Operation and intrinsic error budget of a two-qubit cross-resonance gate

Vinay Tripathi,[1] Mostafa Khezri,[1,2] and Alexander N. Korotkov[1,*]

[1]*Department of Electrical and Computer Engineering, University of California, Riverside, California 92521, USA*
[2]*Department of Electrical Engineering, University of Southern California, Los Angeles, California 90089, USA*

We analyze analytically, semianalytically, and numerically the operation of a cross-resonance (CR) gate for superconducting qubits (transmons). We find that a relatively simple semianalytical method gives accurate results for the controlled-NOT (CNOT) -equivalent gate duration and compensating single-qubit rotations. It also allows us to minimize the CNOT gate duration over the amplitude of the applied microwave drive and find the dependence on the detuning between the qubits. However, full numerical simulations are needed to calculate the intrinsic fidelity of the CR gate. We decompose numerical infidelity into contributions from various physical mechanisms, thus finding the intrinsic error budget. In particular, at small drive amplitudes, the CR gate fidelity is limited by imperfections of the target-qubit rotations, while at large amplitudes it is limited by leakage. The gate duration and fidelity are analyzed numerically as functions of the detuning between qubits, their coupling, drive frequency, relative duration of pulse ramps, and microwave crosstalk. The effect of the echo sequence is also analyzed numerically. Our results show that the CR gate can provide intrinsic infidelity of less than 10^{-3} when a simple pulse shape is used.

I. INTRODUCTION

Two decades have passed since the first superconducting qubit was created [1], and today superconducting quantum computing is a well-developed field with various types and uses of qubits [2–11]. Currently, the most popular type of superconducting qubits is the transmon [12] (including its Xmon and gmon modifications [13,14]), though other types of qubits (see, e.g., [15–17]) are also of interest. Besides sufficiently good coherence of the qubits, quantum computing applications need high-fidelity gates forming a universal set [18]. While single-qubit gates are already considered to be simple and accurate, current fidelity of two-qubit gates also exceeds 99% [5,19,20].

One of the high-fidelity two-qubit gates used for superconducting qubits is the cross-resonance (CR) gate [21,22]. In this gate, two frequency-detuned qubits have a fixed coupling (usually via a resonator) and one of them (called control qubit) is driven by a microwave with frequency of the other one (target qubit). This induces Rabi oscillations of the target qubit, whose frequency depends on the state ($|0\rangle$ or $|1\rangle$) of the control qubit, thus entangling the two qubits and providing a natural way to realize controlled-NOT (CNOT) operation. Since the CR gate uses only microwave control, it permits the use of single-junction transmons, thus avoiding sensitivity to flux noise. However, the drawback is a relatively long gate duration compared with the gates based on tune-detune operation [19].

The idea of the CR gate was proposed in Ref. [23] and then experimentally implemented for flux qubits in Ref. [24] under the name of selective darkening (the difference compared with a simple CR gate is an additional active cancellation pulse applied to the target qubit). The CR terminology was

introduced in the theoretical paper [21] and the first experiment under this name was realized with capacitively shunted flux qubits in Ref. [22] with a fidelity of 81%. The CR gate was applied to transmons in Ref. [25], with a resulting fidelity of 95%. Since then the CR gate has been used in numerous experiments by several groups (e.g., [3,20,26–30]), with a gradual increase of maximum fidelity. An important improvement of the CR operation was achieved by using the echo sequence [26,28], which not only increased the fidelity but also allowed protocols that avoid compensating one-qubit rotations in implementing the CNOT gate. The CR gate with a duration of 160 ns and fidelity of 99.1% reported in Ref. [20] was achieved by using both the echo sequence and active cancellation pulses applied to the target qubit.

In spite of extensive experimental use of the CR gate, its theoretical analysis has been rather limited. Besides the initial papers [21,23] laying out the main idea, the CR gate was analyzed in Ref. [31] with an account of the next level, briefly mentioned in Ref. [32], and analyzed in detail in the recent paper [33]. There were also numerical studies [34,35] and related papers [36,37].

In this paper we analyze the operation of the basic CR gate for transmons (using a simple pulse shape without the echo sequence, which is considered only in the Appendix) at three levels of complexity and accuracy: analytical, semianalytical, and numerical. Some of the goals of our analysis are similar to those of Ref. [33]; however, the approach is very different. After discussing the ideal theory of CR operation for transmons (using the Duffing oscillator model), we develop the next-order approximation, somewhat similar to that of Ref. [33] (which still does not work well, as follows from comparison with full numerics), and then develop the semianalytical approach, based on the numerical solution of a simple one-qubit time-independent Schrödinger equation.

---

*Present address: Google Inc., Venice, CA 90291, USA.

The semianalytical approach gives very accurate results (compared with full numerics) for the CNOT-equivalent gate duration and compensating single-qubit rotations. In particular, it can be used to find the shortest CNOT gate duration, corresponding optimal drive amplitude, and their dependence on detuning between the qubits. However, the semianalytical approach cannot be used for finding the intrinsic fidelity of the gate (neglecting decoherence), for which we use a full numerical simulation. Our numerical simulation includes 7×5 levels in the qubits (we replace qubit coupling via a resonator with an equivalent direct coupling) and is based on Magnus expansion [38] for the evolution matrices. We use a simple pulse shape with cosine-shaped ramps and a flat middle part.

After calculating the gate infidelity, we numerically decompose it into the contributions from various physical mechanisms, thus finding the intrinsic error budget. We show that at small drive amplitudes, the error is dominated by imperfections of the unitary operation within the computational subspace. In contrast, at large drive amplitudes (which correspond to reasonably short gate durations) the infidelity is dominated by leakage. Particular leakage channels depend on detuning between the qubits. In this regime, the analytical estimate for the leakage probability agrees reasonably well with the results for the gate infidelity.

Using numerical results for the CNOT gate duration and fidelity, we analyze their dependence on various parameters, including detuning between qubits, their coupling, drive frequency, smoothness (relative duration of the pulse ramps), and microwave crosstalk. In the Appendix we also analyze the effect of the echo sequence. Our results show that the CR gate can provide an intrinsic infidelity of about 10^{-3} (and even lower, comparable to 10^{-4}) with a simple pulse shape.

The paper is organized as follows. In Sec. II we discuss the system and its Hamiltonian. In Sec. III we first consider the ideal operation of the CR gate, then derive the next-order analytics, and then develop the semianalytical approach. The numerical method is discussed in Sec. IV. Numerical results for the CNOT-equivalent gate duration and compensating single-qubit rotations are discussed in Sec. V. Then in Sec. VI we analyze the error budget for the CNOT gate intrinsic infidelity. In Sec. VII, we discuss the dependence of CNOT duration and infidelity on parameters. We summarize in Sec. VIII. In the Appendix, we analyze the echo-CR gate operation.

II. SYSTEM AND HAMILTONIAN

In the CR gate, the control and target qubits (with frequencies ω_c and ω_t, respectively) are usually detuned by 50–300 MHz and are permanently coupled via a resonator. However, in this paper, for simplicity, we will consider a direct qubit-qubit coupling g (Fig. 1) since the usual analysis of the CR gate [32,33] also reduces the coupling via a resonator to an effective direct coupling. For the CR operation, the control qubit is rf driven at the frequency of the target qubit, $\omega_d \approx \omega_t$. This produces an effective drive (*x* rotation) of the target qubit, with the strength depending on the state of the control qubit. Such a process can be naturally used to realize the CNOT gate by calibrating the target-qubit rotation angle difference
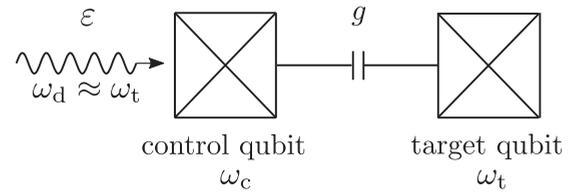

FIG. 1. Schematic of the CR gate. Detuned control and target qubits (transmons with frequencies ω_c and ω_t) have coupling g and the control qubit is microwave driven at the frequency of the target qubit, $\omega_d \approx \omega_t$. The microwave drive amplitude is ε.

(between rotations for the control-qubit states $|0\rangle$ and $|1\rangle$) to be equal to π and somehow compensating the target-qubit rotation for the control-qubit state $|0\rangle$. This compensation can be done, for example, by using the echo sequence [20,26,28] or active cancellation [20,24]; however, in this paper we will assume that the compensation is done afterward [21,22] by applying single-qubit rotations (the echo sequence is considered only in the Appendix). We intentionally consider the simplest case in order to focus on developing a good understanding of the basic operation of the CR gate.

The operating principle of the CR gate can be understood classically, by replacing qubits with classical oscillators (Fig. 2). Since the drive is off-resonance with the control oscillator ($\omega_d \not\approx \omega_c$), it will produce very small forced oscillations at the drive frequency ω_d. However, since the target oscillator is on-resonance with this frequency ($\omega_d \approx \omega_t$), it will still get excited via the coupling g with the control oscillator. Note that if the control oscillator is linear, then its own state (its oscillation with frequency ω_c) does not matter because of linearity. However, if the control oscillator is nonlinear, then its effective frequency depends on its own state (i.e., amplitude of ω_c oscillations); therefore, the amplitude of the small forced oscillations of the control oscillator and consequently the excitation rate of the target oscillator will depend on the control-oscillator state. This simple classical picture explains the basic physical mechanism of the CR gate operation for transmons, which are slightly nonlinear oscillators. It also explains why the CR gate speed depends on nonlinearity of the control qubit and practically does not depend on the target-qubit nonlinearity.
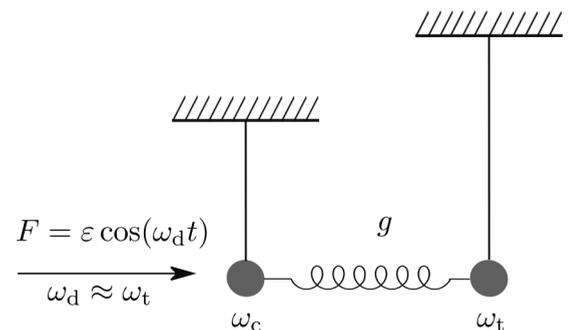

FIG. 2. Classical CR gate counterpart: two coupled nonlinear oscillators, with one oscillator driven by a periodic force F on-resonance with the other oscillator.

For quantum analysis of the CR gate (Fig. 1), let us start with the rotating-frame Hamiltonian (the rotating frame is based on the drive frequency ω_d)

$$H = H_{qb} + H_g + H_\varepsilon, \qquad (1)$$

where H_{qb} describes two uncoupled transmon qubits, H_g describes their coupling, and H_ε describes the microwave drive on the control qubit. The uncoupled-qubit part can be written as

$$H_{qb} = \sum_{n,m} \left(E_n^{(c)} + E_m^{(t)} \right) |n,m\rangle\langle n,m|, \qquad (2)$$

$$E_n^{(c)} = E_n^{(c,lf)} - n\omega_d, \quad E_m^{(t)} = E_m^{(t,lf)} - m\omega_d, \qquad (3)$$

where in the notation $|n,m\rangle$ the control-qubit state is at the left ($n = 0, 1, 2, \ldots$) and the target-qubit state is at the right ($m = 0, 1, 2, \ldots$), the control-qubit energies $E_n^{(c)}$ in the rotating frame are related to the laboratory-frame energies $E_n^{(c,lf)}$ via the drive frequency ω_d, and there is a similar relation for the target-qubit energies $E_m^{(t)}$. We set $E_0^{(c)} = E_0^{(t)} = 0$. For the energies $E_n^{(c)}$ and $E_m^{(t)}$, in this paper we use the Duffing (Kerr) oscillator approximation

$$E_n^{(c)} = n(\Delta + \delta) - \frac{n(n-1)}{2}\eta_c, \qquad (4)$$

$$E_m^{(t)} = m\delta - \frac{m(m-1)}{2}\eta_t, \qquad (5)$$

$$\Delta \equiv \omega_c - \omega_t, \quad \delta \equiv \omega_t - \omega_d \approx 0, \qquad (6)$$

where Δ is the detuning between the qubits, while η_c and η_t are anharmonicities of the control and target qubits, respectively (for transmons $\eta_c > 0$ and $\eta_t > 0$). A small mismatch δ between the drive frequency ω_d and the *bare* frequency ω_t of the target qubit can be used, e.g., to make the drive exactly resonant with the hybridized target qubit for the control-qubit states $|0\rangle$ or $|1\rangle$ (or in between). Note that $\Delta + \delta = \omega_c - \omega_d$. Instead of the approximation (4)–(6), it is possible to use numerical results for the transmon energies or at least the improved approximation [39,40]. However, we prefer the simple approximation for easier comparison with the previous theoretical analyses of the CR gate.

The qubit-qubit coupling Hamiltonian H_g in general couples all pairs of the bare states $|n,m\rangle$ and $|n',m'\rangle$. However, in this paper we use the simplest (traditional) approximation for transmons by keeping only the excitation-preserving terms, i.e., applying the rotating-wave approximation (RWA), and using the matrix elements for linear oscillators,

$$H_g = \sum_{n,m} g\sqrt{nm}\,|n,m-1\rangle\langle n-1,m| + \text{H.c.} \qquad (7)$$

additionally assuming (without loss of generality) that the coupling constant g is real. Similarly, we use the RWA linear-oscillator matrix elements for the drive Hamiltonian (in the rotating frame)

$$H_\varepsilon = \sum_{n,m} \varepsilon(t)\sqrt{n}\,|n,m\rangle\langle n-1,m| + \text{H.c.}, \qquad (8)$$

where the complex amplitude ε of the drive depends on time, so $\varepsilon(t)$ is the pulse shape of the CR gate, with $\varepsilon(t) = 0$ before and after the gate. Instead of Hamiltonians (7) and (8), it is
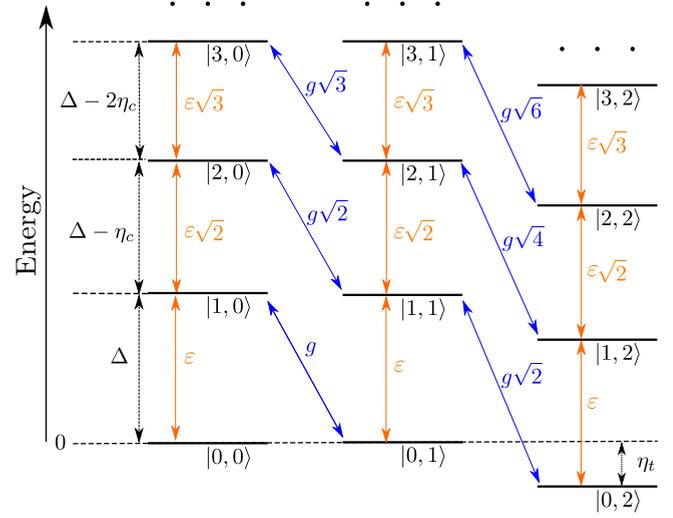

FIG. 3. Diagram of bare energy levels for the CR gate. Each vertical ladder is for all control-qubit states, with a fixed state of the target qubit (we use the notation $|$control, target$\rangle$). Slanted blue lines illustrate coupling between the bare levels due to the qubit-qubit coupling H_g; orange lines are due to the drive Hamiltonian H_ε. On this diagram we assumed $\delta = 0$ (a nonzero δ would produce an energy shift between the ladders; also, in this case, Δ should be replaced with $\Delta + \delta$). Control-qubit states above $|3\rangle$ and target-qubit states above $|2\rangle$ are not shown.

possible to use improved perturbative Hamiltonians [39,40] or numerical matrix elements for transmons, but in this paper we use the simple traditional approximation. Here we do not consider the microwave crosstalk [20,22,25,33]; however, it will be added in Sec. VII.

It is convenient to draw a diagram (Fig. 3) of bare levels $|n,m\rangle$ in which the left ladder of levels corresponds to the target-qubit state $|0\rangle$ ($m = 0$), the next ladder corresponds to the target-qubit state $|1\rangle$, then $|2\rangle$, and so on. Note that for $\delta = 0$, the left two ladders are at exactly equal energies. In Fig. 3 the coupling H_g is represented by slanted blue arrows and the drive H_ε corresponds to vertical orange arrows. For clarity, in Fig. 3 we show the case $\Delta > 2\eta_c$, while in experiments usually $0 < \Delta < \eta_c$. In such a case, all ladders turn down after the states $|1,m\rangle$ and the diagram becomes visually complicated, so for gaining intuition it is easier to use the case of Fig. 3.

Besides the bare states $|n,m\rangle$, we will also use the eigenstates of the Hamiltonian $H_{qb} + H_g$ (without the drive), which we denote by an overline: $\overline{|n,m\rangle}$. The coupling H_g affects the qubit frequencies, so instead of the bare frequency ω_t of the target qubit, we have two eigenfrequencies ω_t^{c0} and ω_t^{c1}, depending on the control-qubit state ($|0\rangle$ and $|1\rangle$, respectively). They can be calculated as

$$\omega_t^{c0} = E_{\overline{|0,1\rangle}}^{(lf)} - E_{\overline{|0,0\rangle}}^{(lf)}, \quad \omega_t^{c1} = E_{\overline{|1,1\rangle}}^{(lf)} - E_{\overline{|1,0\rangle}}^{(lf)}, \qquad (9)$$

where $E_{\overline{|n,m\rangle}}^{(lf)}$ is the laboratory-frame eigenenergy of the state $\overline{|n,m\rangle}$. We will call zz coupling the difference between these frequencies,

$$\omega_{zz} \equiv \omega_t^{c1} - \omega_t^{c0} = E_{\overline{|11\rangle}} + E_{\overline{|00\rangle}} - E_{\overline{|01\rangle}} - E_{\overline{|10\rangle}}, \qquad (10)$$

where this combination of eigenenergies is the same in the laboratory and rotating frames. The zz coupling is mainly due to the repulsion of the energy level $|11\rangle$ from the levels $|02\rangle$ and $|20\rangle$, which gives the approximate value

$$\omega_{zz} \approx \frac{2g^2}{\Delta + \eta_t} - \frac{2g^2}{\Delta - \eta_c}. \tag{11}$$

From Eq. (10) we see that the zz coupling can be also defined as $\omega_{zz} = \omega_c^{t1} - \omega_c^{t0}$, where ω_c^{t0} and ω_c^{t1} are the eigenfrequencies of the control qubit for the target-qubit states $|0\rangle$ and $|1\rangle$, respectively. Nonzero ω_{zz} will be important for numerical results; however, it will be neglected for analytical and semi-analytical results in the next section; in particular, we will not distinguish between ω_t^{c0}, ω_t^{c1}, and ω_t.

III. ANALYTICAL AND SEMIANALYTICAL ANALYSIS

A. Ideal CR gate operation

There is no drive, $\varepsilon = 0$, before and after the CR gate operation. Therefore, the initial and final two-qubit states should be considered in the eigenbasis $\overline{|n, m\rangle}$ of the Hamiltonian $H_{qb} + H_g$. The drive Hamiltonian H_ε couples these eigenstates, providing an evolution used in the CR gate.

As follows from Fig. 3, in the rotating frame based on the drive frequency ω_d, there is a near-resonance condition between states $|n, 0\rangle$ and $|n, 1\rangle$, which leads to a near-resonance between eigenstates $\overline{|n, 0\rangle}$ and $\overline{|n, 1\rangle}$, while other pairs of states are off-resonance. Therefore, as long as the perturbation produced by H_ε is small enough, it effectively couples only states $\overline{|n, 0\rangle}$ and $\overline{|n, 1\rangle}$, and for the ideal effective Hamiltonian H_{CR}^{ideal} of the CR gate we can write

$$H_{CR}^{ideal} - (H_{qb} + H_g) = (\tilde{\varepsilon}_0 \overline{|0, 1\rangle}\,\overline{\langle 0, 0|} + \tilde{\varepsilon}_1 \overline{|1, 1\rangle}\,\overline{\langle 1, 0|}$$
$$+ \tilde{\varepsilon}_2 \overline{|2, 1\rangle}\,\overline{\langle 2, 0|} + \cdots) + \text{H.c.}, \tag{12}$$

where $\tilde{\varepsilon}_0$ is the amplitude of the *effective drive on the target qubit* when the control-qubit state is $|0\rangle$, $\tilde{\varepsilon}_1$ is the effective drive amplitude for the control-qubit state $|1\rangle$, etc. (for small g there is almost no difference between the effective drive in the bare basis or eigenbasis). The effective drive amplitudes $\tilde{\varepsilon}_n$ depend on the actual drive amplitude ε (in the linear approximation being proportional to ε).

Note that if we are interested only in the states $|0\rangle$ and $|1\rangle$ of the control qubit, then in the terminology of Refs. [22,25,29,33] the effective Hamiltonian (12) can be written as

$$\frac{\tilde{\varepsilon}_0 - \tilde{\varepsilon}_1}{2} Z_c X_t + \frac{\tilde{\varepsilon}_0 + \tilde{\varepsilon}_1}{2} I_c X_t,$$

where the Pauli operators Z_c and I_c act on the control qubit and the operator X_t acts on the target qubit (here we assume real $\tilde{\varepsilon}_0$ and $\tilde{\varepsilon}_1$; otherwise, we also need Y_t). The CNOT gate can be realized with this effective interaction by applying the drive pulse with duration τ_p, satisfying the condition

$$\int_0^{\tau_p} [2\tilde{\varepsilon}_1(t) - 2\tilde{\varepsilon}_0(t)]dt = \pi \,(\text{mod}\,2\pi), \tag{13}$$

complemented with two one-qubit rotations. The additional x rotation of the target qubit by the angle $-\int_0^{\tau_p} 2\tilde{\varepsilon}_0(t)dt$

compensates the target-qubit rotation for the control-qubit state $|0\rangle$, also providing x rotation by angle π for the control-qubit state $|1\rangle$. Besides the x rotation of the target qubit, the control qubit should be z rotated by the angle $\pi/2$ (relative to the rotating frame of the control qubit). This is needed because the x rotation of the target qubit by angle π produces the operation $-iX$ instead of the desired (for CNOT) operation X, thus requiring an additional phase factor i for the control-qubit state $|1\rangle$ (the same factor exists in a one-qubit X gate, but it is not important since it is an overall phase, in contrast to the phase difference in a controlled two-qubit operation). Note that the factors of 2 in Eq. (13) are needed because the Rabi frequency is twice larger than the drive matrix element in the Hamiltonian.

The effective drive amplitudes $\tilde{\varepsilon}_n$ in Eq. (12) can be easily found (in the ideal lowest-order case) by comparing Eqs. (1) and (12), which gives

$$\tilde{\varepsilon}_n = \overline{\langle n, 1|}H_\varepsilon \overline{|n, 0\rangle}. \tag{14}$$

To calculate $\tilde{\varepsilon}_n$ to the lowest order, let us assume that $\delta = 0$, i.e., the drive is resonant with the bare target qubit (the difference between bare and eigenfrequencies is not important for these approximate calculations). Then using $\overline{|0, 0\rangle} = |0, 0\rangle$ (see Fig. 3) and the first-order approximation $\overline{|0, 1\rangle} = |0, 1\rangle - (g/\Delta)|1, 0\rangle$ (normalization correction is of the second order), we find the linear approximation

$$\tilde{\varepsilon}_0 = -\frac{g}{\Delta}\varepsilon. \tag{15}$$

Similarly, using the first-order approximations $\overline{|1, 0\rangle} = |1, 0\rangle + (g/\Delta)|0, 1\rangle$ and $\overline{|1, 1\rangle} = |1, 1\rangle - [\sqrt{2}g/(\Delta - \eta_c)]|2, 0\rangle + [\sqrt{3}g/(\Delta + \eta_t)]|0, 2\rangle$ (see Fig. 3), we obtain the approximation

$$\tilde{\varepsilon}_1 = \frac{g}{\Delta}\varepsilon - \frac{\sqrt{2}g}{\Delta - \eta_c}\sqrt{2}\varepsilon = -\frac{g}{\Delta}\frac{\Delta + \eta_c}{\Delta - \eta_c}\varepsilon. \tag{16}$$

Also similarly, using approximations $\overline{|2, 0\rangle} = |2, 0\rangle + [\sqrt{2}g/(\Delta - \eta_c)]|1, 1\rangle$ and $\overline{|2, 1\rangle} = |2, 1\rangle - [\sqrt{3}g/(\Delta - 2\eta_c)]|3, 0\rangle + [2g/(\Delta - \eta_c + \eta_t)]|1, 2\rangle$, we find

$$\tilde{\varepsilon}_2 = \frac{\sqrt{2}g}{\Delta - \eta_c}\sqrt{2}\varepsilon - \frac{\sqrt{3}g}{\Delta - 2\eta_c}\sqrt{3}\varepsilon$$
$$= -\frac{g(\Delta + \eta_c)}{(\Delta - \eta_c)(\Delta - 2\eta_c)}\varepsilon, \tag{17}$$

and for an arbitrary control-qubit state $|n\rangle$, within the model (1)–(8) we obtain

$$\tilde{\varepsilon}_n = \frac{ng\varepsilon}{\Delta - (n-1)\eta_c} - \frac{(n+1)g\varepsilon}{\Delta - n\eta_c}$$
$$= -\frac{g(\Delta + \eta_c)}{[\Delta - (n-1)\eta_c](\Delta - n\eta_c)}\varepsilon. \tag{18}$$

Since the target-qubit rotation for the control-qubit state $|0\rangle$ is usually compensated, most important are the differences of effective drive amplitudes from $\tilde{\varepsilon}_0$, e.g.,

$$\tilde{\varepsilon}_1 - \tilde{\varepsilon}_0 = \frac{2g\eta_c}{\Delta(\eta_c - \Delta)}\varepsilon, \tag{19}$$

$$\tilde{\varepsilon}_2 - \tilde{\varepsilon}_0 = \frac{2g\eta_c(\eta_c - 2\Delta)}{\Delta(\eta_c - \Delta)(2\eta_c - \Delta)}\varepsilon. \tag{20}$$

Note that these formulas depend on anharmonicity η_c of the control qubit but do not depend on the target-qubit anharmonicity η_t. Also, for $\eta_c = 0$ we have $\tilde{\varepsilon}_n = \tilde{\varepsilon}_0 = -(g/\Delta)\varepsilon$. These properties are in agreement with the classical description of the CR gate operation discussed in Sec. II.

Language of virtual-state transitions

Instead of using Eq. (14), we can find the effective drive amplitudes $\tilde{\varepsilon}_n$ (still to the first order) using the ideology of transitions via a virtual state. As seen in Fig. 3, it is possible to go from the state $|0, 0\rangle$ to the resonant state $|0, 1\rangle$ in two jumps, $|0, 0\rangle \rightarrow |1, 0\rangle \rightarrow |0, 1\rangle$, which have transition amplitudes (matrix elements) ε and g, with the intermediate state separated by the energy difference Δ. Therefore, the amplitude of this transition (effective coupling between states $|0, 0\rangle$ and $|0, 1\rangle$) is

$$\tilde{\varepsilon}_0 = \varepsilon \frac{-1}{\Delta} g, \tag{21}$$

which coincides with Eq. (15).

For the transition between states $|1, 0\rangle$ and $|1, 1\rangle$, there are two two-jump paths: via the state $|2, 0\rangle$ (which is higher in energy by $\Delta - \eta_c$) and via $|0, 1\rangle$ (which is lower in energy by Δ). Adding these two amplitudes, we obtain

$$\tilde{\varepsilon}_1 = \sqrt{2}\varepsilon \frac{-1}{\Delta - \eta_c} \sqrt{2}g + g \frac{-1}{-\Delta} \varepsilon, \tag{22}$$

which coincides with Eq. (16). Similarly, adding the amplitudes for the paths $|n, 0\rangle \rightarrow |n + 1, 0\rangle \rightarrow |n, 1\rangle$ and $|n, 0\rangle \rightarrow |n - 1, 1\rangle \rightarrow |n, 1\rangle$, we obtain

$$\tilde{\varepsilon}_n = -\frac{\sqrt{n+1}\varepsilon\sqrt{n+1}g}{\Delta - n\eta_c} + \frac{\sqrt{n}g\sqrt{n}\varepsilon}{\Delta - (n-1)\eta_c}, \tag{23}$$

which coincides with Eq. (18).

B. Next-order analytics

Numerical results for the effective drive amplitudes $\tilde{\varepsilon}_n$ (discussed later) show that $\tilde{\varepsilon}_n$ is proportional to the actual drive amplitude ε [as expected from Eq. (18)] only in some range of ε values. A minor deviation from the linearity at very small ε (discussed later) is due to the dependence of the target-qubit frequency on the control-qubit state [see Eq. (10)]. The deviation from linearity at large ε is much more important for practice since it makes it impossible to shorten the CNOT gate duration beyond some value by simply increasing the drive amplitude.

In order to understand the reason for the deviation from linearity at large ε, in this section we develop the next-order analytics for $\tilde{\varepsilon}_0$ and $\tilde{\varepsilon}_1$, which gives corrections compared with Eqs. (15) and (16). Note that a similar next-order analytics has been developed in Ref. [33], though in a very different way (after a misprint correction, the result of Ref. [33] coincides with our result).

The simple analytics (15)–(18) has been obtained from Eq. (14), which treats the drive Hamiltonian H_ε as a small perturbation. However, for a large drive amplitude ε, the eigenbasis of $H_{qb} + H_g$ is no longer the appropriate eigenbasis; instead, $H_{qb} + H_\varepsilon$ is the main Hamiltonian, while H_g is the perturbation. Note that in the linear approximation, the same

$\tilde{\varepsilon}_n$ as in Eq. (14) can be obtained by exchanging the roles of H_g and H_ε, i.e., by using

$$\tilde{\varepsilon}_n = \overline{{}_\varepsilon\langle n, 1|H_g|n, 0\rangle_\varepsilon}, \tag{24}$$

where $\overline{|n, m\rangle_\varepsilon}$ denotes the eigenstate of $H_{qb} + H_\varepsilon$. This equivalence is clear from the above-discussed approach of virtual-state transitions, which treats H_g and H_ε on equal footing.

For a large ε, Eq. (24) is more appropriate than Eq. (14) to calculate $\tilde{\varepsilon}_n$. Even though the initial and final states should still be treated in the eigenbasis of $H_{qb} + H_g$, during the front and rear ramps of the microwave pulse the appropriate eigenbases essentially transform into each other, leading to Eq. (24). While we do not have a rigorous justification of the approximation (24) (only a general understanding in the spirit of the adiabatic theorem), numerical results confirm its good accuracy.

Since for the eigenstates $\overline{|n, m\rangle_\varepsilon}$ used in Eq. (24) the ladders in Fig. 3 are uncoupled, we can write

$$\overline{|n, 0\rangle_\varepsilon} = \overline{|n\rangle_\varepsilon}|0\rangle_t, \quad \overline{|n, 1\rangle_\varepsilon} = \overline{|n\rangle_\varepsilon}|1\rangle_t, \tag{25}$$

where $\overline{|n\rangle_\varepsilon}$ are the control-qubit eigenstates, which account for the drive. They satisfy the Schrödinger equation

$$H^{(c)}_{qb+\varepsilon}\overline{|n\rangle_\varepsilon} = E_{\overline{|n\rangle_\varepsilon}}\overline{|n\rangle_\varepsilon} \tag{26}$$

with the Hamiltonian for only the control qubit,

$$H^{(c)}_{qb+\varepsilon} = \sum_n E^{(c)}_n |n\rangle\langle n| + \sqrt{n}(\varepsilon|n\rangle\langle n-1| + \varepsilon^*|n-1\rangle\langle n|). \tag{27}$$

Then solving this Schrödinger equation and finding the eigenstates

$$\overline{|n\rangle_\varepsilon} = \sum_k c^{(n)}_k |k\rangle, \tag{28}$$

we find the effective drive amplitudes $\tilde{\varepsilon}_n$ from Eq. (24) as (see Fig. 3)

$$\tilde{\varepsilon}_n = g \sum_k \sqrt{k} c^{(n)}_k \left(c^{(n)}_{k-1}\right)^*. \tag{29}$$

Let us use this approach to find $\tilde{\varepsilon}_0$ up to the order ε^3 [instead of ε^1 in the linear approximation (15)], treating H_ε as a perturbation of H_{qb}. The eigenstate $\overline{|0\rangle_\varepsilon}$ of the control qubit can be written as

$$\overline{|0\rangle_\varepsilon} = \frac{|0\rangle + \alpha|1\rangle + \beta|2\rangle + \gamma|3\rangle + \cdots}{\mathcal{N}}, \tag{30}$$

where $\mathcal{N}$ is a normalization. Substituting this form into the Schrödinger equation (26) and equating the coefficients for the basis states $|0\rangle$, $|1\rangle$, and $|2\rangle$, we obtain

$$\varepsilon\alpha = E, \tag{31}$$

$$E_1\alpha + \sqrt{2}\varepsilon\beta + \varepsilon = E\alpha, \tag{32}$$

$$E_2\beta + \sqrt{2}\varepsilon\alpha + \sqrt{3}\varepsilon\gamma = E\beta, \tag{33}$$

where for brevity $E = E_{\overline{|0\rangle_\varepsilon}}$, $E_n = E^{(c)}_n$, we used $E_0 = 0$ [as in Eq. (4)], and we also assumed that ε is real.

To the lowest order, assuming small ε (therefore $\gamma \ll \beta \ll \alpha$ and $E \approx 0$), we crudely find

$$\alpha \approx \frac{-\varepsilon}{E_1}, \quad \beta \approx \frac{-\sqrt{2}\varepsilon\alpha}{E_2} \approx \frac{\sqrt{2}\varepsilon^2}{E_1 E_2}, \quad E \approx \frac{-\varepsilon^2}{E_1}. \quad (34)$$

Using these values for β and E in Eq. (32), we obtain a better approximation (up to ε^3) for α:

$$\alpha \approx -\frac{\varepsilon(1 + 2\varepsilon^2/E_1 E_2)}{E_1 + \varepsilon^2/E_1} \approx -\frac{\varepsilon}{E_1}\left(1 + \frac{2\varepsilon^2}{E_1 E_2} - \frac{\varepsilon^2}{E_1^2}\right). \quad (35)$$

To find $\tilde{\varepsilon}_0$ with accuracy up to ε^3, we need α with accuracy up to ε^3, β with accuracy up to ε^2, and $\mathcal{N}$ with accuracy up to ε^2, while γ is not needed [see Eq. (29)]. Thus, we use Eq. (35) for α, Eq. (34) for β, and $\mathcal{N} \approx 1 + (\varepsilon/E_1)^2/2$ to obtain

$$\overline{|0\rangle_\varepsilon} \approx \left(1 - \frac{\varepsilon^2}{2E_1^2}\right)|0\rangle - \frac{\varepsilon}{E_1}\left(1 + \frac{2\varepsilon^2}{E_1 E_2} - \frac{3\varepsilon^2}{2E_1^2}\right)|1\rangle$$
$$+ \frac{\sqrt{2}\varepsilon^2}{E_1 E_2}|2\rangle. \quad (36)$$

The energy of state $\overline{|0\rangle_\varepsilon}$ (not needed for this derivation but needed later) is

$$E_{\overline{|0\rangle_\varepsilon}} = \varepsilon\alpha \approx -\frac{\varepsilon^2}{E_1}\left(1 + \frac{2\varepsilon^2}{E_1 E_2} - \frac{\varepsilon^2}{E_1^2}\right). \quad (37)$$

Finally, using Eqs. (29) and (36), we obtain

$$\tilde{\varepsilon}_0 = -g\frac{\varepsilon}{E_1}\left(1 - \frac{2\varepsilon^2}{E_1^2} + \frac{4\varepsilon^2}{E_1 E_2}\right) \quad (38)$$

with accuracy up to ε^3. Note that $E_1 = E_1^{(c)} = \Delta + \delta$, $E_2 = E_2^{(c)} = 2(\Delta + \delta) - \eta_c$, and we can neglect δ (i.e., use $\delta = 0$). For a complex ε, we need to replace ε^2 in parentheses with $|\varepsilon|^2$.

Calculation of $\tilde{\varepsilon}_1$ up to the order ε^3 is similar and requires finding $\overline{|1\rangle_\varepsilon}$. Note that the calculations are easier if the energies are counted from E_1, because then the eigenenergy E in equations similar to Eqs. (31)–(33) is small. The calculations give

$$\overline{|1\rangle_\varepsilon} \approx \left(1 - \frac{\varepsilon^2}{2E_{01}^2} - \frac{\varepsilon^2}{E_{21}^2}\right)|1\rangle$$
$$- \frac{\sqrt{2}\varepsilon}{E_{21}}\left(1 + \frac{3\varepsilon^2}{E_{21}E_{31}} - \frac{\varepsilon^2}{E_{01}E_{21}} - \frac{3\varepsilon^2}{E_{21}^2} - \frac{\varepsilon^2}{2E_{01}^2}\right)|2\rangle$$
$$- \frac{\varepsilon}{E_{01}}\left(1 - \frac{3\varepsilon^2}{2E_{01}^2} - \frac{2\varepsilon^2}{E_{01}E_{21}} - \frac{\varepsilon^2}{E_{21}^2}\right)|0\rangle + \frac{\sqrt{6}\varepsilon^2}{E_{21}E_{31}}|3\rangle, \quad (39)$$

where $E_{nn'} \equiv E_n - E_{n'} = E_n^{(c)} - E_{n'}^{(c)}$. The corresponding energy is

$$E_{\overline{|1\rangle_\varepsilon}} \approx E_1 - \frac{\varepsilon^2}{E_{01}}\left(1 - \frac{\varepsilon^2}{E_{01}^2} - \frac{2\varepsilon^2}{E_{21}E_{01}}\right)$$
$$- \frac{2\varepsilon^2}{E_{21}}\left(1 + \frac{3\varepsilon^2}{E_{21}E_{31}} - \frac{2\varepsilon^2}{E_{21}^2} - \frac{\varepsilon^2}{E_{01}E_{21}}\right). \quad (40)$$
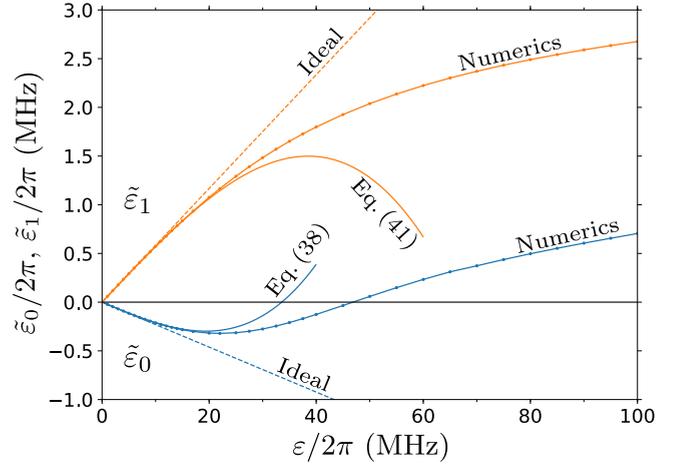

FIG. 4. Effective drive amplitudes $\tilde{\varepsilon}_0$ (blue lines) and $\tilde{\varepsilon}_1$ (orange lines) as functions of the drive amplitude ε, calculated using the ideal-case approximation (15) and (16) (dashed straight lines), the third-order formulas (38) and (41) (solid lines without symbols), and numerically (solid lines with symbols). We used the qubit-qubit coupling $g/2\pi = 3$ MHz, qubit anharmonicity $\eta_c/2\pi = \eta_t/2\pi = 300$ MHz, and detuning $\Delta/2\pi = 130$ MHz.

Using Eqs. (29) and (39), we obtain

$$\tilde{\varepsilon}_1 = -\frac{2\varepsilon g}{E_{21}}\left(1 + \frac{6\varepsilon^2}{E_{21}E_{31}} + \frac{\varepsilon^2}{E_{10}E_{21}} - \frac{4\varepsilon^2}{E_{21}^2} - \frac{\varepsilon^2}{E_{10}^2}\right)$$
$$+ \frac{\varepsilon g}{E_{10}}\left(1 - \frac{2\varepsilon^2}{E_{10}^2} + \frac{2\varepsilon^2}{E_{10}E_{21}} - \frac{2\varepsilon^2}{E_{21}^2}\right) \quad (41)$$

with accuracy up to ε^3 (note the use of E_{10} instead of E_{01} in the preceding formulas). In this formula $E_{10} = \Delta + \delta$, $E_{21} = \Delta + \delta - \eta_c$, $E_{31} = 2(\Delta + \delta) - 3\eta_c$, and we can neglect δ (i.e., $\delta = 0$). For a complex ε, we need to replace ε^2 in parentheses with $|\varepsilon|^2$.

Figure 4 shows the effective drive amplitudes $\tilde{\varepsilon}_0$ and $\tilde{\varepsilon}_1$ as functions of the actual drive amplitude ε calculated in several ways for the following parameters (which are some typical experimental parameters): $g/2\pi = 3$ MHz, $\eta_c/2\pi = 300$ MHz, and $\Delta/2\pi = 130$ MHz. The blue lines (which initially go down) show $\tilde{\varepsilon}_0$ and the orange lines (which initially go up) show $\tilde{\varepsilon}_1$. The solid lines without symbols are calculated using Eqs. (38) and (41) (using $\delta = 0$), while the straight dashed lines represent the simple linear approximation (15) and (16). The solid lines with symbols show the numerical results (the numerical procedure is described later in Sec. IV; for numerics we assume $\eta_t = \eta_c$ and $\omega_d = \omega_t^{c0}$).

We see that the third-order approximation [Eqs. (38) and (41)] correctly describes the deviation of the dependences $\tilde{\varepsilon}_0(\varepsilon)$ and $\tilde{\varepsilon}_1(\varepsilon)$ from ideal straight lines at relatively small ε, but fails to fit well the case of relatively large ε. This is because higher-order terms become important even for moderate values of ε. The problem is similar to the poor performance of the perturbation approach in the analysis of the circuit QED measurement of transmons when the number of photons is comparable to the critical number.

Figure 5 shows the difference $\tilde{\varepsilon}_1 - \tilde{\varepsilon}_0$ as a function of ε (for the same parameters as in Fig. 4), also calculated in several
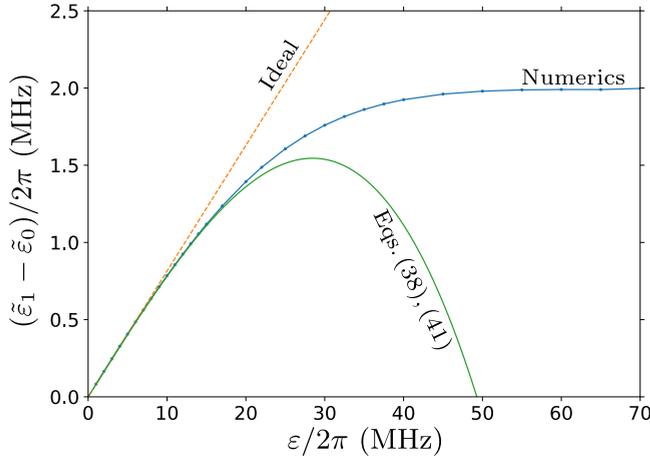
FIG. 5. The CR gate speed $\tilde{\varepsilon}_1 - \tilde{\varepsilon}_0$ as a function of the drive amplitude ε, calculated using the linear-order approximation (19) (dashed straight line), third-order approximation (38) and (41) (solid green line), and numerically (solid blue line with symbols). The parameters are the same as in Fig. 4.

FIG. 6. Effective drive amplitudes $\tilde{\varepsilon}_0$ and $\tilde{\varepsilon}_1$ as functions of ε, calculated numerically (solid lines with symbols) and using the semianalytical approach (26)–(29) (dashed lines, practically coinciding with the solid lines). We used $g/2\pi = 3$ MHz, $\eta_c/2\pi = \eta_t/2\pi = 300$ MHz, and two values for the detuning: $\Delta/2\pi = 130$ and 190 MHz.

ways. The dashed straight line corresponds to the simple formula (19) for the ideal operation. The numerical results are shown by the blue solid line with symbols. The green solid line is calculated using Eqs. (38) and (41). We have checked (analytically and numerically) that this line coincides with the result given by Eq. (4.25) of Ref. [33] (after correction of a misprint in the initial version of Ref. [33]; the translation of notation is $g = J_{[33]}$, $2\varepsilon = \Omega_{[33]}$, $\Delta = \Delta_{[33]}$, $\eta_c = -\delta_{1[33]}$, and $\tilde{\varepsilon}_0 - \tilde{\varepsilon}_1 = (ZX/2)_{\text{coeff}[33]}$). Most importantly, from Fig. 5 we see that the third-order approximation correctly describes the initial deviation of the CR gate speed from the linear-order result, but cannot be used for quantitative analysis in the practically interesting regime of large drive amplitudes.

C. Semianalytical results

The method developed in the preceding section can be naturally extended to arbitrary large drive amplitudes ε. For that the eigenstates $\overline{|n\rangle}_\varepsilon$ of the control-qubit Hamiltonian (27) can be found numerically, and then the effective drive amplitudes $\tilde{\varepsilon}_n$ can be calculated using Eq. (29). Since numerical diagonalization of a Hamiltonian for a few levels is very easy (compared with full numerical simulation of the two-qubit evolution discussed in the next section), we call this method semianalytical.

Figure 6 shows a comparison of the semianalytical results (dashed lines) for $\tilde{\varepsilon}_0$ and $\tilde{\varepsilon}_1$ with the numerical results (solid lines with symbols) (the numerical procedure is discussed in Sec. IV). The parameters are the same as in Figs. 4 and 5, except we use two values of the detuning: $\Delta/2\pi = 130$ and 190 MHz. In the semianalytics, we use seven levels of the control qubit. We see that the numerical results agree with semianalytics very well for all values of the drive amplitude ε (the lines are practically indistinguishable, except for the lowest lines at around 70 MHz, where a minor difference is caused by a resonance between levels $\overline{|0, 1\rangle}_\varepsilon$ and $\overline{|1, 2\rangle}_\varepsilon$). Similarly, we found very good agreement for other values of the parameters as well. Therefore, the semianalytical method
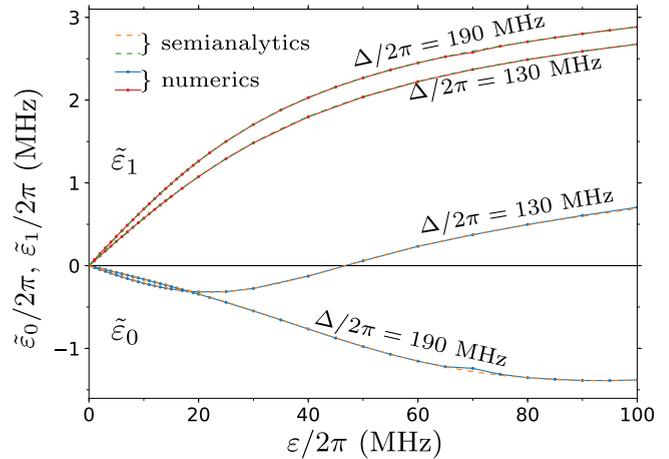
based on Eqs. (26)–(29) seems to be a sufficiently simple and accurate way of analyzing the dependence of the CR gate speed on parameters.

Note that $\tilde{\varepsilon}_n$ in the semianalytical method is proportional to the qubit-qubit coupling g and also depends on two dimensionless ratios Δ/η_c and ε/η_c (assuming $\delta = 0$). In the Duffing (Kerr) approximation (4), these two ratios fully define the eigenstates (28) (in a better approximation [39,40] the results would also depend on the dimensionless parameter η_c/ω_c). Therefore, in our analysis the ratio $\tilde{\varepsilon}_n/g$ is a function of *only two parameters* Δ/η_c and ε/η_c.

Figure 7 shows the dimensionless speed $(\tilde{\varepsilon}_1 - \tilde{\varepsilon}_0)/g$ of the CR gate as a function of the dimensionless drive amplitude ε/η_c for several values of the dimensionless detuning Δ/η_c (the lines are calculated using the semianalytical method). While the behavior at small ε agrees with Eq. (19) (not shown), the behavior at large ε mostly depends on whether the detuning $\Delta = \omega_c - \omega_t$ is negative or positive and on the integer part of the ratio $2\Delta/\eta_c$ for positive Δ. As seen in Fig. 7, at large ε the lines group according to the interval to which Δ belongs: $(-\infty, 0)$, $(0, \eta_c/2)$, $(\eta_c/2, \eta_c)$, $(\eta_c, 3\eta_c/2)$, $(3\eta_c/2, 2\eta_c)$, etc. (in Fig. 7 these groups of lines are labeled as I, II, III, etc.). We do not show the lines for $\Delta/\eta_c = 0$, $1/2$, 1, $3/2$, etc., because at these values there is a resonance between the levels, $E_n^{(c)} = E_0^{(c)}$ and $E_{n-1}^{(c)} = E_1^{(c)}$ for $n = 2\Delta/\eta_c + 1$ [see Eq. (4) for $\delta = 0$], and correspondingly the CR gate does not operate as intended (due to a very large leakage; see below), also leading to computational problems in the semianalytical and numerical calculations.

It is simple to understand why the lines in Fig. 7 group into bands at large ε. In the solution of the Schrödinger equation (26) for the Hamiltonian (27) at large ε, the main effect is a strong level repulsion, which depends on the relative position (topology) of the bare energy levels $E_n^{(c)}$ (i.e., which level is in between which levels; this topology does not change with ε because of the adiabatic theorem). In contrast, the level
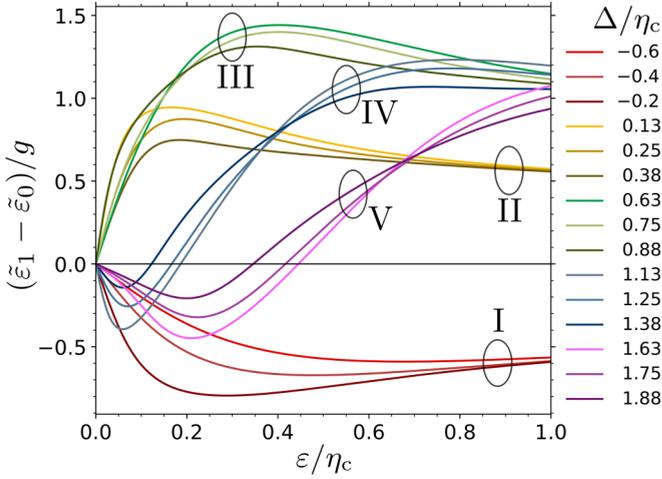
FIG. 7. Dimensionless CR gate speed $(\tilde{\varepsilon}_1 - \tilde{\varepsilon}_0)/g$ as a function of the dimensionless drive amplitude ε/η_c for several values of the dimensionless detuning Δ/η_c. The lines are calculated using the semianalytical method (26)–(29). At large ε, the lines group into bands. Group I is for $\Delta/\eta_c < 0$ and group II is for Δ/η_c in the interval $(0, \frac{1}{2})$. Similarly, groups III, IV, and V are for Δ/η_c in the intervals $(\frac{1}{2}, 1)$, $(1, \frac{3}{2})$, and $(\frac{3}{2}, 2)$, respectively.

repulsion does not depend much on a particular value of the initial bare level difference, since the effect of ε dominates. Therefore, at very large ε the eigenstates (28) do not depend on a particular value of Δ/η_c, but only on the integer part of $2\Delta/\eta_c$ (for $\Delta > 0$), which defines the topological structure of the ladder $E_n^{(c)}$ (as mentioned above, the bare levels $E_n^{(c)}$ intersect at integer values of $2\Delta/\eta_c$). Consequently, at very large ε the effective drive amplitudes (29) depend only on the integer part of $2\Delta/\eta_c$. This is why there is a grouping of lines in Fig. 7.

We see that most of the lines in Fig. 7 (all the lines in the experimentally important groups II and III) have a maximum, with a relatively minor decrease after it (experimental results [22,29] are somewhat similar). Experimentally, faster speed $(\tilde{\varepsilon}_1 - \tilde{\varepsilon}_0)$ means shorter CNOT gate duration and therefore there is no benefit to increase the drive amplitude beyond the maximum in Fig. 7.

The solid blue line in Fig. 8 shows the maximum value $(\tilde{\varepsilon}_1 - \tilde{\varepsilon}_0)_{\max}/g$ of the dimensionless speed (or the minimum value for the negative speed) as a function of the dimensionless detuning Δ/η_c. The dimensionless drive amplitude ε/η_c at which this maximum is reached is shown by the dashed orange line. We see that the maximum CR gate speed is reached for the detuning Δ between $\eta_c/2$ and η_c (group III in Fig. 7). Note that our semianalytical approach cannot be applied in close vicinities of the detunings $\Delta/\eta_c = 0, 0.5, 1, 1.5$, etc. The dependence of $(\tilde{\varepsilon}_1 - \tilde{\varepsilon}_0)_{\max}$ on Δ was measured experimentally [29] and it showed a crudely similar behavior.

IV. NUMERICAL APPROACH

Numerically, we simulate the quantum evolution due to the rotating-frame Hamiltonian (1)–(8), taking into account seven levels in the control qubit (the same number as in the semianalytics) and five levels in the target qubit, so there
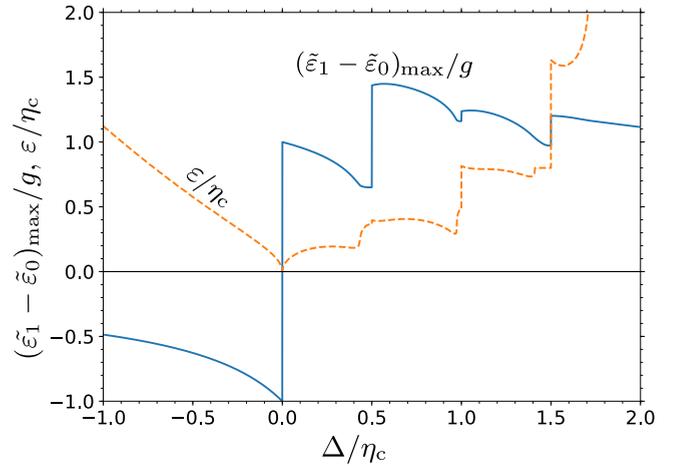

FIG. 8. Maximized (minimized for negative values) dimensionless speed $(\tilde{\varepsilon}_1 - \tilde{\varepsilon}_0)_{\max}/g$ (solid blue line) and corresponding dimensionless drive amplitude ε/η_c (dashed orange line), as functions of the dimensionless detuning Δ/η_c. The lines are calculated using the semianalytical method (26)–(29).

are 35 levels in total. The simulation is based on matrix exponentiation for a time-dependent Hamiltonian, using the second-order Magnus expansion [38]. We also tried to use the fourth-order Runge-Kutta method, but found that for our typical parameters it is almost an order of magnitude slower to reach the same desired accuracy.

We start with diagonalization of the time-independent part $H_{qb} + H_g$ of the Hamiltonian and then the whole simulation is done in the eigenbasis $\overline{|n, m\rangle}$ of $H_{qb} + H_g$, with the time-dependent drive Hamiltonian $H_\varepsilon(t)$ (expanded in the eigenbasis) causing the evolution. In this way, we obtain a 35×35 unitary evolution matrix V (in the eigenbasis $\overline{|n, m\rangle}$) for a given pulse of the drive amplitude $\varepsilon(t)$ with duration τ_p (the pulse shape is discussed later). The matrix V is then projected onto the computational two-qubit subspace, thus producing a 4×4 matrix M, which is no longer unitary (here projection means the simple elimination of all other elements). Note that the reduced matrix M is still defined in the eigenbasis, consisting of states $\overline{|0, 0\rangle} = |0, 0\rangle$, $\overline{|0, 1\rangle}$, $\overline{|1, 0\rangle}$, and $\overline{|1, 1\rangle}$.

To find the *fidelity* of an operation, we compare the reduced matrix M with the desired 4×4 unitary operation, which we denote by U. The fidelity between M and U is defined as [41,42]

$$F_{MU} = \frac{\text{Tr}(M^\dagger M)}{d(d+1)} + \frac{|\text{Tr}(M^\dagger U)|^2}{d(d+1)}, \qquad (42)$$

where $d = 4$ is the dimension of the two-qubit Hilbert space. This definition of the gate fidelity is equal to the final-state fidelity (squared overlap) averaged over all (pure) initial states in the two-qubit subspace; therefore, F_{MU} corresponds to the fidelity in randomized benchmarking (assuming that the states leaked outside the computational subspace never return).

Even though the final goal of the CR gate operation is to produce a CNOT gate (after additional single-qubit rotations), in the numerical procedure the desired U is obviously not the CNOT gate. Instead, for a given pulse $\varepsilon(t)$ (which produces some matrix V and corresponding matrix M), we *define* U

as the closest two-qubit unitary (i.e., which maximizes the fidelity F_{MU}), restricted to the class

$$U = e^{i\theta_0}|0\rangle\langle0|_c e^{-i(\varphi_0/2)X_t} + e^{i\theta_1}|1\rangle\langle1|_c e^{-i(\varphi_1/2)X_t}, \qquad (43)$$

where $|n\rangle\langle n|_c$ acts on the control qubit, while X_t acts on the target qubit (as mentioned above, we use the eigenbasis of $H_{qb} + H_g$ for both M and U). The condition (43) means that the state of the control qubit does not change (in the eigenbasis). Also, for state $|0\rangle$ of the control qubit, the target qubit is rotated about the x axis by angle φ_0; similarly, for control-qubit state $|1\rangle$, the target qubit is rotated about the x axis by angle φ_1. Besides that, in Eq. (43) there are phases θ_0 and θ_1; disregarding the unimportant overall phase, this can be interpreted as z rotation of the control qubit by angle $\theta_1 - \theta_0$. Without loss of generality, we could assume $\theta_0 = 0$ (while keeping the same $\theta_1 - \theta_0$) since this affects only the overall phase, and the definition (42) of the fidelity F_{MU} is insensitive to the overall phase of U. Note that Eqs. (42) and (43) can be easily generalized to include the third state of the control qubit (to consider it as a qutrit); however, here we consider only the two-level subspace.

Thus, to find U for a given pulse of $\varepsilon(t)$, we maximize F_{MU} over the parameters φ_0, φ_1, and $\theta_1 - \theta_0$. Fortunately, these optimal angles are given by analytical formulas in terms of the matrix elements of M:

$$\varphi_0 = -\arg\left(\frac{M_{11} + M_{22} + M_{12} + M_{21}}{M_{11} + M_{22} - M_{12} - M_{21}}\right), \qquad (44)$$

$$\varphi_1 = -\arg\left(\frac{M_{33} + M_{44} + M_{34} + M_{43}}{M_{33} + M_{44} - M_{34} - M_{43}}\right), \qquad (45)$$

$$\theta_0 = \arg[(M_{11} + M_{22})\cos(\varphi_0/2) + i(M_{12} + M_{21})\sin(\varphi_0/2)], \qquad (46)$$

$$\theta_1 = \arg[(M_{33} + M_{44})\cos(\varphi_1/2) + i(M_{34} + M_{43})\sin(\varphi_1/2)], \qquad (47)$$

where the rows (and columns) 1, 2, 3, and 4 of the matrix M correspond to the states $\overline{|00\rangle}$, $\overline{|01\rangle}$, $\overline{|10\rangle}$, and $\overline{|11\rangle}$, respectively. In this way, for a given pulse $\varepsilon(t)$, the CR gate operation is characterized by four resulting parameters: the angles φ_0, φ_1, and $\theta_0 - \theta_1$ of the unitary (43) and also infidelity $1 - F_{MU}$, which is due to leakage outside the computational two-qubit subspace and also due to the computational-space unitary not fitting well the class (43).

We consider the pulse shape $\varepsilon(t)$ of duration τ_p,

$$\varepsilon(t) = \begin{cases} \frac{1 - \cos(\pi t/\tau_r)}{2}\varepsilon_m, & 0 \leqslant t \leqslant \tau_r \\ \varepsilon_m, & \tau_r \leqslant t \leqslant \tau_p - \tau_r \\ \frac{1 - \cos[\pi(\tau_p - t)/\tau_r]}{2}\varepsilon_m, & \tau_p - \tau_r \leqslant t \leqslant \tau_p, \end{cases} \qquad (48)$$

which consists of the flat middle part with the *real* amplitude ε_m of the drive and two symmetric cosine-shaped ramps (so that there are no kinks), each with duration τ_r (see Fig. 9). As discussed later, sufficiently long ramps are needed to reduce leakage outside the computational subspace.

Effective drive amplitudes $\tilde{\varepsilon}_0$ and $\tilde{\varepsilon}_1$ used in Sec. III (Figs. 4–6) have been numerically calculated as the
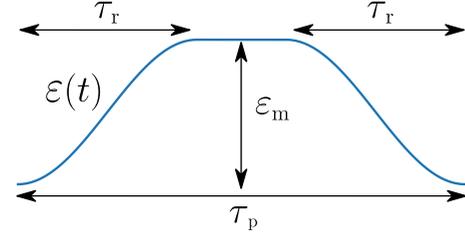

FIG. 9. Pulse shape $\varepsilon(t)$ used in numerical simulations. The total pulse duration is τ_p, each cosine-shaped ramp has duration τ_r, and the drive amplitude in the middle flat part is ε_m.

derivatives,

$$\tilde{\varepsilon}_0 = \frac{1}{2}\frac{\partial\varphi_0}{\partial\tau_p}, \qquad \tilde{\varepsilon}_1 = \frac{1}{2}\frac{\partial\varphi_1}{\partial\tau_p}, \qquad (49)$$

while keeping the ramp duration τ_r, the middle-part amplitude ε_m (which replaces ε in Sec. III), and other parameters fixed.

The small drive frequency detuning δ in numerical simulations is chosen in the following way. We first use the laboratory frame, i.e., $\delta = \omega_t$ (for the Duffing oscillator model we can also use $\delta = \omega_t = 0$), and calculate the eigenfrequencies of the target qubit ω_t^{c0} and ω_t^{c1} [see Eq. (9)]. If we want the drive to be exactly on-resonance with the target qubit when the control qubit is $|0\rangle$, then we need to use $\omega_d = \omega_t^{c0}$, which gives $\delta = \omega_t - \omega_t^{c0}$. Similarly, if we want $\omega_d = \omega_t^{c1}$ (the drive on-resonance with the target qubit when the control qubit is $|1\rangle$), then we use $\delta = \omega_t - \omega_t^{c1}$. If we want the drive frequency exactly in between the two resonances $\omega_d = (\omega_t^{c0} + \omega_t^{c1})/2$, then we use $\delta = \omega_t - (\omega_t^{c0} + \omega_t^{c1})/2$. Note that the frequency differences $\omega_t - \omega_t^{c0}$ and $\omega_t - \omega_t^{c1}$ do not depend on a choice of the rotating frame.

Since experimentally the CR gate is mainly used to realize the CNOT gate, in the numerical simulations we are mostly interested in the operations equivalent to CNOT up to single-qubit rotations. In analyzing the CNOT-equivalent gates, we usually use the pulse shape in which the ramps occupy 30% of the whole pulse duration each, i.e., $\tau_r = 0.3\tau_p$ in Eq. (48) (Fig. 21 is an exception). For a given middle-part amplitude ε_m, we find the shortest pulse duration τ_p, for which

$$\varphi_1 - \varphi_0 = \pi \,(\mathrm{mod}\,2\pi). \qquad (50)$$

This is what we call the CNOT gate duration $\tau_p^{\mathrm{CNOT}}(\varepsilon_m)$, neglecting durations of the additional single-qubit operations (x rotation of the target qubit and z rotation of the control qubit). We assume perfect fidelity of single-qubit operations; therefore, the CNOT gate infidelity is $1 - F_{MU}$ for the pulse with duration $\tau_p^{\mathrm{CNOT}}(\varepsilon_m)$.

In the simulations we fully neglect decoherence. A crude estimate of the fidelity decrease ΔF due to energy relaxation and pure dephasing can be obtained by considering idle qubits, which decohere during time τ_p^{CNOT}. This gives the estimate

$$\Delta F \simeq \frac{1}{5}\frac{\tau_p^{\mathrm{CNOT}}}{T_1^{(c)}} + \frac{1}{5}\frac{\tau_p^{\mathrm{CNOT}}}{T_1^{(t)}} + \frac{2}{5}\frac{\tau_p^{\mathrm{CNOT}}}{T_2^{(c)}} + \frac{2}{5}\frac{\tau_p^{\mathrm{CNOT}}}{T_2^{(t)}}, \qquad (51)$$

where $T_1^{(c)}$ and $T_1^{(t)}$ are the energy relaxation times for the control and target qubits and similarly $T_2^{(c)}$ and $T_2^{(t)}$ are the

dephasing times (which include contributions due to the energy relaxation and pure dephasing $1/T_2 = 1/2T_1 + 1/T_\varphi$). This estimate is obtained by summing the single-qubit Pauli error rates $t/2T_1$ and $t/2T_\varphi$ and then converting the result into the two-qubit average gate fidelity by using the factor $\frac{4}{5}$. Note, however, that the actual fidelity decrease ΔF can be significantly larger than the estimate (51) because the CR gate operation involves a significant population of the level $|2\rangle$ and even higher levels of the control qubit, which have poorer coherence than the level $|1\rangle$.

One run of the evolution simulation for a given pulse duration typically takes a few seconds on a desktop or a laptop computer. Finding τ_p^{CNOT} requires several tens of runs, so a typical time to produce a line showing dependence of the CNOT gate operation on ε_m is a few hours. The simulation time significantly depends on the number of time steps in the pulse ramps; we have used 600 time steps for each ramp, which gives quite good accuracy for the simulations. For quick (and much less accurate) simulations it is possible to use ~100 time steps per ramp.

Let us list the main approximations used in our numerics: (i) neglected decoherence, (ii) a Duffing-oscillator approximation for the transmon energy levels, (iii) a linear-oscillator approximation for the transmon matrix elements, (iv) direct coupling of qubits instead of coupling via resonator, (v) RWA, (vi) using only 7×5 levels (we checked that this is sufficient), (vii) no microwave crosstalk (except in Fig. 22), (viii) a simple pulse shape without distortions, and (ix) the absence of neighboring qubits. In spite of a rather long list of approximations, we believe our simulation results give a reasonably accurate description of the intrinsic operation of the CR gate (neglecting decoherence, which in practice may give the biggest contribution to infidelity).

V. NUMERICAL CNOT GATE DURATION AND SINGLE-QUBIT ROTATIONS

In numerical simulations, we use the qubit-qubit coupling $g/2\pi = 3$ MHz (except in Fig. 20) and transmon anharmonicity $\eta_c/2\pi = \eta_t/2\pi = 300$ MHz. Figure 10 shows the CNOT gate duration τ_p^{CNOT} as a function of the drive amplitude ε_m in the flat middle part of a pulse (with $\tau_r = 0.3\tau_p$), for several values of the qubit-qubit detuning: $\Delta/2\pi = -70$, 70, 130, and 190 MHz. For numerical results (solid lines) we choose $\omega_d = \omega_t^{c0}$, i.e., $\delta = \omega_t - \omega_t^{c0}$. The dashed lines (almost coinciding with the solid lines) show the result of the semianalytical method, in which we use Eq. (13) and integrate over the pulse shape. We see that the semianalytical method works very well; however, there are (barely) visible deviations at both small and large amplitudes ε_m. We guess the slight deviation at large ε_m is because for a short pulse the nonadiabatic evolution during the ramps starts to play a noticeable role. The deviation at small ε_m is because here the zz coupling (10) starts to play a relatively significant role.

For a more detailed analysis of the deviation at small ε_m, solid lines in Fig. 11 show the numerical CNOT time τ_p^{CNOT} for $\Delta/2\pi = 130$ MHz and three values of the drive frequency: $\omega_d = \omega_t^{c0}$ (blue line, drive on-resonance with the target qubit when the control qubit is $|0\rangle$, i.e., $\delta = \omega_t - \omega_t^{c0}$), $\omega_d = \omega_t^{c1}$ (orange line, on-resonance when the control qubit is $|1\rangle$, i.e.,
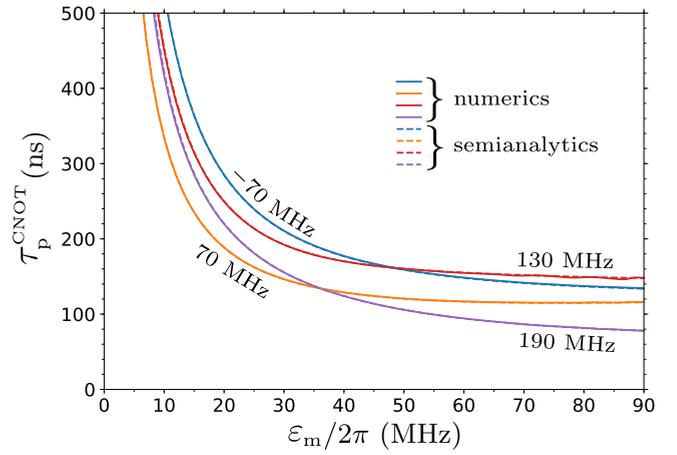

FIG. 10. CNOT gate duration τ_p^{CNOT} (neglecting single-qubit rotations) as a function of the midpulse drive amplitude ε_m for several detunings: $\Delta/2\pi = -70$, 70, 130, and 190 MHz, while $g/2\pi = 3$ MHz, $\eta_c/2\pi = \eta_t/2\pi = 300$ MHz, $\tau_r/\tau_p = 0.3$, and $\omega_d = \omega_t^{c0}$. Solid lines are calculated numerically; dashed lines (almost coinciding with the solid lines) are calculated using the semianalytical method.

$\delta = \omega_t - \omega_t^{c1}$), and exactly in between, $\omega_d = (\omega_t^{c0} + \omega_t^{c1})/2$ (green line). The dashed line shows the semianalytical result (actually, there are three dashed lines for the three values of δ, but they are indistinguishable) and the dotted line shows the ideal result

$$\tau_{p,\text{ideal}}^{\text{CNOT}} = \frac{\pi/2}{0.7\varepsilon_m} \frac{\Delta(\eta_c - \Delta)}{2g\eta_c}, \tag{52}$$

which follows from Eqs. (13) and (19) after integration over the pulse shape (48) with $\tau_r/\tau_p = 0.3$ (this integration gives the factor 0.7). We see that the solid lines noticeably deviate
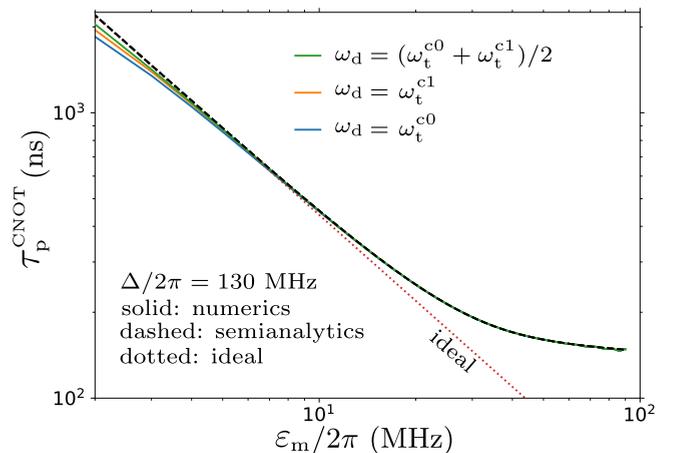

FIG. 11. Solid lines show the dependence of the CNOT gate duration τ_p^{CNOT} on ε_m for three drive frequencies: $\omega_d = \omega_t^{c0}$ (exact resonance for the control-qubit state $|0\rangle$, blue line), $\omega_d = \omega_t^{c1}$ (resonance for the control-qubit state $|1\rangle$, orange line), and $\omega_d = (\omega_t^{c0} + \omega_t^{c1})/2$ (exactly in between, green line). The dashed line shows the semianalytical results and the straight dotted line corresponds to Eq. (52). We use $\Delta/2\pi = 130$ MHz; the other parameters are the same as in Fig. 10.

down from the dashed line for $\varepsilon_m/2\pi$ less than ~ 5 MHz, with the largest deviation for $\omega_d = \omega_t^{c0}$ (blue line). Note that for $\varepsilon/2\pi = 5$ MHz, Eq. (19) gives $(\tilde{\varepsilon}_1 - \tilde{\varepsilon}_0)/2\pi = 0.41$ MHz, while $\omega_{zz}/2\pi = 0.15$ MHz [see Eq. (11)], so the effect of zz coupling is expected to be significant.

In more detail, for $\omega_d = \omega_t^{c0}$ (blue line in Fig. 11) the drive is exactly on-resonance with the target qubit for the control-qubit state $|0\rangle$ and therefore the approximation $\varphi_0 = 0.7\tau_p^{\text{CNOT}} \times 2\tilde{\varepsilon}_0(\varepsilon_m)$ using Eq. (15) should work well (as confirmed by numerics). In contrast, for the control-qubit state $|1\rangle$, the drive is detuned by ω_{zz} from the target-qubit frequency, which leads to Rabi oscillations with frequency $\sqrt{(2\tilde{\varepsilon}_1)^2 + \omega_{zz}^2}$ within the plane tilted by an angle $\arctan(\omega_{zz}/2\tilde{\varepsilon}_1)$ from the zy plane (the frequency and the plane are changing in time because of the pulse shape). The larger Rabi frequency (due to ω_{zz} contribution) leads to a slightly shorter CNOT gate duration than expected analytically, explaining the behavior of solid lines in Fig. 11 at small ε_m. The same effect (rotation within a tilted plane) leads to a significant infidelity of the CNOT gate at small ε_m (see Sec. VI).

As discussed above, in order to realize the CNOT operation, the CR gate with the pulse duration τ_p^{CNOT} should be complemented by single-qubit rotations. The target qubit should be rotated about x axis by the angle $-\varphi_0$ to compensate the operator $e^{-i(\varphi_0/2)X_t}$ in Eq. (43). Similarly, the control qubit should be rotated about z axis to compensate the relative phase $\theta_1 - \theta_0$ and the negative imaginary unit due to the relation $e^{-i(\pi/2)X_t} = -iX_t$. Thus, naively, we would expect that z rotation by the angle $\pi/2 - (\theta_1 - \theta_0)$ is needed. However, the angle $\theta_1 - \theta_0$ is numerically computed in the rotating frame of the drive, while experimental z rotation should be in the rotating frame of the control qubit. For the latter frame based on frequency $\omega_c^{t0} = E_{|1,0\rangle}^{(\text{lf})} - E_{|0,0\rangle}^{(\text{lf})}$ (i.e., when the target qubit is $|0\rangle$, as usually done in experiments), we need to replace the phase difference $\theta_1 - \theta_0$ with

$$\theta_1' - \theta_0' = \theta_1 - \theta_0 + \left(\omega_c^{t0} - \omega_d\right)\tau_p^{\text{CNOT}}, \qquad (53)$$

so in an experiment, the control qubit should be z rotated by the angle $\theta_0' - \theta_1' + \pi/2$. Note that $\omega_c^{t0} - \omega_d = \Delta + \delta + \omega_c^{t0} - \omega_c$.

Figure 12 shows the angle $-\varphi_0$ (normalized by π) as a function of ε_m for two values of the detuning: $\Delta/2\pi = 130$ and 190 MHz (we use $\omega_d = \omega_t^{c0}$). Solid lines show the numerical results, dashed lines show the corresponding semianalytical results (integrating $-2\tilde{\varepsilon}_0$ over the pulse shape), and horizontal dotted lines show the ideal result based on Eqs. (15) and (16): $-\varphi_0 = \pi(\eta_c - \Delta)/2\eta_c$. We see that this ideal result for $-\varphi_0$ is never applicable. The deviation from it at large ε_m is described well by the semianalytics and is due to the deviations from the ideal straight lines in Fig. 6. The deviation from the ideal result in Fig. 12 at small ε_m is due to ω_{zz}, the same effect as discussed above for Fig. 11.

The upper (blue) solid line in Fig. 13 shows the numerical result for the control-qubit rotation angle $\theta_0' - \theta_1' + \pi/2$ (normalized by π) as a function of ε_m for $\Delta/2\pi = 130$ MHz and $\omega_d = \omega_t^{c0}$. We see that it is quite different from the ideally expected value of $\pi/2$ (horizontal brown line). The difference is mainly due to two effects. First, strong drive $\varepsilon(t)$ causes the level repulsion (ac Stark shift) in the control qubit, which
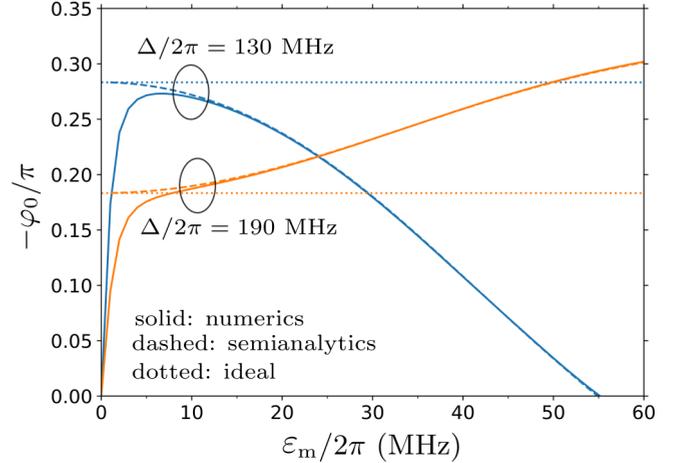

FIG. 12. Angle $-\varphi_0$ of the compensating x rotation of the target qubit to produce the CNOT gate, as a function of the midpulse drive amplitude ε_m. Numerical results are shown by solid lines, semianalytics are represented by dashed lines, and dotted lines show the ideal result $-\varphi_0/\pi = (\eta_c - \Delta)/2\eta_c$. Here we use $\Delta/2\pi = 130$ and 190 MHz, $\omega_d = \omega_t^{c0}$, $g/2\pi = 3$ MHz, $\eta_c/2\pi = \eta_t/2\pi = 300$ MHz, and $\tau_r/\tau_p = 0.3$.

slightly changes the control-qubit frequency and produces the accumulated angle

$$\theta_{\text{rep}} = \int_0^{\tau_p} \left(E_{\overline{|1\rangle}_\varepsilon} - E_{\overline{|0\rangle}_\varepsilon}\right)dt. \qquad (54)$$

(Actually, because the state $|1\rangle$ is assumed to be at the bottom of the Bloch sphere, this produces the z rotation of $-\theta_{\text{rep}}$, so we need to apply the rotation of $+\theta_{\text{rep}}$ to compensate it.) The
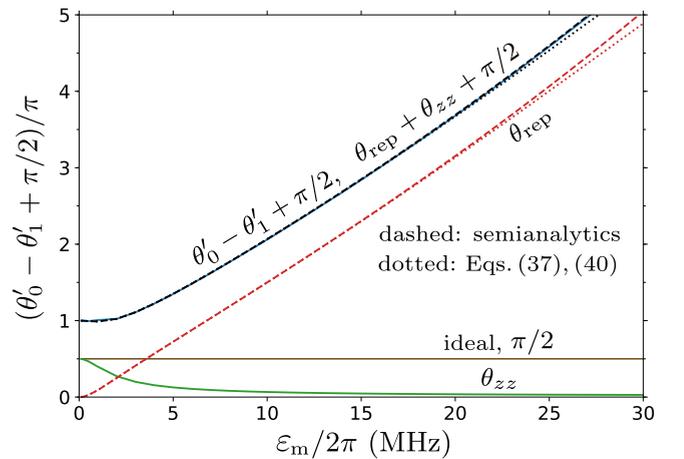

FIG. 13. The upper (blue) solid line shows the numerical result for the angle $\theta_0' - \theta_1' + \pi/2$ of z rotation of the control qubit, needed to produce the CNOT gate. The horizontal (brown) line shows the ideal value $\pi/2$. Dashed and dotted red lines show the contribution θ_{rep} due to ε-induced level repulsion (54), calculated either semianalytically (dashed line) or via Eqs. (37) and (40) (dotted line). The lower (green) solid line shows the contribution θ_{zz} given by Eq. (55). The upper solid and dashed black lines (very close to the solid blue line) show the sum $\theta_{\text{rep}} + \theta_{zz} + \pi/2$. We use $\Delta/2\pi = 130$ MHz and the parameters from Fig. 12.
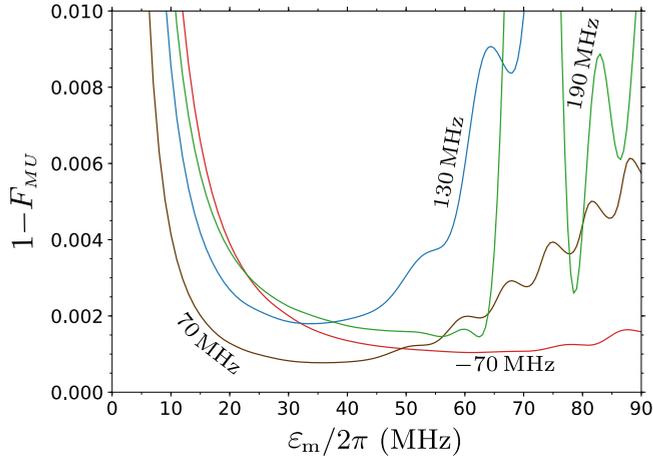
FIG. 14. Infidelity $1 - F_{MU}$ of the CNOT-equivalent gate as a function of midpulse drive amplitude ε_m for several values of detuning: $\Delta/2\pi = -70$ MHz (red line), 70 MHz (brown line), 130 MHz (blue line), and 190 MHz (green line). The other parameters are $g/2\pi = 3$ MHz, $\eta_c/2\pi = \eta_t/2\pi = 300$ MHz, $\omega_d = \omega_t^{c0}$, and $\tau_r/\tau_p = 0.3$.

angle θ_{rep} calculated semianalytically [see Eq. (26)] is shown by the dashed red line in Fig. 13; the same angle calculated using analytical results (37) and (40) is shown by the dotted red line (in both cases we use numerical τ_p^{CNOT} to integrate over the pulse shape). The second contribution into $\theta_0' - \theta_1'$ is due to zz coupling (10), which produces

$$\theta_{zz} = \frac{\omega_{zz}}{2}\tau_p. \qquad (55)$$

This is because the rotating frame here is based on ω_c^{t0}, while $\omega_c^{t1} - \omega_c^{t0} = \omega_{zz}$ and both states of the target qubit participate equally (leading to the factor of 1/2). The angle θ_{zz} is shown by the lower (green) solid line in Fig. 13. Adding the three contributions $\theta_{rep} + \theta_{zz} + \pi/2$, we obtain the dashed and dotted black lines (corresponding to the dashed and dotted red lines for θ_{rep}), which are quite close to the numerical result (solid blue line). This confirms the main physical mechanisms contributing to $\theta_0' - \theta_1'$ and also shows that the approximation based on Eqs. (37) and (40) works quite well.

VI. ERROR BUDGET

In the preceding section we have discussed numerical results for the parameters of a CR-based CNOT gate: the duration τ_p^{CNOT} as a function of the midpulse drive amplitude ε_m and the compensating single-qubit rotation angles $-\varphi_0$ and $\theta_0' - \theta_1' + \pi/2$. We have seen that these parameters can be obtained quite accurately by the semianalytical method. In this section we discuss numerical results for the infidelity $1 - F_{MU}$ of the CR-based CNOT gate (also as a function of ε_m), neglecting decoherence and infidelity of single-qubit rotations (see Sec. IV for the definition of F_{MU} and the calculation method). These results cannot be obtained semianalytically and necessarily require full numerical simulations.

Figure 14 shows the numerical results for the CNOT gate infidelity $1 - F_{MU}$ as a function of the midpulse drive amplitude ε_m for several values of the detuning: $\Delta/2\pi = -70, 70,$

130, and 190 MHz. As in the previous plots, we use $g/2\pi = 3$ MHz, $\eta_c = \eta_t = 300$ MHz, $\omega_d = \omega_t^{c0}$, and $\tau_r = 0.3\tau_p$. Most importantly, we see that the infidelity dependence on ε_m has a minimum, and at this minimum, the infidelity is crudely 10^{-3} for all lines. The second observation is that the minimum is not sharp and is reached for the values of the drive amplitude ε_m above which the CNOT gate duration τ_p^{CNOT} does not become significantly shorter by a further increase of ε_m (see Figs. 7 and 10). Note that we do not take into account the effect of decoherence, which can be crudely (ideally) estimated by Eq. (51). If decoherence were added, then the minima in Fig. 14 would shift to higher values of ε_m; however, since the corresponding decrease of τ_p^{CNOT} is not significant, the benefit for fidelity is also not significant; moreover, using higher drive amplitudes could lead to other experimental problems. Therefore, we think the optimum values of ε_m in Fig. 14 should be somewhat close to experimental optima.

To understand the reason for the minima in Fig. 14 and to understand the physical mechanisms contributing to the infidelity (error budget), we have done the following more detailed calculations. We recall that we calculate the fidelity F_{MU} between the nonunitary 4×4 matrix M (which is the projection of the 35×35 matrix of actual evolution onto the computational subspace) and the closest unitary matrix U, which belongs to the class (43) with $\varphi_1 - \varphi_0 = \pi \pmod{2\pi}$ for a CNOT-equivalent gate. Now let us consider a bigger class of unitary matrices and define $\tilde{M}$ as the matrix, which is closest to M out of unitaries satisfying the condition

$$\tilde{M} = |0\rangle\langle 0|_c \tilde{U}_0^t + |1\rangle\langle 1|_c \tilde{U}_1^t, \qquad (56)$$

where $\tilde{U}_0^t$ and $\tilde{U}_1^t$ are any 2×2 unitary matrices acting on the target qubit (here "closest" means that $\tilde{M}$ maximizes the fidelity $F_{M\tilde{M}}$). From this definition, $\tilde{M}$ does not change the control-qubit states $|0\rangle$ and $|1\rangle$ (i.e., does not allow any leakage from them), but its rotation of the target qubit is arbitrary. The matrix U in Eq. (43) is more restrictive in the sense that it allows only x rotations of the target qubit.

To clarify the error budget, we calculate additional infidelities $1 - F_{M\tilde{M}}$ (between M and $\tilde{M}$) and $1 - F_{\tilde{M}U}$ (between $\tilde{M}$ and U) for the CNOT-equivalent CR operation. The idea is that the infidelity $1 - F_{M\tilde{M}}$ is due to leakages (from the control-qubit states $|0\rangle$ and $|1\rangle$ to any state and also outside the computational subspace for the target qubit). In contrast, the infidelity $1 - F_{\tilde{M}U}$ is due to imperfect unitary rotations of the target qubit. From the physical approach of separation of the error into these different mechanisms, we would expect

$$1 - F_{MU} \approx (1 - F_{M\tilde{M}}) + (1 - F_{\tilde{M}U}). \qquad (57)$$

This is not an exact relation mathematically (the exact relation would require that certain elements of the quantum process tomography matrix for M have exactly zero real and/or imaginary parts). However, numerical results (e.g., Fig. 15) show that this relation is very accurate.

In Fig. 15 the solid blue line, $1 - F_{MU}$, is the same as the blue line in Fig. 14 ($\Delta/2\pi = 130$ MHz). For the same case, the orange line shows $1 - F_{M\tilde{M}}$, the green line shows $1 - F_{\tilde{M}U}$, and the dashed blue line (practically indistinguishable from the solid blue line) shows the sum $(1 - F_{M\tilde{M}}) + (1 - F_{\tilde{M}U})$. We see that the total infidelity $1 - F_{MU}$ can really be
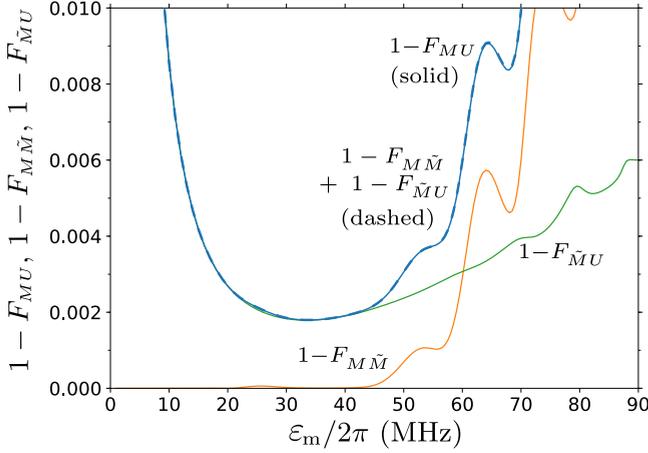
FIG. 15. Decomposition of the CNOT gate infidelity $1 - F_{MU}$ (solid blue line) into the leakage contribution $1 - F_{M\tilde{M}}$ (orange line) and contribution $1 - F_{\tilde{M}U}$ due to imperfect unitaries (green line). The sum $(1 - F_{M\tilde{M}}) + (1 - F_{\tilde{M}U})$ (dashed blue line) is practically indistinguishable from the solid blue line. We use $\Delta/2\pi = 130$ MHz and the parameters from Fig. 14.
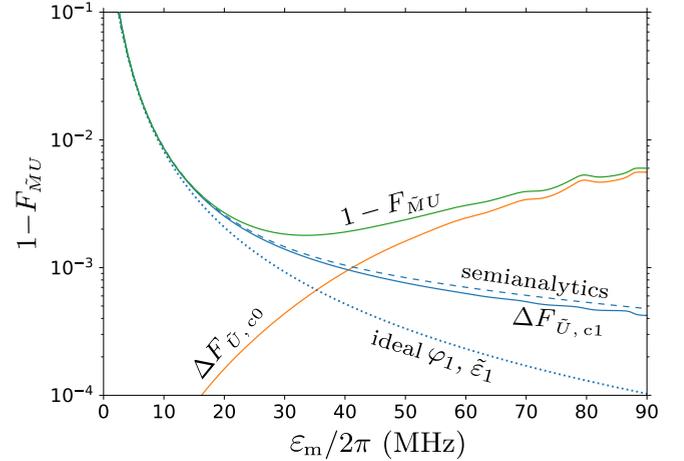

FIG. 16. Further decomposition of the imperfect-unitary infidelity contribution $1 - F_{\tilde{M}U}$ (green line) into contributions $\Delta F_{\tilde{U},c0}$ (orange line) and $\Delta F_{\tilde{U},c1}$ (solid blue line) for the control-qubit states $|0\rangle$ and $|1\rangle$, respectively. The dotted blue line shows the analytical approximation for $\Delta F_{\tilde{U},c1}$ by Eq. (59) with ideal values for φ_1 and $\tilde{\varepsilon}_1$, while for the dashed blue line we use semianalytical values for φ_1 and $\tilde{\varepsilon}_1$ in Eq. (59). The parameters are the same as in Fig. 15.

decomposed into the infidelity $1 - F_{M\tilde{M}}$ due to leakages and infidelity $1 - F_{\tilde{M}U}$ due to imperfect unitaries acting on the target qubit. Similar decomposition into leakages and imperfect unitaries also works well for other lines in Fig. 14 (relative inaccuracy of the decomposition is typically about 10^{-3}).

Besides introducing the error budget via the decomposition (57), we also tried a further decomposition by introducing another 4×4 matrix $\tilde{M}'$, which is a two-qubit unitary closest to M. In this case the infidelity $1 - F_{M\tilde{M}'}$ is due to leakages outside the computational subspace, while the infidelity $1 - F_{\tilde{M}'\tilde{M}}$ is due to unitary transitions between states $|0\rangle$ and $|1\rangle$ of the control qubit (which in our terminology are also leakages for the CR gate). Then the infidelity consists of three components

$$1 - F_{MU} \approx (1 - F_{M\tilde{M}'}) + (1 - F_{\tilde{M}'\tilde{M}}) + (1 - F_{\tilde{M}U}). \quad (58)$$

We have checked that this relation is quite accurate numerically. However, for the cases we checked, the matrix $\tilde{M}'$ was very close to either M or $\tilde{M}$ (depending on the detuning Δ, which determines the strongest leakage process). Therefore, usually only two terms in Eq. (58) are significant, and this is why we will continue using the simpler error decomposition (57) below.

As seen in Fig. 15, at small ε_m the CNOT gate infidelity is dominated by the imperfection of the unitaries $\tilde{U}_0^t$ and $\tilde{U}_1^t$ (contribution $1 - F_{\tilde{M}U}$), while at large ε_m the leakage contribution $1 - F_{M\tilde{M}}$ dominates (the same result for other detunings Δ). This is because at small ε_m the effect of zz coupling is important (as discussed in Sec. V), while at large ε_m the ramps of the pulse become short and high, making the process significantly nonadiabatic and causing leakages.

To clarify the dependence on ε_m of the imperfect-unitary contribution $1 - F_{\tilde{M}U}$ (green line in Fig. 15), we draw this line again in Fig. 16 (now on semilogarithmic scale). We also show the numerical results for the contributions $\Delta F_{\tilde{U},c0}$ and $\Delta F_{\tilde{U},c1}$ to this line from imperfections of the unitaries $\tilde{U}_0^t$ (for the control-qubit state $|0\rangle$)

and $\tilde{U}_1^t$ (for the control-qubit state $|1\rangle$): orange and blue solid lines in Fig. 16, respectively. [Mathematically, the definitions are $\Delta F_{\tilde{U},c0} = 4/5 - (2/5)|\mathrm{Tr}(\tilde{U}_0^t U_0^{t\dagger})|$ and $\Delta F_{\tilde{U},c1} = 4/5 - (2/5)|\mathrm{Tr}(\tilde{U}_1^t U_1^{t\dagger})|$, where U_0^t and U_1^t are obtained from Eq. (43): $U_0^t = e^{-i(\varphi_0/2)\hat{X}_t}$ and $U_1^t = e^{-i(\varphi_1/2)\hat{X}_t}$.] At small ε_m the main contribution comes from imperfect $\tilde{U}_1^t$. This is because we use the drive frequency resonant with the target qubit for the control-qubit state $|0\rangle$ ($\omega_d = \omega_t^{c0}$), so for the control-qubit state $|1\rangle$ the drive is off-resonance by ω_{zz} [see Eq. (10)]. This detuning produces rotation of the target-qubit state about a tilted axis, instead of the desired x axis. To check this explanation, we calculate analytically the corresponding infidelity

$$\Delta F_{\tilde{U},c1} = \frac{2}{5} \sin^2\left(\frac{\varphi_1}{2}\right) \frac{\omega_{zz}^2}{\frac{25}{40}[2\tilde{\varepsilon}_1(\varepsilon_m)]^2 + \omega_{zz}^2}, \quad (59)$$

where the factor $\frac{25}{40}$ comes from the integration of $\varepsilon^2(t)$ over the pulse shape (48) with $\tau_r/\tau_p = 0.3$ (such integration in only the denominator is an approximation). The dotted blue line in Fig. 16 shows this result with φ_1 and $\tilde{\varepsilon}_1$ calculated analytically, $\varphi_1 = (\eta_c + \Delta)/2\eta_c$, and Eq. (16) for $\tilde{\varepsilon}_1$. The dashed blue line shows Eq. (59) with φ_1 and $\tilde{\varepsilon}_1$ calculated using the semianalytical method. We see that both lines fit reasonably well the numerical result (solid blue line) at small ε_m, with a better fit when using the semianalytical values for φ_1 and $\tilde{\varepsilon}_1$. An even better fit (almost perfect, not shown) is when the Bloch-sphere evolution due to ω_{zz} and semianalytical $\tilde{\varepsilon}[\varepsilon(t)]$ is integrated over the pulse shape numerically, instead of using Eq. (59).

Thus, the contribution $\Delta F_{\tilde{U},c1}$ to the gate infidelity due to imperfect unitary $\tilde{U}_t^{c1}$ is well explained quantitatively. In contrast, we do not have a simple analytical way to find the contribution $\Delta F_{\tilde{U},c0}$ due to imperfect $\tilde{U}_t^{c0}$ (solid orange line in Fig. 16). Qualitatively, this contribution appears at large ε_m because the interplay between the level repulsions due to g and
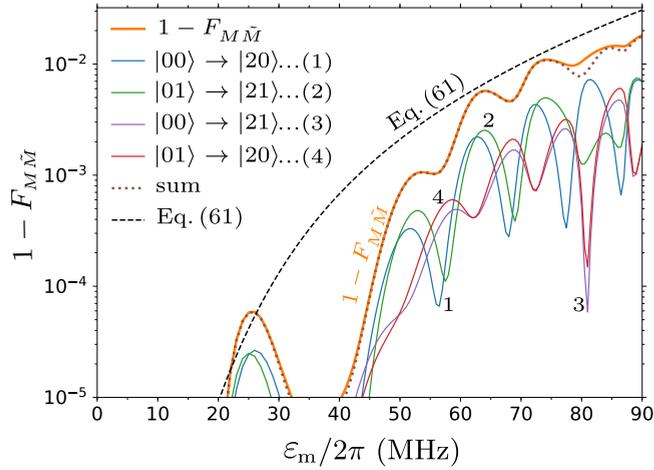
FIG. 17. Numerical results for the leakage. The thick orange line shows the infidelity contribution $1 - F_{M\tilde{M}}$. Four thin solid lines (labeled 1, 2, 3, and 4) show multiplied by the factor $\frac{1}{4}$ probabilities of the main leakage channels: $\overline{|0, 0\rangle} \to \overline{|2, 0\rangle}$, $\overline{|0, 1\rangle} \to \overline{|2, 1\rangle}$, $\overline{|0, 0\rangle} \to \overline{|2, 1\rangle}$, and $\overline{|0, 1\rangle} \to \overline{|2, 0\rangle}$. The sum of these four lines is shown by the dotted brown line; its closeness to the thick orange line verifies that the infidelity $1 - F_{M\tilde{M}}$ is mainly due to these leakage channels. The dashed black line is the leakage probability estimate given by Eq. (61). The parameters are the same as in Fig. 15.

ε in Fig. 3 slightly changes the frequency ω_t^{c0}, so there is no longer exact resonance with the drive, and the Bloch-sphere evolution also becomes tilted, thus producing the infidelity. Note that small oscillations of the solid orange and blue lines in Fig. 16 at large ε_m are apparently related to the oscillations of the orange line in Fig. 15, which are discussed below.

The second contribution to the overall gate infidelity $1 - F_{MU}$ (Fig. 15), which becomes dominating at large ε_m, is the contribution $1 - F_{M\tilde{M}}$ due to leakage produced by nonadiabaticity during the pulse ramps (orange line in Fig. 15). From the fidelity definition (42), we expect an approximate relation

$$1 - F_{M\tilde{M}} \approx P_{\text{leak}}^{\text{out}} + \frac{4}{5} P_{\text{leak}}^{\text{comp}}, \tag{60}$$

where $P_{\text{leak}}^{\text{out}}$ is the probability of leakage to outside the computational subspace, averaged over the initial two-qubit states, and $P_{\text{leak}}^{\text{comp}}$ is the averaged probability of leakage inside the computational subspace, which for our definition (56) means transitions between states $|0\rangle$ and $|1\rangle$ of the control qubit. Note that to average the leakage probability over all initial two-qubit states, it is sufficient to average it over the four basis states $\overline{|0, 0\rangle}$, $\overline{|0, 1\rangle}$, $\overline{|1, 0\rangle}$, and $\overline{|1, 1\rangle}$. We have checked the relation (60) numerically and it works very well; however, usually either the first or the second term in Eq. (60) strongly dominates over the other term (depending on the dominating leakage channel).

Figure 17 shows again the orange line $1 - F_{M\tilde{M}}$ from Fig. 15 (now it is the thick orange line and the scale is semilogarithmic) and also shows the numerical leakage probabilities (multiplied by the factor $\frac{1}{4}$ as in the averaging) for the processes $\overline{|0, 0\rangle} \to \overline{|2, 0\rangle}$, $\overline{|0, 1\rangle} \to \overline{|2, 1\rangle}$, $\overline{|0, 0\rangle} \to \overline{|2, 1\rangle}$, and $\overline{|0, 1\rangle} \to \overline{|2, 0\rangle}$ (thin solid lines). The sum of the thin solid lines is shown as the dotted brown line. It is very close to the thick orange line [as expected from Eq. (60)], indicating that

these are the four dominating leakage channels. A minor difference between the thick orange line and dotted brown line at ε_m close to 80 MHz is due to the significance of the additional leakage channels $\overline{|0, 1\rangle} \to \overline{|1, 2\rangle}$ and $\overline{|0, 0\rangle} \to \overline{|1, 2\rangle}$ (at this frequency the states $\overline{|2, 1\rangle}$ and $\overline{|1, 2\rangle}$ become on-resonance); similarly, a visible difference at about 40 MHz is because the transitions $|0\rangle \leftrightarrow |1\rangle$ in the control qubit become relatively important.

The main leakage channels in Fig. 17 involve the transition $|0\rangle \to |2\rangle$ in the control qubit. This is expected because in the rotating frame $E_0^{(c)} - E_2^{(c)} = \eta - 2\Delta$ is only 40 MHz for the parameters of Fig. 17. The matrix element of this transition via virtual state $|1\rangle$ is $-\sqrt{2}\varepsilon^2/\Delta$. Therefore, the nonadiabatic transition $\overline{|0\rangle_\varepsilon} \to \overline{|2\rangle_\varepsilon}$ during the front ramp with duration τ_r of the pulse (48) has an amplitude proportional to the Fourier transform of $d(\varepsilon^2)/dt$ at the frequency $\eta - 2\Delta$. The standard calculations give an estimate of the leakage probability in the control qubit during the front ramp:

$$P_{\overline{|0\rangle_\varepsilon} \to \overline{|2\rangle_\varepsilon}} = \frac{2\pi^4 \varepsilon_m^4}{\Delta^2 (\eta - 2\Delta)^6 \tau_r^4}. \tag{61}$$

This probability is shown in Fig. 17 by the dashed black line. We see that it gives a reasonable (crude) approximation of the infidelity $1 - F_{M\tilde{M}}$ due to leakage (if also multiplied by the factor of $\frac{1}{4}$, it goes close to the top of oscillating blue and green solid lines). Such a good fit is somewhat surprising because in deriving Eq. (61) we assumed a fixed energy difference $E_0^{(c)} - E_2^{(c)} = \eta - 2\Delta$, while for large ε the difference of eigenenergies $E_{\overline{|0\rangle_\varepsilon}} - E_{\overline{|2\rangle_\varepsilon}}$ becomes significantly larger, e.g., 60.7 MHz for $\varepsilon/2\pi = 60$ MHz and 84.3 MHz for $\varepsilon/2\pi = 80$ MHz. Therefore, we would expect that Eq. (61) should significantly overestimate the actual leakage [note the sixth power of $(\eta - 2\Delta)$ in the denominator]. As we checked, Eq. (61) still works well because the nonadiabatic transition mainly accumulates during the lower half of the ramp, when $\varepsilon(t)$ is not too large.

The oscillations of the solid lines in Fig. 17 (leakage probabilities multiplied by $\frac{1}{4}$) are easily understandable. The nonadiabatic leakage occurs during both front and rear ramps, and the transition amplitudes are added with a nonzero phase due to the energy difference $E_{\overline{|0\rangle_\varepsilon}} - E_{\overline{|2\rangle_\varepsilon}}$, accumulated between the ramps. Thus, the oscillations are due to constructive or destructive interference of the leakage contributions from the two ramps. We have checked that the ε_m difference between the peaks in Fig. 17 is consistent with estimates based on the numerical increase of $E_{\overline{|0\rangle_\varepsilon}} - E_{\overline{|2\rangle_\varepsilon}}$ with ε. The oscillations in the probabilities of individual leakage channels lead to the oscillations of their sum, thus explaining oscillations of $1 - F_{M\tilde{M}}$ in Fig. 15 and oscillations of the overall CNOT gate infidelity $1 - F_{MU}$ in Fig. 14 at large ε_m.

Note that in Fig. 17 the leakage channels $\overline{|0, 0\rangle} \to \overline{|2, 0\rangle}$ and $\overline{|0, 1\rangle} \to \overline{|2, 1\rangle}$ (with a nonchanging state of the target qubit) have higher probabilities than for the leakage channels with a changing state of the target qubit. This is because $|\varphi_0| \ll \pi$ in the interesting range of ε_m (see Fig. 12), so the target-qubit state does not change much during the pulse when the control-qubit state is $|0\rangle$. In the opposite limit $|\pi - \varphi_0| \ll \pi$, we would expect all four leakage channels to have approximately the same strength and also would not
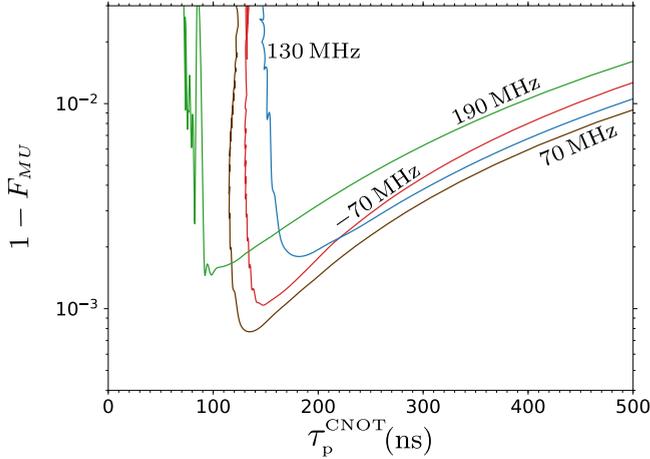
FIG. 18. Parametric plot for the CNOT gate infidelity $1 - F_{MU}$ versus the CNOT gate duration τ_p^{CNOT} (the running parameter is ε_m) for the detunings $\Delta/2\pi = -70, 70, 130$, and 190 MHz. We assume $g/2\pi = 3$ MHz, $\eta_c/2\pi = \eta_t/2\pi = 300$ MHz, $\omega_d = \omega_t^{c0}$, and $\tau_r/\tau_p = 0.3$.

FIG. 19. CNOT gate infidelity $1 - F_{MU}$ versus duration τ_p^{CNOT} (both are functions of ε_m) for $\Delta/2\pi = 70$ and 190 MHz and the drive frequency on-resonance with the target qubit for the control qubit either in the state $|0\rangle$ (solid lines, $\omega_d = \omega_t^{c0}$) or in the state $|1\rangle$ (dotted lines, $\omega_d = \omega_t^{c1}$) or exactly in between [dashed lines, $\omega_d = (\omega_t^{c0} + \omega_t^{c1})/2$]. The other parameters are the same as in Fig. 18.

expect significant oscillations (because the two ramps would mainly contribute to different channels).

Besides the detailed analysis of the leakage contribution $1 - F_{M\tilde{M}}$ for the detuning $\Delta/2\pi = 130$ MHz, we have also analyzed the leakage for other values of the detuning in Fig. 14. The case of $\Delta/2\pi = 190$ MHz is similar (the same dominating leakage channels) and Eq. (61) still works well; however, $|\eta - 2\Delta|$ is larger (80 MHz instead of 40 MHz), so the leakage becomes significant only at larger values of ε_m. Also, a very large value of the green line in Fig. 14 at $\varepsilon_m/2\pi \simeq 70$ MHz is due to an additional leakage channel $\overline{|0, 1\rangle} \rightarrow \overline{|1, 2\rangle}$, which is due to a resonance between these states. For the detuning $\Delta/2\pi$ of 70 MHz and -70 MHz (brown and red lines in Fig. 14), the main leakage is between states $\overline{|0\rangle_\varepsilon}$ and $\overline{|1\rangle_\varepsilon}$ of the control qubit (also due to nonadiabaticity during the ramps).

Thus, in this section we have shown that in our model the error budget of the CR gate consists of two main contributions: the imperfection of the unitary operation at small drive amplitudes and the leakage at large drive amplitudes. At the optimal drive amplitude, the infidelity is on the order of 10^{-3}. However, we did not take into account contributions from decoherence and also from possible problems caused by a strong drive of the control qubit (e.g., due to a resonance between an impurity and $E_{\overline{|0\rangle_\varepsilon}} - E_{\overline{|n\rangle_\varepsilon}}$ or $E_{\overline{|1\rangle_\varepsilon}} - E_{\overline{|n\rangle_\varepsilon}}$). Note that a strong leakage makes the CR gate operation impractical for the detuning Δ close to 0, $\eta_c/2$, η_c, $3\eta_c/2$, etc.

VII. DEPENDENCE ON PARAMETERS

A convenient way to present numerical results [35] is to parametrically plot the infidelity $1 - F_{MU}$ versus CNOT gate duration τ_p^{CNOT}, with both quantities being functions of the drive amplitude ε_m. Figure 18 presents such a plot for the results shown in Figs. 10 and 14 for the detuning values $\Delta/2\pi = -70, 70, 130$, and 190 MHz, while the other parameters are $g/2\pi = 3$ MHz, $\eta_c/2\pi = \eta_t/2\pi = 300$ MHz, $\omega_d = \omega_t^{c0}$, and $\tau_r/\tau_p = 0.3$. An increase of ε_m corresponds to
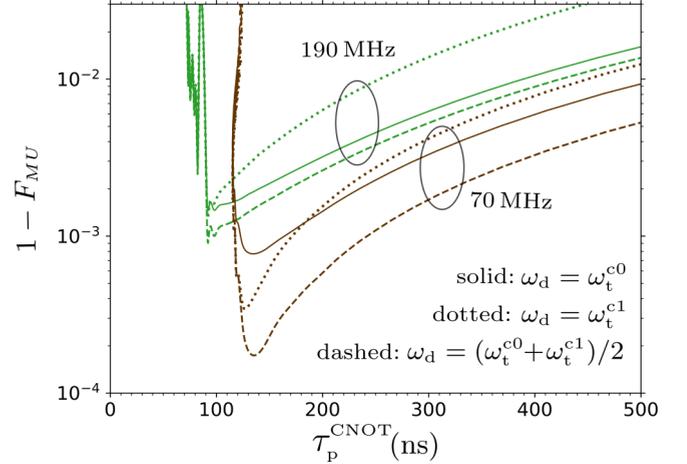
moving from right to left along the lines. The main observation is that for each Δ the line is very steep at the left, which naturally corresponds to a limit on decreasing the duration of the CNOT gate. Among the plotted lines, the line for the detuning of 190 MHz gives the shortest duration (consistent with our discussion in Sec. III C); however, it does not give the best fidelity.

In Fig. 19 we show the results for three slightly different drive frequencies for the qubit detunings of 70 MHz (brown lines) and 190 MHz (green lines). Besides using the drive frequency resonant with the target qubit when the control-qubit state is $|0\rangle$ (solid lines, $\omega_d = \omega_t^{c0}$ and $\delta = \omega_t - \omega_t^{c0}$), we also use the drive resonant with the target qubit when the control-qubit state is $|1\rangle$ (dotted lines, $\omega_d = \omega_t^{c1}$) and also show the case of the drive frequency exactly in between [dashed lines, $\omega_d = (\omega_t^{c0} + \omega_t^{c1})/2$]. Note that $\omega_t^{c1} - \omega_t^{c0} = \omega_{zz}$ is 127 and 200 kHz for detunings of 70 and 190 MHz, respectively [see Eq. (11)]. We see that this small change of the drive frequency practically does not affect the natural limit for the duration τ_p^{CNOT}; however, it may very significantly affect the infidelity. For example, for $\Delta/2\pi = 70$ MHz the minimum infidelity is 1.7×10^{-4} for $\omega_d = (\omega_t^{c0} + \omega_t^{c1})/2$, while it is 7.7×10^{-4} for $\omega_d = \omega_t^{c0}$. We have checked that this small change of the drive frequency practically does not affect the leakage $1 - F_{M\tilde{M}}$, but affects significantly the infidelity contribution $1 - F_{\tilde{M}U}$ due to the imperfection of the unitary operation. This is exactly what is expected from the analysis in Sec. VI, since the unitary imperfection is due to imperfect resonance between the drive and the target qubit. As we see from the numerical results, the drive frequency $\omega_d = (\omega_t^{c0} + \omega_t^{c1})/2$ typically gives a better optimized fidelity than for $\omega_d = \omega_t^{c0}$ or $\omega_d = \omega_t^{c1}$.

In all previous numerical plots, we assumed the qubit-qubit coupling $g/2\pi = 3$ MHz. Figure 20 shows the CNOT gate infidelity for $g/2\pi = 1.5, 3$, and 6 MHz, while $\Delta/2\pi = 190$ MHz and the other parameters are as in Fig. 18. As expected, for minima of the lines in Fig. 20, the CNOT gate
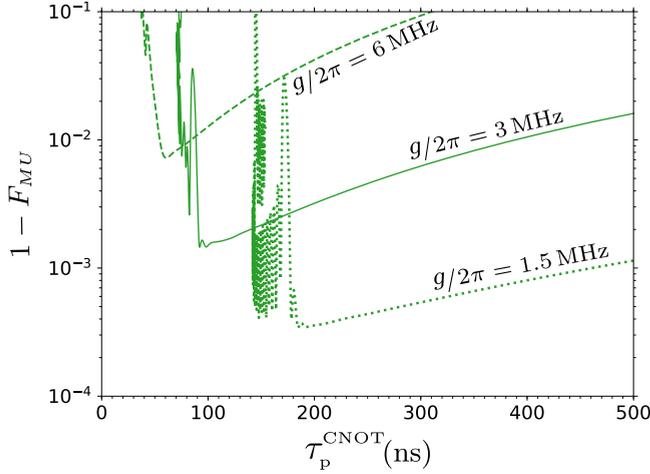
FIG. 20. CNOT gate infidelity $1 - F_{MU}$ versus duration τ_p^{CNOT} for $\Delta/2\pi = 190$ MHz and several values of the qubit-qubit coupling: $g/2\pi = 1.5, 3$, and 6 MHz. The other parameters are the same as in Fig. 18.

FIG. 22. CNOT gate infidelity $1 - F_{MU}$ versus duration τ_p^{CNOT} for $\Delta/2\pi = 190$ MHz and several values of the microwave crosstalk coefficient: $c_{\text{ct}} = 0$ (solid line), 0.05 (dotted line), 0.1 (dashed line), and 0.2 (dash-dotted line). The other parameters are the same as in Fig. 18. The lines are cut at the left for clarity.

duration decreases with increasing g as $\tau_p^{\text{CNOT}} \propto g^{-1}$; however, we see that the infidelity increases crudely as g^2 (consistent with the scaling $\omega_{zz} \propto g^2$, so that $\omega_{zz}\tau_p^{\text{CNOT}} \propto g$). An increase of g also increases undesired zz interaction of idling qubits [43] and therefore cannot be used as a simple way to reduce the CNOT gate duration in an experiment.

In Fig. 21 we numerically analyze the dependence on the relative duration of the pulse ramp by changing the ratio τ_r/τ_p in the pulse shape (48): $\tau_r/\tau_p = 0.1, 0.2, 0.3, 0.4$, and 0.5 (the other parameters are as in Fig. 18 with $\Delta/2\pi = 190$ MHz). For clarity, we do not show all oscillations on the left, cutting the lines at some maxima of the oscillations. We see that at the right side of the graph (long τ_p^{CNOT}) it is better to have the longest possible ramp, $\tau_r/\tau_p = 0.5$. However, in the optimal range of short τ_p^{CNOT} with still small infidelity, the line with
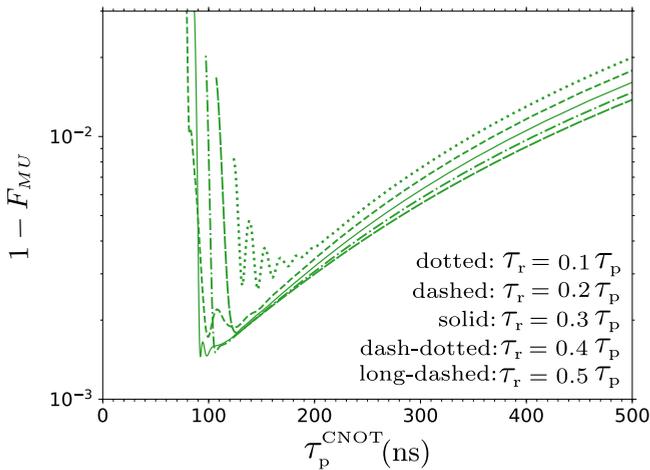

FIG. 21. CNOT gate infidelity $1 - F_{MU}$ versus duration τ_p^{CNOT} for $\Delta/2\pi = 190$ MHz and several values of the relative duration of the pulse ramps: $\tau_r/\tau_p = 0.1$ (dotted line), 0.2 (dashed line), 0.3 (solid line), 0.4 (dash-dotted line), and 0.5 (long-dashed line). The other parameters are the same as in Fig. 18. The lines are cut at the left for clarity.
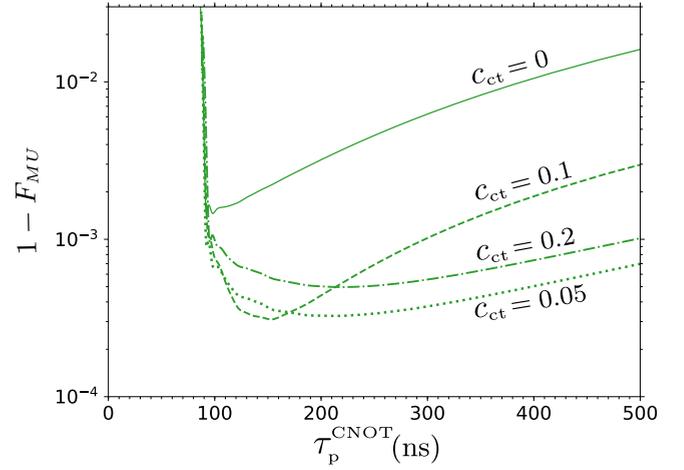
$\tau_r/\tau_p = 0.3$ shows the best performance, which can be understood as the following trade-off. For longer ramps and the same τ_p^{CNOT}, we need to use larger ε_m that increases the leakage (even though the ramp is smoother), while for shorter ramps and the same τ_p^{CNOT}, the leakage is also increased because the ramp is too short and consequently nonadiabatic (even though ε_m is smaller). Thus, the ramps should be sufficiently smooth, but it is still beneficial to have a flat part of the pulse.

Finally, let us discuss the effect of the microwave crosstalk, which is always present in experiments [20,22,25]. To include it, we need to add the crosstalk Hamiltonian

$$H_{\text{ct}} = \sum_{n,m} c_{\text{ct}} \varepsilon(t) \sqrt{m} |n, m\rangle \langle n, m - 1| + \text{H.c.}, \qquad (62)$$

which describes the microwave field applied directly to the target qubit. The crosstalk coefficient c_{ct} can in general be complex, and experimental results seem to indicate a complex c_{ct} [20,33]. However, for simplicity, here we assume a real c_{ct}. In the ideal and semianalytical theory, the crosstalk with real c_{ct} does not affect the results, except adding the phase $\varphi_{\text{ct}} = \int_0^{\tau_p} 2c_{\text{ct}}\varepsilon(t)dt$ to both φ_0 and φ_1. However, it affects the numerical results because the crosstalk changes the Bloch-sphere angle of tilt caused by the effect of ω_{zz}.

Figure 22 shows numerical results for several values of the (real) crosstalk coefficient c_{ct} (the parameters are as in Fig. 18 with $\Delta/2\pi = 190$ MHz). We see that the crosstalk improves fidelity. This improvement can be easily understood. The crosstalk does not affect leakage, but it decreases the imperfect-unitary contribution $1 - F_{M\tilde{M}}$ because now the drive detuning $\omega_d - \omega_t^{c1}$ should be compared with a larger Rabi frequency $\tilde{\varepsilon}_1 + c_{\text{ct}}\varepsilon$ instead of $\tilde{\varepsilon}_1$, and therefore the resulting tilt of the Bloch-sphere evolution is smaller (the same effect for the control-qubit state $|0\rangle$ if the signs of $\tilde{\varepsilon}_0$ and $c_{\text{ct}}\varepsilon$ coincide). Thus, somewhat unexpectedly, the microwave crosstalk can improve the gate fidelity. However, this improvement is significant only if the effect of ω_{zz} is significant. For example, if in an experiment the infidelity

is mainly determined by leakage and decoherence, then the crosstalk will not improve fidelity. Also, we recall that here we considered only real crosstalk coefficients, and the results for a complex c_{ct} may be significantly different.

The effect of the echo sequence is analyzed numerically in the Appendix. The main observation is that the echo-CR gate is typically longer than the basic CR gate for the same level of infidelity (even assuming instantaneous single-qubit gates).

VIII. CONCLUSION

In this paper we have analyzed analytically, semianalytically, and numerically the operation of the cross-resonance gate for superconducting qubits, focusing on using the CR gate to realize the CNOT gate operation [20,22,27]. Our model has been based on the Hamiltonian (1)–(8) and simple pulse shape (no echo sequence).

The analytical theory gives Eq. (18) for the effective drive amplitude $\tilde{\varepsilon}_n$ of the target qubit, which depends on the control-qubit state $|n\rangle$. However, the analytics can be used only for sufficiently small (and not too small) physical drive amplitude ε. The next-order analytics (Sec. III B) slightly widens the applicability range; however, it is still inapplicable in the practically interesting range of ε.

We have found that the speed of the CR gate and the compensating single-qubit rotations can be obtained very accurately by a sufficiently simple semianalytical theory, discussed in Sec. III C. This theory [Eqs. (26)–(29)] is based on solving a one-qubit time-independent Schrödinger equation. The semianalytical theory depends on only two parameters: dimensionless qubit-qubit detuning Δ/η_{c} (normalization is the control-qubit anharmonicity η_{c}) and dimensionless drive amplitude $\varepsilon/\eta_{\mathrm{c}}$. The dimensionless speed of the CR gate as a function of these two parameters is shown in Fig. 7. The CR gate speed cannot be increased indefinitely by increasing the drive amplitude ε; the maximum speed (as a function of Δ/η_{c}) and the corresponding drive amplitude are shown in Fig. 8. As follows from Figs. 7 and 8, the best operation of the CR gate is expected for the detuning within the range $0.5\eta_{\mathrm{c}} < \Delta < \eta_{\mathrm{c}}$.

The full numerical approach (discussed in Sec. IV) is mainly needed to calculate intrinsic fidelity of the CNOT-equivalent CR gate. The numerical results depend on the pulse shape, for which we use the simple cosine-ramp model (48). Most importantly, the infidelity $1 - F_{MU}$ has a minimum as a function of the midpulse drive amplitude ε_{m}. The minimum value depends on the detuning Δ (Figs. 14 and 18), and for typical parameters used in this paper the optimal infidelity is on the order of 10^{-3} (though it approaches 10^{-4} for some parameters and in principle the theoretical infidelity can be arbitrarily small for complicated pulse shapes [35]).

Our model does not include decoherence, so the error budget of the gate consists of two contributions: due to imperfect unitary operations and due to leakage [Eq. (57)]. The imperfect unitary dominates at small ε_{m}. The mechanism of this imperfection is related to the zz interaction of the qubits [Eq. (10)], which makes the target-qubit frequency dependent on the control-qubit state ($\omega_{\mathrm{t}}^{\mathrm{c1}} - \omega_{\mathrm{t}}^{\mathrm{c0}} = \omega_{zz}$) and therefore makes it impossible to use a drive frequency exactly on-resonance with the target qubit in both cases (for the control-qubit states $|0\rangle$ and $|1\rangle$). Because of this contribution

into the error budget, at small ε_{m} the gate infidelity is very sensitive to small changes of the drive frequency (Fig. 19) and the microwave crosstalk can improve fidelity (Fig. 22).

The other contribution to the error budget, which dominates at large drive amplitudes ε_{m}, is due to leakages. The main leakage is in the strongly driven control qubit and is caused by nonadiabaticity during the ramps of the pulse. Depending on the ratio Δ/η_{c} between the detuning and control-qubit anharmonicity, the main leakage mechanism can be either $|0\rangle \leftrightarrow |1\rangle$ (between the control-qubit states $|0\rangle$ and $|1\rangle$) or $|0\rangle \rightarrow |2\rangle$ or some other leakage channel. In particular, for $|\Delta| \ll \eta_{\mathrm{c}}$ there is a strong leakage $|0\rangle \leftrightarrow |1\rangle$ because the drive frequency is near-resonance with the control qubit. Similarly, for $|\Delta - \eta_{\mathrm{c}}/2| \ll \eta_{\mathrm{c}}$ there is a strong leakage $|0\rangle \rightarrow |2\rangle$ because in this case the energy difference between states $|2\rangle$ and $|0\rangle$ of the control qubit is near-resonance with doubled frequency of the drive. In some cases we also observed a significant leakage for the channel $|0, 1\rangle \rightarrow |1, 2\rangle$ (when it becomes near-resonance with the doubled drive frequency).

Strong leakage makes the CR gate operation impractical when detuning Δ is close to 0, $\eta_{\mathrm{c}}/2$, η_{c}, $3\eta_{\mathrm{c}}/2$, etc. Therefore, we would expect the best operation of the CR gate for the detuning within the range $0.6\eta_{\mathrm{c}} < \Delta < 0.8\eta_{\mathrm{c}}$ (see Figs. 7 and 8). Another reasonable range (which requires smaller drive amplitudes) is $0.2\eta_{\mathrm{c}} < \Delta < 0.3\eta_{\mathrm{c}}$.

Crudely, the CNOT-equivalent gate duration $\tau_{\mathrm{p}}^{\mathrm{CNOT}}$ (optimized over ε_{m}) is comparable to π/g, with a coefficient somewhat larger or smaller than 1, depending on Δ/η_{c} (see Fig. 18). The optimized intrinsic infidelity for a simple pulse shape (48) is crudely comparable to $(g/\eta_{\mathrm{c}})^2$ (as follows from scaling of ω_{zz} and $\tau_{\mathrm{p}}^{\mathrm{CNOT}}$), with a coefficient typically about 10^0–$10^{1.5}$, depending on the drive frequency, Δ/η_{c}, crosstalk, etc. We have found that to reduce the nonadiabatic leakage, the ramps of the pulse should occupy a significant fraction of the pulse duration; however, the flat part in the middle of the pulse should not be shortened to zero.

In this paper we have not focused on analyzing the echo sequence. However, some numerical results for the echo-CR gate are presented in the Appendix. In the echo sequence, there are four ramps instead of two; therefore, for the same total pulse duration, the ramps are shorter. This increases nonadiabatic leakages, so for the same infidelity, the echo-CR gate duration is typically longer than for the basic CR gate (Fig. 24).

A crude estimate of the CR gate infidelity contribution due to decoherence (within the computational subspace only) is given by Eq. (51). However, for large drive amplitudes, the bare state $|2\rangle$ and higher states of the control qubit are significantly occupied; their decoherence is typically much faster than in the computational subspace and therefore can significantly increase the gate infidelity. Another potential mechanism for experimental CR gate infidelity is due to two-level systems (TLSs) produced by impurities. A strong drive changes the effective energy levels in the control qubit by about 10^2 MHz and therefore the TLSs can become on-resonance with the control qubit during the ramp of the pulse. Moreover, large drive amplitude makes multiphoton processes easily possible and therefore TLSs can become resonant with combinational frequencies for many channels (similar to the situation for fast measurement of a qubit).

Note that in this paper we did not explicitly take into account the resonator, providing coupling between the qubits; instead we replaced it with an equivalent direct coupling. It would be interesting to repeat our numerical simulations, taking into account the resonator levels explicitly. However, we do not expect a significant modification of the results because the additional nonadiabatic effects should be suppressed by typically large detuning between the resonator and qubits. It would also be interesting to include decoherence in the simulations; however, it is not obvious what a proper model is for decoherence of higher levels, relevant to actual experimental situations.

We hope that some experimental group will carry out detailed measurements of the CNOT gate duration and error budget of the CR gate as functions of the drive amplitude, drive frequency, detuning, and pulse shape. It will be interesting (and important for the CR gate application in quantum computing) to compare experimental results with our theoretical findings.

ACKNOWLEDGMENTS

APPENDIX: EFFECT OF THE ECHO SEQUENCE

In the main text we focused on analysis of the basic CR gate. Here we analyze the echo-CR gate [20,26,28].

The echo-CR gate [26] is essentially a sequence of two same-shape basic CR gates with halved rotation angles $|\varphi_1 - \varphi_0| = \pi/2$, with the control qubit state flipped ($|0\rangle \leftrightarrow |1\rangle$) in between the two halves of the procedure and with the flipped phase ($\varepsilon \to -\varepsilon$) of the applied microwave drive for the second half of the procedure. In this case the x-rotation angles for the target qubit become $\varphi_1 = \varphi_1^{(1)} + \varphi_0^{(2)} = \varphi_1^{(1)} - \varphi_0^{(1)}$ and $\varphi_0 = \varphi_0^{(1)} + \varphi_1^{(2)} = \varphi_0^{(1)} - \varphi_1^{(1)}$, where the superscripts refer to the first or second half of the procedure. Consequently, $\varphi_1 = -\varphi_0$ and also $\varphi_1 - \varphi_0 = \pi$ when $\varphi_1^{(1)} - \varphi_0^{(1)} = \pi/2$. This is what is often called a $ZX_{\pi/2}$ gate [20,22,26]. Note that the relations $\varphi_0^{(2)} = -\varphi_0^{(1)}$ and $\varphi_1^{(2)} = -\varphi_1^{(1)}$ are because of the symmetry of the procedure and the phase shift by π for the drive $\varepsilon(t + \tau_p/2) = -\varepsilon(t)$, where $\tau_p/2$ is the time difference between the two halves of the procedure. Also note that the control qubit should be flipped back after the second pulse (though this flip can sometimes be compiled into the overall sequence of an algorithm).

Because of the symmetry, the echo sequence eliminates the need to apply the compensating x rotation of the target qubit (by $-\varphi_0$, as assumed in the main text). It also eliminates the need to apply compensating z rotation of the control qubit (by $\theta_0 - \theta_1 + \pi/2$, as assumed in the main text). This significantly reduces experimental complexity. Nevertheless, if we want to produce CNOT gate from the echo-CR gate $ZX_{\pi/2}$, we still need to apply additional x rotation by $\pi/2$ for the target qubit and z rotation by $\pi/2$ for the control qubit.

The analytical (Sec. III A) and semianalytical theories (Sec. III C) for the echo-CR gate do not change compared with the basic CR gate because of the symmetry: We can
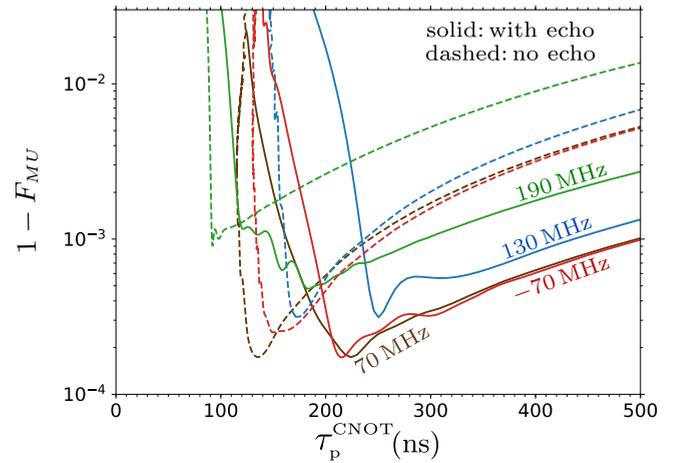

FIG. 23. CNOT gate infidelity $1 - F_{MU}$ versus duration τ_p^{CNOT} for the echo-CR gate (solid lines) and the basic CR gate (dashed lines). The detunings $\Delta/2\pi$ are 190 MHz (green lines), 130 MHz (blue lines), 70 MHz (brown lines), and -70 MHz (magenta lines). The parameters are the same as in Fig. 18, except we use $\omega_d = (\omega_t^{c0} + \omega_t^{c1})/2$. The lines are cut at the left for clarity.

just use the total gate duration with the same drive amplitude. In particular, Figs. 7 and 8 are still applicable without any change. However, numerical results are different because the ramps for the echo-CR gate of the same duration are shorter and correspondingly the leakage is typically bigger.

In the numerical simulations, for the first half of the procedure ($0 \leqslant t \leqslant \tau_p/2$) we use the pulse shape (48) with substitutions $\tau_p \to \tau_p/2$ and $\tau_r \to \tau_r/2$, while for the second half ($\tau_p/2 \leqslant t \leqslant \tau_p$) we use the inverted shape $\varepsilon(t) = -\varepsilon(t - \tau_p/2)$. Therefore, τ_p is the total pulse duration and τ_r is the total duration of the two front ramps (or two rear ramps). We assume ideal instantaneous π rotations of the control qubit (about x axis) at time moments $\tau_p/2$ and τ_p, and also ideal instantaneous rotations converting the $ZX_{\pi/2}$ gate into a CNOT gate: x rotation by $\pi/2$ for the target qubit and z rotation by $\pi/2$ for the control qubit. The fidelity is calculated using Eq. (42), which compares the actual gate with the ideal $ZX_{\pi/2}$ gate, i.e., Eq. (43) with $\theta_0 - \theta_1 = 0$ and $|\varphi_1| = \pi/2$ (the relation $|\varphi_1 - \varphi_0| = \pi$ is achieved by varying τ_p).

Figure 23 shows a comparison between the CR gate performances with and without echo sequence. The dashed lines (without echo) are the same as lines in Fig. 18, except now we use the drive frequency $\omega_d = (\omega_t^{c0} + \omega_t^{c1})/2$. The solid lines show the results for the echo-CR gate with the same parameters (colors correspond to particular detunings Δ). Most importantly, we see that the steep increase at the left for the solid lines occurs at larger τ_p^{CNOT}. This means that for the same infidelity $1 - F_{MU}$, the echo-CR gate is longer than the basic CR gate (even not including the durations of the additional π pulses). As mentioned above, this is because the leakage is a more severe problem for the echo-CR gate: Four ramps instead of two make their durations shorter, and this significantly increases nonadiabaticity during ramps.

Figure 24 summarizes our semianalytical and numerical results for the basic CR gate and the echo-CR gate. It shows the duration of the CNOT-equivalent gate τ_p^{CNOT} versus the detuning
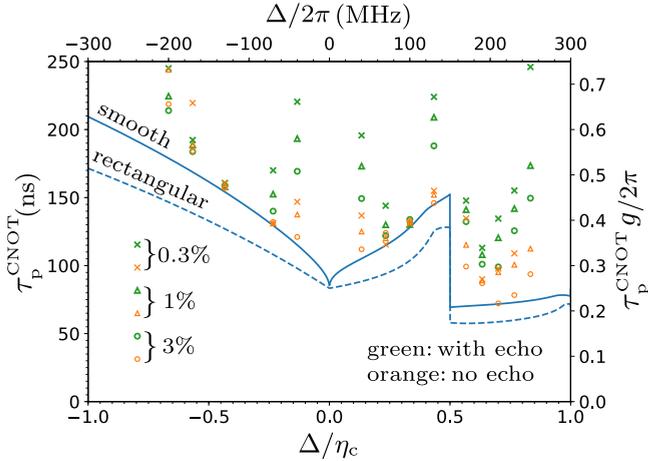
FIG. 24. Symbols show the numerical results for the CNOT gate duration $\tau_{\mathrm{p}}^{\mathrm{CNOT}}$ for the infidelity levels of 0.003 (crosses), 0.01 (triangles), and 0.03 (circles), for several values of the detuning Δ (horizontal axis): -200, -170, -130, -70, -40, 40, 70, 100, 130, 170, 190, 210, 230, and 250 MHz. Green (thicker) symbols are for the echo-CR gate; orange (thinner) symbols are for the basic CR gate. We use the parameters $g/2\pi = 3$ MHz, $\eta_{\mathrm{c}}/2\pi = \eta_{\mathrm{t}}/2\pi = 300$ MHz, and $\tau_{\mathrm{r}}/\tau_{\mathrm{p}} = 0.3$ (results on the dimensionless scales should not depend much on these parameters, except for $\tau_{\mathrm{r}}/\tau_{\mathrm{p}}$). The solid line shows the results of the semianalytical theory (optimized over the drive amplitude) for the same pulse shape. The dashed line shows the semianalytical results for a rectangular pulse (the same as in Fig. 8, but the vertical scale is inverted). Both dimensionless and dimensional scales are given for the axes.

Δ between the control and target qubits, using both the dimensionless and dimensional scales. We use the coupling $g/2\pi = 3$ MHz, anharmonicity $\eta_{\mathrm{c}} = \eta_{\mathrm{t}} = 300$ MHz, relative ramp duration $\tau_{\mathrm{r}}/\tau_{\mathrm{p}} = 0.3$, and drive frequency $\omega_{\mathrm{d}} = (\omega_{\mathrm{t}}^{\mathrm{c0}} + \omega_{\mathrm{t}}^{\mathrm{c1}})/2$. For different values of g, η_{c}, and η_{t}, the numerical results on the dimensionless scale are not expected to change significantly (semianalytical results would remain exactly the same).

The dashed line in Fig. 24 shows the CNOT gate duration (minimized over the drive amplitude ε) as follows from the semianalytical theory with a rectangular pulse shape. This line is the same as the solid line in Fig. 8 on the inverted scale (duration instead of speed). The solid line in Fig. 24 shows the optimized CNOT gate duration for the smooth pulse shape (48) with $\tau_{\mathrm{r}}/\tau_{\mathrm{p}} = 0.3$, also obtained semianalytically. We see that this duration is longer than for the rectangular pulse, but not by the naive factor $1/(1 - \tau_{\mathrm{r}}/\tau_{\mathrm{p}}) = 1.43$ (the ratio is significantly less than this factor because of the pulse-shape integration over the lines in Fig. 7, for which the region near maximum is most important). The semianalytical solid and dashed lines are the same for the echo-CR and basic CR gates.

The symbols in Fig. 24 show numerical durations of the CNOT-equivalent gates for three levels of the infidelity $1 - F_{MU}$, 0.3% (crosses), 1% (triangles), and 3% (circles), using the pulse shape (48) with $\tau_{\mathrm{r}}/\tau_{\mathrm{p}} = 0.3$. Green (thicker) symbols are for the echo-CR gate, while orange (thinner) symbols are for the basic CR gate. The symbols are presented for detunings $\Delta/2\pi$ of -200, -170, -130, -70, -40, 40, 70, 100, 130, 170, 190, 210, 230, and 250 MHz. As expected,

all symbols are above the solid line. For detunings of -130 and 100 MHz, the symbols are quite close to the solid line. This means that at these detunings, the leakage is not yet too strong for the near-optimal values of the drive amplitudes (e.g., compare lines in Figs. 10 and 14 for -70 MHz); the reason for a weak leakage is discussed later. In contrast, for some detunings, the symbols in Fig. 24 are much above the solid line. For example, for a detuning of 170 MHz, the semianalytics predicts a duration of 70 ns, while for 1% infidelity, the numerical results give 115 ns for the basic CR gate and 141 ns for the echo-CR gate. Such a big difference indicates a very significant leakage for the near-optimum drive amplitudes. As expected, in this case the echo-CR gate requires a significantly longer duration than the basic CR gate for the same level of infidelity.

Figure 24 can be used to estimate the range of best detunings, which provide close-to-shortest CNOT gate durations. For the basic CR gate with 1% infidelity, the best detuning range is crudely $0.6 < \Delta/\eta_{\mathrm{c}} < 0.65$, which for our parameters provides the fastest CNOT gate duration of around 90 ns (excluding single-qubit pulses). For the echo-CR gate with 1% infidelity, the best range is the same, and it provides the fastest duration of about 110 ns. Another reasonably well performing range of detunings is around $\Delta/\eta_{\mathrm{c}} \simeq 0.25$; for our parameters it gives a CNOT gate duration of about 120 ns for the basic CR gate and 130 ns for the echo-CR gate (for 1% infidelity). One more reasonable range is around $\Delta/\eta_{\mathrm{c}} \simeq -0.25$; the corresponding CNOT gate durations are 130 ns (without echo) and 150 ns (with echo).

The CNOT gate duration is typically longer if we require a smaller infidelity. Correspondingly, the crosses in Fig. 24 typically are significantly higher than triangles or circles. Moreover, in some cases (e.g., no echo, 250 MHz, and -200 MHz) there are no crosses because the infidelity level of 0.3% is never reached. However, for some detunings (e.g., 100 and -130 MHz) the duration is almost the same for the three considered levels of infidelity; as mentioned above, this indicates low leakage for relatively large drive amplitudes. There is even one weird case (no echo, 70 MHz), where the order of the symbols is reversed (this is because of the unusual behavior of the dashed brown line in Fig. 23 at the left, which relates to the increasing behavior of the orange line in Fig. 10).

As mentioned above, the difference between the symbols and the solid line in Fig. 24 is mainly determined by the leakage. We have checked that the main leakage channel for the detuning within the range $\frac{1}{3} \lesssim \Delta/\eta_{\mathrm{c}} \lesssim \frac{2}{3}$ is the transition $|0\rangle \to |2\rangle$ in the control qubit. This leakage channel is very strong when Δ/η_{c} is close to 0.5 (because of the resonance in the rotating frame), thus making impossible the practical operation of the CR gate at $\Delta/\eta_{\mathrm{c}} \simeq 0.5$. The leakage $|0\rangle \to |2\rangle$ becomes weaker for detunings farther away from this resonance point. For $\Delta/\eta_{\mathrm{c}} \gtrsim \frac{2}{3}$, the leakage channel $|1\rangle \to |2\rangle$ in the control qubit becomes more important (exact resonance at $\Delta/\eta_{\mathrm{c}} = 1$). Similarly, for $\Delta/\eta_{\mathrm{c}} \lesssim \frac{1}{3}$, the leakage channel $|0\rangle \leftrightarrow |1\rangle$ in the control qubit becomes more important (exact resonance at $\Delta/\eta_{\mathrm{c}} = 0$). Thus, the leakage is relatively low for the detuning Δ near $\frac{1}{3}\eta_{\mathrm{c}}$ or $\frac{2}{3}\eta_{\mathrm{c}}$. The trade-off between the lower leakage and shorter semianalytical durations determines the best detuning ranges in Fig. 24.

Note that besides the leakage channels $|0\rangle \rightarrow |2\rangle$, $|1\rangle \rightarrow |2\rangle$, and $|0\rangle \leftrightarrow |1\rangle$ for the control qubit, there are also important leakage channels which involve level $|2\rangle$ of the target qubit. For example, for detunings of 210 and 230 MHz, the leakage is dominated by a near-resonance between levels $|01\rangle$ and $|12\rangle$. Also, for detunings of -200, -170, and -130 MHz, the main leakage is due to a near-resonance between levels $|11\rangle$ and $|02\rangle$. Overall, the interplay between different leakage channels is rather complicated, leading to a rather complicated behavior of numerical results in Fig. 24.

Since the echo-CR gate is more affected by the leakage than the basic CR gate, a natural hypothesis is that its opera-

tion can be improved by using smoother ramps, in particular, by changing the relative duration of the front (and rear) ramps from the value $\tau_{\rm r}/\tau_{\rm p} = 0.3$ (used in Fig. 24) to the maximum possible value $\tau_{\rm r}/\tau_{\rm p} = 0.5$. To check this hypothesis, we have also simulated the echo-CR gate operation with $\tau_{\rm r}/\tau_{\rm p} = 0.5$, but the results were inconclusive: Sometimes this makes CNOT time slightly shorter, sometimes slightly longer. For example, for the detuning of 190 MHz (our shortest-duration point), changing $\tau_{\rm r}/\tau_{\rm p}$ from 0.3 to 0.5 increases the CNOT time by 4 ns for 0.3% infidelity and by 1 ns for 1% infidelity, but decreases it by 12 ns for 3% infidelity. Thus, crudely, we think that the pulse shape ratio $\tau_{\rm r}/\tau_{\rm p} = 0.3$ is still reasonable for the echo-CR gate.

[1] Y. Nakamura, Y. A. Pashkin, and J. S. Tsai, Nature (London) **398**, 786 (1999).

[2] C. Neill, P. Roushan, K. Kechedzhi, S. Boixo, S. V. Isakov, V. Smelyanskiy, A. Megrant, B. Chiaro, A. Dunsworth, K. Arya *et al.*, Science **360**, 195 (2018).

[3] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, and J. M. Gambetta, Nature (London) **549**, 242 (2017).

[4] K. S. Chou, J. Z. Blumoff, C. S. Wang, P. C. Reinhold, C. J. Axline, Y. Y. Gao, L. Frunzio, M. H. Devoret, L. Jiang, and R. J. Schoelkopf, Nature (London) **561**, 7723 (2018).

[5] S. S. Hong, A. T. Papageorge, P. Sivarajah, G. Crossman, N. Dider, A. M. Polloreno, E. A. Sete, S. W. Turkowski, M. P. da Silva, and B. R. Johnson, arXiv:1901.08035.

[6] S. Hacohen-Gourgy, L. S. Martin, E. Flurin, V. V. Ramasesh, K. B. Whaley, and I. Siddiqi, Nature (London) **538**, 491 (2016).

[7] K. J. Satzinger, Y. P. Zhong, H.-S. Chang, G. A. Peairs, A. Bienfait, M.-H. Chou, A. Y. Cleland, C. R. Conner, E. Dumur, J. Grebel *et al.*, Nature (London) **563**, 661 (2018).

[8] A. D. King, J. Carrasquilla, I. Ozfidan, J. Raymond, E. Andriyash, A. Berkley, M. Reis, T. M. Lanting, R. Harris, G. Poulin-Lamarre *et al.*, Nature (London) **560**, 456 (2018).

[9] P. Kurpiers, P. Magnard, T. Walter, B. Royer, M. Pechal, J. Heinsoo, Y. Salathe, A. Akin, S. Storz, J. C. Besse, S. Gasparinetti, A. Blais, and A. Wallraff, Nature (London) **558**, 264 (2018).

[10] Z. K. Minev, S. O. Mundhada, S. Shankar, P. Reinhold, R. Gutierrez-Jauregui, R. J. Schoelkopf, M. Mirrahimi, H. J. Carmichael, and M. H. Devoret, Nature (London) **570**, 200 (2019).

[11] Y. Masuyama, K. Funo, Y. Murashita, A. Noguchi, S. Kono, Y. Tabuchi, R. Yamazaki, M. Ueda, and Y. Nakamura, Nat. Commun. **9**, 1291 (2018).

[12] 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, Phys. Rev. A **76**, 042319 (2007).

[13] R. Barends, J. Kelly, A. Megrant, D. Sank, E. Jeffrey, Y. Chen, Y. Yin, B. Chiaro, J. Mutus, C. Neill, P. O'Malley, P. Roushan, J. Wenner, T. C. White, A. N. Cleland, and J. M. Martinis, Phys. Rev. Lett. **111**, 080502 (2013).

[14] Y. Chen, C. Neill, P. Roushan, N. Leung, M. Fang, R. Barends, J. Kelly, B. Campbell, Z. Chen, B. Chiaro *et al.*, Phys. Rev. Lett. **113**, 220502 (2014).

[15] I. Chiorescu, Y. Nakamura, C. J. P. M. Harmans, and J. E. Mooij, Science **299**, 1869 (2003).

[16] F. Yan, S. Gustavsson, A. Kamal, J. Birenbaum, A. P. Sears, D. Hover, T. J. Gudmundsen, D. Rosenberg, G. Samach, S. Weber, J. L. Yoder, T. P. Orlando, J. Clarke, A. J. Kerman, and W. D. Oliver, Nat. Commun. **7**, 12964 (2016).

[17] Y.-H. Lin, L. B. Nguyen, N. Grabon, J. San Miguel, N. Pankratova, and V. E. Manucharyan, Phys. Rev. Lett. **120**, 150503 (2018).

[18] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).

[19] R. Barends, J. Kelly, A. Megrant, A. Veitia, D. Sank, E. Jeffrey, T. C. White, J. Mutus, A. G. Fowler, B. Campbell *et al.*, Nature (London) **508**, 500 (2014).

[20] S. Sheldon, E. Magesan, J. M. Chow, and J. M. Gambetta, Phys. Rev. A **93**, 060302(R) (2016).

[21] C. Rigetti and M. Devoret, Phys. Rev. B **81**, 134507 (2010).

[22] J. M. Chow, A. D. Córcoles, J. M. Gambetta, C. Rigetti, B. R. Johnson, J. A. Smolin, J. R. Rozen, G. A. Keefe, M. B. Rothwell, M. B. Ketchen, and M. Steffen, Phys. Rev. Lett. **107**, 080502 (2011).

[23] G. S. Paraoanu, Phys. Rev. B **74**, 140504(R) (2006).

[24] P. C. de Groot, J. Lisenfeld, R. N. Schouten, S. Ashhab, A. Lupaşcu, C. J. P. M. Harmans, and J. E. Mooij, Nat. Phys. **6**, 763 (2010).

[25] J. M. Chow, J. M. Gambetta, A. D. Córcoles, S. T. Merkel, J. A. Smolin, C. Rigetti, S. Poletto, G. A. Keefe, M. B. Rothwell, J. R. Rozen, M. B. Ketchen, and M. Steffen, Phys. Rev. Lett. **109**, 060501 (2012).

[26] A. D. Córcoles, J. M. Gambetta, J. M. Chow, J. A. Smolin, M. Ware, J. Strand, B. L. T. Plourde, and M. Steffen, Phys. Rev. A **87**, 030301(R) (2013).

[27] A. Córcoles, E. Magesan, S. J. Srinivasan, A. W. Cross, M. Steffen, J. M. Gambetta, and J. M. Chow, Nat. Commun. **6**, 6979 (2015).

[28] M. Takita, A. D. Córcoles, E. Magesan, B. Abdo, M. Brink, A. Cross, J. M. Chow, and J. M. Gambetta, Phys. Rev. Lett. **117**, 210505 (2016).

[29] M. Ware, Ph.D. thesis, Syracuse University, 2015, Chap. 4.

[30] R. Naik, B. Mitchell, U. Baek, D. Dahlen, J. M. Kreikebaum, V. Ramasesh, M. Blok, and I. Siddiqi, APS March Meeting, Boston, 2019 [B. Am. Phys. Soc. 64, L29.6 (2019)].

[31] P. C. de Groot, S. Ashhab, A. Lupaşcu, L. DiCarlo, F. Nori, C. J. P. M. Harmans, and J. E. Mooij, New J. Phys. **14**, 073038 (2012).

[32] J. M. Gambetta, in *Quantum Information Processing: Lecture Notes of the 44th IFF Spring School*, edited by D. P. DiVincenzo (Forschungszentrum Jülich, Jülich, 2013), Chap. B4.

[33] E. Magesan and J. M. Gambetta, arXiv:1804.04073.

[34] D. Willsch, M. Nocon, F. Jin, H. De Raedt, and K. Michielsen, Phys. Rev. A **96**, 062302 (2017).

[35] S. Kirchhoff, T. Keßler, P. J. Liebermann, E. Assémat, S. Machnes, F. Motzoi, and F. K. Wilhelm, Phys. Rev. A **97**, 042348 (2018).

[36] S. E. Economou and E. Barnes, Phys. Rev. B **91**, 161405(R) (2015).

[37] J. L. Allen, R. Kosut, J. Joo, P. Leek, and E. Ginossar, Phys. Rev. A **95**, 042325 (2017).

[38] W. Magnus, Commun. Pure Appl. Math. **7**, 649 (1954); S. Blanes, F. Casas, J. Oteo, and J. Ros, Phys. Rep. **470**, 151 (2009).

[39] M. Khezri, Ph.D. thesis, University of California, Riverside, 2018, Appendix A.

[40] D. Sank, Z. Chen, M. Khezri, J. Kelly, R. Barends, B. Campbell, Y. Chen, B. Chiaro, A. Dunsworth, A. Fowler *et al.*, Phys. Rev. Lett. **117**, 190503 (2016), Supplemental Material.

[41] P. Zanardi and D. A. Lidar, Phys. Rev. A **70**, 012315 (2004).

[42] L. H. Pedersen, N. M. Møller, and K. Mølmer, Phys. Lett. A **367**, 47 (2007).

[43] A. Galiautdinov, A. N. Korotkov, and J. M. Martinis, Phys. Rev. A **85**, 042321 (2012).