# Robust Quantum Optimal Control

Thomas Propson*
*Department of Physics, University of Chicago, Chicago, Illinois 60637, USA*
(Dated: July 23, 2020)

This is a paper about robust quantum optimal control.

## I. INTRODUCTION

(Existing work) The leading model of universal quantum computation is gate-based. The Quantum Optimal Control (QOC) literature has saught to construct techniques for engineering waveforms to manipulate quantum systems with high fidelity. Analytic methods have been demonstrated to construct quantum gates [1–5]. Most methods rely on solving the time-dependent Schroedinger equation analytically or deriving equations of motion from a suitable Lagrangian.

Some analytical methods have been presented to address realistic experimental constraints. Considering the dynamical and geometric phases of state evolution has led to methods for achieving robustness to pure dephasing in quantum systems [3, 4, 6]. Additionally, suitable choices of bases have been demonstrated that allow the experimentalist to infer tradeoffs in longitudinal and pure dephasing coherence times [2]. To date, however, no analytical method has been presented to construct a gate while satisfying all relevant experimental constraints. Furthermore, most analytic methods have only been demonstrated in the two-level approximation and become cumbersome to solve for larger Hilbert space sizes.

Numerical quantum control techniques have developed in parallel with the analytic frameworks. Numerical quantum control techniques integrate the time-dependent Schroedinger equation and use sensitivity analysis to optimize target metrics of the quantum gate, such as fidelity [7–12]. These techniques typically employ zeroth- or first-order optimizers and make no garauntees of optimality. Further, they formulate the quantum control problem as unconstrained or use projective gradient methods to enforce constraints [11]. The latter approach relies on constructing bases for constraint spaces–which is not always feasible–and may hinder convergence.

(This work) We employ the trajectory optimization literature to formulate the quantum optimal control problem as a constrained optimization problem. The trajectory optimization framework allows us to handle arbitrary constraints on the evolution of the quantum state. Additionally, we introduce methods for constructing gates robust to longitudinal relaxation and pure dephasing, the dominant barriers to experimentally realizing practical quantum computing.

We study the quantum optimal control problem on the fluxonium qubit. We outline experimentally realistic constraints and map them to the trajectory optimization framework. For the device we study we achieve a 1.5x decrease in longitudinal relaxation probability. We present two methods for achieving robustness to pure dephasing, and compare them to existing dynamic decoupling methods. We find that our methods mitigate dephasing by order X over dynamic decoupling.

(Outline) First we formulate the quantum optimal control problem in the trajectory optimization framework. Then, we introduce the dynamics of the fluxonium device and outline experimental considerations relevant to gate construction. Next we outline a method for making the optimization longitudinal relaxation aware. Finally, we present some methods for engineering robustness to pure dephasing and compare them to existing techniques.

## II. QOC + AL-ILQR

(QOC Problem Statement) Here 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 Schroedinger equation (TDSE)

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

The evolution is sometimes cast with the evolution of a density matrix under the Lindblad master equation to model the decoherence of the state explicitly. 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 integration techniques for the TDSE include approximating unitary propagators as well as explicit integration methods, such as Runge-Kutta, of the form $|\psi_{k+1}\rangle \approx |\psi_k\rangle + \frac{d}{dt}|\psi_k\rangle \cdot \Delta t_k$

Quantum optimal control seeks the control parameters that minimize a functional $J(u(t))$. In the simplest case the functional is $J = 1 - |\langle\psi_f|\psi_N(u(t))\rangle|^2$ the infidelity between the inital state evolved to the final knot point ($|\psi_N(u(t))\rangle$) and the target state ($|\psi_f\rangle$). In general $J$ is

* tcpropson@uchicago.edu

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 [7]. Typical quantum optimal control algorithms employ automatic differentiation to compute first order information for the functional ($\nabla_u J(u)$). They employ a first-order optimizer to minimize $J$ with respect to $u$.

(AL-iLQR Problem Statement) The trajectory optimization literature solves a more general class of nonlinear programs that resemble the quantum optimal control problem. The quantum optimal control problem is a specific case of the linear quadratic regulator (LQR). LQR is so called because the dynamics are linear in the state and the functional is quadratic in the state. In the LQR formulation the same functional is evaluated at each knot point

$$J_{\text{iLQR}} = \tilde{x}_N^T Q_N \tilde{x}_N + \sum_{k=0}^{N-1} \tilde{x}_k^T Q_k \tilde{x}_k + u_k^T R_k u_k \quad (2)$$

where $\tilde{x}_k = x_k - x_f$ is the difference between the state at knot point $k$ and final state, $u_k$ are the controls, and $Q_k, R_k$ are matrices that define the penalty metric. The state is propagated using a dynamics function $x_{k+1} = f(x_k, u_k, t_k, \Delta t_k)$. In the case of quantum optimal control $|\psi_k\rangle \subseteq x_k$ and $f$ encodes the TDSE dynamics. In the following we refer to $|\psi\rangle$ as the state. We refer to $x$ and $u$ as the augmented state and augmented controls, respectively.

The advantage of the LQR formulation is that there exists a dynamic programming algorithm to compute the optimal update to the augmented controls ($u_k$) which minimizes the functional ($J_{\text{iLQR},k}$) for each knot point. This algorithm proceeds by deriving a recurrence relation between knot points $k$ and $k+1$ for the optimal feedback law–known as the Ricatti recursion (see Appendix). The iterative LQR (iLQR) algorithm computes $J_{\text{iLQR}}$ and applies the Ricatti recursion to all knot points on multiple executions.

In order to incorporate constraints we employ the augmented Lagrangian method. Constraints are contributions to the functional of arbitrary form $c_k(x_k, u_k)$ which are zero or negative when the constraint is satisfied. The AL-iLQR method associates a penalty multiplier with the functional that estimates the constraint's Lagrange multiplier. The algorithm updates the penalty multiplier between iLQR executions. In this scheme the functional takes the form

$$\begin{aligned} J_{\text{AL-iLQR}} = (\lambda_k + \frac{1}{2} I_{\mu_k} c_k(x_k, u_k))^T c_k(x_k, u_k) \\ + J_{\text{iLQR}} \end{aligned} \quad (3)$$

where $\lambda_k$ is a Lagrange multiplier and $I_{\mu_k}$ is a penalty matrix with $\mu_k$ along the diagonal. $\lambda_k$ and $\mu_k$ are updated after each augmented Lagrangian iteration according to

$$\lambda_{k_i} \leftarrow \max(0, \lambda_{k_i} + \mu_{k_i} c_{k_i}(x_k^*, u_k^*)) \quad (4)$$

$$\mu_{k_i} \leftarrow \phi \mu_{k_i} \quad (5)$$

where $x^*, u^*$ are the optimal augmented state and augmented controls from the iLQR execution, $i$ indicates the $i$-th constraint functional, and $\phi$ is a hyperparameter. With this updated form of the cost functional there still exists a recurrence relation to calculate the optimal control updates, see [13].

## III. QOC ON THE FLUXONIUM

(Fluxonium Device) In the following we study the quantum optimal control problem on the fluxonium qubit. In the two-level approximation the system Hamiltonian takes the form

$$H/h = \omega_q \frac{\sigma_z}{2} + a(t) \frac{\sigma_x}{2} \quad (6)$$

where $\omega_q$ is the qubit frequency at the flux frustration point, $a$ is the flux drive amplitude, $h$ is Planck's constant, and $\sigma_x, \sigma_y$ are Pauli matrices. The flux amplitude $a$ is experimentally realized by modulating the flux threading the device ($\Phi_{\text{ext}}$). The flux amplitude is obtained from the external flux by $a = 4\pi \langle g|\hat{\phi}|e\rangle|_{0.5\Phi_0} E_L \delta\Phi_{\text{ext}}/(h\Phi_0)$ where $\hat{\phi}$ is the phase operator, $E_L$ is the characteristic inductance energy, $\Phi_0$ is a flux quantum, and $\delta\Phi_{\text{ext}} = \Phi_{\text{ext}} - 0.5\Phi_0$ is the flux offset from the flux frustration point.

In the following we compare the gates found with our methods to the $X/2, Y/2, Z/2$ gates reported in [1] for the same device. Note that $Y/2$ and arbitrary $Z$ rotations are sufficient for universal computation.

(Constraints) We now outline the constraints that we require each gate to obey. We require $a(t = 0) = a(t = t_N) = 0$. This constraint ensures gates may be concatenated arbitrarily without inducing AWG ringing due to high-frequency transitions. 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. We require $-300\text{MHz} \leq a(t) \leq 300\text{MHz}$ to ensure the two-level approximation III remains valid. Additionally, we require that each gate achieves the desired state transition $\langle \psi_f | \psi_N \rangle = 1$. In addition to these constraints we penalize the norm of the first and second derivatives of the flux amplitude to mitigate AWG ringing.

The optimization is performed over the second derivative of the flux amplitude ($\frac{d^2}{dt^2} a(t)$) which is contained in the augmented control vector. The first derivative ($\frac{d}{dt} a(t)$), proportional ($a(t)$), and integral ($\int a(t)$) flux amplitude terms are contained in the augmented state vector. They are obtained from the second derivative of the flux amplitude by integration in the dynamics function $f$. Both the zero net flux and target quantum state constraint are then handled by ensuring the augmented

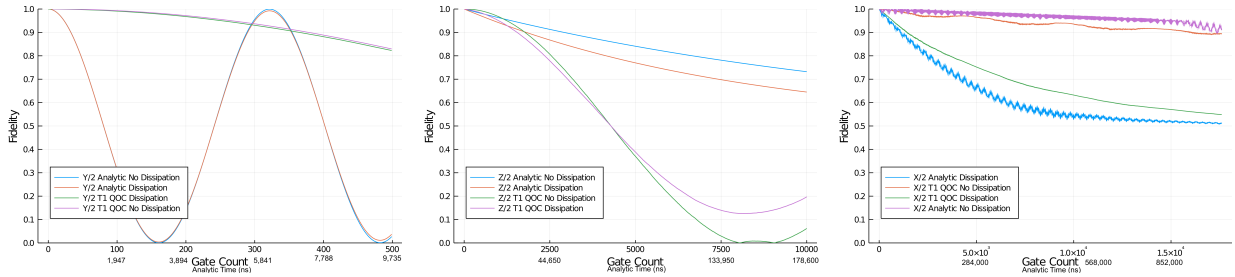


Figure 1: Master equation simulation with $T_1$ dissipation for all basis gates, comparing the performance of the analytic gates and the $T_1$ optimized QOC gates.

final state is reached $x_N - x_f = 0$. The equality and inequality constraints on $a(t)$ are handled with a bound constraint.

($T_1$ and $T_\phi$ noise) In addition to optimizing the cost functional to achieve a gate that obeys experimental constraints and has a high simulated fidelity, we also want to make the gate robust to noise that affects the experimental gate fidelity. 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 $T_\phi$ respectively (see Appendix). 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. There is a tradeoff between the two decoherence processes. In the case of white noise we have that the sum of the noise weights $W_1$ and $W_\phi$ is constant [2]. Our device achieves its best pure dephasing protection at the flux frustration point $T_{2e}(\Phi_{\text{ext}} = 0.5\Phi_0) \sim 300\mu$s where the qubit frequency 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. Conversely, $T_1$ is at a minimum at the flux frustration point $T_1(0.5\Phi_0) = 0.315$ms, and increases away from the flux frustration point $T_1(0.43\Phi_0) = 4.3$ms. Given the nature of the decay processes and the tradeoff, we choose to maximize the longitudinal relaxation time ($T_1$) over the gate duration and employ robust control techniques to mitigate pure dephasing.

## IV. ROBUSTNESS TO $T_1$-TYPE NOISE

(Strategy) We seek to minimize the probability that the qubit decays as a result of longitudinal relaxation. To this end we add the longitudinal relaxation probability to the augmented state vector

$$P_1(t) = \int_0^t \gamma_1(a(t'))dt' \tag{7}$$

where $\gamma_1 = T_1^{-1}$. Setting the target longitudinal relaxation probability to 0 results in a quadratic cost at each knot point of the form $|P_1(t_k)|^2$.

$\gamma_1(a_k)$ is obtained at each knot point by evaluating a spline interpolant fit to experimentally obtained data of the form $\{(\Phi_{\text{ext}}, T_1)\}$. $\Phi_{\text{ext}}$ is a function of $a_k$ with the inverse of the relation given in section 3. Calculating $T_1$ directly from theoretical considerations requires many high-dimensional eigendecompositions, which is computationally expensive. Additionally, $T_1$ values are known to fluctuate greatly with laboratory temperatures [14]. Interpolating $T_1$ from experimental data increases the fridge truth of the simulation.

Furthermore, the probability of longitudinal relaxation is dependent on the gate duration $t_N$. We allow the optimizer to tune the gate duration by adding the time step between knot points $\Delta t_k$ to the augmented control vector $u_k$. Promoting $\Delta t_k$ 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-3 ns $\leq \Delta t_k \leq$ 2e-2 ns. Note that this bound constraint is allowed to be broken for intermediate iterations of the optimization, so we use the absolute value of the time step $|\Delta t_k|$ to ensure that it is non-negative.

(Results) We achieve a factor of 1.5 decrease in the probability of longitudinal relaxation from the analytic gates we benchmark against. We simulate the performance of the gates using the Lindblad master equation (see Appendix). We repeatedly apply the basis gates and measure the fidelity of the resulting state as a function of time as shown in Fig. 1.

| Gate | Analytic $P_1$ $(10^{-5})$ | QOC $P_1$ $(10^{-5})$ | Speedup |
|---|---|---|---|
| Z/2 | 4.615 | 2.576 | 1.792 |
| Y/2 | 2.826 | 1.888 | 1.496 |
| X/2 | 9.562 | 5.080 | 1.882 |

Table I: Probability of longitudinal relaxation for each gate evaluated at the gate's duration.

## V. ROBUSTNESS TO $T_\phi$-TYPE NOISE

(Strategy) For the fluxonium, pure dephasing acts along the flux axis. Low frequency $1/f$ noise, Johnson-Nyquist current noise present in the flux bias lines, and temperature dependent gain fluctuations may cause the flux amplitude to deviate from its desired value by an amount $\delta a$. Additionally, residual calibration errors and fluctuations may cause the qubit frequency to deviate from its measured value by an amount $\delta\omega_q$.

System parameter deviations are typically addressed with dynamic decoupling sequences [6], the DRAG scheme [15], or geometric phase considerations [4] [3]. Dynamic decoupling sequences compose rotations on the Bloch sphere so that erroneous rotations arising due to the parameter deviation are cancelled. The latter two methods are inflexible. We draw on the robust control literature to demonstrate numerical techniques for engineering robustness to parameter deviations.

We propose two methods for engineering robustness to parameter deviations. The first we call the derivative method. We draw on the intuition that making the state evolution insensitive to changes in the parameter $(x)$ is encoded in the derivative of the state with respect to that parameter $\partial_x^l |\psi\rangle$. The derviatives of the state are propgated in the augmented state vector. Their dynamics are found by differentiating the TDSE dynamics with respect to the parameter of interest (see Appendix). Setting the target derivatives of the state to zero vectors results in quadratic costs at each knot point $\left|\partial_x^l |\psi_k\rangle\right|^2$.

The second scheme we analyze is the sampling method, which is well studied in the robust control literature [16] [17]. We add additional states that follow deviant dynamics to the augmented state vector. The parameters of interest are altered by a small deviation, typically chosen to represent the distribution of the parameter. For example, in the 2-point sampling method we propagate $|\psi^{\pm,I}(t)\rangle$ where $\omega_q^\pm = \omega_q \pm \sigma_{\omega_q}$, $\omega_q^I = \omega_q$. The difference between the final states and the target state is penalized, resulting in quadratic costs at each knot point $\left||\psi_k^{\pm,I}\rangle - |\psi_f\rangle\right|^2$

(Results) We compare the robust control schemes to a dynamic decoupling pulse for a single $R_x(\pi/2)$ gate subject to qubit frequency detuning $(\omega_q \leftarrow \omega_q + \delta\omega_q)$ to demonstrate the applicability of our techniques to mitigate system parameter deviations. We use a compensating for off-resonance error (CORPSE) dynamic decoupling pulse that is optimized to mitigate first-order error arising due to the detuning. The pulse is composed of three rotations about the $\hat{x}$, $-\hat{x}$, and $\hat{x}$ axes of the Bloch sphere designed to compensate for erroneous rotations (see Appendix). We simulate the gate fidelity as a function of the qubit frequency detuning for variants of the derivative method, sample method, and CORPSE pulse (see Figure 2a). We also find that increasing the gate duration allows the optimizer to find a more robust pulse (see Figure 2b). Increasing the order of the derivative and sample methods requires propagating an additional quantum state in the augmented state vector, a linear increase in the computational complexity, but finding higher order variants of the CORPSE family is prohibitively difficult.

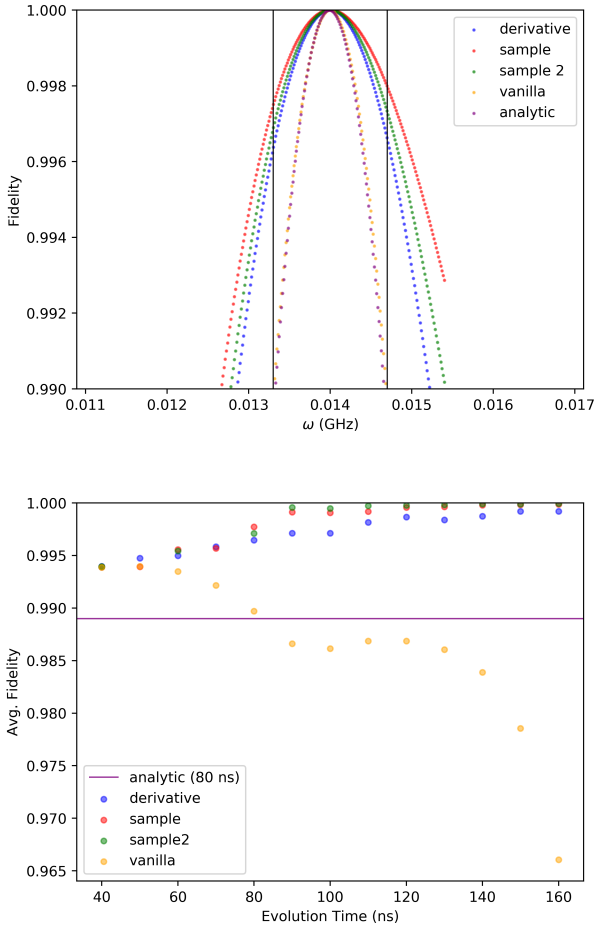


Figure 2: Robustness to qubit frequency detuning for the (2, 4)-point sampling method, ($2^{\text{nd}}$, $3^{\text{rd}}$) order derivative method, and CORPSE pulses for the $R_x(\pi/2)$ gate. (a) Gate fidelity as a function of the qubit frequency detuning. (b) Gate fidelity at one standard deviation from the nominal qubit frequency as a function of time.

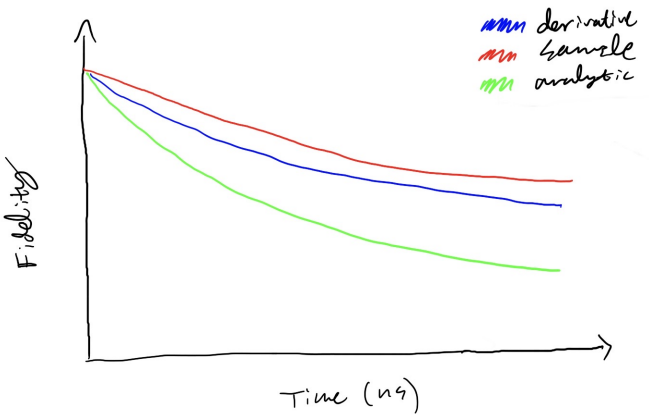


Figure 3: Robustness to $T_\phi$-type noise for the the (2, 4)-point sampling method, ($2^{\text{nd}}$, $3^{\text{rd}}$) order derivative method, and B2CORPSE pulse for the $R_x(\pi/2)$ gate.

Additionally, we compare the robust control schemes to a dynamic decoupling pulse for sequential $R_x(\pi/2)$ gates subject to $T_\phi$ noise to demonstrate the applicability of our techniques to mitigate noise of this type. We simulate the gate fidelity as a function of time for variants of the derivative method, sample method, and B2CORPSE pulse (see Figure 3). The B2CORPSE pulse is constructed to be robust to qubit frequency detuning errors and amplitude scaling errors (see Appendix). The derivative and sampling methods are employed to mitigate error due to qubit frequency detuning ($\omega_q \leftarrow \omega_q + \delta\omega_q$) as well as flux amplitude scaling error ($a \leftarrow a(1+\delta a)$).

## VI. CONCLUSION

We have proposed some schemes and they work well.

### Appendix A: Ricatti Recursion

This will give the reader unfamiliar with trajectory optimization intuition for how the trajectory optimization update scheme works and why it is better than a more naive method.

### Appendix B: Experiment

We measure $T_1$ using the standard experiment and $T_2$ using te Ramsey experiment. We fit with splines and the data looks like fig. 3 in Helin's paper [1]. We measure $\omega_q$ and $\sigma_{\omega_q}$ using X method.

### Appendix C: Dissipation Simulation

We model dissipation using the Lindblad master equation.

$$\frac{d}{dt}\rho = \frac{-i}{\hbar}[H,\rho] + \sum_{i=1}^{N^2-1} \gamma_i (L_i \rho L_i^\dagger - \frac{1}{2}\{L_i^\dagger L_i, \rho\}) \quad \text{(C1)}$$

where $\rho = |\psi\rangle\langle\psi|$ is the density matrix, $N = \dim(\mathcal{H})$, and $[\cdot,\cdot], \{\cdot,\cdot\}$ are the algebraic commutator and anti-commutator. For longitudinal relaxation $\gamma_1 = T_1^{-1} = T_{1,\uparrow}^{-1} + T_{1,\downarrow}^{-1}$ and $L_\uparrow = \sigma^+/2$, $L_\downarrow = \sigma^-/2$ are the ladder operators $\sigma^\pm = \sigma_x \pm i\sigma_y$. For pure dephasing $\gamma_2 = T_2^{-1} = (2T_1)^{-1} + T_\phi^{-1}$ and $L_2 = (I - \sigma_z)/2$.

### Appendix D: Derivative Method Dynamics

To obtain the dynamics for the derivative of the state $\partial_x^l |\psi(t)\rangle$ we differentiate the TDSE dynamics II with respect to the parameter of interest ($x$). Using the fluxonium hamiltonian III we obtain the derivative of the state with respect to the qubit frequency ($\omega_q$) and the flux amplitude ($a$). Here we present the dynamics for the second derivative of the state with respect to the qubit frequency. The case is analogous for the flux amplitude. Both $H$ and $|\psi\rangle$ are functions of $\omega_q$, $a$, and $t$, but we omit the explicit dependence in notation for brevity.

$$\begin{aligned} \partial_{\omega_q}\partial_t |\psi\rangle &= \partial_{\omega_q} H |\psi\rangle \\ &= (\partial_{\omega_q} H)|\psi\rangle + H(\partial_{\omega_q}|\psi\rangle) \\ &= \frac{\sigma_z}{2}|\psi\rangle + H(\partial_{\omega_q}|\psi\rangle) \end{aligned} \quad \text{(D1)}$$

$$\begin{aligned} \partial_{\omega_q}^2\partial_t |\psi\rangle &= \partial_{\omega_q}(\partial_{\omega_q} H |\psi\rangle) \\ &= (\partial_{\omega_q}^2 H)|\psi\rangle + 2(\partial_{\omega_q} H)(\partial_{\omega_q}|\psi\rangle) \\ &\quad + H(\partial_{\omega_q}^2|\psi\rangle) \\ &= \sigma_z(\partial_{\omega_q}|\psi\rangle) + H(\partial_{\omega_q}^2|\psi\rangle) \end{aligned} \quad \text{(D2)}$$

The augmented state vector carries $\partial_{\omega_q}^l |\psi\rangle$ which appears explicitly in its own dynamics. Due to the dependence of $H$ on $\omega_q$ and $a$, the $l^{\text{th}}$ state derivative is coupled to the $l-1^{\text{th}}$ state derivative.

### Appendix E: Dynamic Decoupling Sequences

Here we present the equations for constructing the CORPSE and B2CORPSE sequences.

### Appendix F: Complex Tensor Handling

We use an isomorphism $\mathcal{H}(\mathbb{C}^n) \cong \mathcal{H}(\mathbb{R}^{2n})$ because the software we use does not support complex numbers yet.


[1] 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).

[2] 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).

[3] 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).

[4] 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).

[5] A. Carlini, A. Hosoya, T. Koike, and Y. Okudaira, Quantum brachistochrone, arXiv preprint quant-ph/0511039 (2005).

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

[7] 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).

[8] 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).

[9] P. Doria, T. Calarco, and S. Montangero, Optimal control technique for many-body quantum dynamics, Physical review letters **106**, 190501 (2011).

[10] 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).

[11] 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).

[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] 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.

[14] 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).

[15] 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).

[16] 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.

[17] F. Tronarp, R. Hostettler, and S. Särkkä, Sigma-point filtering for nonlinear systems with non-additive heavy-tailed noise, in *2016 19th International Conference on Information Fusion (FUSION)* (IEEE, 2016) pp. 1859–1866.