Simulated randomized benchmarking of a dynamically corrected cross-resonance gate

R. K. L. Colmenar ⬤,[*] Utkan Güngördü ⬤, and J. P. Kestner

Department of Physics, University of Maryland Baltimore County, Baltimore, Maryland 21250, USA

We theoretically consider a cross-resonance (CR) gate implemented by pulse sequences proposed by Calderon-Vargas and Kestner [Phys. Rev. Lett. **118**, 150502 (2017)]. These sequences mitigate systematic error to first order, but their effectiveness is limited by one-qubit gate imperfections. Using additional microwave control pulses, it is possible to tune the effective CR Hamiltonian into a regime where these sequences operate optimally. This improves the overall feasibility of these sequences by reducing the one-qubit operations required for error correction. We illustrate this by simulating randomized benchmarking for a system of weakly coupled transmons and show that while this pulse sequence does not offer an advantage with the current state of the art in transmons, it does improve the scaling of CR gate infidelity with one-qubit gate infidelity.

I. INTRODUCTION

The ability to implement high-fidelity gates is a necessary requirement for creating a fully functional quantum information processor. To this end, fixed-frequency superconducting transmons [1] show great promise [2,3], as they have been used to theoretically and experimentally demonstrate one-qubit gates [4–8] with fidelities as high as 99.97% [8]. However, generating two-qubit entangling operations with similarly high fidelities remains a challenge. A standard approach to entangling fixed-frequency transmons is through the cross-resonance (CR) effect [9–12]. The CR effect can be observed in a system of two off-resonant fixed-frequency transmons with a small static coupling (e.g., through a quantum bus [13]). By irradiating one transmon at the transition frequency of the other, the coupling is modified by a factor whose magnitude is roughly proportional to the ratio of the microwave drive amplitude and the interqubit detuning.

Theoretical considerations have shown that the CR gate is significantly affected by systematic errors attributed to high-energy excitations of the weakly anharmonic transmon and to crosstalk induced by the CR microwave drive [14,15]. These processes give rise to unwanted terms in the CR effective Hamiltonian. This necessitates the use of control techniques such as composite pulse sequences [16,17] in order to isolate the desired entangling dynamics. In the case of a CR gate, such gate errors can be eliminated by a secondary control pulse on the target qubit which, in conjunction with pulse sequences, can result in CR gate fidelities exceeding 99% [18]. However, the pulse sequence used in Ref. [18] is not capable of addressing all coherent systematic errors to leading order.

In this paper, we analyze how well a different, recently discovered generic composite pulse sequence [19] would perform in the specific application of fixed-frequency transmons coupled via the CR effect as opposed to the conventional approach. This sequence inserts local π rotations between repeated application of an entangling gate to dynamically correct all coherent systematic errors in that entangling gate, but in practice there is a tradeoff between that reduction of error and the introduction of errors coming from the insertion of imperfect local π pulses. The purpose of this paper is to examine this tradeoff for the case of CR-gated transmons and determine the conditions for which there is a net benefit.

We theoretically simulate standard Clifford randomized benchmarking (RB) to assess the CR gate performance and show that, while there is no benefit to using the sequence of Ref. [19] with current transmon noise levels and single-qubit fidelities, as single-qubit fidelities improve the pulse sequence could provide better two-qubit RB fidelities than the currently used dynamical correction scheme.

II. DYNAMICAL ERROR CORRECTION VIA PULSE SEQUENCES

We begin by summarizing the formalism developed in Ref. [19]. We are interested in developing a protocol that allows us to dynamically correct coherent systematic error affecting an arbitrary two-qubit entangling gate. To this end, Ref. [19] presented a family of composite pulse sequences that are composed using repetitions of the nonlocal gate $(\theta)_{ab} = \exp[-i(\theta/2)\sigma_{ab}]$, where $a, b \in \{X, Y, Z\}$, which can be generated from any arbitrary two-qubit coupling along with appropriate one-qubit rotations [20,21]. In practice, the building block $(\theta)_{ab}$ may contain errors, which we only consider up to the leading order. Thus, we have

$$(\theta)_{ab} = \exp\left[-i\frac{\theta}{2}\sigma_{ab}\right]\left(I + i \sum_{i,j \in \{I,X,Y,Z\}} \epsilon_{ij}\sigma_{ij}\right), \quad (1)$$

where ϵ_{ij} is constant in time and is hereafter referred to as the error in the ij error channel. The pulse sequences have the

---

*ralphkc1@umbc.edu

general form,

$$\sigma_{\text{echo}}^{(n)}(\theta)_{ab}\sigma_{\text{echo}}^{(n)}\sigma_{\text{echo}}^{(n-1)}(\theta)_{ab}\sigma_{\text{echo}}^{(n-1)}\ldots\sigma_{\text{echo}}^{(1)}(\theta)_{ab}\sigma_{\text{echo}}^{(1)}$$

$$= \exp\left[-i\frac{\theta}{2}\sum_{l=1}^{n}\xi_l\sigma_{ab}\right]$$

$$\times\left\{I + i\sum_{i,j}\epsilon_{ij}\sigma_{ij}\sum_{m=1}^{n}\zeta_m^{ij}\exp\left[i\frac{\theta}{2}(\chi_{ij}-1)\sum_{l=1}^{m-1}\xi_l\sigma_{ab}\right]\right\},$$

$$(2)$$

where $\sigma_{\text{echo}}^{(l)}$ denotes a local π rotation of the form $\sigma_{cd}\equiv\sigma_c\otimes\sigma_d$ with $c,d\in\{I,X,Y,Z\}$ hereafter referred to as an echo pulse, and

$$\xi_l \equiv \begin{cases} +1, & \text{if}[\sigma_{\text{echo}}^{(l)},\sigma_{ab}]=0, \\ -1, & \text{if}\{\sigma_{\text{echo}}^{(l)},\sigma_{ab}\}=0, \end{cases} \qquad (3)$$

$$\zeta_m^{ij} \equiv \begin{cases} +1, & \text{if}[\sigma_{\text{echo}}^{(l)},\sigma_{ij}]=0, \\ -1, & \text{if}\{\sigma_{\text{echo}}^{(l)},\sigma_{ij}\}=0, \end{cases} \qquad (4)$$

$$\chi_{ij} \equiv \begin{cases} +1, & \text{if}[\sigma_{ij},\sigma_{ab}]=0, \\ -1, & \text{if}\{\sigma_{ij},\sigma_{ab}\}=0. \end{cases} \qquad (5)$$

We refer to a sequence containing n applications of the noisy entangling operation as a "length-n" sequence. To eliminate the effects of the ij error channel to leading order, we require

$$\sum_{m=1}^{n}\zeta_m^{ij}\exp\left[i\frac{\theta}{2}(\chi_{ij}-1)\sum_{l=1}^{m-1}\xi_l\sigma_{ab}\right]=0. \qquad (6)$$

To simplify this robustness condition, Ref. [19] considered two cases: One where only commuting errors are present ($\chi_{ij}=1$) and one where only anticommuting errors are present ($\chi_{ij}=-1$).

Let us first consider the case where we only have errors that commute with the entangling operation $(\theta)_{ab}$. In this case, Eq. (6) reduces to

$$\sum_{m=1}^{n}\zeta_m^{ij}=0. \qquad (7)$$

This immediately suggests that the robustness constraint is satisfied only for even values of n. The robustness condition in Eq. (7) for a length-2 sequence requires $\zeta_1^{ij}=-\zeta_2^{ij}$. Setting $\zeta_1^{ij}=1$ implies that the first echo pulse commutes with all the errors. Without loss of generality, we can choose the first pulse to be the identity operator for simplicity. Note that, in order to have a nonidentity operation, the second pulse must commute with σ_{ab}, i.e., $\xi_2=1$. The second pulse must also anticommute with all the errors in order to satisfy the robustness condition. If all potential commuting errors are present, this is not possible since there is no choice of $\sigma_{\text{echo}}^{(2)}$ that will simultaneously anticommute with all commuting errors, $[\sigma_{\text{echo}}^{(2)},\sigma_{ij}]=0\,\forall\,ij\ni[\sigma_{ij},\sigma_{ab}]=0$. A length-2 sequence can cancel four of the seven commuting error terms while producing an entangling operation, which may be all that is necessary in certain situations, but no more. (This can be quickly verified for any specific choice of σ_{ab} by simply listing all possibilities, but see Appendix A for the general proof.)

Nonetheless, with the exception of error in the ab channel itself, all errors that commute with σ_{ab} can be eliminated to first order by using two nested applications of a length-2 sequence, i.e., a length-4 sequence. For instance, the length-4 sequence,

$$\mathcal{U}^{(4)}[(\theta)_{ab}] \equiv (\theta)_{ab}\sigma_{aI}(\theta)_{ab}\sigma_{aI}\sigma_{cc}(\theta)_{ab}\sigma_{aI}(\theta)_{ab}\sigma_{aI}\sigma_{cc}$$

$$= (\theta)_{ab}\sigma_{aI}(\theta)_{ab}\sigma_{bc}(\theta)_{ab}\sigma_{aI}(\theta)_{ab}\sigma_{bc}$$

$$= \exp\left[-i\frac{4\theta}{2}\sigma_{ab}\right](I+\mathcal{O}(\epsilon^2)), \qquad (8)$$

where $\{\sigma_{cc},\sigma_{ab}\}=0$, eliminates all commuting error channels to first order except for the ab channel itself.

We now consider the complementary case where all the errors instead anticommute with the entangling operation $(\theta)_{ab}$. The robustness constraint in Eq. (6) becomes

$$\sum_{m=1}^{n}\zeta_m^{ij}\exp\left[-i\theta\sum_{l=1}^{m-1}\xi_l\sigma_{ab}\right]=0. \qquad (9)$$

Reference [19] showed that a nontrivial solution can be found when $n=5$, $\xi_l=1$, $\zeta_{(1,2,4,5)}=\pm1$, $\zeta_3=\mp1$, and $\theta=\theta_0\equiv\arccos[(\sqrt{13}-1)/4]\approx0.27\pi$. A set of echo pulses that correspond to these values are $\sigma_{\text{echo}}^{(1,2,4,5)}=I$ and $\sigma_{\text{echo}}^{(3)}=\sigma_{ab}$. Thus, a length-5 sequence that corrects all anticommuting errors to leading order is given by

$$\mathcal{U}^{(5)}[(\theta_0)_{ab}] \equiv (\theta_0)_{ab}(\theta_0)_{ab}\sigma_{ab}(\theta_0)_{ab}\sigma_{ab}(\theta_0)_{ab}(\theta_0)_{ab}$$

$$= \exp\left[-i\frac{5\theta_0}{2}\sigma_{ab}\right]\left[I+\mathcal{O}(\epsilon_{\text{anticomm}}^2)\right]. \qquad (10)$$

The resulting gate in Eq. (10) is nearly maximally entangling, but it is not locally equivalent to a CNOT. We can, however, construct a gate locally equivalent to a CNOT that can serve as a two-qubit Clifford group generator by using two applications of the dynamically corrected gate:

$$\mathcal{U}_{\text{Clif}_2} = \exp\left[-i\frac{\psi}{2}\sigma'\right]\mathcal{U}^{(5)}\exp\left[-i\frac{\phi}{2}\sigma'\right]$$

$$\times\,\mathcal{U}^{(5)}\exp\left[-i\frac{\psi}{2}\sigma'\right], \qquad (11)$$

where $\psi = 2\arctan[(\sqrt{-57+16\sqrt{13}})/(4-\sqrt{13}+2\sqrt{-7+2\sqrt{13}})]\approx0.36\pi$, $\phi = -2\arccos[-1/(2\sqrt{-14+4\sqrt{13}})]\approx-1.56\pi$, and $\sigma'\in\{\sigma_{IX},\sigma_{IY},\sigma_{IZ},\sigma_{XI},\sigma_{YI},\sigma_{ZI}\}$ such that $\{\sigma',\sigma_{ab}\}=0$.

It is possible to combine a length-2 (or length-4) sequence with a length-5 sequence in order to generate a length-10 (or length-20) sequence that can address both commuting and anticommuting error channels simultaneously. Furthermore, all of these pulses can also be combined with a BB1-like pulse sequence in order to correct the ab channel errors. First-order error in this channel can manifest from gate mistiming or fluctuations in the effective interqubit coupling, both of which result in over- or underrotation of the entangling operation. We refer the reader to Ref. [19] for a more detailed discussion.

Finally, we wish to emphasize that although the rest of this manuscript focuses on the application of the length-5 pulse sequence to fixed frequency transmon qubits, similar considerations apply in any other scenario having the key feature that the errors in the entangling gate anticommute

with the entangling operator. For example, in a silicon-based system of two double quantum dots (DQDs), each containing a single spin, coupled through a resonator [22]. The resonator is coupled to only one of the quantum dots, which makes the effective coupling dependent on the magnetic gradient within the DQD. Imperfections on the magnetic field gradient, which can be caused by either an anisotropy in the electron g tensor or misalignment of the local micromagnet, causes systematic commuting and anticommuting errors to emerge. These can be addressed by a length-5 sequence or a combination of a length-2 and a length-5 sequence depending on the severity of the error. However, from this point on we use numbers appropriate for the CR gated transmon case.

III. DYNAMICALLY CORRECTED CR GATE

We now apply the formalism we summarized in Sec. II to a CR gate. We consider a system of two fixed off-resonant transmons that are weakly coupled to a bus resonator. We then apply a constant-amplitude microwave driving field on one qubit, the control qubit, at the transition frequency of the other qubit, the target qubit. In the weak driving limit, a block-diagonal effective Hamiltonian for a CR gate can be perturbatively constructed using the Schrieffer-Wolff transformation [14]:

$$H_{\text{eff}}^{\text{CR}} = \frac{1}{2}h_{ZI}\sigma_{ZI} + \sum_{j\in\{X,Y,Z\}}\left(\frac{1}{2}h_{Ij}\sigma_{Ij} + \frac{1}{2}h_{Zj}\sigma_{Zj}\right), \quad (12)$$

where the expressions for h_{ij} in terms of the physical parameters are given in the Appendix of Ref. [14]. This approach differs from previous derivation of the CR Hamiltonian [9] in that it yields coherent error terms pertaining to higher-energy level leakage. We note that the Hamiltonian belongs in the embedding $\mathfrak{su}(2)\oplus\mathfrak{su}(2)\oplus\mathfrak{u}(1)\subset\mathfrak{su}(4)$. In particular, $\mathfrak{u}(1)$ is generated by σ_{ZI} which yields a factor in the time-evolution operator that can be removed by applying a local Z rotation on the first qubit. For this reason, we will ignore the effects of the σ_{ZI} term. The entangling term here is the σ_{ZX} term which, if factored out, yields

$$U(t) = \exp\left[-\frac{i}{2}t h_{ZX}\sigma_{ZX}\right]$$
$$\times \left[I + i\sum_{j\in\{X,Y,Z\}}(\epsilon_{Ij}\sigma_{Ij} + \epsilon_{Zj}\sigma_{Zj})\right], \quad (13)$$

where ϵ_{ij} can be calculated analytically up to a desired order using the Baker-Campbell-Hausdorff (BCH) formula. The pulse sequence building block, $(\theta)_{ZX}$, can then be obtained by setting $t = \theta/h_{ZX}$.

In experiments, the microwave drive acting on the control qubit often leaks into the target qubit which results in on-resonant crosstalk. This introduces large IX and IY terms in the effective Hamiltonian. Thus, in practice, the commuting errors are in the ZX and IX channels, while the anticommuting ones are in the IY, IZ, ZY, and ZZ channels [18]. Neglecting the ZX error channel for the moment, which is reasonable provided that the errors are static and the evolution time of the entangling gate is properly compensated through calibration, the result in Sec. II suggests that we

apply a length-2 sequence with a σ_{XZ} echo pulse to eliminate the IX channel. In addition, this choice of echo pulse can also eliminate any higher-order ZI channel errors that may accumulate due to the presence of large anticommuting error terms. Although alternatives such as σ_{XY} can serve the same purpose, we choose σ_{XZ} in order to take advantage of novel control methods that allow implementation of near-perfect virtual Z gates via abrupt phase modulation of the microwave control drive [8].

We note that Refs. [17,18] use an operationally distinct pulse sequence called an echoed CR (ECR) gate which has the same effect as the above length-2 sequence with an XZ echo pulse. The key operational difference between the ECR scheme and our length-2 sequence is the sign reversal in the entangling operation,

$$\text{ECR} \equiv \left(\frac{\pi}{4}\right)_{ZX}\sigma_{XI}\left(-\frac{\pi}{4}\right)_{ZX}\sigma_{XI}, \quad (14)$$

which can be implemented experimentally by reversing the signal of the microwave drive ($\Omega \to -\Omega$). Unlike in our length-2 scheme, a σ_{XI} echo pulse, which anticommutes with σ_{ZX}, is applied to avoid implementing a purely local gate. We show in Appendix B that this yields a mathematically equivalent pulse as the length-2 sequence in the case of a CR Hamiltonian. So in the remainder of our discussions, the length-2 sequence and the ECR gate are equivalent.

Another approach involves applying a secondary microwave pulse onto the target qubit so as to eliminate particular terms in the Hamiltonian [18]. This cancellation pulse is calibrated such that it eliminates the h_{IX}, h_{IY}, and h_{ZY} terms. Using the experimental parameters provided in Ref. [18], it can be verified numerically that the remaining terms have different scales, $h_{IZ} \ll h_{ZZ} \ll h_{ZX}$. Thus, the dominant source of remaining error comes from the h_{ZZ} term of the effective Hamiltonian, which translates to errors in the ZZ and IY channels. Although a length-2 sequence with an XZ echo pulse (i.e., ECR) can partially suppress these anticommuting errors, one can instead get complete first-order correction using the length-5 sequence of Eq. (10), $\mathcal{U}^{(5)}[U(\theta_0/h_{ZX})]$, with $ab = ZX$. Then, to obtain a two-qubit Clifford generator, we use Eq. (11) with $\sigma' = \sigma_{IZ}$, where we again make use of virtual Z gates. This yields an entangling Clifford gate compensated for all relevant coherent systematic errors to leading order.

It is important to keep in mind that the theory we summarized in Sec. II assumes that the echo pulses can be implemented perfectly. This is not the case in practice and echo pulse errors can be detrimental to the sequence's efficacy. Even though a longer and theoretically better sequence can be obtained by combining the length-2 and length-5 sequences, the resulting length-10 sequence requires more potentially noisy one-qubit gates to implement. So, depending on the level of one-qubit error, a length-2 or length-5 sequence can be more effective than a length-10 sequence. The supplemental material of Ref. [19] indicates that one-qubit gate errors on the order of, at most, 10^{-5} are required in order to build a CNOT out of a length-20 sequence with gate error below 10^{-3}. This may be very difficult to realize in the near future, which is why we focus our discussion to the length-2 and the length-5 sequence. In principle, a $\left(\frac{\pi}{2}\right)_{ZX}$ gate generated using the
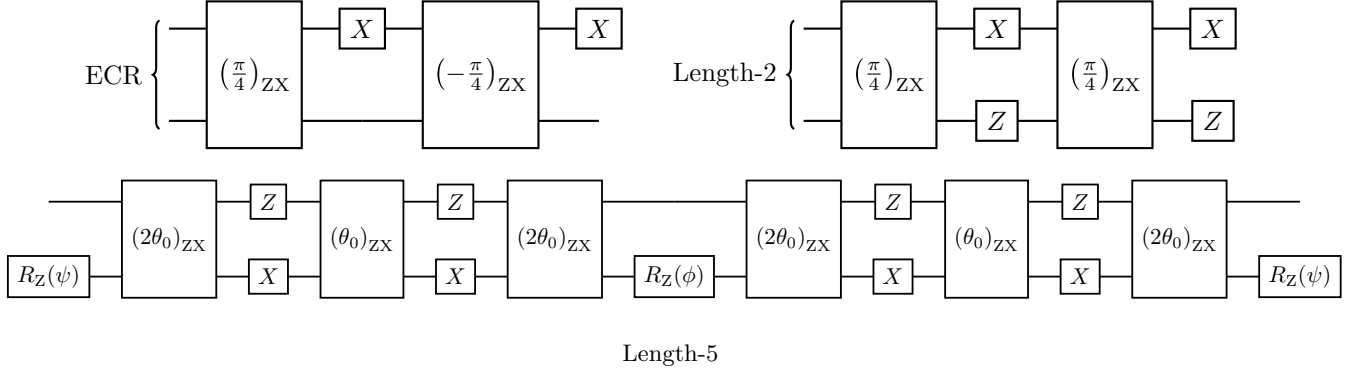
FIG. 1. Circuit diagrams for the two-qubit Clifford generator implemented using the ECR, the length-2 sequence, and the length-5 sequence [see Eq. (11)]. The top (bottom) line corresponds to the control (target) qubit. The X and Z gates are the usual Pauli gates and $R_Z(\theta)$ is a rotation about the Z axis by an angle θ. The generated Clifford gate is locally equivalent to a CNOT gate in all cases.

length-2 sequence requires four one-qubit gates, while $\mathcal{U}_{\text{Cliff}_2}$ requires 14 one-qubit gates. However, by taking advantage of virtual Z gates, we can reduce this to two and four physical one-qubit gates, respectively.

Finally, we note that the two-qubit Clifford gate generated by the length-2 sequence $(\frac{\pi}{2})_{ZX}$, and by the length-5 sequence U_{Cliff_2}, are different up to local Clifford rotations. In both cases, though, the local invariants of the resulting composite gate are equal to that of a CNOT gate. We present in Fig. 1 a circuit diagram for each of the cases we discussed.

IV. SIMULATED RANDOMIZED BENCHMARKING

To assess the performance of our dynamically corrected gate, we simulate standard Clifford randomized benchmarking (RB) [23] using $\{I, X_{\pm\frac{\pi}{2}}, X_{\pm\pi}, Z_{\pm\frac{\pi}{2}}, Z_{\pm\pi}, \mathcal{U}_{\text{Cliff}_2}\}$ as our generating set where, as an example, X_π denotes a π rotation about the X axis. We include quasistatic error in all local X rotations using the following noise model:

$$X_\theta \rightarrow \exp\left[-i\frac{\varepsilon}{2}\frac{r_x\sigma_X + r_y\sigma_Y + r_z\sigma_Z}{\sqrt{r_x^2 + r_y^2 + r_z^2}}\right]X_\theta, \qquad (15)$$

where $\{r_x, r_y, r_z\}$ are sampled uniformly from $[-1, 1]$, and ε is sampled from a normal distribution centered at 0 with standard deviation $\delta\theta$. We present in Appendix C an analytical formula that relates the one-qubit RB infidelity to $\delta\theta$. On the other hand, Z rotations are performed with no error, corresponding to the virtual gate method described in Ref. [8]. We also calculate the trace infidelity, which is not efficiently accessible in experiment, but is a less computationally demanding measure for theory, especially in the limit of very weak noise.

We use the experimental parameters reported in Ref. [18]: $\omega_1/2\pi = 5.114$ GHz, $\omega_2/2\pi = 4.914$ GHz, $\delta_1/2\pi = \delta_2/2\pi = -0.330$ GHz, $\Omega/2\pi = 60$ MHz, and $J/2\pi = 3.8$ MHz. The evolution time for the building block of the length-2 sequence $(\frac{\pi}{4})_{ZX}$ is $t = \pi/(4h_{ZX}) = 49.2$ ns, while that of the length-5 sequence $(\theta_0)_{ZX}$ is $t = \theta_0/h_{ZX} = 54$ ns. In order to simulate the effect of the cancellation pulse, we only include h_{IZ}, h_{ZX}, and h_{ZZ} in our effective Hamiltonian. Furthermore, we ignore relaxation errors in our simulations and focus solely on coherent systematic error. Each point

in the decay curve of our RB simulations are averaged over 1000 different sequences and noise realizations. The sequence length is set just enough to find a good fit for the survival probability function $ap^k + b$, where a, b, and p are fitting parameters and k is the sequence length. The results of our simulations are presented in Fig. 2.

Note that we do not include the initial portion of the decay from 100% down to around 90% (not plotted) in the fit, since there is nonexponential behavior there, particularly in the case of the length-5 decay. Nonexponential decay is commonly attributed to gate-dependent errors or low-frequency time-dependent noise [24], both of which are present in our simulations. Gate-dependent errors are present because we simulated RB with perfect Z gates but noisy X gates, as in experiments. Low-frequency noise effects appear because we perform each individual RB run with a fixed set of randomly generated noisy one-qubit Clifford group, again corresponding to the likely experimental case. We keep generating new sets of noisy Clifford gates until we exhaust all of our RB sequences. This builds statistics consistent with the distribution from which the noisy one-qubit gates are generated. A nonexponential decay is obtained by averaging over this ensemble of RB data.

We find that the length-2 sequence performs similarly to the length-5 sequence when the one-qubit RB error is set to 3×10^{-4} to match Ref. [8]. The length-2 sequence yields a fidelity of 99.7% and the length-5 sequence, which takes about five times as long (540 ns $+4t_{1Q}$ vs 98 ns $+2t_{1Q}$, where t_{1Q} denotes the echo pulse gate time), yields 99.8% [25]. However, if the one-qubit errors were reduced, we see that the length-5 sequence increasingly outperforms the length-2.

We can gain further insight by comparing the performance of the two pulse sequences in the limit where there are no one-qubit errors. For this task we use the trace fidelity since simulated randomized benchmarking requires simulating increasingly long sequences to obtain enough fidelity decay to fit as the one-qubit gate error is reduced. We rearrange Eq. (1) and isolate the error terms:

$$\delta U = \exp\left[\frac{i}{2}th_{ZX}\sigma_{ZX}\right]U(t) - I$$

$$= i\sum_{j\in\{I,X,Y,Z\}} \epsilon_{Ij}\sigma_{Ij} + \epsilon_{Zj}\sigma_{Zj}.$$
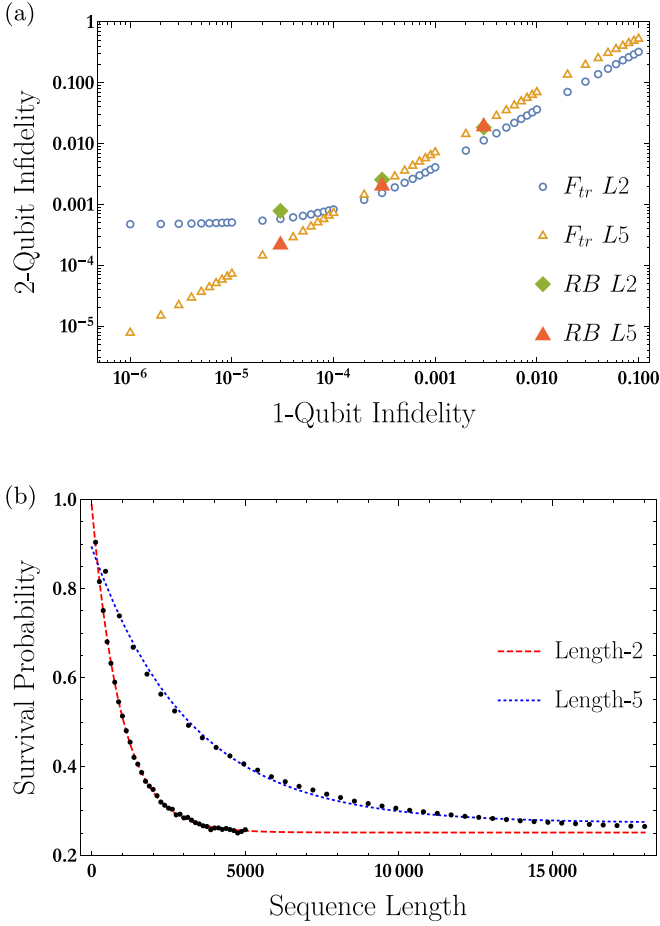
FIG. 2. (a) Solid symbols show randomized benchmarking (RB) two-qubit Clifford gate infidelity as a function of one-qubit RB infidelity for the length-2 and length-5 sequences. A current experimentally attainable value of one-qubit infidelity is 3×10^{-4} [8], corresponding to the RB points in the center of the plot. Open symbols show the trace fidelity. (b) A standard RB decay plot comparing the length-2 and length-5 sequences for the case where one-qubit infidelity is set to 3×10^{-5}.

We numerically calculate this for both schemes, assuming perfect one-qubit gates, and get

$$\delta U_{L2} = -2.4 \times 10^{-4} I + .015 i (\sigma_{IY} - \sigma_{ZZ})$$
$$+ 7.5 i \times 10^{-4} (\sigma_{IZ} + \sigma_{ZY}) + 3.5 i \times 10^{-4} \sigma_{ZX},$$
$$\delta U_{L5} = -2 i \times 10^{-5} \sigma_{IX} - 4.8 i \times 10^{-4} \sigma_{ZX},$$

where we have omitted any error terms with magnitudes below 10^{-5}. Since the gate infidelity is proportional to ϵ_{ij}^2, we see that the length-5 scheme can reach error rates on the order of 10^{-7} at best, while the length-2 scheme can only reach 10^{-4}. The much lower ideal infidelity of the length-5 sequence is because it cancels all the leading order errors in $U(t)$, whereas the length-2 sequence is not capable of eliminating the anticommuting error channels IY and ZZ. Of course, the actual performance of both sequences is highly dependent on the severity of the one-qubit gate imperfections, as is evident in Fig. 2, but one can see there that the trace infidelity of the length-5 sequence keeps decreasing with

decreasing single-qubit error while the length-2 sequence plateaus in the 10^{-4} region. Moreover, the crossing point where the length-5 is predicted to outperform the length-2 sequence occurs when the one-qubit infidelity is roughly 1×10^{-4}. This indicates that the length-5 sequence may be experimentally viable in the near future if one-qubit gate fidelities can be brought above 99.99%.

One caveat to this conclusion is that, as previously noted, the two-qubit Clifford generated from length-5 sequences is about five times slower than its length-2 counterpart. Thus, the length-5 sequence will suffer more from T_1 relaxation error, and its contribution to gate infidelity goes roughly as T_{gate}/T_1 [26]. One could consider increasing the CR drive amplitude Ω to speed up the gate. Numerical analysis of the CR gate in the strong driving regime indicates that the Hamiltonian terms that we considered as systematic error cannot be treated perturbatively when using a naïve cosine ramp model for the drive [15]. These terms can potentially be minimized while reducing the CR gate time by using pulse shapes derived from optimal control schemes which can result in CR gates under 100 ns [27]. Alternatively, one could also consider increasing the coupling between the qubits to speed up the gate, since the corresponding increase in σ_{ZZ} crosstalk due to unwanted excitations to higher energy transmon states would anyways be canceled by the length-5 sequence, but the issue is that the diminished qubit addressability would likely lower *one-qubit* echo pulse fidelities. However, at least the task of engineering a high-fidelity two-qubit gate is then effectively reduced to the problem of engineering high-fidelity local gates.

Without those sort of changes to speed up transmon operations, one has to consider in more detail the trade-off between reduction of coherent error by the length-5 sequence and increased incoherent error due to the longer gate time of the sequence. We aim to quantify this now by analyzing the effects of decoherence. For simplicity, we only consider dephasing and relaxation from the first transmon excited state to the ground state. Using the same parameters as above and setting the ground-state energy to zero, we simulate the evolution by a Lindblad master equation,

$$\dot{\rho} = -i[H_{\mathrm{eff}}^{\mathrm{CR}}, \rho] + \frac{1}{T_1} \sum_{j=1,2} \mathcal{D}[\sigma_j^-]\rho + \frac{1}{T_2^{\mathrm{CPMG}}} \mathcal{D}[\Pi_j^1]\rho, \tag{16}$$

where ρ is the density matrix, T_1 is the relaxation time of the two qubits, T_2^{CPMG} is the dephasing time measured via Carr-Purcell-Meiboom-Gill (CPMG) pulse sequence (used here as a lower bound on T_2), σ_j^- (Π_j^1) is the jth qubit's lowering operator (projection operator to the $|1\rangle$ state), and $\mathcal{D}$ is the damping superoperator,

$$\mathcal{D}[A]\rho = A\rho A^\dagger - \frac{1}{2} A^\dagger A \rho - \frac{1}{2} \rho A^\dagger A. \tag{17}$$

In order to focus on the role of decoherence, we assume that each one-qubit gate in the sequence is implemented without coherent or leakage errors and with an average gate time of 30 ns [18] during which the transmons can relax. The average two-qubit infidelity can then be calculated [28]:

$$\langle F \rangle = \frac{1}{16} \left[4 + \frac{1}{5} \sum_{i,j=I,X,Y,Z} \mathrm{tr}[U \sigma_{ij} U^\dagger \mathcal{M}(\sigma_{ij})] \right], \tag{18}$$
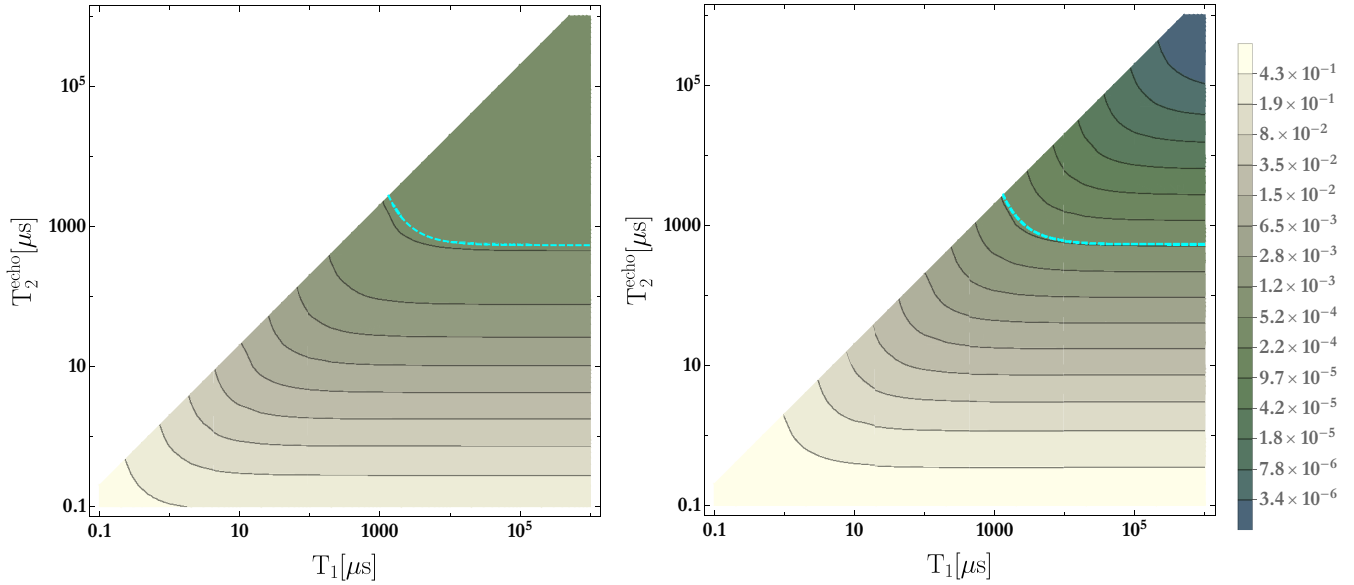
FIG. 3. A contour plot of two-qubit infidelity as a function of relaxation time T_1 and echoed dephasing time T_2^{CPMG} for the length-2 sequence (left) and the length-5 sequence (right), assuming one-qubit gates with no coherent or leakage errors and an average gate time of 30 ns. Unphysical regions where $T_2^{\text{CPMG}} > 2T_1$ are excluded. The dashed cyan contour indicates the crossing point where the two sequences have equal infidelities. The same colorscale is used for both panels.

where U is the ideal unitary time-evolution operator, $\mathcal{M}$ is a trace-preserving linear map, and $\sigma_{ij} = \sigma_i \otimes \sigma_j$ are the 15 nonidentity Kronecker products of Pauli matrices. We plot the results in Fig. 3.

The current state-of-the-art transmons can achieve average coherence times of $T_1 = 0.23$ ms and $T_2^{\text{CPMG}} = 0.38$ ms [29]. For those values, as opposed to the case of purely coherent error considered in Fig. 2, Fig. 3 indicates that even in the absence of one-qubit coherent gate error, the length-5 sequence does not outperform the length-2 sequence due to the effects of incoherent error over the longer gate time. However, at increased coherence times of $T_1, T_2^{\text{CPMG}} \approx 1.6$ms, the fidelity of the length-5 sequence begins to surpass that of the length-2 sequence. For $T_1, T_2^{\text{CPMG}} \gg 1$ ms the performance of the length-2 sequence plateaus at 3.8×10^{-4}, while the length-5 sequence continues to show improvement until it also eventually plateaus at roughly 3×10^{-7}, consistent with what we observed in Fig. 2(a).

Thus, while the length-5 sequence is not currently practical, given the rate of improvement in coherence times in recent years (roughly an order of magnitude every three years) [30] and the amount of attention being devoted to this task [31], it is reasonable to expect the length-5 sequence to become a viable option in the near future.

V. SUMMARY

We have shown how to dynamically correct a CR gate using a recently developed composite pulse sequence and we theoretically simulated a randomized benchmarking protocol for an experimentally accessible comparison of its performance with the standard ECR scheme, which is equivalent to a length-2 pulse sequence. The application of a cancellation pulse onto the target qubit eliminates a significant

amount of coherent systematic error from the effective CR Hamiltonian. The length-2 sequence cannot address all of the remaining dominant errors, all of which anticommute with the entangling operation, but the newly developed length-5 sequence can, at the cost of additional local rotations and a slower entangling gate. We find that both sequences perform similarly against coherent error when using one-qubit gates with currently achievable fidelities. However, we also show that the length-5 sequence performance could scale much better than the length-2 sequence when one-qubit gates are improved.

The pulse sequences we presented can be easily extended to systems with more than two fixed transmon qubits. Ideally, any given pair of control and target qubit must be decoupled from the remaining idle qubits when generating a two-qubit operation. In cases where more than two qubits share the same bus, the static always-on coupling can lead to spurious Z interactions with one or more of the idle qubits. One work-around to this problem is by performing the CR operation on the control and target qubit while decoupling the idle qubits through Hahn-echo-like pulses [32,33]. We can apply the same idea to the length-2 and length-5 sequence in order to simultaneously address entangling gate errors within the control-target subspace and spurious errors with the idle qubits. However, the additional echo pulses required to implement this makes the sequence even longer than it already is.

The long gate time of the length-5 sequence already makes it impractical for current coherence times, as the improvement the sequence is designed to produce against coherent errors is outweighed by the increased susceptibility to incoherent errors. However, once coherence times are increased beyond 1 ms, the sequence we have presented in this paper will become useful for increasing overall two-qubit gate fidelity.

ACKNOWLEDGMENTS

APPENDIX A: ANALYSIS OF THE LENGTH-2 SEQUENCE

In Sec. II we noted that if all potential commuting errors are present then there exists no $\sigma_{\text{echo}} = \sigma_{cd}$ with $c, d \in \{I, X, Y, Z\}$ that can satisfy

$$\{\sigma_{\text{echo}}, \sigma_{ij}\} = 0 \, \forall \, ij \ni [\sigma_{ij}, \sigma_{ab}] = 0. \tag{A1}$$

To prove this, we begin by making the observation that there exists a maximal embedding $\mathfrak{a} \oplus \mathfrak{b} \subset \mathfrak{su}(4)$ [34], where "$\oplus$" implies commutation of elements between the respective subalgebras. Let us take $\mathfrak{a} = \mathfrak{u}(1)$ with σ_{ab} as its generator. This means all the generators of $\mathfrak{b}$, which we denote as $\boldsymbol{b} = \{b_1, b_2, \ldots b_n\}$, commute with σ_{ab}. Thus, not only do error channels that commute with σ_{ab} belong in this embedding, but the echo pulse also must belong to the corresponding group embedding since it also needs to commute with σ_{ab} in order for the sequence to produce a nonidentity operation. Without losing any generality, we can partition $\boldsymbol{b}$ into two subsets as

$$\boldsymbol{b} = \{\overbrace{b_1, b_2, \ldots}^{\{\sigma_{ij}\}} | \overbrace{\ldots b_n}^{\boldsymbol{b} \backslash \{\sigma_{ij}\}}\},$$

where the left partition contains all the commuting errors that are relevant in the system while the right partition is the coset containing the remaining elements of the generating set of the subalgebra. A σ_{echo} which anticommutes with everything on the left partition and commutes with σ_{ab} can only exist in the coset. If all the possible commuting errors are present, then all the elements of $\boldsymbol{b}$ must belong to the left partition, which leaves no possibility for σ_{echo}. In other words, if all commuting errors are present, then there exists no σ_{echo} that can satisfy Eq. (A1) which proves our claim.

When $\mathfrak{b}$ is a semisimple Lie algebra, not all the elements of the coset $\boldsymbol{b} \backslash \{\sigma_{ij}\}$ will necessarily anticommute with all elements of $\{\sigma_{ij}\}$. Nonetheless, if the coset happens to contain a σ_{echo} which anticommutes with all the relevant error channels σ_{ij}, it can be used for error correction. We now show that this is actually the case for the length-2 sequence. But before proceeding further, we first remark that the construction in Sec. II relies on the commutation and anticommutation relations of two-qubit Pauli operators. For this reason, we restrict ourselves to embeddings which contain the subalgebra $\mathfrak{b}$ with spinor representation: $\mathfrak{so}(4) \cong \mathfrak{su}(2) \oplus \mathfrak{su}(2)$ [35]. Given our choice $\mathfrak{a} = \text{span}(\sigma_{ab})$ as the $\mathfrak{u}(1)$ subalgebra, there are two choices for $\mathfrak{b}$.

In the first case with $a = b$, all the elements of the generating set $\{\sigma_{mI}, \sigma_{Im}, \sigma_{nn}, \sigma_{pp}, \sigma_{mp}, \sigma_{pm}\}$, where $m = a$, n, and p are mutually distinct and arranged cyclically (e.g., $m = Z, n = X, p = Y$), commute with σ_{aa}. We can define the generators of the commuting $\mathfrak{su}(2)$ subalgebras as $\sigma_{\widetilde{X}}^{\pm} \equiv (\sigma_{mI} \pm \sigma_{Im})/2$, $\sigma_{\widetilde{Y}}^{\pm} \equiv (\sigma_{np} \pm \sigma_{pn})/2$, and $\sigma_{\widetilde{Z}}^{\pm} \equiv (\sigma_{pp} \mp \sigma_{nn})/2$.

In the second case with $a \neq b$, all the elements of the generating set $\{\sigma_{mI}, \sigma_{In}, \sigma_{nm}, \sigma_{np}, \sigma_{pm}, \sigma_{pp}\}$ commute with σ_{ab}, where now we have $m = a$ and $n = b$. The generators can

be defined as $\sigma_{\widetilde{X}}^{\pm} \equiv (\sigma_{mI} \pm \sigma_{In})/2$, $\sigma_{\widetilde{Y}}^{\pm} \equiv (\sigma_{np} \mp \sigma_{pm})/2$, and $\sigma_{\widetilde{Z}}^{\pm} \equiv (\sigma_{pp} \pm \sigma_{nm})/2$.

In either case, any error channel or echo pulse lies in the subspace spanned by the "+" and "−" generators (e.g., $\sigma_{aI} = \sigma_{\widetilde{X}}^{+} + \sigma_{\widetilde{X}}^{-}$). Therefore, any given echo pulse can only anticommute with errors belonging to a different subspace. As an example, since the σ_{aI} echo pulse belongs to the subspace spanned by the $\sigma_{\widetilde{X}}$ generators, then only errors that belong in the $\sigma_{\widetilde{Y}}$ and $\sigma_{\widetilde{Z}}$ subspaces can be eliminated by a length-2 sequence. Therefore, if all the present commuting errors belong to at most two subspaces only, then the length-2 sequence is sufficient for fixing the errors to first order. The transmon qubit in the main text falls in this category.

A more precise statement of our initial claim is that the length-2 sequence is not capable of correcting errors from all three subspaces. However, placing the initial length-2 sequence inside another length-2 sequence which uses an echo pulse that anticommutes with the initial one allows us to eliminate errors from all three subspaces simultaneously. If we use σ_{aI} for our first sequence's echo pulse, the second echo pulse must be in the $\sigma_{\widetilde{Y}}$ or $\sigma_{\widetilde{Z}}$ subspace in order to satisfy the robustness condition. Clearly, though, the $\mathfrak{u}(1)$ term cannot be corrected by a length-2 sequence since it commutes with every allowable echo.

APPENDIX B: EQUIVALENCE OF THE LENGTH-2 SEQUENCE AND THE ECR SCHEME

In this section we will show that the ECR scheme is mathematically equivalent to a length-2 sequence with a σ_{XZ} echo pulse. We begin by noting that the h_{IZ} and h_{ZZ} terms in the effective Hamiltonian of a CR gate are proportional to Ω^2, whereas the h_{ZX} and h_{IX} are only proportional to Ω. Thus, in the absence of noise, the evolution can be generally expressed as

$$U(\Omega, t) = \exp[-it(\Omega^2(a\sigma_{IZ} + b\sigma_{ZZ}) + \Omega(c\sigma_{ZX} + d\sigma_{IX}))], \tag{B1}$$

where a, b, c, and d are given in Appendix C of Ref. [14]. Using the pulse sequence in Eq. (14), we have

$$U(\Omega, t)\sigma_{XI}U(-\Omega, t)\sigma_{XI}.$$

The change from $\Omega \to -\Omega$ flips the sign of terms that are linearly proportional to Ω. Furthermore, the two σ_{XI} surrounding $U(-\Omega, t)$ flip the sign of terms in the exponential which anticommutes with σ_{XI}. Thus, the cumulative effect of this sequence is

$$\begin{aligned} &U(\Omega, t)\sigma_{XI}U(-\Omega, t)\sigma_{XI} \\ &= \exp[-it(\Omega^2(a\sigma_{IZ} + b\sigma_{ZZ}) + \Omega(c\sigma_{ZX} + d\sigma_{IX}))] \\ &\quad \times \exp[-it(\Omega^2(a\sigma_{IZ} - b\sigma_{ZZ}) - \Omega(-c\sigma_{ZX} + d\sigma_{IX}))]. \end{aligned} \tag{B2}$$

On the other hand, a length-2 sequence with a σ_{XZ} echo pulse yields

$$\begin{aligned} &U(\Omega, t)\sigma_{XZ}U(\Omega, t)\sigma_{XZ} \\ &= \exp[-it(\Omega^2(a\sigma_{IZ} + b\sigma_{ZZ}) + \Omega(c\sigma_{ZX} + d\sigma_{IX}))] \\ &\quad \times \exp[-it(\Omega^2(a\sigma_{IZ} - b\sigma_{ZZ}) + \Omega(c\sigma_{ZX} - d\sigma_{IX}))], \end{aligned} \tag{B3}$$

where now we flip the sign of terms in the second exponential which anticommute with σ_{XZ}. We see that in either case the final products are exactly equivalent.

APPENDIX C: ANALYTICAL EXPRESSION FOR ONE-QUBIT CLIFFORD RB FIDELITY

We now present an analytical expression for the Clifford RB fidelity of a one-qubit gate under the error model given in Eq. (15). In summary, the goal of Clifford RB is to provide a simple, robust, and scalable method for benchmarking the full set of Clifford gates through randomization. The randomization process, also known as twirling, produces a depolarizing channel whose average fidelity can be modeled and experimentally measured. Since the average fidelity of a quantum operation is invariant under the twirling process [36,37], the measured fidelity is representative of the original untwirled operation. For a more detailed discussion of Clifford RB, we refer the reader to Ref. [23].

The key to creating a depolarizing error channel lies in the fact the uniform probability distribution over the Clifford group $\mathcal{C}$ comprises a unitary two-design. By definition, this gives the twirling condition,

$$\frac{1}{|\mathcal{C}|} \sum_{i=1}^{|\mathcal{C}|} (\mathcal{C}_i^\dagger \Lambda \mathcal{C}_i)(\rho) = \int_{U(d)} (U^\dagger \Lambda U)(\rho) dU, \quad \text{(C1)}$$

where $\mathcal{C}_i$ are elements of the Clifford group, Λ is an arbitrary quantum channel acting on the system, and the integral is taken with respect to the Haar measure on $U(2^n)$ with n being the number of qubits. The integral in Eq. (C1) produces a unique depolarizing channel Λ_d with the same average fidelity as Λ [36,37]. The depolarizing channel is modeled by

$$\Lambda_d(\rho) = p\rho + (1-p)\frac{I}{2^n},$$

whose fidelity (as well as Λ's) is given by

$$\mathcal{F} = p + \frac{1-p}{2^n}.$$

To estimate the average fidelity of one qubit under the error model given in Eq. (15), we simply replace Λ accordingly and evaluate the sum:

$$\frac{1}{|\mathcal{C}|} \sum_{i=1}^{|\mathcal{C}|} \left(\mathcal{C}_i^\dagger \exp\left[-i\frac{\varepsilon}{2}\hat{\mathbf{r}} \cdot \vec{\sigma} \right] \mathcal{C}_i \right)(\rho)$$

$$= \frac{1}{|\mathcal{C}|} \sum_{i=1}^{|\mathcal{C}|} \mathcal{C}_i^\dagger \exp\left[-i\frac{\varepsilon}{2}\hat{\mathbf{r}} \cdot \vec{\sigma} \right] \mathcal{C}_i \rho \mathcal{C}_i^\dagger \exp\left[i\frac{\varepsilon}{2}\hat{\mathbf{r}} \cdot \vec{\sigma} \right] \mathcal{C}_i$$

$$= \frac{I}{2} + \frac{1 + 2\cos(\theta)}{3} \frac{\hat{\rho} \cdot \vec{\sigma}}{2}. \quad \text{(C2)}$$

Since $\rho = \frac{I + \hat{\rho}\cdot\vec{\sigma}}{2}$, then we must have

$$p = \frac{1 + 2\cos(\theta)}{3}.$$

Thus, the average Clifford RB fidelity is

$$\mathcal{F} = \frac{2 + \cos(\theta)}{3}. \quad \text{(C3)}$$

In the main text we noted that we used virtual gates in our simulations. This means that any Z gates in the Clifford group are treated as noiseless gates. Thus, we can approximate the fidelity when using virtual Z gates by appropriately weighting the fidelity of the 24 one-qubit Clifford gates:

$$\mathcal{F}_{VZ} = \frac{20\overline{\mathcal{F}} + 4}{24} = \frac{13 + 5\cos(\theta)}{18}, \quad \text{(C4)}$$

where we assumed that we had four noiseless gates $(I, Z, Z_{\pm\frac{\pi}{2}})$. Assuming a Gaussian noise model with a standard deviation $\delta\theta$, we can average over noise realizations and get

$$\overline{\mathcal{F}_{VZ}} = \frac{1}{\sqrt{2\pi\delta\theta^2}} \int_{-\infty}^{\infty} \exp\left[\frac{-\theta^2}{2\delta\theta^2} \right] \frac{13 + 5\cos(\theta)}{18} d\theta$$

$$= \frac{13 + 5\exp\left[-\frac{\delta\theta^2}{2} \right]}{18}. \quad \text{(C5)}$$

[1] 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).

[2] J. A. Schreier, A. A. Houck, J. Koch, D. I. Schuster, B. R. Johnson, J. M. Chow, J. M. Gambetta, J. Majer, L. Frunzio, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, Phys. Rev. B **77**, 180502(R) (2008).

[3] A. A. Houck, J. A. Schreier, B. R. Johnson, J. M. Chow, J. Koch, J. M. Gambetta, D. I. Schuster, L. Frunzio, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, Phys. Rev. Lett. **101**, 080502 (2008).

[4] 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).

[5] S. Sheldon, L. S. Bishop, E. Magesan, S. Filipp, J. M. Chow, and J. M. Gambetta, Phys. Rev. A **93**, 012301 (2016).

[6] F. Motzoi, J. M. Gambetta, P. Rebentrost, and F. K. Wilhelm, Phys. Rev. Lett. **103**, 110501 (2009).

[7] J. M. Chow, L. DiCarlo, J. M. Gambetta, F. Motzoi, L. Frunzio, S. M. Girvin, and R. J. Schoelkopf, Phys. Rev. A **82**, 040305(R) (2010).

[8] D. C. McKay, C. J. Wood, S. Sheldon, J. M. Chow, and J. M. Gambetta, Phys. Rev. A **96**, 022330 (2017).

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

[10] 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).

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

[12] J. Li, K. Chalapat, and G. S. Paraoanu, Phys. Rev. B **78**, 064503 (2008).

[13] J. Majer, J. M. Chow, J. M. Gambetta, J. Koch, B. R. Johnson, J. A. Schreier, L. Frunzio, D. I. Schuster, A. A.

Houck, A. Wallraff, A. Blais, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, Nature (London) **449**, 443 (2007).

[14] E. Magesan and J. M. Gambetta, Phys. Rev. A **101**, 052308 (2020).

[15] V. Tripathi, M. Khezri, and A. N. Korotkov, Phys. Rev. A **100**, 012301 (2019).

[16] L. M. K. Vandersypen and I. L. Chuang, Rev. Mod. Phys. **76**, 1037 (2005).

[17] 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).

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

[19] F. A. Calderon-Vargas and J. P. Kestner, Phys. Rev. Lett. **118**, 150502 (2017).

[20] J. Zhang, J. Vala, S. Sastry, and K. B. Whaley, Phys. Rev. A **67**, 042313 (2003).

[21] M. R. Geller, E. J. Pritchett, A. Galiautdinov, and J. M. Martinis, Phys. Rev. A **81**, 012320 (2010).

[22] A. Warren, E. Barnes, and S. E. Economou, Phys. Rev. B **100**, 161303(R) (2019).

[23] E. Magesan, J. M. Gambetta, and J. Emerson, Phys. Rev. A **85**, 042311 (2012).

[24] M. A. Fogarty, M. Veldhorst, R. Harper, C. H. Yang, S. D. Bartlett, S. T. Flammia, and A. S. Dzurak, Phys. Rev. A **92**, 022326 (2015).

[25] The reason our simulated length-2 sequence fidelity is slightly higher than the experimental one reported in Ref. [18] is simply because we used the more recent lower one-qubit noise value. If we use the conditions of Ref. [18], our trace fidelity calculation yields an error of roughly 6×10^{-4} which is consistent with the experimentally observed values.

[26] J. M. Chow, J. M. Gambetta, L. Tornberg, J. Koch, L. S. Bishop, A. A. Houck, B. R. Johnson, L. Frunzio, S. M. Girvin, and R. J. Schoelkopf, Phys. Rev. Lett. **102**, 090502 (2009).

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

[28] R. Cabrera and W. Baylis, Phys. Lett. A **368**, 25 (2007).

[29] A. P. M. Place, L. V. H. Rodgers, P. S. Mundada, B. Smitham, M. Fitzpatrick, Z. Leng, A. Premkumar, J. Bryon, S. Sussman, G. Cheng, T. Madhavan, H. K. Babla, B. Jaeck, A. Gyenis, N. Yao, R. J. Cava, N. P. de Leon, and A. A. Houck, arXiv:2003.00024.

[30] M. Kjaergaard, M. E. Schwartz, J. Braumüller, P. Krantz, J. I.-J. Wang, S. Gustavsson, and W. D. Oliver, Annu. Rev. Condens. Matter Phys. **11**, 369 (2020).

[31] P. Jurcevic, A. Javadi-Abhari, L. S. Bishop, I. Lauer, D. F. Bogorin, M. Brink, L. Capelluto, O. Günlük, T. Itoko, N. Kanazawa *et al.*, arXiv:2008.08571.

[32] 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).

[33] M. Takita, A. W. Cross, A. D. Córcoles, J. M. Chow, and J. M. Gambetta, Phys. Rev. Lett. **119**, 180501 (2017).

[34] R. Slansky, Phys. Rep. **79**, 1 (1981).

[35] P. Ramond, *Group Theory: A Physicist's Survey* (Cambridge University Press, Cambridge, 2010).

[36] M. A. Nielsen, Phys. Lett. A **303**, 249 (2002).

[37] J. Emerson, R. Alicki, and K. Życzkowski, J. Opt. B: Quantum Sem. Opt. **7**, S347 (2005).