Universal Pulses for Superconducting Qudit Ladder Gates

Boxi Li,[1,2,*] F.A. Cárdenas-López,[1] Adrian Lupascu,[3] and Felix Motzoi[1,2,†]

[1] *Forschungszentrum Jülich, Institute of Quantum Control (PGI-8), D-52425 Jülich, Germany*

[2] *Institute for Theoretical Physics, University of Cologne, D-50937 Cologne, Germany*

[3] *Department of Physics and Astronomy, and Waterloo Institute for Nanotechnology, Institute for Quantum Computing, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada*

(Received 17 January 2025; revised 7 July 2025; accepted 19 August 2025; published 19 September 2025)

Qudits, generalizations of qubits to multilevel quantum systems, offer enhanced computational efficiency by encoding more information per lattice cell, avoiding costly swap operations and providing even exponential speedup in some cases. Utilizing the d-level manifold, however, requires high-speed gate operations because of the stronger decoherence at higher levels. While analytical control methods have proven effective for qubits in achieving fast gates with minimal control errors, their extension to qudits is nontrivial due to the increased complexity of the energy-level structure arising from additional ancillary states. In this work, we present a universal pulse construction for generating rapid, high-fidelity unitary rotations between adjacent qudit levels, thereby providing a prescription for any gate in SU(d). Control errors in these operations are effectively analyzed within a four-level subspace, including two leakage levels with approximately opposite detuning. By identifying the optimal degrees of freedom, we derive concise analytical pulse schemes that suppress multiple control errors and outperform existing methods. Remarkably, our approach achieves consistent coherent error scaling across all levels, approaching the quantum speed limit independently of parameter variations between levels. Numerical validation on transmon circuits demonstrates significant improvements in gate fidelity for various qudit sizes aiming for 10^{-4} error. This method provides a scalable solution for improving qudit control and can be broadly applied to other quantum systems with ladder structures or operations involving multiple ancillary levels.

DOI: 10.1103/9dxw-4c7y

I. INTRODUCTION

Quantum computation and quantum information-processing protocols often rely on qubits, or two-level systems, as the fundamental units of computation due to their simplicity and close analogy to classical computing. However, most quantum systems comprise more than just two levels. These additional quantum levels can also be used as an information register, which is known as a qudit, a generalization of the qubit to a d-level system. Exploring the full Hilbert space of qudits enables more efficient computation by increasing the amount of information stored per quantum unit.

Qudit-based quantum computation offers several known advantages over qubit-based approaches. For instance, a d-level system can encode $\log_2(d)$ qubits [1]. This has been exploited for efficient compilation of arbitrary unitaries, requiring an exponentially reduced number of circuit layers [2–6], to simulate bosonic modes for studying light-matter processes [7,8] and lattice gauge theories [9–11], for enhancing the robustness in quantum cryptography [12, 13], and for simplified implementation of quantum error-correction protocols [14–16]. In general, the larger density of registers means that the connectivity of qubit-based architectures is increased, since neighboring qudits can share up to d^2-level couplings. Qudit processors have been implemented across various physical platforms, including trapped ions [17–20], Rydberg atoms [21], ultracold atomic mixtures [22], molecular spins [23,24], photonic systems [25–29], and superconducting circuits [30–37]; such implementations contribute to significant progress on qudit-based quantum computation.

Despite these advancements, maintaining coherent control of all the qudit levels poses complex new challenges. In transmon superconducting circuits, where the quantum system is represented by a nonlinear oscillator [38], each qudit operation needs to be addressed differently due to the varying surrounding level structure. Compared to a

*Contact author: b.li@fz-juelich.de

†Contact author: f.motzoi@fz-juelich.de

qubit operation, the presence of additional leakage channels significantly limits the gate performance, as shown in Figs. 1(a)–1(c). For instance, it has been reported that the gate time of a single-qutrit gate is around 30 ns [31,33,35], which is 3 times longer than that required for single-qubit gates with state-of-the-art quantum control techniques [39]. Therefore, developing quantum control protocols for qudits is crucial for making qudit computation practical. Of particular relevance is the derivative removal by adiabatic gate (DRAG) method [40–43], successfully employed in superconducting qubit systems to reduce leakage and phase errors. DRAG's simplicity and flexibility allow engineering efficient pulses with easy-to-calibrate parameters, making it ubiquitous in the superconducting qubits platform [39,44–47]. The same advantages remain even with the presence of multiple error sources, whereby multiple DRAG corrections can be combined, offering efficient yet compact solutions [42,48].

In this paper, we investigate the energy structure and dominant error mechanisms of a transmon ladder-type qudit, a weakly anharmonic oscillator, under nearest-level driving. We examine the relationship between the number of usable levels in a transmon qudit and its design parameters, which constrain coherence times and gate design. We demonstrate that transmon qudit control can be studied within a four-level subspace involving two nearest-neighbor interactions and effectively reduced to an enhanced error model in an SU(2) subspace. Through a systematic study across various qudit sizes in transmon circuits, we find that porting the widely used single-derivative DRAG method, as previously suggested in Ref. [41] and experimentally implemented in Refs. [32,49], does not significantly improve fidelity in the qudit case because it is overconstrained by multiple leakage channels.

To overcome this limitation, we apply the DRAG framework to engineer level-independent and high-precision analytical control schemes for qudit systems within a driven ladder structure. We adopt a recursive DRAG approach introduced in Ref. [48] that incorporates higher-order derivatives, providing new degrees of freedom that are used to suppress both single- and multiphoton errors.

Remarkably, despite the presence of multiple parameters in the circuit description, we observe a universal behaviour in the error budget and pulse-specific quantum speed limits, i.e., the minimum gate duration required to achieve a target fidelity. In particular, for a given drive protocol, the speed limit is found to be inversely proportional to the system nonlinearity, and it remains independent of other circuit parameters or the specific qudit transition targeted. However, it is highly sensitive to the drive protocol, particularly to whether certain leakage channels are effectively suppressed. We observe significant improvements in gate performance and successfully reduce gate times to

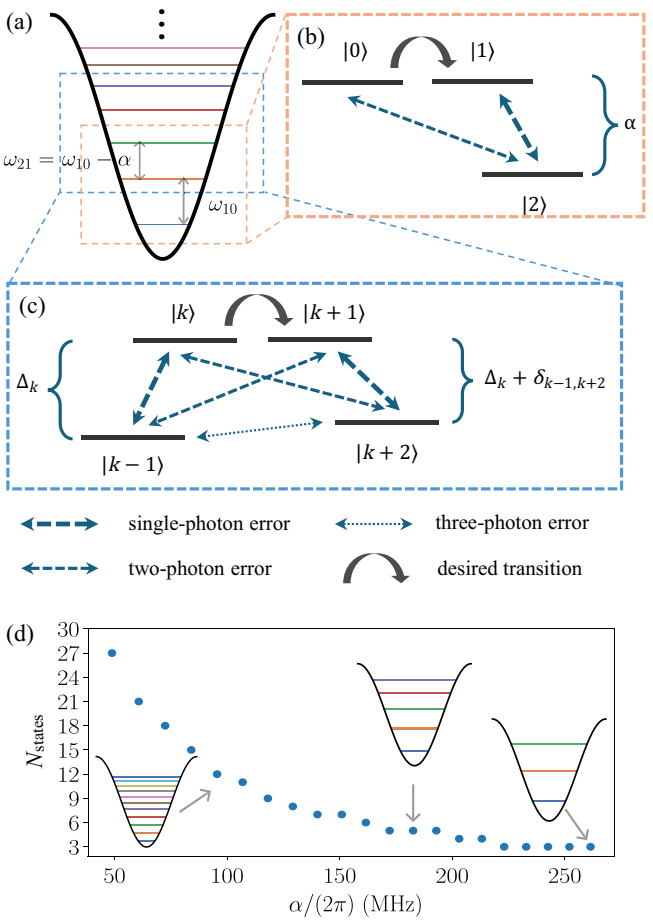

FIG. 1. Energy structure of driving a two-level transition in a qudit system. (a) Typical energy structure of a transmon system. Energy levels and error transitions for (b) the ground and first excited states, and (c) general ladder transition between higher levels in the rotating frame. (d) The number of levels that can be used as a qudit quantum register as a function of the anharmonicity. The upper bound is set by the decoherence introduced by charge fluctuations. The detailed discussion can be found in Appendix B.

mitigate dephasing caused by voltage fluctuations during gate implementation. These results are broadly applicable to any qudit platform with multiple connected ancillary levels.

In the following, we start with the transmon model and derive the four-level effective Hamiltonian in Sec. II. Next, we introduce and explore the recursive DRAG method in detail and perform a systematic study of its performance in Sec. III. Significant improvement in fidelity is observed across a wide range of parameters, with a universal behavior across all levels independent of parameter variations between levels. In Sec. IV, we discuss other potential control errors beyond the two-level transition and provide a summary of our findings in Sec. V.

II. QUDIT MODEL FOR UNIVERSAL QUANTUM GATES

We consider a superconducting transmon qudit, modeled as a weakly anharmonic oscillator with dominant nearest-level interactions, as illustrated in Fig. 1(a). In the rotating frame, the nonlinearity of the system allows for the isolation of effective subspaces that capture the dynamics of specific transitions [Figs. 1(b) and 1(c)]. However, residual couplings to noncomputational states introduce additional error channels. In this section, we derive an effective Hamiltonian for such systems, focusing on selective control of adjacent-level transitions and the constraints imposed by the energy structure and decoherence.

A. Native gate set for superconducting qudit

Our objective is that, for a transmon system, each individual ladder transition between adjacent states, $|k\rangle$ and $|k+1\rangle$, can be selectively controlled. This allows arbitrary unitaries on SU(d) to be implemented. Since a calibrated $\pi/2$ gate combined with a virtual Z gate [50] is a complete native gate set, as we show in Appendix A, our primary focus in the following study is on the $\pi/2$ gate. In addition, we also present results for the π gate, which represents the greatest challenge among Givens rotations at fixed gate duration, as it requires the strongest drive amplitude and typically incurs the highest leakage rate. Overall, these methods can be extended to rotations of arbitrary angles, which are very helpful for reducing circuit compilation depths [51].

B. The transmon Hamiltonian

We start by deriving the effective Hamiltonian for selectively driving the $|k\rangle \leftrightarrow |k+1\rangle$ transition in a superconducting transmon. A transmon nonlinear oscillator is described by the following Hamiltonian [38]:

$$\hat{H} = 4E_C[\hat{n} - n_g(t)]^2 - E_J \cos(\hat{\varphi}), \tag{1}$$

where E_C and E_J represent the charge and Josephson energies, respectively, and $n_g(t)$ is the dimensionless gate voltage. The operator $\hat{n}$ is the charge operator, indicating the number of Cooper pairs on the island, and $\hat{\varphi}$ denotes the phase operator. For implementing single-qudit operations, we capacitively drive the transmon using $n_g(t) = n_0(t)\cos(\omega_d t)$ resulting in the Hamiltonian

$$\hat{H}_{\text{ctrl}} = \Omega(t)\cos(\omega_d t)\hat{n}, \tag{2}$$

where $\Omega(t) = -8E_C n_0(t)$ is the complex drive envelope, and ω_d is the drive frequency.

When only the lowest few energy levels are considered, the transmon can be modeled as an approximate Duffing oscillator [52,53]. In this model, the control operator is expressed as $\hat{n} \propto (\hat{b}^\dagger + \hat{b})$, where $\hat{b}$ is the annihilation operator of a linear oscillator. Consequently, the control Hamiltonian adopts a ladder configuration, connecting states $|k\rangle \leftrightarrow |k \pm 1\rangle$. The transmon Hamiltonian without external drive then simplifies to

$$\hat{H}_0^{\text{duf}} = \omega_q \hat{b}^\dagger \hat{b} + \frac{\alpha}{2}\hat{b}^\dagger \hat{b}^\dagger \hat{b}\hat{b}, \tag{3}$$

with $\omega_q = \sqrt{8E_J E_C} - E_C$ and $\alpha = -E_C$. As long as the states are in the potential well, the dominant coupling is still this ladder coupling between the adjacent levels, as will be shown later. However, the eigenenergies and coupling strengths deviate from the Duffing model at higher levels due to the higher-order expansion of the cosine term in Eq. (1) [52].

An accurate effective model requires exact diagonalization up to a truncation level $N_{\max}$, which gives

$$H_0 = \sum_{k=0}^{N_{\max}} \omega_k |k\rangle \langle k|, \tag{4}$$

and for the charge operator

$$\hat{n} = \sum_{k=0}^{N_{\max}-1} \Big[n_{k,k+1} |k\rangle\langle k+1| + \sum_{j=1} n_{k,k+2j+1} |k\rangle\langle k+2j+1| \Big] + \text{h.c.}, \tag{5}$$

where n 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 $\hat{n}$ operator in the effective frame exhibits additional transitions, these nonzero matrix elements are related to the underlying parity symmetry from the Mathieu functions, the formal solution of the Hamiltonian in Eq. (1). Here, we distinguish between the ladder coupling between $|k\rangle \leftrightarrow |k+1\rangle$ and high-order couplings. The latter are typically orders of magnitude smaller and are suppressed by the rotating-wave approximation, as we will demonstrate later.

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

following total Hamiltonian:

$$\hat{H} = \sum_{k=0}^{N_{\max}} \tilde{\Delta}_k \hat{\Pi}_k + \Omega(t)\cos(\omega_d t) \left[\sum_{k=0}^{N_{\max}-1} \sum_{j=1} n_{k,k+2j+1} |k\rangle\langle k+2j+1| e^{-(2j+1)i\omega_d t} + \sum_{k=0}^{N_{\max}-1} n_{k,k+1} |k\rangle\langle k+1| e^{-i\omega_d t} + \text{h.c.} \right], \quad (6)$$

where $\hat{\Pi}_k = |k\rangle\langle k|$ and $\tilde{\Delta}_k = \omega_k - k\omega_d$ is the detuning between the kth level with the kth driving frequency harmonic. In the rotating frame, all coupling terms, except for $|k\rangle \leftrightarrow |k+1\rangle$, and all the counter-rotating components oscillate rapidly. Additionally, couplings between non-nearest-neighbor levels in a transmon ladder qudit are typically several orders of magnitude weaker, making the associated Stark shifts negligible. Therefore, we can neglect those small contributions, leading to

$$\hat{H}_{\text{RWA}} = \sum_{k=0}^{N_{\max}} \tilde{\Delta}_k \hat{\Pi}_k + \frac{\Omega(t)}{2} \sum_{k=0}^{N_{\max}-1} (n_{k,k+1} |k\rangle\langle k+1| + \text{h.c.}). \quad (7)$$

In this Hamiltonian, we recover the desired ladder coupling, albeit with renormalized eigenenergies and coupling strengths. To target a specific ladder transition $(k, k+1)$, the drive frequency is chosen such that $\tilde{\Delta}_k = \tilde{\Delta}_{k+1}$ for the desired k. The nonlinearity, captured by the remaining $\tilde{\Delta}_k - \tilde{\Delta}_j$ for $j \notin \{k, k+1\}$, permits selective driving of any transition between neighboring levels.

In addition to the Hamiltonian, we also consider the possible decoherence of the higher levels. To use the quantum states as a qudit, we demand that they are robust against charge fluctuations, which increase exponentially up the ladder. This condition sets an upper bound on the maximal number of usable states N_{states}, which *a priori* depends on the ratio E_J/E_C, and is graphed in Fig. 1(d) as a function of anharmonicity $\alpha \approx -E_C$. A higher E_J corresponds to a deeper potential, allowing for more confined states, while a lower E_C reduces charge fluctuations. However, in such a regime we also have decreases in the frequency difference between transitions, making selective driving more challenging. Instead of using the E_J/E_C ratio, it is more natural to select the usable states in terms of their coherence times T_1 and T_ϕ. In our case, as we consider fixed-frequency transmons, the main source of error will correspond to capacitive losses and dephasing due to charge fluctuations. As an optimistic forward-looking estimation, we set the upper bound of T_ϕ for the highest level N_{states} to be around 100 μs such that we are able to potentially achieve gate error below 10^{-4} for a 10-ns gate. For the parameters we choose, this corresponds to a charge dispersion of about 10^{-3} GHz. The details on the calculation of the charge fluctuation and coherence times are presented in Appendix B. In principle, for certain quantum operations, decoherence could be partially mitigated by applying dynamical decoupling methods [54]. In this work, however, we consider more generally using quantum control shaping to speed up the operation time and reduce the irreversible effect of decoherence.

C. Four-level effective model

Although the full qudit contains many levels, transitions that are detuned by more than $2\Delta_k$ from the $|k\rangle \leftrightarrow |k+1\rangle$ transition are far off resonant in the rotating frame relative to the coupling strength and therefore have a negligible influence on the system's dynamics. Therefore, we focus on nearest-neighbor transitions and further simplify the model to a four-level system. This choice is validated by the numerical simulations that follow. We define $\omega_d = \omega_{k+1} - \omega_k - \delta_d$, with δ_d denoting a designed small detuning between the drive frequency and the energy separation. The special case for qubits, $k = 0$, has been studied over the last decade [55]. The primary control error arises from the coupling to the nearest-neighboring levels, $|k-1\rangle$ and $|k+2\rangle$, as illustrated in Fig. 1(c). To simplify the analysis, we truncate the Hamiltonian to a four-level subsystem, described by

$$\hat{H}_k^{(4)} = \begin{pmatrix} \Delta_k & \frac{\lambda_{k-1}\bar{\Omega}}{2} & 0 & 0 \\ \frac{\lambda_{k-1}\Omega}{2} & \delta_d & \frac{\lambda_k\bar{\Omega}}{2} & 0 \\ 0 & \frac{\lambda_k\Omega}{2} & 2\delta_d & \frac{\lambda_{k+1}\bar{\Omega}}{2} \\ 0 & 0 & \frac{\lambda_{k+1}\Omega}{2} & 3\delta_d + \Delta_k + \delta_{k-1,k+2} \end{pmatrix}, \quad (8)$$

where $\bar{\Omega}$ denotes the complex conjugate of the pulse envelope. For clarity, a constant identity operator has been subtracted.

The two middle levels represent the targeted transition, separated by the small drive detuning δ_d. The first and last levels correspond to the potential leakage levels $|k-1\rangle$ and $|k+2\rangle$. An important observation is that, due to the weak nonlinearity, the energy separation between the resonant subspace and the leakage levels, $\Delta_k = \omega_{k-1} - 2\omega_k + \omega_{k+1}$, is always approximately equal to the anharmonicity α and only increases slightly as the levels rise. This is illustrated in Fig. 2(a), with the base case $\Delta_0 = \alpha$. The difference between the two leakage levels is given by $\delta_{k-1,k+2} = -\omega_{k-1} + 3\omega_k - 3\omega_{k+1} + \omega_{k+2}$ [see Fig. 1(c)]. For a harmonic or Duffing oscillator, it is straightforward to verify that $\delta_{k-1,k+2}$ is zero, (see Appendix C). However, for a transmon oscillator, $\delta_{k-1,k+2}$ takes a small but nonzero value, as illustrated in Fig. 2(b), which is plotted as a function of E_J/E_c and the anharmonicity. The curve

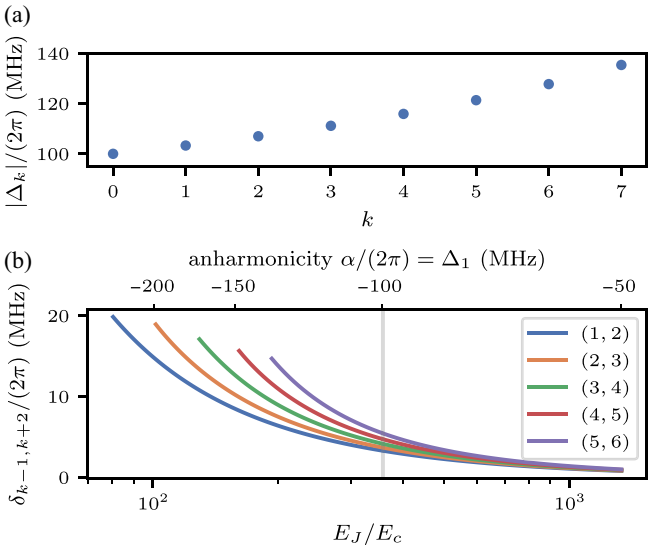

FIG. 2. Properties of transmon qudits. (a) Detuning between the target subspace and the leakage levels in the rotating frame. The qubit frequency and anharmonicity of the ground state are 5 GHz and −100 MHz, corresponding to $E_J/E_C \approx 355$. For $k = 0$, $\Delta_k = \alpha$. (b) The small energy gap between the two leakage levels $|k-1\rangle$ and $|k+2\rangle$ for the first five transitions. This is much smaller than Δ_k, leading to the energy structure shown in Fig. 1. The gray vertical line marks the parameters used in (a).

is truncated when the eigenstate's dispersion noise reaches 10^{-3} GHz, where the qudit coherence time drops below a minimum threshold (see Appendix B). Within this range, $\delta_{k-1,k+2}$ is much smaller than Δ_k.

The off-diagonal coupling term in Eq. (8) shows a similar structure as the Duffing model. The term λ_k denotes the renormalized drive strength between level k and $k+1$, given by $\lambda_k = n_{k,k+1}/|n_{0,1}|$, which equals $\sqrt{k}$ in the Duffing approximation. This model defined in Eq. (8) holds as long as the state remains within the potential well and the eigenenergy's dispersion to charge noise is sufficiently small.

III. RECURSIVE DRAG PULSE FOR QUDIT GATES

To suppress multiple leakage pathways in qudit systems, we employ the recursive DRAG method. The challenge of addressing multiple error channels was first mentioned in Ref. [41], which optimized only single-derivative DRAG parameters to balance competing leakage pathways. The subsequent work, Ref. [42], introduced additional control degrees of freedom, but its main focus was on frequency selectivity in uncoupled qubit systems with linear coupling. A similar approach with linear multiderivative correction is also derived in Ref. [56], motivated by frequency engineering. In contrast, a more systematic extension to systems with connected transitions, such as multilevel architectures, was proposed in Ref. [48], leading to the recursive DRAG method. This approach allows for simultaneous suppression of single- and multiphoton transitions through a hierarchy of higher-order derivative corrections in a compact recursive form, but generally ignores phase and amplitude errors stemming from noncommutativity of the transitions. The key advantage of the recursive framework is its ability to produce compact analytical pulse shapes via a structured expansion, significantly simplifying both theoretical design and experimental calibration. Each error channel can be independently addressed, enabling modular tuning of pulse parameters.

However, the original formulation in Ref. [48] was tailored to suppress fully off-resonant population transfer in a three-level system. This effectively creates quasiadiabatic dynamics, but does not take into account typical diabatic effects present in most gates, such as the qudit system presented here. In fact, there are two quasidiabatic transitions that are present in the qudit error model, namely the resonant $|k\rangle \leftrightarrow |k+1\rangle$ and the near degenerate $|k-1\rangle \leftrightarrow |k+2\rangle$, especially important for high power [see Fig. 1(c)]. The interplay between the off-resonant and on-resonant dynamics thereby entails several new error channels coming from large noncommutative effects, remaining leading order in the expansion, and necessitating accounting for the phase error, which is ignored in Ref. [48] because it commutes with the adiabatic dynamics. Moreover, the presence of these distinct leakage levels in our model offers an opportunity to systematically study calibration strategies and assess how each DRAG coefficient influences specific leakage pathways. In the following, we first analyze the error budget associated with such qudit gates and present an enhanced recursive DRAG protocol designed for this control challenge. Our method concurrently suppresses up to four off-resonant transitions, while maintaining high-fidelity coupling between the target states $|k\rangle$ and $|k+1\rangle$.

A. Error budget

To identify the dominant sources of error in qudit gate operations, we begin by analyzing the error budget associated with driving transitions in a transmon qudit. For the lowest two levels in a transmon, $|0\rangle$ and $|1\rangle$, the system reduces to the well-studied transmon qubit [Fig. 1(b)], well-established techniques like DRAG have been developed to suppress leakage, particularly to $|2\rangle$. In higher-level transitions, however, leakage and control errors become more complex due to residual couplings to additional noncomputational states [Fig. 1(c)]. For a general $|k\rangle \leftrightarrow |k+1\rangle$ gate, adjacent states $|k-1\rangle$ and $|k+2\rangle$ introduce errors such as leakage and Stark shifts, which grow more significant at shorter gate durations.

To quantify these contributions, we consider the Hann envelope,

$$\Omega_{\text{Hann}}(t) = \sin\left[\frac{\pi t}{t_f}\right]^2, \tag{9}$$

a commonly used smooth pulse of duration t_f with a narrow spectral bandwidth that usually performs reasonably well. The corresponding error contributions are shown in Fig. 3(a), with definitions provided in Appendix D. As evident from the figure, both leakage to $|k-1\rangle$ and $|k+2\rangle$ must be addressed. Additionally, residual couplings introduce Stark shifts that alter the resonance condition and effective coupling strength. To ensure high-fidelity rotations within the $|k\rangle, |k+1\rangle$ subspace, it is essential to suppress not only leakage but also phase and amplitude errors. This is different from previous studies, where phase error commutes with the desired dynamics and is automatically removed under the standard calibration routine [48]. As shown in Fig. 3(a), the latter dominates at long gate durations, while leakage errors increase sharply as the gate is sped up. Moreover, using more qudit levels typically requires reducing the qudit nonlinearity Δ_k to suppress charge noise [Fig. 1(d)], further complicating fast gate design.

In the standard qubit scenario, such errors are suppressed using the DRAG pulse [40]:

$$\Omega = \Omega_{\text{Hann}} - ia\frac{\dot{\Omega}_{\text{Hann}}}{\Delta}, \tag{10}$$

where Δ is the detuning to the leakage level, and a is a tunable coefficient to compensate for imperfect Hamiltonian modeling and higher-level corrections. An additional detuning δ_d is often included to correct phase accumulation between driven levels [39].

For transitions involving higher states in qudit [Fig. 1(c)], however, this single-derivative correction becomes insufficient. The standard DRAG form in Eq. (10) cannot simultaneously suppress both, as the associated detunings are approximately equal in magnitude but opposite in sign: $E_{|k\rangle} - E_{|k-1\rangle} \approx -(E_{|k+2\rangle} - E_{|k+1\rangle})$. As a result, the spectral notch introduced by the single derivative cannot align with both frequencies, leading to negligible improvement in gate performance, as seen in Fig. 3(b). This limitation is intrinsic to transitions with $k \geq 1$, due to the level spacing in transmon ladders.

B. General DRAG correction for a *n*-photon transition

To address this limitation, an effective strategy involves incorporating higher-order derivative terms [42,48]. This approach can be interpreted as a superadiabatic transformation [57], wherein a second adiabatic frame is derived, enabling the introduction of new time-dependent control

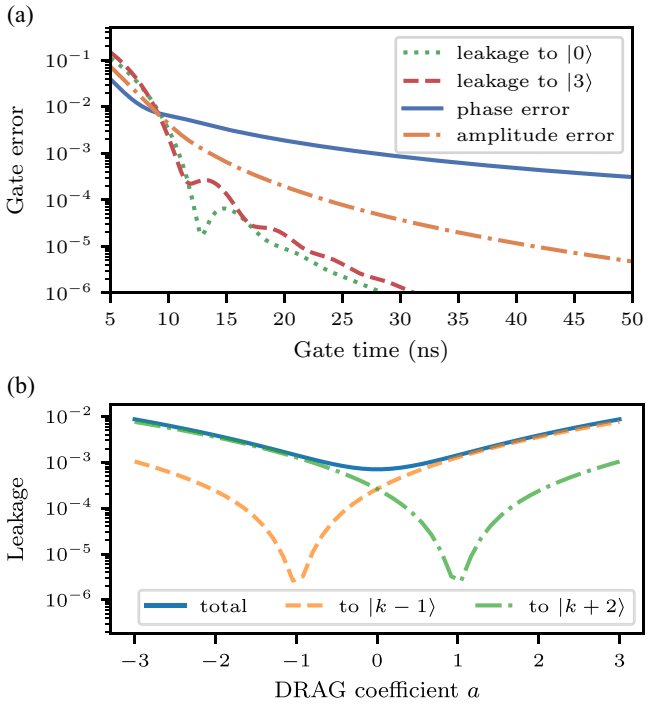

FIG. 3. Control error in driving ladder transitions in a transmon qudit. (a) Estimated error budget of driving a $|1\rangle \leftrightarrow |2\rangle$ π rotation using a Hann pulse with an anharmonicity of $\alpha/(2\pi) = -200$ MHz. The definitions of different errors are given in Appendix D. (b) The leakage error calculated via Eq. (D2) in a regime where the single-derivative DRAG faces limitations and offers no improvement, even with an optimized DRAG coefficient. Parameters used are $\lambda_{k-1} = \lambda_k = \lambda_{k+1} = 1$, $\delta_{k-1,k+2} = 0$, and $t_f = 15$ ns. The pulse is a single-derivative DRAG shape $\Omega - ia\dot{\Omega}/\Delta_k$, with Ω the standard Hann pulse.

functions proportional to the second derivative of the original drive shape. In the presence of multiple leakage levels, DRAG corrections can be tailored for each leakage coupling and chained together. In the following, we first present the general formulation and then derive the specific solution to this problem.

For an nth-order coupling $\Omega^n/\Delta_{\text{eff}}^{n-1}$ between two levels separated by Δ, the Hamiltonian in the two-level subspace is given as

$$\hat{H} = -\frac{\Delta}{2}\hat{\sigma}_{jk}^z + \left(\frac{\Omega^n}{\Delta_{\text{eff}}^{n-1}}\frac{\hat{\sigma}_{jk}^+}{2} + \text{h.c.}\right), \tag{11}$$

where $\hat{\sigma}_{jk}^z = (|k\rangle\langle k| - |j\rangle\langle j|)$ and $\hat{\sigma}_{jk}^+ = |k\rangle\langle j|$. The symbol Δ_{eff} represents an effective prefactor that renormalizes the coupling strength. Assuming $\Omega^n/\Delta_{\text{eff}}^{n-1} \ll \Delta$, we perform a perturbative expansion with the antihermitian generator $\hat{S}(\tilde{\Omega}) = (\tilde{\Omega}^n/2\Delta\Delta_{\text{eff}}^{n-1})\hat{\sigma}_{jk}^+ - \text{h.c.}$ The time-dependent frame transformation is given as

$$\hat{H}'(\Omega) = \hat{V}(\tilde{\Omega})\hat{H}(\Omega)\hat{V}^\dagger(\tilde{\Omega}) + i\dot{\hat{V}}(\tilde{\Omega})\hat{V}^\dagger(\tilde{\Omega}), \tag{12}$$

with $\hat{V}(\tilde{\Omega}) = e^{\hat{S}(\tilde{\Omega})}$. This transformation yields

$$\hat{H}'(\Omega) = i\dot{\hat{S}}(\tilde{\Omega}) + \hat{H}(\Omega) + [\hat{S}(\tilde{\Omega}), \hat{H}(\Omega)] + \cdots \quad (13)$$

$$\approx -\frac{\Delta}{2}\hat{\sigma}^z_{jk} + \frac{1}{\Delta^{n-1}_{\text{eff}}}\left(\Omega^n - \tilde{\Omega}^n + i\frac{\mathrm{d}}{\mathrm{d}t}\frac{\tilde{\Omega}^n}{\Delta}\right)\frac{\hat{\sigma}^+_{jk}}{2} + \text{h.c.},$$

where we keep only the leading-order perturbation. Following from the equation above, the DRAG pulse is given by

$$\Omega^n = \tilde{\Omega}^n - i\frac{\mathrm{d}}{\mathrm{d}t}\frac{\tilde{\Omega}^n}{\Delta}. \quad (14)$$

Therefore, we can derive a drive pulse Ω resistant to this error based on the initial shape $\tilde{\Omega}$ and its derivative.

To ensure that the unitary evolution remains consistent under the frame transformation in Eq. (12), it is crucial that the generator $\hat{S}$ vanishes at the beginning and end of the evolution. To achieve this, we use the following initial pulse shape:

$$\Omega_{\text{I}}(t) = \Omega_{\text{max}}\left[\frac{1}{16}\cos\left[6\pi\frac{t}{t_{\text{f}}}\right] - \frac{9}{16}\cos\left[2\pi\frac{t}{t_{\text{f}}}\right] + \frac{1}{2}\right], \quad (15)$$

with Ω_{max} the drive amplitude. This pulse is designed to ensure that the function and its derivatives up to fourth order vanish at the start and end of the gate, as required for our frame transformation. However, without any correction term, it has more spectral spreading than the Hann pulse and exhibits worse performance.

C. First-order (linearized) solution for qudits

To manage the two different leakage channels with opposite energy gaps as shown in Figs. 1(c) and 3(b), two degrees of freedom are required. The leading-order leakage errors in Eq. (8) is associated with the ladder couplings between $|k-1\rangle \leftrightarrow |k\rangle$ and $|k+1\rangle \leftrightarrow |k+2\rangle$. Both of these are first-order transitions with $n = 1$ in Eq. (11). To address the two errors, two DRAG corrections can be introduced recursively [42,48] as

$$\Omega = \Omega_1 - i\frac{\dot{\Omega}_1}{\Delta_L}, \quad (16)$$

$$\Omega_1 = \Omega_2 - i\frac{\dot{\Omega}_2}{\Delta_H}, \quad (17)$$

where $\Delta_H = E_{|k+2\rangle} - E_{|k+1\rangle}$ and $\Delta_L = E_{|k\rangle} - E_{|k-1\rangle}$ are energy gaps with respect to the upper and lower adjacent levels. For Ω_2 we use Ω_{I} in Eq. (15). The detailed derivation based on perturbation theory is provided in Appendix E. Each of these expressions is designed to address one leakage pathway, and their order is interchangeable due to the linearity of derivatives.

This is different from the high-order perturbative solution proposed in Ref. [41], where no second derivatives were introduced and the result is only a compromise between different errors. This recursive formulation suppresses both errors simultaneously to the leading order and can be extended with additional correction terms if more ancillary levels are involved [42]. Semiclassically, it can be understood as engineering two zero points on the classical spectrum of the pulse. We refer to this DRAG pulse as the DRAG2 pulse. Notably, for a weakly nonlinear oscillator where $\Delta_L \approx -\Delta_H$, the imaginary part of the correction becomes small, and the real part dominates:

$$\Omega \approx \Omega_{\text{I}} + \frac{\dot{\Omega}_{\text{I}}}{\Delta_L^2} \approx \Omega_{\text{I}} + \frac{\dot{\Omega}_{\text{I}}}{\Delta_H^2}. \quad (18)$$

Apart from the leakage error, two other errors, the phase and amplitude errors, must also be addressed to achieve the desired rotation. For a typical qubit operation between $|0\rangle \leftrightarrow |1\rangle$, the phase error comes from both the Stark shift caused by the $|2\rangle$ state and the noncommutativity of the imaginary DRAG correction term. For transitions involving higher levels, the Stark shift is influenced by both the higher and lower adjacent levels. Because the phase accumulation on the states $|k\rangle$ and $|k+1\rangle$ have the same sign, the overall accumulated phase error in this two-level transition is smaller compared to driving $|0\rangle \leftrightarrow |1\rangle$ [45]. Experimentally, this small phase error is often mitigated by applying a constant detuning to the drive [39]. The correction of the drive shape also affects the rotation angle. To compensate for this, a small correction term needs to be added, $\Omega_2 \leftarrow \Omega_2 + \Omega_{\text{amp}}$. Similar to the phase correction, this amplitude error can also be approximately mitigated by calibrating the maximal drive amplitude Ω_{max}.

D. Second-order solution for qudits

Although the two couplings between $|k-1\rangle \leftrightarrow |k\rangle$ and $|k+1\rangle \leftrightarrow |k+2\rangle$ are suppressed by the DRAG2 correction, under a strong drive the higher-order transitions between $|k-1\rangle \leftrightarrow |k+1\rangle$ and $|k\rangle \leftrightarrow |k+2\rangle$ may also play a role. These second-order transitions arise from diagonalizing the direct ladder couplings and are proportional to Ω^2 (see Appendix E). Following the general DRAG expression in Eq. (14), this leads to the chained second-order correction:

$$\Omega_2 = \sqrt{\Omega_3^2 - i\frac{2\Omega_3\dot{\Omega}_3}{\Delta_H}}, \quad (19)$$

$$\Omega_3 = \sqrt{\Omega_4^2 - i\frac{2\Omega_4\dot{\Omega}_4}{\Delta_L}}, \tag{20}$$

where Ω_4 is again taken from Ω_I in Eq. (15). We refer to this pulse, combined with the two corrections in Eqs. (16) and (17) as the DRAG4 pulse.

The second-order corrections introduced above commute with each other but do not commute with the first-derivative corrections. It is important to note that in the recursive DRAG formulation, the second-order correction is applied first to the initial pulse. This ensures that the dynamics in the final effective frame are governed by the initial pulse. This ordering is the reverse of the order of perturbation.

E. Performance benchmarking

To demonstrate the performance of the recursive DRAG pulse, we simulate the time evolution for various gate durations and compare different drive schemes. We use the analytically derived DRAG pulse while numerically calibrating constant detuning δ_d and amplitude Ω_max. The simulation is performed using the full Hamiltonian, truncated at N_max, which is much larger than the qudit size. The average gate fidelity is calculated as [58]

$$F[\hat{\mathcal{U}}_Q] = \frac{\mathrm{Tr}[\hat{\mathcal{U}}_Q\hat{\mathcal{U}}_Q^\dagger]}{d(d+1)} + \frac{\left|\mathrm{Tr}[\hat{\mathcal{U}}_Q\hat{\mathcal{U}}_I^\dagger]\right|^2}{d(d+1)}, \tag{21}$$

with $\hat{\mathcal{U}}_Q$ representing the truncated unitary within the two-level subspace for the targeted transition and $\hat{\mathcal{U}}_I$ the ideal π and $\pi/2$ rotation gates. Note that this fidelity only includes deviations in the gate quality within the two-level subspace and error leakages from the target states $|k\rangle$ and $|k+1\rangle$. Error dynamics that may occur on other levels are discussed in Sec. IV.

In Fig. 4, we compare the gate fidelity between standard Hann pulse, DRAG2 and DRAG4 pulses, for π and $\pi/2$ gates on $|1\rangle \leftrightarrow |2\rangle$. For the Hann pulse defined in Eq. (9), we set the integral of the pulse such that it equals the required gate rotation angle. For short gate durations, below 25 ns, each successive correction improves fidelity by one to two orders of magnitude. For longer gate times, the error is primarily dominated by phase and amplitude errors, which are suppressed by the constant detuning and amplitude optimization for the DRAG pulses. For gate durations shorter than 7 ns, DRAG4 data is omitted, as the pulse amplitude becomes several times larger than the base envelope, violating the perturbative assumptions and is impractical for experimental implementation. This improvement can also be examined by fixing a target fidelity and examining the minimum gate time required to achieve it. For instance, with a target fidelity of 10^{-4}, the DRAG pulse reduces the gate duration to 10 ns from 30 ns for a $\pi/2$ gate, and from more than 40 ns to 15 ns for a π gate.

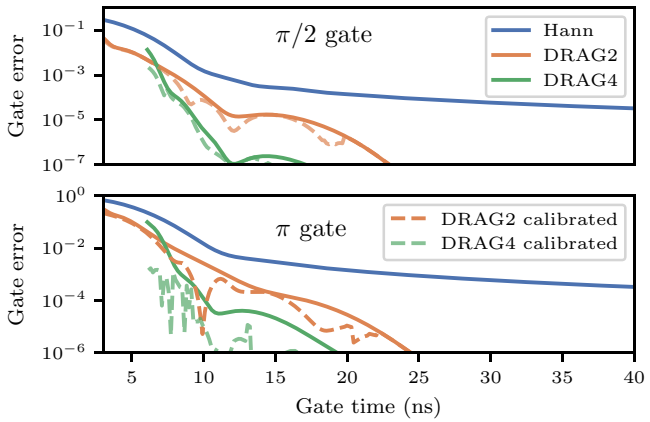

FIG. 4. Gate infidelity as a function of duration for different drive schemes driving the $|1\rangle \leftrightarrow |2\rangle$ transition, with $\alpha/(2\pi) = -200$ MHz and $\omega_{10}/(2\pi) = 5$ GHz. The DRAG2 pulse is defined in Eqs. (16) and (17) and the DRAG4 pulse in Eqs. (19) and (20). The solid line represents the analytically derived DRAG pulse and the dashed line represents gate error with calibrated DRAG coefficients.

Generalizing the analysis to arbitrary $|k\rangle \leftrightarrow |k+1\rangle$ transitions, we apply the same DRAG construction and repeat the benchmarking for different k values. Figure 5(a) shows the fidelity improvement for three different anharmonicities α and gate times. As the anharmonicity decreases, the system more closely resembles a linear oscillator, allowing more levels to be used as quantum registers without significant coupling to environmental noise. However, the energy difference between each level, roughly proportional to the anharmonicity, also decreases. Therefore, we increase the gate time proportionally, inversely to the reduced anharmonicity. The results indicate that the improvements provided by the DRAG corrections for general $|k\rangle$ transitions with varying anharmonicity are analogous to those observed for $|1\rangle \leftrightarrow |2\rangle$ in Fig. 4. This also verifies that the four-level effective model is well suited for studying the transmon ladder transition across different levels. In addition, we observe that control errors decrease as the level k increases, mainly because the leakage coupling to upper and lower levels becomes more symmetric as the level goes up. As the two leakage couplings $|k-1\rangle \leftrightarrow |k\rangle$ and $|k+1\rangle \leftrightarrow |k+2\rangle$ become more symmetric, the phase error is reduced for higher levels, as also observed in Ref. [45].

To further characterize the control of different ladder transitions, we compute the minimal gate duration achievable for an infidelity threshold of 10^{-4}, as depicted in Fig. 5(b). To capture the universality of the pulse solutions, we normalize the time by the energy separation Δ_k for each ladder transition. Remarkably, we see that when comparing different transition indices k, and comparing

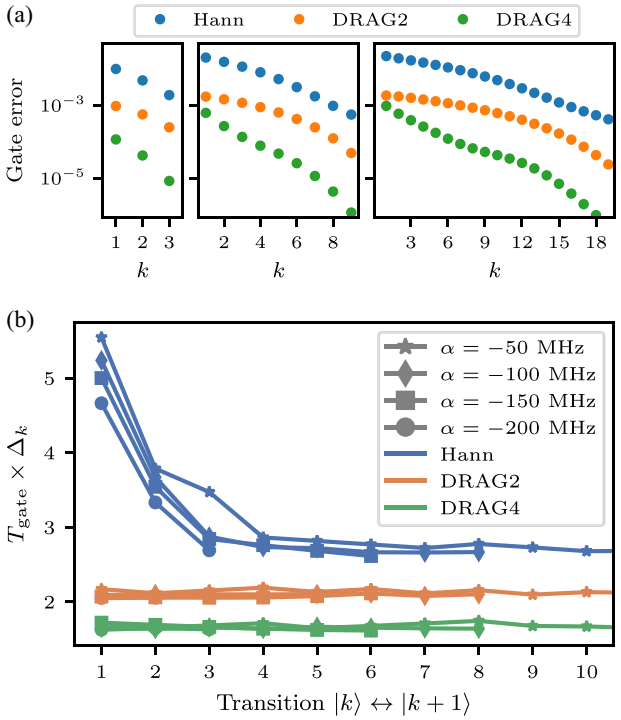

FIG. 5. Controlling the ladder transitions in a transmon qudit. (a) The $\pi/2$ gate error for different ladder transitions $|k\rangle \leftrightarrow |k+1\rangle$ for $\alpha/(2\pi) = -200, -100, -50$ MHz using a fixed gate duration of 8, 15, and 30 ns, respectively. These values of anharmonicity correspond to $E_J/E_C \approx 100$, 355, and 1331, respectively. (b) The minimum gate time required to achieve infidelity of 10^{-4} for different drive schemes and different hardware parameters for a $\pi/2$ gate. The gate duration is multiplied by the corresponding leakage level separation Δ_k, resulting in overlapping outcomes across different hardware parameters and qudit levels.

different values of the anharmonicity, all the values collapse horizontally on the same line for the DRAG family of pulses. When we use DRAG4 pulses instead of DRAG2, we remove two additional weak, two-photon transitions, and these collapse to a yet shorter minimum time line (related to a quantum speed limit for the particular choice of pulse), with apparently even stronger overlap for different parameters. However, this is not the case for the standard Hann pulse, which exhibits a strong dependence on both the anharmonicity (or equivalently E_J/E_C) and the chosen level index. This is due to phase and amplitude errors introduced by asymmetric coupling to neighboring levels. Eliminating those errors, together with the suppression of leakage of different orders, reduces the system to a universal qubitlike model, whose behavior is independent of the specific nature of the suppressed leakage channels.

F. Calibration of the DRAG coefficients

In practice, due to model inaccuracy, the DRAG pulse often needs to be calibrated. A major advantage of the DRAG framework is its modularity and ease of experimental calibration, features that are naturally inherited by the recursive DRAG approach. In this formalism, each error mechanism is addressed by a corresponding derivative correction term, each with its own tunable coefficient $a_{n,l}$. The generalized recursive expression takes the form

$$\Omega^n = \tilde{\Omega}^n - i a_{n,l} \frac{\mathrm{d}}{\mathrm{d}t} \frac{\tilde{\Omega}^n}{\Delta_l}. \tag{22}$$

where n denotes the expansion order and l the target leakage level. For DRAG2 and DRAG4 schemes, this yields two and four independently tunable coefficients, respectively. Ideally, each coefficient targets a distinct error channel.

While the analytical prediction sets all $a_{n,l} = 1$, deviations can arise even in simulation due to coherent interference effects among leakage pathways. As shown in Fig. 4, the DRAG2 pulse with default coefficients already performs well in most regimes, but targeted calibration can further enhance fidelity for specific gate durations.

Figure 6 shows the leakage error landscape for a $|1\rangle \leftrightarrow |2\rangle$ π gate, plotted over the DRAG2 parameters $a_{1,0}$ and $a_{1,3}$ for two gate durations. At $t_f = 15$ ns, the analytical values ($a_{1,0} = a_{1,3} = 1$) are nearly optimal, and each coefficient predominantly affects leakage to its associated level: $a_{1,0}$ governs leakage to $|0\rangle$, and $a_{1,3}$ to $|3\rangle$. Off-target tuning, such as adjusting $a_{1,3}$ when measuring the population of $|0\rangle$, has minimal effect and typically converges toward zero, as smaller corrections reduce unwanted spectral content.

At shorter gate times (e.g., $t_f = 10$ ns), optimal coefficients may deviate from their analytic values due to interference effects, allowing for further cancellation of leakage errors. Nevertheless, the parameter-to-channel mapping remains approximately valid. This observation enables a practical iterative calibration strategy: one may first optimize $a_{1,0}$ by minimizing population in $|0\rangle$, then adjust $a_{1,3}$ to suppress leakage to $|3\rangle$, and finally refine $a_{1,0}$ if needed. The progression of this calibration procedure is illustrated by the arrow markers in the plots.

The DRAG4 case involves a more complex interplay, with multiple coefficients contributing to the same leakage level. As a result, calibration may require joint optimization over two parameters. We discuss this in detail in Appendix F.

IV. ERROR BEYOND THE TARGETED TWO-LEVEL SUBSPACE

In the previous analysis, we focused on the relevant two-level subspace and computed the average gate fidelity of driving a π or $\pi/2$ rotation, using Eq. (21). For a target transition between $|k\rangle \leftrightarrow |k+1\rangle$, this error model includes leakage from the target subspace to $|k-1\rangle$ and $|k+2\rangle$ and

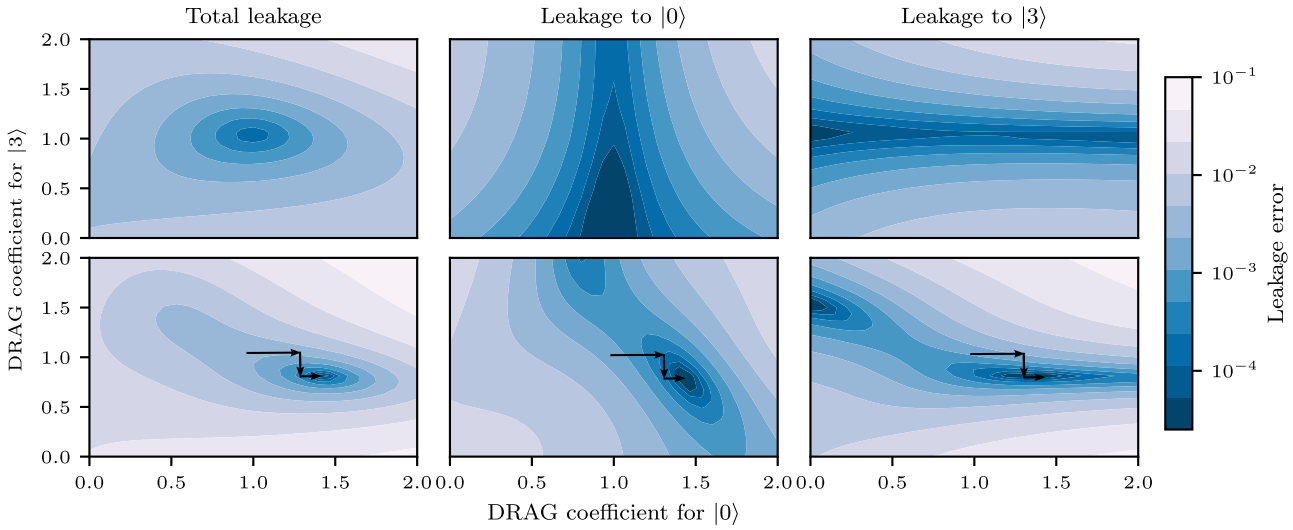

FIG. 6. Calibration of the DRAG2 pulse. Plotted is the leakage error as a function of the DRAG coefficients for DRAG2 pulses $a_{1,0}$ and $a_{1,3}$ for a $|1\rangle \leftrightarrow |2\rangle$ π gate with $\alpha/2\pi = -200$ MHz, with a gate duration of 15 ns (upper row) and 10 ns (lower row). Amplitude and detuning are optimized for each pair of coefficients. While the analytical prediction sets both coefficients to one, the lower row shows deviations due to additional coherent interference effects.

the corresponding phase and amplitude errors. To use it as the building block for universal qudit computational gates, we also need to study its effect on all the K qudit states.

A. Phase error beyond the two target levels

As discussed above, the Stark shift accumulates phases on the affected subspace. The optimized detuning fixes the difference between the phase on $|k\rangle$ and $|k+1\rangle$. However, a phase shift still exists between the subspace and other energy levels. For the target ladder transition, this phase shift is merely a global phase, but for a K-level qudit, it becomes relevant and must be accounted for.

Fortunately, this phase mismatch can be easily calibrated by applying virtual phase gates to each untargeted level [33]. The accumulated phase is calibrated by using a phase-amplification technique. For an operation $\mathrm{RX}_{\pi/2}^{(k,k+1)}$, the state $(|k\rangle + |j\rangle)/\sqrt{2}$ is prepared, where $|j\rangle$ is the state not addressed by the gate. The gate is then applied $8n$ times, followed by a rotation of the system back using $\mathrm{RY}_{\pi/2}^{(k,k+1)}$, similar to a Ramsey experiment. The accumulated phase is then measured on the state $|j\rangle$ and corrected for future use. Virtual phase gate construction is described in Appendix A.

B. Leakage on $|k+2\rangle \leftrightarrow |k+3\rangle$

The DRAG pulse we studied primarily targets leakage involving the target subspace, i.e., leakage from the two target levels to the nearest neighbors, which are separated by approximately Δ_k in the rotating frame. Under a very strong drive, a small population transfer may also appear between nearby states such as $|k+2\rangle \leftrightarrow |k+3\rangle$, which are not directly driven. The level separation between them is about $2|\Delta_k|$. Due to this large separation, the unwanted transition is much smaller, however, it might become non-negligible ($>10^{-4}$) if a DRAG4 pulse is used for a short gate time.

C. Three-photon leakage $|k-1\rangle \leftrightarrow |k+2\rangle$

Another small error that we have not discussed is the three-photon transition between $|k-1\rangle \leftrightarrow |k+2\rangle$. As illustrated in Fig. 1(c), this third-order transition is induced by off-resonant ladder couplings and is proportional to Ω^3. Although the absolute value of this error is accordingly small, due to the very small value of $\delta_{k-1,k+2}$ in a nonlinear

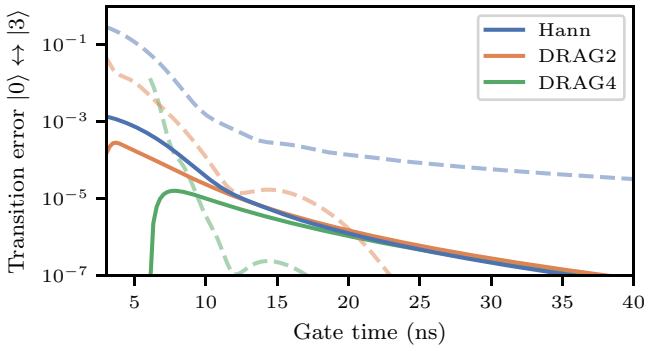

FIG. 7. The three-photon transition error under different drive schemes. The solid lines represent the $|0\rangle \leftrightarrow |3\rangle$ error as a function of the gate time for different drive schemes for a $\pi/2$ gate. The dashed lines represent the gate errors computed for the target subspace for each drive scheme, the same as those shown in Fig. 4(a). The parameters used are also the same.

oscillator, it may still introduce a non-negligible error after DRAG corrections. This error probability, defined by

$$\mathcal{L}_{k-1,k+2} = \frac{1}{4}\left(|\hat{\mathcal{U}}_{k-1,k+2}|^2 + |\hat{\mathcal{U}}_{k+2,k-1}|^2\right) \tag{23}$$

is plotted in Fig. 7 for $k = 1$. Targeting a gate error of 10^{-4}, we see from the plot that this error is generally lower than the two-level gate error calculated in Sec. III, and hence does not pose a significant obstacle.

V. CONCLUSION AND DISCUSSION

Using a universal pulse construction, we have shown how coherent errors can be drastically suppressed in qudit ladder systems. Moreover, they completely predict a universal behavior whereby the minimum gate time to achieve 10^{-4} error can be very accurately calculated, irrespective of the details of the exact energy structure of the nearby surrounding levels.

The method adopts a recursive structure to simultaneously suppress multiple leakage errors. Despite the simple form, the performance benchmarking highlights the substantial error reduction achieved, enabling faster gates and consequently reducing decoherence. This method not only improves gate fidelity for the nonlinear oscillator but also offers a framework that can be adapted to other qudit systems beyond the specific model studied here.

Our results offer valuable insights into the relationship between the number of qudit levels that can be utilized for universal computation, and the corresponding gate time required to achieve a specific fidelity threshold using analytical DRAG pulses. For specific device parameters, the analytical pulses can be further optimized to harness additional coherent interference for better performance or to enhance robustness against system uncertainties. Residual errors beyond those addressed in this work may also be suppressed by further including higher-order derivatives in the DRAG framework or by introducing a weak off-resonant correction drive [59], which we leave for future investigation.

For advanced hardware with high bandwidth waveform generators, where multiplexed frequency is possible, it would be advantageous to implement nonoverlapping ladder transitions in parallel, further improving the efficiency and scalability of qudit-based quantum computing. However, drive-induced Stark shifts and interchannel crosstalk must be carefully accounted for, requiring additional calibration to ensure accurate and high-fidelity gate operations in parallel settings.

ACKNOWLEDGMENTS

The authors would like to thank José Jesus and Shahrukh Chishti for insightful discussion. 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), by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy—Cluster of Excellence Matter and Light for Quantum Computing (ML4Q) EXC 2004/1—390534769, by HORIZON-CL4-2022-QUANTUM-01-SGA Project under Grant 101113946 OpenSuperQPlus100 and the European Union's Horizon Programme (HORIZON-CL4-2021-DIGITALEMERGING-02-10) Grant Agreement 101080085 QCFD. A.L. acknowledges the support of the Natural Sciences and Engineering Research Council of Canada (NSERC), through the Alliance International Catalyst Quantum Grant ALLRP 580928 - 22. The authors gratefully acknowledge the Gauss Centre for Supercomputing e.V. (www.gauss-centre.eu) for funding this project by providing computing time through the John von Neumann Institute for Computing (NIC) on the GCS Supercomputer JUWELS at Jülich Supercomputing Centre (JSC).

DATA AVAILABILITY

The data that support the findings of this article are not publicly available. The data are available from the authors upon reasonable request.

APPENDIX A: UNIVERSALITY OF THE LADDER TRANSITION

In this section, we show that any K-dimensional qudit unitary $\hat{U}$ can be decomposed into $\pi/2$ gates between adjacent levels (ladder gates) $|k\rangle$, $|k+1\rangle$ and virtual phase gates.

1. Decomposition of arbitrary unitary to Givens rotation

In the following, we show that arbitrary qudit unitaries SU(d) can be decomposed into a sequence of Givens rotations and a diagonal phase matrix. We follow the QR decomposition similar to Ref. [60,61] and decompose the unitary by progressively eliminating the left bottom part of the unitary matrix. An upper triangular matrix, which is unitary is easily proved to be a diagonal matrix.

A unitary Givens rotation between two levels j, l is defined as

$$\hat{G}(\gamma,\varphi) = \begin{pmatrix} \ddots & & & & \\ & \cos\frac{\gamma}{2} & \dots & -ie^{i\varphi}\sin\frac{\gamma}{2} & \\ & \vdots & \ddots & \vdots & \\ & -ie^{-i\varphi}\sin\frac{\gamma}{2} & \dots & \cos\frac{\gamma}{2} & \\ & & & & \ddots \end{pmatrix}. \tag{A1}$$

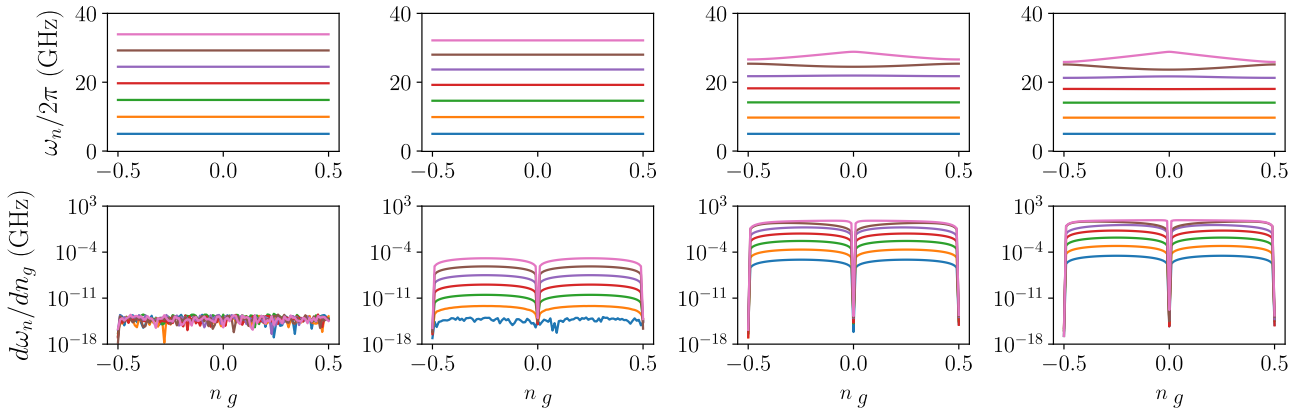

FIG. 8. Energy spectrum of the transmon circuit as a function of the dimensionless gate voltage n_g for four different values of $\alpha/(2\pi) = \{-50, -100, -200, -300\}$ (MHz). For the numerical simulations, we have fixed E_J for obtaining $\omega_{10}/(2\pi) = 5$ (GHz), and as expected, for increasing anharmonicity, the energy spectrum becomes more sensitive to charge fluctuations.

The entries not explicitly defined are filled with identities, i.e., one in the diagonal entries and zero otherwise.

The QR decomposition eliminates each column from left to right and for each row from bottom to the diagonal. This ensures that the eliminated entries will remain zero in later steps. There are in total $M = (K-1)K/2$ Givens rotations. For the mth Givens rotation $G^{(m)}$ designed to eliminate the entry j, l, the parameters are recursively defined by

$$\tan\gamma_m = 2|U^{(m)}_{j,l}/U^{(m)}_{j-1,l}|, \tag{A2}$$

$$\varphi_m = \pi/2 + \arg(U^{(m)}_{j-1,l}) - \arg(U^{(m)}_{j,l}), \tag{A3}$$

where $\hat{U}^{(m)}$ is the remaining unitary after eliminating the first m entries, i.e., $\hat{U}^{(m)} = \hat{G}^{(m-1)}\hat{G}^{(m-2)}\cdots\hat{G}^{(1)}\hat{U}$.

Note that the Givens rotation applied above is always a rotation between two adjacent levels $j, j-1$, which can be directly implemented by ladder coupling. To further simplify the native gate set, it is common to use the *ZXZXZ* decomposition [50] designed for a qubit, which decomposes an arbitrary Givens rotation into three RZ gates and two $\pi/2$ gates

$$\hat{G}(\gamma,\varphi) = \mathrm{RZ}\left(-\varphi - \frac{\pi}{2}\right)\mathrm{RX}_{\frac{\pi}{2}}\mathrm{RZ}(\pi-\gamma)\mathrm{RX}_{\frac{\pi}{2}}\mathrm{RZ}\left(\varphi - \frac{\pi}{2}\right), \tag{A4}$$

where the subscript denotes the rotation angle. The diagonal matrix after the QR decomposition can also be easily written as a series of RZ gates. However note that the DRAG2 and DRAG4 pulses derived in the main text can also be used with any rotation and phase angle.

2. Virtual phase gate in a qudit

The above decomposition resolves arbitrary unitary to $\pi/2$ gates and RZ gates. The former is the main focus of the main text. Here, we show that the *RZ* gate between two arbitrary qudit levels, defined by

$$\mathrm{RZ}(\lambda) = \begin{pmatrix} e^{-i\lambda/2} & 0 \\ 0 & e^{i\lambda/2} \end{pmatrix}, \tag{A5}$$

can be implemented by adjusting a constant phase of the drive [50]. This virtual phase implementation will significantly reduce the duration of the circuit. In contrast to the case of qubits, since there are $K-1$ ladder transitions, we need to track the accumulated phase for each drive, which we denote as θ_k.

A rotation between $|k\rangle$ and $|k+1\rangle$ is implemented by a Hamiltonian

$$\hat{H}^{(k)}_{\Omega}(\theta) = \begin{pmatrix} 0 & \frac{1}{2}e^{i\theta_k}\Omega(t) \\ \frac{1}{2}e^{-i\theta_k}\Omega(t) & 0 \end{pmatrix}, \tag{A6}$$

where $\Omega(t)$ is the time-dependent drive amplitude and θ a constant phase of the drive. This angle θ adjusts the axis in the XY plane, around which the rotation is performed. The corresponding unitary evolution is denoted by $\hat{\mathcal{U}}^{(k)}_{\Omega}(\theta) = e^{-i\hat{H}^{(k)}_{\Omega}(\theta)t}$. It is then straightforward to show that

$$\mathrm{RZ}^{(k,k+1)}(\phi)\hat{\mathcal{U}}^{(k)}_{\Omega}(\theta) = \hat{\mathcal{U}}^{(k)}_{\Omega}(\theta-\phi)\mathrm{RZ}^{(k,k+1)}(\phi). \tag{A7}$$

This relation indicates that by phase shifting the drive Ω by $-\phi$, the RZ phase gate can be effectively moved to the end of the gate operation. Since it appears at the end, it can be neglected, as measurements only capture state populations.

The above is enough for qubit operation. For qudit operation, however, the presence of other computational levels has to be taken into consideration. Therefore, we need

to consider an RZ gate between arbitrary two levels and obtain the following expressions:

$$\mathrm{RZ}^{(j,k)}(\phi)\hat{\mathcal{U}}_{\Omega}^{(k)}(\theta) = \hat{\mathcal{U}}_{\Omega}^{(k)}(\theta+\phi/2)\mathrm{RZ}^{(j,k)}(\phi), \quad \text{(A8)}$$

$$\mathrm{RZ}^{(k,l)}(\phi)\hat{\mathcal{U}}_{\Omega}^{(k)}(\theta) = \hat{\mathcal{U}}_{\Omega}^{(k)}(\theta-\phi/2)\mathrm{RZ}^{(k,l)}(\phi), \quad \text{(A9)}$$

$$\mathrm{RZ}^{(j,k+1)}(\phi)\hat{\mathcal{U}}_{\Omega}^{(k)}(\theta) = \hat{\mathcal{U}}_{\Omega}^{(k)}(\theta-\phi/2)\mathrm{RZ}^{(j,k+1)}(\phi), \quad \text{(A10)}$$

$$\mathrm{RZ}^{(k+1,l)}(\phi)\hat{\mathcal{U}}_{\Omega}^{(k)}(\theta) = \hat{\mathcal{U}}_{\Omega}^{(k)}(\theta+\phi/2)\mathrm{RZ}^{(k+1,l)}(\phi), \quad \text{(A11)}$$

with $j < k$ and $l > k+1$. Notice that the adjusted phase is reduced by half because only one of the levels overlaps between RZ and the transition gate.

APPENDIX B: TRANSMON CIRCUIT IN THE CHARGE REPRESENTATION

In this Appendix, we will discuss a more general description of the transmon circuit that goes beyond the standard Duffing oscillator [53]. The fundamental aspect of this modeling is the representation of both charge and phase operators in Eq. (1). In the charge qubit description, the operator $\hat{n}$ describes the excess of Cooper pair on the superconducting islands, while the cosine operator $\cos(\hat{\varphi})$ describes the tunneling between them along the junction. Explicitly, we write

$$\hat{n} := \sum_{n\in\mathbb{Z}} n\,|n\rangle\langle n|\,, \quad \text{(B1)}$$

$$\exp\left(i\hat{\phi}\right) := \sum_{n\in\mathbb{Z}} |n\rangle\langle n+1|\,. \quad \text{(B2)}$$

In this representation, the eigenstates of the transmon are obtained by numerically diagonalizing the Hamiltonian to a subspace spanned by a few charge states. This modifies the control operator $\hat{n}$ in such a basis so that it exhibits selection rules different from the bosonic oscillator defining the Duffing oscillator [see Eq. (6)].

The gate voltage $n_g(t)$ responsible for driving the transitions on the transmon circuit is susceptible to fluctuations, which could be thermal, due to wiring circuits and quasiparticle tunneling through the junction, or nonthermal, due to impedance mismatching with the signal generator. Thus, we need to quantify the fluctuation of the energy levels of the transmon by varying $n_g(t)$. Figure 8 shows the low-lying energy spectrum as a function of the gate voltage n_g. We have selected E_C and E_J such that the $\omega_{10}/(2\pi) = (\omega_1-\omega_0)/(2\pi) = 5$ GHz, and we vary the anharmonicity $\alpha = \omega_{21} - 2\omega_{10}$ to be in the range $\alpha/(2\pi) = (-50, -300)$ (MHz).

We observe increasing charge dispersion for larger values of α. The main reason for the increase of charge dispersion with decreasing E_J is the constraint imposed to obtain always the same ω_{10} frequency; consequently, fewer states are confined in the cosine potential. This feature is more appreciable when we see the variation of the energy spectrum $\partial\omega_{k+1,k}/\partial n_g$ with respect to the gate voltage, where for smaller α the fluctuations are on the order of KHz.

In this scenario, depending on our transmon parameters, we need to carefully select the workable low-lying energy levels for our qudit gates. In our case, we follow a different approach than Ref. [49]; rather than compute the ratio between the deep potential with the energy spacing, we consider the average of the fluctuation over n_g. In this work, we set a truncation at $\partial\omega_{N_{\max}}/\partial n_g \approx 10^{-3}$ (GHz) and consider any eigenstates with lower dispersion suitable as a qudit level. This results in the number of levels available in the nonlinear oscillator in Fig. 1.

This constraint on dispersion also extends to the dephasing time where we have used $1/f$ noise as the most detrimental source of decoherence, which can be estimated by the relation [38,62]

$$\frac{1}{T_{\phi}^{(k)}} = A_{n_g}\left|\frac{\partial\omega_{k+1,k}}{\partial n_g}\right|\sqrt{2|\ln\left(\omega_{\mathrm{low}}t_{\mathrm{exp}}\right)|}, \quad \text{(B3)}$$

where $\omega_{k+1,k} = \omega_{k+1} - \omega_k$ and $A_{n_g} = 10^{-4}e$ is the noise strength [63–65] with e being the electron charge. Also, $\omega_{\mathrm{low}} = 2\pi/t_{\mathrm{exp}}$ corresponds to the infrared cutoff due to the finite data acquisition time $t_{\mathrm{exp}} = 10^4$ ns [38]

This is illustrated in Fig. 9, where one point corresponds to one qudit eigenstate with a specific hardware parameter. As the anharmonicity decreases, more and more levels with a coherence time longer than 100 μs can be included as quantum information registers.

For amplitude damping, we estimate T_1 assuming that the main loss mechanism corresponds to capacitive losses. In such a way, Fermi's golden rules give the relation [66]

$$\frac{1}{T_1^{(k)}} = |\langle k|\,\hat{n}\,|k+1\rangle|^2 S(\omega_{k+1,k}), \quad \text{(B4)}$$

where the spectral density for the capacitive losses reads [62,67]

$$S(\omega_{k+1,k}) = \frac{4\hbar E_C}{Q_{\mathrm{cap}}(\omega_{k+1,k})}\left[\frac{\coth\left(\frac{\hbar|\omega_{k+1,k}|}{2k_BT}\right)}{1+\exp\left(-\frac{\hbar\omega_{k+1,k}}{k_BT}\right)}\right], \quad \text{(B5)}$$

with $Q_{\mathrm{cap}}(\omega_{k+1,k}) = 10^6(2\pi\times 6\ \mathrm{GHz}/|\omega_{k+1,k}|)^{0.7}$ [68,69] the capacitive quality factor per ladder transition. Also, k_B

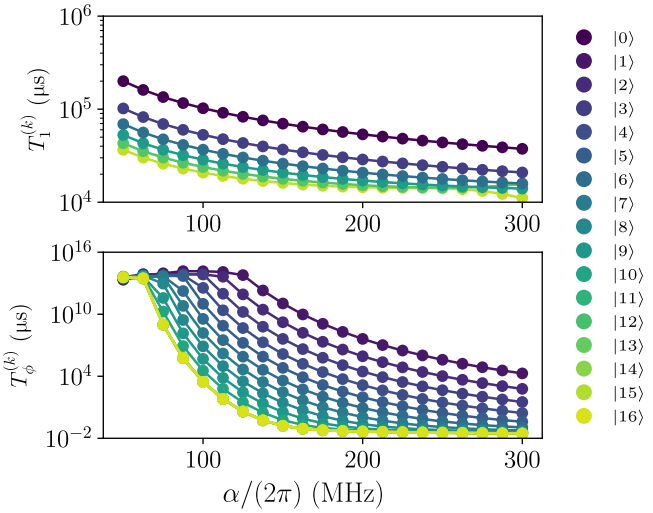

FIG. 9. Coherence times of the transmon circuit as a function of α for different transition frequency $\omega_{k+1,k}$. We set E_J and E_C such that the transition frequency equal to $\omega_{10}/(2\pi) = 5$ (GHz). For the amplitude damping, we assume capacitive losses and dephasing correspond to charge fluctuations.

is the Boltzmann constant, and $T = 15$ mK is the temperature. Since this value is not strongly dependent on the levels in our system studied, we do not use it to truncate the qudit level.

For completeness, we plot the T_1 for the different energy levels in Fig. 9. We should note that improvement of the coherence times could be possible by implementing different fabrication techniques such as surface error mitigation [70,71], changing the niobium with tantalum as the base superconductor [72,73] or mitigating the micromotion of the circuitry [74], among other techniques. Such shielding on the transmon circuit leads to coherence times nearly in the millisecond scale.

APPENDIX C: DERIVATION OF THE LEAKAGE MANIFOLD

Here, we will show that the energy diagram for any qudit gate between the states $(k+1, k)$ is represented as in Fig. 1(c). In other words, if we want to implement this single qudit gate, there appears to be a nearly resonant transition between the states $|k-1\rangle \leftrightarrow |k+2\rangle$. To do so, let us consider the explicit form of the energy of the kth energy level after the frame transformation in Eq. (7)

$$\tilde{\Delta}_k = \omega_k - k(\omega_{k+1} - \omega_k - \delta_d). \tag{C1}$$

For the Duffing oscillator model, we know that $\omega_k = \omega_k - \alpha k(k-1)/2$, where $\omega = \sqrt{8E_C E_J} - E_C$ is the transmon frequency, and $\alpha = -E_C$ is the anharmonicity, respectively. Thus, $\Delta_{k-1} = (k-1)(\alpha(k+2) + 2\delta_d)/2$ while $\Delta_{k+2} = (k+2)(\alpha(k-1) + 2\delta_d)/2$. Thus, the detuning between these energy levels is $\delta_{k-1,k+2} = 3\delta_d$ for all values of k, which is zero if the drive is resonant.

However, such a description of the system Hamiltonian is only valid for larger E_J/E_C. Thus, for obtaining better estimation of the detuning, we consider the eigenenergies of the transmon obtained by numerical diagonalizing Eq. (1). Figure 2(d) shows $\delta_{k-1,k+2}$ as a function of the anharmonicity α for several ladder transitions at $n_g = 0$; from the figure we appreciate an inverse relation between the degeneracy of the leakage state with the anharmonicity, recovering the previous calculation result when $\alpha = -2\pi \times 50$ (MHz). Moreover, we also see an increase of such a discrepancy with the qudit manifold to be addressed, this effect is mainly produced by the sensitivity of the energy spectrum to the charge noise (see Fig. 8).

APPENDIX D: ERROR BUDGET FOR TRANSMON QUDIT

We define the leakage error of a quantum gate as

$$\mathcal{L} = 1 - \mathcal{F}_L = 1 - \frac{\mathrm{Tr}[\hat{\mathcal{U}}_Q \hat{\mathcal{U}}_Q^\dagger]}{2}, \tag{D1}$$

where $\hat{\mathcal{U}}_Q$ denotes the projection of the full unitary evolution onto the two-level target subspace spanned by $|k\rangle, |k+1\rangle$. Correspondingly, leakage to a specific state $|l\rangle$ outside the target subspace is quantified as

$$\mathcal{L}_l = \frac{1}{2} \sum_{j \in \{k,k+1\}} |\mathcal{U}_{l,j}|^2, \tag{D2}$$

where $\mathcal{U}_{l,j}$ is the matrix element of the full unitary operator.

To capture errors that occur within the computational subspace, such as phase and amplitude errors, we define the subspace error:

$$\mathcal{E}_Q = 1 - \mathcal{F}_Q = 1 - \frac{\left|\mathrm{Tr}[\hat{\mathcal{U}}_Q \hat{\mathcal{U}}_I^\dagger]\right|^2}{2\mathrm{Tr}[\hat{\mathcal{U}}_Q \hat{\mathcal{U}}_Q^\dagger]}, \tag{D3}$$

where $\hat{\mathcal{U}}_I$ is the ideal target unitary (e.g., a π or $\pi/2$ gate). This metric captures phase and amplitude errors within the logical subspace. To distinguish phase error from amplitude error, we optimize the drive amplitude to minimize $\mathcal{E}_Q$, and attribute the residual error to phase mismatch.

Using the above definitions, the total gate fidelity $\mathcal{F}$ in Eq. (21) can be expressed as

$$\mathcal{F}[\hat{\mathcal{U}}_Q] = \frac{2}{3}\mathcal{F}_Q\mathcal{F}_L + \frac{\mathcal{F}_L}{3}. \tag{D4}$$

Following the above definition, we also studied the error budget for different drive schemes. Figure 10 complements

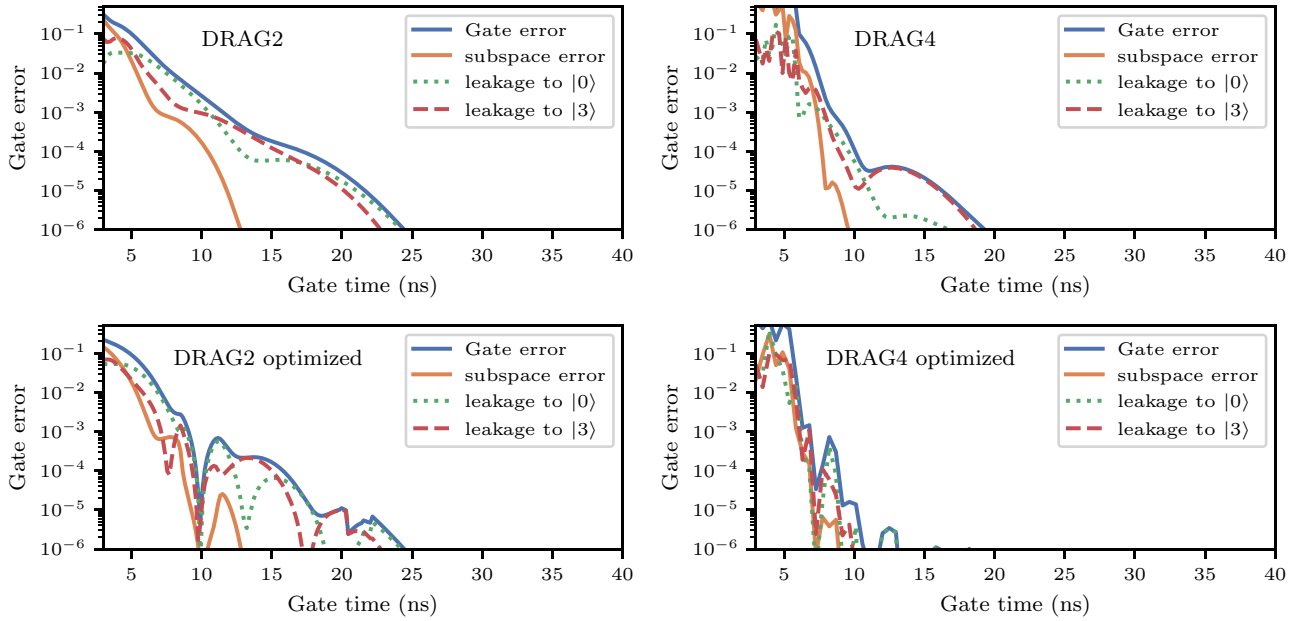

FIG. 10. Error budget for a π rotation between $|1\rangle \leftrightarrow |2\rangle$ in a transmon qudit using different DRAG pulses, with $\alpha/(2\pi) = -200$ MHz. Each panel shows the total gate error alongside contributions from subspace error and leakage to $|0\rangle$ and $|3\rangle$. Top row: uncalibrated DRAG2 and DRAG4 pulses; bottom row: same pulses' ansatz after optimizing the DRAG coefficients.

Fig. 3(a) by providing a detailed breakdown of gate errors under DRAG-based control methods. In the top row, uncalibrated DRAG2 and DRAG4 pulses show that leakage remains the dominant error, which aligns with our observation of the quantum speed limit in Fig. 5. If additional optimization is performed, as shown in the bottom row, both leakage and subspace errors are further reduced, especially for DRAG4, highlighting the effectiveness of recursive DRAG combined with targeted calibration. Here, we group phase and amplitude errors as "subspace error" since both drive detuning and amplitude are optimized, making it no longer possible to isolate phase error. However, we note that full calibration of DRAG4 parameters may increase experimental overhead due to the larger parameter space (see Appendix F).

APPENDIX E: DERIVATION OF RECURSIVE DRAG PULSE

1. Single-photon correction

In the following, we show the derivation of the recursive DRAG pulse shape designed to suppress the two single-photon transitions $|k-1\rangle \leftrightarrow |k\rangle$ and $|k+1\rangle \leftrightarrow |k+2\rangle$. Our general approach is to progressively derive the effective frame and the corresponding drive shapes to minimize the prevalent error. Throughout the calculation, we keep the perturbative correction up to the second order for all the terms with two exceptions: the matrix entry $(0,3)$, which characterizes a three-photon leakage due to the small energy separation, and the entry $(1,2)$, which describes the pulse amplitude correction. For those two, we keep the terms up to the third-order correction.

We start with the rotating-frame Hamiltonian in Eq. (8)

$$\hat{H}_0 = \begin{pmatrix} -\Delta_L & \frac{\lambda_1\bar{\Omega}}{2} & 0 & 0 \\ \frac{\lambda_1\Omega}{2} & \delta_d & \frac{\lambda_2\bar{\Omega}}{2} & 0 \\ 0 & \frac{\lambda_2\Omega}{2} & 2\delta_d & \frac{\lambda_3\bar{\Omega}}{2} \\ 0 & 0 & \frac{\lambda_3\Omega}{2} & \Delta_H + 3\delta_d \end{pmatrix}, \tag{E1}$$

where $\Delta_H = \Delta_k + \delta_{k-1,k+2}$ and $\Delta_L = -\Delta_k$. For ease of notation, we use λ_1, λ_2, λ_3 for λ_{k-1}, λ_k and λ_{k+1} in this section. We define the first transition targeting the single-photon leakage error, $|k+1\rangle \leftrightarrow |k+2\rangle$. For small δ_d, as is typical in the transmon regime, this is the largest leakage source (see Fig. 2). The frame transformation generator is given by

$$\hat{S}_{0\to1} = \begin{pmatrix} 0 & -\frac{\epsilon\lambda_1\bar{\Omega}_1}{2\Delta_H} & 0 & 0 \\ \frac{\epsilon\lambda_1\Omega_1}{2\Delta_H} & 0 & -\frac{\epsilon\lambda_2\bar{\Omega}_1}{2\Delta_H} & 0 \\ 0 & \frac{\epsilon\lambda_2\Omega_1}{2\Delta_H} & 0 & -\frac{\epsilon\lambda_3\bar{\Omega}_1}{2\Delta_H} \\ 0 & 0 & \frac{\epsilon\lambda_3\Omega_1}{2\Delta_H} & 0 \end{pmatrix}, \tag{E2}$$

where ϵ is a small parameter to track the perturbation order. The denominator Δ_H is chosen such that in the effective frame, the matrix entry $(2,3)$ is zero. In addition, $\hat{S}_{0\to1}$ is chosen to be proportional to the control term in $\hat{H}_0$; this is designed in particular such that there is no derivative

$$\hat{H}_1 - \hat{H}_{1,\mathrm{diag}} = \begin{pmatrix} 0 & \frac{1}{2}\epsilon\lambda_{r1}\bar{\Omega}_1 & \frac{-\Delta_L\epsilon^2\lambda_1\lambda_2\bar{\Omega}_1^2}{8\Delta_H^2} & \epsilon^3\bar{\Omega}_{L03}^{(1)} \\ \frac{1}{2}\epsilon\lambda_{r1}\Omega_1 & 0 & \frac{\epsilon\lambda_2\left(\bar{\Omega}_1+\epsilon^2\bar{\Omega}_c^{(1)}\right)}{2} & \frac{\epsilon^2\lambda_2\lambda_3\bar{\Omega}_1^2}{8\Delta_H} \\ \frac{-\Delta_L\epsilon^2\lambda_1\lambda_2\Omega_1^2}{8\Delta_H^2} & \frac{\epsilon\lambda_2\left(\Omega_1+\epsilon^2\Omega_c^{(1)}\right)}{2} & 0 & 0 \\ \epsilon^3\Omega_{L03}^{(1)} & \frac{\epsilon^2\lambda_2\lambda_3\Omega_1^2}{8\Delta_H} & 0 & 0 \end{pmatrix}, \tag{E3}$$

where $\Omega_c^{(1)}$ and $\Omega_{L03}^{(1)}$ denote the third-order error to the drive amplitude in this frame and the three-photon leakage transition, which we do not explicitly use in the following calculation. Notice that in the effective frame, we obtain a renormalized leakage rate $\lambda_{r1}\Omega_1$ between $|k-1\rangle$ and $|k\rangle$, with $\lambda_{r1} = \lambda_1(1-\Delta_L/\Delta_H) \approx 2\lambda_1$ in the limit $\delta_{k-1,k+2} \to 0$. This explains why the leakage increases with only one single derivative DRAG correction in Fig. 3(b). This prefactor also needs to be taken into consideration when making perturbative assumptions. The diagonal energy terms are given by

$$E_{1,|k-1\rangle} = -\Delta_L - \frac{\lambda_1^2\mathrm{Re}\left(\Omega\bar{\Omega}_1\right)}{2\Delta_H} + \frac{\lambda_1^2\Delta_L|\Omega_1|^2}{4\Delta_H^2}, \tag{E4}$$

$$E_{1,|k\rangle} = \delta_d + \left(\frac{\lambda_1^2-\lambda_2^2}{2\Delta_H}\right)\mathrm{Re}\left(\Omega\bar{\Omega}_1\right) - \frac{\lambda_1^2\Delta_L|\Omega_1|^2}{4\Delta_H^2}, \tag{E5}$$

$$E_{1,|k+1\rangle} = 2\delta_d + \left(\frac{\lambda_2^2-\lambda_3^2}{2\Delta_H}\right)\mathrm{Re}\left(\Omega\bar{\Omega}_1\right) + \frac{\lambda_3^2|\Omega_1|^2}{4\Delta_H}, \tag{E6}$$

$$E_{1,|k+2\rangle} = 3\delta_d + \Delta_H + \frac{\lambda_3^2\mathrm{Re}\left(\Omega\bar{\Omega}_1\right)}{2\Delta_H} - \frac{\lambda_3^2|\Omega_1|^2}{4\Delta_H}. \tag{E7}$$

Secondly, we target the single photon leakage between state $|k-1\rangle$ and $|k\rangle$, with the frame transformation generator

$$\hat{S}_{1\to2} = \begin{pmatrix} 0 & -\frac{\epsilon\lambda_{r1}\bar{\Omega}_2}{2\Delta_L} & 0 & 0 \\ \frac{\epsilon\lambda_{r1}\Omega_2}{2\Delta_L} & 0 & -\frac{\epsilon\lambda_2\bar{\Omega}_2}{2\Delta_L} & 0 \\ 0 & \frac{\epsilon\lambda_2\Omega_2}{2\Delta_L} & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \tag{E8}$$

This, together with the substitution $\Omega_1 = \Omega_2 - i(\dot{\Omega}_2/\Delta_L)$, results in the suppression of the transition and gives $\hat{H}_2$

$$\hat{H}_2 - \hat{H}_{2,\mathrm{diag}} = \begin{pmatrix} 0 & 0 & -\frac{\Delta_L\epsilon^2\lambda_2\bar{\Omega}_1^2}{8\Delta_H^2} - \frac{\epsilon^2\lambda_2\lambda_{r1}\bar{\Omega}_2^2}{8\Delta_L} & \bar{\Omega}_{L03}^{(1)} \\ 0 & 0 & \frac{\epsilon\lambda_2\left(\bar{\Omega}_2+\epsilon^2\bar{\Omega}_c^{(2)}\right)}{2} & \frac{\epsilon^2\lambda_2\lambda_3\bar{\Omega}_1^2}{8\Delta_H} \\ -\frac{\Delta_L\epsilon^2\lambda_2\Omega_1^2}{8\Delta_H^2} - \frac{\epsilon^2\lambda_2\lambda_{r1}\Omega_2^2}{8\Delta_L} & \frac{\epsilon\lambda_2\left(\Omega_2+\epsilon^2\Omega_c^{(2)}\right)}{2} & 0 & 0 \\ \Omega_{L03}^{(1)} & \frac{\epsilon^2\lambda_2\lambda_3\Omega_1^2}{8\Delta_H} & 0 & 0 \end{pmatrix}. \tag{E9}$$

In addition to the leakage error, the phase error and the amplitude renormalization also need to be considered to get the desired rotation. The time-dependent phase correction is given by

$$\delta_d = -\mathrm{Re}\left(\Omega_1\bar{\Omega}_2\right)\left(\frac{\lambda_2^2}{\Delta_L} - \frac{\lambda_{r1}^2}{2\Delta_L}\right) - \mathrm{Re}\left(\Omega\bar{\Omega}_1\right)\left(-\frac{\lambda_1^2}{2\Delta_H} + \frac{\lambda_2^2}{\Delta_H} - \frac{\lambda_3^2}{2\Delta_H}\right) - |\Omega_1|^2\left(\frac{\Delta_L\lambda_1^2}{4\Delta_H^2} + \frac{\lambda_3^2}{4\Delta_H}\right) - \frac{|\Omega_2|^2\lambda_{r1}^2}{4\Delta_L}, \tag{E10}$$

where $\lambda_{r1} = \lambda_1(1 - \Delta_L/\Delta_H)$. Note that although $\delta_d(t)$ is promoted to a time-dependent function, this does not affect the derivation of the effective model, as the rotating-frame frequency ω_d remains fixed at the idling transition frequency and is independent of the drive.

Apart from that, the correction on the drive shape also slightly affects the rotation angle. A small correction term needs to be added $\Omega_2 \leftarrow \Omega_2 + \Omega_{\text{amp}}$. The analytical formula of the amplitude correction is written as

$$\begin{aligned}\Omega_{\text{amp}} = {} & \Omega_0 |\Omega_1|^2 \left(\frac{\lambda_1^2}{8\Delta_H^2} + \frac{\lambda_3^2}{8\Delta_H^2} - \frac{\lambda_2^2}{4\Delta_H^2}\right) \\ & + \Omega_1 |\Omega_2|^2 \left(\frac{\lambda_{r1}^2}{8\Delta_L^2} - \frac{\lambda_2^2}{4\Delta_L^2}\right) \\ & + \Omega_2 |\Omega_1|^2 \left(-\frac{\lambda_1^2}{4\Delta_H^2} - \frac{\lambda_3^2}{4\Delta_L\Delta_H}\right) \\ & + \Omega_2 \text{Re}\left(\Omega_0\bar{\Omega}_1\right) \left(\frac{\lambda_1^2}{2\Delta_L\Delta_H} + \frac{\lambda_3^2}{2\Delta_L\Delta_H} - \frac{\lambda_2^2}{\Delta_L\Delta_H}\right) \\ & + \Omega_1^2\bar{\Omega} \left(\frac{\lambda_1^2}{8\Delta_H^2} + \frac{\lambda_3^2}{8\Delta_H^2} - \frac{\lambda_2^2}{4\Delta_H^2}\right) \\ & + \delta_d \left(-\frac{\Omega_1}{\Delta_H} - \frac{\Omega_2}{\Delta_L}\right) \\ & + \Omega_1^2\bar{\Omega}_1 \left(-\frac{\Delta_L\lambda_1^2}{8\Delta_H^3} - \frac{\lambda_3^2}{8\Delta_H^2}\right) \\ & + \Omega_2^2\bar{\Omega}_1 \left(\frac{\lambda_{r1}^2}{8\Delta_L^2} - \frac{\lambda_2^2}{4\Delta_L^2}\right) \\ & - \frac{\lambda_1\Omega_1^2\bar{\Omega}_2\lambda_{r1}}{8\Delta_H^2} - \frac{\Omega_2^2\bar{\Omega}_2\lambda_{r1}^2}{8\Delta_L^2}. \end{aligned} \tag{E11}$$

In our investigation, we neglect the time dependence and numerically optimize a fixed correction of the detuning and the amplitude.

2. Two-photon correction

The first two transitions yield the effective Hamiltonian described in Eq. (E9), where the desired transition between $|k\rangle \leftrightarrow |k+1\rangle$ is preserved, with a renormalized effective coupling strength. The diagonalization of the single-photon coupling introduces new two-photon transitions, $|k-1\rangle \leftrightarrow |k+1\rangle$ and $|k\rangle \leftrightarrow |k+2\rangle$, with the coupling strength proportional to Ω^2. This is also obtained for qubit driving in a nonlinear oscillator, as discussed in Ref. [42]. For very strong drive amplitude, these two-photon transitions become the dominant source of error once the single-photon transitions are sufficiently suppressed.

For simplicity, we do not repeat the full calculation as in the last subsection but note the following properties. We can treat Ω^2 as the new coupling g, then derive the same expression to suppress the two leakages as in Eq. (17) but with the coupling g. Moreover, any perturbative diagonalization of the two-photon transitions will introduce corrections only in the order of ϵ^3 or smaller, which is negligible relative to the truncation order considered. By substituting g back into Ω, we obtain the expression in Eq. (20).

3. Three-photon correction

In principle, based on the DRAG2 pulse, we can follow a similar strategy and use a recursive DRAG design to suppress the transition error between state $|0\rangle$ and $|3\rangle$:

$$\Omega_4 = \sqrt[3]{\Omega_5^3 - i\frac{3\Omega_5^2\dot{\Omega}_5}{\delta_{k-1,k+2}}}. \tag{E12}$$

However, due to the small gap between $|k-1\rangle$ and $|k+2\rangle$ in the transmon regime, the imaginary DRAG correction term is much larger and a constant detuning may not suffice to compensate for the phase error. Nevertheless, even with a not perfectly aligned phase, a $\pi/2$ gate can be implemented with the help of virtual phase gates.

Alternatively, one could explore the direct coupling between the states $|k-1\rangle$ and $|k+2\rangle$ instead of relying on the multiphoton process. However, this would require microwave drive generators with a frequency approximately 3 times that of the qubit frequency.

APPENDIX F: CALIBRATION OF THE DRAG4 PULSE

In Fig. 6, we demonstrated the calibration of DRAG2 pulses. We now extend this analysis to DRAG4, which introduces four tunable parameters: $a_{1,0}$ and $a_{1,3}$ for suppressing single-photon leakage to $|0\rangle$ and $|3\rangle$, and $a_{2,0}$ and $a_{2,3}$ for mitigating two-photon leakage to the same states.

Figure 11 illustrates the leakage error landscape by scanning pairs of DRAG4 coefficients while holding the remaining two fixed at their nominal analytical values. As expected, $a_{1,0}$ and $a_{2,0}$ primarily influence leakage to $|0\rangle$, while $a_{1,3}$ and $a_{2,3}$ control leakage to $|3\rangle$. However, in contrast to the DRAG2 case, the two coefficients associated with the same leakage level (e.g., $a_{1,0}$ and $a_{2,0}$) cannot be calibrated independently, since both require measurement of the same leakage population. This complicates the calibration process.

The global minimum of the leakage error in each parameter space is indicated with a black dot in the left two plots. For $|0\rangle$ leakage, the optimal point deviates significantly from the nominal values ($a_{n,l} = 1$), indicating that independent calibration of each coefficient may result in suboptimal performance or require a more extensive search. Nonetheless, hill-climbing optimization such as Nelder-Mead would be a good option requiring few iterations. A potentially more efficient calibration strategy

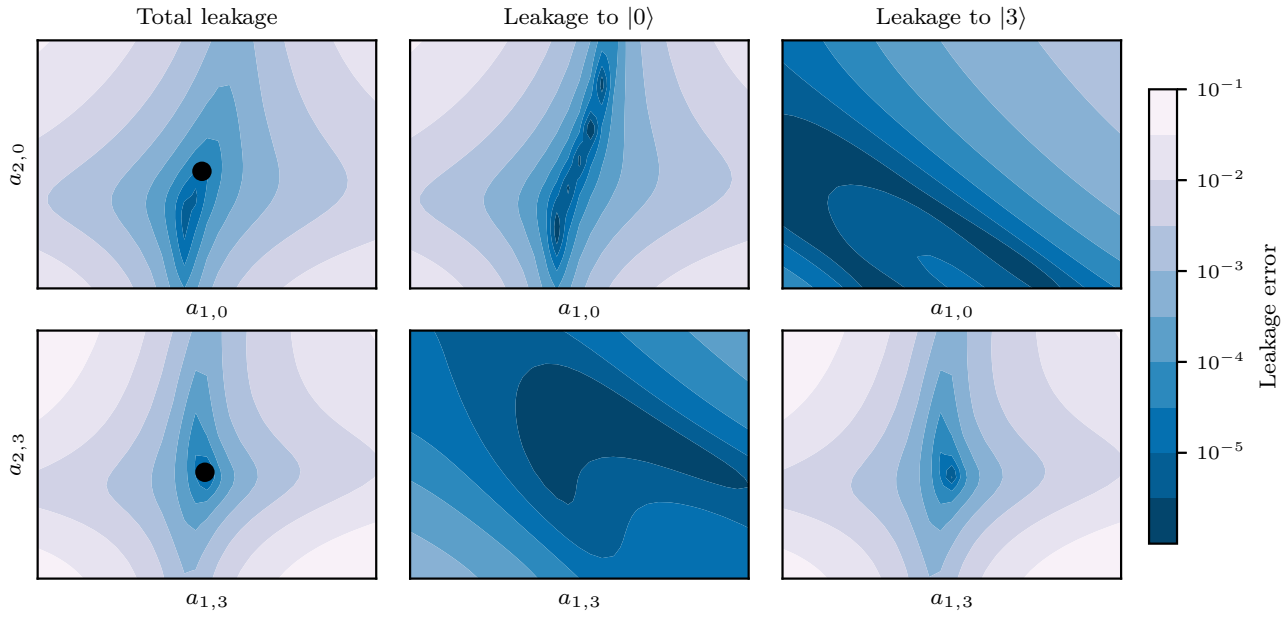

FIG. 11. Calibration of the DRAG4 pulse. Shown is the leakage error as a function of DRAG coefficients for DRAG4 pulses. A $|1\rangle \leftrightarrow |2\rangle$ π gate is calibrated with $\alpha/2\pi = -200$ MHz and a gate duration of 15 ns. The range of all DRAG coefficients is from 0 to 2. In the upper row, the two DRAG coefficients targeting leakage to $|3\rangle$ are fixed at their nominal analytical values (set to one), while the coefficients $a_{1,0}$ and $a_{2,0}$ controlling leakage to $|0\rangle$ are varied. In the lower row, the roles are reversed: $a_{1,0}$ and $a_{2,0}$ are held fixed, and $a_{1,3}$ and $a_{2,3}$ are scanned. The black dot denotes the global minimum of the leakage error across the full parameter space.

might involve fixing part of the coefficients, such as the DRAG2 terms $a_{1,0}$ and $a_{1,3}$, and optimizing the remaining pair. We leave further exploration of such hybrid calibration schemes for future work.

In summary, while the four DRAG4 coefficients offer additional degrees of freedom that can significantly improve gate fidelity, they also increase the complexity of calibration in practice.

[1] D. Gottesman, in *Quantum Computing and Quantum Communications*, edited by G. Goos, J. Hartmanis, J. Van Leeuwen, and C. P. Williams (Springer Berlin Heidelberg, Berlin, Heidelberg, 1999), Vol. 1509, p. 302.

[2] Y.-M. Di and H.-R. Wei, Optimal synthesis of multivalued quantum circuits, Phys. Rev. A **92**, 062317 (2015).

[3] F. Motzoi, M. P. Kaicher, and F. K. Wilhelm, Linear and logarithmic time compositions of quantum many-body operators, Phys. Rev. Lett. **119**, 160503 (2017).

[4] S. Cao, M. Bakr, G. Campanaro, S. D. Fasciati, J. Wills, D. Lall, B. Shteynas, V. Chidambaram, I. Rungger, and P. Leek, Emulating two qubits with a four-level transmon qudit for variational quantum algorithms, Quantum Sci. Technol. **9**, 035003 (2024).

[5] A. Galda, M. Cubeddu, N. Kanazawa, P. Narang, and N. Earnest-Noble, Implementing a ternary decomposition of the Toffoli gate on fixed-frequency transmon qutrits, ArXiv:2109.00558.

[6] B. P. Lanyon, M. Barbieri, M. P. Almeida, T. Jennewein, T. C. Ralph, K. J. Resch, G. J. Pryde, J. L. O'Brien, A. Gilchrist, and A. G. White, Simplifying quantum logic using higher-dimensional Hilbert spaces, Nat. Phys. **5**, 134 (2009).

[7] P. J. Ollitrault, G. Mazzola, and I. Tavernelli, Nonadiabatic molecular quantum dynamics with quantum computers, Phys. Rev. Lett. **125**, 260511 (2020).

[8] A. Miessen, P. J. Ollitrault, and I. Tavernelli, Quantum algorithms for quantum dynamics: A performance study on the spin-boson model, Phys. Rev. Res. **3**, 043212 (2021).

[9] E. Rico, M. Dalmonte, P. Zoller, D. Banerjee, M. Bögli, P. Stebler, and U.-J. Wiese, SO(3) "nuclear physics" with ultracold gases, Ann. Phys. (N. Y.) **393**, 466 (2018).

[10] G. Mazzola, S. V. Mathis, G. Mazzola, and I. Tavernelli, Gauge-invariant quantum circuits for $u(1)$ and yang-mills lattice gauge theories, Phys. Rev. Res. **3**, 043209 (2021).

[11] M. Meth, J. Zhang, J. F. Haase, C. Edmunds, L. Postler, A. J. Jena, A. Steiner, L. Dellantonio, R. Blatt, P. Zoller, T. Monz, P. Schindler, C. Muschik, and M. Ringbauer, Simulating two-dimensional lattice gauge theories on a qudit quantum computer, Nat. Phys. **21**, 570 (2025).

[12] D. Bruß and C. Macchiavello, Optimal eavesdropping in cryptography with three-dimensional quantum states, Phys. Rev. Lett. **88**, 127901 (2002).

[13] H. Bechmann-Pasquinucci and A. Peres, Quantum cryptography with 3-state systems, Phys. Rev. Lett. **85**, 3313 (2000).

[14] M. Grace, C. Brif, H. Rabitz, I. Walmsley, R. Kosut, and D. Lidar, Encoding a qubit into multilevel subspaces, New J. Phys. **8**, 35 (2006).

[15] A. Chiesa, E. Macaluso, F. Petiziol, S. Wimberger, P. Santini, and S. Carretta, Molecular nanomagnets as qubits with

embedded quantum-error correction, J. Phys. Chem. Lett. **11**, 8610 (2020).

[16] E. T. Campbell, Enhanced fault-tolerant quantum computing in d-level systems, Phys. Rev. Lett. **113**, 230501 (2014).

[17] P. J. Low, B. M. White, A. A. Cox, M. L. Day, and C. Senko, Practical trapped-ion protocols for universal qudit-based quantum computing, Phys. Rev. Res. **2**, 033128 (2020).

[18] M. Ringbauer, M. Meth, L. Postler, R. Stricker, R. Blatt, P. Schindler, and T. Monz, A universal qudit quantum processor with trapped ions, Nat. Phys. **18**, 1053 (2022).

[19] P. Hrmo, B. Wilhelm, L. Gerster, M. W. van Mourik, M. Huber, R. Blatt, P. Schindler, T. Monz, and M. Ringbauer, Native qudit entanglement in a trapped ion quantum processor, Nat. Commun. **14**, 2242 (2023).

[20] P. J. Low, B. White, and C. Senko, Control and readout of a 13-level trapped ion qudit, ArXiv:2306.03340.

[21] D. González-Cuadra, T. V. Zache, J. Carrasco, B. Kraus, and P. Zoller, Hardware efficient quantum simulation of non-Abelian gauge theories with qudits on Rydberg platforms, Phys. Rev. Lett. **129**, 160501 (2022).

[22] R. Hussain, G. Allodi, A. Chiesa, E. Garlatti, D. Mitcov, A. Konstantatos, K. S. Pedersen, R. De Renzi, S. Piligkos, and S. Carretta, Coherent manipulation of a molecular ln-based nuclear qudit coupled to an electron qubit, J. Am. Chem. Soc. **140**, 9814 (2018).

[23] M. Chizzini, L. Crippa, L. Zaccardi, E. Macaluso, S. Carretta, A. Chiesa, and P. Santini, Quantum error correction with molecular spin qudits, Phys. Chem. Chem. Phys. **24**, 20030 (2022).

[24] H. Biard, E. Moreno-Pineda, M. Ruben, E. Bonet, W. Wernsdorfer, and F. Balestro, Increasing the Hilbert space dimension using a single coupled molecular spin, Nat. Commun. **12**, 4443 (2021).

[25] M. Kues, C. Reimer, P. Roztocki, L. R. Cortés, S. Sciara, B. Wetzel, Y. Zhang, A. Cino, S. T. Chu, B. E. Little, D. J. Moss, L. Caspani, J. Azaña, and R. Morandotti, On-chip generation of high-dimensional entangled quantum states and their coherent control, Nature **546**, 622 (2017).

[26] M. Erhard, M. Malik, M. Krenn, and A. Zeilinger, Experimental Greenberger–Horne–Zeilinger entanglement beyond qubits, Nat. Photonics **12**, 759 (2018).

[27] Y.-H. Luo, H.-S. Zhong, M. Erhard, X.-L. Wang, L.-C. Peng, M. Krenn, X. Jiang, L. Li, N.-L. Liu, C.-Y. Lu, A. Zeilinger, and J.-W. Pan, Quantum teleportation in high dimensions, Phys. Rev. Lett. **123**, 070505 (2019).

[28] E. J. Davis, G. Bentsen, L. Homeier, T. Li, and M. H. Schleier-Smith, Photon-mediated spin-exchange dynamics of spin-1 atoms, Phys. Rev. Lett. **122**, 010405 (2019).

[29] Y. Chi, *et al.*, A programmable qudit-based quantum processor, Nat. Commun. **13**, 1166 (2022).

[30] M. S. Blok, V. V. Ramasesh, T. Schuster, K. O'Brien, J. M. Kreikebaum, D. Dahlen, A. Morvan, B. Yoshida, N. Y. Yao, and I. Siddiqi, Quantum information scrambling on a superconducting qutrit processor, Phys. Rev. X **11**, 021010 (2021).

[31] P. Liu, R. Wang, J.-N. Zhang, Y. Zhang, X. Cai, H. Xu, Z. Li, J. Han, X. Li, G. Xue, W. Liu, L. You, Y. Jin, and H. Yu, Performing SU(d) operations and rudimentary algorithms in a superconducting transmon qudit for $d = 3$ and $d = 4$, Phys. Rev. X **13**, 021028 (2023).

[32] E. Champion, Z. Wang, R. W. Parker, and M. S. Blok, Efficient control of a transmon qudit using effective spin-7/2 rotations, Phys. Rev. X **15**, 021096 (2025).

[33] A. Morvan, V. V. Ramasesh, M. S. Blok, J. M. Kreikebaum, K. O'Brien, L. Chen, B. K. Mitchell, R. K. Naik, D. I. Santiago, and I. Siddiqi, Qutrit randomized benchmarking, Phys. Rev. Lett. **126**, 210504 (2021).

[34] M. A. Yurtalan, J. Shi, M. Kononenko, A. Lupascu, and S. Ashhab, Implementation of a Walsh-Hadamard gate in a superconducting qutrit, Phys. Rev. Lett. **125**, 180504 (2020).

[35] M. Kononenko, M. A. Yurtalan, S. Ren, J. Shi, S. Ashhab, and A. Lupascu, Characterization of control in a superconducting qutrit using randomized benchmarking, Phys. Rev. Res. **3**, L042007 (2021).

[36] M. Yurtalan, J. Shi, G. Flatt, and A. Lupascu, Characterization of multilevel dynamics and decoherence in a high-anharmonicity capacitively shunted flux circuit, Phys. Rev. Appl. **16**, 054051 (2021).

[37] K. Luo, W. Huang, Z. Tao, L. Zhang, Y. Zhou, J. Chu, W. Liu, B. Wang, J. Cui, S. Liu, F. Yan, M.-H. Yung, Y. Chen, T. Yan, and D. Yu, Experimental realization of two qutrits gate with tunable coupling in superconducting circuits, Phys. Rev. Lett. **130**, 030603 (2023).

[38] 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).

[39] Z. Chen, *et al.*, Measuring and suppressing quantum state leakage in a superconducting qubit, Phys. Rev. Lett. **116**, 020501 (2016).

[40] 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).

[41] 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).

[42] F. Motzoi and F. K. Wilhelm, Improving frequency selection of driven pulses using derivative-based transition suppression, Phys. Rev. A **88**, 062318 (2013).

[43] L. S. Theis, F. Motzoi, S. Machnes, and F. K. Wilhelm, Counteracting systems of diabaticities using DRAG controls: The status after 10 years, EPL (Europhys. Lett.) **123**, 60001 (2018).

[44] J. M. Chow, L. DiCarlo, J. M. Gambetta, F. Motzoi, L. Frunzio, S. M. Girvin, and R. J. Schoelkopf, Optimized driving of superconducting artificial atoms for improved single-qubit gates, Phys. Rev. A **82**, 040305 (2010).

[45] E. Lucero, J. Kelly, R. C. Bialczak, M. Lenander, M. Mariantoni, M. Neeley, A. D. O'Connell, D. Sank, H. Wang, M. Weides, J. Wenner, T. Yamamoto, A. N. Cleland, and J. M. Martinis, Reduced phase error through optimized control of a superconducting qubit, Phys. Rev. A **82**, 042339 (2010).

[46] L. DiCarlo, J. M. Chow, J. M. Gambetta, L. S. Bishop, B. R. Johnson, D. I. Schuster, J. Majer, A. Blais, L. Frunzio, S. M. Girvin, and R. J. Schoelkopf, Demonstration

of two-qubit algorithms with a superconducting quantum processor, Nature **460**, 240 (2009).
[47] K. X. Wei, E. Magesan, I. Lauer, S. Srinivasan, D. F. Bogorin, S. Carnevale, G. A. Keefe, Y. Kim, D. Klaus, W. Landers, N. Sundaresan, C. Wang, E. J. Zhang, M. Steffen, O. E. Dial, D. C. McKay, and A. Kandala, Hamiltonian engineering with multicolor drives for fast entangling gates and quantum crosstalk cancellation, Phys. Rev. Lett. **129**, 060501 (2022).
[48] 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).
[49] Z. Wang, R. W. Parker, E. Champion, and M. S. Blok, High-E_J/E_C transmon qudits with up to 12 levels, Phys. Rev. Appl. **23**, 034046 (2025).
[50] D. C. McKay, C. J. Wood, S. Sheldon, J. M. Chow, and J. M. Gambetta, Efficient Z gates for quantum computing, Phys. Rev. A **96**, 022330 (2017).
[51] F. Preti, T. Calarco, and F. Motzoi, Continuous quantum gate sets and pulse-class meta-optimization, PRX Quantum **3**, 040311 (2022).
[52] 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).
[53] A. Blais, A. L. Grimsmo, S. M. Girvin, and A. Wallraff, Circuit quantum electrodynamics, Rev. Mod. Phys. **93**, 025005 (2021).
[54] V. Tripathi, N. Goss, A. Vezvaee, L. B. Nguyen, I. Siddiqi, and D. A. Lidar, Qudit dynamical decoupling on a superconducting quantum processor, Phys. Rev. Lett. **134**, 050601 (2025).
[55] F. Motzoi, Controlling quantum information devices, PhD thesis, University of Waterloo (2012).
[56] E. Hyyppä, 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).
[57] M. Deschamps, G. Kervern, D. Massiot, G. Pintacuda, L. Emsley, and P. J. Grandinetti, Superadiabaticity in magnetic resonance, J. Chem. Phys. **129**, 204110 (2008).
[58] L. H. Pedersen, N. M. Møller, and K. Mølmer, Fidelity of quantum operations, Phys. Lett. A **367**, 47 (2007).
[59] B. Chiaro and Y. Zhang, Active leakage cancellation in single qubit gates, ArXiv:2503.14731.
[60] V. Ramakrishna, R. Ober, X. Sun, O. Steuernagel, J. Botina, and H. Rabitz, Explicit generation of unitary transformations in a single atom or molecule, Phys. Rev. A **61**, 032106 (2000).
[61] G. K. Brennen, D. P. O'Leary, and S. S. Bullock, Criteria for exact qudit universality, Phys. Rev. A **71**, 052318 (2005).
[62] 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).
[63] O. Astafiev, Y. A. Pashkin, Y. Nakamura, T. Yamamoto, and J. S. Tsai, Quantum noise in the Josephson charge qubit, Phys. Rev. Lett. **93**, 267007 (2004).
[64] A. B. Zorin, F.-J. Ahlers, J. Niemeyer, T. Weimann, H. Wolf, V. A. Krupenin, and S. V. Lotkhov, Background charge noise in metallic single-electron tunneling devices, Phys. Rev. B **53**, 13682 (1996).
[65] B. G. Christensen, C. D. Wilen, A. Opremcak, J. Nelson, F. Schlenker, C. H. Zimonick, L. Faoro, L. B. Ioffe, Y. J. Rosen, J. L. DuBois, B. L. T. Plourde, and R. McDermott, Anomalous charge noise in superconducting qubits, Phys. Rev. B **100**, 140503 (2019).
[66] 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).
[67] 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).
[68] V. Braginsky, V. Ilchenko, and K. Bagdassarov, Experimental observation of fundamental microwave absorption in high-quality dielectric crystals, Phys. Lett. A **120**, 300 (1987).
[69] C. Wang, C. Axline, Y. Y. Gao, T. Brecht, Y. Chu, L. Frunzio, M. H. Devoret, and R. J. Schoelkopf, Surface participation and dielectric loss in superconducting qubits, Appl. Phys. Lett. **107**, 162601 (2015).
[70] A. P. Place, L. V. Rodgers, P. Mundada, B. M. Smitham, M. Fitzpatrick, Z. Leng, A. Premkumar, J. Bryon, A. Vrajitoarea, S. Sussman, *et al.*, New material platform for superconducting transmon qubits with coherence times exceeding 0.3 ms, Nat. Commun. **12**, 1779 (2021).
[71] M. Tuokkola, Y. Sunada, H. Kivijärvi, L. Grönberg, J.-P. Kaikkonen, V. Vesterinen, J. Govenius, and M. Möttönen, Methods to achieve near-millisecond energy relaxation and dephasing times for a superconducting transmon qubit, ArXiv:2407.18778.
[72] C. Wang, *et al.*, Towards practical quantum computers: Transmon qubit with a lifetime approaching 0.5 ms, npj Quantum Inf. **8**, 3 (2022).
[73] M. Bal, A. A. Murthy, S. Zhu, F. Crisa, X. You, Z. Huang, T. Roy, J. Lee, D. V. Zanten, R. Pilipenko, *et al.*, Systematic improvements in transmon qubit coherence enabled by niobium surface encapsulation, npj Quantum Inf. **10**, 43 (2024).
[74] S. Kono, J. Pan, M. Chegnizadeh, X. Wang, A. Youssefi, M. Scigliuzzo, and T. J. Kippenberg, Mechanically induced correlated errors on superconducting qubits with relaxation times exceeding 0.4 ms, Nat. Commun. **15**, 3950 (2024).