Robust Control of a Fluxonium Qubit

Thomas Propson,[1,2,*] Brian Jackson,[3] Zac 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, USA*
[4] *Pritzker School of Molecular Engineering, University of Chicago, Chicago, Illinois 60637, USA*

(Dated: November 24, 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 supress 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 differentation. 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 realisitc 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], and applications now include superconducting circuits [2–6], neutral and ionized atoms [7], nitrogen-vacancy centers in diamond [8], and Bose-Einstien condensates [9]. For quantum computation QOC techniques are employed to achieve high fidelity gates while adhering to experimental constraints. The decision variables of the optimization problem are the time-dependent control parameters, particular to the quantum system, that govern its evolution. Experimental errors may cause the system evolution to deviate from that predicted in optimization, hampering performance. The field of classical control theory has developed robust control techniques to encode experimental errors in optimization objectives, improving experimental performance (citations appreciated). In this work we employ robust control techniques to mitigate realistic system parameter deviations and decoherence that arise when controlling a superconducting fluxonium qubit.

Analytic and numerical techniques for QOC have been remarkably successful in designing high fidelity gates (citations needed). While these methods are able to predict system behavior with high accuracy, they tend to be sensitive to experimental errors such as parameter drift, noise, and finite control resolution. In order to address these issues, recent work has sought to mitigate errors due to parameter deviations. Considering the dynamical and geometric phases of quantum state trajectories has led to analytic methods for mitigating errors due to parameter deviations and pure dephasing [6, 10–12]. Floquet techniques have been experimentally demonstrated to simultaneously mitigate longitudinal relaxation and pure dephasing decoherence [3, 13]. Numerical QOC techniques have been adapted to mitigate longitudinal relaxation by modeling master equations [8] and employing Monte Carlo style quantum trajectories [14]. Additionally, efforts have been made to incorporate experimental feedback into optimization [3]. Comprehensive tools that interleave characterization and pulse design are under active development [15].

In this work we mitigate errors arising from longitudinal relaxation, parameter deviations, and $1/f$ flux noise on the fluxonium qubit [12] by solving a robust trajectory optimization problem. Leveraging recent advances in trajectory optimization within the field of robotics, we solve this optimization problem using the ALTRO solver, which uses iterative LQR (iLQR)—a differential-dynamic programming (DDP)-based indirect method similar to shooting methods such as GOAT [16], GRAPE [1, 5], and Krotov's [17]—within an augmented Lagrangian framework to handle nonlinear equality and inequality constraints at each time step [18]. To mitigate errors due to longitudinal relaxation, we perform time-optimal control and encode the dependence of longitudinal relaxation on the controls in an efficient optimization objective that does not pay the increased computational cost of integrating a master equation. Additionally, to make the quantum state trajectory robust to parameter deviations, we consider three different approaches to making the optimized trajectory less sensitive to model uncertainty:

1. A sampling method, similar to the work of [8, 19, 20].

2. An unscented sampling method, adapted from the

* tcpropson@uchicago.edu

unscented tranform used in the state estimation community [21–24].

3. A derivative-penalization method, which uses efficient mixed-mode differentiation to compute derivative information of parameter deviations with respect to the quantum state trajectory.

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

II. QOC + AL-ILQR

In this section we introduce the notation we will use throughout the paper, review the quantum optimal control problem statement, and introduce the trajectory optimization framework. Quantum optimal control concerns the evolution of a quantum state $|\psi(t)\rangle$ governed by the time-dependent Schrödinger equation (TDSE)

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

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. Numerical quantum optimal control techniques make the problem tractable by discretizing the problem into N knot points (time steps). Typical explicit integration techniques for the TDSE include exponential integrators [25–27] and Runge-Kutta integrators [28].

Quantum optimal control seeks the control parameters that minimize a functional $J(u(t))$. In the simplest case the functional is the infidelity between the inital state evolved to the final knot point and the target state $J = 1 - |\langle\psi_f|\psi_N(u(t))\rangle|^2$. In general J is a linear combinaion of cost functions on the state, e.g. forbidden-state occupation, as well as cost functions on the controls, e.g. the norm of the control amplitudes [5]. Typical quantum optimal control algorithms employ automatic differentiation to compute first order information for the functional $\nabla_u J(u)$. This information is used in a first-order optimizer to minimize J with respect to u. This scheme lends itself to projected gradient methods which restrict the optimization to the constraint manifold [29, 30].

Alternatively, the QOC problem can be formulated as a trajectory optimization problem and solved using any of the state-of-the-art solvers developed by the robotics community. Trajectory optimization problems are typically of the following form:

$$\begin{aligned} \underset{x_{0:N},\,u_{0:N-1}}{\text{minimize}} \quad & \ell_f(x_N) + \sum_{k=0}^{N-1} \ell(x_k,u_k) \\ \text{subject to} \quad & x_{k+1} = f(x_k,u_k), \\ & g_k(x_k,u_k) \le 0, \\ & h_k(x_k,u_k) = 0, \end{aligned} \tag{2}$$

where ℓ_f and ℓ are the final and stage cost functions, $x_k \in \mathbb{R}^n$ and $u_k \in \mathbb{R}^m$ are the state and input 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 (2). Standard methods include direction collocation [31] and differential-dynamic programming (DDP) [32]. Recent state-of-the-art solvers, such as ALTRO [18], have combined principles from both of these approaches.

ALTRO uses iterative LQR (iLQR) [33] 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 closed-loop control policy about the optimal trajectory for free; 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, $c(z) : \mathbb{R}^n \mapsto \mathbb{R}^p$ is the constraint function, $\lambda \in \mathbb{R}^p$ are the Lagrange multipliers, and I_μ is a diagonal matrix of penalty weights, whose magnitude depend on whether the constraint is active or inactive. For ALTRO, $f(z)$ is the objective function of (2), $c(z)$ 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 penalty terms and Lagrange multipliers 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 [18, 34].

III. QOC ON THE FLUXONIUM

In the following we study the quantum optimal control problem on the fluxonium qubit. The fluxonium qubit is a promising building block for superconducting circuits, and the experimental constraints we encode in the optimization reflect those of a realistic device. Furthermore, the accurate two-level approximation of the system Hamiltonian makes it efficient to perform quantum optimal control on a classical computer. In the two-level approximation the system Hamiltonian takes the form

$$H/h = f_q \frac{\sigma_z}{2} + a(t)\frac{\sigma_x}{2} \tag{4}$$

where $f_q = 14$MHz is the qubit frequency at the flux frustration point, a is the flux drive amplitude, h is Planck's constant, and σ_x, σ_y are Pauli matrices. The flux amplitude a is experimentally realized by modulating the flux threading the device. We consider the task of constructing $Z/2$, $Y/2$, and $X/2$ gates for the fluxonium qubit subject to experimental constraints, decoherence, and systematic errors. We compare the gates we obtain with numerical methods to the analytically constructed gates reported in [12] for the same device.

The optimization problem takes the form:

$$\underset{x_{0:N},\, u_{0:N\text{-}1}}{\text{minimize}} \quad \sum_{k=1}^{N} \|x_k\text{-}x_f\|_{Q_k} + \sum_{k=1}^{N-1} \|u_k\|_{R_k} \tag{5a}$$

$$\text{subject to} \quad x_{k+1} = f(x_k, u_k), \tag{5b}$$

$$x_1 = x_{\text{init}}, \tag{5c}$$

$$x_N = x_f \tag{5d}$$

where the augmented state and controls are given by:

$$x = \begin{bmatrix} \psi^{(1)} \\ \psi^{(2)} \\ \int a\ dt \\ a \\ \partial a/\partial t \end{bmatrix} \qquad u = \begin{bmatrix} \partial^2 a/\partial t^2 \end{bmatrix} \tag{6}$$

The initial and final conditions are determined by the desired unitary transformation, which requires two quantum states to specify $\psi^{(1)}$, $\psi^{(2)}$. The ALTRO implementation we use does not currently support complex numbers so we compute in the isomorphism $\mathcal{H}(\mathbb{R}^{2n}) \cong \mathcal{H}(\mathbb{C}^n)$ given in [5]:

$$H\psi \cong \begin{bmatrix} H_{\text{re}} & -H_{\text{im}} \\ H_{\text{im}} & H_{\text{re}} \end{bmatrix} \begin{bmatrix} \psi_{\text{re}} \\ \psi_{\text{im}} \end{bmatrix} \tag{7}$$

The dynamics function $f(x_k, u_k)$ integrates the TDSE dynamics for the quantum states (1), (4) and integrates the moments of the flux drive amplitude. Exposing lower order moments of the flux drive amplitude allows us to penalize their norms, smoothing the flux drive amplitude and mitigating AWG ringing due to high frequency transitions. The matrices Q_k and R_k define the penalty metric. We choose Q to be diagonal because it is computationally efficient. This corresponds to penalizing phase differences between the quantum states, although other phase-insensitive metrics such as infidelity may be employed. The final state transfer is the most important objective which is encoded in the realtive metric weight $Q_N \sim N \cdot Q_k$.

Additionally, we impose constraints that reflect limitations of the apparatus and ones that improve the experimental realization of the control pulse. To ensure gates may be concatenated arbitrarily without inducing AWG ringing due to high-frequency transitions, we require $a(t = 0) = a(t = t_N) = 0$. Furthermore, we require $\int_0^{t_N} a(t)dt = 0$. This constraint ensures the pulse has zero net flux, mitigating the hysteresis ubiquitous in flux bias lines, [12, 35, 36]. We require $-0.5\text{GHz} \leq a(t) \leq 0.5\text{GHz}$ to ensure the two-level approximation (4) remains valid. Additionally, we require that each gate achieves the desired state transition $\psi_N = \psi_f$. Both the zero net flux and target quantum state constraint are then handled by ensuring the target augmented state is reached $x_N = x_f$, the constraint function takes the form $c_k = (x_k - x_f)^T(x_k - x_f)$. The equality and inequality constraints on a are handled with a bound constraint which takes the form $c_k = (x_k - b)^T(x_k - b)$ if $|x_k| \geq b$ and 0 otherwise. In addition the norm of the quantum states are constrained to 1 to ensure discretization error is not exploited $c_k = x_k^T x_k - 1$.

In addition to obeying experimental constraints and achieving low simulated gate errors, we desire to make our gates perform well in the presence of decoherence. Decoherence of the quantum state due to external noise is typically modeled by two phenomena: longitudinal relaxation and pure dephasing. They are modeled using their $1/e$ decay times T_1 and T_ϕ respectively. The main contributions to longitudinal relaxation in our device are dielectric loss in the capacitor, resistive loss in the inductor, and Purcell loss. The main contributions to pure dephasing in our device are $1/f$ flux noise and decay via charge and flux coupling to the control lines. Dissipation to the thermal bath via longitudinal relaxation is an irreversible process that results in information loss. Converesely, pure dephasing is a reversible process. Our device achieves its best pure dephasing protection at the flux frustration point $T_{2e}(a = 0) \sim 300\mu$s where the qubit is first-order insensitive to changes in flux. It becomes more succeptable to pure dephasing as the flux is tuned away from the flux frustration point $|a| > 0$. Conversely, longitudinal relaxation is at a minimum at the flux frustration point $T_1(a = 0) = 0.315$ms, and increases away from the flux frustration point $T_1(a = 0.34\text{GHz}) = 4.3$ms. Given the nature of the decay processes and the tradeoff, we choose to maximize the longitudinal relaxation time directly by tuning the flux amplitude IV and employ robust control techniques to mitigate pure dephasing V, VI.

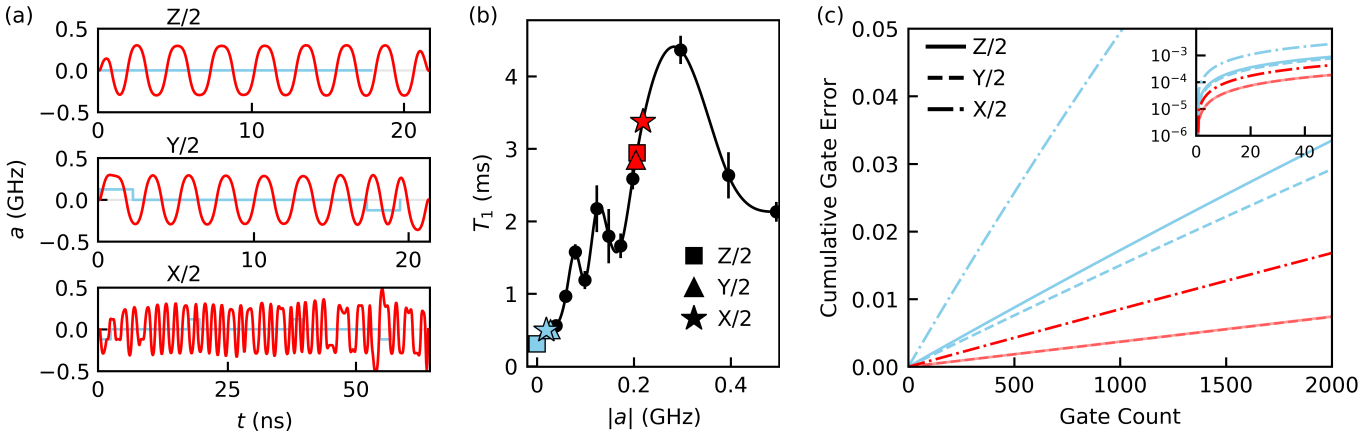

Figure 1: (a) Numerically optimized gates (red) and analytically optimized gates (blue). (b) T_1 interpolation function used in optimization. Markers denote the time-averaged, absolute amplitude of each gate. (c) Lindblad master equation simulation with T_1 dissipation for successive gate applications.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

IV. LONGITUDINAL RELAXATION AWARENESS

The longitudinal relaxation time T_1 varies with control parameters in a range of superconducting circuit platforms. It is advantageous to tune the controls to extend the longitudinal relaxation time. Computing the gate error due to longitudinal relaxation requires propagating density matrices of size $n \times n$ under master equation dynamics, rather than state vectors of size n under the TDSE dynamics. We avoid this increase in computational complexity by penalizing the integrated rate of longitudinal relaxation, i.e. the probability of longitudinal relaxation. Using this probability as 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 duration would favor a lower longitudinal relaxation probability. We allow the optimizer to tune the gate duration in order to minimize the longitudinal relaxation probability. Our scheme for time-optimal control is applicable to any time-optimal problem, not only the one we study here.

TODO: again, I'd recommend writing out either the entire optimization problem or at least point to the parts of the optimization problem defined in the previous section you had to change. Just make it a little more concrete.

The longitudinal relaxation probability is given by

$$P_1(t) = \int_0^t T_1^{-1}(a(t'))dt' \tag{8}$$

P_1 is penalized using a quadratic cost at each knot point $|P_1(t_k)|^2$. $T_1(a_k)$ is obtained at each knot point by evaluating a spline fit to experimental data of the form $\{(a, T_1)\}$. It is also possible to use a spline fit to theoretically obtained data. However, T_1 values are known to fluctuate greatly with laboratory temperatures [37]. Interpolating T_1 from experimental data increases the fridge truth of the simulation.

We allow the optimizer to tune the gate duration 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 Markovian decision structure of the trajectory optimization problem. To ensure numerical integration accuracy is maintained we add a bound constraint at each knot point 5e-2 ns $\leq \Delta t_k \leq$ 2e-1 ns. This bound constraint may be violated for intermediate iterations of the optimization, so we add the square root of the time step $\sqrt{\Delta t_k}$ to the augmented control vector and use the squared root of the time step $|\Delta t_k|$ in optimization.

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 $\sim$ 0.2GHz, see Figure 1. 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 which is 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 target state averaged over 1000 pseudo-randomly generated initial states. The numerical $Z/2$ and $Y/2$ gates perform simlarly in the concatenated gate application comparision, suppressing accumulated gate errors to $8 \cdot 10^{-3}$ over 2000 gate applications $\sim 40\mu$s.

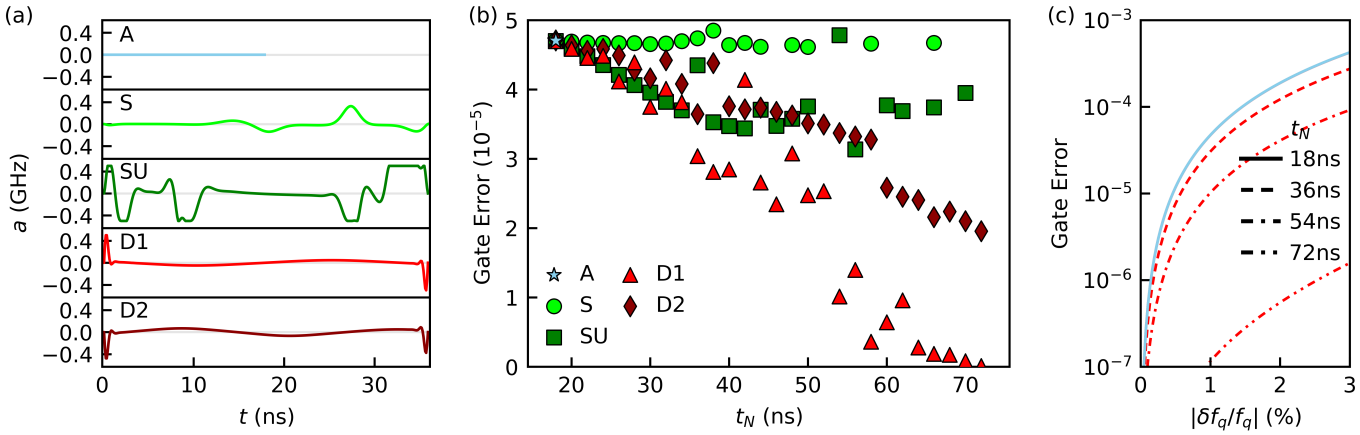

Figure 2: (a) $Z/2$ gates robust to qubit frequency detunings constructed with the analytic, sampling, unscented sampling, and the 1st- and 2nd-order derivative methods. The gates shown for the numerical methods are the solutions at twice the analytic gate time. (b) Single gate error as a function of the gate duration at a one-percent detuning from the nominal qubit frequency 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 detuning from the nominal qubit frequency. The solutions for the analytic and 1st-order derivative methods are shown at multiples of the analytic gate time. The performance of the two methods is indistinguishable at the analytic gate time 18ns.

The numerical $X/2$ gate achieves an accumulated gate error of $1.7 \cdot 10^{-2}$ over 2000 gate applications $\sim 124\mu$s. Both the analytic and numerical gates attain single gate errors sufficient for quantum error correction $< 10^{-4}$, requisite for fault-tolerant quantum computing. The low gate errors achieved by the numerical gates TODO: typo / awkward sentence for extend computations 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 DEVIATIONS

We have formulated the quantum optimal control problem as an open loop optimization problem, i.e. feedback from the experiment is not incorporated 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 $\tilde{f}_q = f_q \pm \sigma_{f_q}$ TODO: maybe reference (4) here?. We consider three robust control techniques. The first is the sampling method, which has been proposed previously in the context of QOC [8, 19, 20]. In the sampling method, multiple states are evolved under distinct deviant dynamics to capture the effect of parameter deviations. We also study the unscented sampling method, which uses the unscented transformation to accurately propagate a distribution representing the uncertainty in an evolving state due to a parameter deviation. The unscented transformation was designed for nonlinear Kalman filtering and is frequently utilized in robust control [21–24]. Finally, we propose the derivative method. Derivative information encoding the sensitivity of the state trajectory with respect to the deviant parameter is used to penalize state trajectory deviations.

TODO: I think you can make this more concrete and simpler to understand by just defining the new augmented state vector and explaining you stack multiple copies of the state vector together, each evolving with a different static parameter. In this case there's only one parameter being changed, so I'd just state it explicitly. In the sampling method, sample states evolve under a Hamiltonian where a parameter is replaced by a deviant value. We propagate the additional states $|\psi^{\pm}\rangle$ with deviant values $\lambda^{\pm} = \lambda \pm \sigma_{\lambda}$. The gate error of each sample state is penalized at each knot point $1-|\langle\psi_k^{\pm}|\psi_f|\psi_k^{\pm}|\psi_f\rangle|^2$. Each initial state is used to initialize two sample states $|\psi^{\pm}\rangle$. The initial states are chosen so that their outer products span the operators on the Hilbert space $\{|0\rangle, |1\rangle, (|0\rangle + i\,|1\rangle)/\sqrt{2}, (|0\rangle - |1\rangle)/\sqrt{2}\}$ [38]. For this method the number of states in the augmented state vector scales as $O(dn^2)$ TODO: is this right? because each of n^2 initial states is represented by $2d$ samples. n is both the dimension of the Hilbert space and the size of the state vector. d is the number of deviant parameters.

The strategy of the unscented sampling method is to propagate an ensemble of sample states (sigma points) which represent a distribution over every element of the initial state. The distribution models the uncertainty in

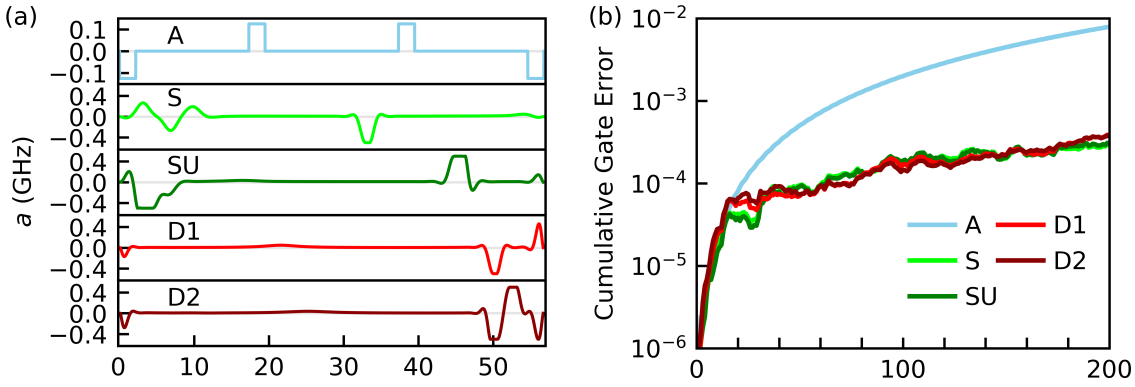

Figure 3: (a) $X/2$ gates robust to flux offsets constructed with the analytic, sampling, unscented sampling, and the 1^{st}- and 2^{nd}-order derivative methods. The gates shown are the solutions at the analytic gate time. (b) Simulation of stochastic $1/f$ flux noise for successive gate applications.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

each state element arising from the parameter deviation. The ensemble consists of $4n + 2d$ states. Each state in the ensemble is propagated to the next knot point using separate deviant dynamics. Then, the mean and covariance of the propagated ensemble is calculated. New states are sampled from the distribution given by the calculated statistics for propagation to the next knot point. This resampling procedure, the unscented transformation, accurately propagates first and second moments of the distribution and ensures the points lie on the $\sqrt{2n^{\text{th}}}$ covariance contour at each knot point. A detailed update procedure is given in Appendix B. Each initial state from the operator basis is represented with a distribution of $4n + 2d$ samples. Hence, the number of states in the augmented state vector scales as $O(n^3 + dn^2)$. The gate error of each sample state is penalized as in the sampling method. TODO: I think this section needs some more detailed explanation.

The derivative method draws on the intuition that the sensitivity of the state evolution to a parameter λ is encoded in the l^{th}-order derivative of the state with respect to that parameter $\left|\partial_\lambda^l \psi\right\rangle$. The m^{th}-order derivative method minimizes the norm of the first m state derivatives with a quadratic cost at each knot point $\left|\left\langle \partial_\lambda^l \psi | \partial_\lambda^l \psi | \partial_\lambda^l \psi | \partial_\lambda^l \psi \right\rangle\right|^2$, $l \in \{1, \ldots, m\}$. The state derivatives could be obtained with backwards mode differentiation. Naive automatic differentiation would compute the state derivative at all $1, \ldots, k-1$ knot points to obtain the state derivative at knot point k. For a single state derivative and N knot points this requires $O(N^2)$ matrix multiplications. Instead, we forward propagate the state derivatives in the augmented state vector under coupled dynamics, resulting in $O(N)$ matrix multiplications. For example, the dynamics for the 1^{st}-order derivative method are

$$i\hbar \frac{d}{dt} \left|\psi\right\rangle = H \left|\psi\right\rangle \tag{9}$$

$$i\hbar \frac{d}{dt} \left|\partial_\lambda \psi\right\rangle = H \left|\partial_\lambda \psi\right\rangle + \left(\partial_\lambda H\right) \left|\psi\right\rangle \tag{10}$$

TODO: inconsistent use of $\hbar$ vs h. Exponential integrators that account for the non-linear term may be used to efficiently integrate the coupled dynamics, see Appendix C. For this method m state derivatives are associated with each initial state from the operator basis. So, the number of states in the augmented state vector scales as $O(dmn^2)$. Additionally, we note that the number of initial states and samples we have employed for each of the three methods follows standard perscriptions. However, knowledge of the problem can typically be used to reduced the number of required samples. TODO: I think this also needs some more explanation. It's not clear exactly what you're doing. You're penalizing the derivatives of the dynamics with respect to the uncertain parameter in the objective, so why do you need to augment the state vector? Or are you adding the derivatives to the state vector so that you have standard quadratic penalty terms in the objective? Again, I think this confusion would be easily cleared up by explicitly writing down your new dynamics, objective function, and any additional constraints you impose on the system. Since this is really the meat of the paper, it's worth spending the time to make sure this is clear and easy to understand, especially for those who aren't very familiar with trajectory optimization or optimization in general. Also, I think you should explicitly state the size of the problem you're actually solving. Might be worth putting in a table, along with some timing results, to give an idea of how well these methods scale.

To demonstrate the applicability of these techniques to mitigate parameter deviations, we consider the task of achieving a single $Z/2$ gate subject to a constant qubit

frequency detuning $f_q \leftarrow f_q + \delta f_q$. We take $\sigma_{f_q}/f_q = 1\%$ to be one standard devation, and equip the sampling methods accordingly. For each method we compute the gate error for one simulated gate application subject to the deviant dynamics given by the stated qubit frequency detuning δf_q.

We compare the numerical methods to an analytically derived $Z/2$ gate, see Figure 2a TODO: LaTeX suggetion: don't hard-code the sublabels. You can directly reference them.. This gate corresponds to idling at the flux frustration point $a = 0$. The analytic gate 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 detutning. At a one-percent qubit frequency detuning the analytic gate achieves a gate error $\sim 4.5 \cdot 10^{-5}$, which is sufficient for quantum error correction, see Figure 2b. Although the analytic $Z/2$ gate performs well, it works only at the gate time $t_{Z/2}$. The ability to perform Z rotations in arbitrary times is critical for operating multi-qubit experiments in the lab frame. Each numerical method is able to find solutions at all gate times above $t_{Z/2}$, but is unable to find solutions at shorter times. These numerical methods offer an effective scheme for synchronizing qubits operating at different frequencies $f_{q,i} \neq f_{q,j}$.

The sampling and unscented sampling methods converge on qualitatively similar solutions which combine idling periods with fast ramps to the maximum amplitude. The gate error at a one-percent detuning from the nominal qubit frequency achieved by the sampling method does not improve substantially over the range of gate durations. The unscented sampling method achieves linear decreases in its gate error with longer gate durations 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 traingle 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 the gate duration. The gate error for the 1^{st}-order derivative method approaches zero at the larmor period $1/f_q$, see Figure 2c. We believe the 1^{st}-order method outperforms the 2^{nd}-order method due to the low contribution of second-order terms to the gate error in this deviation regime, see Appendix C.

Additionally, there are multiple analytic methods to mitigate errors due to parameter deviations including composite pulse sequences [11], the DRAG scheme [36], and geometric phase considerations [6, 10]. Composite pulse sequences are derived by computing the erroneous rotation arising from the parameter deviation with a finite order magnus expansion. Pulses are then suitably composed to eliminate the error. In principle the error may be eliminated to arbitrary order with sufficiently many pulses [**?**], which mirrors the negative correlation between gate error and gate duration we observe with the derivative method. It is difficult to choose an appropriate composite pulse for the problem studied here, so we propose comparisons between composite pulses and numerical techniques for future work.

VI. ROBUSTNESS TO STOCHASTIC PARAMETER DEVIATIONS

An additional source of experimental error arises from stochastic, time-varying 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 qubit's amplitude from its nominal value by an amount δa. It is well studied that the spectral density of δa follows a $1/f$ distribution for a range of devices, consisting primarily of low frequency noise (citations needed). 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. This is the strategy employed by the analytic gate considered here.

We compare the analytic gate and those produced by the numerical methods discussed in the previous section on the task of realizing a $X/2$ gate subject to $1/f$ flux noise. The flux noise is generated by filtering white noise sampled from a standard normal distribution with a finite impulse response filter [39] [40]. It is then scaled by the flux noise amplitude of our device $A_\Phi = 5.21\mu\Phi_0 \implies \delta a \sim 2.5 \cdot 10^{-5}$GHz. The unscented sampling method is modified so that its sampled deviations follow a $1/f$ distribution by carrying the state of a finite impulse response filter in the augmented state vector. In principle the basic sampling method could be modified similarly but we choose to sample statically at δa for comparision. The derivative methods require no modification from the static case. TODO: can you provide a little more detail on how this noise factors into the dynamics? Are you just adding some white noise to $a(t)$? Are you also considering uncertainty in f_q at the same time?

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 comparision. 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 signficiant source of coherent errors in NISQ applications and these numerical techniques offer effective avenues to mitigate them.

VII. CONCLUSION

In conclusion, we have demonstrated techniques for achieving robustness to systematic errors and mitigating decoherence on a quantum system using state-of-the-art trajectory optimization. We have proposed a scheme for mitigating longitudinal relaxation with time-optimal control and an efficient optimization metric that comes at a constant computational cost as opposed to integrating a master equation which scales quadratically with the dimension of the Hilbert space. We have proposed the derivative method for robust control which achieves super-linear gate error reductions in the gate duration for the problem we studied here. We have shown that robust control techniques can be used to mitigate decoherence due to $1/f$ flux noise, a dominant source of coherent errors for flux controlled qubits. The numerical techniques we have studied here can be used to perform phase gates in arbitrary times, which will be critical for synchronizing multi-qubit systems TODO: it might be worth mentioning the scaling properties of your method here, since you suggest scaling it up to systems with more states. How may the different methods you suggest perform at scale? Maybe there will be an advantage to using the sampling-based methods as the problems grow? Also you don't mention the sampling methods in the conclusion? The derivative methods only outperformed the sampling methods on the simpler of your 2 test cases, so I'd be hesitant to imply that the derivative method is the better choice. Additionally, interleaving the error models with existing characterization methods will improve their effectiveness in experiments. These techniques will be employed 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 and Tanay Roy for useful discussions.

Appendix A: Longitudinal Relaxation

We comment on the longitudinal relaxation metrics and then give our procedure for integrating the Lindblad master equation. The longitudal relaxation probability and the gate error due to longitudinal relaxation metrics are compared in Table I for the experiment shown in Figure 1c. The relative performance of the analytic and numerical techniques is similar across the two metrics.

To compute the gate error due to longitudinal relaxation, we require the final state of the quantum system subject to longitudinal relaxation. To obtain the final state we employ the Lindblad master equation. This equation takes the form

Gate	$P_{1\mathrm{A}}$ (10^{-5})	$P_{1\mathrm{N}}$ (10^{-5})	$P_{1\mathrm{A}}/P_{1\mathrm{N}}$	GE_{A} (10^{-5})	GE_{N} (10^{-5})	$\mathrm{GE}_{\mathrm{A}}/\mathrm{GE}_{\mathrm{N}}$
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 ratios and single gate error due to longitudinal relaxation ratios. Values are reported for the analytic (A) and numerical (N) techniques presented in the main text.

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

Here $\rho = |\psi\rangle\langle\psi|$ is the density matrix, $n = \dim(\mathcal{H})$, and $[\cdot,\cdot],\{\cdot,\cdot\}$ are the algebraic commutator and 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. The T_1 values in this spline are obtained experimentally by driving the qubit at the desired flux amplitude and monitoring the resultant decay. For more details consult [12].

Exponential integrators can be employed to integrate the Lindblad master equation using the Vectorization/Choi-Jamiolkowski isomorphism [41]

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

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

Here $\rho = \sum_{i,j}\alpha_{i,j}|i\rangle\langle j|$ and $\mathrm{vec}(\rho) = \sum_{i,j}\alpha_{i,j}|i\rangle\otimes|j\rangle$. We use zero-order hold on the controls so the integration is exact $\mathrm{vec}(\rho_{k+1}) = \exp\left(\Delta t_k\hat{\mathcal{L}}_k\right)\mathrm{vec}(\rho_k)$. 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 held 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. The unscented sampling method concerns a quantum state $\psi \in \mathbb{R}^{2n}$ with deviant parameters $\lambda \in \mathbb{R}^d$ and dynamics $\psi_{k+1} = f(\psi_k, \lambda_k)$. The nominal initial quantum state is given by ψ_0 with an associated positive-definite covariance matrix $P_0 \in \mathbb{R}^{n\times n}_{++}$ which describes the uncertainty in the initial state. P_0 is typically chosen to be non-zero even if the state preparation error is negligible. The deviant parameter is assumed to have 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 dynamics f.

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

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

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 [22]. Other prescriptions may be chosen to reduce the number of required sigma points [42].

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 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 could be completed for future problems to predict the efficacy of the derivative method at each order.

Method	$\left\|\langle \partial_{f_q}\psi_N \vert \partial_{f_q}\psi_N \vert \partial_{f_q}\psi_N \vert \partial_{f_q}\psi_N \rangle\right\|^2$ (10^3)	$\left\|\left\langle \partial^2_{f_q}\psi_N \vert \partial^2_{f_q}\psi_N \right.\right.$	$\partial^2_{f_q}$
D-1	0.436		57.8
D-2	1.702		9.03

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's duration $t_N = 60$ns and are averaged over the four state derivatives.

The dynamics for the derivative methods can be integrated efficiently using exponential integrators. General exponential integrators break the dynamics into a linear term and a non-linear term. For example, consider integrating the dynamics of the first state derivative $\frac{d}{dt} |\partial_\lambda \psi\rangle = H |\partial_\lambda \psi\rangle + (\partial_\lambda H) |\psi\rangle$ in units of $i\hbar = 1$. 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 propagation is

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

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 [26]

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

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.

[1] 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).

[2] 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).

[3] 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).

[4] 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).

[5] 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).

[6] 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).

[7] 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).

[8] 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).

[9] 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).

[10] 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).

[11] J. T. Merrill and K. R. Brown, Progress in compensating pulse sequences for quantum computation, Quantum Information and Computation for Chemistry , 241 (2014).

[12] 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).

[13] 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).

[14] 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).

[15] 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].

[16] 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).

[17] 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).

[18] 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.

[19] 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).

[20] P. Reinhold, *Controlling Error-Correctable Bosonic Qubits*, Ph.D. thesis, Yale University (2019).

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

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

[23] 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.

[24] 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.

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

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

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

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

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

[30] 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.

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

[32] D. Q. Mayne, A second-order gradient method of optimizing non- linear discrete time systems, Int J Control

3, 8595 (1966).
[33] 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).
[34] 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.
[35] 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).
[36] 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).
[37] 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).
[38] 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).
[39] J. O. Smith, *Spectral Audio Signal Processing* (`http://ccrma.stanford.edu/ jos/sasp/`, 2020) online book, 2011 edition.
[40] `https://ccrma.stanford.edu/~jos/sasp/Example_Synthesis_1_F_Noise.html`.
[41] G. T. Landi, Lecture notes on quantum information and quantum noise (2018).
[42] 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.