Robust Control of a Fluxonium Qubit

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: December 10, 2020)

The ability to engineer high-fidelity gates on quantum processors in the presence of systematic errors and decoherence remains the primary challenge requisite to achieving quantum advantage. Quantum optimal control techniques 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 deviations and noise. We propose a method that takes advantage of deviant parameter derivative information while maintaining computational efficiency by utilizing mixed-mode differentiation. Additionally, completely modeling decoherence effects due to longitudinal relaxation requires integrating the Lindblad master equation, which is computationally expensive. We propose a computationally efficient metric and utilize time-optimal control to achieve high fidelity gates in the presence of longitudinal relaxation. We demonstrate these techniques numerically on a fluxonium qubit with realistic experimental parameters and constraints, achieving orders of magnitude gate error reductions from our baseline gate set.

I. INTRODUCTION

Quantum optimal control (QOC) techniques are 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-Einstien condensates [35, 36]. For quantum computation, optimal control techniques are 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 techniques improve upon standard optimal control techniques by encoding model uncertainties in optimization objectives, yielding performance guarantees over a range of parameters [37–39]. We adapt robust control techniques from the robotics community to mitigate parameter deviation errors for a superconducting fluxonium qubit.

QOC techniques have had tremendous success in engineering high fidelity quantum gates [13, 28, 29, 40–51]. While standard QOC techniques can predict system behavior with high accuracy, they are sensitive to experimental errors such as parameter drift, noise, finite control resolution, and decoherence. Multiple techniques have been developed to address these shortcomings. Analytic techniques to mitigate parameter deviation errors include composite pulses [52–55], dynamic and geometric phase considerations [16, 56], and the DRAG scheme [57]. To mitigate decoherence, Floquet techniques have been employed [12, 58]. Numerical techniques to mitigate parameter deviation errors include closed-loop methods [12, 41, 59, 60] and open-loop methods [33, 61–64]. Numerical techniques to mitigate decoherence include modeling master equations [33] and employing Monte Carlo style quantum trajectories [65].

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

1. A sampling method, similar to the work of [33, 61–63].
2. An unscented sampling method adapted from the unscented transform used in the state estimation community [66–69].
3. A derivative-penalization method, which uses efficient mixed-mode differentiation to compute the sensitivity of the quantum state trajectory to deviant parameters.

We apply these techniques to the fluxonium qubit presented in [70]. We also show that QOC can solve important problems for fluxonium-based qubits: mitigating longitudinal relaxation type decoherence by taking advantage of the T_1-dependence of the controls and performing phase gates in arbitrary times. To mitigate errors due to longitudinal relaxation, we perform time-optimal control and utilize an efficient optimization objective that does not pay the increased computational cost of integrating a master equation. Leveraging recent advances in trajectory optimization within the field of robotics, we solve these optimization problems using the ALTRO solver, which uses iterative LQR (iLQR)—a differential-dynamic programming (DDP)-based indirect method similar to shooting methods such as GOAT [71], GRAPE [3, 14], and Krotov's [72]—within an augmented

* tcpropson@uchicago.edu

Lagrangian framework to handle nonlinear equality and inequality constraints at each time step [73].

This paper is organized as follows. First, we introduce the ALTRO method in the context of QOC in Section II. We describe realistic constraints for the fluxonium and map them to the ALTRO method in Section III. Then, we outline a method for making the optimization aware of longitudinal relaxation in Section IV. Next, we outline three methods for achieving robustness to static parameter deviations in Section V. Finally, we employ the robust control techniques to mitigate $1/f$ flux noise in Section VI.

II. QOC + ALTRO

In this section we review the QOC problem statement and introduce the ALTRO solver. QOC concerns the evolution of a (quantum) state $|\psi(t)\rangle$ governed by the time-dependent Schrödinger equation (TDSE),

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

The Hamiltonian has an arbitrary dependence on the possibly multi-valued controls $u(t)$. The controls are so-called because they are the means the experimentalist has to act on the system. To make the problem numerically tractable, the controls and quantum state are discretized into N knot points (time steps). The optimization problem is to find the controls that minimize a functional $J(u)$. In the simplest case, the functional is the infidelity between the initial state evolved to the final knot point and the desired final state $J = 1 - |\langle\psi_f|\psi_N(u)\rangle|^2$. In general, J is a linear combination of cost functions on the state as well as cost functions on the controls [14]. Standard QOC solvers employ automatic differentiation to compute gradients of the functional $\nabla_u J(u)$, which can easily be used to implement first-order optimization methods [3, 14, 71, 72].

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 [73–76]. Trajectory optimization problems are typically of the following form:

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

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

$$g_k(x_k,u_k) \le 0, \tag{2c}$$

$$h_k(x_k,u_k) = 0 \tag{2d}$$

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.

Many techniques have been proposed for solving (2a)–(2d). Standard methods include direction collocation [77] and differential-dynamic programming (DDP) [78]. Recent state-of-the-art solvers, such as ALTRO [73], have combined principles from both of these approaches.

ALTRO uses iterative LQR (iLQR) [79] 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:

$$\underset{z}{\text{minimize}} \quad f(z) + \lambda^T c(z) + \frac{1}{2}c(z)^T I_\mu c(z) \tag{3}$$

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 (2a), $c(z)$ 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 (3) using iLQR, the Lagrange multiplier estimates are updated according to,

$$\lambda \leftarrow \lambda + \mu c(z) \tag{4}$$

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

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 [73, 80].

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 [81, 82]. Conversely, the ALTRO solver delays projecting onto the constraint manifold defined by (2c) and (2d) 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 constraints must be satisfied in addition to minimizing auxiliary metrics.

III. QOC ON THE FLUXONIUM

In the following, we study the QOC problem on the fluxonium qubit. The fluxonium qubit is a promising building block for superconducting circuits, 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:

$$H/h = f_q \frac{\sigma_z}{2} + a(t)\frac{\sigma_x}{2} \tag{5}$$

where $f_q = 14\text{MHz}$ is the qubit frequency at the flux frustration point, $a(t)$ is the flux drive amplitude, h is Planck's constant, and σ_x, σ_y are Pauli matrices. We consider the task of constructing $Z/2$, $Y/2$, and $X/2$ gates for the fluxonium qubit subject to experimental constraints, decoherence, and Hamiltonian parameter deviations. We compare the gates we obtain with numerical methods to the analytically constructed gates reported in [70] for the same device.

The optimization problem takes the form:

$$\begin{aligned}
\underset{x_{1:N},\, u_{1:N\text{-}1}}{\text{minimize}} \quad & \sum_{k=1}^{N} \|x_k\text{-}x_f\|_{Q_k} + \sum_{k=1}^{N-1} \|u_k\|_{R_k} && \text{(6a)}\\
\text{subject to} \quad & x_{k+1} = f(x_k, u_k), && \text{(6b)}\\
& \left|\psi_1^0\right\rangle = |0\rangle\,, \left|\psi_1^1\right\rangle = |1\rangle\,, && \text{(6c)}\\
& \left|\psi_N^i\right\rangle = \left|\psi_f^i\right\rangle, && \text{(6d)}\\
& \left|\left\langle\psi_k^i|\psi_k^i\right\rangle\right|^2 = 1, && \text{(6e)}\\
& \textstyle\int_0^{t_1} a \ \mathrm{d}t = \partial a_1/\partial t = 0, && \text{(6f)}\\
& \textstyle\int_0^{t_N} a \ \mathrm{d}t = 0, && \text{(6g)}\\
& a_1 = a_N = 0, && \text{(6h)}\\
& |a_k| \le 0.5 \text{ GHz} && \text{(6i)}
\end{aligned}$$

The augmented state and augmented controls are:

$$x = \begin{bmatrix} \psi^0 \\ \psi^1 \\ \int_0^t a \ \mathrm{d}t \\ a \\ \partial a/\partial t \end{bmatrix} \quad u = \left[\partial^2 a/\partial t^2\right] \tag{7}$$

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.

We impose constraints that reflect the physical limitations of the apparatus and improve the experimental realization of the control pulse. The discrete dynamics (6b) integrates the states according to the TDSE (1) and the fluxonium Hamiltonian (5) and integrates the moments of the flux amplitude. Exposing lower order moments of the flux amplitude allows us to penalize their norms in (6a), smoothing the flux amplitude and mitigating AWG ringing due to high frequency transitions. We ensure the desired gate is achieved by imposing the constraints (6c) and (6d). The initial states in (6c) are chosen to span the Hilbert space. The final states in (6d) 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],

$$H\,|\psi\rangle \cong \begin{bmatrix} H_{\text{re}} & -H_{\text{im}} \\ H_{\text{im}} & H_{\text{re}} \end{bmatrix} \begin{bmatrix} \psi_{\text{re}} \\ \psi_{\text{im}} \end{bmatrix} \tag{8}$$

We ensure the solver does not take advantage of discretization error by imposing the normalization constraint (6e). The constraint (6f) is an initial condition. We ensure the pulse has zero net flux by imposing (6g), which mitigates the hysteresis ubiquitous in flux bias lines [70, 83, 84]. We ensure gates may be concatenated arbitrarily without inducing AWG ringing due to high-frequency transitions by imposing the boundary condition (6h). Finally, we ensure the two-level approximation (5) remains valid by imposing the amplitude constraint (6i). All of the gates presented in this work achieve a maximum individual constraint violation of 10^{-8} for (6b)–(6i).

IV. LONGITUDINAL RELAXATION AWARENESS

The strength of longitudinal relaxation type decoherence varies with the control parameters in a range of superconducting circuit platforms, including the fluxonium qubit. Because dissipation to the thermal bath via longitudinal relaxation is an irreversible process that results in information loss, it is advantageous to tune the controls to extend the longitudinal relaxation time T_1– which represents the $1/e$ relaxation time of the quantum state. Typical approaches completely model the gate error due to longitudinal relaxation by performing optimal control on n^2 density matrices of size $n \times n$ that evolve under a master equation, rather than n state vectors of

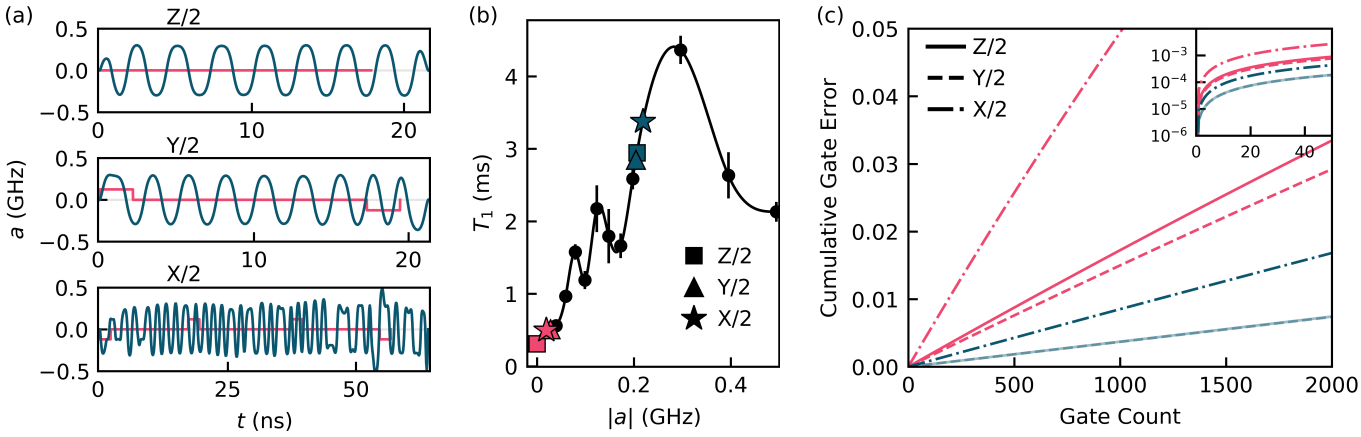

Figure 1: (a) Numerically optimized gates (dark) and analytically optimized gates (light). (b) T_1 interpolation function used in optimization. Markers denote the time-averaged, absolute amplitude of each gate. (c) Lindblad master equation simulation with T_1 dissipation for successive gate applications. The cumulative gate error is computed after each gate application. The numerically optimized $Z/2$ and $Y/2$ gate errors are indistinguishable.

size n under the TDSE (1). We avoid this increase in computational complexity by penalizing the probability (integrated rate) of longitudinal relaxation. Using this probability as a proxy for the gate error incurred is reasonable because losses due to longitudinal relaxation increase monotonically in time. This technique can be extended to error channels which share the monotonically increasing property. Additionally, for a constant T_1 time, a shorter gate time would favor a lower longitudinal relaxation probability. We allow the optimizer to tune the gate time to minimize the longitudinal relaxation probability. Our scheme for time-optimal control is suitable for any time-optimal problem, not only the one we study here.

The longitudinal relaxation probability is given by,

$$P_1(t) = \int_0^t T_1^{-1}(a(t'))dt' \tag{9}$$

This value is appended to the augmented state vector (7) and penalized according to (6a). $T_1(a_k)$ 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 [85]. 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 vector (7) and the squared root of the time step $|\Delta t_k|$ is used for integration in the discrete dynamics function (6b). To ensure numerical integration accuracy is maintained we add a bound constraint on the time step at each knot point.

We compare the numerical method we have developed to the analytic gates on the task of achieving low gate errors in the presence of longitudinal relaxation for the $Z/2$, $Y/2$, and $X/2$ gates. The numerically optimized gates converge on similar solutions, a periodic waveform with amplitude ~ 0.2GHz, see Figure 1a. They extend their gate times beyond their analytic counterparts, trading longer gate times for access to higher amplitudes and therefore higher T_1 times. All numerical gates reduce their single gate errors by a factor of 5 over their analytic counterparts, commensurate to their probability of longitudinal relaxation reductions, see Appendix A. The gate error reported in this text is the infidelity of the evolved state and the final state averaged over 1000 pseudo-randomly generated initial states. The numerical $Z/2$ and $Y/2$ gates perform similarly in the concatenated gate application comparison, suppressing cumulative gate errors to $8 \cdot 10^{-3}$ over 2000 gate applications $\sim 40\mu$s, see Figure 1c. The numerical $X/2$ gate achieves a cumulative gate error of $1.7 \cdot 10^{-2}$ over 2000 gate applications $\sim 124\mu$s. These low cumulative gate errors for high gate counts are critical for noisy, intermediate-scale quantum (NISQ) applications. Both the analytic and numerical gates attain single gate errors sufficient for quantum error correction ($< 10^{-4}$), which are required for fault-tolerant quantum computing [86–88]. 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.

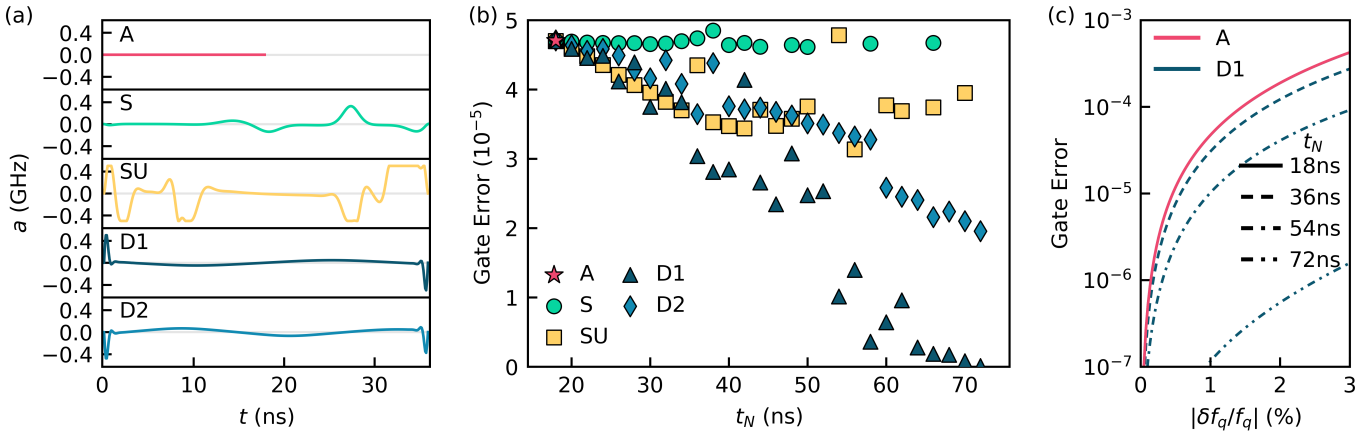

Figure 2: (a) $Z/2$ gates robust to qubit frequency detunings constructed with the analytic, sampling, unscented sampling, and the 1st- and 2nd-order derivative methods. The gates shown for the numerical methods are the solutions at twice the analytic gate time. (b) Single gate error as a function of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

V. ROBUSTNESS TO STATIC PARAMETER DEVIATIONS

We have formulated the QOC problem as an open-loop optimization problem, i.e. 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 deviations. As an example, we mitigate errors arising from the drift and finite measurement precision of the qubit frequency which modifies (5) by $f_q \leftarrow f_q + \delta f_q$. We consider three robust control techniques: the sampling method, the unscented sampling method, and the derivative method.

The sampling method simultaneously optimizes over copies of a state, each of which evolves with a distinct deviant parameter value. We append four sets of two sample states $\psi^{\pm}$ to the augmented state vector (7). We initialize each pair of samples in the same state. The four initial states are chosen so that their outerproducts span the operators on the Hilbert space $\{|0\rangle, |1\rangle, (|0\rangle + i|1\rangle)/\sqrt{2}, (|0\rangle - |1\rangle)/\sqrt{2}\}$ [89]. The discrete dynamics (6b) is modified so that the samples evolve under the Hamiltonian given in (5) with $f_q \leftarrow f_q \pm \sigma_{f_q}$ for a fixed σ_{f_q}. The final state for each sample is the image of its initial state under the desired gate. The infidelity between the samples and their final states are penalized by adding a term $\sum_{k=1}^{N} q_k(1 - |\langle \psi_k^{\pm} | \psi_f \rangle|^2)$ to the objective (6a).

For the unscented sampling method, we append $2(2n+d)$ samples to the augmented state vector (7) for each initial state in the operator basis. Here $2n = 4$ is twice the dimension of the Hilbert space, resulting from the isomorphism (8), and $d = 1$ is the number of deviant parameters. The samples represent a Gaussian distribution over the $2n$ elements of the initial state, modeling the uncertainty in the state as a result of the uncertainty in the deviant parameter. Each sample evolves under the Hamiltonian given in (5) with $f_q \leftarrow f_q + \delta f_q$. We determine the qubit frequency detuning δf_q at each knot point using the statistics of the samples. We modify the discrete dynamics (6b) to evolve the samples under the deviant Hamiltonian and then apply the unscented transformation to the ensemble of samples, accurately preserving the first and second moments of the distribution. We penalize the infidelity between the samples and their final states by adding an infidelity contribution to (6a) as in the sampling method. A detailed procedure for the unscented transformation is given in Appendix B.

The derivative method penalizes the sensitivty of the state to the deviant parameter, which is encoded in the (l^{th}-order) derivative $\partial_{f_q}^l \psi$. In the m^{th}-order derivative method, we append all state derivatives of order $1, \ldots, m$ to the augmented state vector (7) for each initial state in the operator basis. We penalize the norms of the state derivatives in (6a) by setting the corresponding final states to zero. We could obtain the state derivatives in the discrete dynamics (6b) with backwards mode differentiation. Naive automatic differentiation would differentiate all dynamics at knot points $1, \ldots, k-1$ to obtain the state derivative at knot point k, requiring $O(N^2)$ matrix multiplications. Instead, we employ forwards mode differentiation on the TDSE to obtain coupled dynamics which require $O(N)$ matrix multiplications to integrate.

For example, the dynamics for the 1st-order derivative method are:

$$i\hbar\frac{d}{dt}|\psi\rangle = H|\psi\rangle \tag{10}$$

$$i\hbar\frac{d}{dt}\left|\partial_{f_q}\psi\right\rangle = H\left|\partial_{f_q}\psi\right\rangle + (\partial_{f_q}H)|\psi\rangle \tag{11}$$

Exponential integrators that account for the non-linear term in (11) are used to efficiently integrate the coupled dynamics, see Appendix C. For runtimes and complexity analyses for the three robust control methods, consult Appendix D.

We challenge these methods to achieve a $Z/2$ gate subject to a static qubit frequency detuning. We take $\sigma_{f_q}/f_q = 1\%$ to be one standard devation, and equip the sampling methods accordingly. We compute gate errors for each method by evolving the system under the Hamiltonian in (5) with the optimized flux amplitude and $f_q \leftarrow f_q + \delta f_q$ at the stated qubit frequency detuning.

We compare the numerical methods to an analytically derived $Z/2$ gate, see Figure 2a. The analytic gate corresponds to idling at the flux frustration point $a = 0$. It is at the device's speed limit for a $Z/2$ gate $t_{Z/2} = 1/4f_q$ and is simple to derive. Its erroneous rotation angle $2\pi t_{Z/2}\delta f_q$ is linearly sensitive to the qubit frequency detuning, resulting in a gate error that is quadratically sensitive to the qubit frequency detuning. At a one-percent qubit frequency detuning ($\delta f_q/f_q = 1\%$) the analytic gate achieves a gate error $\sim 4.5\cdot 10^{-5}$, which is sufficient for quantum error correction. Although the analytic $Z/2$ gate performs well, it only works at the gate time $t_{Z/2}$. The ability to perform Z rotations in arbitrary times is critical for multi-qubit experiments, where the qubits operate at different frequencies $f_{q,i} \neq f_{q,j}$. Each numerical method can find solutions at all gate times above $t_{Z/2}$, see Figure 2b. These numerical methods offer an effective scheme for synchronizing multi-qubit experiments.

The sampling and unscented sampling methods combine idling periods with time-anti-symmetric ramps. The gate error at $\delta f_q/f_q = 1\%$ for the sampling method does not improve substantially over the range of gate times. The unscented sampling method achieves linear decreases in its gate error at $\delta f_q/f_q = 1\%$ with longer gate times until half the Larmor period $1/2f_q$ after which it achieves 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 balance time on either side of the flux-frustration point symmetrically at low amplitudes. Both methods achieve a super-linear scaling in their gate error as a function of their gate time. We believe the 1st-order method outperforms the 2nd-order method due to the low contribution of second-order terms to the gate error in this deviation regime, see Appendix C. The gate error 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 deviation 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 DEVIATIONS

An additional source of experimental error arises from time-dependent Hamiltonian parameter deviations. For many flux-biased and inductively-coupled superconducting circuit elements, magnetic flux noise is a significant source of coherent errors. Magnetic flux noise modifies the fluxonium Hamiltonian (5) by $a(t) \leftarrow a(t) + \delta a(t)$. The spectral density of $\delta a(t)$ follows a $1/f$ distribution for a range of devices, consisting primarily of low frequency noise [90–93]. Analytic methods to combat flux noise take advantage of the low-frequency characteristic and treat the noise as quasi-static, performing generalizations of the spin-echo technique to compensate for erroneous drift. The analytic gate considered here follows this strategy.

We challenge the analytic method and the robust control methods of the previous section to achieve an $X/2$ gate subject to $1/f$ flux noise. We compute gate errors for each method by evolving the system under the Hamiltonian in (5) where the optimized flux amplitude is modified $a(t) \leftarrow a(t) + \delta a(t)$. The flux noise $\delta a(t)$ is generated by filtering white noise sampled from a standard normal distribution with a finite impulse response filter [94]. It is then scaled by the flux noise amplitude of our device $A_\Phi = 5.21\mu\Phi_0 \implies \delta a(t) \sim \sigma_a = 2.5\cdot 10^{-5}$GHz. The unscented sampling method is modified so that its samples are subject to $1/f$ flux noise by carrying the state of a finite impulse response filter in the augmented state vector (7). In principle, we could modify the sampling method similarly; however, we choose to sample statically at σ_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 (11).

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 methods achieve single gate errors sufficient for quantum error correction. Despite converging on qualitatively different solutions, the numerical methods perform similarly in the concatenated gate application comparison. They achieve a two order of magnitude cumulative gate error reduction over the analytic method after 200 gate applications $\sim 11\mu$s. $1/f$ flux noise is a significiant source of coherent errors in NISQ applications and these numerical techniques offer effective avenues to mitigate it.

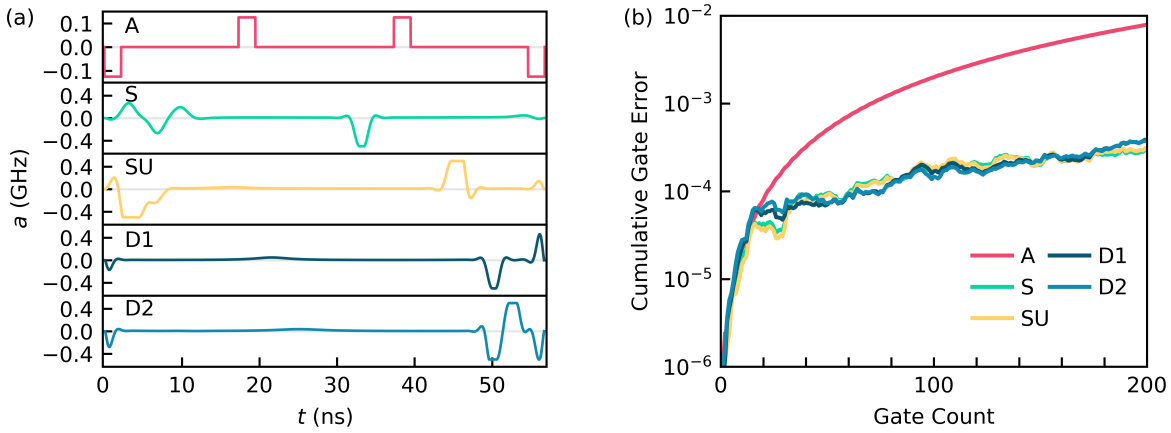

Figure 3: (a) $X/2$ gates robust to flux amplitude offsets constructed with the analytic, sampling, unscented sampling, and the 1st- and 2nd-order derivative methods. The gates shown are the solutions at the analytic gate time. (b) Simulation of stochastic $1/f$ flux noise for successive gate applications. The cumulative gate error is computed after each gate application.

VII. CONCLUSION

In conclusion, we have applied state-of-the-art trajectory optimization techniques to achieve robustness to parameter deviations and mitigate decoherence on a quantum system. We have proposed a scheme for mitigating longitudinal relaxation with time-optimal control and the probability of longitudinal relaxation metric, avoiding the optimal control problem for density matrices which scales quartically with the dimension of the Hilbert space. We have proposed the derivative method for robust control which achieves super-linear gate error reductions in its gate time for the static parameter deviation 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. 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 without models. These techniques 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 Daniel Weiss, Taylor Howell, and Tanay Roy for useful discussions. TODO: cite software packages. TODO: funding.

Appendix A: Longitudinal Relaxation

We comment on the longitudinal relaxation metrics and then give our procedure for integrating the Lindblad master equation. The longitudinal relaxation probability and the gate error due to longitudinal relaxation 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.

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 longitudinal relaxation probability (P) ratios and single gate error due to longitudinal relaxation (GE) ratios. Values are reported for the analytic (A) and numerical (N) techniques.

To compute the gate error due to longitudinal relaxation 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 longitudinal relaxation $\gamma_\uparrow = T_{1,\uparrow}^{-1}$, $\gamma_\downarrow = T_{1,\downarrow}^{-1}$, $L_\uparrow = \sigma^+/2$, 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 ob-

tain 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 [70].

Exponential integrators can be employed to integrate the Lindblad master equation using the Vectorization/Choi-Jamiolkowski isomorphism [95],

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

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

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

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

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 $i \in \{1, \ldots, 2n + d\}$ and $(-)$ for $i \in \{2n + d + 1, \ldots, 4n + 2d\}$. 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,

$$\Psi_0^i \leftarrow \frac{\Psi_0^i}{\sqrt{\Psi_0^{iT}\Psi_0^i}} \tag{B2}$$

The sigma points are propagated to the next knot point,

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

The mean and covariance of the sigma points are computed,

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

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

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 [67]. Prescriptions that require fewer sigma points exist [96].

Appendix C: Derivative Method

We comment on the optimization metrics of the derivative methods and then outline how to efficiently integrate their dynamics. The 1^{st}-order derivative method tends to outperform the 2^{nd}-order derivative method in the small, static detuning regime, see Figure 2b. The norms of the state derivatives for the 1^{st}- and 2^{nd}-order methods are provided in Table II for the solutions to the numerical experiment detailed in Section V at $t_N = 60$ns. The 2^{nd}-order method is able to decrease the 2^{nd}-order state derivate norm relative to the 1^{st}-order method at the expense of increasing its 1^{st}-order derivative norm. In the problem we have studied it is likely the 2^{nd}-order derivative norm has a smaller contribution to the gate error than the 1^{st}-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	$\left\|\partial_{f_q}\psi_N\right\|_2$ (10^3)	$\left\|\partial^2_{f_q}\psi_N\right\|_2$ (10^6)
D1	0.436	57.817
D2	1.702	9.030

Table II: Norm of state derivates 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 $t_N = 60$ns and are averaged over the four state derivatives.

The dynamics for the derivative methods can be integrated efficiently using exponential integrators [97–99]. 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:

$$|\partial_\lambda\psi_{k+1}\rangle = \exp(\Delta t_k L_k)|\partial_\lambda\psi_k\rangle + \int_0^{\Delta t_k} \exp\Big((\Delta t_k - t^{'})L_k\Big)N(t_k + t^{'})dt^{'} \tag{C1}$$

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 [98],

$$|\partial_\lambda\psi_{k+1}\rangle \approx \exp(\Delta t_k L_k)|\partial_\lambda\psi_k\rangle + \exp(\Delta t_k L_k)N_k \tag{C2}$$

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.

	Wall Time (±0.001 s)					
t_N (ns)	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: 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. ToÅ¡ner, 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.emrs

[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. ToÅ¡ner, 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. Heinzel, 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. Heinzel, 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 Funnels, 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. Liebermann 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 multiqubit gates, New Journal of Physics **20**, 113009 (2018).

[51] 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).

[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] 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).

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

[67] S. J. Julier and J. K. Uhlmann, Unscented filtering and nonlinear estimation, Proceedings of the IEEE **92**, 401 (2004).

[68] 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.

[69] 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.

[70] 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).

[71] S. Machnes, E. Assémat, D. J. Tannor, and F. K. Wilhelm, Gradient optimization of analytic controls: the route to high accuracy quantum optimal control, arXiv preprint arXiv:1507.04261 (2015).

[72] 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).

[73] 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.

[74] 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.

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

[76] 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).

[77] C. R. Hargraves and S. W. Paris, Direct Trajectory Optimization Using Nonlinear Programming and Collocation, J. Guidance **10**, 338 (1987).

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

[79] 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).

[80] B. E. Jackson, T. Punnoose, D. Neamati, K. Tracy, R. Jitosho, 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.

[81] K. L. Clarkson, Coresets, sparse greedy approximation, and the frank-wolfe algorithm, ACM Transactions on Algorithms (TALG) **6**, 1 (2010).

[82] 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.

[83] 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).

[84] 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).

[85] 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).

[86] 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.

[87] A. G. Fowler, A. M. Stephens, and P. Groszkowski, High-threshold universal quantum computation on the surface code, Phys. Rev. A **80**, 052312 (2009).

[88] D. Gottesman, Stabilizer codes and quantum error correction, arXiv preprint quant-ph/9705052 (1997).

[89] 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).

[90] 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).

[91] 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).

[92] 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).

[93] 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).

[94] J. O. Smith, *Spectral Audio Signal Processing* (`http://ccrma.stanford.edu/ jos/sasp/`, 2020) online book, 2011 edition.

[95] G. T. Landi, Lecture notes on quantum information and quantum noise (2018).

[96] 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.

[97] N. Auer, L. Einkemmer, P. Kandolf, and A. Ostermann, Magnus integrators on multicore cpus and gpus, Computer Physics Communications **228**, 115 (2018).

[98] H. Berland and B. Skaflestad, *Solving the nonlinear Schrödinger equation using exponential integrators*, Tech. Rep. (Norwegian University of Science and Technology, 2005).

[99] L. Einkemmer, M. Tokman, and J. Loffeld, On the performance of exponential integrators for problems in magnetohydrodynamics, Journal of Computational Physics **330**, 550 (2017).