

Analytical Blueprint for 99.999% Fidelity X-Gates on Present Superconducting Hardware Under Strong Driving

José Diogo Da Costa Jesus^{1,2,*} Boxi Li¹ Yuan Gao^{3,4} Rami Barends,^{3,4}
Francisco Andrés Cárdenas-López¹ and Felix Motzoi^{1,2}

¹*Institute for Quantum Control (PGI-8), Forschungszentrum Jülich, 52425 Jülich, Germany*

²*Institute for Theoretical Physics, University of Cologne, 50937 Cologne, Germany*

³*Institute for Functional Quantum System (PGI-13), Forschungszentrum Jülich, 52425 Jülich, Germany*

⁴*Department of Physics, RWTH Aachen University, 52074 Aachen, Germany*

 (Received 24 February 2026; revised 22 May 2026; accepted 16 July 2026; published 27 August 2026)

Achieving ultrafast single-qubit gates that approach the limits set by decoherence requires operating in the strong-driving regime, where conventional semi-classical descriptions and single-leakage models break down, and multi-photon transitions emerge as dominant error channels that grow rapidly as the gate time is reduced. This makes it essential to develop systematic methods to characterize and suppress these errors using pulse shapes that remain simple and experimentally straightforward to calibrate. In contrast to standard derivative removal by adiabatic gate (DRAG) approaches for single-qubit gates, which suppress leakage within an effective three-level description, our framework systematically treats multi-photon transitions that arise in the strong-driving regime due to the participation of higher energy levels. In particular, we identify and suppress the two-photon $|1\rangle \leftrightarrow |3\rangle$ channel, incorporate time-ordering corrections, and derive a recursive analytical construction that captures the dominant higher-order error processes. We term the resulting pulse families R1D and R2D. These constructions extend and unify DRAG-type approaches by incorporating higher-order processes and time-ordering effects, yielding analytical expressions for optimal quadrature corrections, detunings, and prefactors. In addition, we clarify the optimal choice of DRAG prefactors and constant detuning when time-ordering effects are properly taken into account. This provides a systematic prescription for calibrating additional correction terms to further improve gate performance. Using this framework, we numerically demonstrate gate infidelities below 10^{-5} for a 7 ns full π -rotation under realistic decoherence rates.

DOI: [10.1103/sdxb-v39h](https://doi.org/10.1103/sdxb-v39h)

I. INTRODUCTION

Superconducting quantum circuits [1,2] have emerged as a leading platform for quantum computation due to their tunability and the high degree of control achievable over individual qubits. State-of-the-art quantum processors based on transmon circuits [3] are well described as weakly anharmonic oscillators. However, this weak anharmonicity allows the control drive to induce unwanted transitions to higher energy levels, resulting in population leakage out of the computational subspace and phase errors, which ultimately limit single-qubit gate fidelities [4].

A widely used strategy to mitigate these errors is pulse shaping via the derivative removal by adiabatic gate

(DRAG) protocol [5–7]. In this approach, an additional quadrature component is introduced to suppress off-resonant single-photon transitions and associated phase errors within a perturbative, few-level description [5,8]. However, as gate times decrease and the drive amplitude approaches the anharmonicity, the perturbative regime breaks down and the conventional three-level description becomes insufficient, with multiple leakage channels simultaneously activated, thus leading to increasingly complex dynamics that are not captured by conventional DRAG leakage modeling [9,10]. While alternative approaches based on spectral shaping have been proposed to suppress specific transitions [11,12], these methods rely on semi-classical approximations that are not systematically valid beyond the perturbative regime.

In this work, we develop a framework that goes beyond these limitations by explicitly accounting for multi-photon processes exiting the computational space and the population of higher energy levels from the transmon circuit. We show that, once single-photon leakage is suppressed, two-photon transitions—particularly the

*Contact author: j.da.costa.jesus@fz-juelich.de

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

$|0\rangle \rightleftharpoons |2\rangle$ the $|1\rangle \rightleftharpoons |3\rangle$ channels—become the dominant sources of error. To address this, we construct recursive DRAG [9]-type corrections that simultaneously suppress single- and two-photon leakage processes through a sequence of frame transformations. This leads to fully analytical pulse families that eliminate the leading error terms in the strong-driving regime, i.e. where the pulse amplitude cannot be considered small compared to the detuning. In particular, we demonstrate that a minimum four-level description is required to capture the relevant dynamics. Using these analytical pulses without additional calibration, we achieve gate errors at the level of 10^{-5} for an 11.8 ns pulse.

Beyond this analytical construction, we provide a systematic understanding of commonly used calibration parameters. While additional control prefactors are often introduced empirically to improve performance, their role and interdependence remain not yet understood [4]. Using a combination of toggling-frame transformations and a Magnus expansion, we explain the observed correlations between these parameters. In particular, we clarify the relationship between the DRAG prefactor and detuning, and show why, counterintuitively, a constant detuning can outperform time-dependent detuning in relevant regimes. By incorporating additional prefactors to the R2D pulses we reduce the gating time from 11.8 to 6.8 ns achieving gate fidelity errors of 10^{-5} .

The remainder of the paper is organized as follows: Section II A introduces the DRAG formalism and the

first-order errors on single-qubit gates. Section II C describes the recursive DRAG approach, termed as R1D, for correcting two-photon errors $|0\rangle \rightleftharpoons |2\rangle$. Section II D generalizes to a second recursive set of frame transformations, termed as R2D, that further remove errors from the higher excited states of the ladder. Section III provides analytical predictions and analysis for the drive amplitude and constant detuning to remove additional computational-space errors. In Sec. IV, we analyze the effect of decoherence on the gate performance. Finally, in Sec. V we provide our conclusions.

II. ERROR BUDGET AND CORRESPONDING CORRECTIONS

A. Linear leakage error and standard DRAG correction

We start by revisiting the DRAG method introduced in Ref. [5], but considering a weakly anharmonic oscillator with four levels, $|0\rangle$, $|1\rangle$, $|2\rangle$ and $|3\rangle$, as depicted in Fig. 1(a). The first two levels are used for information processing—they form the computational subspace—with the rest representing leakage states. DRAG aims to implement operations between the lowest two levels while suppressing leakage and phase errors by engineering the drive envelope. Let us start with our system Hamiltonian in the rotating frame with respect to the drive frequency, taking $\hbar = 1$ (see Appendix A for a full derivation):

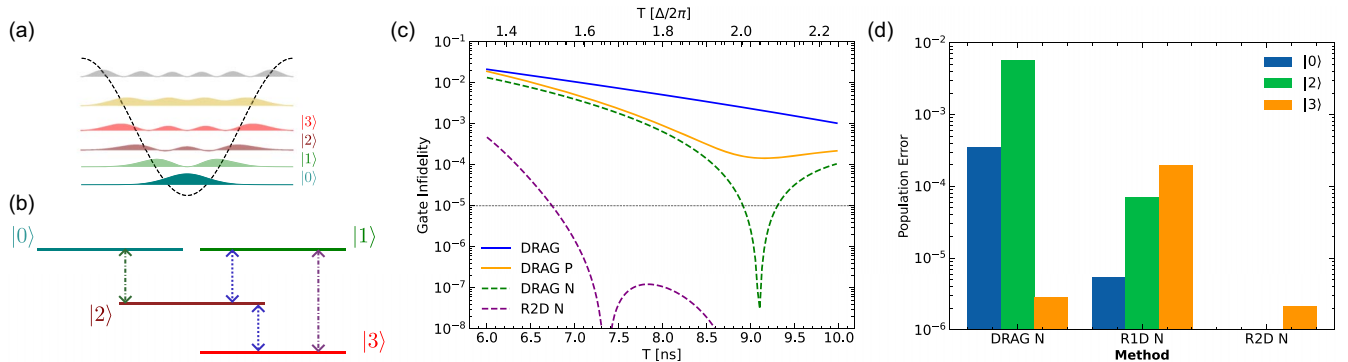


FIG. 1. Pulse performance for a transmon qubit. (a) Low-lying energy structure of a transmon system represented as a particle in a cosine potential, where different opacity refers to the importance of higher energy levels in the performance of the gate. Only the lowest 4 levels are considered in the analytical expressions. (b) Structure of the energy levels of the transmon system in the rotating frame with respect to the transition ω_{10} . The dashed arrows connecting different states represent single-photon transitions, while dotted dashed connects states through two-photon processes. (c) Gate infidelity as a function of gate time for analytical linear DRAG (blue, Sec. II F), linear DRAG using the predictions for α , β and δ_c (DRAG P, orange, Sec. III D), linear DRAG with optimized parameters (DRAG N, dashed green, Sec. III A) and R2D with optimized parameters, α_{12} , α_{02} , α_{13} , β and δ_c (R2D N, dashed purple, Sec. III E). The parameterized R2D pulse outperforms all others for all gate times considered, achieving a fidelity of 99.999% at 6.75 ns and beyond. (d) Histogram with the final population of the relevant leakage states for different calibrated pulses when implementing an X gate at 7 ns. The main source of error in DRAG N, Sec. III A, can be addressed by using R1D N, Appendix E, which will suppress 2 photon transitions to $|2\rangle$. However, for R1D N the leakage to $|3\rangle$ is a major source of error. This can be suppressed by more than an order of magnitude by R2D N, Sec. III E). The simulations were performed by modeling the transmon as an anharmonic oscillator with anharmonicity $\Delta_2 = -2\pi \times 0.225$ GHz and $\Delta_3 = 3\Delta_2$.

$$\hat{\mathcal{H}}(t) = \sum_{j=1}^M \left[(\Delta_j + j\delta(t))|j\rangle\langle j| + \frac{\lambda_j}{2} \sum_{p=x,y} \Omega_p(t) \hat{\sigma}_{j-1,j}^p \right] \quad (1)$$

where Δ_j is the anharmonicity of the j th level, $\delta(t)$ is a time-dependent detuning to be determined later and λ_j denotes the relative drive strength between $|j\rangle \rightleftharpoons |j+1\rangle$. Without loss of generality, we assume $\lambda_1 = 1$. For transmon circuits [3,13], we can approximate $\lambda_j \equiv \sqrt{j}$. The symbols $\Omega_x(t)$ [$\Omega_y(t)$] are the in-phase (out-of-phase) drive envelopes. Finally $\hat{\sigma}_{j,k}^x = |j\rangle\langle k| + |k\rangle\langle j|$ and $\hat{\sigma}_{j,k}^y = i(|j\rangle\langle k| - |k\rangle\langle j|)$ are the generalized Pauli operators acting on multi-level systems.

The goal is to implement an on-resonant interaction between the lowest two levels ($\Delta_1 = 0$) while avoiding any population transfer to higher levels. DRAG achieves this by moving to an effective frame in which the ancillary levels decouple from the computational subspace, and by determining the corresponding corrections to the drive envelopes. The Hamiltonian in this frame reads $\hat{\mathcal{H}}_1(t) = \hat{V}(t)\hat{\mathcal{H}}(t)\hat{V}^\dagger(t) + i\dot{\hat{V}}(t)\hat{V}^\dagger(t)$, where overdot represents derivative with respect to time. Ideally, the frame transformation $\hat{V}$ is chosen such that $\hat{V}(t)\hat{\mathcal{H}}(t)\hat{V}^\dagger(t)$ becomes block diagonal, eliminating couplings between the computational subspace and the ancillary levels [14].

Because computing the exact $\hat{V}$ is typically expensive, a perturbative Schrieffer-Wolff approach [15] is commonly used, though in some cases non-perturbative transformations are possible [6,9,15,16]. Let $\hat{S}_1(t)$ denote the generator of $\hat{V}(t)$, $\hat{V}(t) = \exp(\hat{S}_1(t))$. The transformed Hamiltonian to second order is

$$\begin{aligned} \hat{\mathcal{H}}_1(t) &\approx \hat{\mathcal{H}}(t) + [\hat{S}_1(t), \hat{\mathcal{H}}(t)] \\ &+ \frac{1}{2!} [\hat{S}_1(t), [\hat{S}_1(t), \hat{\mathcal{H}}(t)]] + i\dot{\hat{S}}_1(t). \end{aligned} \quad (2)$$

The generator $\hat{S}_1(t)$ generally corresponds to an anti-hermitian version of the interaction that we want to remove [14]. For the first order it is

$$\hat{S}_1(t) = -i \sum_{j=1}^M \lambda_j \frac{\Omega_x(t)}{2\Delta_2} \hat{\sigma}_{j-1,j}^y. \quad (3)$$

Substituting this into Eq. (2), expanding to second order in Ω_x/Δ_2 , truncating to the lowest four levels yields:

$$\begin{aligned} \hat{\mathcal{H}}_1(t) &\approx \frac{\Omega_x}{2} \hat{\sigma}_{0,1}^x + \left[\frac{\Omega_y}{2} + \frac{\dot{\Omega}_x}{2\Delta_2} \right] (\hat{\sigma}_{0,1}^y + \lambda_2 \hat{\sigma}_{1,2}^y) \\ &+ [\delta(t) + \mathcal{E}_{11} \Omega_x^2] |1\rangle\langle 1| + \Omega_x^2 [\mathcal{E}_{02}^x \hat{\sigma}_{0,2}^x + \mathcal{E}_{13}^x \hat{\sigma}_{1,3}^x] \\ &+ [\Delta_2 + 2\delta(t) + \mathcal{E}_{22} \Omega_x^2] |2\rangle\langle 2| + \Omega_x \mathcal{E}_{23}^x \hat{\sigma}_{2,3}^x \\ &+ [\Delta_3 + 3\delta(t) + \mathcal{E}_{33} \Omega_x^2] |3\rangle\langle 3|, \end{aligned} \quad (4)$$

where to avoid cumbersome notation we omit time-dependency in Ω and use general prefactors $\mathcal{E}_{ij}$ for error in the i, j -th matrix element. Maintaining the approximation order $O(\Omega_x^3/\Delta_2^2)$, here and throughout the rest of this section, yields $\mathcal{E}_{11} = (4 - \lambda_2^2)/4\Delta_2$, $\mathcal{E}_{02}^x = \lambda_2/8\Delta_2$, $\mathcal{E}_{13}^x = \lambda_2\lambda_3(\Delta_3 - 2\Delta_2)/8\Delta_2^2$, $\mathcal{E}_{23}^x = \lambda_3(2\Delta_2 - \Delta_3)/2$, $\mathcal{E}_{22} = (\lambda_3^2(\Delta_2 - 3\Delta_3) + \Delta_2(2 + \lambda_2^2))/(4\Delta_2^2)$, $\mathcal{E}_{33} = (2\Delta_2 + (3\Delta_2 - \Delta_3)\lambda_3^2)/(4\Delta_2^2)$.

Equation (4) is central to our analysis, as it identifies the dominant error mechanisms. The leading error arises from the linear coupling term $\sigma_{1,2}^y$ with transition frequency offset Δ_2 which motivates the standard DRAG correction $\Omega_y(t) = -\dot{\Omega}_x(t)/\Delta_2$ [5] as depicted in Fig. 1(b), where we see the dominant error channels in the rotating frame with respect to the transition frequency ω . The phase accumulated on $|1\rangle$ can be removed by dynamically modulating the detuning following the relation $\delta(t) = -\Omega_x^2(4 - \lambda_2^2)/4\Delta_2$. In practice, this is often approximated by a constant detuning shift [4,8], which we formally explain in subsequent sections. Applying these corrections yields for the computational subspace and its associated transitions:

$$\hat{\mathcal{H}}_1(t) = \frac{1}{2} \Omega_x \hat{\sigma}_{0,1}^x + \Omega_x^2 [\mathcal{E}_{02}^x \hat{\sigma}_{0,2}^x + \mathcal{E}_{13}^x \hat{\sigma}_{1,3}^x] + \dots$$

The remaining leakage channels are thus the two-photon processes $\sigma_{0,2}^x$ and $\sigma_{1,3}^x$. These contributions are negligible for weak drives but become significant as the gate time is reduced and Ω_x/Δ_2 increases, as illustrated in Fig. 1(d), where a significant leakage to $|2\rangle$ can be observed for a calibrated DRAG pulse at 7 ns. In the following subsections, we focus on eliminating these higher-order leakage terms.

B. Rotation angle and fidelity metric

Before proceeding, we first consider the constraints on the initial pulse shape prior to applying the DRAG correction. In general, we require the pulse to be smooth and vanish at both its beginning and end. This leads to the conditions

$$\int_0^T \Omega_x(t) dt = \theta, \quad (5a)$$

$$\Omega_x(0) = \Omega_x(T) = \dot{\Omega}_x(0) = \dot{\Omega}_x(T) = 0, \quad (5b)$$

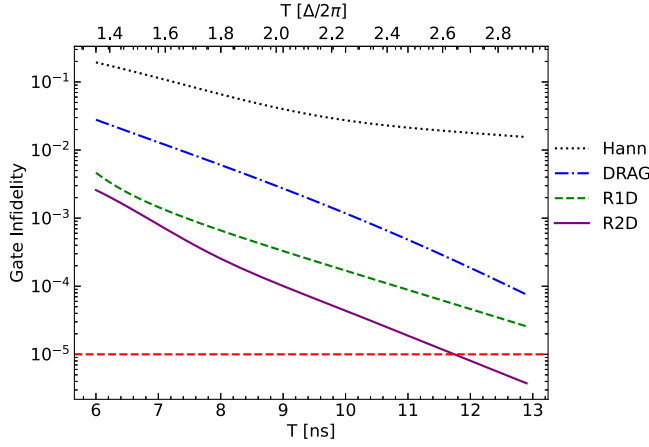


FIG. 2. Gate infidelity for implementing an X gate under different pulses. The black line corresponds to the Hann pulse, the blue line is for standard DRAG, and green and violet lines stand for recursive pulses R1D and R2D, respectively. These pulses are simulated analytically without calibration. R2D outperforms every method and achieves a gate fidelity of 99.99% at 9.1 ns and a fidelity of 99.999% at 11.8 ns.

which guarantee that the frame generated by $\hat{V}(t)$ coincides with the laboratory frame at the beginning and at the end of the dynamics. To benchmark gate performance, we compute the average gate infidelity defined as $\mathcal{E}(\hat{U}_Q) = 1 - \mathcal{F}[\hat{U}_Q]$ where $\mathcal{F}[\hat{U}_Q]$ can be found in Appendix B, following closely [17].

We compare different implementations using the Hann pulse $\Omega_{\text{Hann}}(t) = \Omega_0 \sin^2(\pi t/T)$ and the DRAG pulse $\Omega_{\text{DRAG}}(t) = \Omega_{\text{Hann}}(t) - i\dot{\Omega}_{\text{Hann}}(t)/\Delta_2$ for different gate times T . In Fig. 2, we can see the gate infidelity for both protocols, for an anharmonicity of $\Delta_2/2\pi = -0.225$ GHz. This value for anharmonicity is used for all simulations throughout this paper. Although we only considered 4 levels for the analytical derivations, all simulations were performed considering 6 levels to confirm the validity of the model considered.

We observe that linear DRAG outperforms the Hann pulse at every gate time, obtaining errors below 10^{-4} for times longer than ≈ 12.7 ns. However, for gate times shorter than ≈ 10 ns we see that the performance of DRAG falls sharply with fidelities below 99.9%, demonstrating that at shorter times additional leakage channels (two-photon transitions) are the leading sources of errors.

An illustration of the pulse shape obtained using DRAG can be found in Fig. 3 as well as the Fourier spectrum of the pulse, $\Omega = \Omega_x + i\Omega_y$, and the Fourier spectrum of the square of its real component Ω_x^2 . Although we have suppressed the $|1\rangle \rightleftharpoons |2\rangle$ transition, as can be seen by the hole in the Fourier spectrum of Ω at Δ_2 , transitions $|0\rangle \rightleftharpoons |2\rangle$ and $|1\rangle \rightleftharpoons |3\rangle$ are left unsuppressed, as seen by the Fourier spectrum of Ω_x^2 .

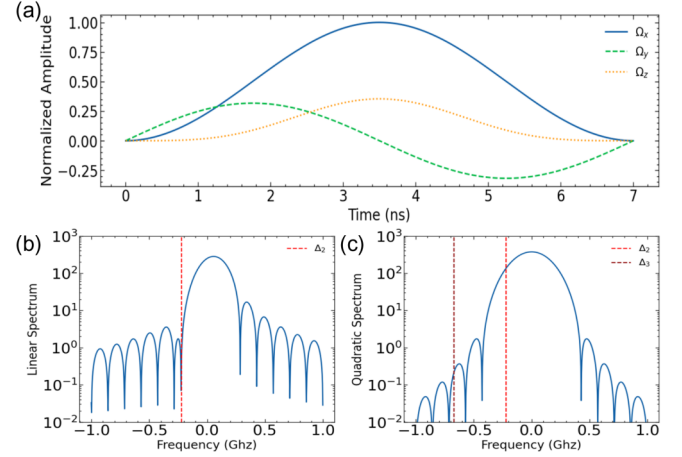


FIG. 3. DRAG pulse shape and Fourier spectrum of Ω and Ω_x^2 . (a) Pulse shape of a DRAG pulse starting with $\Omega_x = \Omega_0 \sin^2(2\pi t/T)$ for a gate time of $T = 7$ ns. (b) Fourier spectrum of a DRAG pulse, $\Omega = \Omega_x + i\Omega_y$. As can be seen, a whole in the spectrum has been added to the frequency Δ_2 , suppressing the $|1\rangle \rightleftharpoons |2\rangle$ transition. (c) Fourier spectrum of Ω_x^2 . As can be seen, we have significant powers at Δ_2 and Δ_3 , the corresponding frequencies in Ω_x^2 for the $|0\rangle \rightleftharpoons |2\rangle$ and $|1\rangle \rightleftharpoons |3\rangle$ transition, which need to be suppressed.

C. Virtual leakage $|0\rangle \rightleftharpoons |2\rangle$ and R1D pulse

As mentioned in the previous subsection, to obtain a shorter gate time without compromising on fidelity, we need to address multiple leakage pathways arising from two-photon transitions. To simultaneously mitigate these errors, we employ the DRAG method recursively [6,7,9], where each successive frame transformation targets a specific leakage channel. This method was first employed in [6] to target the two-photon process $\hat{\sigma}_{0,2}^x$, in a simplified 3-level ladder transmon model. Here we review this approach in the context of 4-level ladder system leakage.

To eliminate the first two-photon process $\hat{\sigma}_{0,2}^x$, we move to a new frame generated by $\hat{S}_2(t) = -i\Omega_1^2 \mathcal{E}_{02}^x \hat{\sigma}_{0,2}^y / \Delta_2$, where Ω_1 is now a new trial function. This generator was chosen as in Sec. II B, i.e., to eliminate a generic transition element $\Omega\sigma^x$, we use the generator proportional to $-i\Omega\sigma^y$. This transforms the Hamiltonian to

$$\hat{\mathcal{H}}_2(t) \approx \hat{\mathcal{H}}_1(t) + \mathcal{E}_{02}^x \left[\frac{2\Omega_1 \dot{\Omega}_1}{\Delta_2} \hat{\sigma}_{0,2}^y - \Omega_1^2 \hat{\sigma}_{0,2}^x \right]. \quad (6)$$

In principle, one may introduce here a complex field correction to Ω_x , as in Ref. [9]; however, this will generate a large out-of-phase component and complicates the gate calibration due to large dynamical phases. Instead, we apply an additional virtual correction proportional to $\hat{\sigma}_{0,2}^y$, via $\hat{S}_3 = i\mathcal{E}_{02}^x 2\Omega_1 \dot{\Omega}_1 \hat{\sigma}_{0,2}^y / \Delta_2^2$, that enables the resulting error term to be real [6,18], leading to

$$\hat{\mathcal{H}}_3(t) \approx \hat{\mathcal{H}}_1(t) - \mathcal{E}_{02}^x \left[2 \frac{\Omega_1 \ddot{\Omega}_1 + \dot{\Omega}_1^2}{\Delta_2^2} + \Omega_1^2 \right] \hat{\sigma}_{0,2}^x. \quad (7)$$

Collecting all terms involving $\hat{\sigma}_{0,2}^x$ and setting them to zero yields the first recursive correction (R1D) pulse:

$$\Omega_x(t) = \sqrt{\Omega_1^2 + \frac{2}{\Delta_2^2} (\dot{\Omega}_1^2 + \ddot{\Omega}_1 \Omega_1)}. \quad (8)$$

We require that the R1D pulse satisfies the boundary conditions given in Eq. (5b). By direct inspection, $\Omega_x(t)$ vanishes at $t = 0$ and $t = T$, provided that Ω_1 obeys the same boundary conditions. However, $\dot{\Omega}_x(t)$ also needs to vanish at $t = 0$ and $t = T$, which imposes a new constraint

$$\text{on } \Omega_1(t). \text{ Evaluating } \lim_{t \rightarrow 0} \dot{\Omega}_x = \lim_{t \rightarrow 0} \frac{\Omega_1 + \frac{1}{\Delta_2^2} (3\dot{\Omega}_1 + \ddot{\Omega}_1 \frac{\Omega_1}{\dot{\Omega}_1})}{\sqrt{(\frac{\Omega_1}{\dot{\Omega}_1})^2 + \frac{2}{\Delta_2^2} (1 + \ddot{\Omega}_1 \frac{\Omega_1}{\dot{\Omega}_1})}}$$

and using $\lim_{t \rightarrow 0} \Omega_1(t)/\dot{\Omega}_1(t) = 0$, the expression reduces to $\lim_{t \rightarrow 0} \dot{\Omega}_x = 3\ddot{\Omega}_1/\sqrt{2\Delta_2^2}$. Thus satisfying the boundary conditions requires $\ddot{\Omega}_1(0) = \ddot{\Omega}_1(T) = 0$. The Hann pulse does not meet this additional condition, so instead we use $\Omega_1(t) = \Omega_0 \sin^3(\pi t/T)$. Relying on a higher power in the sinusoidal envelope does however sharpen the pulse in the time domain, which in turn broadens its frequency spectrum and potentially increases spectral content in unwanted transitions if no correction terms are added, as can be seen in Fig. 4(c).

Using R1D, we implement a π -rotation along the x axis. It's infidelity as a function of gate time T is shown in Fig. 2. As can be seen, this technique outperforms standard DRAG for all times, but especially for extremely fast gates, achieving a gate fidelity above 99.99% for gate times longer than ≈ 10.9 ns, a 14% improvement in minimum gate time.

In Fig. 4 we can see the pulse shape for a R1D pulse as well as its Fourier spectrum and the Fourier spectrum of the square of its real component. We have maintained the suppression of $|1\rangle \rightleftharpoons |2\rangle$ and suppressed the $|0\rangle \rightleftharpoons |2\rangle$ transition, as seen by the hole in the frequency spectrum of Ω_x^2 at Δ_2 . However, $|1\rangle \rightleftharpoons |3\rangle$ not only remains unaddressed, but slightly worsened, and needs to be suppressed.

D. Virtual leakage $|1\rangle \rightleftharpoons |3\rangle$ and R2D pulse

The remaining bottleneck arises from leakage to $|3\rangle$, a two-photon process that has not been previously analyzed. After inserting the first recursion, the Hamiltonian restricted to the computational subspace and its transitions reads

$$\hat{\mathcal{H}}_3(t) \approx \frac{1}{2} \Omega_x \sigma_{0,1}^x + \Omega_x \mathcal{E}_{13}^x \hat{\sigma}_{1,3}^x. \quad (9)$$

To suppress this remaining transition, we introduce a second recursive DRAG (R2D), targeting the $|1\rangle \rightleftharpoons |3\rangle$ transition.

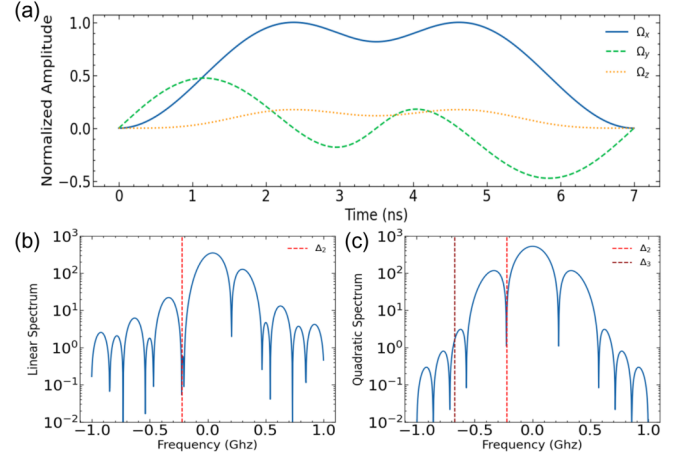


FIG. 4. R1D pulse shape and Fourier spectrum of Ω and Ω_x^2 . (a) Pulse shape of a R1D pulse starting with $\Omega_x = \Omega_0 \sin^3(2\pi t/T)$ for a gate time of $T = 7$ ns. (b) Fourier spectrum of a R1D pulse, $\Omega = \Omega_x + i\Omega_y$. The hole in the frequency spectrum associated with the $|1\rangle \rightleftharpoons |2\rangle$ transition at Δ_2 is preserved. (c) Fourier spectrum of Ω_x^2 . As can be seen, we have introduced a hole at Δ_2 , suppressing the $|0\rangle \rightleftharpoons |2\rangle$ transition. However the $|1\rangle \rightleftharpoons |3\rangle$ transition has not yet been addressed and the increase in the power of the sin, from $\sin^2$ to $\sin^3$ has increased the power at this transition and therefore its importance.

Because Ω_x has already been defined as a functional of Ω_1 , the next recursion must update Ω_1 itself. Directly doing so is rather challenging due to its derivative terms introduced by Eq. (8). We therefore perform a sequence of virtual transformations to eliminate these derivative contributions, which are frame transformations that do not modify the pulse shape. We introduce Schrieffer-Wolff generators with yet a new trial waveform Ω_V so that

$$\hat{\mathcal{S}}_4(t) = -i \frac{\Omega_V^2 \mathcal{E}_{13}^x}{\Delta_3} \hat{\sigma}_{1,3}^y, \quad \hat{\mathcal{S}}_5(t) = -i \frac{2\Omega_V \dot{\Omega}_V \mathcal{E}_{13}^x}{\Delta_3^2} \hat{\sigma}_{1,3}^x, \quad (10)$$

which together yield:

$$\hat{\mathcal{H}}_4(t) \approx \hat{\mathcal{H}}_3(t) - \mathcal{E}_{13}^x \left[2 \frac{\dot{\Omega}_V^2 + \ddot{\Omega}_V \Omega_V}{\Delta_3^2} + \Omega_V^2 \right] \hat{\sigma}_{1,3}^x. \quad (11)$$

The role of these transformations is to eliminate derivatives of Ω_1 in the $|1\rangle \rightleftharpoons |3\rangle$ transition, which leads to the constraint

$$\frac{2}{\Delta_2^2} (\dot{\Omega}_1^2 + \ddot{\Omega}_1 \Omega_1) - \frac{2}{\Delta_3^2} (\dot{\Omega}_V^2 + \ddot{\Omega}_V \Omega_V) + \Omega_V^2 - \Omega_1^2 = 0. \quad (12)$$

Choosing $\Omega_V = \Delta_3 \Omega_1 / \Delta_2$ reduces the remaining coupling in $\mathcal{H}_4(t)$ to a single Ω_1^2 term.

$$\hat{\mathcal{H}}_4(t) \approx \frac{\Omega_x}{2} \sigma_{0,1}^x + \Omega_1^2 \mathcal{E}_{13}^x \left(\frac{\Delta_3^2}{\Delta_2^2} - 1 \right) \hat{\sigma}_{1,3}^x. \quad (13)$$

We now eliminate this term using a second recursive transformation using a new trial waveform, Ω_2 , given by the generators:

$$\hat{\mathcal{S}}_6(t) = -i \frac{\Omega_2^2}{\Delta_3} \mathcal{E}_{13}^x \left(\frac{\Delta_3^2}{\Delta_2^2} - 1 \right) \hat{\sigma}_{1,3}^y, \quad (14a)$$

$$\hat{\mathcal{S}}_7(t) = i \frac{2\Omega_2 \dot{\Omega}_2}{\Delta_3} \mathcal{E}_{13}^x \left(\frac{\Delta_3^2}{\Delta_2^2} - 1 \right) \hat{\sigma}_{1,3}^x, \quad (14b)$$

yielding the second recursive (R2D) DRAG relation:

$$\Omega_1(t) = \sqrt{\Omega_2^2 + \frac{2}{\Delta_3^2} (\dot{\Omega}_2^2 + \ddot{\Omega}_2 \Omega_2)} \quad (15)$$

As before, we must determine the boundary conditions for the new trial pulse Ω_2 . We obtain that $\lim_{t \rightarrow 0} \ddot{\Omega}_1 = 3\ddot{\Omega}_2 / \sqrt{2\Delta_3^2}$ which requires $\ddot{\Omega}_1(0) = \ddot{\Omega}_1(T) = 0$. A possible solution is $\Omega_2(t) = \Omega_0 \sin^4(\pi t/T)$.

E. Starting pulse shape

Because of the sharpness of the new pulse envelope, we expect its frequency spectrum to be broader and increasingly address other leakage transitions. Likewise, the higher the power in the pulse envelope, the larger the amplitude and consequently the higher order terms revealed by the DRAG frame may result in no improvement on the gate fidelity. However, this is not a limiting factor of the framework, instead it forces us to make an adequate selection of the initial pulse envelope that satisfies all the required boundary conditions and at the same time limits the maximal drive amplitude as much as possible. To this end, we use the pulse proposed in Ref. [10],

$$\Omega_2(t) = \Omega_I(t) = \Omega_0 \left[\frac{1}{16} \cos \frac{6\pi}{T} t - \frac{9}{16} \cos \frac{2\pi}{T} t + \frac{1}{2} \right]. \quad (16)$$

It is straightforward to demonstrate that $\Omega_I(t)$ satisfies all the boundary conditions required by both R1D and R2D. Moreover, this pulse has the advantage that the maximal amplitude is smaller than the extension of the Hann pulse as $\sin^4(\pi t/T)$.

Figure 1(d) shows the error budget for the different pulses, and that for ultrafast gates the need becomes apparent to remove the $|1\rangle \rightleftharpoons |3\rangle$ transition. If we do not remove it, this becomes the limiting factor. By using the second square root we can obtain a decrease by more than an order of magnitude in the population of $|3\rangle$.

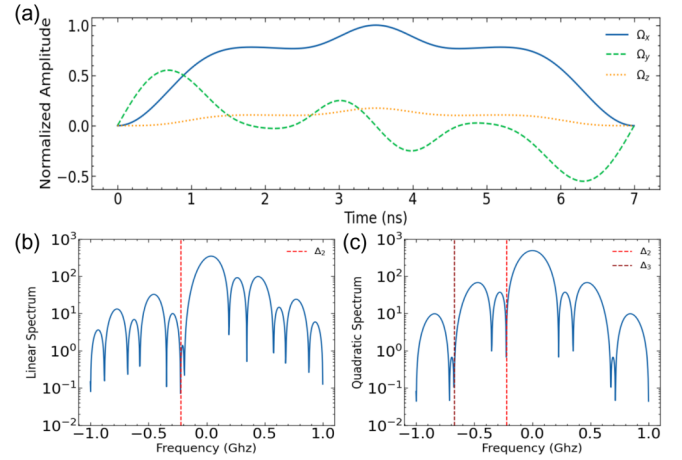


FIG. 5. R2D pulse shape and Fourier spectrum of Ω and Ω_x^2 . (a) Pulse shape of a R2D pulse starting with $\Omega_x = \Omega_I$ for a gate time of $T = 7$ ns. (b) Fourier spectrum of a R2D pulse, $\Omega = \Omega_x + i\Omega_y$. The whole in the frequency spectrum associated with the $|1\rangle \rightleftharpoons |2\rangle$ transition at Δ_2 is preserved. (c) Fourier spectrum of Ω_x^2 . As can be seen, we now have a hole at Δ_2 and Δ_3 suppressing both $|0\rangle \rightleftharpoons |2\rangle$ and $|1\rangle \rightleftharpoons |3\rangle$.

In Fig. 5 we can see the pulse shape for a R2D pulse as well as its Fourier spectrum and the Fourier spectrum of the square of its real component. We have maintained the suppression of $|1\rangle \rightleftharpoons |2\rangle$ and the suppression of $|0\rangle \rightleftharpoons |2\rangle$, as seen by their respective holes in the frequency spectrum of Ω and Ω_x^2 . Additionally, we introduced a hole at Δ_3 in the frequency spectrum of Ω_x^2 suppressing $|1\rangle \rightleftharpoons |3\rangle$.

F. Superlinear corrections within the qubit manifold

We have demonstrated that it is possible to systematically reduce the error for single-qubit gates by including different recursions of the DRAG framework leading to significant error reduction. However, it is possible to reduce the error even more, achieving fidelities around 99.999%, by including corrections that involve higher order contributions to the X and Y component of the pulses. These corrections are obtained by expanding $\hat{\mathcal{H}}_4(t)$ Eq. (13) up to third order in Ω_x/Δ_2 , whose form is shown explicitly in Appendix C, Eq. (C1), or in simplified notation:

$$\begin{aligned} \hat{\mathcal{H}}_5(t) \approx & \left[\frac{\Omega_x}{2} + \mathcal{E}_{01}^x \Omega_x^3 \right] \hat{\sigma}_{0,1}^x \\ & + \Omega_x (\mathcal{E}_{12}^{y1} \Omega_1 \dot{\Omega}_1 + \mathcal{E}_{12}^{y2} \Omega_2 \dot{\Omega}_2) \hat{\sigma}_{1,2}^y \\ & + \Omega_x (\mathcal{E}_{12}^{x0} \Omega_x^2 + \mathcal{E}_{12}^{x1} \Omega_1^2 + \mathcal{E}_{12}^{x2} \Omega_2^2) \hat{\sigma}_{1,2}^x \\ & + \Omega_x (\mathcal{E}_{03}^{y1} \Omega_1 \dot{\Omega}_1 + \mathcal{E}_{03}^{y2} \Omega_2 \dot{\Omega}_2) \hat{\sigma}_{0,3}^y \\ & + \Omega_x (\mathcal{E}_{03}^{x0} \Omega_x^2 + \mathcal{E}_{03}^{x1} \Omega_1^2 + \mathcal{E}_{03}^{x2} \Omega_2^2) \hat{\sigma}_{0,3}^x + O(\Omega^4) \end{aligned} \quad (17)$$

To avoid over/under rotation on the computational subspace, we need to modify the pulse envelope to cancel the higher order correction $\mathcal{E}_{01}^x \Omega_x^3$ by updating the pulse with

$$\Omega'_x(t) = \Omega_x(t) - \frac{(4 - \lambda_2^2)\Omega_x^3(t)}{8\Delta_2^2}. \quad (18)$$

The remaining higher order terms in $|1\rangle \rightleftharpoons |2\rangle$ can be removed by changing the Y component of the pulse, which needs to be updated according to the effective Hamiltonian obtained through two transformations. The first is a modification of initial DRAG transformation (3):

$$\hat{S}_8(t) = -i \sum_{j=1}^m \frac{\lambda_j}{2} \left[\frac{\Omega_x}{\Delta_2} + \frac{A_{0,1}(t)}{\Delta_2^3} \right] \hat{\sigma}_{j-1,j}^y \quad (19)$$

where $A_{0,1}(t)$ is an auxiliary variable defined in Appendix C Eq. (C13) and must be applied as the first step in the analysis. This frame transformation will induce an extra term $-\frac{\lambda_2}{2\Delta_2^2} A_{0,1}(t) \hat{\sigma}_{1,2}^x$ in Eq. (C1) that can be used to remove the extra terms in $\hat{\sigma}_{1,2}^x$. The second frame transformation changes the terms in $\hat{\sigma}_{1,2}^y$ to terms in $\hat{\sigma}_{1,2}^x$ so that all superlinear terms may be corrected and it is defined as:

$$\hat{S}_9(t) = -\frac{i}{\Delta_2} \Omega_x (\mathcal{E}_{12}^{y1} \Omega_1 \dot{\Omega}_1 + \mathcal{E}_{12}^{y2} \Omega_2 \dot{\Omega}_2) \hat{\sigma}_{1,2}^x \quad (20)$$

To remove both the Y component introduced in the pulse and the Stark-shift we perform the following pulse updates:

$$\Omega'_y(t) = -\frac{\dot{\Omega}'_x(t)}{\Delta_2} - \frac{\dot{A}_{0,1}(t)}{\Delta_2^3}, \quad \delta' = \delta - \frac{A_{0,1}(t)\Omega_x}{\Delta_2^3}. \quad (21)$$

These corrections in Eqs. (18) and (21) are not exclusive to R2D and can be performed to every DRAG framework. In practice, they can also be corrected by optimizing the amplitude of pulses in experiments. Therefore, the corrections were included in all frameworks in Fig. 2 to demonstrate a fair comparison of all methods under such higher order optimization. Note that for linear DRAG these expressions simplify substantially and $A_{0,1}(t) = (4 - \lambda_2^2)\Omega_x^3(t)/(8\Delta_2^2)$. This can be implemented by using Ω'_x and simply taking $\Omega'_y(t) = -\dot{\Omega}'_x(t)/\Delta_2$ and therefore no Y correction is necessary.

Using the corrected R2D pulse, we obtain errors an order of magnitude smaller than standard DRAG above 6 ns and R1D above 12 ns. As can be seen in Fig. 2(c), a fidelity of 99.999% is obtained for pulses from 12 ns.

After the correction, we obtain a remaining leakage term involving the three-photon process in $|0\rangle \rightleftharpoons |3\rangle$:

$$\begin{aligned} \hat{\mathcal{H}}_6(t) \approx & \left[\frac{\Omega'_x}{2} \right] \hat{\sigma}_{0,1}^x + \Omega_x (\mathcal{E}_{03}^{y1} \Omega_1 \dot{\Omega}_1 + \mathcal{E}_{03}^{y2} \Omega_2 \dot{\Omega}_2) \hat{\sigma}_{0,3}^y \\ & + \Omega_x (\mathcal{E}_{03}^{x0} \Omega_x^2 + \mathcal{E}_{03}^{x1} \Omega_1^2 + \mathcal{E}_{03}^{x2} \Omega_2^2) \hat{\sigma}_{0,3}^x + O(\Omega^4) \end{aligned} \quad (22)$$

These terms are difficult to remove and cannot be easily simplified by virtual transitions due to the fact that all terms are proportional to Ω_x and Ω_x is dependent on $\dot{\Omega}_1$, $\dot{\Omega}_2$, $\ddot{\Omega}_2$.

G. Minimum gate time

In Eq. (8) we have expressed the real component of our pulse, Ω_x , in terms of a square root. The terms in this square root have to be positive during the gate duration, since the imaginary component of the pulse, Ω_y , has already been defined. Due to the second derivative, we introduce negative components that will establish a minimum gate time for a full rotation of the R1D pulse considered, $T_{\min} = \sqrt{6}\pi/\Delta_2$.

We can generalize this expression to several orders of recursion by assuming that the initial pulse has as shape $\Omega_1(t) = \Omega_0 \sin^n(\pi t/T)$ with the condition $\Omega_1^2(T_{\min}) + 2[\dot{\Omega}_1^2(T_{\min}) + \ddot{\Omega}_1(T_{\min})\Omega_1(T_{\min})]/\Delta_2^2 = 0$. By replacing the argument of $\Omega_x(t)$ from Eq. (8), we get

$$\left[1 - \frac{2n\pi^2}{\Delta_2^2 T_{\min}^2} \right] \sin^2 \left[\frac{\pi t}{T_{\min}} \right] + \frac{(4n^2 - 2n)\pi^2}{\Delta_2^2 T_{\min}^2} \cos^2 \left[\frac{\pi t}{T_{\min}} \right] = 0. \quad (23)$$

We need to guarantee that this expression is positive for all times. However, we know that the minimum of Eq. (23) is achieved when $t = T_{\min}/2$ (so that $\cos^2[\frac{\pi t}{T_{\min}}] = 0$ and we no longer have the terms proportional to $\cos^2$, which are always positive), reducing the expression to

$$T_{\min} = \sqrt{2n} \frac{\pi}{\Delta_2}. \quad (24)$$

For the R2D pulse, we require that both square roots remain positive. Following the same procedure as above we obtain the following:

$$T_{\min} = \pi \sqrt{\frac{n}{\Delta_3^2} + \frac{n}{\Delta_2^2} + \sqrt{\frac{n^2}{\Delta_3^4} + \frac{n^2}{\Delta_2^4} + \frac{4n - 10n^2}{\Delta_2^2 \Delta_3^2}}} \quad (25)$$

which reduces to the previous equation under the assumption $\Delta_3 \gg \Delta_2$. This leads us to the interesting and unexpected finding that, for the same Δ_2 , the minimum gate time $T_{\min}$ for R2D increases with increasing Δ_3 , that is, the bigger the separation between $|2\rangle$ and $|3\rangle$ for the

same anharmonicity, the slower the gate (for the same n). This is a consequence of the first recursion reducing the amplitude of Ω_1 , when compared to Ω_2 , which allows for greater values of $\ddot{\Omega}_2$. However, if $\Delta_3 \gg \Delta_2$, $\Omega_1 \approx \Omega_2$ and the first recursion has no effect on pulse shape.

The minimum gate time for R1D and R2D using the pulse in Ref. [10] is obtained by ensuring both square roots are positive. This can be solved numerically and, for an anharmonicity of $\Delta_2/2\pi = -0.225$ GHz gives us a gate time of ≈ 4.45 ns.

III. PREFACTOR VALUES FROM ANALYTICAL TIME-ORDERING AND NUMERICAL OPTIMIZATION

It is now common practice to optimize the coefficient for linear DRAG in single qubit gates [4,19]. Yet the underlying physics governing these optimized corrections is seldom clarified. This gap arises both from discrepancies between the idealized model and the actual device, such as parameter drift [20–22], distortion on the pulses [23,24] and crosstalk [18], and from the presence of additional destructive interference pathways among different error channels. While treating the calibration as a black-box optimization has enabled state-of-the-art fidelities, further progress entails a deeper understanding of the dynamics.

In this section, we analyze the role of each free parameter and develop a systematic framework for Hamiltonian engineering. Our analysis reveals how these parameters interact, identifies which error channels they influence, and provides analytical expressions for their near-optimal values.

A. Parameterization of DRAG solution families

For linear DRAG, a set of tunable parameters is given by the triad $\{\beta, \alpha, \delta_c\}$, first introduced in [8], that appears in the control as

$$\Omega \rightarrow \beta(\Omega_x + \alpha\Omega_y), \quad \delta \rightarrow \delta_c. \quad (26)$$

where β parameterizes changes in the total amplitude and α the ratio for the Y component. The parameter δ_c is a constant detuning correction due to frequency mismatch between the carrier frequency of the drive and the real transition frequency of the system, introduced by the ac-Stark shift. This set of parameters follows the one used in Ref. [4].

Figure 1(c) compares the gate infidelity $\mathcal{E}(\hat{\mathcal{U}}_Q)$ as a function of the gate time for different pulses. We observe that the optimized version of linear DRAG (represented by the acronym DRAG N, the green dashed line) always outperforms the analytical one for gate times longer than 6 ns (the blue line). We find that around 9.1 ns we have a destructive interference between the remaining

uncorrected leakage channels, such that we can eliminate all leakage remaining. This way, between $T = 8.93$ ns and $T = 9.33$ ns we achieve the threshold of error below 10^{-5} , assuming sufficiently low decoherence. Moreover, we observe that this numerical optimization exhibits quasi-periodic peaks that are monotonically decreasing in the gate fidelity, though these are outside the range of the region of interest that is plotted, indicating an effective recurrence time. Note however that picking the precise time duration for a low-infidelity gate is known [25] to be problematic for a number of reasons. Aside from pulse timing errors, the calibration process becomes much more cumbersome as in addition to the calibration of the pulse parameters, the gate duration must also be calibrated. In a multi-qubit processor, these optimal durations will moreover vary from qubit to qubit, which prohibits a common clock frequency for the processor, rendering quantum circuits highly cumbersome. Therefore, one cannot in general count on such an approach for ultra-low error.

In the next sections, we now aim to further understand the underlying physics behind these parameters, and why such an increase in performance is gained. Therefore, we modify Eq. (3) to consider a continuous set of adiabatic transformation given by [8]

$$\hat{S}_D(t) = -i\lambda_1 \frac{\alpha\Omega_x}{2\Delta_2} \hat{\sigma}_{0,1}^y - i \sum_{j=2}^m \lambda_j \frac{\Omega_x}{2\Delta_2} \hat{\sigma}_{j-1,j}^y. \quad (27)$$

We use this frame to ensure no terms proportional to $\hat{\sigma}_{0,1}^y$ remain in the computational subspace while still eliminating the leakage term $\hat{\sigma}_{1,2}^x$.

B. Optimal detuning

We update the triad $\{\beta, \alpha, \delta_c\}$ such that both leakage and phase errors cancel at the same time. To do so, we transform the starting Hamiltonian, Eq. (1), and expand to third order in $O(\Omega_x/\Delta)$, using the new frame generator, Eq. (27). Additionally, we move to the *toggling frame* [26,27] given by the unitary which will represent ideal dynamics:

$$\hat{V}_{\text{tog}}(\theta) = \exp(-i\theta\sigma_{0,1}^x + \phi_1|2\rangle\langle 2| + \phi_3|3\rangle\langle 3|) \quad (28)$$

where the rotation angle is given by $\theta(t) = \int_0^t \Omega_x(s)ds/2$ and ϕ_1 and ϕ_2 allow for free precession beyond the qubit subspace. Within the qubit subspace, this gives

$$\begin{aligned} \hat{\mathcal{H}}_{\text{tog}}^{(\text{qub})} = & ([\delta_c + \mathcal{E}_{11}\Omega_x^2] \cos \theta + \mathcal{E}_{11}^{(\alpha)} \Omega_x \dot{\Omega}_x \sin \theta) |1\rangle\langle 1| \\ & + (\Omega_x - \dot{\theta} + i[\delta_c + \mathcal{E}_{11}\Omega_x^2] \sin \theta - \mathcal{E}_{01}^x \Omega_x^3 \\ & + \mathcal{E}_{01}^x \Omega_x \delta_c + i\mathcal{E}_{11}^{(\alpha)} \Omega_x \dot{\Omega}_x \cos \theta) \frac{\sigma_{0,1}^x}{2}, \end{aligned} \quad (29)$$

where the $\mathcal{E}$ prefactors are given in Appendix D, as well as remaining transitions to higher levels. It is important to note that $\hat{V}_{\text{tog}}(T) \neq I$, and in fact $\hat{V}_{\text{tog}}(T)$ generates the desired π rotation. Therefore, the transformed toggling Hamiltonian will represent the error of our gate. If $H_{\text{tog}} = 0$ for all time, the π rotation can be ideally implemented. For now we ignore the effect of β noting that for $\beta = 1$ we have $\Omega_x = \dot{\theta}$.

The goal of the toggling transformation is to separate the error part of the dynamics H_{tog} , which is typically small. This ensures fast convergence of the Magnus expansion [25,28,29] which allows us to express the propagator as $\hat{T} \exp(-i \int_0^T \hat{\mathcal{H}}_{\text{tog}}(t) dt) = \exp(-i \sum_{\ell} \hat{\mathcal{Z}}_{\ell}(t))$, where $\hat{\mathcal{Z}}_{\ell}(t)$ are the different orders of the expansion defined as

$$\hat{\mathcal{Z}}_1(t) = \int_0^T d\tau \hat{\mathcal{H}}_{\text{tog}}(t), \quad (30a)$$

$$\hat{\mathcal{Z}}_2(t) = \int_0^T dt \int_0^t d\tau [\hat{\mathcal{H}}_{\text{tog}}(t), \hat{\mathcal{H}}_{\text{tog}}(\tau)], \quad (30b)$$

Thus, we shall find the conditions for the parameters introduced in the toggling frame to make $\hat{\mathcal{Z}}_1(t)$ close to 0 as possible.

For a π rotation, the phase error on $|1\rangle\langle 1|$ turns out to be trivially zero due to symmetry. This can be seen because each of the terms in the first line of Eq. (29) is anti-symmetric around the half-time of the operation when θ undergoes the gate period (see also Appendix D). However, in general, one would not be able to make both equations zero using constant detuning. A time-dependent detuning, like the one used in Sec. II, would be required. This also explains why using the DRAG formula with constant detuning for partial rotation is often seen to have

larger error than one would naively expect given the proportionate reduction in pulse amplitude.

The requirement that the $\sigma_{0,1}^x$ term also integrates to zero gives us the tools to predict the value of the constant detuning and the amplitude of the pulse. Zeroing the imaginary part [c.f. Eq. (D2)], we get the equation for δ_c :

$$\delta_c = -\frac{\int_0^T \mathcal{E}_{11} \Omega_x^2 \sin \theta}{\int_0^T \sin \theta} - \frac{\int_0^T \dot{\Omega}_x \mathcal{E}_{11}^{(\alpha)} \Omega_x \cos \theta}{\int_0^T \sin \theta}. \quad (31)$$

Note that this is not just an average of the accumulated time-dependent Stark shift, but is renormalized by the rotation angle θ . Assuming a pulse shape of the form $\Omega_x = 2\pi \sin^2(\pi t/T)/T$ we get:

$$\delta_c \approx 0.712 \frac{\lambda_2^2 - 4\alpha \pi^2}{\Delta_2 T^2}. \quad (32)$$

Figure 6(a) compares the analytical expression using Eq. (32) with numerically optimized results for different DRAG coefficient α and two different gate times $T = 8$ ns and $T = 15$ ns. The level of agreement is governed by the Magnus expansion: as the gate time increases, higher-order terms (whose terms have more commutators of $\hat{\mathcal{H}}_{\text{tog}}$ and terms proportional to higher powers of Ω_x/Δ_2) become negligible, and the analytical and numerical results converge. The constant detuning expression also explains the physics observed in the experimental realization in Ref. [4], where a linear relationship between α and δ_c was reported with $\delta_c = 0$ roughly at $\alpha = 0.5$ (for $\lambda_2 \approx \sqrt{2}$) [8]. The experiment also observed scaling $\delta_c \propto 1/T^2$, which is verified by our analysis. Recent results in Ref. [30] further confirm the strong agreement between this analytical expression and calibrated constant-detuning values.

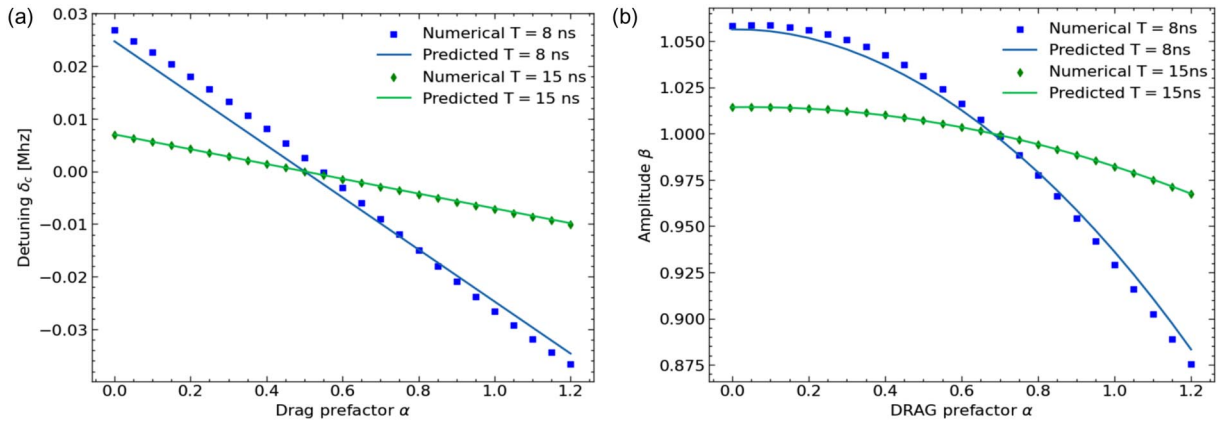


FIG. 6. Amplitude and constant detuning as a function of α . (a) Comparison between the optimized values of constant detuning δ_c for the DRAG pulse and the analytical approximation, for different values of α . The analytical expression can accurately predict the values of constant detuning. The longer the gate time, the better the prediction. (b) Comparison between the optimized values of β for the DRAG pulse and the analytical approximation, for different values of α . Once again, the analytical expression can accurately predict the values of constant detuning.

C. Optimal pulse amplitude

Taking the real part of the integral [Eq. (D2)], we have the following term for the amplitude of the pulse:

$$\int_0^T dt [\beta \Omega_x - \dot{\theta} - \beta \Omega_x (\mathcal{E}_{01}^{x1} \delta_c + \mathcal{E}_{01}^{x0} \beta^2 \Omega_x^2)] \quad (33)$$

which is a cubic polynomial of β and must be set to zero. Despite Eq. (33) having an analytical solution, it is quite cumbersome to solve. Instead, assuming a mean-field approximation as $\beta = 1 + \eta$ by considering a small variation of the expected optimal amplitude, we can linearize Eq. (33) as $\beta^3 \approx 1 + 3\eta$, leading to

$$\beta = 1 + \frac{16\alpha\delta_c\pi\Delta_1 + (4\alpha^2 + \lambda_2^2 - 2\alpha\lambda_2^2)\frac{5\pi^3}{T^2}}{16\pi\Delta_1^2 - 16\alpha\delta_c\pi\Delta_1 - 3(4\alpha^2 + \lambda_2^2 - 2\alpha\lambda_2^2)\frac{5\pi^3}{T^2}}, \quad (34)$$

corresponding to the correction of the drive amplitude as a function of the DRAG coefficient α for fixed gate time T . It also shows how one can correct the X superlinear correction identified in Sec. II F alternatively, through prefactor calibration.

Figure 6(b) compares the amplitude correction β obtained in Eq. (34) with the one obtained through numerical optimization at two different gate times. Just as in the previous case, the agreement between the analytical and the numerical results is conditioned on the validity of the first-order Magnus expansion and improves with gate times, and is generally extremely good.

D. Linear DRAG prefactor

Finally, to understand and predict the value of α we must look into the $|1\rangle\langle 2|$ term of $\hat{\mathcal{Z}}_1$, $\hat{\mathcal{Z}}_1[1, 2]$ (and the term $|2\rangle\langle 1|$, $\hat{\mathcal{Z}}_1[2, 1]$) and the $|0\rangle\langle 2|$ term of $\hat{\mathcal{Z}}_1$, $\hat{\mathcal{Z}}_1[0, 2]$ (and the term $|2\rangle\langle 0|$, $\hat{\mathcal{Z}}_1[2, 0]$) which represent the main leakage channels in our error Hamiltonian (and whose contributions came from a mixture of the $|1\rangle \rightleftharpoons |2\rangle$ and $|0\rangle \rightleftharpoons |2\rangle$ transitions in the lab frame). These elements can be found in Appendix D, Eqs. (D3) and (D4).

In order to minimize the leakage we now want the α that minimizes $l = |\hat{\mathcal{Z}}_1[1, 2]|^2 + |\hat{\mathcal{Z}}_1[0, 2]|^2$. This can be found by taking the derivative $\partial l / \partial \alpha = 0$. To solve this equation, we approximate $\phi_\ell = \Delta_\ell t$, and we once again perform a mean-field approximation, this time over the alphas $\alpha = (1 + \delta\alpha)$ such that $\alpha^n = (1 + n\delta\alpha)$.

Solving this equation we obtain the following expression for α :

$$\alpha \approx 1 - \frac{x_1 + \bar{x}_1}{x_2 + \bar{x}_2}, \quad (35)$$

with x_1 and x_2 defined in Appendix D.

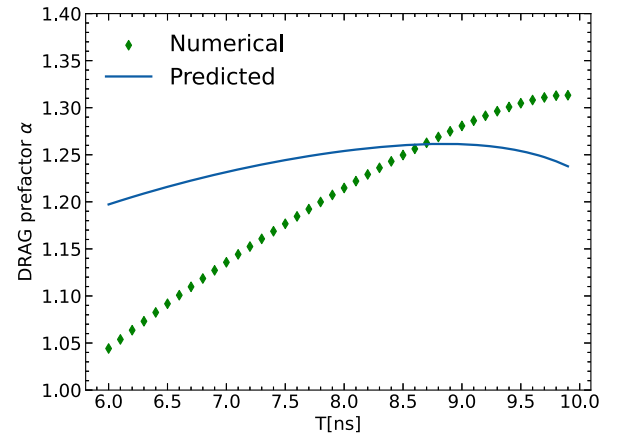


FIG. 7. Analytically vs numerically optimized α parameters. Comparison between the optimized values of α for linear DRAG pulse and the analytical expression of α , as a function of gate time T . Although our expression captures the general behavior of α , we find large discrepancies, most likely due to higher orders.

Figure 7 compares the analytical expression for α with values obtained from full numerical optimization. In principle, $\alpha = 1$ should completely compensate for the single-photon leakage channel $|1\rangle \rightleftharpoons |2\rangle$. The analytical prediction lies around 1.2, reflecting the trade-off between this channel and the concurrent $|0\rangle \rightleftharpoons |2\rangle$ transition and explicitly demonstrating why $\alpha \neq 1$. Numerical optimization reveals a steeper slope for the α -dependence, ranging from 1.0 to 1.3, which arises from higher-order corrections not included in the analytical calculation. The fact that the optimal α is not 1 was first experimentally demonstrated in [4] and it is here explained.

With the prediction of all three parameters, we can now compare the performance between the predicted prefactors and those from full numerical optimization. We plot the infidelity obtained for the DRAG pulse using our predictions of the 3 constants, α , β and δ_c , in Fig. 1(c), respectively termed DRAG P. As seen, despite the discrepancy in the α , the predicted prefactors provide large improvements to gate fidelity, with more than 1 order of magnitude. The largest deviation in fidelity is observed around 9 ns, where destructive interference is present among different leakage channels. In this regime, numerical optimization achieves further gains because it fully explores these interference pathways by optimizing the final unitary directly. Nevertheless, the expressions here obtained can be used as a warm start for further calibration in an experimental setting.

E. Parameter optimization for recursive pulses

The same procedure can be applied to R2D. We replace α with 3 constants, α_{12} , α_{02} and α_{13} , each allowing for higher order corrections in the 1–2, 0–2 and 1–3 transitions, respectively. The introduction of these constants corresponds to replacing α in Eq. (27) with α_{12} and

changing the frame transformations: $\hat{S}_3 \rightarrow \alpha_{02}\hat{S}_3$ and $\hat{S}_7 \rightarrow \alpha_{13}\hat{S}_7$. This changes our pulses to:

$$\Omega_1(t) = \sqrt{\Omega_2^2 + \frac{2\alpha_{13}}{\Delta_3^2}(\dot{\Omega}_2^2 + \ddot{\Omega}_2\Omega_2)} \quad (36)$$

$$\Omega_x(t) = \sqrt{\Omega_1^2 + \frac{2\alpha_{02}}{\Delta_2^2}(\dot{\Omega}_1^2 + \ddot{\Omega}_1\Omega_1)}. \quad (37)$$

and

$$\Omega \rightarrow \beta(\Omega_x + \alpha_{12}\Omega_y), \quad \delta' \rightarrow \delta_c. \quad (38)$$

with $\Omega_2(t) = \sin^4 \pi t/T$. To second order in the error, $O(\Omega_x^2/\Delta_2^2)$, these parameters will be orthogonal and are free to eliminate higher order errors.

We plot the infidelity of such a calibrated pulse, termed R2D N in Fig. 1(c) as a function of time. As can be seen, we achieve significant improvements in fidelity, with a fidelity of 99.999% achieved at $T \approx 6.75$ ns, demonstrating once again the capacity of correcting for higher order terms through calibrating constants. Contrary to linear DRAG, the fidelity stays above this threshold for all times beyond.

Unfortunately, analytical prediction of the three DRAG prefactors becomes intractable, as deviations in any one parameter feed back into the others through higher-order corrections, resulting in a coupled set of nonlinear equations. Nevertheless, we find a good agreement between our simulations and the experimental results obtained in [30]. The parameters can be found in Fig. 8. In contrast to α_{12} and α_{02} , α_{13} becomes bigger as time goes on because the

maximum amplitude of the pulse becomes smaller with increasing α_{13} . This reduces all sources of error at higher orders (and therefore compensates the increase in the 2 photon leakage from $|1\rangle \rightleftharpoons |3\rangle$).

For the amplitude and constant detuning, Eqs. (31) and (33) remain valid to the same order in $O(\Omega_x/\Delta)$, with the replacement rule $\alpha \rightarrow \alpha_{12}$, as can be seen in Figs. 9(a) and 9(b). Interestingly enough, this means that the correction to the amplitude of the pulse and the constant detuning are mostly defined by α_{12} .

One final note regarding the optimization of R2D pulses— α_{02} and α_{13} will change the minimum gate time. In general, a numerical verification is necessary to extract a minimum gate time. However, if $\alpha_{02} > \alpha_{13}$, the expression for minimum gate time can be reduced to:

$$T_{\min} = \frac{\pi \sqrt{n(\alpha_{13}\Delta_2^2 + \alpha_{02}\Delta_3^2) + \nu}}{\Delta_2\Delta_3}, \quad (39)$$

with

$$\nu = \sqrt{n^2(\alpha_{13}^2\Delta_2^4 + \alpha_{02}^2\Delta_3^4) + 2(2-5n)n\alpha_{13}\alpha_{02}\Delta_2^2\Delta_3^2}. \quad (40)$$

IV. DISSIPATIVE DYNAMICS

We calculate the system dynamics considering loss mechanisms including energy relaxation and dephasing. The evolution is described by the master equation in Lindblad form [31]

$$\dot{\rho}(t) = -i[\mathcal{H}(t), \rho(t)] + \gamma \mathcal{D}[\hat{a}]\rho(t) + \gamma_\phi \mathcal{D}[\hat{a}^\dagger \hat{a}]\rho(t), \quad (41)$$

where $\rho(t)$ is the density matrix of the system, $\mathcal{D}[\hat{O}]\bullet = \mathcal{O} \bullet \mathcal{O}^\dagger - \{\mathcal{O}^\dagger \mathcal{O}, \bullet\}/2$ is the Lindbladian super-operator and $\gamma = 1/T_1$ and $\gamma_\phi = 1/T_2 - 1/(2T_1)$ correspond to the decay rates for energy losses and dephasing, respectively. We have chosen 3 sets of T_1 and T_2 , respectively (40, 56) μ s, (350, 350) μ s and (500, 500) μ s, representing a range of values achievable today [32,33]. We also add a fourth set, (1000, 1000) μ s, values which we hope will be reached in the near future. Finally, $\hat{a}$ is the annihilation operator describing the transmon system. As the dissipative dynamics are state-dependent instead of unitary-dependent, we need to evaluate the gate performance for all the states of the Bloch sphere of the two-level system and average them [34].

As can be seen in Fig. 10, R2D is capable of achieving a fidelity of 99.999% even under decoherence, for a gate time of $T = 6.88$ ns, with a minimum in infidelity at $T = 7.33$ ns. This demonstrates that single qubit gates with infidelities below 10^{-5} are possible in today's hardware.

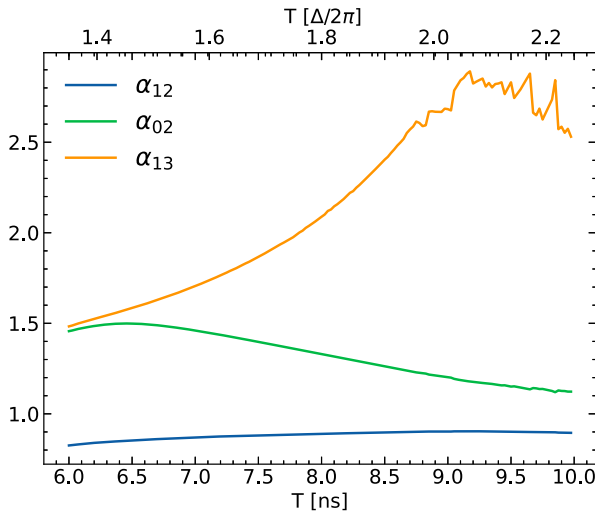


FIG. 8. Numerically optimized α parameters. Optimal values of α_{12} , α_{02} and α_{13} for the R2D pulse as a function of gate time T . These results generally match the ones obtained in experiment [30], providing validation of the theory.

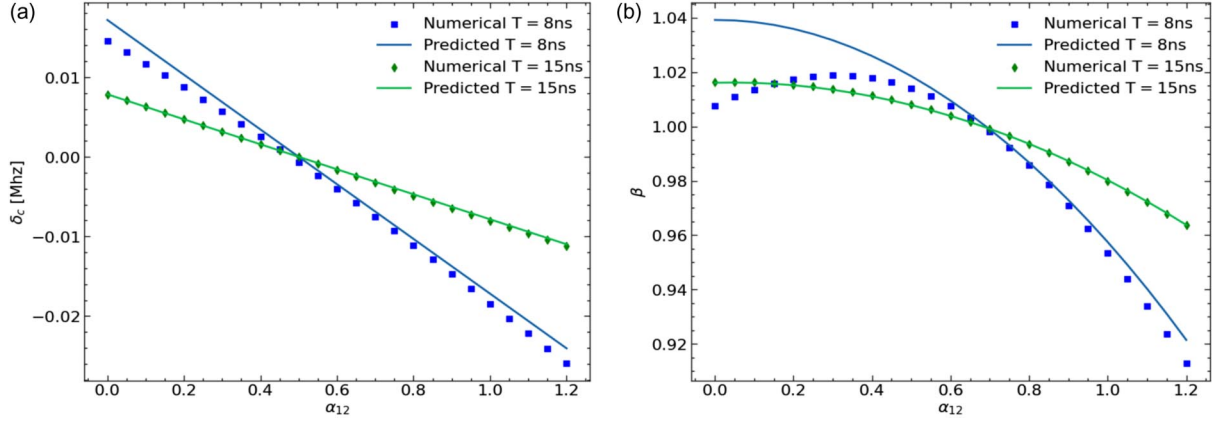


FIG. 9. Amplitude and constant detuning as a function of α_{12} . (a) Comparison between the optimized values of constant detuning for the R2D and the analytical approximation, for different values of α_{12} . To make this example, we chose $\alpha_{02} = 1.3$ and $\alpha_{13} = 1.7$. The analytical expression can accurately predict the values of constant detuning. The longer the gate time, the better the prediction. (b) Comparison between the optimized values of β for the R2D pulse and the analytical approximation, for different values of α_{12} . Once again, the analytical expression can accurately predict the values of β .

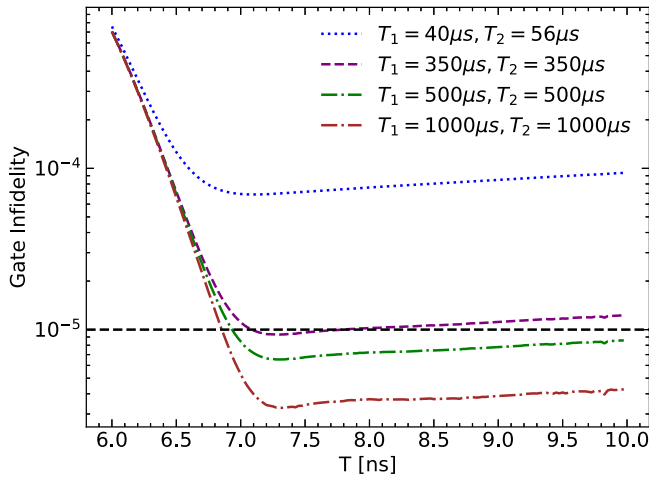


FIG. 10. Gate averaged infidelities for a π gate using optimized parameters under decoherence. As T_1 and T_2 improve, R2D pulses manage to reach 99.999% fidelity.

V. SUMMARY AND CONCLUSION

In this work, we expand on the DRAG framework by including higher levels not usually considered under weak driving and we show that these levels have a real impact on gate fidelity through multi-photon processes. Taking these in consideration, we extract new analytical expressions for pulses that show that it is possible to implement high-fidelity single-qubit gates at faster speeds and stronger driving strengths than previously considered.

We also explore numerical optimization of DRAG and R2D pulses. For the DRAG pulses we obtain theoretical approximations for the calibrating parameters, expressions that we hope will simplify and speed up normal calibration procedures. The obtained expressions for constant detuning are experimentally validated in [30]. For R2D pulses

we show that a moderate optimization allows us to achieve high-fidelity pulses faster, even under the effects of system decoherence, a critical step in enabling future quantum computation.

Experimental implementation of this theoretical framework has since been achieved by Y. Gao *et al.* [30], where R2D pulses were implemented and used to implement single qubit gates with suppressed two-photon leakage channels at world record speeds (6.8 ns) and drive strengths, Ω_x/Δ_2 , while maintaining a fidelity of 99.98%. This gate is mainly limited by decoherence. The difference between the measured value and the simulated is a consequence of heating of the chip during the action of the gate, leading to a large thermal population ($n \approx 0.1$), and the usage of 1 ns of padding time to account for pulse distortion effects. Given that these parameters are chip dependent they were not included in the simulation of Fig. 10.

Although our equations were motivated by the transmon model, R2D is applicable for any ladder system modelled under the rotating wave approximation [6]. Future work is required to go beyond this approximation to remove the remaining three-photon processes in the system, including the $|1\rangle - |4\rangle$ that will require an inclusion of the $|4\rangle$ level. Moreover, going to shorter times and gate times below 5 ns requires better accounting of waveform generation electronics where distortion can be significant, meanwhile the present optimization procedure for R2D may find negative values inside the square root [30]. More generally, one may derive higher order and multi-photon corrections to similar strong driving dynamics in entangling gates [29].

ACKNOWLEDGMENTS

We thank Manuel Guatto, Dimitrios Georgiadis, Shahrukh Chishti, and Asier Galicia for valuable discussions. This work was funded by the Federal Ministry of

Education and Research (BMBF) within the framework programme “Quantum technologies—from basic research to market” (Project QSolid, Grant No. 13N16149) and by Horizon Europe program via project QCED (101080085, HORIZON-CL4-2021-DIGITAL-EMERGING02-10), project OpenSuperQPlus100 (101113946, HORIZON-CL4-2022-QUANTUM-01-SGA) and by the Cluster of Excellence Matter and Light for Quantum Computing (ML4Q2) EXC 2004/2–390534769.

DATA AVAILABILITY

The data that support the findings of this article are openly available [35], embargo periods may apply.

APPENDIX A: CHARGE BASIS HAMILTONIAN

We start by deriving the effective Hamiltonian of a transmon circuit being driven on resonance with the $|0\rangle \rightleftharpoons |1\rangle$ transition. The transmon circuit oscillator is described by the Hamiltonian [3]

$$\hat{\mathcal{H}} = 4E_C[\hat{n} - n_g(t)]^2 - E_J \cos(\hat{\phi}) \quad (\text{A1})$$

where E_C and E_J represent the charge and Josephson energies, and $n_g(t)$ is the dimensionless gate voltage. The operator $\hat{n}$ is the charge operator, indicating the number of Cooper pairs, and $\hat{\phi}$ denotes the phase operator. We can implement single-qubit operations by driving the circuit with an external source using $n_g(t) = n_0(t) \cos(\omega_d t)$ resulting in the Hamiltonian

$$\hat{\mathcal{H}}_{\text{ctrl}} = \Omega(t) \cos(\omega_d t) \hat{n} \quad (\text{A2})$$

where $\Omega(t) = -8E_C n_0(t)$ is the complex drive envelope, and ω_d is the drive frequency. For the low-lying energy spectrum of this device, we can model it as an approximate Duffing oscillator [2,13], where the controls now are given by $\hat{n} \propto (\hat{b}^\dagger + \hat{b})$, where $\hat{b}$ is the annihilation operator of a linear oscillator. For the free energy term, we expand the cosine up to fourth order leading to

$$\hat{\mathcal{H}}_0^{\text{duf}} = \omega_q \hat{b}^\dagger \hat{b} + \frac{\Delta}{2} \hat{b}^\dagger \hat{b}^\dagger \hat{b} \hat{b}, \quad (\text{A3})$$

with $\omega_q = \sqrt{8E_J E_C} - E_C$ and $\Delta = -E_C$. However, the eigenenergies and coupling strengths deviate from the Duffing model at higher levels due to the higher-order expansion of the cosine potential [13]. The accurate description of the transmon circuit requires us to diagonalize $\hat{\mathcal{H}}$ up to a truncation level N_{max} , which gives

$$H_0 = \sum_{k=0}^{N_{\text{max}}} \omega_k |k\rangle \langle k|, \quad (\text{A4})$$

whereas the charge operator takes the form

$$\hat{n} = \sum_{k=0}^{N_{\text{max}}-1} \left[\lambda_{k,k+1} |k\rangle \langle k+1| + \sum_{j=1} \lambda_{k,k+2j+1} |k\rangle \langle k+2j+1| \right] + \text{H.c.} \quad (\text{A5})$$

where λ denotes the corresponding matrix element. The maximal j is chosen such that $k+2j+1$ falls within the truncated levels. Unlike the Duffing oscillator model, the charge operator contains additional transitions corresponding to the underlying parity symmetry from the Mathieu functions. One can show that the coupling terms beyond the ladder structure are orders of magnitude smaller and are suppressed by the rotating-wave approximation.

From this point, it is more convenient to express the Hamiltonian in the rotating frame defined by the transformation $R = \exp(-i\omega_d t \sum_k k |k\rangle \langle k|)$, where ω_d is the selected driving transition frequency. This leads to the following total Hamiltonian:

$$\begin{aligned} \hat{H} = & \sum_{k=0}^{N_{\text{max}}} \tilde{\Delta}_k |k\rangle \langle k| + \Omega(t) \cos(\omega_d t) \\ & \times \left[\sum_{k=0}^{N_{\text{max}}-1} \sum_{j=1} \lambda_{k,k+2j+1} |k\rangle \langle k+2j+1| e^{-(2j+1)i\omega_d t} \right. \\ & \left. + \sum_{k=0}^{N_{\text{max}}-1} \lambda_{k,k+1} |k\rangle \langle k+1| e^{-i\omega_d t} + \text{H.c.} \right], \quad (\text{A6}) \end{aligned}$$

where $\Delta_k = \omega_k - k\omega_d$ is the detuning between the k -th level with the k -th driving frequency harmonic. In this frame, all coupling terms, except for $|k\rangle \leftrightarrow |k+1\rangle$, and all the counter-rotating components oscillate rapidly. Therefore, we can neglect those small contributions, leading to the Hamiltonian in Eq. (1).

APPENDIX B: FIDELITY EXPRESSIONS

In this appendix, we define what are the relevant metric to benchmark the performance of our single-qubit gate. In the case of dynamics without dissipative effects, the gate infidelity is defined as $\mathcal{E}(\hat{\mathcal{U}}_Q) = 1 - \mathcal{F}[\hat{\mathcal{U}}_Q]$ where $\mathcal{F}[\hat{\mathcal{U}}_Q]$ is the gate fidelity defined as

$$\mathcal{F}[\hat{\mathcal{U}}_Q] = \frac{\text{Tr}[\hat{\mathcal{U}}_Q \hat{\mathcal{U}}_Q^\dagger]}{d(d+1)} + \frac{|\text{Tr}[\hat{\mathcal{U}}_Q \hat{\mathcal{U}}_I^\dagger]|^2}{d(d+1)}, \quad (\text{B1})$$

where $\hat{\mathcal{U}}_Q$ is the time evolution operator at the final time projected in the computational subspace and $\hat{\mathcal{U}}_I = \hat{\sigma}_{0,1}^x$ is the target π -rotation.

In presence of dissipative effects, the metric must change because the master equation strongly depends

on the initial state of the system, requiring sampling to compute the gate error. Nevertheless, it is possible to avoid sampling by solving the master equation for different initial states given by the poles of the Bloch sphere spanned by the eigenstates of the Pauli matrices projected in the qubit subspace $\{\hat{\sigma}_{0,1}^x, \hat{\sigma}_{0,1}^y, \hat{\sigma}_{0,1}^z\}$

$$\mathcal{F}_d[\hat{\mathcal{U}}_Q] = \frac{1}{6} \sum_{j=\pm, \pm x, \pm y} \text{Tr}[\hat{\mathcal{U}}_Q \rho_j \hat{\mathcal{U}}_Q^\dagger \mathcal{M}(\rho_j)], \quad (\text{B2})$$

where ρ_j are the eigenstates of the Pauli matrices projected in the qubit subspace $\{\sigma_{0,1}^x, \sigma_{0,1}^y, \sigma_{0,1}^z\}$, moreover, $\mathcal{M}(\rho_j)$ is the state obtained by solving Eq. (41) considering ρ_j as its initial state.

APPENDIX C: THIRD-ORDER HAMILTONIAN AND CORRECTIONS

Expanding $\hat{\mathcal{H}}_4(t)$ [Eq. (13)] up to third order in Ω_x/Δ_2 and keeping only the qubit subspace and its transitions we obtain:

$$\begin{aligned} \hat{\mathcal{H}}_5(t) = & \left[\frac{\Omega_x}{2} + \mathcal{E}_{01}^x \Omega_x^3 \right] \hat{\sigma}_{0,1}^x + \Omega_x (\mathcal{E}_{12}^{y1} \Omega_1 \dot{\Omega}_1 + \mathcal{E}_{12}^{y2} \Omega_2 \dot{\Omega}_2) \hat{\sigma}_{1,2}^y \\ & + \Omega_x (\mathcal{E}_{12}^{x0} \Omega_x^2 + \mathcal{E}_{12}^{x1} \Omega_1^2 + \mathcal{E}_{12}^{x2} \Omega_2^2) \hat{\sigma}_{1,2}^x \\ & + \Omega_x (\mathcal{E}_{03}^{y1} \Omega_1 \dot{\Omega}_1 + \mathcal{E}_{03}^{y2} \Omega_2 \dot{\Omega}_2) \sigma_{0,3}^y \\ & + \Omega_x (\mathcal{E}_{03}^{x0} \Omega_x^2 + \mathcal{E}_{03}^{x1} \Omega_1^2 + \mathcal{E}_{03}^{x2} \Omega_2^2) \sigma_{0,3}^x + O(\Omega^4) \end{aligned} \quad (\text{C1})$$

The individual constants are given by

$$\mathcal{E}_{01}^x = \frac{(4 - \lambda_2^2)}{16\Delta_2^2} \quad (\text{C2})$$

$$\mathcal{E}_{12}^{y1} = -\frac{\lambda_2 \lambda_3^2 (\Delta_2^2 - \Delta_3^2) (2\Delta_2 - \Delta_3)^2}{8\Delta_2^5 \Delta_3^2}, \quad (\text{C3})$$

$$\mathcal{E}_{12}^{y2} = -\frac{\lambda_2 (\Delta_2^2 + (2\Delta_2 - \Delta_3)^2 \lambda_3^2)}{8\Delta_2^5}, \quad (\text{C4})$$

$$\mathcal{E}_{12}^{x0} = \frac{\lambda_2 (-3\Delta_3 \lambda_3^2 + \Delta_2 (31 - 14\lambda_2^2 + 10\lambda_3^2))}{48\Delta_2^3}, \quad (\text{C5})$$

$$\mathcal{E}_{12}^{x1} = \frac{\lambda_2 (\Delta_2^3 + \Delta_3 \lambda_3^2 (2\Delta_2 - \Delta_3)^2)}{16\Delta_2^5}, \quad (\text{C6})$$

$$\mathcal{E}_{12}^{x2} = \frac{\lambda_2 \lambda_3^2 (\Delta_2^2 - \Delta_3^2) (2\Delta_2 - \Delta_3)^2}{16\Delta_2^5 \Delta_3}, \quad (\text{C7})$$

$$\mathcal{E}_{03}^{y1} = \frac{\lambda_2 \lambda_3 (2\Delta_2 - \Delta_3)}{4\Delta_2^4}, \quad (\text{C8})$$

$$\mathcal{E}_{03}^{y2} = \frac{\lambda_2 \lambda_3 (2\Delta_2 - \Delta_3) (\Delta_2^2 - \Delta_3^2)}{8\Delta_2^4 \Delta_3}, \quad (\text{C9})$$

$$\mathcal{E}_{03}^{x0} = \frac{\lambda_2 \lambda_3 (3\Delta_2 - \Delta_3)}{48\Delta_2^3}, \quad (\text{C10})$$

$$\mathcal{E}_{03}^{x1} = \frac{\lambda_2 \lambda_3 (\Delta_3 - 2\Delta_2) (\Delta_2 + \Delta_3)}{16\Delta_2^4}, \quad (\text{C11})$$

$$\mathcal{E}_{03}^{x2} = \frac{\lambda_2 \lambda_3 (\Delta_3 - 2\Delta_2) (\Delta_2^2 - \Delta_3^2)}{16\Delta_2^4 \Delta_3}. \quad (\text{C12})$$

To correct higher order terms in $\sigma_{0,1}^x$ we use the linear shift in Eq. (18). To correct the higher order terms in $\sigma_{1,2}^x$ and $\sigma_{1,2}^y$ we follow the transformation generated by Eq. (19) and (20) introducing an auxiliary variable defined as:

$$\begin{aligned} A_{0,1}(t) = & \frac{2\Delta_2^2}{\lambda_2} \left(\Omega_x (\mathcal{E}_{12}^{x0} \Omega_x^2 + \mathcal{E}_{12}^{x1} \Omega_1^2 + \mathcal{E}_{12}^{x2} \Omega_2^2) \hat{\sigma}_{1,2}^x \right. \\ & - \dot{\Omega}_x (\mathcal{E}_{12}^{y1} \Omega_1 \dot{\Omega}_1 + \mathcal{E}_{12}^{y2} \Omega_2 \dot{\Omega}_2) \\ & - \Omega_x (\mathcal{E}_{12}^{y1} \Omega_1 \dot{\Omega}_1^2 + \mathcal{E}_{12}^{y2} \Omega_2 \dot{\Omega}_2^2) \\ & \left. - \Omega_x (\mathcal{E}_{12}^{y1} \Omega_1 \ddot{\Omega}_1 + \mathcal{E}_{12}^{y2} \Omega_2 \ddot{\Omega}_2) \right) \end{aligned} \quad (\text{C13})$$

APPENDIX D: MAGNUS EXPANSION OF TOGGLING FRAME

Given $\mathcal{H}_{\text{tog}}$, where we have with $\mathcal{E}_{11}^{(a)} = \frac{(a-1)\alpha\lambda_2^2}{8\Delta_2^3}$, $\mathcal{E}_{01}^{x0} = \frac{4\alpha^2 + \lambda_2^2(1-2\alpha)}{8\Delta_2^2}$, and $\mathcal{E}_{01}^{x1} = \frac{\alpha}{\Delta_2}$, in order to obtain β and δ_c we have to compute the computational subspace of $\hat{\mathcal{Z}}$, where the element $|i\rangle\langle j|$ of $\hat{\mathcal{Z}}$ is represented by $\hat{\mathcal{Z}}[i, j] = \int_0^T dt \langle i | \hat{\mathcal{H}}_{\text{tog}}(t) | j \rangle$:

$$\hat{\mathcal{Z}}_1[1, 1] = i \int_0^T ([\delta_c + \mathcal{E}_{11} \Omega_x^2] \cos \theta + \mathcal{E}_{11}^{(a)} \Omega_x \dot{\Omega}_x \sin \theta) dt \quad (\text{D1})$$

$$\begin{aligned} \hat{\mathcal{Z}}_1[0, 1] = & \frac{1}{2} \int_0^T dt (\Omega_x - \dot{\theta} + i[\delta_c + \mathcal{E}_{11} \Omega_x^2] \sin \theta \\ & - \mathcal{E}_{01}^{x0} \Omega_x^3 + \mathcal{E}_{01}^{x1} \Omega_x \delta_c + i\mathcal{E}_{11}^{(a)} \Omega_x \dot{\Omega}_x \cos \theta), \end{aligned} \quad (\text{D2})$$

Note that both Ω_x and θ are time-dependent functions in the above equations.

Given the rotation angle $\theta = \int_0^T \sin^2(\pi\tau/T) d\tau$, that smoothly increases from 0 to π , then $\cos \theta$ will be an odd function over the period of integration $[0, T]$ around $T/2$ (it corresponds to $\cos([0, \pi])$ which is odd). This simplifies the first term in Eq. (D1) to zero—we only have products of constants and odd functions or even functions

and odd functions. The same logic applies to the second component where now $\sin\theta$ is even but $\dot{\Omega}_x$ is odd. This way we can use the constant detuning δ_c and the amplitude of the pulse to make Eq. (D1) = 0. It is important to note that this is only true for the π rotation.

In order to compute the analytical DRAG prefactor, we need the elements of the Magnus expansion of the Toggling frame of transitions $|1\rangle \rightleftharpoons |2\rangle$ and $|0\rangle \rightleftharpoons |2\rangle$. To the second order, these are:

$$\begin{aligned} \hat{Z}_1[1, 2] = \int_0^T \frac{\lambda_2 e^{-i\phi_1}}{4\Delta_2} & \left(i \left(\Omega_x^2 \sin \frac{\theta}{2} - 2\dot{\Omega}_x \cos \frac{\theta}{2} \right) \right. \\ & - \alpha \frac{-4i\Delta_2 \dot{\Omega}_x \cos \frac{\theta}{2} + \Omega_x (2\dot{\Omega}_x + i\Delta_2 \Omega_x) \sin \frac{\theta}{2}}{2\Delta_2} \\ & \left. + \alpha^2 \frac{\Omega_x \dot{\Omega}_x \sin \frac{\theta}{2}}{\Delta_2} \right) dt \end{aligned} \quad (D3)$$

and

$$\begin{aligned} \hat{Z}_1[0, 2] = \int_0^T \frac{\lambda_2 e^{-i\phi_1}}{4\Delta_2} & \left(\left(\Omega_x^2 \cos \frac{\theta}{2} + 2\dot{\Omega}_x \sin \frac{\theta}{2} \right) \right. \\ & - \alpha \frac{4\Delta_2 \dot{\Omega}_x \sin \frac{\theta}{2} + \Omega_x (-2i\dot{\Omega}_x + \Delta_2 \Omega_x) \cos \frac{\theta}{2}}{2\Delta_2} \\ & \left. - i\alpha^2 \frac{\Omega_x \dot{\Omega}_x \cos \frac{\theta}{2}}{\Delta_2} \right) dt \end{aligned} \quad (D4)$$

with $\hat{Z}_1[2, 1] = \hat{Z}_1[1, 2]$ and $\hat{Z}_1[2, 0] = \hat{Z}_1[0, 2]$.

To obtain (35), we define:

$$\begin{aligned} x_1 &= (h_{12}^1 + 2h_{12}^2)(\bar{h}_{12}^0 + \bar{h}_{12}^1 + \bar{h}_{12}^2) \\ &+ (h_{02}^1 + 2h_{02}^2)(\bar{h}_{02}^0 + \bar{h}_{02}^1 + \bar{h}_{02}^2), \\ x_2 &= h_{12}^1 \bar{h}_{12}^1 + 2h_{12}^1 \bar{h}_{12}^2 + 2h_{12}^2(\bar{h}_{12}^0 + 2\bar{h}_{12}^1 + 3\bar{h}_{12}^2) \\ &+ h_{02}^1 \bar{h}_{02}^1 + 2h_{02}^1 \bar{h}_{02}^2 + 2h_{02}^2(\bar{h}_{02}^0 + 2\bar{h}_{02}^1 + 3\bar{h}_{02}^2), \end{aligned} \quad (D5)$$

where the symbol h_{ij}^k represents the integral from 0 to T of the k th α power in $\hat{\mathcal{H}}_{\text{tog}}[i, j]$, i.e. $\hat{Z}_1[1, 2] = \int_0^T \hat{\mathcal{H}}_{\text{tog}}[1, 2] dt = h_{12}^0 + h_{12}^1 \alpha + h_{12}^2 \alpha^2$ and $\bar{x}_1$ ($\bar{x}_2$) is the complex conjugate of x_1 (x_2).

APPENDIX E: R1D PREFACTORS OPTIMIZATION

We can also optimize R1D pulses by using Eq. (37) and starting with $\Omega_1 = \sin^3 \pi t/T$. As can be seen in Fig. 11 although we see an improvement over optimized linear DRAG, R1D does not match R2D. Plotting population evolution for a gate time of $T = 7$ (ns) in Fig. 12 explains why—we see that leakage to $|3\rangle$ clearly limits our gate performance. This lends even more strength to the need for R2D pulses.

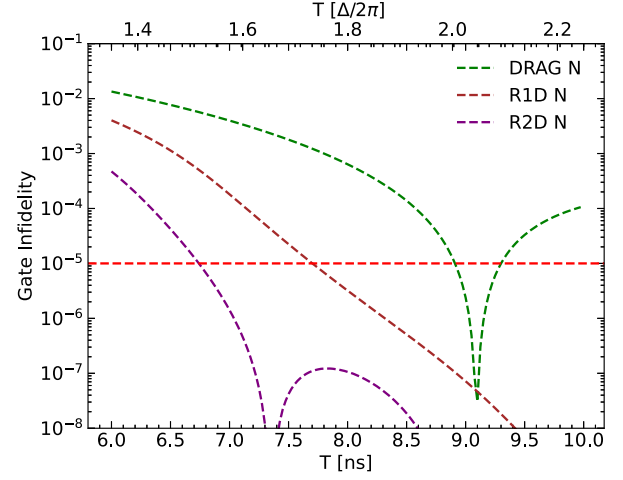


FIG. 11. Gate Infidelities for calibrated linear DRAG pulses (green), calibrated R1D pulses (brown) and calibrated R2D pulses (purple). Although R1D performs better than linear DRAG, it fails to achieve fidelities above 99.999% for gate times shorter than $T \approx 7.7$ ns.

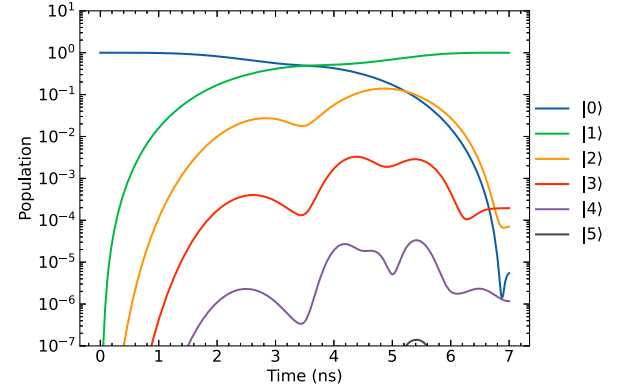


FIG. 12. Evolution of population for a π gate using a calibrated R1D pulse. As can be seen, leakage to $|3\rangle$ is substantial and hampers gate performance.

APPENDIX F: ANSATZ SELECTION

Here, we will show how to obtain the pulse from Eq. (16). Our starting point is the boundary conditions given in Eq. (5b) together with the additional one provided by both R1D and R2D, thus, we can choose a Fourier ansatz of the form

$$\Omega(t) = \frac{1}{2} + \frac{1}{2(k-1)} \left[\cos \left[j \frac{2\pi}{T} t \right] - k \cos \left[n \frac{2\pi}{T} t \right] \right], \quad (F1)$$

this functional form ensures that the first and third derivatives are zero at $t = 0$ and $t = T$. Then, we need to find the condition for the second derivative at both times, such that the pulse from Eq. (16) achieves lower

TABLE I. Optimal gate times T achieving an error below 10^{-4} for different choices of the ansatz parameters $\{n, j, k\}$. The highlighted entry corresponds to the fastest configuration.

n	j	k	$T_{99.99}$ (ns)
1	2	4	10.43
1	3	9	8.42
1	4	16	8.45
1	5	25	8.60
2	3	9/4	18.0
2	4	8	18.0
2	5	25/4	18.0

errors for a gate faster than 10 ns.

$$\ddot{\Omega} = -\left[\frac{2\pi}{T}\right]^2 \left[j^2 \cos\left[j\frac{2\pi}{T}t\right] - kn^2 \cos\left[n\frac{2\pi}{T}t\right] \right]. \quad (\text{F2})$$

To obtain null derivative at $t = 0$ and $t = T$, we obtain the relation

$$k = \frac{j^2}{n^2}, \quad \forall k \neq 1. \quad (\text{F3})$$

Table I shows the optimal gate time T to obtain gate error below 10^{-4} for different value of the triad $\{n, j, k\}$.

-
- [1] P. Krantz, M. Kjaergaard, F. Yan, T.P. Orlando, S. Gustavsson, and W. D. Oliver, *A quantum engineer's guide to superconducting qubits*, *Appl. Phys. Rev.* **6**, 021318 (2019).
- [2] A. Blais, A. L. Grimsmo, S. M. Girvin, and A. Wallraff, *Circuit quantum electrodynamics*, *Rev. Mod. Phys.* **93**, 025005 (2021). [10.1103/RevModPhys.93.025005](#)
- [3] 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).
- [4] Z. Chen *et al.*, *Measuring and suppressing quantum state leakage in a superconducting qubit*, *Phys. Rev. Lett.* **116**, 020501 (2016).
- [5] 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).
- [6] F. Motzoi and F. K. Wilhelm, *Improving frequency selection of driven pulses using derivative-based transition suppression*, *Phys. Rev. A* **88**, 062318 (2013).
- [7] L. S. Theis, F. Motzoi, S. Machnes, and F. K. Wilhelm, *Counteracting systems of diatomicities using DRAG controls: The status after 10 years*, *Europhys. Lett.* **123**, 60001 (2018).
- [8] 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).
- [9] B. Li, T. Calarco, and F. Motzoi, *Experimental error suppression in Cross-Resonance gates via multi-derivative pulse shaping*, *npj Quantum Inf.* **10**, 1 (2024).
- [10] B. Li, F. A. Cárdenas-López, A. Lupascu, and F. Motzoi, *Universal pulses for superconducting qudit ladder gates*, *PRX Quantum* **6**, 030357 (2025).
- [11] E. Hyppä, A. Vepsäläinen, M. Papič, C. F. Chan, S. Inel, A. Landra, W. Liu, J. Luus, F. Marxer, C. Ockeloen-Korppi, S. Orbell, B. Tarasinski, and J. Heinsoo, *Reducing leakage of single-qubit gates for superconducting quantum processors using analytical control pulse envelopes*, *PRX Quantum* **5**, 030353 (2024).
- [12] B. Chiaro and Y. Zhang, *Active leakage cancellation in single qubit gates*, *Phys. Rev. Lett.* **135**, 130601 (2025).
- [13] B. Khani, J. M. Gambetta, F. Motzoi, and F. K. Wilhelm, *Optimal generation of Fock states in a weakly nonlinear oscillator*, *Phys. Scr.* **2009**, 014021 (2009).
- [14] B. Li, T. Calarco, and F. Motzoi, *Nonperturbative analytical diagonalization of Hamiltonians with application to circuit QED*, *PRX Quantum* **3**, 030313 (2022).
- [15] J. R. Schrieffer and P. A. Wolff, *Relation between the Anderson and Kondo Hamiltonians*, *Phys. Rev.* **149**, 491 (1966).
- [16] 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).
- [17] L. H. Pedersen, N. M. Møller, and K. Mølmer, *Fidelity of quantum operations*, *Phys. Lett. A* **367**, 47 (2007).
- [18] R. Wang, Y. Feng, Y. Zhang, J. Ding, B. Li, F. Motzoi, Y. Gao, H. Xu, Z. Yang, W. Nuerbolati, H. Yu, W. Sun, and F. Yan, *Suppressing spurious transitions using spectrally balanced pulse*, *Phys. Rev. Lett.* **135**, 160804 (2025).
- [19] F. Marxer *et al.*, *Above 99.9 single-qubit gates, two-qubit gates, and readout in a single superconducting quantum device*, [arXiv:2508.16437](#).
- [20] J. J. Burnett, A. Bengtsson, M. Scigliuzzo, D. Niepce, M. Kudra, P. Delsing, and J. Bylander, *Decoherence benchmarking of superconducting qubits*, *npj Quantum Inf.* **5**, 54 (2019).
- [21] S. E. de Graaf, L. Faoro, L. B. Ioffe, S. Mahashabde, J. J. Burnett, T. Lindström, S. E. Kubatkin, A. V. Danilov, and A. Y. Tzalenchuk, *Two-level systems in superconducting quantum devices due to trapped quasiparticles*, *Sci. Adv.* **6**, eabc5055 (2020).
- [22] N. Wu, J. Lin, C. Xie, Z. Guo, W. Huang, L. Zhang, Y. Zhou, X. Sun, J. Zhang, W. Guo, X. Linpeng, S. Liu, Y. Liu, W. Ren, Z. Tao, J. Jiang, J. Chu, J. Niu, Y. Zhong, and D. Yu, *In situ mixer calibration for superconducting quantum circuits*, *Appl. Phys. Lett.* **125**, 204003 (2024).
- [23] S. Gustavsson, O. Zwiern, J. Bylander, F. Yan, F. Yoshihara, Y. Nakamura, T. P. Orlando, and W. D. Oliver, *Improving quantum gate fidelities by using a qubit to measure microwave pulse distortions*, *Phys. Rev. Lett.* **110**, 040502 (2013).
- [24] L.-L. Guo, P. Duan, S. Zhang, X.-X. Yang, C. Zhang, L. Du, H.-F. Zhang, H.-R. Tao, T.-L. Wang, Z.-L. Jia, Z.-Y. Chen, and G.-P. Guo, *Universal scalable characterization and correction of pulse distortions in controlled quantum systems*, *Phys. Rev. Appl.* **21**, 064060 (2024).

- [25] W. S. Warren, *Effects of arbitrary laser or NMR pulse shapes on population inversion and coherence*, *J. Chem. Phys.* **81**, 5437 (1984).
- [26] U. Haeberlen, *High Resolution NMR in Solids: Selective Averaging*, *Advances in Magnetic Resonance*, Supplement 1 (Academic Press, New York, 1976).
- [27] A. Brinkmann, *Introduction to average Hamiltonian theory. I. Basics*, *Concepts Magn. Reson., Part A* **45A**, e21414 (2016).
- [28] S. Blanes, F. Casas, J. Oteo, and J. Ros, *The Magnus expansion and some of its applications*, *Phys. Rep.* **470**, 151 (2009).
- [29] S. Kirchhoff, F. K. Wilhelm, and F. Motzoi, *Correction formulas for the Mølmer-Sørensen gate under strong driving*, *PRX Quantum* **6**, 010328 (2025).
- [30] Y. Gao, A. Galicia, J. D. D. C. Jesus, Y. Liu, Y. Haddad, D. A. Volkov, J. R. Guimarães, H. Bhardwaj, M. Jerger, M. Neis, B. Li, F. A. Cárdenas-López, F. Motzoi, P. A. Bushev, and R. Barends, *Ultrafast single qubit gates through multi-photon transition removal*, [arXiv:2511.22365](https://arxiv.org/abs/2511.22365).
- [31] D. Manzano, *A short introduction to the Lindblad master equation*, *AIP Adv.* **10**, 025106 (2020).
- [32] IBM Quantum, *IBM quantum platform* (2025), https://quantum.cloud.ibm.com/computers?system=ibm_pittsburgh, accessed: 2025-12-10.
- [33] C. Wang, X. Li, H. Xu, Z. Li, J. Wang, Z. Yang, Z. Mi, X. Liang, T. Su, C. Yang *et al.*, *Towards practical quantum computers: Transmon qubit with a lifetime approaching 0.5 milliseconds*, *npj Quantum Inf.* **8**, 3 (2022).
- [34] M. D. Bowdrey, D. K. L. Oi, A. J. Short, K. Banaszek, and J. A. Jones, *Fidelity of single qubit maps*, *Phys. Lett. A* **294**, 258 (2002).
- [35] J. D. D. C. Jesus, Y. Gao, B. Li, R. Barends, F. A. Cárdenas-López, and F. Motzoi, *Data produced for paper “Analytical blueprint for 99.999% fidelity X-gates on present superconducting hardware under strong driving” version 2*, Zenodo (2026), [10.5281/zenodo.18744365](https://zenodo.org/record/18744365).