Superconducting qubit gates via analytically-derived accelerated adiabatic pulses

F. Setiawan,[1,*] Peter Groszkowski,[1] Hugo Ribeiro,[2] and Aashish A. Clerk[1]

[1] *Pritzker School of Molecular Engineering, University of Chicago, 5640 South Ellis Avenue, Chicago, Illinois 60637, USA*

[2] *Max Planck Institute for the Science of Light, Staudtstraße 2, 91058 Erlangen, Germany*

(Dated: February 5, 2021)

Shortcuts to adiabaticity (STA) is a general methodology for speeding up adiabatic quantum protocols, and has many potential applications in quantum information processing. Unfortunately, analytically constructing STAs for systems having complex interactions and more than a few levels is a challenging task. This is usually overcome by assuming an idealized Hamiltonian (e.g., only a limited subset of energy levels are retained, and the rotating-wave approximation (RWA) is made). Here, we develop an *analytic* approach that allows one to go beyond these limitations. Our method is general and results in analytically-derived pulse shapes that correct both non-adiabatic errors as well as non-RWA errors. We also show that our approach can yield pulses requiring a smaller driving power than conventional non-adiabatic protocols. We show in detail how our ideas can be used to analytically design high-fidelity single-qubit "tripod" gates in a realistic superconducting fluxonium qubit.

I. INTRODUCTION

Quantum gates based on adiabatic evolution [1–7] are generally desirable because of their intrinsic robustness against imperfections in control pulses, and have been implemented in a variety of platforms (see e.g., [8–10]). They, however, require extremely long evolution times, making them potentially susceptible to dissipation and noise. An intriguing possibility is to try to "accelerate" adiabatic gates using techniques drawn from the field of shortcuts to adiabaticity (STA) [11–15]. STA protocols seek to modify pulses to completely cancel non-adiabatic errors. They are usually developed for simple evolutions that do not correspond to a true quantum gate, as their form is tied to a specific choice of initial state. However, they can be adapted for true gates. Ref. [16] showed how a particular shortcut approach (superadiabatic transitionless driving (SATD) [17]) could be used to accelerate a true (arbitrary) single-qubit gate, the paradigmatic "tripod" adiabatic gate introduced in Refs. [3, 4] (see Fig. 1). Other STA approaches to gates have been presented in Refs. [18–22]. Note that STA-accelerated gates are conceptually distinct from the so-called non-adiabatic holonomic gates (see e.g., Ref. [23]), and have distinct advantages [16].

While the above results are promising, they are limited by a constraint that plagues most STA approaches to quantum control: accelerated protocols can only be derived analytically for few-level systems (or systems that reduce to uncoupled few-level systems). Further, one needs to ignore fast-oscillating non-resonant terms (i.e., one must necessarily make the rotating-wave approximation (RWA)). In many realistic settings, unwanted non-RWA dynamics and couplings to higher levels cannot be neglected, and will limit the operation fidelity even if non-adiabatic errors are suppressed. For this reason, the utility of analytic STA protocols for high-fidelity operations in complex systems have remained unclear.

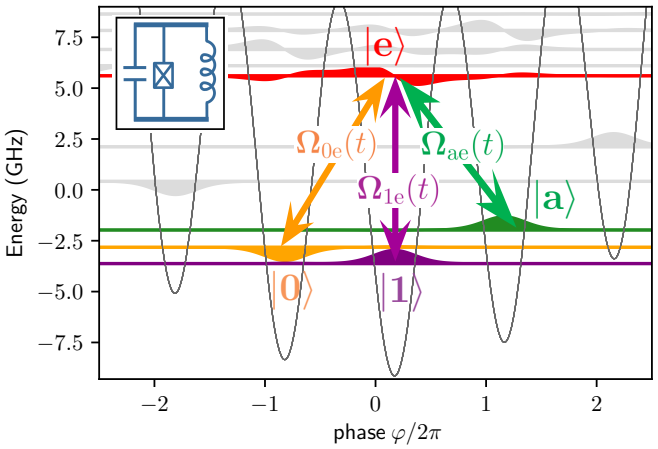


FIG. 1. A single-qubit gate realized using a 4-level tripod energy structure consisting of three lower energy levels ($|0\rangle$, $|1\rangle$ and $|a\rangle$) resonantly coupled to an excited state $|e\rangle$ by three different driving tones: each with envelope $\Omega_{0e}(t)$, $\Omega_{1e}(t)$, and $\Omega_{ae}(t)$, respectively. This tripod-energy structure can be engineered in a multi-energy spectrum (shown above are 9 lowest energy levels) of a fluxonium qubit whose device schematic is shown in the inset (upper left corner). The driving fields can drive unwanted transitions which give rise to coherent errors, e.g., transitions between the energy levels outside the tripod structure (shown using light gray) and the energy levels in the tripod structure as well as the transitions inside the tripod structure driven by non-resonant tones.

In this paper, we present a generic approach for improving STA protocols in settings where assuming an idealized dynamics is not possible, e.g., non-RWA terms cannot be neglected. The result is a general method for *analytically* deriving pulse sequences that both fully cancel non-adiabatic errors, and partially mitigate non-RWA

* setiawan@uchicago.edu

errors. To highlight the efficacy of our approach, we focus on a specific, experimentally relevant setting: an accelerated geometric tripod gate implemented in a fluxonium-style superconducting circuit [24–27]. The isolated qubit levels of this system make conventional approaches to gates problematic, motivating new ideas that do not require a direct coupling of qubit levels. The so-called tripod gate [3, 4] is a natural candidate. However, as shown in Fig. 1, a fluxonium circuit has a complex level structure, implying that non-resonant, non-RWA corrections will be important. Using realistic parameters compatible with experiment, we show (via full master-equation solutions including dissipation, non-RWA effects, and power constraints from cavity-based driving) that our enhanced STA gate achieves gate fidelities of ≈ 0.9997 with a gate time $t_g = 100$ ns. This is roughly comparable to results obtained using fully numerical optimal control on a related superconducting circuit [28]. It also demonstrates that the accelerated tripod-gate can be advantageous, despite the ability to realize a perfectly isolated tripod level structure. Note that the use of our enhanced protocol is crucial: if one simply uses the STA derived without corrections, the fidelity error for the same gate time is orders of magnitude worse.

The paper is organized as follows. We begin in Sec. II by introducing the most general version of our problem: how can one analytically design STA protocols in complex multi-level systems? In Sec. III, we briefly review the basic (RWA) geometric tripod gate [3, 4] as well as its accelerated version [16]. In Sec. IV, we go beyond the RWA, and discuss how in general STA approaches can be further enhanced to mitigate non-resonant errors. In Sec. V, we explore the utility of these methods by applying them to design an accelerated gate in a realistic fluxonium superconducting circuit. Results for gate performance are presented in Sec. VI, and comparison against a simpler "direct-driving" gate is presented in Sec. VII. We summarize our results in Sec. VIII.

II. GENERAL PROBLEM: SHORTCUT-TO-ADIABATICITY APPROACHES FOR COMPLEX DRIVEN SYSTEMS

We begin by considering a generic driven multilevel system whose Hamiltonian in the lab frame has the form

$$\hat{H}(t) = \sum_k \varepsilon_k |k\rangle\langle k| + \left(V(t) \sum_{\substack{(k,l)\,| \\ l>k}} n_{kl} |k\rangle\langle l| + \text{H.c.} \right), \quad (1)$$

where ε_k, $|k\rangle$ are the eigenenergies and eigenstates of the undriven system, and $n_{kl} = \langle k|\hat{n}|l\rangle$ is an effective dipole matrix element. The full control pulse $V(t)$ consists of several distinct drive tones ω_j, each associated with a slowly-varying complex envelope $V_j(t)$, i.e.,

$$V(t) = \frac{1}{2} \sum_j \left(V_j(t) e^{i\omega_j t} + \text{H.c.} \right). \quad (2)$$

We next move to an interaction picture defined by $\hat{U}_{\text{diag}} = e^{-i\hat{H}_{\text{diag}}t}$ where $\hat{H}_{\text{diag}} = \sum_k \varepsilon_k |k\rangle\langle k|$. The Hamiltonian in this frame takes the general form

$$\hat{H}(t) = \hat{H}_0(t) + \hat{H}_{\text{err}}(t). \quad (3)$$

Here $\hat{H}_0(t)$ describes resonant processes, and is time-dependent only through its dependence on the envelope functions $V_j(t)$. Defining $\varepsilon_{kl} = \varepsilon_l - \varepsilon_k$, we have

$$\hat{H}_0(t) = \frac{1}{2} \sum_j \sum_{\substack{(k,l)\,| \\ l>k \,\&\, \varepsilon_{kl}=\omega_j}} \Big(V_j(t) n_{kl} |k\rangle\langle l| + \text{H.c.} \Big). \quad (4)$$

In contrast, $\hat{H}_{\text{err}}(t)$ describes all non-resonant processes

$$\hat{H}_{\text{err}}(t) = \frac{1}{2} \Bigg(\sum_j \sum_{\substack{(k,l)\,| \\ l>k \,\&\, \varepsilon_{kl}\neq\omega_j}} V_j(t) e^{-i(\varepsilon_{kl}-\omega_j)t} n_{kl} |k\rangle\langle l| + \sum_j \sum_{\substack{(k,l)\,| \\ l>k \,\&\, \varepsilon_{kl}\neq-\omega_j}} V_j^*(t) e^{-i(\varepsilon_{kl}+\omega_j)t} n_{kl} |k\rangle\langle l| \Bigg) + \text{H.c.} \quad (5)$$

The standard next step in most analytic STA approaches to quantum control is to make the rotating wave approximation (RWA): one assumes that the energy detunings in $\hat{H}_{\text{err}}(t)$ are sufficiently large that this error Hamiltonian can be approximated as zero. The result is a much simpler Hamiltonian that only involves the slowly varying amplitudes $V_j(t)$, and which typically only couples a small subset of levels. It is in this context that many exact STA protocols have been derived; these protocols yield a perfect, error-free evolution within the RWA. Examples range from accelerated versions of the two-level Landau-Zener problem [11, 13, 14, 29], to more complex three-[17, 30–33] and four-[16] level protocols.

Despite the power of the above STA approaches, they only address non-adiabatic errors associated with the RWA Hamiltonian $\hat{H}_0(t)$. A crucial question is whether they can also be adapted to address additional errors arising from the non-RWA dynamics described by $\hat{H}_{\text{err}}(t)$. Corrections to the RWA are important in many physical systems if one is interested in realizing truly high-fidelity operations. One could in principle try to derive an STA for the full multilevel fast-oscillating Hamiltonian $H(t)$. But in most cases, this is completely infeasible. Not only does this involve dealing with a large-dimensional Hilbert space, but it also involves working with a starting Hamiltonian that has extremely fast time-dependent terms (i.e., in $\hat{H}_{\text{err}}(t)$), and hence is nowhere close to an adiabatic limit.

A central goal of this work is to present a much more tractable approach to adapting exact STA protocols so that they also mitigate non-resonant, non-RWA errors. Our method is ultimately perturbative, and amounts to modifying the original (RWA) STA protocol to correct the leading effects of $\hat{H}_{\rm err}(t)$. We stress that our approach retains the crucial feature of the original STA pulse sequence of being described and derived fully analytically, i.e., no recourse is made to numerical optimal-control approaches. While our method is extremely general, we will focus in what follows on a particularly promising protocol involving accelerated geometric gates using a tripod level structure [3, 4] [see Fig. 2(a)]. Note that the general approach we present is completely distinct from recent works [34, 35] which deliberately introduce additional high-frequency oscillatory terms to a RWA Hamiltonian to approximately engineer desired STA protocols.

III. REVIEW: ACCELERATED ADIABATIC QUANTUM GATES

In this section we briefly review the basic geometric tripod gate introduced in Refs. [3, 4], and its accelerated version [16]. All these analyses were done in the context of a simplified four-level RWA Hamiltonian. Our review here will set the stage for our following discussion on how these approaches can be modified and effectively implemented in a realistic multi-level superconducting circuit where non-RWA effects play a crucial role.

A. Double STIRAP protocol in an ideal tripod system

Starting with the full driven Hamiltonian in Eq. (1), we assume a situation where within the RWA, we realize a so-called tripod level configuration [see Fig. 2(a)]. An ideal tripod system consists of three lower levels (labeled by $|0\rangle, |1\rangle$ and $|\mathrm{a}\rangle$) which are controllably coupled to a common excited state $|\mathrm{e}\rangle$ [see Fig. 2(a)]. Denoting these (complex) couplings as $\Omega_{j\mathrm{e}}(t)$ $(j = 0, 1, \mathrm{a})$, the tripod Hamiltonian is ($\hbar = 1$):

$$\hat{H}_0(t) = \frac{1}{2}\left[\Omega_{0\mathrm{e}}(t)|0\rangle\langle \mathrm{e}| + \Omega_{1\mathrm{e}}(t)|1\rangle\langle \mathrm{e}| + \Omega_{\mathrm{ae}}(t)|\mathrm{a}\rangle\langle \mathrm{e}| + \mathrm{H.c.}\right], \tag{6}$$

where

$$\Omega_{j\mathrm{e}}(t) = V_{j\mathrm{e}}(t) n_{j\mathrm{e}}, \tag{7}$$

for $j = 0, 1, \mathrm{a}$. We will take the states $|0\rangle$ and $|1\rangle$ to encode a logical qubit, while $|\mathrm{a}\rangle$ and $|\mathrm{e}\rangle$ serve as auxiliary states used to perform gate operations.

The basic idea of the gate is that $\hat{H}_0(t)$ always has two degenerate zero-energy adiabatic eigenstates, and hence cyclic adiabatic evolution can result in a non-trivial geometric 2×2 unitary in this subspace. To understand this more concretely, we follow Ref. [16], and consider control pulses of the form

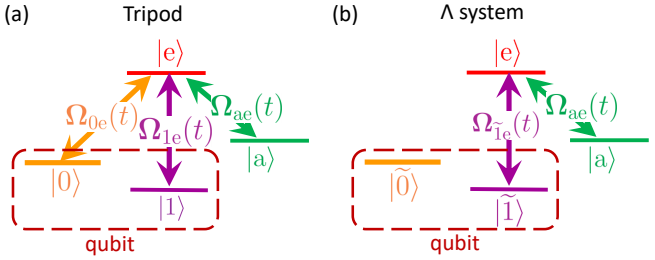


FIG. 2. (a) An ideal tripod system consisting of three lower energy levels (the qubit states $|0\rangle$ and $|1\rangle$ as well as the auxiliary state $|\mathrm{a}\rangle$) resonantly coupled to an excited state $|\mathrm{e}\rangle$ by three different driving tones: each with envelope $\Omega_{0\mathrm{e}}(t)$, $\Omega_{1\mathrm{e}}(t)$, $\Omega_{\mathrm{ae}}(t)$, respectively. (b) An effective Λ system (consisting of the state $|\tilde{1}\rangle$ and $|\mathrm{a}\rangle$ resonantly coupled to the state $|\mathrm{e}\rangle$) used to describe the dynamics of the tripod system. The control pulses $\Omega_{\tilde{1}\mathrm{e}}(t)$ and $\Omega_{\mathrm{ae}}(t)$ perform a double STIRAP protocol which cyclically evolves the zero-energy dark states in the Λ system.

$$\Omega_{0\mathrm{e}}(t) = \Omega_0 \cos(\alpha)\sin[\theta(t)], \tag{8a}$$
$$\Omega_{1\mathrm{e}}(t) = \Omega_0 \sin(\alpha)\sin[\theta(t)] e^{i\beta}, \tag{8b}$$
$$\Omega_{\mathrm{ae}}(t) = \Omega_0 \cos[\theta(t)] e^{i\gamma(t)}. \tag{8c}$$

The angles α and θ control the relative magnitudes of the pulses, while β and γ control relative phases; we will only require $\theta(t)$ and $\gamma(t)$ to be time dependent. The overall amplitude Ω_0 sets the instantaneous adiabatic gap of $\hat{H}_0(t)$, which we have chosen to keep constant:

$$\Omega_{\rm ad}(t) \equiv \frac{1}{2}\sqrt{|\Omega_{0\mathrm{e}}(t)|^2 + |\Omega_{1\mathrm{e}}(t)|^2 + |\Omega_{\mathrm{ae}}(t)|^2} = \frac{\Omega_0}{2}. \tag{9}$$

At every instant in time, $\hat{H}_0(t)$ has two zero-energy dark states (orthogonal to $|\mathrm{e}\rangle$), and bright states at energy $\pm\Omega_0/2$.

The tripod system in Fig. 2(a) can be related to an effective three-level Λ system [36, 37] [Fig. 2(b)]. In this representation, the gate corresponds to a "double STIRAP protocol" in the Λ system, where the dark state

$$|\mathrm{d}(t)\rangle = \cos[\theta(t)]|\tilde{1}\rangle - e^{i\gamma(t)}\sin[\theta(t)]|\mathrm{a}\rangle \tag{10}$$

undergoes a cyclic adiabatic evolution $|\tilde{1}\rangle \to |\mathrm{a}\rangle \to |\tilde{1}\rangle$. This requires an appropriate cyclic variation of the pulse parameter $\theta(t)$ (see Ref. [16] and Appendix A for details). This cyclic evolution can result in a Berry phase. We take the gate to start at $t = 0$ and end at $t = t_g$. For the case where the pulse parameter $\gamma(t)$ is chosen as

$$\gamma(t) = \gamma_0 \Theta\left(t - \frac{t_g}{2}\right), \tag{11}$$

with $\Theta(t)$ being the Heaviside step function, this geometric phase is simply γ_0 [16].

In the adiabatic limit $\dot{\theta}(t)/\Omega_0 \to 0$, the gate unitary in the qubit subspace is given by [16]

$$\hat{U}_{\mathrm{G},01} = \exp\left(-i\gamma_0/2\right) \exp\left(-i\frac{\gamma_0}{2}\boldsymbol{n}\cdot\hat{\boldsymbol{\sigma}}_{01}\right), \tag{12a}$$

$$\boldsymbol{n} = [\sin(2\alpha)\cos(\beta), \sin(2\alpha)\sin(\beta), \cos(2\alpha)], \tag{12b}$$

where $\hat{\boldsymbol{\sigma}}_{01} = (|0\rangle\langle 1| + \mathrm{H.c.}, -i|0\rangle\langle 1| + \mathrm{H.c.}, |0\rangle\langle 0| - |1\rangle\langle 1|)$ is the vector of the Pauli matrices in the qubit subspace. For example, the X gate can be realized by using the angle parameters $\alpha = \pi/4$, $\beta = 0$ and $\gamma_0 = \pi$.

B. Accelerated tripod gates

The geometric tripod gate yields a perfect gate fidelity in the adiabatic limit where the protocol time is infinitely longer than $1/\Omega_0$. In many realistic systems, dissipative effects involving the lower tripod levels make such long evolution times infeasible. It would thus be desirable to reduce the gate time without introducing non-adiabatic errors. This is exactly the goal of *shortcuts to adiabaticity* (STA) methods [11–13, 16, 17, 29].

Following Ref. [16], we consider an STA protocol based on the SATD method [17], where non-adiabatic errors are mitigated by having the system follow a dressed adiabatic eigenstate (see Appendix B). The accelerated protocol is implemented by simply modifying the complex envelope of the original control pulse [16, 17, 38, 39]. Specifically, the SATD protocol requires that one corrects the original pulses via

$$\begin{aligned}
\Omega_{0\mathrm{e}}(t) \to \tilde{\Omega}_{0\mathrm{e}}(t) &\equiv \Omega_0\cos(\alpha)\left[\sin[\theta(t)] + 4\frac{\cos[\theta(t)]\ddot{\theta}(t)}{\Omega_0^2 + 4\dot{\theta}^2(t)}\right], \\
\Omega_{1\mathrm{e}}(t) \to \tilde{\Omega}_{1\mathrm{e}}(t) &\equiv \Omega_0\sin(\alpha)e^{i\beta}\left[\sin[\theta(t)] + 4\frac{\cos[\theta(t)]\ddot{\theta}(t)}{\Omega_0^2 + 4\dot{\theta}^2(t)}\right], \\
\Omega_{\mathrm{ae}}(t) \to \tilde{\Omega}_{\mathrm{ae}}(t) &\equiv \Omega_0 e^{i\gamma(t)}\left[\cos[\theta(t)] - 4\frac{\sin[\theta(t)]\ddot{\theta}(t)}{\Omega_0^2 + 4\dot{\theta}^2(t)}\right],
\end{aligned} \tag{13}$$

where the angle $\gamma(t)$ [Eq. (11)] remains unchanged. It can be shown that [16] the resulting accelerated protocol obtained using correction in Eq. (13) achieves the same unitary $\hat{U}_{\mathrm{G},01}$ in the qubit subspace as in the adiabatic limit [c.f. Eq. (12)]. In what follows, we will use a tilde throughout to denote SATD-corrected pulse parameters.

C. Infinite family of perfect RWA protocols

For our ideal (RWA) tripod systems, our SATD approach yields an *infinite* number of perfect protocols (i.e., pulse sequences) that realize a given gate in a fixed gate time t_g. These protocols are indexed by Ω_0 (c.f. Eq. (9)), which is the scale of amplitudes of the *uncorrected* pulse (and the corresponding time-independent adiabatic gap).

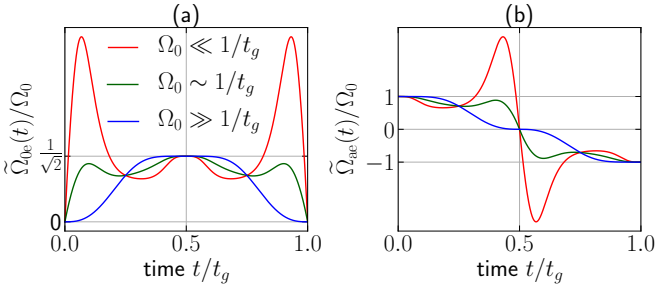


FIG. 3. Time profiles of the envelopes of the three driving tones used to realize X gates in an ideal tripod system: (a) $\tilde{\Omega}_{0\mathrm{e}}(t) = \tilde{\Omega}_{1\mathrm{e}}(t)$, (b) $\tilde{\Omega}_{\mathrm{ae}}(t)$. Pulses are calculated using Eq. (13) with $\theta(t)$ given in Eq. (A5) of Appendix A and other angle parameters given by $\alpha = \pi/4$, $\beta = 0$, and $\gamma_0 = \pi$.. Shown here are pulses for a fixed gate time t_g but with different uncorrected gap frequencies: $\Omega_0/2\pi = 4/t_g$ (red curve), $\Omega_0/2\pi = 2/t_g$ (green curve), and $\Omega_0/2\pi = 0.2/t_g$ (blue curve). The SATD correction to the adiabatic pulse becomes larger as Ω_0 decreases.

For every choice of Ω_0, there is a corresponding SATD protocol [given by Eq. (13)] that yields a pulse sequence with a perfect gate fidelity. At a heuristic level, for $\Omega_0 \gg 1/t_g$ the uncorrected protocol is already almost in the adiabatic limit, meaning that the additional SATD modification of pulses will be minimal. In contrast, for $\Omega_0 < 1/t_g$, the SATD correction to the original pulse shape will be extremely large (in order to cancel non-adiabatic errors). Figure 3 shows the time profiles of pulse envelopes used to realize perfect X gates in an ideal tripod system, for different choices of $\Omega_0 t_g$. We note that besides the degeneracy in choosing Ω_0, there is also a degeneracy resulting from different choices of the pulse-shape function $P(t/t_g)$ [Eq. (A6)] that determines $\theta(t)$ [Eq. (A5)].

IV. ENHANCING STA PROTOCOLS TO MITIGATE NON-RESONANT ERRORS

We now return to the central question of this work: can STA approaches still be effective in settings where non-resonant, non-RWA processes also degrade fidelity? In this section, we present a general strategy for improving STA protocols to partially mitigate non-RWA errors. For concreteness, we will do this in the specific context of the accelerated tripod gate introduced above. To achieve a gate in a fixed time t_g, our strategy has two basic steps:

1. We first use the degeneracy of perfect STA protocols that exists in the RWA limit (c.f. Sec. III C) to pick a protocol that minimizes the "size" (appropriately defined) of our control fields.

2. Next, we modify the SATD pulse shape by chirping the control fields to offset frequency-shifts arising from non-RWA terms. The form of the required

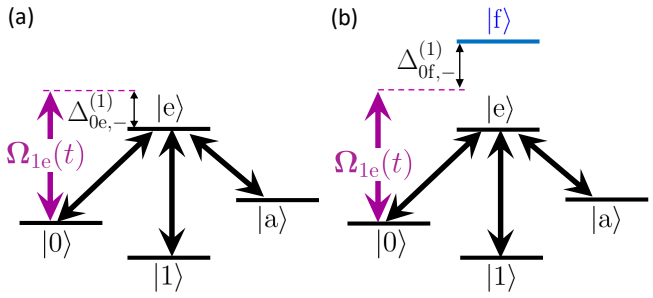


FIG. 4. Spurious non-RWA processes giving rise to coherent errors. (a) Crosstalk where the driving tone $\Omega_{1e}(t)$ drives the $|0\rangle \leftrightarrow |e\rangle$ transition. The frequency of the driving tone $\Omega_{1e}(t)$ is detuned by $\Delta^{(1)}_{0e,-}$ from the $|0\rangle \leftrightarrow |e\rangle$ transition. (b) Coupling to spurious level $|f\rangle$ where the driving tone $\Omega_{1e}(t)$ drives the $|0\rangle \leftrightarrow |f\rangle$ transition. The frequency of the driving tone $\Omega_{1e}(t)$ is detuned by $\Delta^{(1)}_{0f,-}$ from the $|0\rangle \leftrightarrow |f\rangle$ transition.

chirp can be found *analytically* using a perturbative approach.

As we will see, this two-pronged, *fully analytic* approach results in a modified set of pulses that yield exceptional gate performance even when non-RWA effects are included. We now discuss each step of our general method in more detail, focusing on the specific case of our accelerated tripod gate.

A. Step 1: Power minimization

The non-resonant, non-RWA processes described by $\hat{H}_{\rm err}(t)$ (c.f. Eq. (5)) yield new unwanted coherent dynamics that will degrade the performance of our gate; examples processes are sketched in Fig. 4. One effect of these terms is to generate effective time-dependent energy shifts of the four levels involved in our tripod gate. We define $\Delta^{(j)}_{kl,\pm} = \varepsilon_l - \varepsilon_k \pm \omega_{je}$ as the detuning associated with the transition between the energy level $|k\rangle$ and the level $|l\rangle$ associated with the drive tone ω_{je}. Recall that in our tripod gate, there is a drive tone for each ground-state level (c.f. Fig. 2), hence $j = 0, 1, \mathrm{a}$. Using a Magnus-based approach (see Ref. [40]), one can derive the leading-order, time-dependent energy shift $\delta\varepsilon_k(t)$ of energy level $|k\rangle$ due to $\hat{H}_{\rm err}(t)$. This energy shift has the usual form expected from the second-order perturbation theory, i.e.,

$$\delta\varepsilon_k(t) = \sum_{\substack{j=0,1,\mathrm{a}\\ \sigma=\pm}} \sum_{l\,|\,\Delta^{(j)}_{kl,\sigma}\neq 0} \frac{|\tilde{V}_{je}(t) n_{kl}|^2}{4\Delta^{(j)}_{kl,\sigma}}. \tag{14}$$

The sum here is over all non-resonant processes that involve the state $|k\rangle$. We will be interested in energy shifts of the four tripod levels, i.e. $k = 0, 1, \mathrm{a}, \mathrm{e}$. Note that the intermediate states l in Eq. (14) includes both the four tripod levels (i.e., "crosstalk" process, Fig. 4(a)), or could involve non-tripod states (i.e., couplings to "leakage" levels, Fig. 4(b)). Formally, Eq. (14) is valid in the perturbative limit where $|\tilde{V}_{je}(t) n_{kl}| \ll |\Delta^{(j)}_{kl,\pm}|$ and the quasistatic limit $|\int_0^{t_g} \partial_{t'} \tilde{V}_{je}(t') n_{kl} e^{-i\Delta^{(j)}_{kl,\pm} t'} dt'| \ll |\Delta^{(j)}_{kl,\pm}|$.

The simplest way to mitigate errors associated with the above non-RWA generated energy shifts is to minimize their size by minimizing the SATD-corrected pulse amplitudes $\tilde{V}_{je}(t)$. We would like to find a simple metric to characterize the size of these amplitudes in a meaningful manner. We see that at each instant in time, the relevant quantity is the square of these amplitudes (as the energy shifts are a second-order effect). This motivates characterizing the "size" of our control pulses by the root-mean-square (RMS) voltage of the control field, i.e.,

$$\tilde{V}_{\rm RMS} \equiv \sqrt{\frac{1}{t_g}\int_0^{t_g} \left[\tilde{V}(t)\right]^2 dt}. \tag{15}$$

Here, $\tilde{V}(t)$ is the total real-valued control pulse function (including the SATD correction), c.f. Eq. (2).

Using the specific form of the SATD pulses in Eqs. (13), we have

$$\begin{aligned}\tilde{V}_{\rm RMS} &\simeq \sqrt{\frac{1}{2t_g}\int_0^{t_g} dt \sum_{j=0,1,\mathrm{a}} \left|\tilde{V}_{je}(t)\right|^2} \\ &= \sqrt{\frac{\cos^2\alpha}{|n_{0e}|^2} + \frac{\sin^2\alpha}{|n_{1e}|^2} + \frac{1}{|n_{ae}|^2}}\,\frac{\tilde{\Omega}_{\rm RMS}(t_g)}{2},\end{aligned} \tag{16}$$

where n_{0e}, n_{1e}, and n_{ae} are the tripod matrix elements, and

$$\begin{aligned}\tilde{\Omega}_{\rm RMS}(t_g) &\equiv \sqrt{\frac{1}{t_g}\int_0^{t_g} dt \left[|\tilde{\Omega}_{0e}(t)|^2 + |\tilde{\Omega}_{1e}(t)|^2 + |\tilde{\Omega}_{ae}(t)|^2\right]} \\ &= \Omega_0 \sqrt{\frac{1}{t_g}\int_0^{t_g} dt \left[1 + \left(\frac{\ddot{\theta}(t)}{\dot{\theta}^2(t) + \Omega_0^2/4}\right)^2\right]}.\end{aligned} \tag{17}$$

In the first line of Eq. (16), we have used the fact that the terms involving differences of tone frequencies almost exactly average to 0. Further, in the second line, we have use the fact that our protocol is symmetric about $t = t_g/2$, i.e., $\theta(t) = \pi/2 - \theta(t - t_g/2)$ for $t_g/2 < t \leq t_g$ [Eq. (A5) in Appendix A].

From Eq. (16), we can see that $\tilde{V}_{\rm RMS}$ is related to a more fundamental metric $\tilde{\Omega}_{\rm RMS}/2$ [Eq. (17)]: the time-averaged RMS value of the instantaneous gap of our Hamiltonian $\hat{H}_0(t)$ for a SATD-corrected pulse sequence. Note that this metric is solely a property of the ideal tripod Hamiltonian $\hat{H}_0(t)$ and our SATD pulse sequence. The SATD correction makes the adiabatic gap time dependent, and necessarily increases $\tilde{\Omega}_{\rm RMS}(t_g)$ above Ω_0.

We finally can make use of the property discussed in Sec. III C: in the RWA-limit, there are an infinite number of SATD protocols that yield a perfect gate fidelity

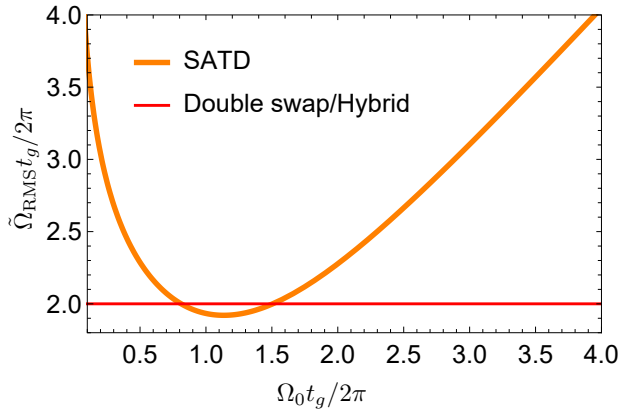


FIG. 5. Plot of $\tilde{\Omega}_{\rm RMS}t_g/2\pi$ vs $\Omega_0 t_g/2\pi$ for SATD (orange curve) and double-swap/hybrid protocols (red line). For the SATD protocol, $\tilde{\Omega}_{\rm RMS} = \Omega_0$ in the adiabatic limit ($\Omega_0 t_g \to \infty$). Away from the adiabatic limit, the SATD correction increases $\tilde{\Omega}_{\rm RMS}$ above Ω_0. The minimum of the SATD plot is $\tilde{\Omega}_{\rm RMS}t_g/2\pi = 1.92$ occuring at $\Omega_0 t_g/2\pi = 1.135$. The red line $\tilde{\Omega}_{\rm RMS}t_g/2\pi = 2$ is the value of $\tilde{\Omega}_{\rm RMS}$ for the double-swap protocol and hybrid scheme.

for a given gate time t_g. Out of this set, we will choose the protocol that minimizes $\tilde{\Omega}_{\rm RMS}$ (and hence will approximately minimize the phase errors arising from the non-RWA energy shifts $\delta\varepsilon_k(t)$ in Eq. (14)). As discussed, the different SATD protocols are indexed by Ω_0, the adiabatic gap associated with the *uncorrected* pulses. We thus seek to identify the value of $\Omega_0 t_g$ that minimizes the SATD energy cost $\tilde{\Omega}_{\rm RMS}$. The behavior of this quantity (obtained numerically for the specific smooth pulse shape function $P(t)$ given in Eq. (A6)) is shown in Fig. 5. We find that $\tilde{\Omega}_{\rm RMS}/2\pi$ has a minimum value $\simeq 1.92/t_g$, occurring for $\Omega_0/2\pi = 1.135/t_g$. We have confirmed that attempting to further optimize by using more complex pulse-shape functions $P(t)$ does not yield appreciable improvements. We have thus completed step one of our two-step approach to enhancing accelerated gates to minimize non-RWA errors: use the degeneracy of STA protocols to pick a minimum-energy pulse.

Before proceeding, we pause to note something somewhat remarkable: by optimizing the parameter Ω_0, the SATD pulses are able to achieve an energy-time tradeoff (as quantified by the product $\tilde{\Omega}_{\rm RMS}t_g$) that is *better* than more obviously fast, non-adiabatic population-transfer protocols (see red curve in Fig. 5). These are the so-called double-swap protocols (see e.g., Refs. [41, 42]) (where one sequentially moves population through the three levels), and the "hybrid" protocol [42] (a non-adiabatic holonomic protocol). This point and connections to a formal quantum speed limit valid for time-dependent Hamiltonians [43] are discussed in Appendix C.

B. Step 2: Modifying accelerated protocol pulses to cancel non-RWA phase errors

The next step of our method is to go beyond simply minimizing the non-RWA energy shifts $\delta\varepsilon_k(t)$ in Eq. (14), and actually *cancel* them by slightly modifying our SATD pulses. In keeping with our general philosophy, we will derive an analytic prescription for how to do this (as opposed to resorting to a brute-force numerical optimization).

The methodology here is conceptually simple: to offset the unwanted, time-dependent non-RWA energy shifts, we introduce a time-dependent variation of the central tones in our pulse (i.e., a generalized chirp). Concretely, this means that we introduce a time-dependent shift in each of the three center frequencies ω_{je} ($j = 0, 1, \mathrm{a}$) that appear in our pulse:

$$\omega_{je} \to \tilde{\omega}_{je}(t) = \omega_{je} - \delta\varepsilon_j(t) + \delta\varepsilon_{\rm e}(t), \tag{18}$$

where $\delta\varepsilon_j$ is given by Eq. (14). These frequency shifts ensure that at each instant in time, each tone is resonant (to leading order) with the appropriate transition it is intended to drive.

The net result of our approach is thus a two-step correction to the original pulse in Eq. (8). For a given desired gate time t_g, we first pick an optimal value of Ω_0 as per Sec. IV A and add the SATD correction to the pulses as per Eq. (13). Subsequently, we chirp each of the three center frequencies as per Eq. (18). We can write the overall modification of each control tone in Eq. (2) as

$$V_{je}(t)\exp\left(i\omega_{je}t\right) \to \tilde{V}_{je}(t)\exp\left(i\int_0^t \tilde{\omega}_{je}(t')dt'\right), \tag{19}$$

where $\tilde{V}_{je}(t) = \tilde{\Omega}_{je}(t)/n_{je}$ are the SATD-corrected pulse envelopes, and $\tilde{\omega}_{je}(t)$ are the chirp-corrected central drive frequencies.

C. Leakage errors and connections to DRAG

Our discussion so far has only focused on errors arising from energy shifts generated by the non-RWA terms in $\hat{H}_{\rm err}(t)$. There is another generic kind of error to consider: the non-RWA terms can drive transitions out of the tripod subspace, leading to a final population of non-tripod levels. This kind of error is commonly referred to as leakage, and has been discussed extensively in many other settings (e.g., in discussing gate errors in weakly-anharmonic transmon-style superconducting qubits [40, 44, 45]).

The general approach we have taken to mitigating non-RWA errors partially minimizes leakage by minimizing the size of the control pulses (c.f. Sec. IV A). However, we do not make any additional modifications of our pulses to further reduce leakage errors. This is in contrast to our treatment of phase errors (which we further mitigate through frequency chirping). It is also in contrast

to the well-known DRAG technique [44, 45] for dealing with leakage errors in superconducting circuits driven by a single control tone.

There are two key rationales for our apparent neglect of pure leakage errors. The first is purely pragmatic: in general there are many equally important leakage levels, and there is no simple way to modify our pulses (using the Magnus strategy of Ref. [40]) to simultaneously correct all of these error channels. This is stark contrast to the usual DRAG problem in superconducting circuits, where there is just a single relevant leakage level (i.e., the second excited state of the qubit). The second rationale is that in the perturbative limit (where non-RWA errors are small), leakage errors are much smaller than the phase errors associated with energy shifts. As shown in Refs. [39, 46] in the long-gate time limit ($t_g \gg 1/\Delta^{(j)}_{kl,\sigma}$), phase errors scales as $1/(\Delta^{(j)}_{kl,\pm} t_g)^2$, while leakage errors are much weaker, scaling as $1/(\Delta^{(j)}_{kl,\pm} t_g)^4$.

As we will see in the next section (where we implement our ideas in a realistic multi-level superconducting fluxonium circuit), our approach yields extremely good results despite the lack of any specific leakage correction.

V. FLUXONIUM QUBIT

A. Basic setup

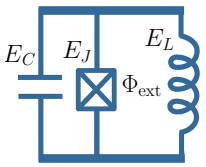
The tripod gate we have analyzed is ideally suited to systems where it is difficult to directly drive transitions between the qubit levels $|0\rangle$ and $|1\rangle$. This is often the case in qubits that have long T_1 relaxation times. A paradigmatic example that has received considerable recent attention is a fluxonium-style superconducting qubit [24–27]. A fluxonium circuit consists of a single Josephson junction (energy E_J) in parallel with both a capacitor (charging energy E_C) and a superinductor (inductive energy energy E_L, typically implemented using a chain of Josephson junctions) [Fig. 6(a)]. Experiments demonstrate that these qubits can possess exceptionally long relaxation times (on the order of milliseconds) [25–27, 47]. Moreover, they can also be made first-order insensitive to the dephasing from $1/f$ charge noise [48]. These properties make fluxonium an attractive quantum computing platform. A complication however is that the relative isolation of qubit levels (which yields long T_1 times) also makes conventional approaches to gates challenging. This makes our accelerated tripod gate especially attractive.

The fluxonium Hamiltonian is

$$\hat{H}_f = 4E_C\hat{n}^2 - E_J\cos(\hat{\varphi} - 2\pi\Phi_{\rm ext}/\Phi_0) + \frac{1}{2}E_L\hat{\varphi}^2, \quad (20)$$

where $\hat{n}$ and $\hat{\varphi}$ are (respectively) the charge and phase operators. They obey the commutation relation $[\hat{n}, \hat{\varphi}] = i$. $\Phi_{\rm ext}$ is the external magnetic flux biasing the loop formed by the Josephson junction and its shunting inductance

(a)



(b)

kl	$\omega_{kl}/2\pi$ (GHz)	$\lvert n_{kl}\rvert$
01	0.81	0.02
0e	8.42	0.27
1e	9.23	0.46
ae	7.58	0.16

FIG. 6. (a) Circuit diagram of a fluxonium. (b) The qubit and tripod transition frequencies (ω_{jk}) as well as the magnitude of the charge matrix elements (n_{jk}) of the fluxonium whose energy spectrum is given in Fig. 1. The computational states in the tripod system are labeled by $|0\rangle$, $|1\rangle$, $|a\rangle$, and $|e\rangle$. The fluxonium parameters used to get the values of ω_{jk} and n_{jk} in panel (b) are $E_L/h = 0.063$ GHz, $E_J/h = 9.19$ GHz, $E_C/h = 2$ GHz and $\Phi_{\rm ext} = 0.17\Phi_0$.

and $\Phi_0 = h/2e$ is the flux quantum. Eq. (20) shows that the effective phase potential consists of a cosine potential superimposed on a parabolic background [see Fig. 1]. The highly tunable, anharmonic nature of the fluxonium allows us to engineer a variety of different candidate four-level tripod systems.

B. Optimal parameters for a tripod gate

A first question is to identify parameters yielding an "optimal" tripod configuration, meaning that we both have a long T_1 time as well as an accelerated SATD gate with small (non-RWA) coherent errors. This leads to the following selection criteria:

1. The ground states should be well isolated from each other (i.e., small charge matrix element connecting them) and be nondegenerate. Strong isolation ensures a long T_1 time.

2. The charge matrix elements coupling the excited state to the ground states of the tripod system must be large and have the same order of magnitude. The latter helps minimize coherent errors arising from non-RWA processes.

Since the above requirements cannot be perfectly satisfied simultaneously (see Appendix D), we choose parameters that strike an optimal balance between them. To this end, we performed a numerical search through parameter space to identify optimal regimes. The result is the following near-ideal parameter set for realizing high-fidelity SATD tripod gates in a T_1 protected regime: $E_L/h = 0.063$ GHz, $E_J/h = 9.19$ GHz, $E_C/h = 2$ GHz and $\Phi_{\rm ext} = 0.17\Phi_0$ (see Appendix D for parameter justification). We note that the small inductive energy here puts our device in same regime as the "Blochnium" circuit recently realized in experiment [49].

For the above parameter set, the qubit and tripod transition frequencies together with their corresponding

charge matrix elements are shown in Fig. 6(b). The corresponding energy spectrum and wavefunctions are plotted in Fig. 1. We have labeled the energy levels used for the tripod system by $|0\rangle$, $|1\rangle$, $|\mathrm{a}\rangle$, and $|\mathrm{e}\rangle$. The charge matrix element connecting qubit states $|0\rangle$ and $|1\rangle$ is extremely small as desired: $|n_{01}| \equiv |\langle 0|\hat{n}|1\rangle| = 0.02$. In contrast, the charge matrix elements for the desired tripod transitions are much larger and comparable in magnitude to one another: $|n_{0\mathrm{e}}| = 0.27$, $|n_{1\mathrm{e}}| = 0.46$ and $|n_{\mathrm{ae}}| = 0.16$ (right column of the table in Fig. 6(b)).

C. SATD protocols for tripod gates in fluxonium

To realize our accelerated tripod gate, we drive the fluxonium circuit with a microwave pulse (described by a voltage $V(t)$) that couples to the charge operator $\hat{n}$. In the eigenbasis of the bare fluxonium Hamiltonian $\hat{H}_f$ [Eq. (20)], we can write the Hamiltonian of the driven fluxonium exactly in the general form given in Eq. (1). To realize the accelerated tripod gate, the driving voltage $V(t)$ [Eq. (2)] consists of three driving tones $V_{j\mathrm{e}}(t) = \Omega_{j\mathrm{e}}(t)/n_{j\mathrm{e}}$ for $j = 0, 1, \mathrm{a}$ [see Eq. (7)]. Here, $\Omega_{j\mathrm{e}}(t)$ is the complex coupling given in Eq. (8) for the uncorrected pulse and Eq. (13) for the SATD corrected pulse.

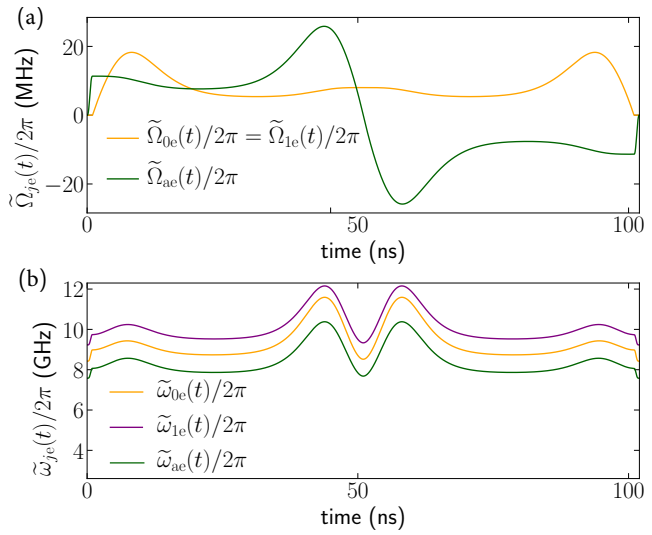


FIG. 7. Time profiles of (a) the envelopes and (b) chirped frequencies of the three driving tones used to realize SATD X gates in a realistic fluxonium. The pulse sequences in panel (a) are given in Eq. (E1). They are obtained by sandwiching the pulses in Eq. (13) by a ramp time t_{ramp} at the beginning and end of the protocol during which the $\tilde{\Omega}_{\mathrm{ae}}(t)$ pulse (green curve in panel (a)) is turned on and off, respectively, using a smooth polynomial function given in Eq. (A6). The parameters used are $t_g = 100$ ns, turn-on/off time $t_{\mathrm{ramp}} = 1$ ns, $\Omega_0/2\pi = 1.135/t_g = 11.35$ MHz. The parameter set for the fluxonium is the same as that used in Fig. 6.

The envelopes of the driving tones $\tilde{\Omega}_{j\mathrm{e}}(t)$ of the SATD tripod gate pulse (for optimal $\Omega_0 t_g$) are shown in Fig 7(a). We have slightly modified the pulses derived in Sec. III B so that $V(t)$ goes smoothly to zero at the start and end of the protocol (as would be in the case in experiment). We do this by sandwiching the ideal pulses with a short ramp up (down) of duration $t_{\mathrm{ramp}} = 0.01 t_g$ at the beginning (end) of the protocol. During this ramp, the $\tilde{\Omega}_{\mathrm{ae}}(t)$ tone is smoothly turned on and off, respectively [see Appendix E]. We specifically use Eq. (A6) for a smooth ramp function that turns the pulse on and off. Note that including these ramps does not appreciably change our results.

VI. GATE PERFORMANCE: COMPARING DIFFERENT ERROR CHANNELS

To quantify the performance of our accelerated tripod gate, we calculate the state-averaged fidelity of the gate. This is given by [50]

$$\bar{F} = \frac{1}{6} \sum_{m=\pm x,\pm y,\pm z} \mathrm{Tr}\left[\hat{U}_{\mathrm{q}} \hat{\rho}_m \hat{U}_{\mathrm{q}}^{\dagger} \hat{\rho}_m(t_g)\right], \qquad (21)$$

where $\hat{\rho}_m$ is an axial pure state on the qubit's Bloch sphere with $m \in \{\pm x, \pm y, \pm z\}$ (e.g. $\hat{\rho}_x = 1/2(|0\rangle + |1\rangle)(\langle 0| + \langle 1|)$). $\hat{\rho}_m(t_g)$ is the lab-frame density matrix of the system at the end of the protocol ($t = t_g$) for the initial state $\hat{\rho}_m$. Here, $\hat{U}_{\mathrm{q}}$ is a product of the ideal target unitary gate operation in the qubit subspace $\hat{U}_{\mathrm{G},01}$ [c.f. Eq. (12)] and an innocuous phase factor corresponding to dynamical phases in the lab-frame:

$$\hat{U}_{\mathrm{q}} = \hat{U}_{\mathrm{diag},01}(t_g) \hat{U}_{\mathrm{G},01}, \qquad (22)$$

with $\hat{U}_{\mathrm{diag},01}(t_g) = \sum_{k=0,1} e^{-i\int_0^{t_g} \varepsilon_k(t)dt} |k\rangle\langle k|$. Here, $\varepsilon_k(t) = \varepsilon_k$ for unchirped protocols and $\varepsilon_k(t) = \varepsilon_k + \delta\varepsilon_k(t)$ for chirped protocols. In what follows, we use this standard metric to characterize a target X qubit gate in the presence of both coherent (non-RWA) errors and dissipation.

A. Effects of coherent errors only

Consider first the case where dissipation is neglected, and the only source of gate errors are the non-RWA terms in Eq. (5). To calculate gate performance in this limit, we numerically evolve initial states as per the lab-frame Hamiltonian [Eq. (1)] using the Python package QuTiP [51, 52]. We performed the simulation by including the 18 lowest energy levels of the fluxonium and all charge matrix elements in this space. We have checked that including more energy levels in the simulations does not change the results.

In Fig. 8, we plot the gate errors $\bar{\varepsilon} = 1 - \bar{F}$ for a target tripod gate $\hat{U}_{\mathrm{G},01} = -\hat{\sigma}_{x,01}$ (an X-gate) obtained for different choices of pulses. For all curves, the uncorrected gap frequency at each protocol time t_g is picked to be the

optimal value $\Omega_0/2\pi = 1.135/t_g$ (see Sec. IV A). If we use the uncorrected adiabatic pulses [c.f. Eq. (8), green curve], errors arise both from non-adiabatic transitions as well as non-RWA processes. If instead, we use the SATD-corrected pulses [purple curve, c.f. Eq. (13)], non-adiabatic errors are completely canceled, leaving only non-RWA errors. Finally, if we also frequency chirp the SATD pulses as per Eq. (18) (red curve), we further reduce gate errors by reducing the leading non-RWA errors. This yields a dramatic improvement over the unchirped SATD protocol at long gate times.

As discussed, our corrections do not specifically cancel pure leakage errors. To characterize these, we calculate the state-averaged population outside the tripod subspace (leakage population) at the end of the protocol. This is given by

$$1-\bar{\mathcal{P}}_{\text{tripod}} = 1-\frac{1}{6}\sum_{m=\pm x,\pm y,\pm z} \text{Tr}\left[\hat{P}_{\text{tripod}}\hat{\rho}_m(t_g)\right], \quad (23)$$

where $\bar{\mathcal{P}}_{\text{tripod}}$ and $\hat{P}_{\text{tripod}}$ are respectively the state-averaged population and projector in the tripod subspace. The final state-averaged leakage population is plotted as a gray dashed curve in Fig. 8. As discussed, leakage makes a minimal contribution to error at modest to long gate times.

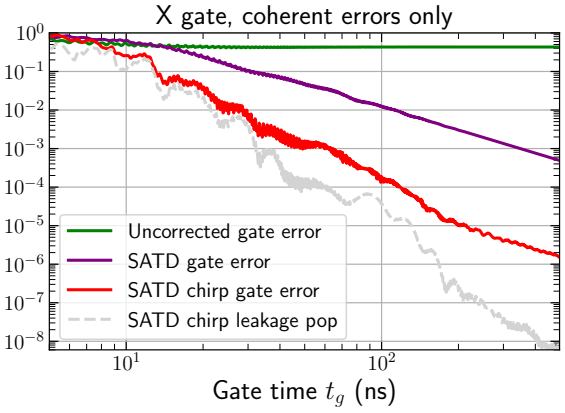


FIG. 8. State-averaged gate error $\bar{\varepsilon} = 1 - \bar{F}$ (solid curves) vs gate time t_g for different realizations of a tripod X gate in a fluxonium qubit, in the absence of dissipation. We keep the lowest 18 levels of the circuit, and include all non-RWA error channels. Curves correspond to uncorrected adiabatic pulses (green), SATD pulses without frequency chirping (purple) and SATD pulses with frequency chirping (solid red). For all pulses, we use the optimal value of the uncorrected gap frequency $\Omega_0/2\pi = 1.135/t_g$ that minimizes the RMS voltage. Also shown is the leakage population outside the tripod subspace for the chirped SATD protocol (dashed gray).

B. Effects of $1/f$ dephasing noise only

1. *Effective modelling of non-Markovian noise*

We now turn to modelling additional gate errors arising from dissipation. Given our operating point (isolated qubit states, but not at the flux sweet spot), $1/f$ flux noise will be the dominant form of dissipation. As is common [53–57], we will model this non-Markovian noise using an approximate Markovian description that qualitatively captures the relevant dephasing timescales correctly. Letting $\hat{\rho}(t)$ denote the fluxonium reduced density matrix, we model our system by the Lindblad-form master equation

$$\partial_t\hat{\rho}(t) = -i[\hat{H}(t),\hat{\rho}(t)] + \left(\hat{Z}\hat{\rho}\hat{Z} - \frac{1}{2}\{\hat{Z}^2,\hat{\rho}\}\right). \quad (24)$$

$\hat{H}(t)$ is the driven fluxonium Hamiltonian [Eq. (1)], and the Hermitian operator $\hat{Z}$ has the general form:

$$\hat{Z} = \sum_k \text{sgn}\left(\frac{\partial\varepsilon_k}{\partial\Phi_{\text{ext}}}\right)\sqrt{2\Gamma_k}|k\rangle\langle k|. \quad (25)$$

Heuristically, this describes the fact that each fluxonium energy level ε_k depends on the bias flux, and hence flux noise causes each energy to fluctuate. We have written the coupling constant associated with each level $|k\rangle$ in terms of an overall sign (which captures whether ε_k increases or decreases with increasing flux), and a magnitude Γ_k.

To fix the couplings Γ_k, we use the fact that for classical, Gaussian $1/f$ noise, free induction decay of a given coherence $\rho_{kl} \equiv \langle k|\hat{\rho}|l\rangle$ has a decay envelope $\exp\left[-(t/T_{\varphi,kl})^2\right]$ (up to logarithmic corrections). A standard calculation (see e.g., Ref. [55]) yields:

$$1/T_{\varphi,kl} = A_{\Phi_{\text{ext}}}|\partial_{\Phi_{\text{ext}}}(\varepsilon_k-\varepsilon_l)|\sqrt{|\ln D|}. \quad (26)$$

Properties of the flux-noise spectral density only enter through $A_{\Phi_{\text{ext}}}$ (the standardly-defined flux-noise amplitude) and D (the product of a measurement time and the low-frequency cutoff of the noise, see Refs. [28, 55]).

We will construct our approximate Markovian master equation by picking the couplings Γ_k to ensure that for free induction decay over a time t_g, the *final* decay of a large set of coherences is captured correctly. In particular, we take $|1\rangle$ as a reference level (i.e., the lowest energy level of the fluxonium), and insist that the net decay of any coherence ρ_{k1} ($k \neq 1$) after an evolution time t_g is the *same* for our Markovian dynamics as it would be following the non-Markovian, Gaussian-lineshape decay described by Eq. (26). This leads to the choice $\Gamma_1 = 0$, and for $k \neq 1$:

$$\Gamma_k = \Gamma_k(t_g) = t_g/(T_{\varphi,k1})^2. \quad (27)$$

We stress that Γ_k (and hence our master equation) depends only on the choice of total evolution time t_g; during the evolution the dephasing superoperator is constant.

The above choice guarantees that at the end of evolution for a time t_g, the overall free induction decay of any coherence involving the qubit level $|1\rangle$ (the lowest energy state of the circuit) is captured correctly. As discussed in Appendix F, there is no way to make a choice for the Γ_k that captures the decay of all coherences correctly. Nonetheless, as shown in the appendix, our approach if anything overestimates the dominant dephasing within the tripod subspace (see Table I of Appendix F). Note also that our approach overestimates dephasing compared to alternate approximations that use an explicitly time-dependent Γ_k [56, 57].

In what follows we take $A_{\Phi_{\rm ext}} = 3\mu\Phi_0$ (which is typical for state-of-the-art experiments [26, 27, 58]), and also choose (following Refs. [28, 55, 59]) the constant $D = (2\pi \times 1\,\text{Hz}) \times 10\,\mu s$. With these choices and for the circuit parameters used here, we find dephasing times $T_{\varphi,kl}$ that are of the order $\sim 1 - 100\mu$s [Table I of Appendix F]. The dephasing of the qubit levels is $T_{\varphi,01} = 7$ μs. Note that we have confirmed that our results are largely unchanged if one adds realistic T_1 decay processes (e.g., dielectric loss with dielectric quality factor $Q_{\rm diel} \gtrsim 10^6$). This is not surprising. The tripod gate involves the dynamics of eigenstates of our driven Hamiltonian; as they are superpositions of fluxonium eigenstates $|k\rangle$, dephasing processes are able to cause transitions between them.

2. *RWA gate performance in the presence of dephasing noise*

We numerically evolve the master equation Eq. (24) (with the above form for the dephasing superoperator), and use the results to calculate the state-averaged gate error $\bar{\varepsilon} = 1 - \bar{F}$ of a tripod X gate (c.f. Eq. (21)). We first consider the case where all non-RWA terms in the coherent Hamiltonian are neglected, so that errors are only the result of dissipation or non-adiabaticity. The results as a function of gate time t_g are shown in Fig. 9, where the performance of the uncorrected and accelerated (SATD) pulses are compared. The simulations of the gate dynamics for all gate times t_g are done for a fixed $\Omega_0/2\pi = 100$ MHz.

The behavior here is generic and as expected. In the adiabatic regime where $t_g \gg 1/\Omega_0$, the uncorrected and SATD gates have almost identical performance. In this regime the error $\bar{\varepsilon}$ is dominated by dephasing, and grows quadratically with t_g (reflecting the quadratic loss of coherence expected from $1/f$ noise at short times). This quadratic-in-t_g error scaling continues down to small values of t_g for the SATD curve, as in this case, dephasing is the only error mechanism. In contrast, the uncorrected curve has much larger errors at short time, corresponding to non-adiabatic errors. There are special sharply defined values of t_g where these non-adiabatic errors constructively cancel; as discussed in Ref. [16], these are difficult to exploit experimentally as they require extreme fine tuning.

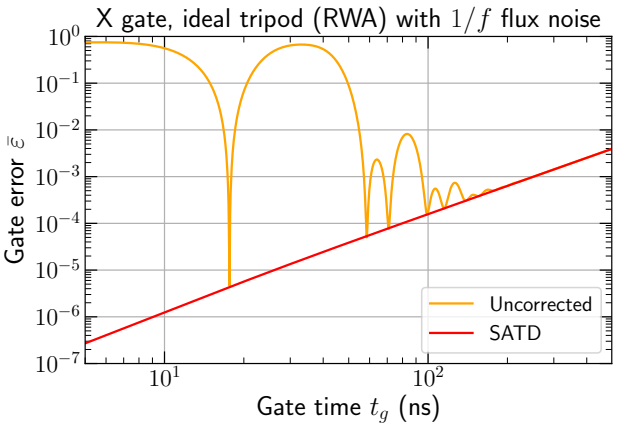


FIG. 9. State-averaged gate error $\bar{\varepsilon} = 1 - \bar{F}$ as a function of gate time t_g for ideal (4 level, RWA) tripod X gates. We compare uncorrected adiabatic pulses (orange) against the accelerated SATD pulses (red), in the presence of $1/f$ flux noise. The scale of the uncorrected pulses is set by $\Omega_0/2\pi = 100$ MHz, and the $1/f$ flux noise has a strength parameterized by $A_{\Phi_{\rm ext}} = 3\mu\Phi_0$ (yielding a qubit dephasing time $T_{\varphi,01} = 7$ μs). Fluxonium parameters match those given in Fig. 6.

VII. FULL GATE PERFORMANCE AND COMPARISON AGAINST BRUTE-FORCE DIRECT DRIVING

A. Direct driving gate and power scaling

Having investigated the impact of different error channels (non-adiabatic errors, non-RWA errors, dephasing), we are now ready to study the accelerated tripod gate with all error channels present. To properly understand the advantages of our accelerated tripod gate, it is instructive to compare performance against the simplest possible gate. This is the "direct driving" (DD) gate, where one resonantly drives the qubit transition between $|0\rangle$ and $|1\rangle$. As we are working with a T_1 protected qubit, the magnitude of the matrix element for this transition $|n_{10}|$ is small; however, for a large enough power one could in principle still achieve a given gate. To compare our tripod approach against this approach, we will want to compare not only the gate error $\bar{\varepsilon}$, but also the *power* required to achieve the gate. As we will see in what follows, in many experimental systems additional constraints limit the magnitude of pulses that can be used. This will provide a strong advantage in many regimes for the tripod gate.

We begin by writing a simple pulse shape that can be used to realize a DD gate [39]:

$$\tilde{V}_{\rm DD}(t) = \frac{\chi}{t_g |n_{01}|}\left[1 - \cos\left(\frac{2\pi t}{t_g}\right)\right]\cos\left[\int_0^t dt'\,\tilde{\omega}_{01}(t')\right], \tag{28}$$

where $\chi = \pi$ for an X gate. Similar to our accelerated adiabatic gate, we will chirp to partially correct for non-RWA errors, and hence $\tilde{\omega}_{01}(t) = \omega_{01} + \delta\varepsilon_1(t) - \delta\varepsilon_0(t)$ [c.f. Eq. (18)].

Using the definition in Eq. (30), we can calculate the RMS time-averaged voltage of the DD gate. We can also use our previous result [Eq. (16)] for the RMS time-averaged voltage of the SATD gate (using the optimal value of $\Omega_0 t_g$ discussed in Sec. IV A). These two RMS voltages are given by

$$\tilde{V}_{\mathrm{RMS,DD}} = \frac{\sqrt{3}\chi}{2|n_{01}|t_{g,\mathrm{DD}}}, \tag{29a}$$

$$\tilde{V}_{\mathrm{RMS,SATD}} = \frac{1.92\pi}{t_{g,\mathrm{SATD}}}\sqrt{\frac{\cos^2\alpha}{|n_{0e}|^2} + \frac{\sin^2\alpha}{|n_{1e}|^2} + \frac{1}{|n_{ae}|^2}}. \tag{29b}$$

In both cases, the RMS voltage scales inversely with the gate time t_g, but note the crucial dependence on matrix elements and gate type.

It follows from Eq. (29a) that an X gate is the most energy-consuming (and hence problematic) gate if one uses the DD approach. We will thus focus on this gate in what follows, and compare against our accelerated adiabatic approach. Substituting in the matrix elements for the fluxonium parameters used throughout this paper [see Fig. 6(b)], we get the $\tilde{V}_{\mathrm{RMS}}$ for the X gate ($\chi = \pi$ for DD and $\alpha = \pi/4$ for tripod) as

$$\tilde{V}_{\mathrm{RMS,DD}} = \frac{136}{t_{g,\mathrm{direct}}}, \tag{30a}$$

$$\tilde{V}_{\mathrm{RMS,SATD}} = \frac{42.1}{t_{g,\mathrm{SATD}}}. \tag{30b}$$

The equations already describe a crucial advantage of the tripod approach over direct driving: for a fixed gate time, the SATD tripod approach requires drive amplitudes that are 3.2 times smaller, corresponding to a factor of ≥ 9 savings in power. An immediate consequence is that if we fix the RMS voltage to be the same for both protocols, the SATD tripod gate will be faster.

We stress that a constraint on the size of drive amplitudes emerges naturally in typical superconducting circuits, where drives are applied via a highly-detuned microwave cavity mode which couples to the circuit as

$$\hat{H}_c = \sum_{kl} g n_{kl} |k\rangle\langle l| (\hat{a}^\dagger + \hat{a}). \tag{31}$$

Here g is the cavity-fluxonium coupling, n_{kl} is a charge matrix element, and $\hat{a}$ is the cavity photon annihilation operator. A standard constraint in such setups is that the time-averaged intracavity photon number $\bar{n}_{\mathrm{cav}}$ should not exceed some small value [58, 60] to avoid additional dissipative mechanisms; here, we will require $\bar{n}_{\mathrm{cav}} \leq 0.05$. This in turn directly constrains the RMS voltage via the relation (see Appendix G):

$$\tilde{V}^2_{\mathrm{RMS}} = 2\bar{n}_{\mathrm{cav}} g^2 \leq (0.1) g^2. \tag{32}$$

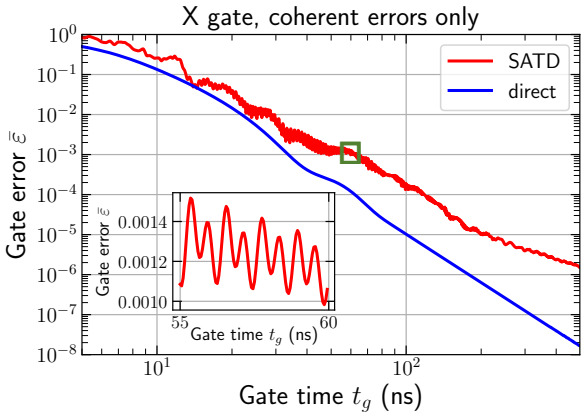


FIG. 10. State-averaged gate error $\bar{\varepsilon} = 1 - \bar{F}$ as a function of gate time t_g, for SATD tripod (red curve) and direct-driving (blue curve) realizations of a fluxonium X gate. Dissipation is not included here, but non-RWA error channels are. Inset: zoom-in of the SATD gate error. The gates are calculated using the 18 lowest energy levels of the circuit. For the SATD gate, we use the optimal uncorrected gap frequency $\Omega_0/2\pi = 1.135/t_g$ that minimizes the RMS voltage at each gate time t_g. The fluxonium parameters are as in Fig. 6.

In what follows, we will compare the SATD tripod gate against the DD gate for fixed values of g (and hence fixed maximum possible $\tilde{V}_{\mathrm{RMS}}$). We will see that this physically-motivated power constraint gives the accelerated adiabatic gate an important advantange. Further details about driving via cavity (including the driving pulse applied to the cavity and constraints that allow to ignore cavity-induced dissipation) are given in Appendix G.

B. Comparison of gate performance, coherent errors only

We first compare the accelerated tripod gate to the direct driving (DD) gate in the absence of dissipation, but including all non-RWA terms. Figure 10 shows the state-averaged gate errors $\bar{\varepsilon}$ for an X gate as a function of gate time t_g for our chosen fluxonium parameters. To mitigate non-RWA errors, the SATD protocol is implemented using the optimal value of Ω_0 (see Sec. IV A) and is frequency chirped (see Sec. IV B); the DD gate is also chirped.

We see that for both approaches, errors increase as t_g is reduced. This simply reflects the higher drive amplitudes needed at shorter times (which in turn increase non-RWA errors). For all gate times, the coherent errors are larger for SATD versus DD. This is a result of the SATD protocol using multiple drive tones, and being subject to more near-resonant non-RWA error channels (including those involving higher-energy excited states). Note that interference between its multiple drive tones causes the SATD error curve to exhibit fast, low-amplitude oscillations as

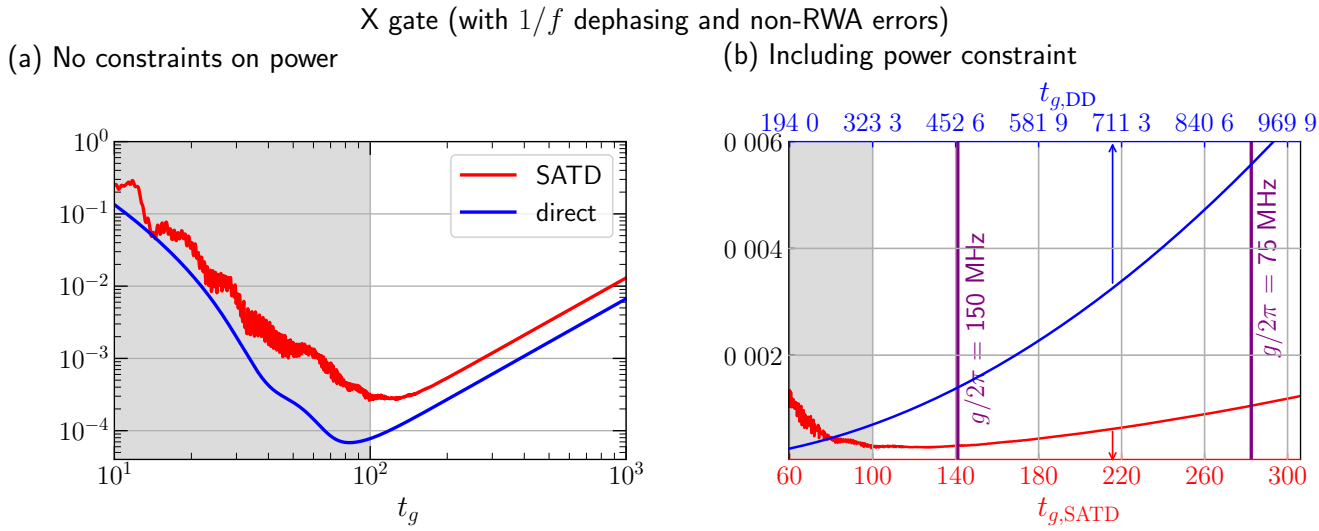


FIG. 11. Comparison between state-averaged X-gate error $\bar{\varepsilon} = 1 - \bar{F}$ for SATD (red curve) and direct-driving (blue curve) protocols in our fluxonium system, including effects of $1/f$ flux noise dephasing as well as non-RWA errors; 18 fluxonium levels are included in the simulations. Dephasing is treated as per Sec. VI B 1, with a $1/f$ flux noise amplitude $A_{\Phi_{\mathrm{ext}}} = 3\mu\Phi_0$ corresponding to the qubit dephasing time $T_{\varphi,01} = 7$ μs. (a) Log-log plot of gate error $\bar{\varepsilon}$ vs gate time t_g. For each time, the DD protocol has a smaller error but uses significantly larger pulse amplitudes. (b) Linear plot of gate error $\bar{\varepsilon}$ for both protocols at equal power levels $\tilde{V}_{\mathrm{RMS}}$ [c.f. Eq. (30)]. The shaded (unshaded) region corresponds to the time regime where the gate errors are dominated by non-RWA errors (dephasing). For the same value of $\tilde{V}_{\mathrm{RMS}}$, the gate time for the direct driving (DD) protocol is longer than the SATD protocol (hence the two distinct x axes). Vertical lines indicate the shortest gate times possible (corresponding to the maximum allowed $\tilde{V}_{\mathrm{RMS}}$, see Eq. (32)) for cavity-based driving for time-averaged cavity photon number $\bar{n}_{\mathrm{cav}} = 0.05$ and a fixed cavity-fluxonium coupling g (purple vertical lines). Fluxonium parameters are the same as in Fig. 6.

a function of t_g (see inset of Fig. 10).

The results of Fig. 10 may seem depressing. However, as we will argue in what follows, they are misleading. For a given gate time, we have already seen (c.f. Sec. VII) that DD requires considerably larger pulse amplitudes than SATD. Once we enforce power constraints associated with realistic driving via a cavity (and also include dissipation), the accelerated tripod gate will have a marked advantage.

C. Qubit gates with $1/f$ flux noise

We next compare the DD gate and accelerated SATD tripod gate including $1/f$ flux noise (modelled as per Sec. VI B 1), as well as non-RWA errors; results are shown in Fig. 11(a). For both protocols, the error is non-monotonic with gate time t_g. For short times (grey shaded region) errors are dominated by non-RWA effects and decrease with increasing t_g. For longer times dephasing dominates, causing error to increase with increasing t_g. In this latter regime, both curves increase quadratically, but the SATD curve has the higher error. This corresponds to a shorter decoherence timescale for tripod coherences (as used in SATD) versus the qubit 01 coherence (see Table I in Appendix F).

In Fig. 11(b) we replot these results in a way that now accounts for power constraints that arise when driving through a cavity (c.f. Sec.VII). Each vertical cut corresponds to a fixed value of RMS voltage (c.f. Eq. (30)); given that DD uses more power, fixing voltage thus results in a t_g for SATD that is smaller by ~ 3.2 (for our parameter choices) than that for DD (hence the different t_g axis for DD versus SATD). We thus see that once power is constrained (through $\tilde{V}_{\mathrm{RMS}}$), the accelerated tripod gate has a marked advantage over DD for $t_g \gtrsim 100$ ns; this is the regime where dephasing dominates non-RWA errors, and hence the faster speed of SATD is advantageous.

We can quantify this relative advantage by the ratio of DD and SATD errors for the same fixed $\tilde{V}_{\mathrm{RMS}}$:

$$\zeta \equiv \frac{\bar{\varepsilon}_{\mathrm{DD}}(\tilde{V}_{\mathrm{RMS}})}{\bar{\varepsilon}_{\mathrm{SATD}}(\tilde{V}_{\mathrm{RMS}})}. \tag{33}$$

In the dephasing-limited regime, the gate errors for the SATD and DD gates both exhibit a quadratic scaling with t_g. For our chosen parameters and $1/f$ flux noise strength, we find

$$\zeta = \frac{\bar{\varepsilon}_{\mathrm{DD}}}{\bar{\varepsilon}_{\mathrm{SATD}}} = 5.3. \tag{34}$$

We thus have a central conclusion of our work: our analytically designed accelerated tripod gate allows one to suppress gate errors (associated with an X gate) by

more than a factor of 5 compared to more simplistic DD approach. This conclusion holds for $\tilde{V}_{\rm RMS}$ values small enough that errors are dominated by dephasing. As discussed in Sec.VII, the maximum value of $\tilde{V}_{\rm RMS}$ is determined by the cavity-qubit coupling g and the requirement that the time-averaged intracavity photon number $\bar{n}_{\rm cav} \leq 0.05$. As shown in Fig. 11 (purple vertical lines), realistic choices of g put us squarely in this dephasing-dominated regime where we have a strong advantage.

While we have focused on comparing our accelerated tripod gate against DD, it is also worthwhile to consider its absolute performance. As shown in Fig. 11(a), we are able to achieve a fidelity of approximately 0.9997 in a gate time $t_g = 100$ ns (including $1/f$ flux noise with $A_{\Phi_{\rm ext}} = 3\mu\Phi_0$ and non-RWA error channels). This compares well to gates designed for similar systems using state-of-the-art numerical optimal control methods. For example, Ref. [28] used numerical optimal control to design gates in a fluxonium circuit operated in a T_1-protected regime. This study constrained drive power in a similar manner to our approach, but used slightly different fluxonium parameters and a lower level of flux noise ($A_{\Phi_{\rm ext}} = 1\mu\Phi_0$, 3 times smaller than in our work). Ref. [28] achieved an X gate with $\bar{F} \approx 0.996$ of in a time $t_g = 60$ ns. The main difference in parameters is that Ref. [28] used a flux bias $\Phi_{\rm ext} = 0.45\Phi_0$ (versus $\Phi_{\rm ext} = 0.17\Phi_0$ in our work), and had more isolated qubit states: $|n_{01}| = 0.01$ in their work, a factor of two smaller than in our model system.

VIII. CONCLUSIONS

In this paper we have presented a general strategy showing how analytic shortcuts-to-adiabaticity techniques can be employed even in complex multi-level systems where the rotating-wave approximation is not valid. Our strategy to mitigate non-RWA errors involved first exploiting the degeneracy of perfect STA protocols in the RWA limit, and then correcting pulses with (analytically-derived) frequency chirps. As a demonstration of our technique, we theoretically analyzed the implementation of an accelerated adiabatic "tripod" gate in a realistic multilevel superconducting circuit (a driven fluxonium qubit). We focused on parameter regimes where the qubit levels are highly isolated, yielding T_1 protection but making traditional gates more problematic. Our analysis revealed that our techniques combined with a judicious choice of system parameters yield an accelerated adiabatic gate having competitive performance. Including realistic levels of $1/f$ flux noise dephasing, we achieve a gate fidelity of 0.9997 in a gate time of 100 ns. We also showed that our approach can compare favorably against a more straightforward direct-driving approach to gates in this system. Using power constraints arising from a realistic setup as if the qubit is driven through a cavity, we found that for an X gate, our approach yields errors more than 5 times smaller than the direct driving approach.

While our test system consisted of a fluxonium qubit, our discussion and methods are very general and can be readily applied to various other architectures. This, along with the fact that the pulses are described analytically, while still providing good performance, will hopefully prove to be very useful in various quantum-control-related applications.

ACKNOWLEDGMENTS

This work is supported by the Army Research Office under Grant Number W911NF-19-1-0328. We thank Helin Zhang, Srivatsan Chakram, Brian Baker and Jens Koch for fruitful discussions. F.S. is grateful to Long B. Nguyen for enlightening discussions and pointing out Refs. [58, 60]. We acknowledge the University of Chicago Research Computing Center for support of this work.

Appendix A: Basic working of the tripod gate

In this section, we give a brief overview of the basic 4-level geometric tripod gate introduced in Refs. [3, 4] [see Fig. 2(a)]; the discussion here follows Ref. [16]. The tripod Hamiltonian $\hat{H}_0(t)$ [Eq. (6)] has two instantaneous zero-energy dark states which span the dark-state manifold and are orthogonal to $|{\rm e}\rangle$. The basic tripod gate uses the geometric evolution of states in the dark state manifold. In this manifold, there is always one (time-independent) state that is purely qubit-like:

$$|\tilde{0}\rangle = \sin(\alpha)|0\rangle - \exp(i\beta)\cos(\alpha)|1\rangle. \tag{A1}$$

The qubit state orthogonal to this dark state is

$$|\tilde{1}\rangle = \cos(\alpha)|0\rangle + \exp(i\beta)\sin(\alpha)|1\rangle. \tag{A2}$$

Expressing $\hat{H}_0(t)$ in these new qubit basis states, we have

$$\hat{H}_0(t) = \frac{1}{2}\left[\Omega_{\tilde{1}{\rm e}}(t)|\tilde{1}\rangle\langle{\rm e}| + \Omega_{\rm ae}(t)|{\rm a}\rangle\langle{\rm e}| + {\rm H.c.}\right], \tag{A3}$$

where $\Omega_{\tilde{1}{\rm e}} = \Omega_0 \sin[\theta(t)]$ and $\Omega_{\rm ae} = \Omega_0 \cos[\theta(t)]e^{i\gamma(t)}$ [Eq. (8c) of the main text]. In this new basis, the qubit state $|\tilde{1}\rangle$, together with states $|{\rm a}\rangle$ and $|{\rm e}\rangle$, form a three-level Λ system [36, 37] [see Fig. 2(b)]. As is well known, one can write a geometric phase [61, 62] onto the $|\tilde{1}\rangle$ by performing the "double STIRAP protocol", where one slowly varies control pulses to realize the cyclic adiabatic evolution $|\tilde{1}\rangle \rightarrow |{\rm a}\rangle \rightarrow |\tilde{1}\rangle$. The resulting phase is the basis of the tripod adiabatic single-qubit gate [3, 4]. One can perform an arbitrary single qubit gate in this manner, without requiring precise pulse timing, and without requiring direct couplings between the logical qubit states.

To understand the above double-STIRAP protocol in more detail, note that the dark state relevant to our Λ system (and orthogonal to $|\tilde{0}\rangle$) is:

$$|{\rm d}(t)\rangle = \cos[\theta(t)]|\tilde{1}\rangle - e^{i\gamma(t)}\sin[\theta(t)]|{\rm a}\rangle. \tag{A4}$$

The required cyclic adiabatic evolution is achieved by varying the pulse parameter $\theta(t)$, which brings the dark state $|\mathrm{d}(t)\rangle$ from the state $|\tilde{1}\rangle$ at $t=0$ to $|\mathrm{a}\rangle$ at $t=t_g/2$ and back to $|\tilde{1}\rangle$ at the final gate time $t=t_g$. To do this, we use a symmetric form for $\theta(t)$, i.e.,

$$\theta(t)=\begin{cases}\dfrac{\pi}{2}P(t/t_g) & 0\le t\le \dfrac{t_g}{2}\\ \dfrac{\pi}{2}\left[1-P\left(\dfrac{t}{t_g}-\dfrac{1}{2}\right)\right] & \dfrac{t_g}{2}<t\le t_g,\end{cases} \tag{A5}$$

where $P(x)$ is a function that increases monotonically from $P(0)=0$ to $P(1/2)=1$. Furthermore, to ensure a smooth turn on and off of the control fields, we choose a polynomial which gives $\dot{\theta}(0)=\dot{\theta}(t_g/2)=\dot{\theta}(t_g)=\ddot{\theta}(0)=\ddot{\theta}(t_g/2)=\ddot{\theta}(t_g)=0$. In particular, we use the simplest polynomial satisfying the above criteria which is given by [16]

$$P(x)=6\,(2x)^5-15\,(2x)^4+10\,(2x)^3\,. \tag{A6}$$

One also needs a non-trivial relative pulse phase $\gamma(t)$ to obtain a net Berry phase. Following Ref. [16], we use the simple form of $\gamma(t)$ as given in Eq. (11) of the main text:

$$\gamma(t)=\gamma_0\Theta\left(t-\frac{t_g}{2}\right), \tag{A7}$$

where $\Theta(t)$ is the Heaviside step function.

In the adiabatic limit $\dot{\theta}(t)/\Omega_0\to 0$, one can show that [16] the dark state $|\mathrm{d}(t)\rangle$ accumulates a geometric phase γ_0 at $t=t_g$, and qubit subspace evolves independently of the auxiliary-level subspace. The net result is a geometric single-qubit gate controlled by the pulse parameters α,β, and γ_0. It is described by the unitary in the qubit subspace $\hat{U}_{\mathrm{G},01}$ given in Eq. (12) of the main text. The full adiabatic limit unitary has the form $\hat{U}_{\mathrm{G}}=\hat{U}_{\mathrm{G},01}\oplus\hat{U}_{\mathrm{G,ae}}$ where $\hat{U}_{\mathrm{G},01}$ and $\hat{U}_{\mathrm{G,ae}}$ are the unitaries acting in the qubit and auxilary subspaces, respectively (see Ref. [16]).

Appendix B: SATD dressing for STA protocols

We review briefly how the "dressed state" approach to constructing STA protocols [16, 17] can be used to accelerate the tripod gate [16]. The general goal is to have the system follow a "dressed" version of the original adiabatic eigenstate which coincides with the original state at the start and end of the protocol. This can be achieved by using a time-dependent dressing function $\nu(t)$ which vanishes at $t=0$ and $t=t_g$. Following Refs. [16, 17], we introduce $|\mathrm{d}_\nu\rangle$, a dressed version of the original dark state $|\mathrm{d}\rangle$ [Eq. (10)]:

$$|\mathrm{d}_\nu(t)\rangle=\exp\left[-i\nu(t)\hat{J}_x\right]|\mathrm{d}(t)\rangle. \tag{B1}$$

Here $\hat{J}_x=(|\mathrm{b}_+(t)\rangle\langle\mathrm{d}(t)|+|\mathrm{b}_-(t)\rangle\langle\mathrm{d}(t)|+\mathrm{H.c.})/\sqrt{2}$, and $|\mathrm{b}_\pm(t)\rangle$ denote the bright adiabatic eigenstates of $\hat{H}_0(t)$ with energies $\pm\Omega_0/2$.

As discussed in Ref. [16], for the phase accumulated by this state to be purely geometric and equal to the adiabatic-limit geometric phase γ_0, we require $\nu(t_g/2)=0$. A particular dressing that satisfies this constraint is the "superadiabatic transitionless driving" (SATD) [16, 17] where the dressing angle ν is given by

$$\nu(t)=\nu_{\mathrm{SATD}}(t)\equiv\arctan\left[\frac{2\dot{\theta}(t)}{\Omega_0}\right]. \tag{B2}$$

Using the SATD dressing function, one can show that [16, 17] the accelerated protocol is implemented by modifying the original uncorrected pulse according to Eq. (13) of the main text.

Appendix C: Comparing SATD against simple non-adiabatic protocols

In Sec. IV A, we discussed how $\tilde{\Omega}_{\mathrm{RMS}}$, the time-averaged adiabatic gap of our Hamiltonian, was a relevant metric of our protocol's energy cost. We also discussed that by optimizing SATD, one could achieve an energy cost $\tilde{\Omega}_{\mathrm{RMS}}/2\pi=1.92/t_g$ where t_g is the gate time. In this appendix, we show that this compares surprisingly favorably to simpler, non-adiabatic pulse protocols.

First, consider the double-swap and hybrid scheme protocols (see e.g., Refs. [41, 42]). To understand these protocols, consider the effective three-level Λ system as shown in Fig. 2(b) [with the Hamiltonian given in Eq. (A3) of Appendix A]. The double-swap protocol involves two sequential swap operations for each STIRAP process where in the first half of the protocol, the pulse $\Omega_{\tilde{1}\mathrm{e}}(t)$ is first turned on with a constant value of Ω_0 for half of the time and then turned off with a simultaneous turn on of the pulse $\Omega_{\mathrm{ae}}(t)$ with a constant value of Ω_0 for another half of the time. The whole sequence is then reversed for the second half of the protocol. It follows that to generate a geometric quantum gate which cyclically evolves the state $|\tilde{1}\rangle\to|\mathrm{e}\rangle\to|\mathrm{a}\rangle\to|\mathrm{e}\rangle\to|\tilde{1}\rangle$, the protocol must be executed for a total gate time $t_g=2\pi/(\Omega_0/2)$. So, we have $\tilde{\Omega}_{\mathrm{RMS}}=\Omega_0=4\pi/t_g$. As shown in Fig. 5, this is slightly larger than the energy cost of the optimized SATD protocol.

Alternatively, consider the hybrid scheme. This has both pulses $[\Omega_{\tilde{1}\mathrm{e}}(t)$ and $\Omega_{\mathrm{ae}}(t)]$ turned on for the whole protocol with a constant value of Ω_0. For a cyclic evolution of the state, the protocol must then be carried out for a total gate time of $t_g=2\pi/(\sqrt{2}\Omega_0/2)$. Since both pulses are turned on for the whole protocol, we have $\tilde{\Omega}_{\mathrm{RMS}}=\sqrt{2}\Omega_0=4\pi/t_g$ which is the same as the double-swap protocol.

1. Proof for the lower bound of $\tilde{\Omega}_{\mathrm{RMS}}$

Finally, we establish a rigorous lower bound on $\tilde{\Omega}_{\mathrm{RMS}}$ using the quantum speed limit of Ref. [43], which generalized previous works [63, 64]. The lower bound on

$\tilde{\Omega}_{\rm RMS}$ that we are going to derive below holds for any generic dressing function $\nu(t)$ [Eq. (B1)] including the SATD dressing [Eq. (B2)]. We begin by applying the bound of Ref. [43] to the first half of our gate protocol. Letting $\hat{\rho}_0$ ($\hat{\rho}_{t_g/2}$) denote the initial system state (state after evolution for a time $t_g/2$), we have

$$\mathcal{L}^{\rm QF}(\hat{\rho}_0,\hat{\rho}_{t_g/2}) \leq \frac{1}{\hbar}\int_0^{t_g/2} dt\sqrt{\langle \hat{H}_0^2(t)\rangle - \langle \hat{H}_0(t)\rangle^2}, \quad \text{(C1)}$$

where $\mathcal{L}^{\rm QF}(\rho_0,\rho_{t_g/2}) = \arccos\left[\sqrt{F(\rho_0,\rho_{t_g/2})}\right]$, and $F(\rho_0,\rho_{t_g/2}) = {\rm Tr}\left[\sqrt{\sqrt{\rho_0}\rho_{t_g/2}\sqrt{\rho_0}}\right]$ is the Uhlmann fidelity. The LHS is the distance of initial and final states according to the quantum Fisher information metric. The RHS of this inequality is the time-integrated instantaneous energy uncertainty of our Hamiltonian $\hat{H}_0(t)$ with the accelerated protocol. Note that the symmetry of our protocol implies the LHS also bounds the energy uncertainty over the interval $(t_g/2, t_g)$.

To apply this to our accelerated protocol, note that during the first half of the evolution, the zero-energy dark state $|{\rm d}(t)\rangle\langle{\rm d}(t)|$ [Eq. (10)] evolves from the initial state $\rho_0 = |\tilde{1}\rangle\langle\tilde{1}|$ to an orthogonal state $\rho_{t_g/2} = |{\rm a}\rangle\langle{\rm a}|$. As a result, the LHS of Eq. C1 becomes $\pi/2$.

For the RHS, we note that for a generic dressing function $\nu(t)$, the accelerated Hamiltonian $\hat{H}_0(t)$ can be obtained from the adiabatic Hamiltonian [Eq. (A3)] by modifying the original pulse angle and amplitude via [17]

$$\theta(t) \to \tilde{\theta}(t) = \theta(t) + \arctan\left(\frac{\dot{\nu}(t)}{\dot{\theta}(t)/\tan[\nu(t)]}\right), \quad \text{(C2a)}$$

$$\Omega_0 \to \tilde{\Omega}(t) = \sqrt{\dot{\nu}^2(t) + \left(\frac{\dot{\theta}(t)}{\tan[\nu(t)]}\right)^2}. \quad \text{(C2b)}$$

Using the fact that the accelerated protocol guarantees that the system's state follows (at all times) the "dressed" dark state $|{\rm d}_\nu(t)\rangle$ defined in Eq. (B1), we can calculate the instantaneous energy uncertainty (RHS of Eq. (C1)) from the probability $p_\pm$ of the dressed dark state being in the instantaneous eigenstates $|{\rm b}_\pm(t)\rangle$ of $\hat{H}_0(t)$. Here, $|{\rm b}_\pm(t)\rangle$ are the bright states of $\hat{H}_0(t)$ with eigenenergies $\pm\tilde{\Omega}/2$:

$$|{\rm b}_\pm(t)\rangle = \frac{1}{\sqrt{2}}\left(\pm\sin[\tilde{\theta}(t)]|\tilde{1}\rangle \pm e^{i\gamma(t)}\cos[\tilde{\theta}(t)]|{\rm a}\rangle + |{\rm e}\rangle\right). \quad \text{(C3)}$$

The probabilities $p_\pm$ are given by

$$\begin{aligned} p_\pm(t) &= |\langle {\rm d}_\nu(t)|{\rm b}_\pm(t)\rangle|^2 \\ &= \frac{1}{2}\left[\left(\frac{\dot{\nu}^2(t)}{\tilde{\Omega}^2(t)}\right)\cos^2[\nu(t)] + \sin^2[\nu(t)]\right]. \end{aligned} \quad \text{(C4)}$$

Using Eq. (C4), we calculate the instantaneous value of $\langle \hat{H}_0(t)\rangle$ and $\langle \hat{H}_0^2(t)\rangle$ as

$$\langle \hat{H}_0(t)\rangle = \frac{\tilde{\Omega}(t)}{2}(p_+ - p_-) = 0, \quad \text{(C5a)}$$

$$\langle \hat{H}_0^2(t)\rangle = \frac{[\tilde{\Omega}(t)]^2}{4}(p_+ + p_-) = \frac{1}{4}\left\{\dot{\nu}^2(t) + \frac{\dot{\theta}^2(t)}{\sec^2[\nu(t)]}\right\}. \quad \text{(C5b)}$$

Substituting Eq. (C5) into Eq. (C1), we then have

$$\begin{aligned} \frac{\pi}{2} &\leq \frac{1}{2}\int_0^{t_g/2} dt\sqrt{\dot{\nu}^2(t) + \frac{\dot{\theta}^2(t)}{\sec^2[\nu(t)]}} \\ &\leq \frac{1}{2}\sqrt{\int_0^{t_g/2} dt}\sqrt{\int_0^{t_g/2} dt\left(\dot{\nu}^2(t) + \frac{\dot{\theta}^2(t)}{1+\tan^2[\nu(t)]}\right)} \\ &< \frac{t_g}{4}\sqrt{\frac{2}{t_g}\int_0^{t_g/2} dt\left(\dot{\nu}^2(t) + \frac{\dot{\theta}^2(t)}{\tan^2[\nu(t)]}\right)} = \frac{t_g}{4}\tilde{\Omega}_{\rm RMS}. \end{aligned} \quad \text{(C6)}$$

Note that in going to the second line of Eq. (C6), we have used the Cauchy-Schwarz inequality and the relation $\sec^2(x) = 1 + \tan^2(x)$. From Eq. (C6), we can write the bound for $\tilde{\Omega}_{\rm RMS}$ as

$$\tilde{\Omega}_{\rm RMS} > 2\pi/t_g. \quad \text{(C7)}$$

Appendix D: Optimal fluxonium parameter regime for a tripod gate in the T_1 protected regime

To get a tripod gate with small non-RWA errors as well as a qubit with a long T_1 coherence time, we use the following criteria in choosing the fluxonium circuit parameters:

1. The ground states, i.e., low-lying energy levels of the tripod which are labeled by $|0\rangle$ $|1\rangle$ and $|{\rm a}\rangle$, should be well isolated from each other and be non-degenerate. The requirement of strong isolation is necessary to obtain a T_1 protected qubit. On the other hand, the nondegeneracy of ground states is required to ensure sufficiently large detuning of the spurious crosstalk transitions from the driving frequencies in order to reduce the coherent errors due to crosstalks.

2. The charge matrix elements coupling the excited state to the ground states of the tripod system should be large and have the same order of magnitude. This requirement helps to minimize coherent errors arising from non-RWA processes.

The criterion (1) requires us to choose circuit parameters that satisfy $E_J \gg E_C$ and $E_L \ll E_J$ (for well-localized ground states) as well as $0 < \Phi_{\rm ext} \leq \Phi_0/4$ (to lift the degeneracy of the ground states). The criterion

(2), on the other hand, requires that we choose the excited state $|\mathrm{e}\rangle$ of the tripod gate to be the first excited state of the central well that is delocalized over the potential wells where the ground states $|0\rangle$, $|1\rangle$ and $|\mathrm{a}\rangle$ reside, but yet somewhat separated from the much more densely-spaced higher energy levels. This implies that the state $|\mathrm{e}\rangle$ must lie in the vicinity of the top edge of the cosine potential, which requires $\sqrt{8E_CE_J} \gg 2E_J$. It is clear that the criteria (1) and (2) cannot be satisfied simultaneously in a standard fluxonium circuit. That forces us to seek a balanced parameter set, ensuring that we can both end up with a long-lived qubit, as well as with a tripod that allows for a high-fidelity SATD gate.

In order to achieve this balance, we initially do a numerical search over experimentally realizable circuit parameters, enforcing conditions on the energy-level structure that do not strongly violate our desired selection criteria outlined in points (1) and (2) above. These include, for example, requiring that the charge matrix element between levels $|0\rangle$ and $|1\rangle$ to be small (which maximizes T_1), and that the tripod transition energies $\omega_{j\mathrm{e}}$, with $j = 0, 1, \mathrm{a}$, are not degenerate (to minimize effects of crosstalk), and so on. After applying this procedure, we end up with a much reduced parameter space, which in turn is used to run a more focused search that optimizes over the fidelity of an actual SATD gate (although without the frequency chirping). Taking all of the above into account, we end up with a final choice of parameters: $E_L/h = 0.063$ GHz, $E_J/h = 9.19$ GHz, $E_C/h = 2$ GHz and $\Phi_{\mathrm{ext}} = 0.17\Phi_0$. We stress that the small inductive energy of our fluxonium puts our E_L in same regime as the Blochnium device that was recently realized experimentally [49]. The corresponding energy level structure along with the potential energy landscape, are shown in Fig. 1. Due to the small matrix element between the qubit levels $|0\rangle$ and $|1\rangle$ ($|n_{01}| = 0.02$), our qubit is not T_1 limited, and due the positioning of the excited level $|\mathrm{e}\rangle$, a tripod with relatively strong tripod matrix elements ($|n_{0\mathrm{e}}| = 0.27$, $|n_{1\mathrm{e}}| = 0.46$ and $|n_{\mathrm{ae}}| = 0.16$) can be realized.

Appendix E: Accounting for a smooth turn on and off of the pulse at the beginning and end of the protocol

Since realistic pulses are off in the beginning and the end of the protocol, we sandwich the pulse in Eq. (8) by a ramp time t_{ramp} during which the pulse $\tilde{\Omega}_{\mathrm{ae}}(t)$ is smoothly turned on (off) at the beginning (end) of the protocol. The full driving pulses can then be written as smooth piecewise continuous functions which can be separated in three time regions [(I): $0 \le t < t_{\mathrm{ramp}}$, (II): $t_{\mathrm{ramp}} \le t \le t_g + t_{\mathrm{ramp}}$, (III): $t_g + t_{\mathrm{ramp}} < t \le t_g + 2t_{\mathrm{ramp}}$)] as

$$\frac{\tilde{\Omega}_{0\mathrm{e}}(t)}{\Omega_0} = \begin{cases} 0 & \text{(I)}, \\ \cos\alpha \left\{ \sin[\theta(t_-)] + \dfrac{\cos[\theta(t_-)]\ddot{\theta}(t_-)}{\dot{\theta}^2(t_-) + \Omega_0^2/4} \right\} & \text{(II)}, \\ 0 & \text{(III)}, \end{cases} \tag{E1a}$$

$$\frac{\tilde{\Omega}_{1\mathrm{e}}(t)}{\Omega_0} = \begin{cases} 0 & \text{(I)}, \\ e^{-i\beta}\sin\alpha \left\{ \sin[\theta(t_-)] + \dfrac{\cos[\theta(t_-)]\ddot{\theta}(t_-)}{\dot{\theta}^2(t_-) + \Omega_0^2/4} \right\} & \text{(II)}, \\ 0 & \text{(III)}, \end{cases} \tag{E1b}$$

$$\frac{\tilde{\Omega}_{\mathrm{ae}}(t)}{\Omega_0} = \begin{cases} P\left(\dfrac{t}{2t_{\mathrm{ramp}}}\right) & \text{(I)}, \\ e^{-i\gamma t_-} \left\{ \cos[\theta(t_-)] - \dfrac{\sin[\theta(t_-)]\ddot{\theta}(t_-)}{\dot{\theta}^2(t_-) + \Omega_0^2/4} \right\} & \text{(II)}, \\ 1 - P\left(\dfrac{t_- - t_g}{2t_{\mathrm{ramp}}}\right) & \text{(III)}, \end{cases} \tag{E1c}$$

where $t_- = t - t_{\mathrm{ramp}}$. Specifically, we choose $t_{\mathrm{ramp}} = 0.01t_g$ to be short enough compared to the overall pulse length such that it will not significantly affect the whole dynamics but long enough such that there is no sharp jump in the pulse when it is turned on or off. Note that, in Eq. (E1c) we have used the function $P(x)$ given in Eq. (A6) to ensure a smooth turn on and off of the pulse $\tilde{\Omega}_{\mathrm{ae}}(t)$ during a time duration t_{ramp} at the beginning and end of the protocol. The time profile of the driving pulse with the inclusion of the ramp function is shown in Fig. 7(a).

Appendix F: Approximating non-Markovian noise with Markovian dynamics

In this Appendix, we show that approximating non-Markovian dynamics of $1/f$ noise with Markovian master equation cannot capture the decay rate of all coherences correctly. In the main text, we construct our Linblad master equation by picking the decay rate Γ_k of the coherences $\rho_{k1} \equiv \langle k|\hat{\rho}|1\rangle$ (coherences which involve the reference level, i.e., the qubit state $|1\rangle$)) such that the final decay of the coherences ρ_{k1}, $\forall k \neq 1$ at the gate time t_g is the same for our Markovian dynamics as it would be following the non-Markovian, Gaussian-lineshape decay with envelope $\exp\left[-(t/T_{\varphi,kl})^2\right]$ where $T_{\varphi,kl}$ is the free induction decay time given by Eq. (26). (Note that we can equivalently choose the qubit state $|0\rangle$ as the reference level as both choices will result in the same value of state-averaged fidelities.) While the decay of ρ_{k1} coherences can be captured correctly, the dephasing times of coherences that do not involve the reference level (ρ_{kl},

$\forall k, l \neq 1$) in our Markovian dynamics are in general not the same as the free induction decay times calculated from Eq. (26). To see this, we can write the Lindblad master equation [Eq. (24)] in terms of the density matrix elements as

$$\begin{aligned}\dot{\rho}_{kl}(t) &= -i[\hat{H}(t), \hat{\rho}(t)]_{kl} \\ &\quad - \left(\mathrm{sgn}\left(\frac{\partial \varepsilon_k}{\partial \Phi_{\mathrm{ext}}}\right)\sqrt{\Gamma_k} - \mathrm{sgn}\left(\frac{\partial \varepsilon_l}{\partial \Phi_{\mathrm{ext}}}\right)\sqrt{\Gamma_l}\right)^2 \rho_{kl} \\ &= -i[\hat{H}(t), \hat{\rho}(t)]_{kl} - \frac{t_g}{\left(T^{\mathrm{eff}}_{\varphi,kl}\right)^2}\rho_{kl}, \end{aligned} \tag{F1}$$

where we have identified

$$\frac{1}{T^{\mathrm{eff}}_{\varphi,kl}} \equiv \left|\mathrm{sgn}\left(\frac{\partial \varepsilon_k}{\partial \Phi_{\mathrm{ext}}}\right)\frac{1}{T_{\varphi,k1}} - \mathrm{sgn}\left(\frac{\partial \varepsilon_l}{\partial \Phi_{\mathrm{ext}}}\right)\frac{1}{T_{\varphi,l1}}\right|. \tag{F2}$$

We can see that the effective dephasing time $T^{\mathrm{eff}}_{\varphi,kl}$ [Eq. (F2)], $\forall k, l \neq 1$ calculated from the Lindblad master equation is in general not the same as the free induction dephasing time $T_{\varphi,kl}$ calculated from Eq. (26). In fact, as shown in Table I, our approach if anything overestimates the dominant dephasing processes within the tripod subspace, e.g., $T^{\mathrm{eff}}_{\varphi,\mathrm{a0}}$ and $T^{\mathrm{eff}}_{\varphi,\mathrm{ae}}$.

kl	$T_{\varphi,kl}$ (μs)	$T^{\mathrm{eff}}_{\varphi,kl}$ (μs)
01	7.03	7.03
a1	6.97	6.97
e1	53.43	53.43
a0	3.50	1.75
e0	8.09	17.31
ae	6.16	3.76

TABLE I. Dephasing times for transitions between computational states $|k\rangle$ and $|l\rangle$ calculated using two different methods: (Middle column): directly from the free induction decay formula [Eq. (26)] and (Right column): indirectly from the Lindblad Master equation [Eq. (F2)].

Appendix G: Indirect qubit driving through a coupled cavity

In order to provide Purcell protection, in practice, superconducting qubits are often driven indirectly through a coupled cavity. In this Appendix, we discuss the details of such driving of our fluxonium qubit, especially in the context where the drive power is constrained such that additional dissipative mechanisms due to the cavity can be avoided. In the following, we show how the driving field applied to the qubit-coupled cavity is related to the driving field seen by the qubit $V(t)$ [Eq. (2) of the main text].

1. Driving field for the cavity

In this subsection we derive an explicit relationship between the field that drives the cavity $u(t)$ and the voltage $V(t)$ [Eq. (2)] seen by the qubit. We begin by writing the Hamiltonian for the driven cavity-coupled fluxonium as

$$\begin{aligned}\hat{H}_{\mathrm{JC}}(t) &= \sum_k \varepsilon_k |k\rangle\langle k| + \omega_{\mathrm{cav}}\hat{a}^\dagger\hat{a} + \sum_{k,l} g n_{kl}|k\rangle\langle l|(\hat{a}^\dagger + \hat{a}) \\ &\quad + u(t)(\hat{a}^\dagger + \hat{a}), \end{aligned} \tag{G1}$$

where ε_k, $|k\rangle$ are the fluxonium eigenenergies and eigenstates, respectively, and ω_{cav} is the cavity frequency. The qubit-cavity coupling strength is g, while $n_{kl} = \langle k|\hat{n}|l\rangle$ is the fluxonium charge matrix element. The driving field on the cavity is $u(t)$ and the operators $\hat{a}^\dagger$ and $\hat{a}$ are the cavity photon creation and annihilation operators, respectively.

We consider operating the cavity-qubit system in the dispersive regime where $gn_{kl} \ll |\varepsilon_l - \varepsilon_k - \omega_{\mathrm{cav}}|$, and applying a highly off-resonant drive to limit the cavity photon population. We also consider the drive strength to be weak enough that the cavity photon number is small so to avoid cavity-induced dissipations. In this weak driving power regime, we can estimate the relationship between the applied cavity field $u(t)$ and the driving field on the qubit $V(t)$, by first solving for the classical cavity field independently (i.e., in the limit $g \to 0$). To this end, we model the cavity as a stand-alone, damped, driven harmonic oscillator. The Heisenberg equation of motion for the cavity-photon annihilation $\hat{a}(t)$ operator can be written as [65]

$$\begin{aligned}\dot{\hat{a}}(t) &= i[\hat{H}_{\mathrm{JC}}(t), \hat{a}(t)] - \frac{\kappa}{2}\hat{a}(t) - \sqrt{\kappa}\hat{b}_{\mathrm{in}}(t) \\ &= -i\omega_{\mathrm{cav}}\hat{a}(t) - iu(t) - \frac{\kappa}{2}\hat{a}(t) - \sqrt{\kappa}\hat{b}_{\mathrm{in}}(t). \end{aligned} \tag{G2}$$

Here κ is the cavity photon decay rate (due to the photon leakage to the bath), $\hat{b}_{\mathrm{in}}(t)$ is the standard bath annihilation operator which represents the noise [65]. The equation of motion for $\hat{a}^\dagger(t)$ can be obtained by taking the Hermitian conjugate of Eq. (G2). By treating the cavity field classically and writing the equation of motion in terms of the mean value of the photon field displacement $x = \langle \hat{a} + \hat{a}^\dagger\rangle/\sqrt{2}$, we have

$$\ddot{x}(t) = -\left(\omega^2_{\mathrm{cav}} + \frac{\kappa^2}{4}\right)x(t) - \kappa\dot{x}(t) - \sqrt{2}\omega_{\mathrm{cav}}u(t), \tag{G3}$$

where we have used $\langle \hat{b}_{\mathrm{in}}(t)\rangle = \langle \hat{b}^\dagger_{\mathrm{in}}(t)\rangle = 0$. The solution of Eq. (G3) is

$$x(t) = x_p(t) + x_h(t), \tag{G4}$$

where

$$x_h(t) = Ae^{-\frac{\kappa}{2}t}\sin(\omega_{\mathrm{cav}}t + \phi), \tag{G5}$$

is the homogeneous solution and $x_p(t)$ is the inhomogeneous solution. To solve for $x_p(t)$, we first define the Fourier transform of $x(\omega)$ and $u(\omega)$ as

$$x(\omega) = \int_{-\infty}^{\infty} x(t)e^{-i\omega t}dt,$$
$$u(\omega) = \int_{-\infty}^{\infty} u(t)e^{-i\omega t}dt, \tag{G6}$$

and their inverse Fourier transform as

$$x(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} x(\omega)e^{i\omega t}d\omega,$$
$$u(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} u(\omega)e^{i\omega t}d\omega. \tag{G7}$$

Taking the Fourier transform of Eq. (G3), we have

$$\left[-\omega^2 + i\kappa\omega + \omega_{\rm cav}^2 + \frac{\kappa^2}{4}\right]x(\omega) = -\sqrt{2}\omega_{\rm cav}u(\omega), \tag{G8}$$

which gives

$$x(\omega) = \frac{\sqrt{2}\omega_{\rm cav}u(\omega)}{\omega^2 - i\kappa\omega - (\omega_{\rm cav}^2 + \kappa^2/4)}. \tag{G9}$$

We can relate the mean-photon displacement $x(t)$ with the qubit driving voltage $V(t)$ by replacing the cavity degree of freedom in the cavity-qubit coupling Hamiltonian by its classical value, i.e.,

$$\hat{H}_{\rm c} \approx \sum_{kl} g\langle\hat{a}^\dagger + \hat{a}\rangle n_{kl}|k\rangle\langle l|$$
$$= \sqrt{2}gx(t)\sum_{kl} n_{kl}|k\rangle\langle l|. \tag{G10}$$

Comparing Eq. (G10) with the qubit driving term in the Hamiltonian of Eq. (1), we can identify

$$\sqrt{2}gx(t) = V(t), \tag{G11}$$

where in the frequency domain, it can be written as

$$\sqrt{2}gx(\omega) = \int_{-\infty}^{\infty} V(t')e^{-i\omega t'}dt'. \tag{G12}$$

Substituting Eq. (G9) into Eq. (G12), we have

$$u(\omega) = \frac{(\omega^2 - i\kappa\omega - (\omega_{\rm cav}^2 + \kappa^2/4))}{2g\omega_{\rm cav}}\int_{-\infty}^{\infty} V(t')e^{-i\omega t'}dt'. \tag{G13}$$

Finally, taking the inverse Fourier transform, we arrive at an expression for the time-domain form of the cavity pulse $u(t)$ that is required to generate a qubit drive $V(t)$ as

$$u(t) = -\frac{1}{2g\omega_{\rm cav}}\left[\frac{d^2}{dt^2} + \kappa\frac{d}{dt} + \omega_{\rm cav}^2 + \frac{\kappa^2}{4}\right]V(t). \tag{G14}$$

2. Relation between the RMS voltage and the time-averaged cavity photon number

In this section, we establish the relationship between the RMS voltage $\tilde{V}_{\rm RMS}$ of the qubit driving field and the time-averaged cavity photon number $\bar{n}_{\rm cav}$ where

$$\bar{n}_{\rm cav} = \frac{1}{t_g}\int_0^{t_g}\langle\hat{a}^\dagger\hat{a}\rangle dt$$
$$\approx \frac{1}{t_g}\int_0^{t_g}\langle\hat{a}^\dagger\rangle\langle\hat{a}\rangle dt = \frac{1}{t_g}\int_0^{t_g} dt|\eta(t)|^2. \tag{G15}$$

Here, we treat the photon field classically, which allows us to replace $\langle\hat{a}^\dagger\rangle$ and $\langle\hat{a}^\dagger\rangle$ by the classical field amplitudes $\eta^*(t)$ and $\eta(t)$, respectively. Using this classical field in Eq. (G10), we can identify the qubit driving voltage as

$$V(t) = 2g{\rm Re}[\eta(t)]. \tag{G16}$$

The RMS voltage of the qubit driving field is then given by

$$\tilde{V}_{\rm RMS} = \sqrt{\frac{1}{t_g}\int_0^{t_g}|V(t)|^2}$$
$$= g\sqrt{2\bar{n}_{\rm cav}}, \tag{G17}$$

where in evaluating the second line, we have used the relation

$$\frac{1}{t_g}\int_0^{t_g}[{\rm Re}(\eta(t))]^2 = \sum_j \frac{1}{t_g}\int_0^{t_g}|\eta_j(t)|^2\cos^2(\omega_j t + \phi_j)$$
$$\simeq \frac{1}{2t_g}\int_0^{t_g}|\eta(t)|^2 = \frac{\bar{n}_{\rm cav}}{2}. \tag{G18}$$

Without loss of generality, in Eq. (G18), we have written the photon field as a multicomponent field, i.e., $\eta(t) = \sum_j \eta_j e^{-i(\omega_j t+\phi_j)}$. To avoid cavity-induced dissipation mechanisms, it is preferable to have a small cavity photon number [58, 60], i.e., $\bar{n}_{\rm cav} = \langle\hat{a}^\dagger\hat{a}\rangle \ll 1$; for our simulations in the main text, we specifically set $\bar{n}_{\rm cav} = 0.05$. This constrains the maximum RMS voltage for a fixed cavity-qubit coupling strength g as shown in Eq. (32) of the main text. Each vertical cut of the plot in Fig. 11(b) corresponds to different maximum allowed values of $\tilde{V}_{\rm RMS}$ where the purple vertical lines show explicitly two different fixed values of the cavity-qubit coupling strength g corresponding to two different maximum allowed values of $\tilde{V}_{\rm RMS}$.

3. Effects of the cavity on qubit's T_1 and T_2 times

In this section we briefly outline the effects of the coupled cavity on the coherence times of the qubit, and show that they do not limit the performance of our gates. In particular, we consider a regime where the cavity photon decay rate $\kappa \ll g^2/|\Delta_{kl,\pm}^{\rm cav}|$, $\forall k, l$ where

$\Delta^{\rm cav}_{kl,\pm} = \varepsilon_l - \varepsilon_k \pm \omega_{\rm cav}$ is the detuning of the cavity frequency $\omega_{\rm cav}$ from the $|k\rangle \leftrightarrow |l\rangle$ energy transition of the qubit. The first noise channel we consider is the Purcell relaxation time, which we denote as $T_{1,\rm cav}$ and can be calculated as [66]

$$(T_{1,\rm cav})_{jk} = (\Delta^{\rm cav}_{kl,\pm})^2/(\kappa g^2 |n_{kl}|^2). \tag{G19}$$

Since the all the relevant transitions are highly detuned from the cavity frequency, we expect this kind of process to be weak. Another relevant noise channel is due to photon shot noise, and leads to pure dephasing. We denote the corresponding time scale as $T_{2,\rm cav}$, which can be approximated by [67, 68]

$$T_{2,\rm cav} = 1/(\kappa \bar{n}_{\rm cav}). \tag{G20}$$

We stress that Eq. (G20) is valid in the limit of small cavity photon decay relative to all the relevant dispersive shifts ($\kappa \ll g^2/|\Delta^{\rm cav}_{kl,\pm}|$), a regime that we are interested in. Moreover, we also consider operating in the dispersive limit ($g n_{kl} \ll |\Delta^{\rm cav}_{kl,\pm}|, \forall k, l$) and small-cavity-photon-number regime ($\bar{n}_{\rm cav} \ll 1$) such that the cavity response can be treated independently of the qubit [see Sec. G 1]. A parameter set that fulfills the above criteria is $\omega_{\rm cav}/2\pi = 2$ GHz (chosen to be sufficiently detuned from all of the transition frequencies in our fluxonium [Fig. 1]), $\kappa/2\pi = 10$ kHz, $g/2\pi = 250$ MHz and $\bar{n}_{\rm cav} = 0.05$. Substituting the above parameters in Eqs. (G19) and (G20), we obtain $(T_{1,\rm cav})_{01} \approx 0.15$ s and $T_{2,\rm cav} \approx 0.32$ ms. Since the cavity-induced relaxation $(T_{1,\rm cav})_{01}$ and dephasing times $T_{2,\rm cav}$ are much larger than the the dephasing times due to the $1/f$ flux noise (see Table I of Appendix F), we can ignore the effect of cavity-induced dissipation on the gate dynamics.

[1] P. Zanardi and M. Rasetti, Holonomic quantum computation, Phys. Lett. A **264**, 94 (1999).
[2] J. Pachos, P. Zanardi, and M. Rasetti, Non-abelian berry connections for quantum computation, Phys. Rev. A **61**, 010305 (1999).
[3] L.-M. Duan, J. I. Cirac, and P. Zoller, Geometric manipulation of trapped ions for quantum computation, Science **292**, 1695 (2001).
[4] Z. Kis and F. Renzoni, Qubit rotation by stimulated raman adiabatic passage, Phys. Rev. A **65**, 032318 (2002).
[5] L. Faoro, J. Siewert, and R. Fazio, Non-abelian holonomies, charge pumping, and quantum computation with josephson junctions, Phys. Rev. Lett. **90**, 028301 (2003).
[6] P. Solinas, P. Zanardi, N. Zanghì, and F. Rossi, Holonomic quantum gates: A semiconductor-based implementation, Phys. Rev. A **67**, 062315 (2003).
[7] A. Frees, S. Mehl, J. K. Gamble, M. Friesen, and S. Coppersmith, Adiabatic two-qubit gates in capacitively coupled quantum dot hybrid qubits, npj Quantum Information **5**, 73 (2019).
[8] H. Wu, E. M. Gauger, R. E. George, M. Möttönen, H. Riemann, N. V. Abrosimov, P. Becker, H.-J. Pohl, K. M. Itoh, M. L. W. Thewalt, and J. J. L. Morton, Geometric phase gates with adiabatic control in electron spin resonance, Phys. Rev. A **87**, 032326 (2013).
[9] K. Toyoda, K. Uchida, A. Noguchi, S. Haze, and S. Urabe, Realization of holonomic single-qubit operations, Phys. Rev. A **87**, 052307 (2013).
[10] Y.-Y. Huang, Y.-K. Wu, F. Wang, P.-Y. Hou, W.-B. Wang, W.-G. Zhang, W.-Q. Lian, Y.-Q. Liu, H.-Y. Wang, H.-Y. Zhang, L. He, X.-Y. Chang, Y. Xu, and L.-M. Duan, Experimental realization of robust geometric quantum gates with solid-state spins, Phys. Rev. Lett. **122**, 010503 (2019).
[11] M. Demirplak and S. A. Rice, Adiabatic population transfer with control fields, The Journal of Physical Chemistry A **107**, 9937 (2003).
[12] M. Demirplak and S. A. Rice, Assisted adiabatic passage revisited, The Journal of Physical Chemistry B **109**, 6838 (2005).
[13] M. V. Berry, Transitionless quantum driving, Journal of Physics A: Mathematical and Theoretical **42**, 365303 (2009).
[14] S. Ibáñez, X. Chen, E. Torrontegui, J. G. Muga, and A. Ruschhaupt, Multiple schrödinger pictures and dynamics in shortcuts to adiabaticity, Phys. Rev. Lett. **109**, 100403 (2012).
[15] D. Guéry-Odelin, A. Ruschhaupt, A. Kiely, E. Torrontegui, S. Martínez-Garaot, and J. G. Muga, Shortcuts to adiabaticity: Concepts, methods, and applications, Rev. Mod. Phys. **91**, 045001 (2019).
[16] H. Ribeiro and A. A. Clerk, Accelerated adiabatic quantum gates: Optimizing speed versus robustness, Phys. Rev. A **100**, 032323 (2019).
[17] A. Baksic, H. Ribeiro, and A. A. Clerk, Speeding up adiabatic quantum state transfer by using dressed states, Phys. Rev. Lett. **116**, 230503 (2016).
[18] J. Zhang, T. H. Kyaw, D. Tong, E. Sjöqvist, and L.-C. Kwek, Fast non-abelian geometric gates via transitionless quantum driving, Scientific reports **5**, 18414 (2015).
[19] X.-K. Song, H. Zhang, Q. Ai, J. Qiu, and F.-G. Deng, Shortcuts to adiabatic holonomic quantum computation in decoherence-free subspace with transitionless quantum driving algorithm, New Journal of Physics **18**, 023001 (2016).
[20] T. Wang, Z. Zhang, L. Xiang, Z. Jia, P. Duan, W. Cai, Z. Gong, Z. Zong, M. Wu, J. Wu, *et al.*, The experimental realization of high-fidelity 'shortcut-to-adiabaticity'quantum gates in a superconducting xmon qubit, New Journal of Physics **20**, 065003 (2018).
[21] T. Yan, B.-J. Liu, K. Xu, C. Song, S. Liu, Z. Zhang, H. Deng, Z. Yan, H. Rong, K. Huang, M.-H. Yung, Y. Chen, and D. Yu, Experimental realization of nonadiabatic shortcut to non-abelian geometric gates, Phys. Rev. Lett. **122**, 080501 (2019).
[22] A. C. Santos, A. Nicotina, A. M. Souza, R. S. Sarthour, I. S. Oliveira, and M. S. Sarandy, Optimizing nmr quantum information processing via generalized transitionless quantum driving, EPL (Europhysics Letters) **129**, 30008

(2020).
[23] E. Sjöqvist, D.-M. Tong, L. M. Andersson, B. Hessmo, M. Johansson, and K. Singh, Non-adiabatic holonomic quantum computation, New Journal of Physics **14**, 103035 (2012).
[24] V. E. Manucharyan, J. Koch, L. I. Glazman, and M. H. Devoret, Fluxonium: Single cooper-pair circuit free of charge offsets, Science **326**, 113 (2009).
[25] N. Earnest, S. Chakram, Y. Lu, N. Irons, R. K. Naik, N. Leung, L. Ocola, D. A. Czaplewski, B. Baker, J. Lawrence, J. Koch, and D. I. Schuster, Realization of a Λ system with metastable states of a capacitively shunted fluxonium, Phys. Rev. Lett. **120**, 150504 (2018).
[26] L. B. Nguyen, Y.-H. Lin, A. Somoroff, R. Mencia, N. Grabon, and V. E. Manucharyan, High-coherence fluxonium qubit, Phys. Rev. X **9**, 041041 (2019).
[27] 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).
[28] 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).
[29] M. Demirplak and S. A. Rice, On the consistency, extremal, and global properties of counterdiabatic fields, The Journal of chemical physics **129**, 154111 (2008).
[30] B.-H. Huang, Y.-H. Chen, Q.-C. Wu, J. Song, and Y. Xia, Fast generating greenberger–horne–zeilinger state via iterative interaction pictures, Laser Physics Letters **13**, 105202 (2016).
[31] Y.-H. Kang, Y.-H. Chen, Q.-C. Wu, B.-H. Huang, J. Song, and Y. Xia, Fast generation of w states of superconducting qubits with multiple schrödinger dynamics, Scientific reports **6**, 36737 (2016).
[32] X.-K. Song, Q. Ai, J. Qiu, and F.-G. Deng, Physically feasible three-level transitionless quantum driving with multiple schrödinger dynamics, Phys. Rev. A **93**, 052324 (2016).
[33] J.-L. Wu, S.-L. Su, X. Ji, and S. Zhang, Superadiabatic scheme for optimizing the fast generation of tree-type 3d entanglement, Annals of Physics **386**, 34 (2017).
[34] E. Boyers, M. Pandey, D. K. Campbell, A. Polkovnikov, D. Sels, and A. O. Sushkov, Floquet-engineered quantum state manipulation in a noisy qubit, Phys. Rev. A **100**, 012341 (2019).
[35] P. W. Claeys, M. Pandey, D. Sels, and A. Polkovnikov, Floquet-engineering counterdiabatic protocols in quantum many-body systems, Phys. Rev. Lett. **123**, 090602 (2019).
[36] K. Bergmann, H. Theuer, and B. W. Shore, Coherent population transfer among quantum states of atoms and molecules, Rev. Mod. Phys. **70**, 1003 (1998).
[37] N. V. Vitanov, A. A. Rangelov, B. W. Shore, and K. Bergmann, Stimulated raman adiabatic passage in physics, chemistry, and beyond, Rev. Mod. Phys. **89**, 015006 (2017).
[38] B. B. Zhou, A. Baksic, H. Ribeiro, C. G. Yale, F. J. Heremans, P. C. Jerger, A. Auer, G. Burkard, A. A. Clerk, and D. D. Awschalom, Accelerated quantum control using superadiabatic dynamics in a solid-state lambda system, Nature Physics **13**, 330 (2017).
[39] T. F. Roque, A. A. Clerk, and H. Ribeiro, Engineering fast high-fidelity quantum operations with constrained interactions, arXiv:2003.12096 (2020).
[40] H. Ribeiro, A. Baksic, and A. A. Clerk, Systematic magnus-based approach for suppressing leakage and nonadiabatic errors in quantum dynamics, Phys. Rev. X **7**, 011021 (2017).
[41] Y.-D. Wang and A. A. Clerk, Using dark modes for high-fidelity optomechanical quantum state transfer, New Journal of Physics **14**, 105010 (2012).
[42] Y.-D. Wang and A. A. Clerk, Using interference for high fidelity quantum state transfer in optomechanics, Phys. Rev. Lett. **108**, 153603 (2012).
[43] D. P. Pires, M. Cianciaruso, L. C. Céleri, G. Adesso, and D. O. Soares-Pinto, Generalized geometric quantum speed limits, Phys. Rev. X **6**, 021031 (2016).
[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] J. M. Gambetta, F. Motzoi, S. T. Merkel, and F. K. Wilhelm, Analytic control methods for high-fidelity unitary operations in a weakly nonlinear oscillator, Phys. Rev. A **83**, 012308 (2011).
[46] N. Wiebe and N. S. Babcock, Improved error-scaling for adiabatic quantum evolutions, New Journal of Physics **14**, 013024 (2012).
[47] Y.-H. Lin, L. B. Nguyen, N. Grabon, J. San Miguel, N. Pankratova, and V. E. Manucharyan, Demonstration of protection of a superconducting qubit from energy decay, Phys. Rev. Lett. **120**, 150503 (2018).
[48] J. Koch, V. Manucharyan, M. H. Devoret, and L. I. Glazman, Charging effects in the inductively shunted josephson junction, Phys. Rev. Lett. **103**, 217004 (2009).
[49] I. V. Pechenezhskiy, R. A. Mencia, L. B. Nguyen, Y.-H. Lin, and V. E. Manucharyan, The superconducting quasicharge qubit, Nature **585**, 368 (2020).
[50] M. D. Bowdrey, D. K. Oi, A. J. Short, K. Banaszek, and J. A. Jones, Fidelity of single qubit maps, Physics Letters A **294**, 258 (2002).
[51] J. R. Johansson, P. D. Nation, and F. Nori, Qutip: An open-source python framework for the dynamics of open quantum systems, Computer Physics Communications **183**, 1760 (2012).
[52] J. R. Johansson, P. D. Nation, and F. Nori, Qutip 2: A python framework for the dynamics of open quantum systems, Computer Physics Communications **184**, 1234 (2013).
[53] Y. Makhlin and A. Shnirman, Dephasing of solid-state qubits at optimal points, Phys. Rev. Lett. **92**, 178301 (2004).
[54] G. Ithier, E. Collin, P. Joyez, P. J. Meeson, D. Vion, D. Esteve, F. Chiarello, A. Shnirman, Y. Makhlin, J. Schriefl, and G. Schön, Decoherence in a superconducting quantum bit circuit, Phys. Rev. B **72**, 134519 (2005).
[55] P. Groszkowski, A. Di Paolo, A. Grimsmo, A. Blais, D. Schuster, A. Houck, and J. Koch, Coherence properties of the 0-π qubit, New Journal of Physics **20**, 043053 (2018).
[56] N. Didier, E. A. Sete, J. Combes, and M. P. da Silva, ac flux sweet spots in parametrically modulated superconducting qubits, Phys. Rev. Applied **12**, 054015 (2019).
[57] A. Di Paolo, T. E. Baker, A. Foley, D. Sénéchal, and A. Blais, Efficient modeling of superconducting quantum circuits with tensor networks, arXiv:1912.01018 (2019).

[58] F. Yan, S. Gustavsson, A. Kamal, J. Birenbaum, A. P. Sears, D. Hover, T. J. Gudmundsen, D. Rosenberg, G. Samach, S. Weber, *et al.*, The flux qubit revisited to enhance coherence and reproducibility, Nature communications **7**, 12964 (2016).

[59] J. Koch, T. M. Yu, J. Gambetta, A. A. Houck, D. I. Schuster, J. Majer, A. Blais, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, Charge-insensitive qubit design derived from the cooper pair box, Phys. Rev. A **76**, 042319 (2007).

[60] Z. Wang, S. Shankar, Z. Minev, P. Campagne-Ibarcq, A. Narla, and M. Devoret, Cavity attenuators for superconducting qubits, Phys. Rev. Applied **11**, 014031 (2019).

[61] S. Pancharatnam, Generalized theory of interference and its applications, in *Proceedings of the Indian Academy of Sciences-Section A*, Vol. 44 (Springer, 1956) pp. 398–417.

[62] M. V. Berry, Quantal phase factors accompanying adiabatic changes, Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences **392**, 45 (1984).

[63] L. Mandelstam and I. Tamm, The uncertainty relation between energy and time in non-relativistic quantum mechanics, in *Selected Papers* (Springer, 1991) pp. 115–123.

[64] N. Margolus and L. B. Levitin, The maximum speed of dynamical evolution, Physica D: Nonlinear Phenomena **120**, 188 (1998).

[65] A. A. Clerk, M. H. Devoret, S. M. Girvin, F. Marquardt, and R. J. Schoelkopf, Introduction to quantum noise, measurement, and amplification, Rev. Mod. Phys. **82**, 1155 (2010).

[66] A. A. Houck, J. A. Schreier, B. R. Johnson, J. M. Chow, J. Koch, J. M. Gambetta, D. I. Schuster, L. Frunzio, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, Controlling the spontaneous emission of a superconducting transmon qubit, Phys. Rev. Lett. **101**, 080502 (2008).

[67] P. Bertet, I. Chiorescu, G. Burkard, K. Semba, C. J. P. M. Harmans, D. P. DiVincenzo, and J. E. Mooij, Dephasing of a superconducting qubit induced by photon noise, Phys. Rev. Lett. **95**, 257002 (2005).

[68] C. Rigetti, J. M. Gambetta, S. Poletto, B. L. T. Plourde, J. M. Chow, A. D. Córcoles, J. A. Smolin, S. T. Merkel, J. R. Rozen, G. A. Keefe, M. B. Rothwell, M. B. Ketchen, and M. Steffen, Superconducting qubit in a waveguide cavity with a coherence time approaching 0.1 ms, Phys. Rev. B **86**, 100506 (2012).