Analytic Design of Accelerated Adiabatic Gates in Realistic Qubits: General Theory and Applications to Superconducting Circuits

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*

Shortcuts to adiabaticity is a general method for speeding up adiabatic quantum protocols, and has many potential applications in quantum information processing. Unfortunately, analytically constructing shortcuts to adiabaticity 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 nonadiabatic errors and non-RWA errors. We also show that our approach can yield pulses requiring a smaller driving power than conventional nonadiabatic 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 [1–8] 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., Refs. [9–11]). 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) [12–16]. STA protocols seek to modify pulses to completely cancel nonadiabatic 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. Ribeiro and Clerk [17] showed how a particular shortcut approach [superadiabatic transitionless driving (SATD) [18]] 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 were presented in Refs. [19–23]. Note that STA-accelerated gates

are conceptually distinct from the so-called nonadiabatic holonomic gates (see e.g., Ref. [24]) and have distinct advantages [17].

While the above results are promising, they are limited by a constraint that plagues most STA approaches to quantum control: accelerated protocols can be derived analytically for only few-level systems (or systems that reduce to uncoupled few-level systems). Furthermore, one needs to ignore fast-oscillating nonresonant 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 ignored and will limit the operation fidelity even if nonadiabatic errors are suppressed. For this reason, the utility of analytic STA protocols for high-fidelity operations in complex systems has remained unclear.

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 ignored). The result is a general method for analytically deriving pulse sequences that both fully cancel nonadiabatic errors and partially mitigate non-RWA 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 [25–28]. The isolated qubit levels of this system make conventional approaches to gates problematic, providing motivation for new ideas that do not require a direct coupling of qubit levels. The so-called
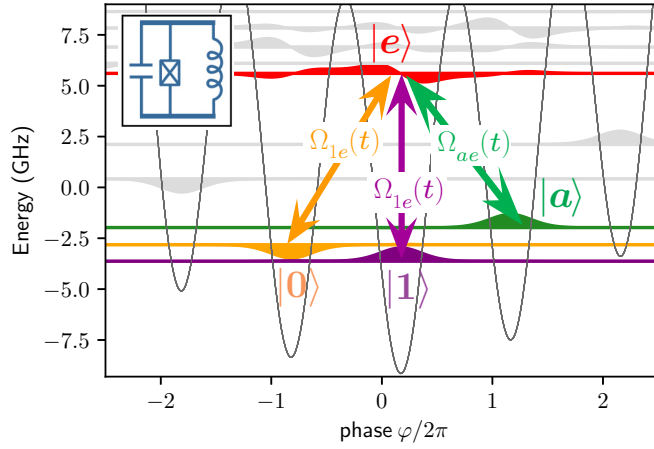
---

*setiawan@uchicago.edu

FIG. 1. A single-qubit gate realized using a four-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, with envelopes $\Omega_{0e}(t)$, $\Omega_{1e}(t)$, and $\Omega_{ae}(t)$, respectively. This tripod-energy structure can be engineered in a multienergy spectrum (shown above are the ten lowest energy levels) of a fluxonium qubit, whose device schematic is shown in the inset. The driving fields can drive unwanted transitions that 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 nonresonant tones.

tripod gate [3,4] is a natural candidate. However, as shown in Fig. 1, a fluxonium circuit has a complex level structure, implying that nonresonant, 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, subject to realistic noise, our enhanced STA gate achieves gate fidelities of 0.9991–0.9997 with a gate time of $t_g = 100$ ns.

Our gate performance is roughly comparable to results obtained using fully numerical optimal control on a related superconducting circuit [29]. 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 higher.

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 multilevel systems? In Sec. III, we briefly review the basic (RWA) geometric tripod gate [3,4] as well as its accelerated version [17]. In Sec. IV, we go beyond the RWA, and discuss how in general STA approaches can be further enhanced to mitigate nonresonant 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 laboratory 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 and $|k\rangle$ are the eigenenergies and eigenstates of the undriven system, respectively, 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

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

In contrast, $\hat{H}_{\text{err}}(t)$ describes all nonresonant processes:

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

The standard next step in most analytic STA approaches to quantum control is to make the 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 involves only the slowly varying amplitudes $V_j(t)$ and which typically couples only 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 [12,14,15,30] to more complex three-level [18,31–34] and four-level [17] protocols.

Despite the power of the above STA approaches, they address only nonadiabatic 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 nonresonant, non-RWA errors. Our method is ultimately perturbative, and amounts to modifying the original (RWA) STA protocol to correct the leading effects of $\hat{H}_{\text{err}}(t)$. 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 focus in what follows on a particularly promising protocol involving accelerated geometric gates using a tripod level structure [3,4] [see Fig. 2(a)]. The general approach we present is completely distinct from recent work [35,36] that deliberately introduced 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. [3,4,7,37] and its accelerated version [17]. All these analyses were done in the context of a simplified four-level RWA Hamiltonian. Our review here sets the stage for our following discussion on how these approaches can be modified and effectively implemented in a realistic multilevel superconducting circuit where non-RWA effects play a crucial role.
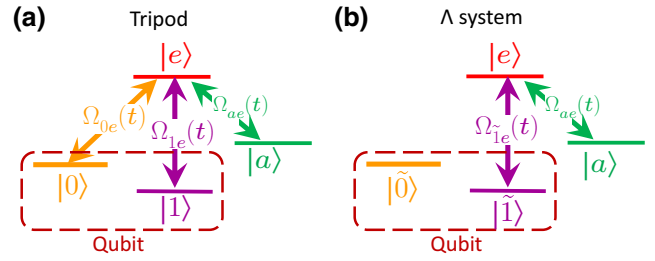

FIG. 2. (a) An ideal tripod system consisting of three lower energy levels (the qubit states $|0\rangle$ and $|1\rangle$ as well as the auxiliary state $|a\rangle$) resonantly coupled to an excited state $|e\rangle$ by three different driving tones, with envelopes $\Omega_{0e}(t)$, $\Omega_{1e}(t)$,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

A. Double stimulated Raman adiabatic passage 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 $|a\rangle$) that are controllably coupled to a common excited state $|e\rangle$ [see Fig. 2(a)]. Denoting these (complex) couplings as $\Omega_{je}(t)$ ($j = 0, 1, a$), we can write the tripod Hamiltonian as ($\hbar = 1$)

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

where

$$\Omega_{je}(t) = V_{je}(t)n_{je} \quad (7)$$

for $j = 0, 1, a$. We take the states $|0\rangle$ and $|1\rangle$ to encode a logical qubit, while $|a\rangle$ and $|e\rangle$ 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 nontrivial geometric 2×2 unitary in this subspace. To understand this more concretely, we follow Ref. [17], and consider control pulses of the form

$$\Omega_{0e}(t) = \Omega_0 \cos(\alpha) \sin[\theta(t)], \quad (8a)$$

$$\Omega_{1e}(t) = \Omega_0 \sin(\alpha) \sin[\theta(t)]e^{i\beta}, \quad (8b)$$

$$\Omega_{ae}(t) = \Omega_0 \cos[\theta(t)]e^{i\gamma(t)}. \quad (8c)$$

The angles α and θ control the relative magnitudes of the pulses, while β and γ control relative phases; α and β are

time independent and we 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 choose to keep constant:

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

At every instant in time, $\hat{H}_0(t)$ has two zero-energy dark states (orthogonal to $|e\rangle$), 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 by moving into a suitable frame defined by the time-independent control pulse parameters α and β [38,39] [Fig. 2(b)]. In this representation, the gate corresponds to a "double stimulated Raman adiabatic passage (STIRAP) protocol" in the Λ system, where the dark state

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

undergoes a cyclic adiabatic evolution $|\tilde{1}\rangle \rightarrow |a\rangle \rightarrow |\tilde{1}\rangle$. This requires an appropriate cyclic variation of the pulse parameter $\theta(t)$ (see Ref. [17] 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 [17].

In the adiabatic limit $\dot{\theta}(t)/\Omega_0 \rightarrow 0$, the gate unitary in the qubit subspace is given by [17]

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

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

where $\hat{\boldsymbol{\sigma}}_{01} = (|0\rangle\langle1| + \text{H.c.}, -i|0\rangle\langle1| + \text{H.c.}, |0\rangle\langle0| - |1\rangle\langle1|)$ 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 nonadiabatic errors. This is exactly the goal of STA methods [12–14,17,18,30].

Following Ref. [17], we consider an STA protocol based on the SATD method [18], where nonadiabatic 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 [17,18,40,41]. Specifically, the SATD protocol requires that one corrects the original pulses via

$$\Omega_{0e}(t) \rightarrow \tilde{\Omega}_{0e}(t)$$

$$\equiv \Omega_0\cos(\alpha)\left(\sin[\theta(t)] + 4\frac{\cos[\theta(t)]\ddot{\theta}(t)}{\Omega_0^2 + 4\dot{\theta}^2(t)}\right), \tag{13a}$$

$$\Omega_{1e}(t) \rightarrow \tilde{\Omega}_{1e}(t)$$

$$\equiv \Omega_0\sin(\alpha)e^{i\beta}\left(\sin[\theta(t)] + 4\frac{\cos[\theta(t)]\ddot{\theta}(t)}{\Omega_0^2 + 4\dot{\theta}^2(t)}\right), \tag{13b}$$

$$\Omega_{ae}(t) \rightarrow \tilde{\Omega}_{ae}(t)$$

$$\equiv \Omega_0 e^{i\gamma(t)}\left(\cos[\theta(t)] - 4\frac{\sin[\theta(t)]\ddot{\theta}(t)}{\Omega_0^2 + 4\dot{\theta}^2(t)}\right), \tag{13c}$$

where the angle $\gamma(t)$ [Eq. (11)] remains unchanged. It can be shown [17] that the resulting accelerated protocol obtained using correction in Eq. (13) achieves the same unitary $\hat{U}_{G,01}$ in the qubit subspace as in the adiabatic limit [see Eq. (12)]. In what follows, we use a tilde 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 [see Eq. (9)], which is the scale of amplitudes of the *uncorrected* pulse (and the corresponding time-independent adiabatic gap). 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 (to cancel nonadiabatic 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 $\Omega_0 t_g$. 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)].
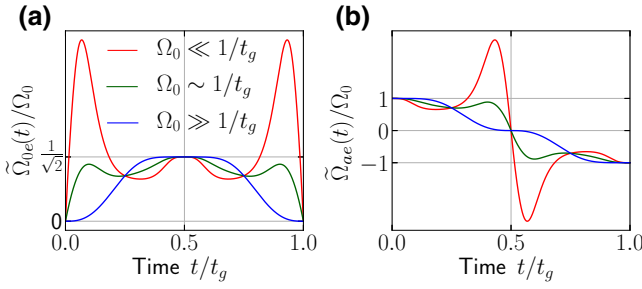
(a)

(b)

FIG. 3. Time profiles of the envelopes of the three driving tones used to realize X gates in an ideal tripod system: (a) $\tilde{\Omega}_{0e}(t) = \tilde{\Omega}_{1e}(t)$; (b) $\tilde{\Omega}_{ae}(t)$. Pulses are calculated with use of Eq. (13) with $\theta(t)$ given in Eq. (A5) 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.

IV. ENHANCING STA PROTOCOLS TO MITIGATE NONRESONANT ERRORS

We now return to the central question of this work: can STA approaches still be effective in settings where nonresonant, non-RWA processes also degrade fidelity? The non-RWA terms can in general induce energy shifts of the main tripod levels, cause leakage to nontripod levels, and also drive additional higher-order processes; these all represent error mechanisms. In this section, we present a general strategy for improving STA protocols to partially mitigate non-RWA errors. For concreteness, we 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 (see Sec. III C) to pick a protocol that minimizes the "size" (appropriately defined) of our control fields. Since the non-RWA errors increase with increasing pulse amplitude, this step mitigates all non-RWA errors.
2. Next, we use a perturbative approach to partially correct non-RWA errors. We focus on correcting the leading-order error mechanism, which is unwanted energy shifts of the computational states (i.e., Stark shifts). As we see in Sec. IV C, there is in general a wide range of gate times where leakage (and higher-order processes) does not play a significant role, meaning that this perturbative approach is very beneficial. To this end, 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 chirp can be found *analytically* using a perturbative approach.

As we see later, 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. Thus, our correction strategy is well suited for obtaining high-fidelity gates. 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 nonresonant, non-RWA processes described by $\hat{H}_{err}(t)$ [see Eq. (5)] yield new unwanted coherent dynamics that will degrade the performance of our gate; example 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 $\Delta_{kl,\pm}^{(j)} = \varepsilon_l - \varepsilon_k \pm \omega_{je}$ 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 (see Fig. 2), and hence $j = 0, 1, a$. Using a Magnus-based approach (see Ref. [42]), one can derive the leading-order, time-dependent energy shift $\delta\varepsilon_k(t)$ of energy level $|k\rangle$ due to $\hat{H}_{err}(t)$. This energy shift has the usual form expected from second-order perturbation theory, i.e.,

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

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

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 us to characterize the "size" of our control pulses by the root-mean-square (rms) voltage of the control field, i.e.,

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

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

Spurious non-resonant (non-RWA) processes

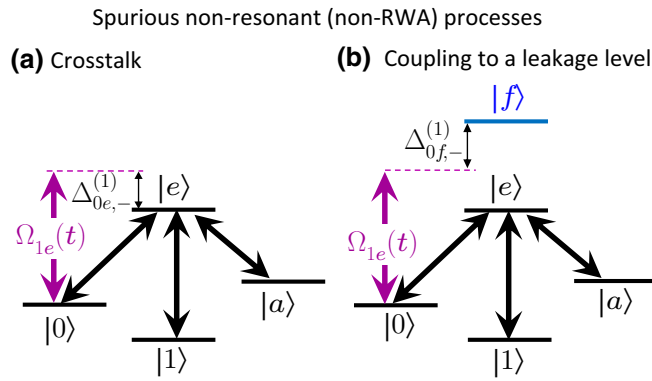**(a)** Crosstalk

(b) Coupling to a leakage level

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 $\Delta_{0e,-}^{(1)}$ 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 $\Delta_{0f,-}^{(1)}$ from the $|0\rangle \leftrightarrow |f\rangle$ transition.

In general, coherent errors increase with increasing drive strength (i.e., pulse amplitude), which we characterize by the rms voltage of our pulses. By minimizing $\tilde{V}_{\rm rms}$, we can therefore limit the effects of these errors on the performance of our gate. As discussed in Sec. VII A, there are also other factors that contribute to a desire to limit the overall amplitude of the drive pulses. For example, when the qubit is indirectly driven through a cavity, it is preferable to use a smaller driving power to constrain the cavity photon number. High drive-induced cavity photon numbers are believed to cause heating effects, which may result in a coherence loss of the qubit, potentially further limiting the performance of our gate.

Using the specific form of the SATD pulses in Eqs. (13), we can write the rms voltage in Eq. (15) as

$$
\tilde{V}_{\rm rms} \simeq \sqrt{\frac{1}{2t_g} \int_0^{t_g} dt \sum_{j=0,1,a} |\tilde{V}_{je}(t)|^2}
$$

$$
= \sqrt{\frac{\cos^2 \alpha}{|n_{0e}|^2} + \frac{\sin^2 \alpha}{|n_{1e}|^2} + \frac{1}{|n_{ae}|^2} \frac{\tilde{\Omega}_{\rm rms}(t_g)}{2}}, \quad (16)
$$

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

$$
\tilde{\Omega}_{\rm rms}(t_g) \equiv \sqrt{\frac{1}{t_g} \int_0^{t_g} dt \left[|\tilde{\Omega}_{0e}(t)|^2 + |\tilde{\Omega}_{1e}(t)|^2 + |\tilde{\Omega}_{ae}(t)|^2 \right]}
$$

$$
= \Omega_0 \sqrt{\frac{1}{t_g} \int_0^{t_g} dt \left[1 + \left(\frac{\ddot{\theta}(t)}{\dot{\theta}^2(t) + \Omega_0^2/4} \right)^2 \right]}. \quad (17)
$$

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. Furthermore, in the second line, we have used the fact that our protocol is symmetric about $t = t_g/2$, i.e., $\theta(t) = \pi/2 - \theta(t - t_g/2)$ for $t_g/2 < t \le t_g$ [Eq. (A5)].

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. 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.

Finally we can use 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 for a given gate time t_g. Out of this set, we choose the protocol that minimizes $\tilde{\Omega}_{\rm rms}$ [and hence approximately minimizes 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 of approximately $1.92/t_g$, occurring for $\Omega_0/2\pi = 1.135/t_g$. We have confirmed that attempting further optimization by using more complex pulse-shape functions $P(t)$ does not yield appreciable improvements. We have thus completed step 1 of our two-step approach to enhancing accelerated gates to minimize non-RWA
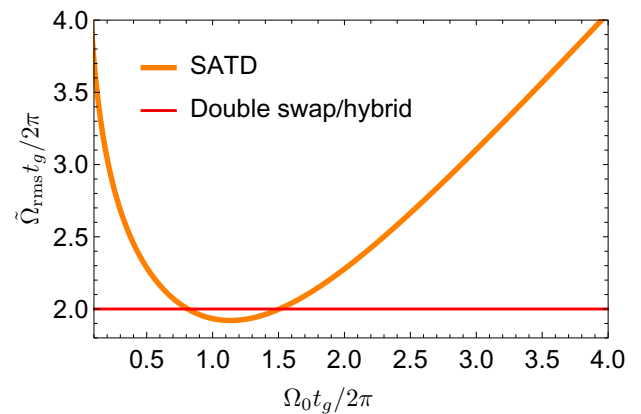

FIG. 5. $\tilde{\Omega}_{\rm rms}t_g/2\pi$ versus $\Omega_0 t_g/2\pi$ for the SATD protocol (orange curve) as well as the double swap protocol and the hybrid protocol (red line). For the SATD protocol, $\tilde{\Omega}_{\rm rms} = \Omega_0$ in the adiabatic limit ($\Omega_0 t_g \to \infty$). Away from the adiabatic limit, the SATD correction increases $\tilde{\Omega}_{\rm rms}$ above Ω_0. The minimum of the SATD plot is $\tilde{\Omega}_{\rm rms}t_g/2\pi = 1.92$, occurring 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 the hybrid protocol.

errors: use the degeneracy of STA protocols to pick a minimum-energy pulse.

Before proceeding, we pause to note something somewhat remarkable: by our optimizing the parameter Ω_0, the SATD pulses are able to achieve an energy-time trade-off (as quantified by the product $\tilde{\Omega}_{\text{rms}} t_g$) that is *better* than more obviously fast, nonadiabatic population-transfer protocols (see the red curve in Fig. 5). These are the so-called double-swap protocols (see e.g., Refs. [43,44]) (where one sequentially moves a population through the three levels) and the "hybrid" protocol [44] (a nonadiabatic holonomic protocol). This point and connections to a formal quantum speed limit valid for time-dependent Hamiltonians [45] 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 derive an analytic prescription for how to do this (as opposed to resorting to a brute-force numerical optimization).

The method 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} ($j = 0, 1, a$) that appear in our pulse:

$$\omega_{je} \to \tilde{\omega}_{je}(t) = \omega_{je} - \delta\varepsilon_j(t) + \delta\varepsilon_e(t), \quad (18)$$

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. Our chirping procedure does not require any additional calibration procedure: the matrix elements required to calculate the frequency chirps [see Eq. (14)] can be computed using the circuit parameters, parameters that must be accurately estimated even in the absence of chirping.

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

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

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 derivative removal by adiabatic gate

Our discussion so far has focused only on errors arising from energy shifts generated by the non-RWA terms in $\hat{H}_{\text{err}}(t)$. 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 nontripod 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 [42,46,47]).

The general approach we take in mitigating non-RWA errors partially minimizes leakage by minimizing the size of the control pulses (see 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 derivative removal by adiabatic gate (DRAG) technique [46,47] for dealing with leakage errors in superconducting circuits driven by a single control tone.

There are two key rationales for our apparent ignoring 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. [42]) to simultaneously correct all of these error channels. This is because the corrections needed to mitigate leakage are not additive in a simple way: adding a correction to cancel one leakage transition could make another leakage transition even worse. This is in 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. [41,48] in the long-gate time limit $(t_g \gg 1/\Delta_{kl,\sigma}^{(j)})$, phase errors scale as $1/(\Delta_{kl,\pm}^{(j)}t_g)^2$, while leakage errors are much weaker, scaling as $1/(\Delta_{kl,\pm}^{(j)}t_g)^4$.

As we see in the next section (where we implement our ideas in a realistic multilevel superconducting fluxonium circuit), our approach yields extremely good results despite the lack of any specific leakage correction. That said, it would be an interesting topic for future work to devise new methods for mitigating leakage in truly multilevel systems; for example, the general semianalytic method in Ref. [41] may provide a route for doing this.

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 qubit levels $|0\rangle$ and $|1\rangle$. This is often the case in
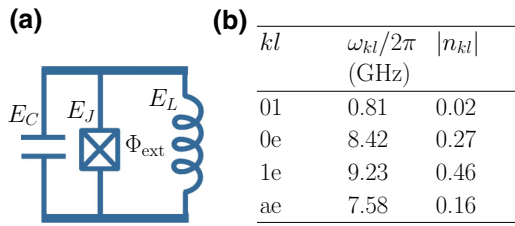
(a)   **(b)**

kl	$\omega_{kl}/2\pi$ (GHz)	$|n_{kl}|$
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_{kl}) as well as the magnitude of the charge matrix elements (n_{kl}) 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_{kl} and n_{kl} in (b) are $E_L/h = 0.063$ GHz, $E_J/h = 9.19$ GHz, $E_C/h = 2$ GHz, and $\Phi_{\text{ext}} = 0.17\Phi_0$.

qubits that have long T_1 relaxation times. A paradigmatic example that has received considerable attention recently is a fluxonium-style superconducting qubit [25–28]. 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) [26–28,49]. Moreover, they can also be made first-order insensitive to the dephasing from $1/f$ charge noise [50]. These properties make a 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

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

where $\hat{n}$ and $\hat{\varphi}$ are, respectively, the charge and phase operators. They obey the commutation relation $[\hat{\varphi}, \hat{n}] = i$. Φ_{ext} is the external magnetic flux biasing the loop formed by the Josephson junction and its shunting inductance, and $\Phi_0 = h/2e$ is the flux quantum. Equation (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 have both a long T_1 time and 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.

Criterion 1 requires us to choose circuit parameters satisfying $E_J \gg E_C$ and $E_L \ll E_J$ (for well-localized ground states) as well as $0 < \Phi_{\text{ext}} \lesssim \Phi_0/4$ (to lift the degeneracy of the ground states). Criterion 2, on the other hand, implies that we have to pick the excited state $|e\rangle$ of the tripod gate to be the first excited state of the central well. Furthermore, this state should be delocalized over the potential wells where the ground states $|0\rangle$, $|1\rangle$, and $|a\rangle$ are located, but should still be somewhat separated from higher-lying energy levels. This implies that the state $|e\rangle$ must lie near the top edge of the cosine potential, which requires $\sqrt{8E_CE_J} \gg 2E_J$. It is obvious that we cannot fulfill criteria 1 and 2 simultaneously in a standard fluxonium circuit.

Since the above requirements cannot be perfectly satisfied simultaneously, we choose parameters that strike an optimal balance, ensuring that we can end up with both a long-lived qubit and a tripod that allows a high-fidelity SATD gate. To this end, we perform a numerical search through parameter space to identify optimal regimes (see Appendix D). 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 $\Phi_{\text{ext}} = 0.17\Phi_0$ (see Appendix D for parameter justification). The small inductive energy here puts our device in same regime as the "Blochnium" circuit recently realized in experiment [51].

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 wave functions are plotted in Fig. 1. We label the energy levels used for the tripod system by $|0\rangle$, $|1\rangle$, $|a\rangle$, and $|e\rangle$. 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: $|n_{0e}| = 0.27$, $|n_{1e}| = 0.46$ and $|n_{ae}| = 0.16$ [right column of the table in Fig. 6(b)].

C. SATD protocols for tripod gates in a 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 $V_{je}(t) = \Omega_{je}(t)/n_{je}$ for $j = 0, 1, a$ [see Eq. (7)]. Here $\Omega_{je}(t)$ is the complex coupling given in Eq. (8) for the uncorrected pulse and Eq. (13) for the SATD-corrected pulse.

The envelopes of the driving tones $\tilde{\Omega}_{je}(t)$ 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 $t_{\rm ramp} = 0.01t_g$ at the beginning
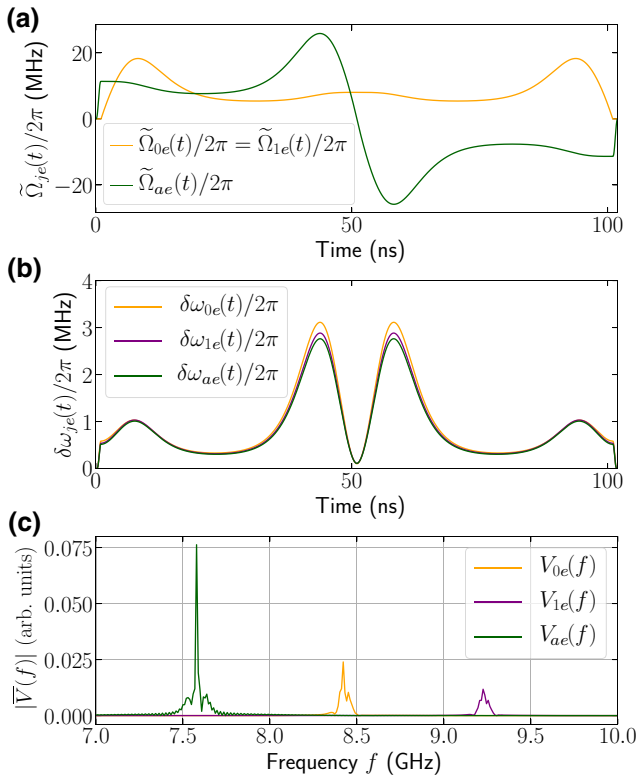
(a)

(b)

(c)

FIG. 7. Driving tones used to realize the SATD X gate in a realistic fluxonium. (a) Time profiles of the driving tone envelopes. (b) Time profiles of the frequency chirps $\delta\omega_{je}(t) = \delta\varepsilon_e(t) - \delta\varepsilon_j(t)$ of each driving tone. (c) The Fourier transform of the driving tones $|\bar{V}(f)|$ showing distinct peaks corresponding to the tripod transition frequencies shown in Fig. 6(b). The pulse sequences in (a) are given in Eq. (E1). They are obtained by our sandwiching the pulses in Eq. (13) by a ramp time $t_{\rm ramp}$ at the beginning and end of the protocol, during which the $\tilde{\Omega}_{ae}(t)$ pulse [green curve in (a)] is turned on and off, respectively, with use of the smooth polynomial function given in Eq. (A6). The parameters used are $t_g = 100$ ns, turn-on and turn-off time $t_{\rm ramp} = 1$ ns, and $\Omega_0/2\pi = 1.135/t_g = 11.35$ MHz. The parameter set for the fluxonium is the same as that used for Fig. 6.

(end) of the protocol. During this ramp, the $\tilde{\Omega}_{ae}(t)$ 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. Including these ramps does not appreciably change our results.

Figure 7(b) shows the size of the frequency chirps $\delta\omega_{je}(t)$ applied to each driving tone [$\omega_{je} \rightarrow \omega_{je} + \delta\omega_{je}(t)$ for $j = 0, 1, a$], to correct for the energy shift of the computational levels due to the non-RWA terms. The sizes of these corrections $\delta\omega_{je}(t)$ are on the order of megahertz, while the base frequency ω_{je} is on the order of gigahertz. The Fourier transform of the driving pulse is shown in Fig. 7(c), where each distinct peak in the plot corresponds to one of the tripod transition frequencies whose values are given in the table in Fig. 6(b).

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 [52]

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

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 laboratory-frame density matrix of the system at the end of the protocol ($t = t_g$) for the initial state $\hat{\rho}_m$. Here $\hat{U}_q$ is a product of the ideal target unitary gate operation in the qubit subspace $\hat{U}_{G,01}$ [see Eq. (12)] and an innocuous phase factor corresponding to dynamical phases in the laboratory frame:

$$\hat{U}_q = \hat{U}_{\text{diag},01}(t_g)\hat{U}_{G,01}, \quad (22)$$

where $\hat{U}_{\text{diag},01}(t_g) = \sum_{k=0,1} e^{-i\int_0^{t_g} \varepsilon_k(t)dt}|k\rangle\langle k|$. 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 ignored and the only sources 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 laboratory-frame Hamiltonian [Eq. (1)] using the PYTHON package QuTiP [53,54]. We perform the simulation by including the 18 lowest energy levels of the fluxonium and all charge matrix elements in this space. We check that including more energy levels in the simulations does not change the results.
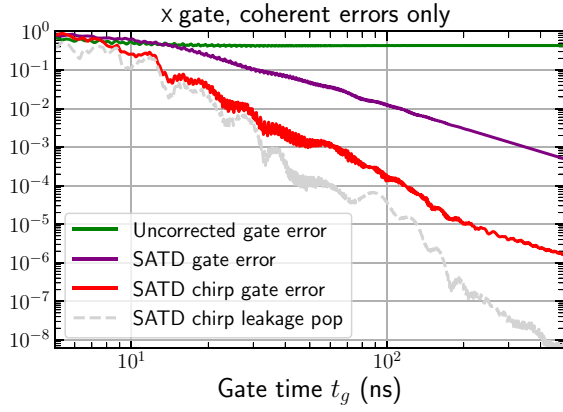
FIG. 8. State-averaged gate error $\bar{\varepsilon} = 1 - \bar{F}$ (solid curves) versus gate time t_g for different realizations of a tripod X gate in a fluxonium qubit in the absence of dissipation. We keep the lowest 18 levels of the circuit and include all non-RWA error channels. The curves correspond to uncorrected adiabatic pulses (green curve), SATD pulses without frequency chirping (purple curve), and SATD pulses with frequency chirping (red curve). 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 curve).

In Fig. 8, we plot the gate errors $\bar{\varepsilon} = 1 - \bar{F}$ for a target tripod gate $\hat{U}_{G,01} = -\hat{\sigma}_{x,01}$ (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 [see Eq. (8); green curve], errors arise both from nonadiabatic transitions and from non-RWA processes. If instead we use the SATD-corrected pulses [see. Eq. (13); purple curve], nonadiabatic 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 the projector in the tripod subspace. The final state-averaged leakage population is plotted as a dashed gray curve in Fig. 8. As discussed, leakage

makes a minimal contribution to the error at moderate to long gate times.

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

1. Effective modeling of non-Markovian noise

We now turn to modeling additional gate errors arising from dissipation. This will give rise to a nontrivial competition: mitigating coherent non-RWA errors favors long gate times (and hence low powers), whereas minimizing dissipative errors favors short gate times. Given our operating point (isolated qubit states, but not at the flux sweet spot), $1/f$ flux noise will often be the dominant dissipation mechanism. In what follows, we thus focus on dephasing dissipation and the resulting competition between coherent and dissipative errors. In the relevant basis of dressed states, dephasing noise can cause transitions; hence, for our scheme, there is no strong qualitative difference between the effects of dephasing and T_1 relaxation. That said, the additional effect of explicit T_1 relaxation is analyzed in detail in Appendix F.

As is common [55–59], we model the $1/f$ 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 of the form $\exp\left[-(t/T_{\varphi,kl})^2\right]$ (up to logarithmic corrections). A standard calculation (see e.g., Ref. [57]) 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 enter only through $A_{\Phi_{\text{ext}}}$ (the standardly defined flux-noise amplitude) and D (the product of the measurement time and the low-frequency cutoff of the noise; see Refs. [29,57]).

We 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 $k \neq 1$:

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

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 G, there is no way to make a choice for Γ_k that captures the decay of all coherences correctly. Nonetheless, as shown in Appendix G, our approach if anything overestimates the dominant dephasing within the tripod subspace (see Table II in Appendix G). Note also that our approach overestimates dephasing compared with alternative approximations that use an explicitly time-dependent Γ_k [58,59]. Moreover, our modeling of dephasing as being Markovian is also a worst-case scenario, as there is spectral weight for arbitrary transitions (which would not be true for realistic $1/f$ noise, where there is little spectral weight to drive transitions between levels with large energy detunings).

In what follows we take $A_{\Phi_{ext}} = 3\mu\Phi_0$ (which is typical for state-of-the-art experiments [27,28,60]), and also choose (following Refs. [29,57,61]) $D = (2\pi \times 1 \text{ Hz}) \times 10 \ \mu s$. With these choices and for the circuit parameters used here, we find dephasing times $T_{\varphi,kl}$ that are on the order of approximately $1 - 100 \ \mu s$ (see Table II in Appendix G). The dephasing of the qubit levels is $T_{\varphi,01} = 7 \ \mu s$. Our results are largely unchanged if one adds T_1 decay processes (e.g., dielectric loss with dielectric quality factor $Q_{diel} \gtrsim 5 \times 10^6$); see Appendix F.

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 [see Eq. (21)]. We first consider the case where all non-RWA terms in the coherent Hamiltonian are ignored so that errors are the result of only dissipation or nonadiabaticity. The results as a function of gate time t_g are shown in Fig. 9, where the performances of
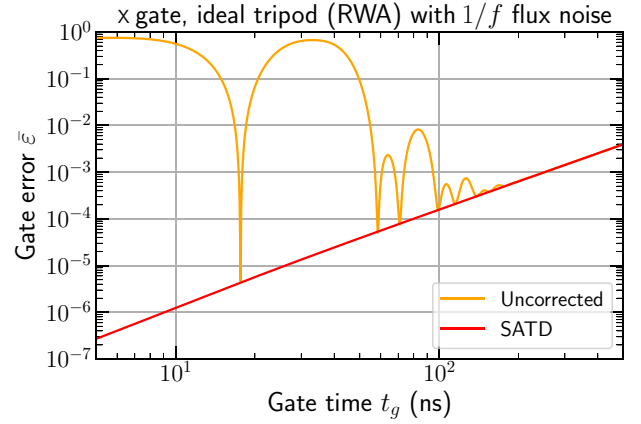

FIG. 9. State-averaged gate error $\bar{\varepsilon} = 1 - \bar{F}$ as a function of gate time t_g for ideal (four-level, RWA) tripod X gates. We compare uncorrected adiabatic pulses (orange curve) against the accelerated SATD pulses (red curve) in the presence of $1/f$ flux noise. The scale of the uncorrected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

the uncorrected and accelerated (SATD) pulses are compared. The simulations of the gate dynamics for all gate times t_g are done for 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 nonadiabatic errors. There are special sharply defined values of t_g where these nonadiabatic errors constructively cancel; as discussed in Ref. [17], these are difficult to exploit experimentally as they require extreme fine-tuning.

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 (nonadiabatic errors, non-RWA errors, and 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 its performance against that of gates realized

using other approaches. For example, one can consider comparing it against a traditional Raman gate (realized by indirectly driving the qubit states via off-resonant coupling to a common excited state [39]) or against "direct driving" (DD) gates (where one resonantly drives the qubit transition between $|0\rangle$ and $|1\rangle$). One key advantage of the SATD tripod approach over these alternatives is their intrinsic robustness: the SATD gate inherits the resilience to control-pulse imperfections that would be expected by a purely adiabatic gate. For example, there is a marked resilience to errors in the magnitude of Ω_0, which could arise either from imperfect pulse calibration or from uncertainties in matrix elements. This advantage persists even in the accelerated adiabatic regime [17].

In what follows we demonstrate *additional* advantages of our tripod gate by specifically comparing its performance against that of the simple DD gate (see Appendix H for comparisons with Raman gates). To this end, we note that as we are working with a T_1-protected qubit, the magnitude of the matrix element used to drive the qubit transition in the DD gate ($|n_{10}|$) is small; however, for a large enough power, one could, in principle, still achieve a given gate. To compare our tripod approach against the DD gate, we want to compare not only the gate error $\bar{\varepsilon}$ but also the *power* required to achieve the gate. As we 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 [41]:

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

where $\chi = \pi$ for an X gate. Similarly to our accelerated adiabatic gate, we perform chirping to partially correct for non-RWA errors, and hence $\tilde{\omega}_{01}(t) = \omega_{01} + \delta\varepsilon_1(t) - \delta\varepsilon_0(t)$ [see 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

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

$$\tilde{V}_{\mathrm{rms,SATD}} = \frac{1.92\pi}{t_{g,\mathrm{SATD}}}\sqrt{\frac{\cos^2\alpha}{|n_{0e}|^2} + \frac{\sin^2\alpha}{|n_{1e}|^2} + \frac{1}{|n_{ae}|^2}}.$$
(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 thus focus on this gate in what follows, and compare it against our accelerated adiabatic approach. Substituting the matrix elements for the fluxonium parameters used throughout this paper [see Fig. 6(b)] in Eqs. (29a) and (29b), we get $\tilde{V}_{\mathrm{rms}}$ for the X gate ($\chi = \pi$ for the DD approach and $\alpha = \pi/4$ for the tripod approach) as

$$\tilde{V}_{\mathrm{rms,DD}} = \frac{136}{t_{g,\mathrm{direct}}},$$
(30a)

$$\tilde{V}_{\mathrm{rms,SATD}} = \frac{42.1}{t_{g,\mathrm{SATD}}}.$$
(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 or greater saving 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 sufficiently detuned microwave cavity mode (e.g., $\omega_{\mathrm{cav}}/2\pi = 2$ GHz; see Appendix I 3 for a justification) that couples to the circuit as

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

Here g is the cavity-fluxonium coupling strength, 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 to avoid additional dissipative mechanisms; here we require $\bar{n}_{\mathrm{cav}} \leq 0.05$. This in turn directly constrains the rms voltage via the relation (see Appendix I)

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

Instead of constraining the cavity photon number, one could consider minimizing some other quantities that are sensitive to the power delivered to the system. This is expected to yield results qualitatively similar to those obtained with our chosen constraint on $\bar{n}_{\mathrm{cav}}$, which is motivated by experimental observations as well as other theoretical studies (e.g., Ref. [29]). We stress that the motivation for limiting power is empirical, and is not based on some rigorous theory of large-drive dissipation.

A variety of experiments exhibit performance degradation when high-amplitude drive pulses are used (see, e.g., Refs. [26,62–64]). Many mechanisms could contribute to these observations, including a simple heating of the circuit environment induced by high drive power. The specific nature of the mechanism is not relevant to the results we present below.

Finally, we remark that since the qubit frequency is lower than the drive frequencies of our tripod gate, instead of driving the qubit via a cavity, one could consider coupling the fluxonium circuit directly to a transmission line that includes a high-pass filter [65]. Such a scenario would not only be less constraining on the amplitudes of the high-frequency tones required to perform our tripod gate but would also still protect the qubit subspace from the undesired transitions caused by the low(er)-frequency fluctuations of the transmission line. Nevertheless, here we concentrate on studying the case of driving the qubit indirectly via a cavity, as that is the most widely used approach in experiments, and hence might be the easiest thing to try in a first experimental realization of our gate.

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

B. Comparison of gate performance, coherent errors only

We first compare the accelerated tripod gate with the 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 with use of 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). Interference between its multiple drive tones causes the SATD error curve to exhibit fast, low-amplitude oscillations as a function of t_g (see the inset in Fig. 10).

The results in Fig. 10 may seem depressing. However, as we argue in what follows, they are misleading. For a given gate time, we have already seen (see Sec. VII) that DD
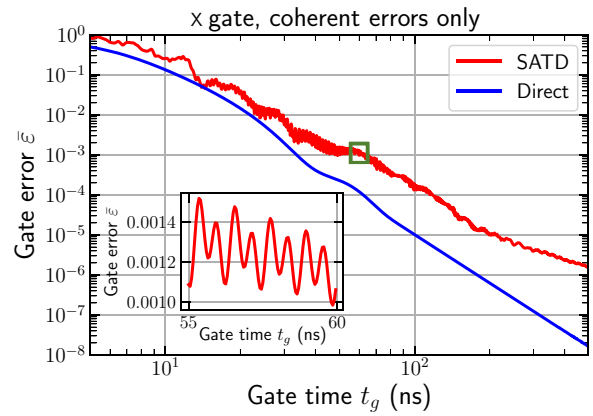

FIG. 10. State-averaged gate error $\bar{\varepsilon} = 1 - \bar{F}$ as a function of gate time t_g, for SATD tripod (red curve) and direct-driving (blue curve) realizations of a fluxonium X gate. Dissipation is not included here, but non-RWA error channels are. The inset shows an enlargement of the SATD gate error. The gates are calculated with 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 for Fig. 6.

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 (modeled as per Sec. VI B 1) as well as non-RWA errors; the results are shown in Fig. 11(a). For both protocols, the error is nonmonotonic with gate time t_g. For short times (gray shaded region), the errors are dominated by non-RWA effects and decrease with increasing t_g. For longer times, dephasing dominates, causing the 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 II in Appendix G).

In Fig. 11(b) we replot these results in a way that now accounts for power constraints that arise when driving is performed through a cavity (see Sec. VII). Each vertical cut corresponds to a fixed value of the rms voltage [see Eq. (30)]; given that DD uses more power, fixing the voltage thus results in t_g approximately 3.2 smaller for SATD (for our parameter choices) than for DD (hence the different t_g axes for DD versus SATD). We thus see that once power is constrained (through $\tilde{V}_{\rm rms}$), 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.
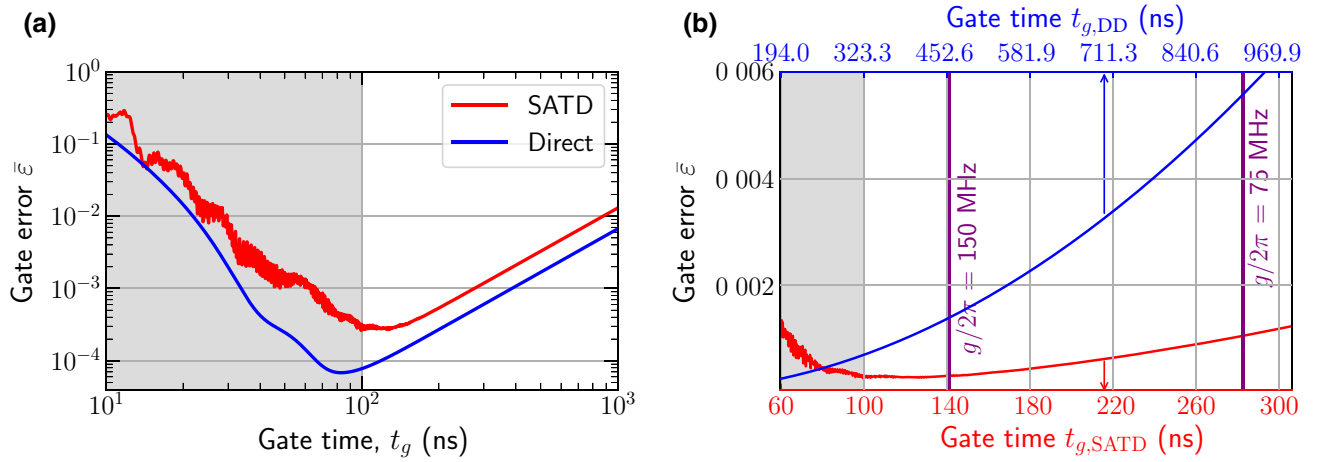
FIG. 11. Comparison between state-averaged x-gate error $\bar{\varepsilon} = 1 - \bar{F}$ for SATD (red curve) and direct-driving (blue curve) protocols in our fluxonium system, including effects of $1/f$ flux noise dephasing as well as non-RWA errors; 18 fluxonium levels are included in the simulations. Dephasing is treated as per Sec. VI B 1 with a $1/f$ flux noise amplitude $A_{\Phi_{ext}} = 3\mu\Phi_0$ corresponding to the qubit dephasing time $T_{\varphi,01} = 7\,\mu$s. (a) Log-log plot of gate error $\bar{\varepsilon}$ versus 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 $\tilde{V}_{rms}$ [see 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 $\tilde{V}_{rms}$, the gate time for the DD protocol is longer than for the SATD protocol (hence the two distinct x axes). Vertical lines indicate the shortest gate times possible [corresponding to the maximum allowed $\tilde{V}_{rms}$; see Eq. (32)] for cavity-based driving for time-averaged cavity photon number $\bar{n}_{cav} = 0.05$ and a fixed cavity-fluxonium coupling g (purple vertical lines). Fluxonium parameters are the same as for Fig. 6.

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

$$\zeta \equiv \frac{\bar{\varepsilon}_{DD}(\tilde{V}_{rms})}{\bar{\varepsilon}_{SATD}(\tilde{V}_{rms})}. \qquad (33)$$

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

$$\zeta = \frac{\bar{\varepsilon}_{DD}}{\bar{\varepsilon}_{SATD}} = 5.3. \qquad (34)$$

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 with the more simplistic DD approach. This conclusion holds for $\tilde{V}_{rms}$ values small enough that errors are dominated by dephasing. As discussed in Sec. VII, the maximum value of $\tilde{V}_{rms}$ is determined by the cavity-qubit coupling g and the requirement that the time-averaged intracavity photon number $\bar{n}_{cav} \leq 0.05$. 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 gate time

$t_g = 100$ ns (including $1/f$ flux noise with $A_{\Phi_{ext}} = 3\mu\Phi_0$ and non-RWA error channels). This compares well with gates designed for similar systems using state-of-the-art numerical optimal control methods. For example, Abdelhafez et al. [29] used numerical optimal control to design gates in a fluxonium circuit operated in a T_1-protected regime. They constrained the drive power in a manner similar to our approach but used slightly different fluxonium parameters and a lower level of flux noise ($A_{\Phi_{ext}} = 1\mu\Phi_0$, 3 times smaller than in our work). They achieved an x gate with $\bar{F} \approx 0.996$ in time $t_g = 60$ ns. The main difference in parameters is that Abdelhafez et al. [29] used flux bias $\Phi_{ext} = 0.45\Phi_0$ (versus $\Phi_{ext} = 0.17\Phi_0$ in our work) and had more isolated qubit states: $|n_{01}| = 0.01$ in their work, a factor of 2 smaller than in our model system.

Our results thus show that even with the inclusion of the effects of $1/f$ dephasing, our SATD tripod gate can achieve excellent performance. These results are essentially unchanged if one now also includes T_1 dissipation, as long as the relevant T_1 times are long enough. Appendix F provides a thorough analysis of our gate performance obtained by our including realistic T_1 effects due to dielectric losses (which is believed to be the dominant mechanism in a fluxonium [27,28]). We find that high gate fidelities are still possible: for example, the gate fidelity at $t_g = 100$ ns is 0.9991–0.9997 for reasonably high but realistic [66,67] values of dielectric quality factor $Q_{diel} \geq 10^6$.

VIII. CONCLUSIONS

We have presented a general strategy showing how analytic shortcuts-to-adiabaticity techniques can be used even in complex multilevel systems where the rotating-wave approximation is not valid. Our strategy to mitigate non-RWA errors involves 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 and T_1 dissipation noise, we achieve a gate fidelity of 0.9991–0.9997 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 consists 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.

APPENDIX A: BASIC WORKING OF THE TRIPOD GATE

In this section, we give a brief overview of the basic four-level geometric tripod gate introduced in Refs. [3,4] [see Fig. 2(a)]; the discussion here follows Ref. [17]. The tripod Hamiltonian $\hat{H}_0(t)$ [Eq. (6)] has two instantaneous zero-energy dark states that span the dark-state manifold and are orthogonal to $|e\rangle$. 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

(defined by the time-independent control pulse parameters α and β) that is purely qubitlike:

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

The qubit state orthogonal to this dark state is

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

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

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

where $\Omega_{\tilde{1}e} = \Omega_0 \sin[\theta(t)]$ and $\Omega_{ae} = \Omega_0 \cos[\theta(t)]e^{i\gamma(t)}$ [Eq. (8c)]. In this new basis, the qubit state $|\tilde{1}\rangle$ and the states $|a\rangle$ and $|e\rangle$ form a three-level Λ system [38,39] [see Fig. 2(b)]. One can write a geometric phase [68,69] onto the state $|\tilde{1}\rangle$ by performing the "double STIRAP protocol," where one slowly varies control pulses to realize the cyclic adiabatic evolution $|\tilde{1}\rangle \to |a\rangle \to |\tilde{1}\rangle$. 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

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

The required cyclic adiabatic evolution is achieved by varying the pulse parameter $\theta(t)$, which brings the dark state $|d(t)\rangle$ from the state $|\tilde{1}\rangle$ at $t = 0$ to $|a\rangle$ 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.,

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

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 turn off of the control fields, we choose a polynomial that 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 [17]

$$P(x) = 6(2x)^5 - 15(2x)^4 + 10(2x)^3. \quad (A6)$$

One also needs a nontrivial relative pulse phase $\gamma(t)$ to obtain a net Berry phase. Following Ref. [17], we use the

simple form of $\gamma(t)$ as given in Eq. (11):

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

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

In the adiabatic limit $\dot{\theta}(t)/\Omega_0 \to 0$, one can show [17] that the dark state $|d(t)\rangle$ accumulates a geometric phase γ_0 at $t = t_g$, and the 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 $\hat{U}_{G,01}$ given in Eq. (12). The full adiabatic limit unitary has the form $\hat{U}_G = \hat{U}_{G,01} \oplus \hat{U}_{G,ae}$, where $\hat{U}_{G,01}$ and $\hat{U}_{G,ae}$ are the unitaries acting in the qubit and auxiliary subspaces, respectively (see Ref. [17]).

APPENDIX B: SATD DRESSING FOR STA PROTOCOLS

We review briefly how the "dressed state" approach to constructing STA protocols [17,18] can be used to accelerate the tripod gate [17]. The general goal is to have the system follow a "dressed" version of the original adiabatic eigenstate that coincides with the original state at the start and end of the protocol. This can be achieved by using a time-dependent dressing function $\nu(t)$ that vanishes at $t = 0$ and $t = t_g$. Following Refs. [17,18], we introduce $|d_\nu\rangle$, a dressed version of the original dark state $|d\rangle$ [Eq. (10)]:

$$|d_\nu(t)\rangle = \exp\left(-i\nu(t)\hat{J}_x\right)|d(t)\rangle. \qquad (B1)$$

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

As discussed in Ref. [17], 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 SATD [17,18], where the dressing angle ν is given by

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

Using the SATD dressing function, one can show [17,18] that the accelerated protocol is implemented by modifying the original uncorrected pulse according to Eq. (13).

APPENDIX C: COMPARING SATD AGAINST SIMPLE NONADIABATIC PROTOCOLS

In Sec. IV A, we discussed how $\tilde{\Omega}_{\text{rms}}$, the time-averaged adiabatic gap of our Hamiltonian, is a relevant metric of

our protocol's energy cost. We also discussed that by optimizing SATD, one can achieve an energy cost $\tilde{\Omega}_{\text{rms}}/2\pi = 1.92/t_g$, where t_g is the gate time. In this appendix, we show that this compares surprisingly favorably with simpler, nonadiabatic pulse protocols.

First, consider the double-swap and hybrid scheme protocols (see, e.g., Refs. [43,44]). To understand these protocols, consider the effective three-level Λ system as shown in Fig. 2(b) [with the Hamiltonian given in Eq. (A3)]. The double-swap protocol involves two sequential swap operations for each STIRAP process, where in the first half of the protocol the pulse $\Omega_{\tilde{1}e}(t)$ is first turned on with a constant value of Ω_0 for half of the time and is then turned off with a simultaneous turn on of the pulse $\Omega_{ae}(t)$ with a constant value of Ω_0 for the other 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 that cyclically evolves the state $|\tilde{1}\rangle \to |e\rangle \to |a\rangle \to |e\rangle \to |\tilde{1}\rangle$, the protocol must be executed for a total gate time $t_g = 2\pi/(\Omega_0/2)$. So we have $\tilde{\Omega}_{\text{rms}} = \Omega_0 = 4\pi/t_g$. 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 $[\Omega_{\tilde{1}e}(t)$ and $\Omega_{ae}(t)]$ 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 performed for a total gate time $t_g = 2\pi/(\sqrt{2}\Omega_0/2)$. Since both pulses are turned on for the whole protocol, we have $\tilde{\Omega}_{\text{rms}} = \sqrt{2}\Omega_0 = 4\pi/t_g$, which is the same as for the double-swap protocol.

1. Proof for the lower bound of $\tilde{\Omega}_{\text{rms}}$

Finally, we establish a rigorous lower bound on $\tilde{\Omega}_{\text{rms}}$ using the quantum speed limit of Ref. [45], which generalized previous work [70,71]. The lower bound on $\tilde{\Omega}_{\text{rms}}$ that we derive below holds for any generic dressing function $\nu(t)$ [Eq. (B1)], including the SATD dressing function [Eq. (B2)]. We begin by applying the bound in Ref. [45] 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

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

where $\mathcal{L}^{\text{QF}}(\rho_0, \rho_{t_g/2}) = \arccos\sqrt{F(\rho_0, \rho_{t_g/2})}$, where $F(\rho_0, \rho_{t_g/2}) = \text{Tr}\sqrt{\sqrt{\rho_0}\rho_{t_g/2}\sqrt{\rho_0}}$ is the Uhlmann fidelity. The left-hand side is the distance of the initial and final states according to the quantum Fisher information metric. The right-hand side of this inequality is the time-integrated instantaneous energy uncertainty of our Hamiltonian $\hat{H}_0(t)$

with the accelerated protocol. The symmetry of our protocol implies the left-hand side 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 $|d(t)\rangle\langle d(t)|$ [Eq. (10)] evolves from the initial state $\rho_0 = |\tilde{1}\rangle\langle\tilde{1}|$ to an orthogonal state $\rho_{t_g/2} = |a\rangle\langle a|$. As a result, the left-hand side of Eq. (C1) becomes $\pi/2$.

For a generic dressing function $\nu(t)$, we can obtain the accelerated Hamiltonian $\hat{H}_0(t)$ in the right-hand side of Eq. (C1) from the adiabatic Hamiltonian [Eq. (A3)] by modifying the original pulse angle and amplitude via [18]

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

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

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

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

The probabilities $p_\pm$ are given by

$$p_\pm(t) = |\langle d_\nu(t)|b_\pm(t)\rangle|^2$$
$$= \frac{1}{2}\left[\left(\frac{\dot{\nu}^2(t)}{\tilde{\Omega}^2(t)}\right)\cos^2[\nu(t)] + \sin^2[\nu(t)]\right]. \quad \text{(C4)}$$

Using Eq. (C4), we calculate the instantaneous values 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

$$\frac{\pi}{2} \leq \frac{1}{2}\int_0^{t_g/2} dt \sqrt{\dot{\nu}^2(t) + \frac{\dot{\theta}^2(t)}{\sec^2[\nu(t)]}}$$

$$\leq \frac{1}{2}\sqrt{\int_0^{t_g/2} dt}\sqrt{\int_0^{t_g/2} dt\left(\dot{\nu}^2(t) + \frac{\dot{\theta}^2(t)}{1 + \tan^2[\nu(t)]}\right)}$$

$$< \frac{t_g}{4}\sqrt{\frac{2}{t_g}\int_0^{t_g/2} dt\left(\dot{\nu}^2(t) + \frac{\dot{\theta}^2(t)}{\tan^2[\nu(t)]}\right)} = \frac{t_g}{4}\tilde{\Omega}_{\text{rms}}. \quad \text{(C6)}$$

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

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

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 that are labeled by $|0\rangle$, $|1\rangle$, and $|a\rangle$) should be well isolated from each other and be nondegenerate. 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 to reduce the coherent errors due to crosstalk.

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.

Criterion 1 requires us to pick circuit parameters that satisfy $E_J \gg E_C$ and $E_L \ll E_J$ (for well-localized ground states) as well as $0 < \Phi_{\text{ext}} \lesssim \Phi_0/4$ (to lift the degeneracy of the ground states). Criterion 2, on the other hand, requires that we pick the excited state $|e\rangle$ 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 $|a\rangle$ reside but be somewhat separated from the much more densely spaced higher energy levels. This means that the state $|e\rangle$ must lie in the vicinity of the top edge of the cosine potential, requiring $\sqrt{8E_CE_J} \gg 2E_J$. It is clear that 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 end up with both a long-lived qubit and a tripod that allows a high-fidelity SATD gate.

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 criteria 1 and 2 above. These include, for example, requiring that the charge matrix element between levels $|0\rangle$ and $|1\rangle$ be small (which maximizes T_1), and that the tripod transition energies ω_{je}, with $j = 0, 1, a$, are not degenerate (to minimize effects of crosstalk). After applying this procedure, we end up with a much reduced parameter space, which in turn is used to perform a more focused search on the circuit parameter set that optimizes the fidelity of an actual SATD gate (although without the frequency chirping). Taking all of the above factors 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 $\Phi_{\text{ext}} = 0.17\Phi_0$. 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 [51]. The corresponding energy level structure and the potential energy landscape are shown in Fig. 1. Because of the small matrix element between qubit levels $|0\rangle$ and $|1\rangle$ ($|n_{01}| = 0.02$), our qubit is not T_1 limited, and because of the positioning of the excited level $|e\rangle$, a tripod with relatively strong tripod matrix elements ($|n_{0e}| = 0.27$, $|n_{1e}| = 0.46$, and $|n_{ae}| = 0.16$) can be realized.

APPENDIX E: ACCOUNTING FOR A SMOOTH TURN ON AND TURN OFF OF THE PULSE AT THE BEGINNING AND END OF THE PROTOCOL

Since realistic pulses are off at 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 $\tilde{\Omega}_{ae}(t)$ 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 that can be separated into three time regions (region I, $0 \leq t < t_{\text{ramp}}$; region II, $t_{\text{ramp}} \leq t \leq t_g + t_{\text{ramp}}$; region III, $t_g + t_{\text{ramp}} < t \leq t_g + 2t_{\text{ramp}}$) as

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

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

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

where $t_- = t - t_{\text{ramp}}$. Specifically, we choose $t_{\text{ramp}} = 0.01t_g$ to be short enough compared with 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. In Eq. (E1c) we have used the function $P(x)$ given in Eq. (A6) to ensure a smooth turn on and turn off of the pulse $\tilde{\Omega}_{ae}(t)$ during a 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: EFFECTS OF THE T_1 DECAY PROCESS ON THE GATE ERROR

In this section, we discuss the effects of T_1 decay on the gate error. We focus only on the most dominant T_1 relaxation process that involves the computational states (i.e., the relaxation from the state $|e\rangle$ to the state $|1\rangle$). We have checked that the other T_1 relaxation processes affecting the computational states are much longer and hence do not impact the gate performance significantly. We consider the T_1 decay process to be due to the dielectric loss in capacitors since it is typically the most significant relaxation process in experiments [27,28]. The T_1 relaxation time for any $|k\rangle \rightarrow |l\rangle$ transition due to the dielectric loss is given by

$$1/(T_1)_{kl} = \frac{|\omega_{lk}|\omega_{lk}}{8E_C Q_{\text{diel}}}\left[\coth\left(\frac{\omega_{lk}}{2k_B T}\right) + 1 \right]|\langle l|\hat{\varphi}|k\rangle|^2, \tag{F1}$$

where Q_{diel} is the dielectric quality factor, E_C is the capacitive energy, $\omega_{lk} \equiv \varepsilon_k - \varepsilon_l$ is the energy transition between the state $|l\rangle$ and the state $|k\rangle$, and $\hat{\varphi}$ is the phase operator. Table I shows the $(T_1)_{e1}$ relaxation time, corresponding to the $|e\rangle \rightarrow |1\rangle$ decay, calculated at zero temperature ($T = 0$) for different values of Q_{diel}. We take into account this relaxation process by adding the relaxation Lindblad operator

$$\hat{Z} = \sqrt{1/(T_1)_{e1}}|1\rangle\langle e| \tag{F2}$$

to the master equation [Eq. (24)].

Figure 12 shows a comparison between the state-averaged X-gate error $\bar{\varepsilon}$ for SATD protocols calculated with and without inclusion of the T_1 relaxation process. For all plots, we take into account the effects of both $1/f$

TABLE I.   Relaxation times $(T_1)_{e1}$ for the $|e\rangle \rightarrow |1\rangle$ decay. The relaxation times are calculated at zero temperature with use of Eq. (F1) for different values of the dielectric quality factor Q_{diel}. For comparison, the qubit dephasing time $T_{\varphi,01} = 7.03\ \mu$s. The fluxonium parameters used are as for Fig. 6.

Q_{diel}	$(T_1)_{e1}$ (μs)
5×10^5	11.9
1×10^6	23.8
2×10^6	47.6
1×10^7	238

flux noise and the coherent errors. As seen in Fig. 12, for $(T_1)_{e1}$ values slightly larger than $T_{\varphi,01}$, e.g., $Q_{\mathrm{diel}} \approx 10^6$ [see Table I for $(T_1)_{e1}$ values calculated for different Q_{diel} values], the T_1 relaxation mechanism can still have an appreciable effect on the gate error. This is because our Markovian modeling of the non-Markovian $1/f$ dephasing noise (as discussed in Sec VI B 1 and Appendix G) gives an effective dephasing rate $t_g/(T_{\varphi,kl})^2$ [Eq. (27)], which is smaller than the decay rate $1/(T_1)_{kl}$ of a T_1 Markovian decay process when $(T_1)_{kl} \approx T_{\varphi,kl}$. Here $T_{\varphi,kl}$ is the pure dephasing time calculated from the standard free-induction-decay calculation for the state $|k\rangle$ with use of the state $|l\rangle$ as the reference state. On the other hand, for high values of Q_{diel} (i.e., $Q_{\mathrm{diel}} \gtrsim 5 \times 10^6$), the T_1 relaxation process does not significantly change the gate error. We note

that the use of capacitors with high values of Q_{diel} (e.g., $Q_{\mathrm{diel}} \gtrsim 10^6$) is typical for present experiments [67,72].

For cases where the T_1 dissipation dominates, we can reoptimize the gate error by using a value of Ω_0 that is larger than the power-optimal Ω_0 [i.e., $\Omega_0 = 1.135(2\pi)/t_g$]. This is because for a fixed t_g, SATD protocols populate the excited state $|e\rangle$ less for larger values of Ω_0 [18], and hence they are less susceptible to the T_1 decay process. However, one cannot use an arbitrarily large value of Ω_0, as beyond a certain value of Ω_0 the coherent errors will start to dominate over the T_1 decay process. To this end, we perform a numerical minimization of gate errors in
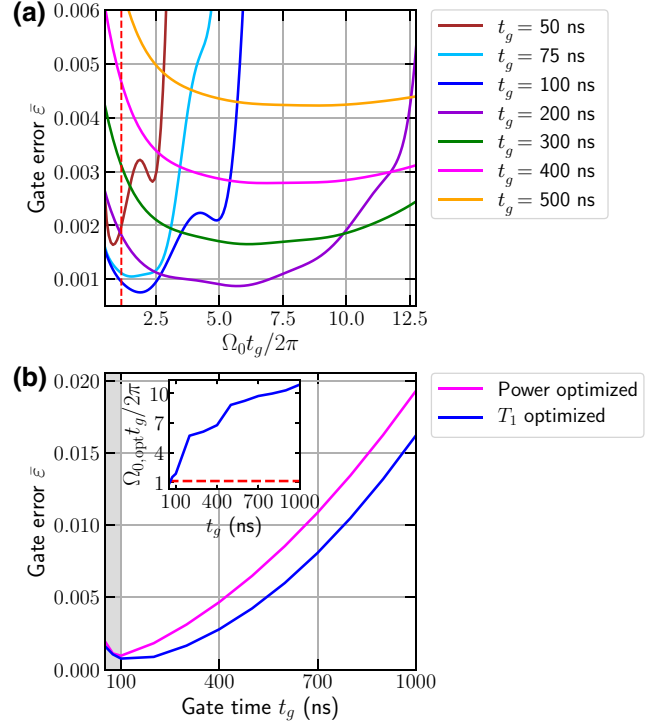
FIG. 13.   State-averaged SATD X-gate errors $\bar{\varepsilon} = 1 - \bar{F}$ calculated in the presence of the $(T_1)_{e1}$ relaxation process, $1/f$ flux noise dephasing, and non-RWA errors. (a) Gate error $\bar{\varepsilon}$ versus $\Omega_0 t_g/2\pi$ for different gate times t_g. The dashed red line indicates the value of power-optimal Ω_0, i.e., $\Omega_0 t_g/2\pi = 1.135$. (b) Gate error $\bar{\varepsilon}$ versus gate time t_g calculated with the power-optimal Ω_0 (magenta curve) and optimal values of Ω_0 (inset) that minimize the gate errors in the presence of the T_1 decay process (blue curve). The shaded (unshaded) region corresponds to the time regime where the gate errors are dominated by non-RWA errors (T_1 dissipation). The inset shows the optimal value of Ω_0 that minimizes the gate error in the presence of the T_1 decay process as a function of gate time t_g, where the dashed red line corresponds to the power-optimal value $\Omega_0 t_g/2\pi = 1.135$. The relaxation time $(T_1)_{e1}$ used is 23.8 μs, which corresponds to dielectric quality factor $Q_{\mathrm{diel}} = 1 \times 10^6$. Eighteen fluxonium levels are included in the simulations. Fluxonium parameters and the dephasing amplitude (corresponding to the qubit dephasing time $T_{\varphi,01} = 7.03\ \mu$s) are the same as for Fig. 11.
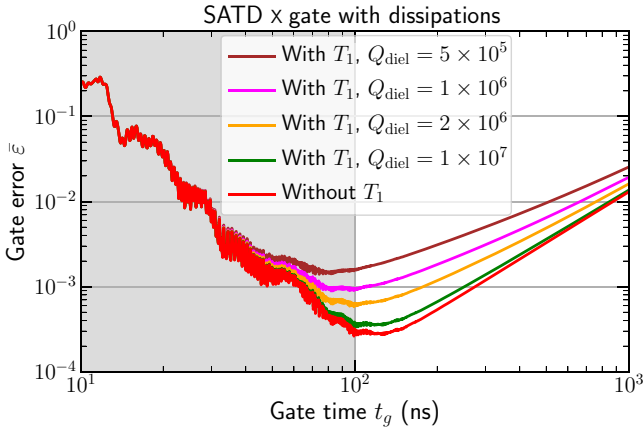

FIG. 12.   Comparison between the state-averaged X-gate error $\bar{\varepsilon} = 1 - \bar{F}$ for SATD protocols with and without T_1 relaxation rate. We take into account only the most dominant relaxation process $(T_1)_{e1}$, i.e., the relaxation from the state $|e\rangle$ to the state $|1\rangle$. The relaxation time $(T_1)_{e1}$ is calculated with Eq. (F1) for different values of the dielectric quality factor Q_{diel} [see Table I for the values of $(T_1)_{e1}$]. The effects of $1/f$ flux noise dephasing as well as non-RWA errors are taken into account; 18 fluxonium levels are included in the simulations. Fluxonium parameters and the dephasing amplitude (corresponding to the qubit dephasing time $T_{\varphi,01} = 7.03\ \mu$s) are the same as for Fig. 11.

the T_1-dominated regime by doing a brute-force optimization of Ω_0. Figure 13(a) shows the gate error versus $\Omega_0 t_g$ calculated for different gate times t_g. In the coherent-error-dominated regime (i.e., $t_g \lesssim 100$ ns), the value of Ω_0 that gives the minimum gate error is close to the power-optimal Ω_0 [shown as a red dashed line in Fig. 13(a) and inset of Fig. 13(b)]. However, for the T_1-dissipation-dominated regime (i.e., $t_g \gtrsim 100$ ns), the optimal value of Ω_0 that minimizes the gate error at fixed t_g is larger than the power-optimal Ω_0 [see the inset in Fig. 13(b)]. In Fig. 13(b), we compare the gate errors calculated with the power-optimal Ω_0 (magenta curve) with those obtained with optimal values of Ω_0 (one for each t_g value) that minimize the gate errors in the presence of the T_1 decay process (blue curve). In the coherent-error-dominated regime (shaded region), the minimum gate errors are approximately the same as those obtained with the power-optimal Ω_0. However, in the T_1-decay-dominated regime, there is a reduction in the gate errors obtained with use of the reoptimized values of Ω_0.

APPENDIX G: APPROXIMATING NON-MARKOVIAN NOISE WITH MARKOVIAN DYNAMICS

In this appendix, we show that approximating non-Markovian dynamics of $1/f$ noise with the Markovian master equation cannot capture the decay rate of all coherences correctly. In the main text, we constructed our Linblad master equation by picking the decay rate Γ_k of the coherences $\rho_{k1} \equiv \langle k|\hat{\rho}|1\rangle$ (coherences that involve the reference level, i.e., the qubit state $|1\rangle$) such that the final decay of the coherences ρ_{k1}, for all $k \neq 1$, 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 an envelope of the form $\exp\left[-(t/T_{\varphi,kl})^2\right]$, 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}, for all $k, l \neq 1$) 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

$$\dot{\rho}_{kl}(t) = -i[\hat{H}(t), \hat{\rho}(t)]_{kl}$$
$$- \left[\mathrm{sgn}\left(\frac{\partial \varepsilon_k}{\partial \Phi_{\mathrm{ext}}}\right)\sqrt{\Gamma_k} - \mathrm{sgn}\left(\frac{\partial \varepsilon_l}{\partial \Phi_{\mathrm{ext}}}\right)\sqrt{\Gamma_l}\right]^2 \rho_{kl}$$
$$= -i[\hat{H}(t), \hat{\rho}(t)]_{kl} - \frac{t_g}{\left(T_{\varphi,kl}^{\mathrm{eff}}\right)^2}\rho_{kl}, \quad \text{(G1)}$$

where we have identified

TABLE II. Dephasing times for transitions between computational states $|k\rangle$ and $|l\rangle$ calculated by two different methods: directly from the free induction decay formula [Eq. (26)] (middle column) and indirectly from the Lindblad master equation [Eq. (G2)] (right column).

kl	$T_{\varphi,kl}$ (μs)	$T_{\varphi,kl}^{\mathrm{eff}}$ (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

$$\frac{1}{T_{\varphi,kl}^{\mathrm{eff}}} \equiv \left|\mathrm{sgn}\left(\frac{\partial \varepsilon_k}{\partial \Phi_{\mathrm{ext}}}\right)\frac{1}{T_{\varphi,k1}} - \mathrm{sgn}\left(\frac{\partial \varepsilon_l}{\partial \Phi_{\mathrm{ext}}}\right)\frac{1}{T_{\varphi,l1}}\right|. \quad \text{(G2)}$$

We can see that the effective dephasing time $T_{\varphi,kl}^{\mathrm{eff}}$ [Eq. (G2)], for all $k, l \neq 1$, 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). As shown in Table II, our approach if anything overestimates the dominant dephasing processes within the tripod subspace, e.g., $T_{\varphi,a0}^{\mathrm{eff}}$ and $T_{\varphi,ae}^{\mathrm{eff}}$.

APPENDIX H: COMPARISON WITH RAMAN GATES

In this section, we compare the performance of our tripod gate against that of the Raman gate. In the Raman protocol, detuned drives couple the excited state $|e\rangle$ to the qubit states $|0\rangle$ and $|1\rangle$, leading to an effective direct coupling between the qubit states [39]. The requirement that the drives in a Raman approach be highly detuned means that (for fixed drive amplitudes) it will result in a much slower gate than is possible, for example, by using our SATD tripod approach (which uses resonant drive).

To make the comparison more explicit, let Δ denote the detunings of both drives used in a Raman protocol and let Ω_0 denote the drive amplitudes. The Raman gate is based on adiabatically eliminating the excited state to generate an effective qubit-only Hamiltonian. This elimination necessarily requires $\Omega_0 \ll \Delta$, leading to a Raman gate time $t_g = 2\pi\Delta/\Omega_0^2 \gg 1/\Omega_0$. In contrast, using our tripod gate, we are able to achieve much shorter gate times $t_g \sim 1/\Omega_0$ for the same drive amplitude. As a result, we expect the Raman gate to be more susceptible to dissipation than our tripod gate, which gives an enormous advantage for the tripod gate if dissipation prevents the use of long gate times. Moreover, unlike the direct-driving and Raman gates, our SATD gate is robust against imperfections in control pulses [17].

APPENDIX I: INDIRECT QUBIT DRIVING THROUGH A COUPLED CAVITY

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)].

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

$$\hat{H}_{JC}(t) = \sum_k \varepsilon_k |k\rangle\langle k| + \omega_{cav}\hat{a}^\dagger\hat{a}$$
$$+ \sum_{k,l} gn_{kl}|k\rangle\langle l|(\hat{a}^\dagger + \hat{a}) + u(t)(\hat{a}^\dagger + \hat{a}), \quad (I1)$$

where ε_k and $|k\rangle$ 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 $gn_{kl} \ll |\varepsilon_l - \varepsilon_k - \omega_{cav}|$, 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 operator $\hat{a}(t)$ can be written as [73]

$$\dot{\hat{a}}(t) = i[\hat{H}_{JC}(t), \hat{a}(t)] - \frac{\kappa}{2}\hat{a}(t) - \sqrt{\kappa}\hat{b}_{in}(t)$$
$$= -i\omega_{cav}\hat{a}(t) - iu(t) - \frac{\kappa}{2}\hat{a}(t) - \sqrt{\kappa}\hat{b}_{in}(t). \quad (I2)$$

Here κ is the cavity photon decay rate (due to photon leakage to the bath) and $\hat{b}_{in}(t)$ is the standard bath annihilation operator, which represents the noise [73]. The

equation of motion for $\hat{a}^\dagger(t)$ can be obtained by taking the Hermitian conjugate of Eq. (I2). 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

$$\ddot{x}(t) = -\left(\omega_{cav}^2 + \frac{\kappa^2}{4}\right)x(t) - \kappa\dot{x}(t) - \sqrt{2}\omega_{cav}u(t), \quad (I3)$$

where we have used $\langle\hat{b}_{in}(t)\rangle = \langle\hat{b}_{in}^\dagger(t)\rangle = 0$. The solution of Eq. (I3) is

$$x(t) = x_p(t) + x_h(t), \quad (I4)$$

where

$$x_h(t) = Ae^{-\frac{\kappa}{2}t}\sin(\omega_{cav}t + \phi) \quad (I5)$$

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

$$x(\omega) = \int_{-\infty}^{\infty} x(t)e^{-i\omega t}dt,$$
$$u(\omega) = \int_{-\infty}^{\infty} u(t)e^{-i\omega t}dt \quad (I6)$$

and their inverse Fourier transforms as

$$x(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} x(\omega)e^{i\omega t}d\omega,$$
$$u(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} u(\omega)e^{i\omega t}d\omega. \quad (I7)$$

Taking the Fourier transform of Eq. (I3), we have

$$\left(-\omega^2 + i\kappa\omega + \omega_{cav}^2 + \frac{\kappa^2}{4}\right)x(\omega) = -\sqrt{2}\omega_{cav}u(\omega), \quad (I8)$$

which gives

$$x(\omega) = \frac{\sqrt{2}\omega_{cav}u(\omega)}{\omega^2 - i\kappa\omega - (\omega_{cav}^2 + \kappa^2/4)}. \quad (I9)$$

We can relate the mean photon displacement $x(t)$ to 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.,

$$\hat{H}_c \approx \sum_{kl} g \langle \hat{a}^\dagger + \hat{a} \rangle n_{kl} |k\rangle \langle l|$$

$$= \sqrt{2} g x(t) \sum_{kl} n_{kl} |k\rangle \langle l|. \qquad (I10)$$

Comparing Eq. (I10) with the qubit driving term in the Hamiltonian in Eq. (1), we can identify

$$\sqrt{2} g x(t) = V(t), \qquad (I11)$$

where in the frequency domain it can be written as

$$\sqrt{2} g x(\omega) = \int_{-\infty}^{\infty} V(t') e^{-i\omega t'} dt'. \qquad (I12)$$

Substituting Eq. (I9) into Eq. (I12), we have

$$u(\omega) = \frac{[\omega^2 - i\kappa\omega - (\omega_{cav}^2 + \kappa^2/4)]}{2g\omega_{cav}} \int_{-\infty}^{\infty} V(t') e^{-i\omega t'} dt'. \qquad (I13)$$

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

$$u(t) = -\frac{1}{2g\omega_{cav}} \left(\frac{d^2}{dt^2} + \kappa \frac{d}{dt} + \omega_{cav}^2 + \frac{\kappa^2}{4} \right) V(t). \qquad (I14)$$

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

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

$$\bar{n}_{cav} = \frac{1}{t_g} \int_0^{t_g} \langle \hat{a}^\dagger \hat{a} \rangle dt$$

$$\approx \frac{1}{t_g} \int_0^{t_g} \langle \hat{a}^\dagger \rangle \langle \hat{a} \rangle dt = \frac{1}{t_g} \int_0^{t_g} dt |\eta(t)|^2. \qquad (I15)$$

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. (I10), we can identify the qubit driving voltage as

$$V(t) = 2g \, \mathrm{Re}\, \eta(t). \qquad (I16)$$

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

$$\tilde{V}_{rms} = \sqrt{\frac{1}{t_g} \int_0^{t_g} |V(t)|^2}$$

$$= g\sqrt{2\bar{n}_{cav}}, \qquad (I17)$$

where in evaluating the second line, we used the relation

$$\frac{1}{t_g} \int_0^{t_g} [\mathrm{Re}\, \eta(t)]^2 = \sum_j \frac{1}{t_g} \int_0^{t_g} |\eta_j(t)|^2 \cos^2(\omega_j t + \phi_j)$$

$$\simeq \frac{1}{2t_g} \int_0^{t_g} |\eta(t)|^2 = \frac{\bar{n}_{cav}}{2}. \qquad (I18)$$

Without loss of generality, in Eq. (I18), 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, i.e., $\bar{n}_{cav} = \langle \hat{a}^\dagger \hat{a} \rangle \ll 1$; for our simulations in the main text, we specifically set $\bar{n}_{cav} = 0.05$. This constrains the maximum rms voltage for a fixed cavity-qubit coupling strength g as shown in Eq. (32). Each vertical cut of the plot in Fig. 11(b) corresponds to different maximum allowed values of $\tilde{V}_{rms}$, 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 $\tilde{V}_{rms}$.

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

In this section we briefly outline the effects of thermal photons in the 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 $\kappa \ll g^2/|\Delta_{kl,\pm}^{cav}|$, for all k,l, where $\Delta_{kl,\pm}^{cav} = \varepsilon_l - \varepsilon_k \pm \omega_{cav}$ is the detuning of the cavity frequency ω_{cav} from the $|k\rangle \leftrightarrow |l\rangle$ energy transition involving the computational levels. The first noise channel we consider is the Purcell relaxation time, which we denote as $T_{1,cav}$ and can be calculated as [74]

$$(T_{1,cav})_{jk} = (\Delta_{kl,\pm}^{cav})^2 / (\kappa g^2 |n_{kl}|^2). \qquad (I19)$$

Since all the relevant transitions are sufficiently detuned from the cavity frequency, we expect the Purcell relaxation 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 $T_{2,cav}$, which can be approximated by [75,76]

$$T_{2,cav} = 1/\kappa \bar{n}_{th}. \qquad (I20)$$

We stress that Eq. (I20) is valid in the limit of small cavity photon decay rate relative to all the relevant dispersive shifts ($\kappa \ll g^2/|\Delta_{kl,\pm}^{cav}|$), a regime that we are interested in. Moreover, we also consider operating in the dispersive limit ($g n_{kl} \ll |\Delta_{kl,\pm}^{cav}|$, for all k,l) and small cavity-thermal-photon-number regime ($\bar{n}_{th} \ll 1$) such that the cavity response can be treated independently of the qubit (see Sec. I 1). As an example, we can pick $\omega_{cav}/2\pi = 2$ 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 $\bar{n}_{th} = 0.05$ (corresponding to a temperature $T \approx 30$ mK). Substituting the above parameters into Eqs. (I19) and (I20), we obtain $(T_{1,cav})_{01} = 0.15$ s and $T_{2,cav} = 0.32$ ms. Since the cavity-induced relaxation time $(T_{1,cav})_{01}$ and dephasing time $T_{2,cav}$ are much larger than the dephasing times due to the $1/f$ flux noise (see Table II), we can ignore the effect of cavity-thermal-photon-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(R) (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] D. Møller, L. B. Madsen, and K. Mølmer, Geometric phases in open tripod systems, Phys. Rev. A **77**, 022306 (2008).

[8] 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 Inf. **5**, 73 (2019).

[9] 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).

[10] K. Toyoda, K. Uchida, A. Noguchi, S. Haze, and S. Urabe, Realization of holonomic single-qubit operations, Phys. Rev. A **87**, 052307 (2013).

[11] 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).

[12] M. Demirplak and S. A. Rice, Adiabatic population transfer with control fields, J. Phys. Chem. A **107**, 9937 (2003).

[13] M. Demirplak and S. A. Rice, Assisted adiabatic passage revisited, J. Phys. Chem. B **109**, 6838 (2005).

[14] M. V. Berry, Transitionless quantum driving, J. Phys. A: Math. Theor. **42**, 365303 (2009).

[15] 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).

[16] 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).

[17] H. Ribeiro and A. A. Clerk, Accelerated adiabatic quantum gates: Optimizing speed versus robustness, Phys. Rev. A **100**, 032323 (2019).

[18] A. Baksic, H. Ribeiro, and A. A. Clerk, Speeding Up Adiabatic Quantum State Transfer by Using Dressed States, Phys. Rev. Lett. **116**, 230503 (2016).

[19] J. Zhang, T. H. Kyaw, D. Tong, E. Sjöqvist, and L.-C. Kwek, Fast non-abelian geometric gates via transitionless quantum driving, Sci. Rep. **5**, 18414 (2015).

[20] 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 J. Phys. **18**, 023001 (2016).

[21] 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 J. Phys. **20**, 065003 (2018).

[22] 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).

[23] 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 (Europhys. Lett.) **129**, 30008 (2020).

[24] E. Sjöqvist, D.-M. Tong, L. M. Andersson, B. Hessmo, M. Johansson, and K. Singh, Non-adiabatic holonomic quantum computation, New J. Phys. **14**, 103035 (2012).

[25] 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).

[26] 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).

[27] 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).

[28] 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).

[29] 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).

[30] M. Demirplak and S. A. Rice, On the consistency, extremal, and global properties of counterdiabatic fields, J. Chem. Phys. **129**, 154111 (2008).

[31] 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 Phys. Lett. **13**, 105202 (2016).

[32] 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, Sci. Rep. **6**, 36737 (2016).

[33] 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).

[34] J.-L. Wu, S.-L. Su, X. Ji, and S. Zhang, Superadiabatic scheme for optimizing the fast generation of tree-type 3D entanglement, Ann. Phys. **386**, 34 (2017).

[35] 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).

[36] 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).

[37] B.-J. Liu, Z.-H. Huang, Z.-Y. Xue, and X.-D. Zhang, Superadiabatic holonomic quantum computation in cavity qed, Phys. Rev. A **95**, 062308 (2017).

[38] K. Bergmann, H. Theuer, and B. W. Shore, Coherent population transfer among quantum states of atoms and molecules, Rev. Mod. Phys. **70**, 1003 (1998).

[39] 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).

[40] 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, Nat. Phys. **13**, 330 (2017).

[41] T. F. Roque, A. A. Clerk, and H. Ribeiro, Engineering fast high-fidelity quantum operations with constrained interactions, npj Quantum Inf. **7**, 28 (2021).

[42] 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).

[43] Y.-D. Wang and A. A. Clerk, Using dark modes for high-fidelity optomechanical quantum state transfer, New J. Phys. **14**, 105010 (2012).

[44] Y.-D. Wang and A. A. Clerk, Using Interference for High Fidelity Quantum State Transfer in Optomechanics, Phys. Rev. Lett. **108**, 153603 (2012).

[45] 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).

[46] 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).

[47] 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).

[48] N. Wiebe and N. S. Babcock, Improved error-scaling for adiabatic quantum evolutions, New J. Phys. **14**, 013024 (2012).

[49] 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).

[50] 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).

[51] I. V. Pechenezhskiy, R. A. Mencia, L. B. Nguyen, Y.-H. Lin, and V. E. Manucharyan, The superconducting quasicharge qubit, Nature **585**, 368 (2020).

[52] M. D. Bowdrey, D. K. Oi, A. J. Short, K. Banaszek, and J. A. Jones, Fidelity of single qubit maps, Phys. Lett. A **294**, 258 (2002).

[53] J. R. Johansson, P. D. Nation, and F. Nori, Qutip: An open-source python framework for the dynamics of open quantum systems, Comput. Phys. Commun. **183**, 1760 (2012).

[54] J. R. Johansson, P. D. Nation, and F. Nori, Qutip 2: A python framework for the dynamics of open quantum systems, Comput. Phys. Commun. **184**, 1234 (2013).

[55] Y. Makhlin and A. Shnirman, Dephasing of Solid-State Qubits at Optimal Points, Phys. Rev. Lett. **92**, 178301 (2004).

[56] 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).

[57] P. Groszkowski, A. Di Paolo, A. Grimsmo, A. Blais, D. Schuster, A. Houck, and J. Koch, Coherence properties of the 0-π qubit, New J. Phys. **20**, 043053 (2018).

[58] N. Didier, E. A. Sete, J. Combes, and M. P. da Silva, AC Flux Sweet Spots in Parametrically Modulated Superconducting Qubits, Phys. Rev. Appl. **12**, 054015 (2019).

[59] 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).

[60] 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, Nat. Commun. **7**, 12964 (2016).

[61] 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).

[62] D. Sank, *et al.*, Measurement-Induced State Transitions in a Superconducting Qubit: Beyond the Rotating Wave Approximation, Phys. Rev. Lett. **117**, 190503 (2016).

[63] Z. K. Minev, S. O. Mundhada, S. Shankar, P. Reinhold, R. Gutiérrez-Jáuregui, R. J. Schoelkopf, M. Mirrahimi, H. J. Carmichael, and M. H. Devoret, To catch and reverse a quantum jump mid-flight, Nature **570**, 200 (2019).

[64] R. Lescanne, L. Verney, Q. Ficheux, M. H. Devoret, B. Huard, M. Mirrahimi, and Z. Leghtas, Escape of a Driven Quantum Josephson Circuit Into Unconfined States, Phys. Rev. Appl. **11**, 014030 (2019).

[65] We thank the anonymous referee for pointing out this scenario.

[66] Z. Wang, S. Shankar, Z. K. Minev, P. Campagne-Ibarcq, A. Narla, and M. H. Devoret, Cavity Attenuators for Superconducting Qubits, Phys. Rev. Appl. **11**, 014031 (2019).

[67] W. Smith, A. Kou, X. Xiao, U. Vool, and M. Devoret, Superconducting circuit protected by two-cooper-pair tunneling, npj Quantum Inf. **6**, 8 (2020).

[68] S. Pancharatnam, in *Proceedings of the Indian Academy of Sciences-Section A*, Vol. 44 (Springer, 1956), p. 398.

[69] M. V. Berry, Quantal phase factors accompanying adiabatic changes, Proc. R. Soc. London, A: Math. Phys. Sci. **392**, 45 (1984).

[70] L. Mandelstam and I. Tamm, in *Selected Papers* (Springer, Berlin, 1991), p. 115.

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

[72] C. Wang, C. Axline, Y. Y. Gao, T. Brecht, Y. Chu, L. Frunzio, M. Devoret, and R. J. Schoelkopf, Surface participation and dielectric loss in superconducting qubits, Appl. Phys. Lett. **107**, 162601 (2015).

[73] 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).

[74] 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).

[75] 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).

[76] 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(R) (2012).