, 221 (1982).
- [88] C. V. Rao, S. J. Wright, and J. B. Rawlings, Application of interior-point methods to model predictive control, *Journal of Optimization Theory and Applications* **99**, 723 (1998).
- [89] K. L. Clarkson, Coresets, sparse greedy approximation, and the frank-wolfe algorithm, *ACM Transactions on Algorithms (TALG)* **6**, 1 (2010).
- [90] A. Hauswirth, S. Bolognani, G. Hug, and F. Dörfler, Projected gradient descent on riemannian manifolds with applications to online power system optimization, in *2016 54th Annual Allerton Conference on Communication, Control, and Computing (Allerton)* (2016) pp. 225–232.
- [91] B. E. Jackson, T. Punnoose, D. Neamati, K. Tracy, R. Jitsho, and Z. Manchester, ALTRO-C: A fast solver for conic model-predictive control, in *International Conference on Robotics and Automation ICRA* (Virtual, 2021) in Review.
- [92] D. Bertsekas, *Constrained optimization and lagrange multiplier methods* (second edn.) athena scientific: Belmont (1996).
- [93] 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).
- [94] P. Krantz, M. Kjaergaard, F. Yan, T. P. Orlando, S. Gustavsson, and W. D. Oliver, A quantum engineer's guide to superconducting qubits, *Applied Physics Reviews* **6**, 021318 (2019).
- [95] P. Klimov, J. Kelly, Z. Chen, M. Neeley, A. Megrant, B. Burkett, R. Barends, K. Arya, B. Chiaro, Y. Chen, *et al.*, Fluctuations of energy-relaxation times in superconducting qubits, *Physical review letters* **121**, 090502 (2018).
- [96] D. Aharonov and M. Ben-Or, Fault-tolerant quantum computation with constant error rate, *SIAM Journal on Computing* **38**, 1207 (2008), <https://doi.org/10.1137/S0097539799359385>.
- [97] A. G. Fowler, A. M. Stephens, and P. Groszkowski, High-threshold universal quantum computation on the surface code, *Phys. Rev. A* **80**, 052312 (2009).
- [98] D. Gottesman, Stabilizer codes and quantum error correction, *arXiv preprint quant-ph/9705052* (1997).
- [99] J. Chow, J. M. Gambetta, L. Tornberg, J. Koch, L. S. Bishop, A. A. Houck, B. Johnson, L. Frunzio, S. M. Girvin, and R. J. Schoelkopf, Randomized benchmarking and process tomography for gate errors in a solid-state qubit, *Physical review letters* **102**, 090502 (2009).
- [100] 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).
- [101] 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).
- [102] 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).
- [103] 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).
- [104] J. O. Smith, *Spectral Audio Signal Processing* (<http://ccrma.stanford.edu/jos/sasp/>, 2020) online book, 2011 edition.
- [105] G. T. Landi, *Lecture notes on quantum information and quantum noise* (2018).
- [106] S. J. Julier and J. K. Uhlmann, Reduced sigma point filters for the propagation of means and covariances through nonlinear transformations, in *Proceedings of the 2002 American Control Conference (IEEE Cat. No. CH37301)*, Vol. 2 (IEEE, 2002) pp. 887–892.