Letter

Characterization of control in a superconducting qutrit using randomized benchmarking

M. Kononenko,[1,2,*] M. A. Yurtalan,[1,2,†] S. Ren,[1,2] J. Shi,[1,2] S. Ashhab,[3,4] and A. Lupascu[1,2,5,‡]

[1]*Institute for Quantum Computing, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1*

[2]*Department of Physics and Astronomy, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1*

[3]*Qatar Environment and Energy Research Institute, Hamad Bin Khalifa University, Qatar Foundation, Ar-Rayyan, Qatar*

[4]*Advanced ICT Research Institute, National Institute of Information and Communications Technology (NICT), 4-2-1, Nukui-Kitamachi, Koganei, Tokyo 184-8795, Japan*

[5]*Waterloo Institute for Nanotechnology, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1*

(Received 6 September 2020; revised 10 September 2021; accepted 14 September 2021; published 22 October 2021)

We characterize control of a qutrit implemented in the lowest three energy levels of a capacitively shunted flux-biased superconducting circuit. Randomized benchmarking over the qutrit Clifford group yields an average fidelity of $98.89 \pm 0.05\%$. For a selected subset of the Clifford group, we perform quantum process tomography and observe the behavior of repeated gate sequences. Each qutrit gate is generated using only two-state rotations via a method applicable to any unitary. We find that errors are due primarily to decoherence and have a significant contribution from level shifts. This work demonstrates high-fidelity qutrit control and outlines avenues for future work on the optimal control of superconducting qudits.

DOI: 10.1103/PhysRevResearch.3.L042007

Recent advances in large-scale quantum information processors have relied on manipulating quantum information using two-level systems as qubits [1–4]. Theoretical work shows that using multilevel systems as qudits offers performance advantages in quantum error correction [5–8], quantum sensing [9,10], and quantum communication [11]. Efficient universal qudit control for these applications follows from an extension of the Solovay-Kitaev theorem from the qubit unitary group SU(2) to the qudit group SU(d), where d is the dimension of the qudit's Hilbert space [12]. However, implementation of such control brings new challenges including mapping qudit gates to experimentally accessible controls and understanding how control errors translate into errors in the qudit gates. Characterizing qudit gates is also more resource intensive than characterizing qubit gates because the larger Hilbert space allows more complex states to form.

In this Letter, we characterize qutrit control using randomized benchmarking (RB), which is a protocol that yields the average fidelity for the elements of the Clifford group. We implement qutrit gates using a universal decomposition method that can be used to generate any unitary for a qutrit and more generally for qudits of any dimension. The measured average fidelity is $\bar{\mathcal{F}} = 98.89 \pm 0.05\%$ for members of the qutrit Clifford group $\mathcal{C}_3$. In addition, we characterize a subset of the Clifford group using quantum process tomography (QPT), which provides an independent verification of the unitary synthesis method and yields results in agreement with RB. Analysis of this experiment revealed errors due to decoherence, level shifts, and leakage. Level shifts, which do not lead to significant errors in the usual case of resonant control in a two-dimensional subspace, are a significant source of error in this experiment, pointing to relevant future work on shaped control pulses for qudits.

The experiment is performed on a variant of the capacitively shunted flux qubit that combines relatively long coherence times with high anharmonicity [13]. Anharmonicity here is defined as the difference between the second and first transition frequencies. A qutrit is encoded in the lowest three energy states of the device, denoted by 0, 1, and 2. The large anharmonicity enables fast selective driving of the 0-1 and 1-2 transitions. Qutrit control is done using microwave pulses sent via a coplanar waveguide capacitively coupled to the device [see Fig. 1(a)]. Application of a microwave pulse resonant with the m-n transition implements a rotation $R(\theta)^{mn}_{\phi}$ in the two-dimensional space formed by states m and n, where θ is the rotation angle and ϕ is the rotation axis phase [14]. Control pulses are synthesized using mixers Marki IQ1545 and IQ0307, for the 0-1 and 1-2 transitions, respectively, fed by two Agilent E8527D PSG signal generators and quadrature control signals from a Tektronix AWG5014 arbitrary waveform generator. The device state is measured using the heterodyne readout of a coplanar waveguide resonator also coupled capacitively to the device. The readout voltage, averaged over many repetitions, corresponds to the expectation value of the operator $V = V_0 |0\rangle\langle 0| + V_1 |1\rangle\langle 1| + V_2 |2\rangle\langle 2|$. State preparation is done by waiting for the device to relax to the thermal state $\rho_{\text{th}} = P_{\text{th},0} |0\rangle\langle 0| + P_{\text{th},1} |1\rangle\langle 1|$, where $P_{\text{th},n} = \langle n|\rho_{\text{th}}|n\rangle$. We assume that higher state populations are

*mkononen@uwaterloo.ca
†mayurtalan@uwaterloo.ca
‡alupascu@uwaterloo.ca

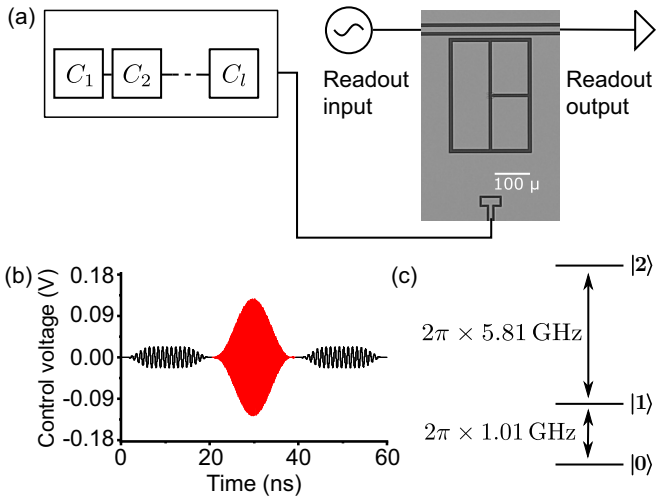

FIG. 1. (a) A representation of the experiment setup showing a scanning electron micrograph of the device, capacitively coupled to a control line and to a readout resonator. The control setup synthesizes a sequence of Clifford gates C_1 to C_l for an RB experiment. (b) The control waveform for a qutrit Walsh-Hadamard gate H_3. Black (red) lines are signals resonant with the 0-1 (1-2) transition. Control voltage refers to the voltage at the output of the control setup. (c) The level diagram for the qutrit, with transition frequencies shown.

negligible, in line with the large transition frequency between states 1 and 2. The thermal excited state population $P_{\text{th},1}$ and the readout voltage levels V_0, V_1, and V_2 are measured using a protocol based on observing Rabi oscillations on the 0-1 transition starting with initial states based on the thermal state with additional population swaps [13,14]. Based on the measured values of V_0, V_1, and V_2, the populations in any state can be characterized by measuring $\langle V \rangle$ in combination with suitable premeasurement analyzer pulses [14].

Performing a qudit gate U requires a decomposition of U that can be mapped to the available controls, which becomes more difficult as d increases. In this experiment, each gate U is decomposed into the product of a set of Givens rotations and a diagonal unitary, using an approach that is universal for any qudit of dimension d as long as $d-1$ different transitions involving all the d states can be controlled. The Givens rotations are determined based on the procedure in Ref. [15]; a maximum number of $\frac{1}{2}d(d-1)$ Givens rotations are required. Each Givens rotation $R(\theta)_\phi^{mn}$ is then mapped to a control pulse with the envelope area proportional to θ, frequency resonant with the m-n transition, and phase given by ϕ. The control Hamiltonian for $R(\theta)_\phi^{mn}$ in the rotating frame is $H_{\text{drive}}(t) = \frac{1}{2}\Omega_{mn}(t)e^{-i\phi}\,|m\rangle\,\langle n| + \text{H.c.}$, where $\Omega_{mn}(t)$ is the drive strength for the m-n transition. We use Givens rotations in subspaces 0-1 and 1-2. Rotations in the 0-1 subspace have cosine-shaped rise and fall envelopes with a total duration of 18.4 ns. Rotations in the 1-2 subspace have derivative removal by adiabatic gate (DRAG) [16] cosine-shaped pulses with a duration of 16.8 ns [14]. We use DRAG for the 1-2 transition in order to suppress leakage to state 3 arising from the small difference between the 1-2 and 2-3 transition frequencies [17]. The diagonal component of U is tracked, and the phases of the Givens rotations in the subsequent gate are modified in order to implement phase gates at the software level [13,14,18].

We verify this approach of implementing qutrit gates using QPT, which is a standard technique for finding the process matrix of a black box [19]. We implement QPT using the procedure from Ref. [20]. The gate fidelity [21] calculated using the process matrix determines whether a qutrit gate is synthesized as intended. We performed QPT for two representative gates, the Walsh-Hadamard gate H_3, which together with the phase gate S_3 generates the Clifford group, and the Pauli X_3 gate, which together with the Pauli Z_3 gate generates the Pauli group [22,23]. We note that S_3 and Z_3 are diagonal gates, and are therefore software defined. The process fidelity for the Walsh-Hadamard gate H_3 is 97.45% and the generalized Pauli gate X_3 is 98.47%. QPT shows that the experimental implementation of the gate decomposition synthesizes the intended gates.

Next, the average fidelity $\bar{\mathcal{F}}$ of $\mathcal{C}_3$ is characterized using RB [23]. RB relies on the fact that the application of a sequence of l gates, of which the first $l-1$ are chosen randomly from $\mathcal{C}_3$ and the last is chosen as the inverse of the product of the first $l-1$ gates, behaves on average as a depolarizing channel with depolarizing coefficient p^l [23–25]. Specifically, the population P_n for each level n decays with l as

$$P_n = (P_{in} - 1/3)p^l + 1/3, \tag{1}$$

where P_{in} is the initial population of state n. The average gate fidelity $\bar{\mathcal{F}}$ is related to the depolarization coefficient p via $\bar{\mathcal{F}} = p + (1-p)/3$. Randomized benchmarking is more efficient than QPT for measuring $\bar{\mathcal{F}}$ while being immune to state preparation and measurement errors, at the expense of not giving information about individual gates in $\mathcal{C}_3$ [25]. Based on the fact that H_d and S_d generate $\mathcal{C}_d$ for any prime d [26], we find all 216 elements of $\mathcal{C}_3$ by calculating all distinct products of H_3 and S_3. Each Clifford gate is decomposed into Givens rotations as explained above. We measure the result of applying a set of random gates of length l ranging from $l=2$ to $l=987$, with 25 different randomizations for each value of l. The measurement is repeated $N_{\text{rep}} = 8192$ times for each different random sequence. Figure 2(a) shows the results of a RB experiment, with the initial state set to be the thermal state. The fit of the populations versus the sequence length, given by $P_n = (P_{in} - P_{\text{fn}})p_n^l + P_{\text{fn}}$, with P_{in}, P_{fn}, and p_n as free parameters, is in excellent agreement with the data. The polarization decay coefficients are the same within the experimental errors, with an average $p = 0.9833 \pm 5 \times 10^{-4}$ and a corresponding fidelity $\bar{\mathcal{F}} = 98.89 \pm 0.05\%$. Similarly, the final values $P_{\text{f0}} = 0.341 \pm 0.003$, $P_{\text{f1}} = 0.333 \pm 0.001$, and $P_{\text{f2}} = 0.325 \pm 0.003$ are close to the expected value of $\frac{1}{3}$ corresponding to the fully depolarized state of a qutrit. Repeating the experiment with 50 randomizations for each l gave the same result to within experimental error.

Since each element in $\mathcal{C}_3$ has a finite order, repeated application of the same gate leads to a periodic result versus the number of repetitions N. This is tested experimentally by measuring the populations after repeated application of H_3, S_3, X_3, and Z_3, starting with the thermal state as an initial state. For S_3 and Z_3, an H_3 gate is prepended and appended to produce a change in the measured populations. Figure 3 shows the populations P_n for these gates. The experiments confirm the

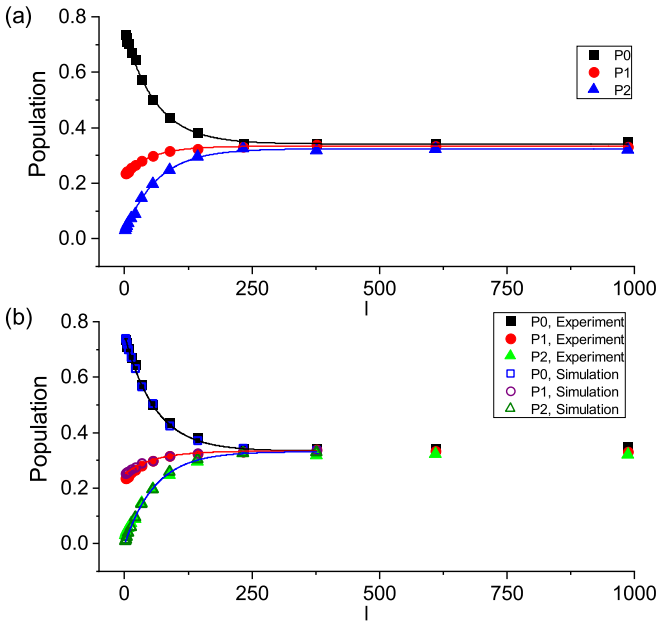

FIG. 2. (a) The experimentally measured average populations P_0 (black squares), P_1 (red circles), and P_2 (blue triangles) vs sequence length l. Solid lines show exponential fits. For P_0, P_1, and P_2, the p_n values are $p_0 = 0.9839 \pm 5.4 \times 10^{-4}$, $p_1 = 0.9814 \pm 1.1 \times 10^{-3}$, and $p_2 = 0.9846 \pm 6.9 \times 10^{-4}$, respectively. (b) The experimentally measured P_0 (black squares), P_1 (red circles), and P_2 (green triangles) vs the sequence length l, compared to the simulated P_0 (open blue squares), P_1 (open purple circles), and P_2 (open dark green triangles). Solid lines show the fits of the RB model to the simulated P_0, P_1, and P_2.

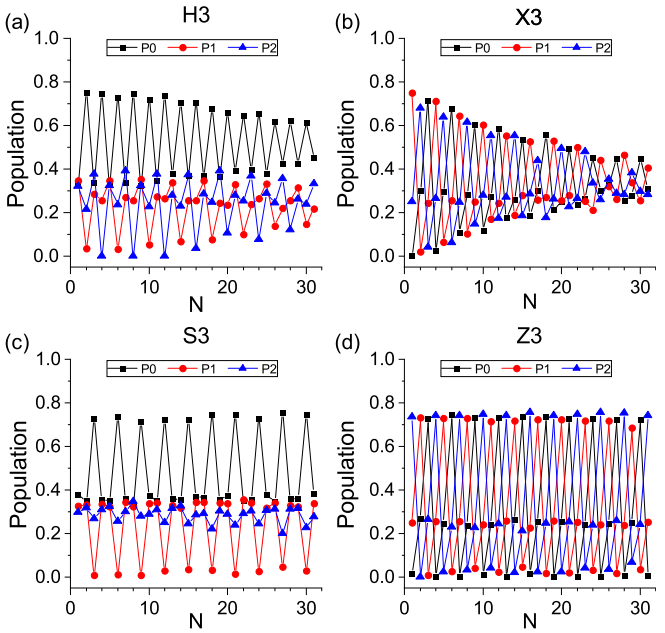

FIG. 3. The experimentally measured populations P_0 (black squares), P_1 (red circles), and P_2 (blue triangles) vs the number N of repetitions of gates (a) H_3, (b) X_3, (c) S_3, and (d) Z_3.

expected periodicity of these elements of the Clifford group, which is 4, 3, 3, and 3 for H_3, S_3, X_3, and Z_3, respectively. The effect of control errors and decoherence is visible as N increases, in good agreement with numerical simulations of the dynamics [14]. The lack of decay in S_3 and Z_3 signals is due to these gates being entirely defined in software following the decomposition above. In the S_3 and Z_3 sequence, errors in control arise only from the preparation and post H_3 pulses.

We now discuss the sources of error in the RB experiment. Numerical simulations of the dynamics are done with two models, based on coherent and incoherent evolution. The device Hamiltonian is truncated to the seven lowest-energy states, which was found to be sufficient to properly explain level shifts in previous work including two-photon driving [14,20]. The superconducting circuit parameters are extracted based on a model fit against the measured spectroscopy data. In addition, calibration of the pulses is done in the simulation in a manner similar to the experiment [14]. We model the dynamics under the randomized benchmarking sequences used in the experiment, by numerically solving the Schrödinger equation with driving [27]. With coherent evolution, state populations extracted from the simulations give a fidelity of $\bar{\mathcal{F}} = 99.91 \pm 0.02\%$, significantly higher that the experimental result. In addition, simulations indicate that leakage to states outside the qutrit space is negligibly small ($2.29 \times 10^{-4}\%$ at $l = 377$).

The role of decoherence is analyzed using a model that includes the measured relaxation and excitation rates and dephasing in the qutrit space. The decoherence model consists of Lindblad operators of the form $\sqrt{\Gamma_1^{mn}}\,|n\rangle\langle m|$ for relaxation/excitation, and $\sqrt{\frac{1}{2}\Gamma_2^{mn}}(|m\rangle\langle m| - |n\rangle\langle n|)$ for dephasing [14]. Dynamics are simulated using a master equation solver [27,28]. Figure 2(b) shows the numerically calculated populations with decoherence. The fit of this simulation gives a fidelity of $\bar{\mathcal{F}} = 98.9 \pm 0.03\%$, which is within the fit error of the experimental result. As for the coherent case, the total population outside the qutrit space is negligible, reaching 0.244% at $l = 377$. Numerical simulation of each gate in $\mathcal{C}_3$ individually, and then averaging their respective fidelities, gives $\bar{\mathcal{F}} = 98.9\%$ with a standard deviation of 0.3% and a worst-case fidelity of 98.5%. This compares well with experimental values, and indicates that in addition to the average fidelity being high, each gate in $\mathcal{C}_3$ is synthesized with high fidelity as well. The range in fidelity is also comparable to the fidelities measured using QPT. From fitting the model in Ref. [29] to the numerically calculated populations, we determine the leakage L_1 per Clifford gate [17], the seepage L_2, and the adjusted average fidelity $\bar{\mathcal{F}}_L$. Whereas leakage measures population transfer out of the qutrit space, seepage measures population transfer into the qutrit space. From simulation, we find $L_1 \lesssim 6.26 \times 10^{-5}$, $L_2 \lesssim 6.78 \times 10^{-5}$, and $\bar{\mathcal{F}}_L = 98.9 \pm 0.2\%$. Since $\bar{\mathcal{F}}_L$ is not significantly smaller than the simulated $\bar{\mathcal{F}}$, leakage does not contribute significantly to the error in the simulation.

The analysis discussed above indicates that the dominant source of error in the current experiment is decoherence. Faster control using stronger driving can help to mitigate this error, but coherent control errors must be understood, since these errors increase with faster driving. To analyze coherent control errors in the qutrit RB experiment, it is useful to connect the errors of Clifford group unitaries to

errors in their component $R(\theta)_{\phi}^{mn}$. For a noisy implementation $\tilde{C} = \prod_n \tilde{R}_n$ of a qutrit gate $C = \prod_n R_n$, with $\tilde{R}_n$ of noisy versions of ideal Givens rotations R_n, the gate error $r(\tilde{C}, C) = 1 - \mathcal{F}(\tilde{C}, C)$ is approximately $r(\tilde{C}, C) \approx \sum_n r(\tilde{R}_n, R_n)$, with $r(\tilde{R}_n, R_n)$ the error for a Givens rotation. The approximation follows from modeling the error for a Givens rotation as an operator $K_n = \tilde{R}_n R_n^{\dagger}$, with $K_n = \alpha_n I + \beta_n M_n$, where α_n the complex number minimizing $||I_3(K_n - \alpha_n I)I_3||_{\infty}$ and $I_3 = |0\rangle\langle 0| + |1\rangle\langle 1| + |2\rangle\langle 2|$, $\beta_n = ||I_3(K - \alpha I_n)I_3||_{\infty}$, and $M_n = \frac{1}{\beta_n}(K_n - \alpha_n I)$. $||A||_{\infty}$ is the magnitude of the largest eigenvalue for an operator A. Relating α_n and β_n to $\mathcal{F}(\tilde{R}_n, R_n)$ and $\mathcal{F}(\tilde{C}, C)$ using

$$\mathcal{F}(\tilde{A}, A) = \frac{\sum_{U_j \in \mathcal{P}_d} \mathrm{Tr}(A^{\dagger} U_j A \tilde{A} U_j \tilde{A}^{\dagger}) + d^2}{d^2(d+1)} \tag{2}$$

for two unitary operators A and $\tilde{A}$ yields $r(\tilde{C}, C) \approx \sum_n r(\tilde{R}_n, R_n)$, assuming products of β_n are small [14]. Figure 4(a) shows that the approximation holds numerically.

To understand the errors in individual Givens rotations, the effective Hamiltonian $H_{\mathrm{eff}} = -\frac{i}{\tau} \ln \tilde{R}_n$ is calculated for each numerically calculated $\tilde{R}_n$, where τ is the effective time to implement R_n and $\ln \tilde{R}_n$ is the matrix logarithm. The difference between H_{eff} and the ideal control Hamiltonian H_{drive} can be connected to errors introduced by the failure of the rotating wave approximation. Figure 4(b) shows the error of the Givens rotations versus drive strength, showing that the error scales quadratically with the drive strength. The effective Hamiltonian is well approximated by $H_{\mathrm{drive}} + \sum_{m,n} s_{mn} \Omega_{mn}^2 \sigma_z^{mn}$, with $\sigma_z^{mn} = |m\rangle\langle m| - |n\rangle\langle n|$, and s_{mn} being a real coefficient. The quadratic scaling indicates that the main errors are caused by driving-induced level shifts. Figures 4(c) and 4(d) show the total frequency shifts versus the drive strength, i.e., $\mathrm{Tr}(-H_{\mathrm{eff}}\sigma_z^{mn})$, for $mn = 01$ and $mn = 12$ transitions induced by $R(\pi)_0^{01}$ and $R(\pi)_0^{12}$ pulses, respectively. The calculated total frequency shifts are in excellent agreement with the analytical formula given in Ref. [20]. The leakage error identified in the RB analysis is also visible in simulations of the Givens rotations as non-negligible values of $\langle 2|H_{\mathrm{eff}}|3\rangle$, $\langle 1|H_{\mathrm{eff}}|6\rangle$, and $\langle 2|H_{\mathrm{eff}}|6\rangle$. However, these terms are small compared to $\mathrm{Tr}(H_{\mathrm{eff}}\sigma_z^{mn})$, confirming that the contribution of leakage to the control error is small compared to level shifts. Note that level shift errors are much more significant for qutrit control than for qubit control, since the level shift is a coherent error on levels used to store information in qutrits, compared to an incoherent error on levels not used to store information in qubits.

In conclusion, we demonstrated control sufficient to synthesize the qutrit Clifford group $\mathcal{C}_3$ with $98.89 \pm 0.05\%$ fidelity, using a universal method for gate decomposition into Givens rotations. This fidelity is in agreement with QPT data. The experimental errors are dominated by decoherence, but level shifts due to off-resonant coupling to states outside the driven two-dimensional subspace of each Givens rotation contribute to control error as well. In future work it will be interesting to explore the application of level shift corrections, as done in Ref. [20], and more generally design optimal control pulses that mitigate both level shifts and leakage. These results establish randomized benchmarking as a tool to understand superconducting qutrit control and pave the way towards using superconducting qudits in quantum information tasks.

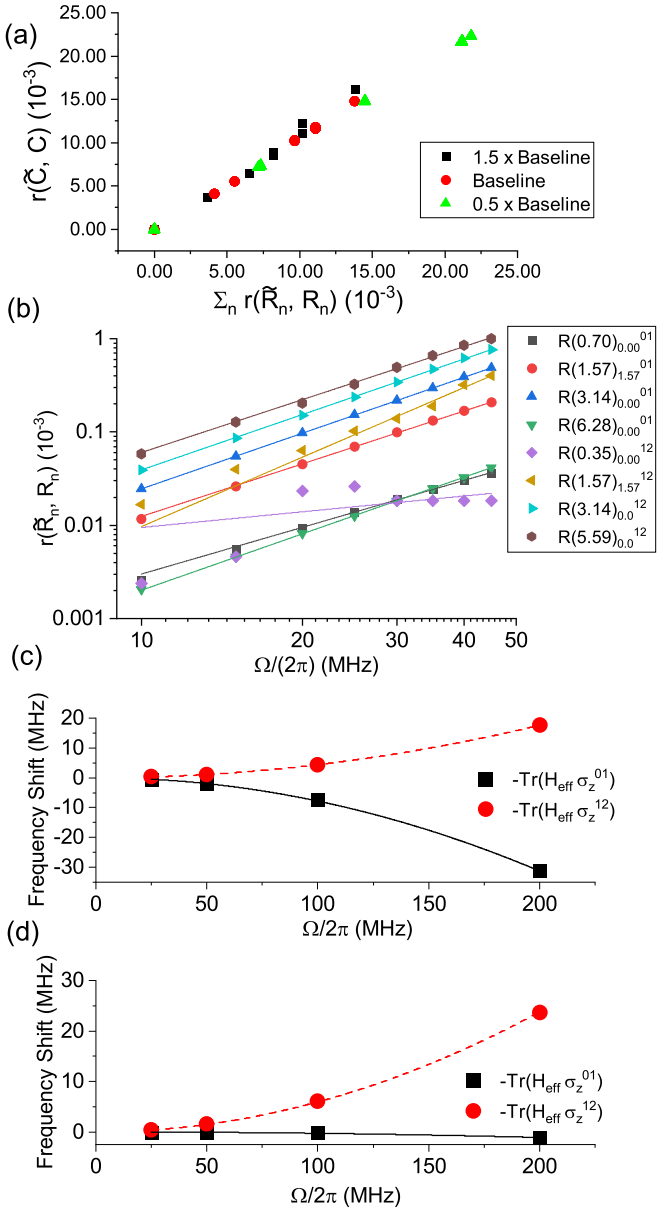

FIG. 4. (a) The error of the Clifford gate vs the sum of the errors of its constituent Givens rotations for all gates in $\mathcal{C}_3$. The black squares, red circles, and green triangles correspond to amplitudes multiplied by 1.5, 1.0, and 0.5 relative to the experiment pulses, respectively. Pulse durations for each multiplication case are adjusted to preserve the intended rotation angles. (b) The error $r = 1 - \mathcal{F}$ vs the drive strength for a set of Givens rotations. Different symbols correspond to the different rotations. (c), (d) The value of the indicated component of the effective Hamiltonian for (c) $R(\pi)_0^{01}$ and (d) $R(\pi)_0^{12}$ vs drive strength. Black squares (red circles) indicate the 01 (12) level shift. The black solid (red dashed) line shows a quadratic fit to the 01 (12) level shift vs drive frequency.

Note added. Recently, we became aware of a related manuscript in qutrit randomized benchmarking [30].

We thank the University of Waterloo Quantum Nanofab team members for their help during device fabrication. We

acknowledge support from NSERC, the Canada Foundation for Innovation (CFI), the Ontario Ministry of Research and Innovation, Industry Canada, and the Canadian Microelectronics Corporation (CMC). S.A. was supported in part by the Ministry of Education, Culture, Sports, Science and Technology (MEXT) Quantum Leap Flagship Program Grant No. JPMXS0120319794.

M.K. and A.Y. contributed equally to this work.

[1] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. S. L. Brandao, D. A. Buell *et al.*, Quantum supremacy using a programmable superconducting processor, Nature (London) **574**, 505 (2019).

[2] J. S. Otterbach, R. Manenti, N. Alidoust, A. Bestwick, M. Block, B. Bloom, S. Caldwell, N. Didier, E. Schuyler Fried, S. Hong *et al.*, Unsupervised machine learning on a hybrid quantum computer, arXiv:1712.05771.

[3] A. D. Córcoles, A. Kandala, A. Javadi-Abhari, D. T. McClure, A. W. Cross, K. Temme, P. D. Nation, M. Steffen, and J. M. Gambetta, Challenges and opportunities of near-term quantum computing systems, Proc. IEEE **108**, 1338 (2020).

[4] K. Wright, K. M. Beck, S. Debnath, J. M. Amini, Y. Nam, N. Grzesiak, J.-S. Chen, N. C. Pisenti, M. Chmielewski, C. Collins *et al.*, Benchmarking an 11-qubit quantum computer, Nat. Commun. **10**, 5464 (2019).

[5] E. T. Campbell, Enhanced Fault-Tolerant Quantum Computing in d-Level Systems, Phys. Rev. Lett. **113**, 230501 (2014).

[6] H. Anwar, B. J. Brown, E. T. Campbell, and D. E. Browne, Fast decoders for qudit topological codes, New J. Phys. **16**, 063038 (2014).

[7] A. Krishna and J.-P. Tillich, Towards Low Overhead Magic State Distillation, Phys. Rev. Lett. **123**, 070507 (2019).

[8] S. Prakash, Magic state distillation with the ternary Golay code, Proc. R. Soc. A **476**, 20200187 (2020).

[9] A. R. Shlyakhov, V. V. Zemlyanov, M. V. Suslov, A. V. Lebedev, G. S. Paraoanu, G. B. Lesovik, and G. Blatter, Quantum metrology with a transmon qutrit, Phys. Rev. A **97**, 022115 (2018).

[10] M. V. Suslov, G. B. Lesovik, and G. Blatter, Quantum abacus for counting and factorizing numbers, Phys. Rev. A **83**, 052317 (2011).

[11] F. Bouchard, R. Fickler, R. W. Boyd, and E. Karimi, High-dimensional quantum cloning and applications to quantum hacking, Sci. Adv. **3**, e1601915 (2017).

[12] C. M. Dawson and M. A. Nielsen, The Solovay-Kitaev algorithm, Quantum Inf. Comput. **6**, 81 (2006).

[13] M. A. Yurtalan, J. Shi, G. J. K. Flatt, and A. Lupascu, Characterization of multi-level dynamics and decoherence in a high-anharmonicity capacitively shunted flux circuit, arXiv:2008.00593.

[14] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevResearch.3.L042007 for additional information related to qubit parameters, simulations, and a proof of the Givens error approximation.

[15] S. G. Schirmer, A. D. Greentree, V. Ramakrishna, and H. Rabitz, Constructive control of quantum systems using factorization of unitary operators, J. Phys. A: Math. Gen. **35**, 8315 (2002).

[16] F. Motzoi, J. M. Gambetta, P. Rebentrost, and F. K. Wilhelm, Simple Pulses for Elimination of Leakage in Weakly Nonlinear Qubits, Phys. Rev. Lett. **103**, 110501 (2009).

[17] Z. Chen, J. Kelly, C. Quintana, R. Barends, B. Campbell, Y. Chen, B. Chiaro, A. Dunsworth, A. G. Fowler, E. Lucero *et al.*, Measuring and Suppressing Quantum State Leakage in a Superconducting Qubit, Phys. Rev. Lett. **116**, 020501 (2016).

[18] D. C. McKay, C. J. Wood, S. Sheldon, J. M. Chow, and J. M. Gambetta, Efficient Z gates for quantum computing, Phys. Rev. A **96**, 022330 (2017).

[19] I. L. Chuang and M. A. Nielsen, Prescription for experimental determination of the dynamics of a quantum black box, J. Mod. Opt. **44**, 2455 (1997).

[20] M. A. Yurtalan, J. Shi, M. Kononenko, A. Lupascu, and S. Ashhab, Implementation of a Walsh-Hadamard Gate in a Superconducting Qutrit, Phys. Rev. Lett. **125**, 180504 (2020).

[21] M. A. Nielsen, A simple formula for the average gate fidelity of a quantum dynamical operation, Phys. Lett. A **303**, 249 (2002).

[22] A. N. Glaudell, N. J. Ross, and J. M. Taylor, Canonical forms for single-qutrit Clifford+T operators, Ann. Phys. (NY) **406**, 54 (2019).

[23] M. Jafarzadeh, Y.-D. Wu, Y. R. Sanders, and B. C. Sanders, Randomized benchmarking for qudit Clifford gates, New J. Phys. **22**, 063014 (2020).

[24] J. Emerson, R. Alicki, and K. Zyczkowski, Scalable noise estimation with random unitary operators, J. Opt. B: Quantum Semiclass. Opt. **7**, S347 (2005).

[25] E. Magesan, J. M. Gambetta, and J. Emerson, Scalable and Robust Randomized Benchmarking of Quantum Processes, Phys. Rev. Lett. **106**, 180504 (2011).

[26] D. Gottesman, Fault-tolerant quantum computation with higher-dimensional systems, in *Quantum Computing and Quantum Communications*, edited by C. P. Williams, Lecture Notes in Computer Science (Springer, Berlin, 1999), pp. 302–313.

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

[28] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, UK, 2013).

[29] C. J. Wood and J. M. Gambetta, Quantification and characterization of leakage errors, Phys. Rev. A **97**, 032306 (2018).

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