# Robust Trajectory Optimization on a Quantum System

Thomas Propson,[1, 2, *] Brian Jackson,[3] Zachary Manchester,[3] and David I. Schuster[1, 2, 4]

[1] *James Franck Institute, University of Chicago, Chicago, Illinois 60637, USA*
[2] *Department of Physics, University of Chicago, Chicago, Illinois 60637, USA*
[3] *Robotics Institute, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213, USA*
[4] *Pritzker School of Molecular Engineering, University of Chicago, Chicago, Illinois 60637, USA*

(Dated: January 26, 2021)

The ability to engineer high-fidelity gates on quantum processors in the presence of systematic errors remains the primary challenge requisite to achieving quantum advantage. Quantum optimal control methods have proven effective in experimentally realizing high-fidelity gates, but they require exquisite calibration to be performant. We apply robust trajectory optimization techniques to suppress gate errors arising from system parameter uncertainty. We propose a method that takes advantage of uncertain parameter derivative information while maintaining computational efficiency by transforming high-order differential equations to coupled, first-order ODEs. Additionally, the effect of depolarization on a gate is most accurately modeled by integrating the Lindblad master equation, which is computationally expensive. We propose a computationally efficient model and utilize time-optimal control to achieve high fidelity gates in the presence of depolarization. We apply these techniques to a fluxonium qubit and suppress simulated gate errors due to parameter uncertainty below $10^{-7}$ for static parameter deviations on the order of 1%.

## I. INTRODUCTION

Quantum optimal control (QOC) is a class of optimization algorithms for accurately and efficiently manipulating quantum systems. Early techniques were proposed for nuclear magnetic resonance experiments [1–7], and applications now include superconducting circuits [8–14], neutral atoms and ions [15–25], nitrogen-vacancy centers in diamond [26–32], and Bose-Einstein condensates [33, 34]. In the context of quantum computation, optimal control is employed to achieve high-fidelity gates while adhering to experimental constraints. Experimental errors such as parameter drift, noise, and finite control resolution cause the system to deviate from the model used in optimization, leading to poor experimental performance. Robust control improves upon standard optimal control by encoding model parameter uncertainties in optimization objectives, yielding performance guarantees over a range of parameter values [35–37]. We adapt robust control techniques from the robotics community to mitigate parameter uncertainty errors for a superconducting fluxonium qubit.

Analytically-dervied control pulses that mitigate parameter uncertainty errors include composite pulses [38–41], pulses designed by considering dynamic and geometric phases [42, 43], and pulses obtained with the DRAG scheme [44]. These techniques are not amenable to arbitrary experimental constraints, so recent work has sought to achieve robustness in quantum optimal control frameworks using closed-loop methods [45–47] and open-loop methods [31, 48–51].

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

1. A sampling method, similar to the work of [3, 31, 48–50, 52].

2. An unscented sampling method adapted from the unscented transform used in state estimation [53–56].

3. A derivative method, which penalizes the sensitivity of the quantum state trajectory to uncertain parameters.

We apply these techniques to the fluxonium qubit presented in [57]. We also show that QOC can solve important problems associated with fluxonium-based qubits: mitigating depolarization by taking advantage of the $T_1$-dependence of the controls and synchronizing qubits with distinct frequencies by performing phase gates in arbitrary times. To mitigate depolarization, we perform time-optimal control and employ an efficient depolarization model for which the computational cost is independent of the Hilbert space dimension. Leveraging recent advances in trajectory optimization within the field of robotics, we solve these optimization problems using ALTRO (Augmented Lagrangian TRajectory Optimizer) [58], which can enforce constraints on the quantum state trajectory and control parameters.

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

* tcpropson@uchicago.edu

## II. BACKGROUND

In this section, we review the QOC problem statement and describe the ALTRO solver [58]. QOC concerns a vector of time-dependent control parameters $u(t)$ that steer the evolution of a (quantum) state $|\psi(t)\rangle$. The evolution of the state is governed by the time-dependent Schrödinger equation (TDSE),

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

The Hamiltonian $H(u(t), t)$ is determined by the quantum system. The QOC problem is to find the controls that minimize a functional $J(u(t))$. To make the problem numerically tractable, the controls and state are discretized into $N$ knot points (time steps). In the case of a single state transfer problem, the functional is the infidelity of the desired final state and the initial state evolved to the final knot point, $J(u) = 1 - |\langle\psi_f|\psi_N(u)\rangle|^2$. In general, $J(u)$ is a linear combination of cost functions on the state as well as cost functions on the controls. Standard QOC solvers compute derivatives of the functional $\nabla_u J(u)$, which can easily be used to implement first-order optimization methods [3, 13, 59, 60].

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 [58, 61–63]. The functional $J(u)$ is divided into its constituent cost functions at each knot point $\ell_k(x_k, u_k)$, where $k \in \{1, ..., N\}$ denotes the knot point, $x_k$ is the augmented state vector, and $u_k$ is the augmented control vector. The augmented state contains all relevant variables that are dependent upon the controls, for example the state $|\psi_k\rangle \subseteq x_k$. The augmented state trajectory obeys the physics of the system if the dynamics constraint is satisfied $x_{k+1} = f(x_k, u_k)$. For QOC, the discrete dynamics function $f(x_k, u_k)$ propagates the state by integrating the TDSE 1 with Runge-Kutta methods [64] or exponential integrators [65–68].

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

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

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

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

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

Standard techniques for solving (2a)-(2d) typically fall into two categories: direct methods [69, 70] and indirect methods [71]. For indirect methods, the augmented controls are the decision variables–the variables the optimizer adjusts to solve the problem. The augmented states are obtained from the augmented controls using the discrete dynamics function, and are then used to evaluate derivatives of the cost functions (2a). Then, the derivative information is employed to update the augmented controls. This is the approach taken by standard QOC solvers such as GOAT [59], GRAPE [3, 13], and Krotov's method [60]. Conversely, direct methods treat both the augmented controls and the augmented states as decision variables. In addition to minimizing the cost functions, the optimizer uses derivative information for the discrete dynamics function to satisfy the dynamics constraint (2b) to a specified tolerance. In this sense, the TDSE (1) is a constraint which may be violated for intermediate steps of the optimization, where the states need not be physical. The direct approach lends itself to a nonlinear program formulation, for which a variety of general purpose solvers exist [72, 73]. Recent state-of-the-art solvers, such as ALTRO, have combined principles from both of these approaches. ALTRO uses an iterative linear-quadratic regulator (iLQR) algorithm [74] as the internal solver of an augmented Lagrangian method (ALM) [75, 76] and employs a projected Newton method [77, 78] in its final solving stage.

iLQR is an indirect method for solving the dynamically constrained trajectory optimization problem (2a)-(2b), and its update procedure is based on the differential-dynamic-programming approach [79]. First, iLQR uses an initial guess for the augmented controls to obtain the augmented states with the discrete dynamics function. iLQR then constructs quadratic models for each cost function using their zeroth-, first- and second-order derivatvies in a Taylor expansion about the current augmented controls and augmented states. These models are used to derive a recurrence relation between knot points which gives the locally optimal update for the augmented controls. Finally, a line search [80] is performed in the direction of this update to ensure a decrease in the sum of the cost functions. This procedure is repeated until convergence.

Indirect solvers such as iLQR are popular because they are very computationally efficient and maintain high accuracy for the discrete dynamics throughout optimization. However, standard implementations have no ability to handle nonlinear equality and inequality constraints (2c)-(2d). Projected gradient methods are a typical approach to handle constraints [81–84]. Unfortunately, within the indirect framework, they can only be used for constraints on the augmented controls, not the augmented states. Another technique, which is popular for QOC [13], is to add the constraint functions to the objective (2a). This strategy does not guarantee that the constraints are satisfied as the solver trades minimization of the cost functions and constraint functions against each other. ALM remedies this issue by adaptively adjusting a Lagrange multiplier estimate for each constraint function to ensure the constraints are satisfied. ALM adds terms that are linear and quadratic in the constraint functions to the objective. The new objective is then minimized with iLQR. If the solution produced by iLQR does not satisfy the constraints, the prefactors for the constraint terms in the objective are increased intelligently and the procedure is repeated.

ALM converges superlinearly, but poor numerical conditioning may lead to small decreases in the constraint violations near the locally optimal solution [85]. To address this shortcoming, ALTRO projects the solution from the ALM stage onto the constraint manifold using a projected Newton method, achieving ultra-low constraint violations $\sim 10^{-8}$. For more information on the details of the ALTRO solver, see [58, 86].

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

## III. QOC FOR THE FLUXONIUM

In the following, we optimize quantum gates (unitary transformations) for the fluxonium qubit. The fluxonium is a promising building block for quantum computers, and the accurate two-level approximation of its Hamiltonian makes QOC on a classical computer inexpensive. To high accuracy, we approximate the Hamiltonian near the flux frustration as a spin-1/2 system:

$$H/h = f_q \frac{\sigma_z}{2} + a(t)\frac{\sigma_x}{2}, \tag{3}$$

where $f_q$ is the qubit frequency at the flux frustration point, $a(t)$ is the flux bias, $h$ is Planck's constant, and $\sigma_z, \sigma_x$ are Pauli matrices. Although we use an approximation for the Hamiltonian, our noise model is realistic and considers the full Hilbert space. We optimize $X/2$, $Y/2$, and $Z/2$ gates for the fluxonium presented in [57], and compare them to the analytically constructed gates for that device.

First, we outline the constraints for the fluxonium gate problem. We formulate this problem as a multi-state-transfer problem. The initial conditions on the states are $\left|\psi_1^0\right\rangle = |0\rangle$, $\left|\psi_1^1\right\rangle = |1\rangle$ (6c) where the superscript is an index $i \in \{0, 1\}$, and the subscript indicates the first knot point $k = 1$. The states at the final knot point are constrained to be the image of the initial states under the desired gate $U$, $\left|\psi_N^i\right\rangle = U\left|\psi_1^i\right\rangle\ \forall\ i$ (6d). Furthermore, we impose the normalization constraint $\left|\left\langle\psi_k^i\middle|\psi_k^i\right\rangle\right|^2 = 1\ \forall\ i, k$ (6e) to ensure the solver does not take advantage of discretization errors in numerical integration. To refer to the discrete moments the flux, we introduce the notation $\int_{t_1}^{t_k} a(t)\ \mathrm{d}t \equiv \int_t a_k$, $a(t = t_k) \equiv a_k$, $\partial^n a(t)/\partial t^n \mid_{t=t_k} \equiv \partial_t^n a_k$. We impose the zero net flux constraint $\int_t a_N = 0$ (6f) which mitigates the inductive drift ubiquitous in flux-bias lines [57, 87, 88]. The flux is constrained by $|a_k| \leq 0.5$ GHz $\forall\ k$ (6g) to ensure the two-level approximation (3) remains valid. We also enforce the boundary condition $a_1 = a_N = 0$ (6h) so the gates may be concatenated arbitrarily. Additionally, we have the initial condition $\int_t a_1 = \partial_t a_1 = 0$ (6i).

Next, we introduce the augmented control and augmented state:

$$u_k = \begin{bmatrix}\partial_t^2 a_k\end{bmatrix}, \quad x_k = \begin{bmatrix}\left|\psi_k^0\right\rangle \\ \left|\psi_k^1\right\rangle \\ \int_t a_k \\ a_k \\ \partial_t a_k\end{bmatrix}. \tag{4}$$

The variables in the augmented state are derived from the decision variable in the augmented control via coupled, first-order differential equations in the the discrete dynamics function (6b). We integrate the states according to the TDSE (1) and the fluxonium Hamiltonian (3) and integrate the second-derivative of the flux amplitude to obtain the first-derivative, proportional, and integral moments. The ALTRO implementation we use does not currently support complex numbers, so we repesent the states in the isomorphism $\mathcal{H}(\mathbb{C}^n) \cong \mathcal{H}(\mathbb{R}^{2n})$ given in [13],

$$H\left|\psi\right\rangle \cong \begin{bmatrix} H_{\mathrm{re}} & -H_{\mathrm{im}} \\ H_{\mathrm{im}} & H_{\mathrm{re}}\end{bmatrix}\begin{bmatrix}\left|\psi\right\rangle_{\mathrm{re}} \\ \left|\psi\right\rangle_{\mathrm{im}}\end{bmatrix}. \tag{5}$$

The cost function at each knot point is $\ell_k(x_k, u_k) = (x_k - x_f)^T Q_k (x_k - x_f) + u_k^T R_k u_k$ where $Q_k$ and $R_k$ are diagonal matrices we supply. The $Q_k$ term penalizes deviations from the final augmented state $x_f$, which is given by the constraints we have imposed on $\left|\psi_N^i\right\rangle$, $\int_t a_N$, and $a_N$ in addition to $\partial_t a_f = 0$. The $R_k$ term penalizes the norm of $\partial_t^2 a_k$, mitigating AWG ringing due to high-frequency transitions. Stated succinctly, the optimization problem takes the form:

$$\underset{x_{1:N},\, u_{1:N\text{-}1}}{\text{minimize}} \quad \sum_{k=1}^{N} (x_k\text{-}x_f)^T Q_k (x_k\text{-}x_f) + \sum_{k=1}^{N-1} {u_k}^T R_k u_k \tag{6a}$$

$$\begin{aligned}
\text{subject to} \quad & x_{k+1} = f(x_k, u_k), && \text{(6b)}\\
& \left|\psi_1^0\right\rangle = |0\rangle, \left|\psi_1^1\right\rangle = |1\rangle, && \text{(6c)}\\
& \left|\psi_N^i\right\rangle = U\left|\psi_1^i\right\rangle\ \forall\ i, && \text{(6d)}\\
& \left|\left\langle\psi_k^i\middle|\psi_k^i\right\rangle\right|^2 = 1\ \forall\ i, k, && \text{(6e)}\\
& \textstyle\int_t a_N = 0, && \text{(6f)}\\
& |a_k| \leq 0.5\ \text{GHz}\ \forall\ k, && \text{(6g)}\\
& a_1 = a_N = 0, && \text{(6h)}\\
& \textstyle\int_t a_1 = \partial_t a_1 = 0. && \text{(6i)}
\end{aligned}$$

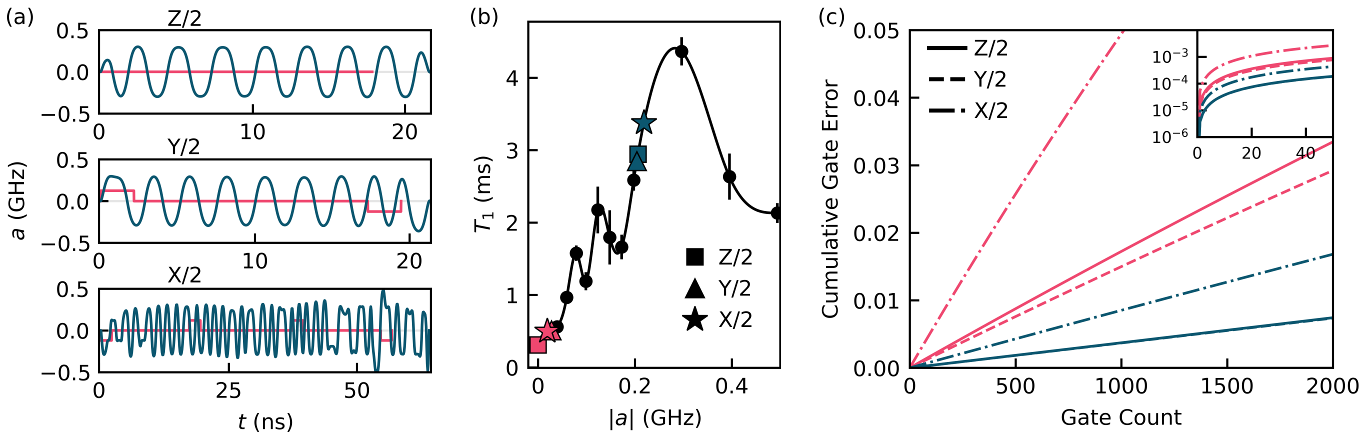


Figure 1: (a) Flux pulses for the numerical gates (dark blue) and the analytic gates (light pink). (b) $T_1$ interpolation function used in optimization. Circle markers indicate measured $T_1$ times. Non-circle markers indicate the time-averaged, absolute amplitude of each flux pulse. (c) Cumulative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

## IV. DEPOLARIZATION MITIGATION

In this section, we outline a method for optimizing the flux to mitigate depolarization. For many superconducting circuits, the $1/e$ depolarization time ($T_1$) is independent of the control parameters, so the fastest possible gate incurs the least depolarization error [89]. For the fluxonium, however, $T_1$ is strongly dependent on the flux. We enable the optimizer to trade longer gate times for lower depolarizaiton times or vice-versa by making the gate time a decision variable. Additionally, previous work has modeled the gate error due to depolarization by evolving density matrices under a master equation [31, 89] or evolving a large number of states in a quantum trajectory approach [90]. We avoid the increase in computational complexity required for these techniques by penalizing the integrated rate (probability) of depolarization in optimization. Using this probability as a proxy for the gate error incurred is reasonable because depolarization errors are incoherent and therefore increase monotonically in time without interference.

The depolarization probability is given by,

$$P_1(t) = \int_0^t T_1^{-1}(a(t'))dt'. \tag{7}$$

This value is appended to the augmented state (4) and its norm is penalized in the objective (6a) by setting the corresponding element of the final augmented state to zero. $T_1(a_k)$ 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, experimental data often deviates from theoretical predictions because it is difficult to account for all loss channels [57], and $T_1$ values often fluctuate on hour time scales [91].

We allow the optimizer to tune the gate time by making the time step between each knot point $\Delta t_k$ a decision variable. Promoting $\Delta t_k$ 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 (4) and the squared root of the time step $|\Delta t_k|$ is used for numerical integration in the discrete dynamics function (6b). To ensure numerical integration accuracy is maintained, we constrain the bounds of the time step at each knot point.

We analyze the effect of depolarization on the $X/2$, $Y/2$, and $Z/2$ gates obtained with our numerical method and the analytic gates. We use the Lindblad master equation to simulate $T_1$ dissipation for successive gate applications, and compute the cumulative gate error after each application, see Appendix A. TODO: Jens, Danny, Ziwen, any idea what the correct way to simulate depolarization is when $T_1$ is time dependent?. 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 flux pulses for the numerical gates are approximately periodic with amplitudes $\sim 0.2$GHz, see Figure 1a. They are reminiscent of the analytically determined Floquet operations for a fluxonium described in [92] and realized in [93]. The numerical gate times are greater than the analytic gate times, but the numerical flux pulses spend more time at higher amplitudes, achieving higher $T_1$ times. The single gate errors for both the analytic and numerical gates are less than $10^{-4}$, which makes them sufficient for quantum error correction–a prerequisite for fault-tolerant quantum computing [94–96]. However, the numerical gates achieve single gate errors $\sim 5$ times less than those for the analytic gates, which tracks closely with their relative improvement on the depolarization probability metric, see Appendix A. This single gate error advantage corresponds to a significant reduction in error correction resources. Furthermore, for successive gate applications, the gate error due to depolarization is approximately linear in the gate count, which we expect for $t \ll T_1$, see Figure 1c. The gate error reduction for large gate counts is important for noisy, intermediate-scale quantum (NISQ) applications. These improvements are significant for the constraints we have imposed on the gates, and do not represent a fundamental limit to the optimization methods we have employed.

## V. ROBUSTNESS TO STATIC PARAMETER UNCERTAINTY

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

The sampling method incentivizes the optimizer to ensure multiple instances of a state, each of which evolves with a distinct value of the uncertain parameter, achieve the same target state. Variants of this technique have been proposed in the context of QOC [3, 31, 48, 50, 52] and applied experimentally [49]. For each initial state, we add two sample states $|\psi^{\pm}\rangle$ to the augmented state (4). The discrete dynamics function (6b) is modified so the sample states evolve under the fluxonium Hamiltonian (3) with $f_q \to f_q \pm \sigma_{f_q}$ for a fixed hyperparameter $\sigma_{f_q}$ which is the standard deviation of the qubit frequency. We penalize the infidelities of the sample states and their target state by adding a cost function to the objective (6a) of the form $\sum_{k,\pm} q_k(1 - |\langle\psi_k^{\pm}|U\psi_1\rangle|^2)$ where $q_k$ is a constant we supply. For this method, the standard orthonormal basis states are an insufficient choice for the initial states. As an example, a $Z/2$ gate achieved by idling at the flux frustration point $a_k = 0 \; \forall \; k$ will be robust to qubit frequency detunings for the initial states $|0\rangle$ or $|1\rangle$ because the infidelity metric is insensitive to global phases, but this gate will not be robust for any other initial states. Therefore, we choose four initial states so that their outer products span the operators on the Hilbert space $\{|0\rangle, |1\rangle, (|0\rangle + i\,|1\rangle)/\sqrt{2}, (|0\rangle - |1\rangle)/\sqrt{2}\}$ [97], which we refer to as the operator basis.

Whereas the sampling method penalizes the deviations of the sample states from the target state, the unscented sampling method penalizes the deviations of the sample states from the nominal state [53, 55, 56]. Accordingly, the cost function we add to the objective (6a) takes the form $\sum_{k,j} q_k(\psi_k^j - \psi_k)^T(\psi_k^j - \psi_k)$, where $q_k$ is a constant we supply, $\psi_k$ is the evolved initial state (nominal state), and $\psi_k^j$ is a sample state that evolves under a modified Hamiltonian similar to that in the sampling method. We omit bra-ket notation here to emphasize that the states are real vectors, and are given by the right-hand-side of the isomorphism (5). Additionally, the unscented sampling method employs the unscented transform [54, 98] to prevent the sample states from drifting too far or close to the nominal state during evolution, which would result in exceedingly large or small deviations in the cost function. In particular, the sample states are chosen to encode a unimodal distribution over the $2n$ elements of the nominal state, modeling the uncertainty in the state as a result of the uncertainty in the parameter. The unscented transform accurately propagates the mean and covariance of this distribution between knot points, or equivalently, through the transformation of the TDSE. For this method, we sample on one initial state $(|0\rangle + i\,|1\rangle)/\sqrt{2}$, and do not observe a performance increase using more initial states, for example those from the operator basis. A detailed procedure for the unscented transformation is given in Appendix B.

The derivative method penalizes the sensitivity of the state to the uncertain parameter, which is encoded in the $l^{\text{th}}$-order state derivative $\partial_{f_q}^l |\psi\rangle$. In the $m^{\text{th}}$-order derivative method, we append all state derivatives of order $1, \ldots, m$ to the augmented state vector (4) for each initial state. We use the initial state $|0\rangle$ for this method, and observe no advantage to using more initial states. We penalize the norms of the state derivatives in the objective (6a) by setting the corresponding elements of the final augmented state to zero. We could obtain the state derivatives at each knot point with backward-mode differentiation. In a naive automatic differentiation scheme, the discrete dynamics function at knot points $1, \ldots, k-1$ would be differentiated to obtain the state derivative at knot point $k$, requiring $O(N^2)$ matrix multiplications. Instead, we employ forward-mode differentiation on the TDSE (1) to obtain coupled, first-order ODEs which require $O(N)$ matrix multiplications to integrate. For example, the dynamics for the $1^{\text{st}}$-order derivative method are:

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

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

We integrate the coupled ODEs with exponential integrators, see Appendix C, in the discrete dynamics function (6b). For runtimes and complexity analyses for the three robust control techniques, consult Appendix D.

We examine the gate errors due to a static qubit fre-

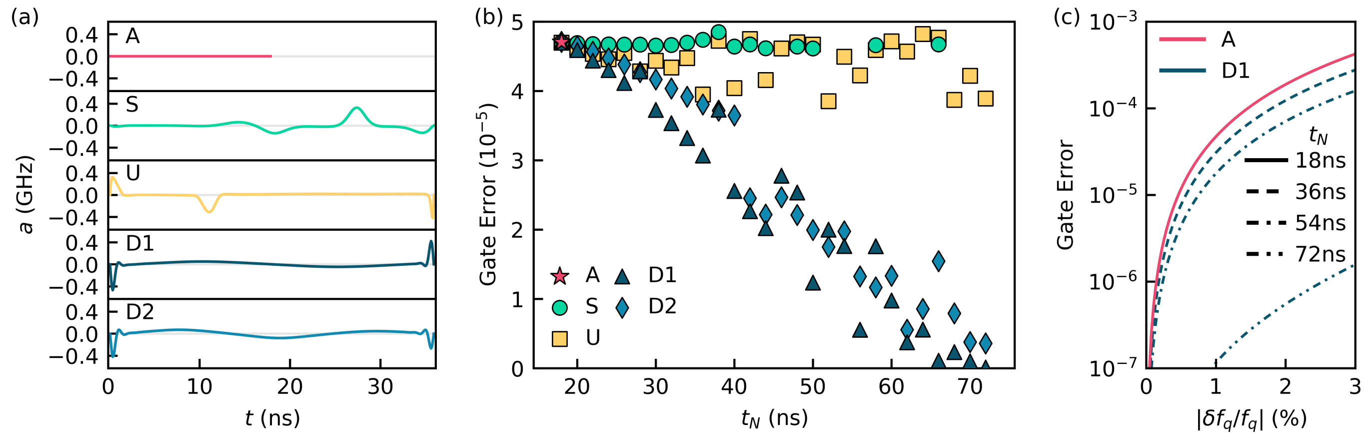


Figure 2: (a) Flux pulses for $Z/2$ gates robust to qubit frequency detunings constructed with the analytic (A), sampling (S), unscented sampling (U), and the $1^{\text{st}}$- and $2^{\text{nd}}$-order derivative methods (D1, D2). The flux pulses shown for the sampling, unscented sampling, and derivative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

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

The analytic gate corresponds to idling at the flux frustration point $a_k = 0\ \forall\ k$, see Figure 2a. Its gate time $1/4f_q \sim 18$ns is the shortest possible for a $Z/2$ gate on the device. The gate's erroneous rotation angle $2\pi\delta f_q/4f_q$ is linear in the qubit frequency detuning, resulting in a gate error that is quadratic in the detuning. At a one-percent detuning ($|\delta f_q|/f_q = 1\%$), the gate error is $\sim 4.5 \cdot 10^{-5}$, which is sufficient for quantum error correction. Although the gate performs well, it cannot be extended to gate times other than $1/4f_q$. The ability to perform phase gates in any given time is critical for multi-qubit experiments, where the qubits operate at different frequencies $f_{q,i} \neq f_{q,j}$. We can find solutions using the numerical methods at all gate times above 18ns, see Figure 2b. These numerical methods offer an effective scheme for synchronizing multi-qubit experiments.

The flux pulse produced by the sampling method transitions smoothly away from idles at the flux frustration point, whereas that for the unscented sampling method employs three distinct transitions. The gate error at a one-percent qubit frequency detuning for the sampling method does not improve substantially over the range of gate times. Conversely, the gate error at a one-percent detuning for the unscented sampling method reaches a minimum of $\sim 3.9 \cdot 10^{-5}$ near fractions of the Larmor period $2/4f_q \sim 36$ns, $3/4f_q \sim 54$ns, $4/4f_q \sim 72$ns.

The two derivative methods converge on qualitatively similar flux pulses that use fast triangle pulses at the boundaries and idle near the flux frustration point, similar to the flux pulse produced by the unscented sampling method. The gate errors at a one-percent qubit frequency detuning for both derivative methods decrease super-linearly in their gate times. The $2^{\text{nd}}$-order method does not offer a substantial improvement over the $1^{\text{st}}$-order method for most gate times. The gate error at a one-percent detuning for the $1^{\text{st}}$-order method reaches $10^{-7}$ at the Larmor period $1/f_q \sim 72$ns, see Figure 2c. This result mimics the ability of composite pulses to mitigate parameter uncertainty errors to arbitrary order with sufficiently many pulses [41]. It is difficult to choose an appropriate composite pulse for the problem studied here due to our Hamiltonian and experimental constraints. We propose comparisons between composite pulses and numerical techniques for future work.

## VI. ROBUSTNESS TO TIME-DEPENDENT PARAMETER UNCERTAINTY

An additional source of experimental error arises from time-dependent Hamiltonian parameter uncertainty. For many flux-biased and inductively-coupled superconducting circuit elements, magnetic flux noise is a significant source of coherent errors. Flux noise modifies the fluxonium Hamiltonian (3) by $a(t) \to a(t) + \delta a(t)$. The spectral density of the flux noise $\delta a(t)$ is observed to follow

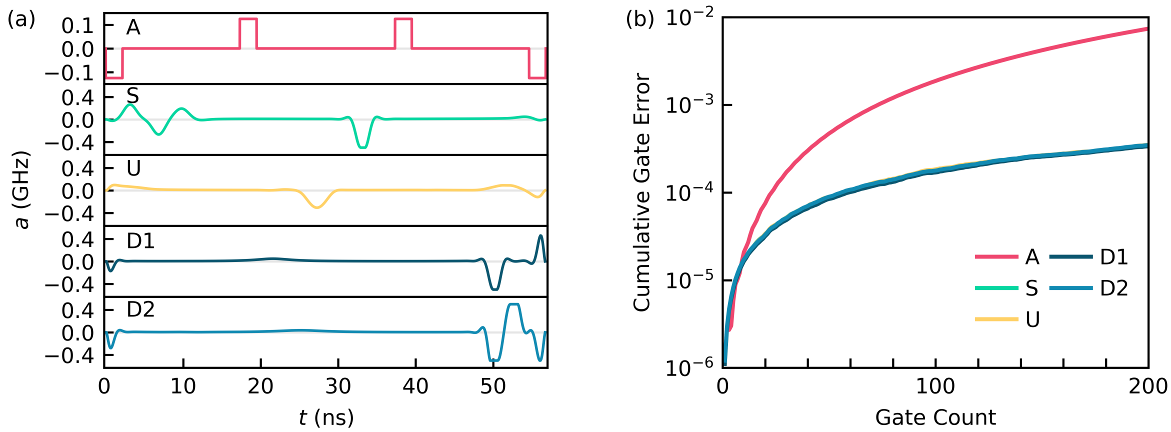


Figure 3: (a) Flux pulses for $X/2$ gates robust to flux offsets constructed with the analytic (A), sampling (S), unscented sampling (U), and the 1$^{\text{st}}$- and 2$^{\text{nd}}$-order derivative methods (D1, D2). (b) Cumulative gate error due to $1/f$ flux noise for successive gate applications. The cumulative gate errors for the sampling, unscented sampling, and the derivative methods are indistinguishable.

a $1/f$ distribution [99–102], so the noise is dominated by low-frequency components. The analytic gate considered here takes advantage of the low-frequency characteristic and treats the noise as quasi-static, performing a generalization of the spin-echo technique to compensate for erroneous drift [103, 104].

Additionally, we modify the robust control techniques presented in the previous section to combat $1/f$ flux noise. The unscented sampling method is modified so that the sample states are subject to $1/f$ flux noise. The noise is generated by filtering white noise sampled from a standard normal distribution with a finite impulse response filter [105]. The noise is then scaled by the flux noise amplitude of our device $A_\Phi = 5.21\mu\Phi_0 \implies \sigma_a = 2.5 \cdot 10^{-5}\text{GHz}$. In principle, we could modify the sampling method similarly; however, we choose to subject the sample states to static noise $a(t) \to a(t) \pm \sigma_a$ for comparison. The derivative methods require no algorithmic modification from the static case, but the TDSE is now differentiated with respect to $a(t)$ instead of $f_q$ as in (9).

We analyze the gate errors due to $1/f$ flux noise for the $X/2$ gates constructed with the robust control techniques and the analytic $X/2$ gate. To compute the gate error, an initial state is evolved under the fluxonium Hamiltonian (3) where the optimized flux is modified $a(t) \to a(t) + \delta a(t)$. The flux noise is generated as we described for the unscented sampling method. The reported gate error is the infidelity averaged over 1000 pseudorandomly generated initial states, each of which is subject to a distinct pseudorandomly generated flux noise instance. To observe the effect of interfering coherent errors, we simulate successive applications of the gate constructed by each method and compute the cumulative gate error after each application, see Figure 3. Both the analytic and numerical gates yield single gate errors sufficient for quantum error correction. Despite converging on qualitatively different solutions, the numerical gates perform similarly in the concatenated gate application comparison. Their gate errors after 200 gate applications $\sim 11\mu$s are two orders of magnitude less than the gate error produced by the analytic gate. $1/f$ flux noise is a significant source of coherent errors in NISQ applications and these numerical techniques offer effective avenues to mitigate it.

## VII. CONCLUSION

In conclusion, we have applied state-of-the-art trajectory optimization techniques to mitigate decoherence and achieve robustness to parameter uncertainty errors on a quantum system. We have proposed a scheme for suppressing depolarization with time-optimal control and the depolarization probability model. The computational cost of this model is independent of the dimension of the Hilbert space, enabling inexpensive optimization on high-dimensional quantum systems. We have also proposed the derivative method for robust control which achieves super-linear gate error reductions in the gate time for the static parameter uncertainty problem we studied. We have shown that the derivative, sampling, and unscented sampling methods can mitigate $1/f$ flux noise errors–which dominate coherent errors for flux controlled qubits. These robust control techniques can be applied to any Hamiltonian, allowing experimentalists in all domains to engineer robust operations on their quantum systems. These methods will be used to achieve the low gate errors required for fault-tolerant quantum computing applications. Our implementations of the techniques described in this work are available at `https://github.com/SchusterLab/rbqoc`. TODO:

static code?

## ACKNOWLEDGMENTS

We thank Helin Zhang for experimental assistance and Taylor Howell, Jens Koch, Tanay Roy, Colm Ryan, and Daniel Weiss for useful discussions. This work is funded in part by EPiQC, an NSF Expedition in Computing, under grant CCF-1730449. This work was supported by the Army Research Office under Grant No. W911NF1910016. TODO: funding. TODO: cite software packages.

## Appendix A: Depolarization

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

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

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

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

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

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 depolarization $\gamma_\uparrow = T_{1,\uparrow}^{-1}$, $\gamma_\downarrow = T_{1,\downarrow}^{-1}$, $L_\uparrow = \sigma^+/2$, and $L_\downarrow = \sigma^-/2$ where $\sigma^\pm = \sigma_x \pm i\sigma_y$. Both $T_{1,\uparrow}$ and $T_{1,\downarrow}$ are obtained from the spline shown in Figure 1b. We obtain the $T_1$ values in this spline by driving the qubit at the desired flux amplitude and monitoring the resultant decay. For more details on these measurements, consult [57].

So that we may use exponential integrators, we employ the Vectorization/Choi-Jamiolkowski isomorphism [106],

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

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

where $\rho = \sum_{i,j} \alpha_{i,j} |i\rangle\langle j|$ and $\mathrm{vec}(\rho) = \sum_{i,j}\alpha_{i,j}|i\rangle \otimes |j\rangle$. We use zero-order hold on the controls so the exact solution is $\mathrm{vec}(\rho_{k+1}) = \exp(\Delta t_k \hat{\mathcal{L}}_k)\mathrm{vec}(\rho_k)$. 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 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. In the unscented sampling method, we consider a state $\psi \in \mathbb{R}^{2n}$, an uncertain parameter $\lambda \in \mathbb{R}^d$, and discrete dynamics $\psi_{k+1} = f(\psi_k, \lambda_k)$. The state is obtained with an isomorphism (5) and we omit bra-ket notation to emphasize that the state is real. The nominal initial state is given by $\bar{\psi}_1$ with an associated positive-definite covariance matrix $P_1 \in \mathbb{S}_{++}^{2n}$ which describes the uncertainty in the initial state. $P_1$ is typically non-zero even if the state preparation error is negligible. The uncertain parameter has zero-mean and its distribution is given by the covariance matrix $L_k \in \mathbb{S}_{++}^{d}$ at 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 function $f(\psi_k, \lambda_k)$.

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

$$\begin{bmatrix} \psi_1^j \\ \lambda_1^j \end{bmatrix} = \begin{bmatrix} \bar{\psi}_1 \\ 0 \end{bmatrix} \pm \beta \sqrt{\begin{bmatrix} P_1 & 0 \\ 0 & L_1 \end{bmatrix}}^{\,j}. \tag{B1}$$

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

$$\psi_1^j \rightarrow \frac{\psi_1^j}{\sqrt{{\psi_1^j}^T \psi_1^j}}. \tag{B2}$$

The sample states are propagated to the next knot point,

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

The mean and covariance of the sample states are computed,

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

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

The sample states are then resampled and propagated to the next knot point using (B1), (B2), and (B3). Our choice of sample states (sigma points) follows equation 11 of [54]. Prescriptions that require fewer sigma points exist [107].

### Appendix C: Derivative Method

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

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

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 [66],

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

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 depolarization optimization in Section IV, and the robust optimizations in Section V are presented in Table II for a $Z/2$ gate at gate times which are multiples of $1/4f_q \sim 18$ns. We performed optimizations on a single core of an AMD Ryzen Threadripper 3970X 32-Core Processor. Future work will parallelize the robustness methods using GPUs [13], which will enable fast optimizations on high-dimensional Hilbert spaces.

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

Table II: Average runtimes for $Z/2$ optimizations using the base, depolarization, sampling (S), unscented Sampling (U), and the $1^{\text{st}}$- and $2^{\text{nd}}$-order derivative methods (D1, D2).

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 uncertain 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 $m^{\text{th}}$-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.emrstm1043.

[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. 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).

[13] 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).

[14] 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.

[15] 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).

[16] 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).

[17] 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).

[18] 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).

[19] 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).

[20] 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).

[21] 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).

[22] 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).

[23] 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).

[24] 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).

[25] 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).

[26] 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).

[27] 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).

[28] 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).

[29] 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).

[30] F. Poggiali, P. Cappellaro, and N. Fabbri, Optimal control for one-qubit quantum sensing, Phys. Rev. X **8**, 021059 (2018).

[31] 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).

[32] 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).

[33] 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).

[34] 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).

[35] K. Zhou, *Essentials of Robust Control*, 1st ed. (Pearson, 1997).

[36] J. Morimoto and C. Atkeson, Minimax differential dynamic programming: An application to robust biped walking, Advances in neural information processing systems **15**, 1563 (2002).

[37] Z. Manchester and S. Kuindersma, Robust Direct Trajectory Optimization Using Approximate Invariant Funnels, Autonomous Robots 10.1007/s10514-018-9779-5 (2018).

[38] 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).

[39] 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).

[40] Ä. Kupce and R. Freeman, Stretched adiabatic pulses for broadband spin inversion, Journal of Magnetic Resonance, Series A **117**, 246 (1995).

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

[42] 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).

[43] 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).

[44] 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).

[45] D. J. Egger and F. K. Wilhelm, Adaptive hybrid optimal quantum control for imprecisely characterized systems, Phys. Rev. Lett. **112**, 240503 (2014).

[46] 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).

[47] 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].

[48] J. Allen, *Robust Optimal Control of the Cross-Resonance Gate in Superconducting Qubits*, Ph.D. thesis, University of Surrey (2019).

[49] 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).

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

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

[52] H. Ball, M. J. Biercuk, A. Carvalho, R. Chakravorty, J. Chen, L. A. de Castro, S. Gore, D. Hover, M. Hush, P. J. Liebermann, *et al.*, Software tools for quantum control: Improving quantum computer performance through noise and error suppression, arXiv preprint arXiv:2001.04060 (2020).

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

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

[55] 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.

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

[57] H. Zhang, S. Chakram, T. Roy, N. Earnest, Y. Lu, Z. Huang, D. K. Weiss, J. Koch, and D. I. Schuster, Universal fast-flux control of a coherent, low-frequency qubit, Phys. Rev. X **11**, 011010 (2021).

[58] 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.

[59] S. Machnes, E. Assémat, D. Tannor, and F. K. Wilhelm, Tunable, flexible, and efficient optimization of control pulses for practical qubits, Phys. Rev. Lett. **120**, 150401 (2018).

[60] 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).

[61] 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.

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

[63] 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).

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

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

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

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

[68] R. Shillito, J. A. Gross, A. D. Paolo, Élie Genois, and A. Blais, Fast and differentiable simulation of driven quantum systems (2020), arXiv:2012.09282 [quant-ph].

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

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

[71] J. T. Betts, Survey of numerical methods for trajectory optimization, Journal of guidance, control, and dynamics **21**, 193 (1998).

[72] P. E. Gill, W. Murray, and M. A. Saunders, Snopt: An sqp algorithm for large-scale constrained optimization, SIAM review **47**, 99 (2005).

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

[74] 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).

[75] G. Lantoine and R. P. Russell, A hybrid differential dynamic programming algorithm for constrained optimal control problems. part 1: Theory, Journal of Optimization Theory and Applications **154**, 382 (2012).

[76] B. Plancher, Z. Manchester, and S. Kuindersma, Constrained unscented dynamic programming, in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (IEEE, 2017) pp. 5674–5680.

[77] D. P. Bertsekas, Projected newton methods for optimization problems with simple constraints, SIAM Journal on control and Optimization **20**, 221 (1982).

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

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

[80] L. Zhang, W. Zhou, and D. Li, Global convergence of a modified fletcher–reeves conjugate gradient method with armijo-type line search, Numerische Mathematik **104**, 561 (2006).

[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] O. V. Morzhin and A. N. Pechen, Minimal time generation of density matrices for a two-level quantum system driven by coherent and incoherent controls, International Journal of Theoretical Physics 10.1007/s10773-019-04149-w (2019).

[84] M. S. Nikol'skii, Convergence of the gradient projection method in optimal control problems, Computational Mathematics and Modeling **18**, 148 (2007).

[85] D. Bertsekas, Constrained optimization and lagrange multiplier methods (second edn.) athena scientific: Belmont (1996).

[86] 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.

[87] 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).

[88] 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).

[89] T. Schulte-HerbrÃŒggen, A. SpÃ¶rl, N. Khaneja, and S. J. Glaser, Optimal control for generating quantum gates in open dissipative systems, Journal of Physics B: Atomic, Molecular and Optical Physics **44**, 154013 (2011).

[90] 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).

[91] 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).

[92] 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).

[93] 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).

[94] 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.

[95] E. Knill, Quantum computing with realistically noisy devices, Nature **434**, 39–44 (2005).

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

[97] 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).

[98] J. K. Uhlmann, *Dynamic map building and localization: New theoretical foundations*, Ph.D. thesis, University of Oxford Oxford (1995).

[99] 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).

[100] 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).

[101] 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).

[102] 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).

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

[104] S. Meiboom and D. Gill, Modified spin-echo method for measuring nuclear relaxation times, Review of scientific instruments **29**, 688 (1958).

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

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

[107] 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.