

Exploring ququart computation on a transmon using optimal control

Lennart Maximilian Seifert ^{1,*} Ziqian Li ^{2,3,4} Tanay Roy ^{2,3} David I. Schuster ^{2,3,4,5}
Frederic T. Chong ¹ and Jonathan M. Baker ^{1,6}

¹*Department of Computer Science, University of Chicago, Chicago, Illinois 60637, USA*

²*James Franck Institute, University of Chicago, Chicago, Illinois 60637, USA*

³*Department of Physics, University of Chicago, Chicago, Illinois 60637, USA*

⁴*Department of Applied Physics, Stanford University, Stanford, California 94305, USA*

⁵*Pritzker School of Molecular Engineering, University of Chicago, Chicago, Illinois 60637, USA*

⁶*Duke Quantum Center, Duke University, Durham, North Carolina 27701, USA*

 (Received 21 April 2023; revised 25 June 2023; accepted 18 October 2023; published 11 December 2023)

Contemporary quantum computers encode and process quantum information in binary qubits ($d = 2$). However, many architectures include higher energy levels that are left as unused computational resources. We demonstrate a superconducting ququart ($d = 4$) processor and combine quantum optimal control with efficient gate decompositions to implement high-fidelity ququart gates. We distinguish between viewing the ququart as a generalized four-level qubit and an encoded pair of qubits, and characterize the resulting gates in each case. In randomized benchmarking experiments we observe gate fidelities $\geq 95\%$ and identify coherence as the primary limiting factor. Our results validate ququarts as a viable tool for quantum information processing.

DOI: [10.1103/PhysRevA.108.062609](https://doi.org/10.1103/PhysRevA.108.062609)

I. INTRODUCTION

While current efforts to build quantum computers mostly focus on the fabrication and high-fidelity control of two-level qubits, many proposed implementations like superconducting transmons or trapped ions provide a much larger Hilbert space with more energy levels present. Their influence is typically aimed to be suppressed through careful device engineering; however, these states are readily available for computation.

Qudits, the d -level generalization of qubits, have gained a lot of interest as they provide an exponential increase in Hilbert space dimension (d^N versus 2^N), allowing for the adaption and simplification of a variety of algorithms [1–5]. Extensive research has been conducted for the qutrit case $d = 3$. While theoretical studies have shown benefits for quantum compilation [6,7] and improved schemes for quantum error correction [8–11], experimental implementations of qutrits have been presented on several architectures [12–20].

Ququarts ($d = 4$) can be explored from a similar perspective, and they furthermore offer the alternative interpretation of storing the information content of two qubits. Theoretical works have looked at the encoding of qubit pairs into ququarts to study advantages for quantum circuit compilation [21,22]. Increased hardware utilization, efficient internal two-qubit gates, and reduced routing costs render this design highly promising. Recent experimental results have shown the efficient realization of a variational quantum algorithm [23] and the verification of the entropic inequality [24] under this scheme.

Quantum optimal control comprises a set of methods to find hardware-specific control pulses to make a quantum

system undergo a desired transformation. A variety of frameworks have been developed [25–33] and applied in both theory and experiment to solve problems like state transfer [34–36] and the implementations of gates [37–40].

In this work we study the realization of single-ququart operations on a superconducting transmon under both the qudit perspective as well as the encoded-qubit perspective. We compare two approaches for each gate: an optimized decomposition into natively supported gates on our system, and direct implementation through quantum optimal control. Using the frameworks JUQBOX [25] and BOULDER OPAL [30], we find optimal control pulses that drive all three transitions in parallel while respecting hardware constraints. We perform quantum-process tomography (QPT) as well as randomized benchmarking (RB) to characterize the fidelity of different gate implementations, where we distinguish between the ququart Clifford group \mathcal{C}_4 and the two-qubit Clifford group $\mathcal{C}_2^{\otimes 2}$. We focus on the ququart Hadamard H_4 and the two-qubit Hadamard tensor product $H \otimes H$, due to their similar structure, and include further results in the Appendix.

II. ONE QUQUART OR TWO QUBITS

The concept of the ququart extends the computational unit of a qubit by two levels and describes the special qudit case $d = 4$. Single-ququart gates are then given by 4×4 unitary matrices, and some of them can thus be seen as generalizations of familiar single-qubit gates [41]:

$$X_4 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad H_4 = \frac{1}{2} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & i & -1 & -i \\ 1 & -1 & 1 & -1 \\ 1 & -i & -1 & i \end{bmatrix},$$

$$Z_4 = \text{diag}(1, i, -1, -i). \quad (1)$$

*Corresponding author: lmseifert@uchicago.edu

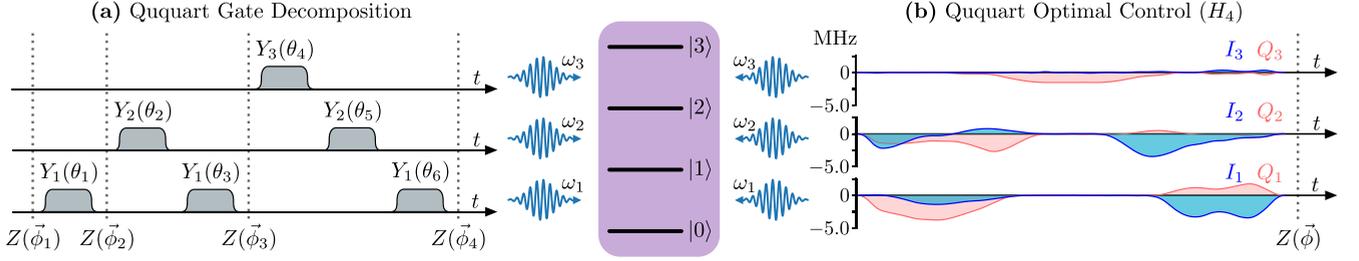


FIG. 1. Two ways of realizing arbitrary quantum gates on a four-level ququart, which is controlled by three modulated carrier waves with resonant frequencies ω_j . (a) Any ququart gate can be decomposed into a series of at most six qubitlike Y rotations interleaved with ququart phase gates. Given a target unitary, the parameters θ_i and $\vec{\phi}_i$ can be calculated numerically. (b) Lifting the gate abstraction, quantum optimal control can be used to determine a control signal that operates on all subspaces in parallel and manipulate the entire ququart state. A final virtual Z gate adjusts the relative phases. The shown pulses together with $\vec{\phi} = (4.265, 4.590, 0.000)$ (radians) implement a ququart Hadamard H_4 in $T = 350$ ns.

Here X_4 and Z_4 represent the generalizations of the qubit Pauli operators X and Z , respectively, and H_4 denotes the ququart Hadamard.

Due to the special Hilbert space dimension $d = 4 = 2^2$, another way to think about a ququart is as a pair of qubits. One possible encoding is given by mapping the qubit bit strings to their corresponding decimal numbers:

$$|q_0 q_1\rangle = \left\{ \begin{array}{l} |00\rangle \leftrightarrow |0\rangle \\ |01\rangle \leftrightarrow |1\rangle \\ |10\rangle \leftrightarrow |2\rangle \\ |11\rangle \leftrightarrow |3\rangle \end{array} \right\} = |q^{(4)}\rangle. \quad (2)$$

Therefore two-qubit operators can be directly translated to ququart gates given their matrix representation. For example, simultaneous bit flips on q_0 and q_1 described by $X \otimes X$ can be implemented by flipping $|0\rangle \leftrightarrow |3\rangle$ and $|1\rangle \leftrightarrow |2\rangle$. Furthermore, such an encoding provides efficient qubit CX and $SWAP$ gates through state flips $|2\rangle \leftrightarrow |3\rangle$ and $|1\rangle \leftrightarrow |2\rangle$, respectively. Past theoretical and experimental works have studied qubit-based computation under this scheme to find potential advantages in resource requirements and circuit speed [21–23].

A. Gate decomposition

We use the lowest four energy levels $\{|0\rangle, |1\rangle, |2\rangle, |3\rangle\}$ of a superconducting transmon [42] to represent the ququart. The elementary gates available in the transmon are elements of $SU(2)$, meaning qubitlike rotations $R_j(\theta, \phi) = Z_j^\dagger(\phi)Y_j(\theta)Z_j(\phi)$ between two adjacent levels $|j-1\rangle$ and $|j\rangle$, with $j \in \{1, 2, 3\}$. The rotations Y_j around the y axis are implemented using microwave pulses resonant with the $|j-1\rangle \leftrightarrow |j\rangle$ transitions, while the phase rotations Z_j are realized in software (known as “virtual Z ”) by appropriately updating the phase of the respective carrier wave [43] and are therefore near perfect. The concept of virtual Z gates can be generalized to the ququart case, allowing the direct implementation of relative phase rotations $Z(\vec{\phi}) = \text{diag}(1, e^{i\phi_1}, e^{i(\phi_1+\phi_2)}, e^{i(\phi_1+\phi_2+\phi_3)})$ by updating all three carrier waves at once, where $\vec{\phi} = (\phi_1, \phi_2, \phi_3)$, and ϕ_j denotes the phase shift to the j th microwave drive.

It has been shown that a sequence of at most six Y_j rotations interleaved with phase rotations Z suffices to realize any ququart unitary [44], which is visualized in Fig. 1(a). In

our work we choose to modify the decomposition protocol outlined in Ref. [44] such that only one Y_3 rotation is required. This is because we do not observe the expected \sqrt{j} scaling of the Rabi rate in experiment but instead find the transition $|2\rangle \leftrightarrow |3\rangle$ to have the smallest achievable Rabi rate. Given a ququart unitary U , we can numerically optimize the parameters of the sequence, where we formulate the optimization problem such that we find a sequence with minimal duration. Appendix B provides more details on the Rabi rate calibration, while Appendix C describes the decomposition optimization and includes Table II, showing decompositions for H_4 , $H \otimes H$, and a few more special gates.

B. Quantum optimal control

Instead of constructing a target gate from high-level building blocks like precalibrated gates, one can take a more low-level approach and directly search for control pulses which make the quantum system undergo the desired transformation. Quantum optimal control (QOC) deals with finding such optimized pulses that implement desired operations, where the optimization can be based exclusively on a simulated model (open-loop) or incorporate real hardware feedback (closed loop). Here we use the open-source software package JUQBOX [25,45] by Lawrence Livermore National Laboratory (LLNL) as well as the proprietary tool BOULDER OPAL [30,31] from Q-CTRL to solve open-loop tasks.

To this end, we model our superconducting transmon with the Hamiltonian

$$\mathcal{H} = \mathcal{H}_0 + \mathcal{H}_c(t) = \sum_{n=0}^{\tilde{d}-1} \epsilon_n |n\rangle\langle n| - i\gamma(t)(a - a^\dagger), \quad (3)$$

where \mathcal{H}_0 and \mathcal{H}_c describe the drift and the control Hamiltonian, respectively. $\gamma(t)$ is the control pulse, which we write as

$$\begin{aligned} \gamma(t) &= \sum_{j=1}^3 \Gamma_j(t) \cos(\omega_j t + \varphi_j(t)) \\ &= \text{Re} \left[\sum_{j=1}^3 (I_j(t) + iQ_j(t)) e^{i\omega_j t} \right], \end{aligned} \quad (4)$$

where for each subspace we define the quadratures $I_j(t) = \Gamma_j(t) \cos(\varphi_j(t)) = \sum_{s=1}^{N_s} \alpha_{j,s}^I B_s(t)$ and $Q_j(t) = \Gamma_j(t) \sin(\varphi_j(t)) = \sum_{s=1}^{N_s} \alpha_{j,s}^Q B_s(t)$. In this work we follow the parametrization implemented in JUQBOX, where pulse envelopes are decomposed into N_s time-local B-spline basis functions $B_s(t)$ with optimizable coefficients $\alpha_{j,s}^{I/Q}$, which allows a low-dimensional representation of long-duration pulses.

We calibrate the resonant carrier frequencies $\omega_j = \epsilon_j - \epsilon_{j-1}$ for the ququart subspace from Rabi and Ramsey experiments and extrapolate the energies of higher levels as $\epsilon_n = \epsilon_3 + (n-3)\omega_3 + \frac{(n-2)(n-3)}{2}(\omega_3 - \omega_2)$ for $n > 3$. This method essentially defines an anharmonicity $\xi = \omega_3 - \omega_2$ and extends the calibrated spectrum according to the anharmonic oscillator model. The Hamiltonian is truncated at $\tilde{d} = 5$ levels for simulation, which includes a guard level outside our computational subspace to capture leakage effects.

The open-loop optimal control task consists of adjusting the pulse parameters $\alpha_{j,s}^{I/Q}$ such that one obtains a pulse that minimizes the gate infidelity

$$1 - F_{\vec{\phi}} = 1 - \frac{1}{d^2} |\text{Tr}(U_T^\dagger Z(\vec{\phi}) V_{\vec{\alpha}})|^2. \quad (5)$$

Here U_T denotes the target unitary at time T , and $V_{\vec{\alpha}}$ describes the unitary the pulse with parameter vector $\vec{\alpha}$ realizes. We search for a pulse that implements the target up to a trailing phase gate $Z(\vec{\phi})$, thus we also optimize over the phase vector $\vec{\phi}$. This effectively creates a manifold of equally desirable optimization targets, some of which might be easier to realize than the original target. Since virtual phase gates come at zero cost in the experiment, this increases the freedom of the optimization and generally allows for faster pulse implementations. This trailing Z gate has the same effect as $Z(\vec{\phi}_4)$ in the gate decomposition approach shown in Fig. 1(a). Note that an explicit leading phase gate is not required, as its effect is already captured by the phase relations of $I_j + iQ_j$.

We further introduce additional cost terms to the optimization problem to ensure a narrow Fourier spectrum of the pulse γ , as well as to respect the power constraint $\sum_{j=1}^3 \frac{|\Gamma_j(t)|}{r_j} \leq (1 \text{ AWG})$ determined by the output limit of the arbitrary waveform generator (AWG). r_j corresponds to the drive amplitude the transmon experiences when we solely drive the j th transition at full power, which produces a frequency-dependent response we calibrate in advance. Figure 1(b) shows optimized quadrature envelopes I_j and Q_j that modulate the carrier waves to directly realize the ququart Hadamard H_4 up to a final phase correction $\vec{\phi} = (4.265, 4.590, 0.000)$ (radians). We similarly obtain controls for the Hadamard tensor product $H \otimes H$. For this gate the freedom of a final phase correction did not help to achieve a reduced gate duration, so we set $\vec{\phi} = \vec{0}$. More details on the drive calibration can be found in Appendix B, and further information on the optimal control setup as well as visualizations of pulse results are included in Appendix D.

III. EXPERIMENTAL RESULTS

We use the two-transmon processor presented in Refs. [18,20], which is built from two superconducting transmons with a flux-tunable coupler, and focus solely on transmon Q_1 . Q_2 remains in the ground state and can be disregarded when the coupler is not modulated and biased at the sweet spot. Details of the device are presented in Appendix A.

A. Quantum process tomography

We start by estimating the quality of our optimal control pulses via quantum-process tomography [46], which allows us to fully characterize the corresponding quantum channel $\mathcal{E}(\rho) = \sum_{j,k} \chi_{jk} B_j \rho B_k^\dagger$. We choose the operator basis $\{B_j\}$ depending on the scenario. While for true ququart gates like the H_4 we use products $Z_4^m X_4^n$ of the generalized Pauli operators X_4 and Z_4 , in the encoded-qubits case we follow the common choice of tensor products of qubit Pauli gates $\sigma_m \otimes \sigma_n$ and apply the encoding scheme to map them to ququart operators ($m, n \in \{0, 1, 2, 3\}$). We extract the process matrix χ of a specific gate by applying it to 16 differently prepared states $\{|k\rangle, (|l\rangle + |k\rangle)/\sqrt{2}, (|l\rangle - i|k\rangle)/\sqrt{2}\}$ ($k, l \in \{0, 1, 2, 3\}, k > l$) and reconstructing the resulting state using maximum likelihood estimation.

Figure 2 shows the process matrices for both gates of interest H_4 and $H \otimes H$, comparing between ideal theory and implementation on our transmon. Defining the experimental process fidelity as $F = |\text{Tr}[\chi_{\text{ideal}} \chi_{\text{expt}}]|$, we obtain fidelities of 86.19% and 84.06%, respectively. As a reference, we note that performing no gate (thus, characterizing the identity $\mathbb{1}_4$) leads to a process fidelity of 89.88%. We trace back this loss in fidelity to measurement errors as the primary source, resulting from the challenge of implementing a four-state single-shot readout. A well-known disadvantage of quantum-process tomography for benchmarking quantum gates is that it cannot separate gate errors from state preparation and measurement (SPAM) errors, thus yielding a lower bound on the true gate performance. We expand on the QPT configuration as well as the measurement setup in Appendix E and include further experimental results in Table IV.

B. Randomized benchmarking

Randomized benchmarking (RB) allows extracting the average error per Clifford gate r_C without including SPAM errors [47]. By averaging over randomly sampled sequences of gates from the Clifford group \mathcal{C} , the error channel becomes effectively depolarizing; analyzing the exponential decay over sequences of different depths m allows extracting the depolarization parameter p_C . The error per gate is then given as $r_C = \frac{d-1}{d}(1 - p_C)$ and the average gate fidelity defined as $F_C = 1 - r_C$.

Again, we have to carefully distinguish both scenarios, as the ququart Clifford group \mathcal{C}_4 and the two-qubit Clifford group $\mathcal{C}_2^{\otimes 2}$ differ. While \mathcal{C}_4 can be generated from $\{H_4, Z_4, S_4 = \text{diag}(1, \sqrt{i}, i, \sqrt{i})\}$, $\mathcal{C}_2^{\otimes 2}$ is generated by the qubit gates $\{H, S, CX\}$ and contains 11 520 elements (up to scalars) [48]. \mathcal{C}_4 is not a subgroup of $\mathcal{C}_2^{\otimes 2}$ as $H_4 \notin \mathcal{C}_2^{\otimes 2}$, for instance.

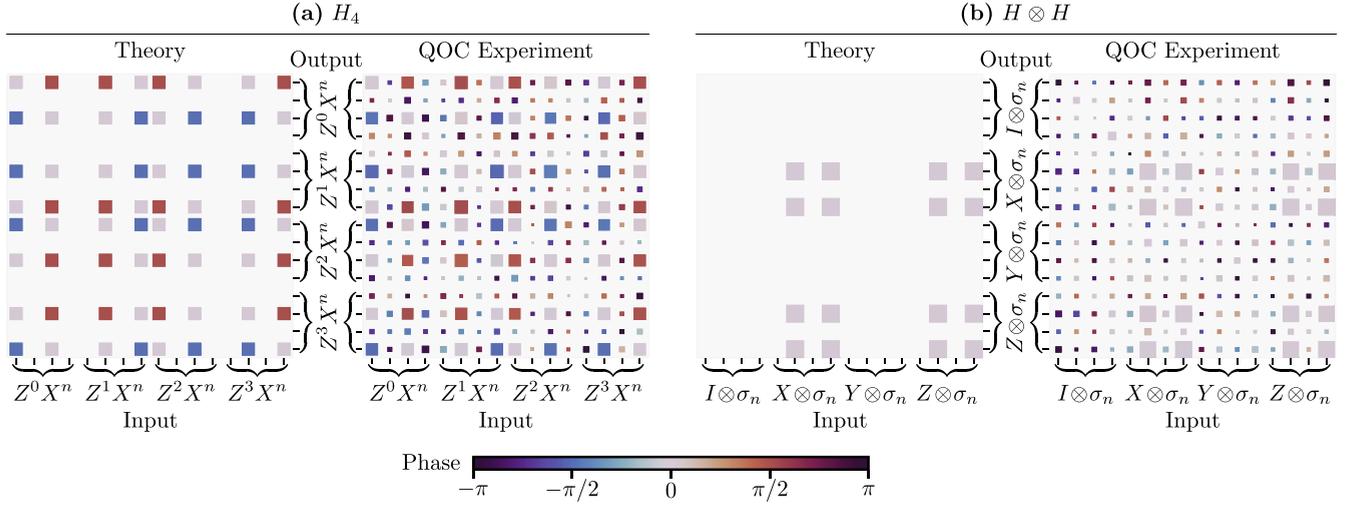


FIG. 2. Process matrices for quantum optimal control implementations of the (a) ququart H_4 and (b) two-qubit $H \otimes H$. The marker area encodes the magnitude of each element while the color (shade) encodes the phase. The H_4 process matrix is represented in the generalized Pauli basis, while $H \otimes H$ is represented in the Pauli basis for two qubits. The exponent or index $n \in \{0, 1, 2, 3\}$ takes on increasing values from left to right or top to bottom for each tick in a label group.

In experiment, we perform ququart RB by sampling gates from \mathcal{C}_4 , as well as two-qubit RB by sampling and encoding gates from $\mathcal{C}_2^{\otimes 2}$, using sequences of depth m up to 100. Each sampled gate is decomposed into elementary operations as outlined in Sec. II A. The measured survival probabilities are visualized in Fig. 3 and clearly show the exponential decay. We obtain RB fidelities $F_{\mathcal{C}_4} = 96.22(14)\%$ and $F_{\mathcal{C}_2^{\otimes 2}} = 95.84(05)\%$ for the ququart RB and two-qubit RB, respectively.

Interleaved randomized benchmarking

Interleaved RB (IRB) builds upon standard RB to characterize the error of a specific Clifford group element G , which is achieved by interleaving the gate of interest between the randomly sampled RB gates [49]. Similar to standard RB, this yields the depolarization parameter p_{C+G} . Together with the previously determined parameter p_C , the gate fidelity can be computed as

$$F_G = 1 - r_G = 1 - \frac{d-1}{d} \left(1 - \frac{p_{C+G}}{p_C} \right). \quad (6)$$

We apply IRB to characterize the ququart Hadamard H_4 and the Hadamard tensor product $H \otimes H$, where we interleave them between gates sampled from the appropriate Clifford groups. We do this for both the gate-level decomposition as well as the optimal control implementation of each gate and juxtapose the results in Fig. 3. The open-loop optimized pulses achieve an improvement over the gate-based realizations in both cases: the H_4 QOC pulse is benchmarked at 96.41(32)%, while the decomposition achieves 96.18(32)% fidelity; the $H \otimes H$ QOC pulse yields 95.98(19)% fidelity compared to the decomposition with 95.17(26)%. This observation suggests a good agreement between the Hamiltonian model and the device. The main limitation is now caused by decoherence, dominated by the increased dephasing in the $\{|2\rangle, |3\rangle\}$ subspace. We include more details on ququart RB as well as

additional results for special gates in Appendix F, and present error simulations in Appendix G.

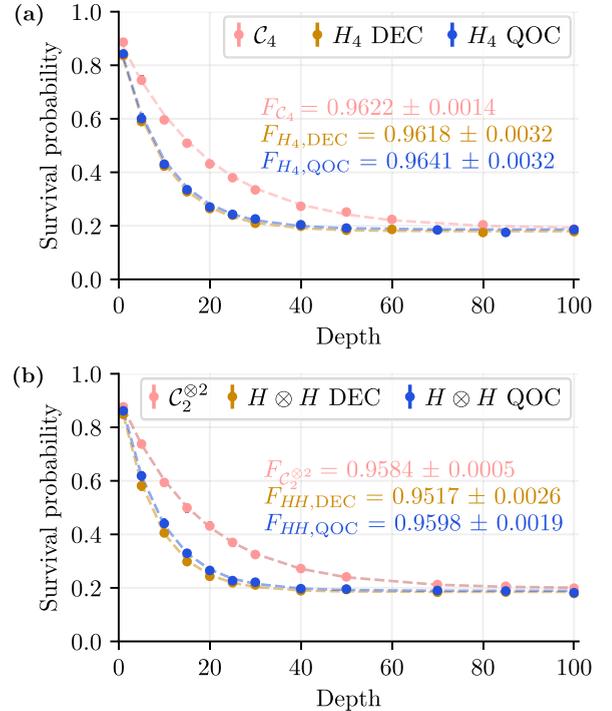


FIG. 3. Experimental results for (a) ququart RB [\mathcal{C}_4 Clifford group, red] and (b) two-qubit RB [$\mathcal{C}_2^{\otimes 2}$ Clifford group, red (light gray)] on a four-level transmon. Using IRB we benchmark (a) H_4 and (b) $H \otimes H$ in the appropriate setting, for both their gate-level decomposition [DEC, yellow (medium gray)] and their quantum optimal control [QOC, blue (dark gray)] implementation. Markers represent the mean survival probabilities, and error bars (smaller than the marker size) are the standard deviations of these means.

IV. CONCLUSION AND OUTLOOK

We demonstrate the coherent control of the lowest four levels in a superconducting transmon, where we explore the implementation of both high-level optimized gate sequences and low-level optimized control pulses to realize target unitaries. While in the first approach we drive only one transition at a time, the latter approach enables manipulating all internal states at once by using open-loop quantum optimal control tools.

We study the quantum system from two angles: from the ququart perspective, as the four-dimensional generalization of a qubit, and from the perspective of two encoded qubits. We characterize gates from either perspective in QPT and RB experiments, focusing in particular on the ququart H_4 and the two-qubit $H \otimes H$ gate. They have similar matrix representations but belong to different Clifford groups \mathcal{C}_4 and $\mathcal{C}_2^{\otimes 2}$, respectively, and we observe IRB fidelities $\geq 95\%$ for both, finding that the optimal control implementations perform better.

The fundamental challenges of increased noise on higher levels as well as the difficulty of performing a four-state read-out limit the capabilities of our device. Future work should explore the application of data-driven methods like presented in Ref. [32] to incorporate hardware feedback into the pulse calibration.

While recently presented related work also studied the two-qubit encoding scheme [23,24], we explore optimally controlled quantum gates on a transmon ququart and further perform ququart randomized benchmarking using the Clifford group \mathcal{C}_4 .

ACKNOWLEDGMENTS

We thank A. Goldschmidt and J. D. Chadwick for helpful discussions on optimal control, as well as N. A. Petersson and S. Günther for assistance with its implementation in JUQBOX. We further thank Q-CTRL for providing us with a license for their software BOULDER OPAL and André Carvalho for his support in incorporating it into our work. L.M.S., F.T.C., and J.M.B. acknowledge funding in part by EPiQC, an NSF Expedition in Computing, under Award No. CCF-1730449; in part by STAQ under Award No. NSF Phy-1818914; in part by the U.S. Department of Energy Office of Advanced Scientific Computing Research, Accelerated Research for Quantum Computing Program; and in part by the NSF Quantum Leap Challenge Institute for Hybrid Quantum Architectures and Networks (NSF Award No. 2016136) and in part based upon work supported by the U.S. Department of Energy, Office of Science, National Quantum Information Science Research Centers. F.T.C. is Chief Scientist for Quantum Software at Infleqion and an advisor to Quantum Circuits, Inc. Z.L., T.R., and D.I.S. acknowledge support by AFOSR Grant No. FA9550-19-1-0399 and ARO Grant No. W911NF-17-S-0001. The device was fabricated in the Pritzker Nanofabrication Facility at the University of Chicago, which receives support from Soft and Hybrid Nanotechnology Experimental (SHyNE) Resources (NSF ECCS-1542205), a node of the National Science Foundation's National Nanotechnology Coordinated Infrastructure. This work also made use of the

TABLE I. Ququart characterization. Relaxation times T_1 and Ramsey dephasing times T_R between neighboring energy levels $|j-1\rangle \leftrightarrow |j\rangle$ ($j = 1, 2, 3$) are listed, as well as transition frequencies ω_j and on-chip drive responses r_j to unit power AWG driving.

Transitions	$ 0\rangle \leftrightarrow 1\rangle$	$ 1\rangle \leftrightarrow 2\rangle$	$ 2\rangle \leftrightarrow 3\rangle$
T_1 (μs)	22.62	25.96	10.19
T_R (μs)	40.60	42.32	2.68
Frequency $\omega_j/2\pi$ (GHz)	3.2222	3.1021	2.9717
Drive response r_j (MHz/AWG)	5.5528	5.6712	1.6749

shared facilities at the University of Chicago Materials Research Science and Engineering Center.

APPENDIX A: DEVICE PARAMETERS AND FABRICATION

The substrate used for the device fabrication is a 430- μm -thick C-plane sapphire that was annealed at 1200 °C for 2 h. A tantalum film, 200-nm thick, was sputtered at 800 °C. Large features are written with optical lithography using Heidelberg MLA 150 aligner and wet-etched using Transene tantalum etchant 111. Ebeam lithography with Ratih EBPG5000 Plus e-beam writer was used to create the junction mask with a double-layer resist consisting of MMA EL11-950 and PMMA A7. The Dolan-bridge junctions were evaporated in the Plassys electron-beam evaporator through double-angle evaporation. 7×7 mm chips are diced and lifted off. The resistance of the on-chip test junctions are measured at room temperature to help preselection. The selected chip was mounted on a printed circuit board, wire bounded, and mounted inside a double-shielded μ -metal can. The coherence times and frequency parameters are listed in Table I. We note that we observe a larger T_1 relaxation time for the $|2\rangle \rightarrow |1\rangle$ transition than for the $|1\rangle \rightarrow |0\rangle$ transition, which seems to contradict the expected scaling $T_1[|j\rangle \rightarrow |j-1\rangle] \sim T_1/j$ [12]. This is due to our choice of centering the filter stop bands around the $|1\rangle \leftrightarrow |2\rangle$ transition, which leads to less protection for the other transitions. Figure 4 shows our device's false-colored optical image. Unused charge lines coupling to the transmon pad are visible, which cause Purcell-limited relaxation times between qudit levels. The cryogenic and room-temperature measurement setup is illustrated in Fig. 8.

APPENDIX B: AWG DRIVE CALIBRATION AND NATIVE GATES

In experiments, all components in the measurement and control chain have their unique frequency response matrix. To map the actual ququart drive strength observed in the experiment to the drive strength specified by the input signal, we need to extract the frequency response matrix between the arbitrary waveform generator (AWG) and the transmon around the frequency band used in the control pulse. In our case we only need to consider the frequencies around the three transitions $|0\rangle \leftrightarrow |1\rangle$, $|1\rangle \leftrightarrow |2\rangle$, and $|2\rangle \leftrightarrow |3\rangle$, each within a 25-MHz frequency span. We further assume that under such a narrow span the frequency response matrix is

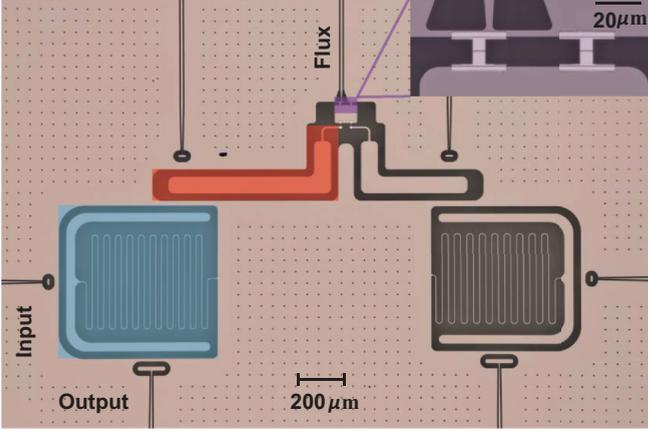


FIG. 4. False-colored optical image of the device. The left transmon (red, left “L” shape) is used as the ququart in the experiment. Single-ququart rotations are sent through the input line coupled to the resonator (blue, bottom left). The SQUID coupler (purple, zoomed-in region) is kept biased to the flux sweet spot throughout the experiment.

constant for each transition $|j-1\rangle \leftrightarrow |j\rangle$, which allows us to directly extract the on-chip drive response factors r_j : r_j is the experimentally measured effective drive strength at the maximum AWG output power for each transition, determined from the observed Rabi rate. For arbitrary transition rates, as they appear in the optimal control pulses, the corresponding AWG output power can be scaled accordingly. The unit of these parameters is $[r_j] = \text{MHz}/\text{AWG}$, where we overload the abbreviation AWG to also refer to the amplitude unit of the waveform generator. We include the calibrated r_j values in Table I. We note that we do not observe the expected scaling $r_j \sim \sqrt{j}$, which we suspect is due to the filter choices and undesired couplings of the input line to external modes on the chip. Especially for the $|2\rangle \leftrightarrow |3\rangle$ transition, we suspect that the unused control lines, which are directly coupled to the transmon pads, cause a stray on-chip mode which blocks this input signal. We do not study this more within the scope of this work but instead adjust our gate decomposition method as outlined in the main text to adapt to these circumstances. No specific changes are required for the optimal control design as the optimization directly takes the drive responses r_j into account, as discussed in Appendix D below.

For all physical rotations in the decomposed gate sequence, the pulse shape is a Gaussian flat-top pulse with a $2\sigma = 5$ ns ramp and a $2\sigma = 5$ ns tail, where σ is the standard deviation of the Gaussian edge. In order to implement an arbitrary rotation angle θ in the j th subspace, we tune the duration of the constant part τ_j of the flat-top section while keeping the power unchanged. The overall gate duration $T_j(\theta) = \tau_j(\theta) + 4\sigma$ is then given by

$$T_j(\theta) = \frac{\theta}{2\pi\sqrt{j}r_j \times 1 \text{ AWG}} - \sqrt{2\pi}\text{erf}(\sqrt{2})\sigma + 4\sigma \quad (\text{B1})$$

$$\approx \frac{\theta}{2\pi\sqrt{j}r_j \times 1 \text{ AWG}} + 4 \text{ ns.}$$

APPENDIX C: NUMERICAL GATE DECOMPOSITION

We elaborate on finding the parameters of the gate decomposition outlined in Sec. II A given a target unitary U . Our native operation set consists of rotations $R_j(\theta, \phi) = Z_j^\dagger(\phi)Y_j(\theta)Z_j(\phi)$ in the subspaces $\{|j-1\rangle, |j\rangle\}$, $j = \{1, 2, 3\}$, which explicitly read

$$R_1(\phi, \theta) = \begin{bmatrix} \cos \frac{\theta}{2} & -e^{i\phi} \sin \frac{\theta}{2} & 0 & 0 \\ e^{-i\phi} \sin \frac{\theta}{2} & \cos \frac{\theta}{2} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad (\text{C1a})$$

$$R_2(\phi, \theta) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \frac{\theta}{2} & -e^{i\phi} \sin \frac{\theta}{2} & 0 \\ 0 & e^{-i\phi} \sin \frac{\theta}{2} & \cos \frac{\theta}{2} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \quad (\text{C1b})$$

$$R_3(\phi, \theta) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos \frac{\theta}{2} & -e^{i\phi} \sin \frac{\theta}{2} \\ 0 & 0 & e^{-i\phi} \sin \frac{\theta}{2} & \cos \frac{\theta}{2} \end{bmatrix}. \quad (\text{C1c})$$

As derived in Ref. [44] and shown in Fig. 1, every single-ququart gate U can be decomposed in the following way:

$$U = Z(\vec{\phi}_4)Y_1(\theta_6)Y_2(\theta_5)Y_3(\theta_4)Z(\vec{\phi}_3) \times Y_1(\theta_3)Y_2(\theta_2)Z(\vec{\phi}_2)Y_1(\theta_1)Z(\vec{\phi}_1). \quad (\text{C2})$$

Here $Z(\vec{\phi}_i)$ is a virtual ququart phase rotation which implements all subspace phase shifts $\vec{\phi}_i = (\phi_{i,1}, \phi_{i,2}, \phi_{i,3})$ at once. This decomposition holds up to a global phase, which is inconsequential in the experiment. While at first it looks like 4×3 phases + 6 angles = 18 real degrees of freedom, which contradicts the 15 real degrees of freedom for an element of $\text{SU}(4)$, it is actually the case that one can fix $\phi_{1,2} = \phi_{1,3} = \phi_{2,3} = 0$ [44], leaving the right number of parameters. We changed the indexing of the subspaces compared to Ref. [44] such that the decomposition uses the fewest gates in the third subspace. This was motivated by the fact that we experimentally observed a significantly reduced achievable Rabi rate in this subspace as presented in Appendix B, which would likely cause much slower gate implementations when using the original decomposition under the same experimental constraints. We emphasize that Eq. (C2) represents the most general decomposition, and in special cases not all gates might be needed. For instance, we have $X_4 = Z(\pi, 0, 0)Y_1(\pi)Y_2(\pi)Y_3(\pi)$, but H_4 requires all gates as shown in Table II.

Given a target unitary U , we determine the decomposition parameters $\theta_i \in [0, \pi]$ and $\vec{\phi}_i \in [0, 2\pi]^3$ via numerical optimization, where we primarily minimize the gate infidelity,

$$1 - F = 1 - \frac{1}{d^2} |\text{Tr}(U^\dagger V)|^2, \quad (\text{C3})$$

TABLE II. Numerically obtained decomposition parameters for several gates according to Eq. (C2), together with sequence execution durations T computed using Eq. (B1). All angles are measured in radians. We recognize decimal approximations of parameters as multiples of $\frac{\pi}{2}$ or special analytical expressions which we show in the lower table.

Gate	$\theta_{1,2,3}$	$\theta_{4,5,6}$	$\vec{\phi}_1$	$\vec{\phi}_2$	$\vec{\phi}_3$	$\vec{\phi}_4$	T [ns]
H_4	$\frac{\pi}{2}$	$\tilde{\theta}_1$	$\frac{3}{2}\pi$	$\tilde{\phi}_1$	$\tilde{\phi}_1$	$\frac{\pi}{2}$	343
	$\tilde{\theta}_2$	$\tilde{\theta}_2$	0	$\frac{3}{2}\pi$	$\tilde{\phi}_2$	$\frac{\pi}{2}$	
	$\tilde{\theta}_3$	$\frac{\pi}{2}$	0	0	$\frac{3}{2}\pi$	$\frac{\pi}{2}$	
$H \otimes H$	$\frac{\pi}{2}$	$\tilde{\theta}_1$	π	π	π	0	365
	$\tilde{\theta}_2$	$\tilde{\theta}_2$	0	0	π	π	
$\mathbb{1} \otimes H$	$\frac{\pi}{2}$	$\frac{\pi}{2}$	π	0	$\frac{\pi}{2}$	$\frac{3}{2}\pi$	139
	0	0	0	0	$\frac{\pi}{2}$	$\frac{\pi}{2}$	
	0	0	0	0	π	0	
$H \otimes \mathbb{1}$	0	$\frac{\pi}{2}$	0	π	0	π	272
	π	π	0	π	π	π	
	$\frac{\pi}{2}$	0	0	0	0	0	
$CX_{c=q_0}^{t=q_1}$	0	π	0	0	0	0	176
	0	0	0	0	π	0	
	0	π	0	π	π	π	
$CX_{c=q_1}^{t=q_0}$	0	0	0	$\frac{3}{2}\pi$	π	$\frac{\pi}{2}$	309
	0	0	0	0	$\frac{\pi}{2}$	$\frac{3}{2}\pi$	
	$\tilde{\theta}_1$	$\tilde{\theta}_2$	$\tilde{\theta}_3$	$\tilde{\phi}_1$	$\tilde{\phi}_2$		
Value	2.09440	1.91063	1.31812	5.17604	5.03414		
$e^{i\tilde{\theta}/2}$	$\frac{1+\sqrt{3}i}{2}$	$\frac{1+\sqrt{2}i}{\sqrt{3}}$	$\frac{\sqrt{5+\sqrt{3}i}}{\sqrt{8}}$	—	—		
$e^{i\tilde{\phi}}$	—	—	—	$\frac{\sqrt{2}-\sqrt{8}i}{\sqrt{10}}$	$\frac{1-3i}{\sqrt{10}}$		

between the target and the decomposition $V = V(\theta_i, \vec{\phi}_i)$. We speak of a *feasible decomposition* if $1 - F \leq 10^{-8}$.

For a specific target, there may be many feasible decompositions; however, not all of them need to be reasonable with respect to experimental implementation. Consider that the identity operation $\mathbb{1} = Y_1(\pi)Y_1(\pi)$ would be compatible with the general formulation (C2) but unnecessarily expensive to implement. Therefore we also want to minimize the estimated overall gate duration

$$T^*(\theta_i) = \sum_{i=1}^6 \frac{\theta_i}{2\pi \sqrt{j} r_{s(i)} \times 1 \text{ AWG}}, \quad (\text{C4})$$

which is determined by the rotation angles. For simplicity, we assume rectangular pulse shapes in this model and consider driving at the maximum AWG output like in the experiment. The function s maps the angle index i to the corresponding subspace index j . However, this approach is not ideal because it could lead to the unnecessary splitting of operations. For instance, if $U = Y_1(\pi)$, feasible decompositions would be $U = Y_1(\pi)$ (ideal) or $U = Y_1(\frac{\pi}{3})Y_1(\frac{\pi}{3})Y_1(\frac{\pi}{3})$ (not ideal). Therefore we actually use a different penalty of the

form

$$\tilde{T}(\theta_i) = \sum_{i=1}^6 \sqrt{\frac{\theta_i}{2\pi \sqrt{j} r_{s(i)} \times 1 \text{ AWG}}}. \quad (\text{C5})$$

This makes use of the fact that for $a, b \geq 0$, $\sqrt{a+b} \leq \sqrt{a} + \sqrt{b}$, with equality if and only if a or b equals zero. Thus splitting of native rotations is discouraged. Other than that, \tilde{T} is still monotonically increasing in each θ_i and minimizing \tilde{T} also leads to reducing T^* .

Additionally, it may be the case that (infinitely) many sets of phase angles $\vec{\phi}_i$ lead to feasible decompositions, even for optimal rotation angles. In order to reduce this degeneracy and furthermore get a better intuition of the action of the gate sequence, we steer the optimization towards phase angles that are multiples of $\frac{\pi}{2}$. In conjunction with Y_j gates, these special phase shifts effectively lead to rotations about the $\pm x$ axis and $\pm y$ axis in the j th subspace's Bloch sphere. We achieve this by introducing another cost term

$$P(\vec{\phi}_i) = \frac{1}{12} \sum_{i=1}^4 \sum_{j=1}^3 \sin^2(2\phi_{i,j}). \quad (\text{C6})$$

Overall, for a given target U we solve the minimization problem

$$\hat{\theta}_i, \hat{\phi}_i = \arg \min_{\theta_i, \phi_i} I + c_{\tilde{T}} \tilde{T} + c_P P \quad (\text{C7})$$

to find a feasible decomposition. We optimize a batch of ten initial guesses for the optimization variables and check if the best-obtained infidelity is below the desired threshold; if not, we keep optimizing batches until we do. For a suitable choice of penalty coefficients $c_{\tilde{T}}$ and c_P , this method is guaranteed to succeed eventually since any unitary can be realized with decomposition (C2). We set $c_{\tilde{T}} = 0.2$ to strongly encourage fast sequences and $c_P = 0.05$ to gently push the optimizer towards preferred phase values and find this setup to work consistently well in practice.

We show the decomposition for the ququart Hadamard H_4 as well as several gates acting on two qubits in Table II and further list the execution time computed by summing up the individual pulse times according to Eq. (B1), rounded to the full nanosecond. The numerically obtained parameter values suggest that they approximate multiples of $\frac{\pi}{2}$ or more complicated analytical expressions. For example, for the gate H_4 we find $\theta_1 \approx 2.09440$, which has the property $e^{i\theta_1/2} \approx \frac{1+\sqrt{3}i}{2}$. This analytical nature is reasonable given the structure of the gates we consider to factor unitaries, which motivates studying qudit gate decompositions further and finding better ways to obtain these expressions directly.

APPENDIX D: OPTIMAL CONTROL

Our native as well as optimal-control pulses implement a desired gate U_T (duration T) in the interaction frame defined by $|\psi_{\text{rot}}(t)\rangle = e^{i\mathcal{H}_0 t} |\psi_{\text{lab}}(t)\rangle$ with the Hamiltonian

$$\mathcal{H}_{\text{int}} = \sum_{j=1}^3 \frac{\sqrt{j}}{2} [I_j(t)Y_j + Q_j(t)X_j],$$

where I_j and Q_j are the quadrature signals for the j th subspace. In the optimal-control setting, they are built from time-local B-spline basis functions, as defined in Sec. II B and used in JUQBOX [25]:

$$I_j(t) = \sum_{s=1}^{N_s} \alpha_{j,s}^I B_s(t), \quad Q_j(t) = \sum_{s=1}^{N_s} \alpha_{j,s}^Q B_s(t). \quad (\text{D1})$$

However, computation in this frame cannot take the leakage into higher energy levels into account; therefore we consider a different rotating frame transformation $|\psi_{\text{rot}}(t)\rangle = e^{i\omega_{\text{rot}}\hat{n}}|\psi_{\text{lab}}(t)\rangle$ to solve the open-loop optimal control problem instead. This yields the Hamiltonian

$$\mathcal{H}_{\text{rot}}(t) = \sum_{n=0}^{\tilde{d}-1} (\epsilon_n - n\omega_{\text{rot}})|n\rangle\langle n| + \frac{1}{2}[I(t)(-i)(a - a^\dagger) + Q(t)(a + a^\dagger)], \quad (\text{D2})$$

with $I(t) + iQ(t) = \sum_{j=1}^3 [I_j(t) + iQ_j(t)]e^{i(\omega_j - \omega_{\text{rot}})t}$. The Hamiltonian is truncated at $\tilde{d} = 5$ levels to account for leakage, and we find this to be sufficient in simulation (see Appendix G). Accordingly, the target gate has to be transformed to $U'_T = W_{\text{trans},T} U_T$, where $W_{\text{trans},T} = e^{i\omega_{\text{rot}}T\hat{n}} e^{-iH_0T}$. The infidelity cost term becomes

$$1 - F_{\vec{\phi}} = 1 - \frac{1}{d^2} |\text{Tr}(U_T'^{\dagger} Z(\vec{\phi}) V_{\vec{a}})|^2, \quad (\text{D3})$$

with the phase corrections $\vec{\phi}$ and pulse parameters $\vec{a} = (\alpha_{j,s}^I, \alpha_{j,s}^Q)_{j,s}$ that are optimized over.

We choose $\omega_{\text{rot}} = \omega_1$ to reduce the oscillating behavior of the functions I and Q , which allows the choice of larger time steps in the numerical solution of the dynamics. Typically we set the time-step size $\Delta t = 0.03 \dots 0.3$ ns. In most optimal control scenarios the gate time T needs to be fixed *a priori*, and infidelity convergence is an indicator if the chosen duration was sufficient to realize the gate. For the pulse parametrization at hand, we choose the number of B-splines N_s proportional to the gate time, $N_s = \lfloor \frac{T}{10\text{ns}} \rfloor$, to ensure similar B-spline densities for pulses of different durations.

We try to find optimal control pulses of short durations as decoherence limits the capabilities of our device, especially due to increased dephasing in the $\{|2\rangle, |3\rangle\}$ subspace (see Table I). In this work we perform the duration optimization by hand, which is achieved by reducing the gate time further and further while making sure the numerical infidelity stays below the threshold $1 - F_{\vec{\phi}} \leq 10^{-4}$. We note that this is a rather simple method that could be improved. One alternative approach would be to make the duration T an optimization variable, rescale all time variables with $\frac{1}{T}$ and all frequency variables with T , respectively, and then solve the optimal control task in the time interval $[0, 1]$. However, we observe worse convergence and higher obtained gate durations in this case. Despite its simpler nature, our method is sufficient for our purposes and manages to show the impact of performing final phase corrections using a virtual Z (VZ) on the gate time T . Figure 5 shows H_4 pulse implementations both with (solid lines) and

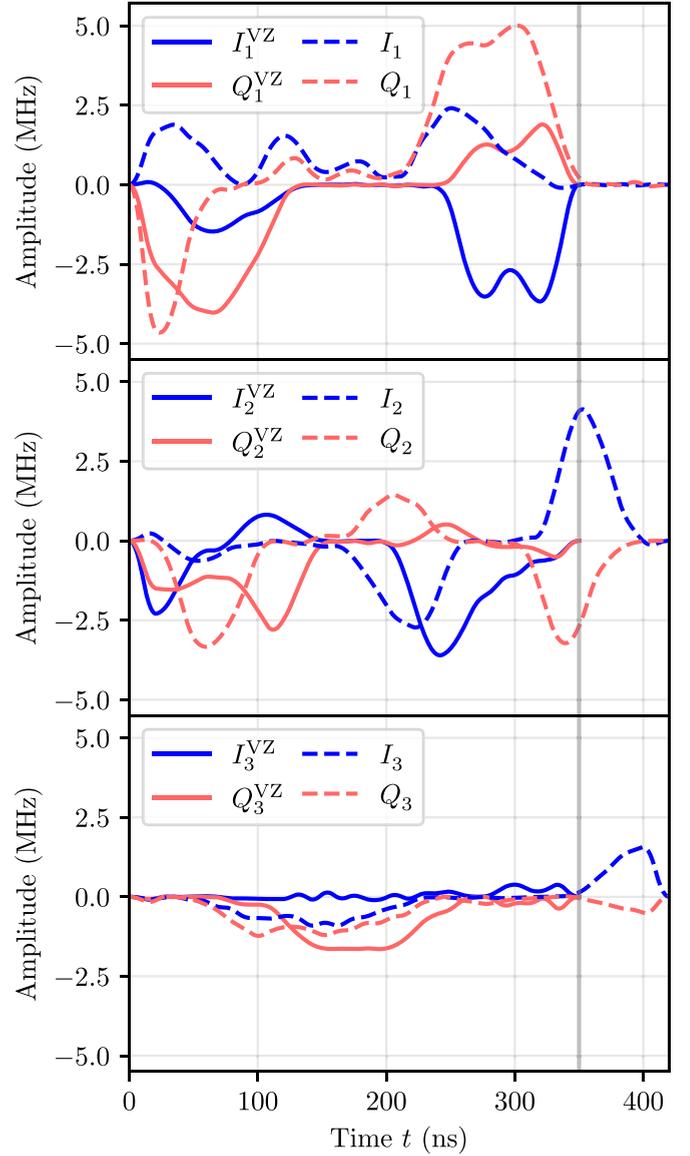


FIG. 5. Optimal control pulses for ququart Hadamard H_4 , with final VZ correction (solid lines, $T = 350$ ns) and without final VZ correction (dashed lines, $T = 420$ ns). The vertical gray line marks the duration of the former.

without a final VZ (dashed lines), where the difference in duration is highlighted by the solid gray line. Interestingly, we find that for the $H \otimes H$ tensor product optimizing with a final VZ does not lead to a reduced gate duration; therefore we choose the optimization result without a VZ in our experiments.

In the following, we provide details on the additionally introduced cost terms like mentioned in Sec. II B. One important hardware constraint the optimized control pulse needs to obey is the power limit of the arbitrary waveform generator, which imposes $|\gamma(t)| \leq 1$ AWG. Here we overload the abbreviation AWG to use it as the corresponding power unit and further reiterate that the conversion from the drive strength on-chip

(in MHz) to AWG power is frequency dependent. With Eq. (4) this constraint can be ensured by imposing

$$|\gamma(t)| \leq \sum_{j=1}^3 \frac{|\Gamma_j(t)|}{r_j} \equiv \tilde{\Gamma}(t) \stackrel{!}{\leq} 1 \text{ AWG}, \quad (\text{D4a})$$

$$\Rightarrow \sum_{j=1}^3 \frac{1}{r_j} \sqrt{(\alpha_{j,s}^I)^2 + (\alpha_{j,s}^Q)^2} \leq 1 \text{ AWG}. \quad (\text{D4b})$$

At this time JUQBOX does not support constraints that include multiple optimization variables. Without any additional knowledge of the structure of the target gate, the most general approach is assigning one-third of the power budget to each of the three carrier waves, which can be done by enforcing $|\alpha_{j,s}^I|, |\alpha_{j,s}^Q| \leq \frac{r_j}{3\sqrt{2}}$. However, this may drastically reduce the potential speed of the gate implementation, as no transition is allowed to be driven at more than a third of the maximum power.

We transition to using BOULDER OPAL [30] while sticking to the JUQBOX pulse parametrization to overcome this limitation. While BOULDER OPAL currently cannot directly implement constraints of the form (D4b) either, it allows for a more flexible formulation of the quantum control task, construction of signals, and definition of cost terms. Let us consider the functions

$$f(x) = [1 - g(x - a)]x + g(x - a)a, \quad (\text{D5})$$

$$g(x) = 1 - \frac{\tanh\left(-\frac{1000x}{a}\right) + 1}{2}, \quad (\text{D6})$$

where g serves as a differentiable version of the Heaviside step function $\Theta(x)$ that can be implemented in BOULDER OPAL's framework. f can be seen as a suppression function with threshold a , which means $f(x) \approx x$ for $x < a$ and $f(x) \approx a$ for $x \geq a$. Using this we can construct a suppressed signal

$$\tilde{\Gamma}^{\text{sup}}(t) = f(\tilde{\Gamma}(t)) \quad (\text{D7})$$

and a penalty term

$$P_{\text{amp}} = \int_0^T |\tilde{\Gamma}(t) - \tilde{\Gamma}^{\text{sup}}(t)| dt, \quad (\text{D8})$$

which captures how much the original signal $\tilde{\Gamma}$ violates the power constraint. We set the threshold amplitude to $a = 0.95$ AWG to allow the optimizer some slack to the critical value of 1 AWG.

The second cost term addresses the assumption of a constant frequency response matrix in the range $[-25 \text{ MHz}, 25 \text{ MHz}]$ around the three transition frequencies, which we discuss in Appendix B. In order to be consistent with this assumption in the optimal control pulse design, we introduce a filter penalty to steer the optimization towards pulses with narrow peaks in the Fourier spectrum:

$$P_{\text{fil}} = \sum_{j=1}^3 \int_0^T |\Gamma_j(t) - \Gamma_j^{\text{fil}}(t)| dt, \quad (\text{D9})$$

where Γ_j^{fil} is the convolution of Γ_j with a sinc kernel to limit the bandwidth around each transition to $[-25 \text{ MHz}, 25 \text{ MHz}]$. The computation of the convolution is a built-in feature in BOULDER OPAL.

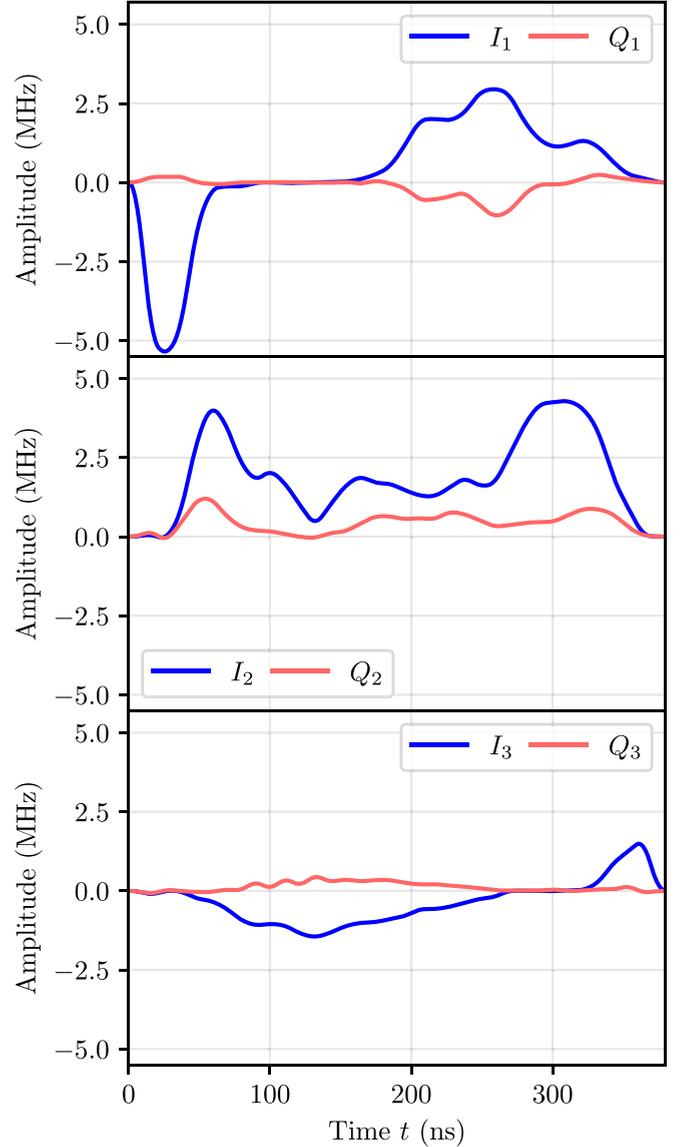


FIG. 6. Optimal control pulses for the tensor product $H \otimes H$ without a final VZ correction, $T = 380$ ns. Allowing a correction did not lead to a significantly reduced gate duration.

The full optimal control problem then reads

$$\hat{\alpha}_{j,s}, \hat{\phi} = \arg \min_{\alpha_{j,s}, \phi} I_{\hat{\phi}} + c_{\text{amp}} P_{\text{amp}} + c_{\text{fil}} P_{\text{fil}}, \quad (\text{D10})$$

where we empirically choose $c_{\text{amp}} = c_{\text{fil}} = \frac{1}{2}$. We show the obtained pulse implementations for our gates of interest H_4 and $H \otimes H$ in Figs. 5 and 6, respectively.

APPENDIX E: TOMOGRAPHY AND READOUT HISTOGRAM

To perform single-ququart quantum-state tomography (QST), we apply one of 16 different postrotation gates prior to the single-shot readouts. This set \mathcal{T} is shown in Table III, where $\mathbb{1}$ is the identity gate, and the rotations R_1 , R_2 , and R_3 are defined in Appendix C.

TABLE III. Sets of prerotation (\mathcal{S}) and postrotation (\mathcal{T}) gates to implement quantum-state tomography and quantum-process tomography.

Prerotation set \mathcal{S}	Postrotation set \mathcal{T}
$\mathbb{1}_4$	$\mathbb{1}_4$
$R_1(0, \frac{\pi}{2})$	$R_1(0, \frac{\pi}{2})$
$R_1(\frac{\pi}{2}, \frac{\pi}{2})$	$R_1(\frac{\pi}{2}, \frac{\pi}{2})$
$R_1(0, \pi)$	$R_1(0, \pi)$
$R_2(0, \frac{\pi}{2})$	$R_2(0, \frac{\pi}{2})$
$R_2(\frac{\pi}{2}, \frac{\pi}{2})$	$R_2(\frac{\pi}{2}, \frac{\pi}{2})$
$R_2(0, \pi)R_1(0, \frac{\pi}{2})$	$R_2(0, \frac{\pi}{2})R_1(0, \pi)$
$R_2(0, \pi)R_1(\frac{\pi}{2}, \frac{\pi}{2})$	$R_2(\frac{\pi}{2}, \frac{\pi}{2})R_1(0, \pi)$
$R_2(0, \pi)R_1(0, \pi)$	$R_2(0, \pi)R_1(0, \pi)$
$R_3(0, \frac{\pi}{2})$	$R_3(0, \frac{\pi}{2})$
$R_3(\frac{\pi}{2}, \frac{\pi}{2})$	$R_3(\frac{\pi}{2}, \frac{\pi}{2})$
$R_3(0, \pi)R_2(0, \frac{\pi}{2})$	$R_3(0, \frac{\pi}{2})R_2(0, \pi)$
$R_3(0, \pi)R_2(\frac{\pi}{2}, \frac{\pi}{2})$	$R_3(\frac{\pi}{2}, \frac{\pi}{2})R_2(0, \pi)$
$R_3(0, \pi)R_2(0, \pi)R_1(0, \frac{\pi}{2})$	$R_3(0, \frac{\pi}{2})R_2(0, \pi)R_1(0, \pi)$
$R_3(0, \pi)R_2(0, \pi)R_1(\frac{\pi}{2}, \frac{\pi}{2})$	$R_3(\frac{\pi}{2}, \frac{\pi}{2})R_2(0, \pi)R_1(0, \pi)$
$R_3(0, \pi)R_2(0, \pi)R_1(0, \pi)$	$R_3(0, \pi)R_2(0, \pi)R_1(0, \pi)$

The single-shot readouts are collected after applying the postrotations. The readout histogram and heat map of the assignment fidelity matrix are visualized in Fig. 7. We apply the confusion matrix, which is the inverse matrix of the normalized readout heat map, to the single-shot data to reduce the readout error. Maximum likelihood estimation (MLE) is then used to reconstruct the physical density matrix.

The single-ququart process tomography (QPT) [46] is performed by sandwiching the unknown quantum channel \mathcal{E} with different combinations of prerotation and postrotation gates, where the prerotation set \mathcal{S} is shown in Table III.

After applying each prerotation and the gate for characteristics, the density matrix is reconstructed through MLE QST. The full quantum channel $\mathcal{E}(\rho) = \sum_{i,j} \chi_{ij} B_i \rho B_j^\dagger$ is calculated based on Ref. [50] with either the basis $\{B_i\} = \{Z_4^m X_4^n\}$ (ququart scenario) or $\{B_i\} = \{\sigma_m \otimes \sigma_n\}$ (two-qubit scenario, ququart encoding), $m, n = \{0, 1, 2, 3\}$. Here Z_4 and X_4 are the single-ququart Pauli operators, and σ_m denotes one of the qubit Pauli operators $\{\mathbb{1}, X, Y, Z\}$. The operator basis elements are orthogonal under the Hilbert-Schmidt inner product: $\text{Tr}[B_i^\dagger B_j] = d \delta_{ij}$. The process fidelity of two chan-

TABLE IV. Single-ququart gate fidelity characterized by QPT, for decomposed gates (DEC) and optimally controlled gates (QOC). In the QOC implementation of $H \otimes H$ we only considered the case without a final VZ, see Appendix D.

Gate	DEC	QOC	
		No VZ	With VZ
$\mathbb{1}_4$	—	89.88%	—
Z_4	—	88.85%	—
H_4	84.30%	81.35%	86.18%
$H \otimes H$	83.53%	84.06%	—*
$X \otimes H$	83.47%	—	—
$H \otimes \mathbb{1}$	84.80%	—	—
$CX_{c=q_1}^{t=q_0}$	85.36%	—	—

nels with process matrices χ_1 and χ_2 is then given by

$$F = |\text{Tr}[\chi_1 \chi_2]|, \quad (\text{E1})$$

which in the case of unitary channels recovers the expression (C3). Table IV shows our experimental QPT results.

APPENDIX F: QUQUART RB

In the RB protocol [47] one constructs sequences of m Clifford gates $C_m C_{m-1} \cdots C_1$ and further computes inversion gates C_{inv} , which invert the effect of the sequences: $C_{\text{inv}} C_m C_{m-1} \cdots C_1 = \mathbb{1}$. C_{inv} can be computed efficiently due to the Gottesman-Knill theorem [51]. In the experiment, one initializes the system in the ground state, applies the sequence, and measures the population in the ground state (survival probability) P_m .

Averaging over different sequences effectively corresponds to a depolarizing channel, and the depolarization parameter p can be extracted from a fit to the exponential decay $\langle P_m \rangle = A p^m + B$. The error per Clifford gate (or for a specific Clifford gate in the IRB setting) is then determined as described in Sec. III B.

Since RB relies on sampling from the Clifford group, choosing the appropriate Clifford group for a certain scenario is important. The Clifford group \mathcal{C} is defined as the group which normalizes the Pauli group \mathcal{P} . In the ququart case, the Pauli group \mathcal{P}_4 consists of products of X_4 and Z_4 , which are normalized by H_4 and $S_4 = \text{diag}(1, \sqrt{i}, i, \sqrt{i})$ [52]. Z_4 trivially is an element of the ququart Clifford group \mathcal{C}_4 and we check numerically that $\{H_4, Z_4, S_4\}$ can generate the predicted number of 768 elements in \mathcal{C}_4 [53]. In the case of two qubits, the Pauli group $\mathcal{P}_2^{\otimes 2}$ is given by the tensor product of the single-qubit operators $\mathbb{1}, X, Y$ and Z , and the gates $\{H, S, CX\}$ generate the Clifford group $\mathcal{C}_2^{\otimes 2}$ with 11 520 elements (up to scalars) [48].

For the ququart case, we sample from \mathcal{C}_4 which we construct explicitly, and for the two-qubit case we use QISKIT [54] to produce RB sequences. In either case we decompose the matrices for each gate in a sequence according to decomposition (C2) in order to implement the RB experiment on our transmon. RB and IRB results are summarized in Table V.

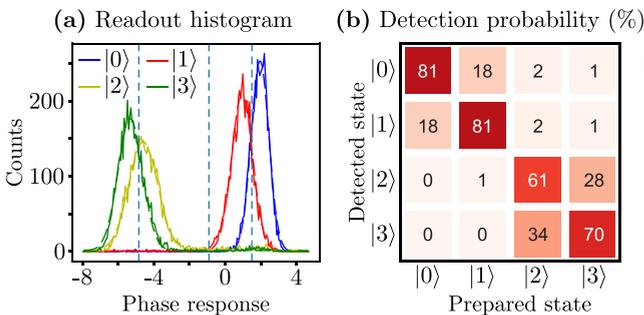


FIG. 7. Readout characterization. (a) Single-shot histogram after phase rotation and (b) heat map of detection probabilities (in %). Single-ququart basis states are prepared and measured 5000 times.

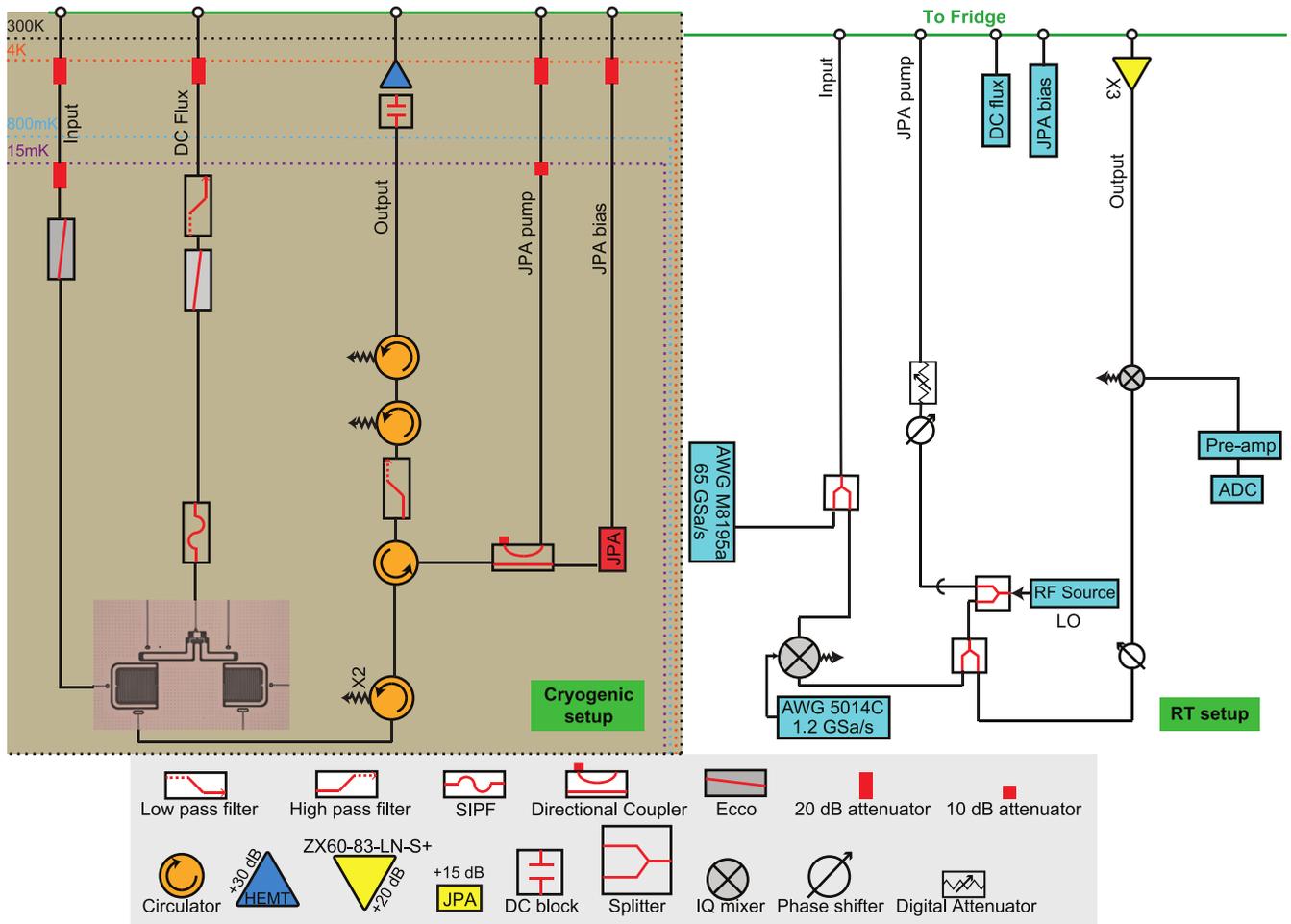


FIG. 8. Measurement setup.

APPENDIX G: ERROR BUDGET

Table VI shows the infidelities of different ququart gates extracted from simulating the Lindblad master equation. We compare the noise-free case with the noisy case, where for the latter we include decay and dephasing collapse operators for each subspace using the experimentally measured lifetimes. The simulation time step is set to $dt = 0.001$ ns for high

TABLE V. Single-ququart gate fidelity characterized by different types of RB experiments, for decomposed gates (DEC) and quantum optimally controlled gates (QOC). The H_4 gate fidelity is characterized using the Clifford group C_4 , other gates are characterized using the Clifford group $C_2^{\otimes 2}$. Base fidelity for RB with C_4 is 96.22(14)%, and with $C_2^{\otimes 2}$ is 95.84(5)%. We only consider the VZ implementation for the QOC H_4 pulse.

Gate	DEC	QOC
H_4	96.18(32)%	96.41(32)%
$H \otimes H$	95.17(19)%	95.98(19)%
$\mathbb{1} \otimes H$	98.51(13)%	—
$H \otimes \mathbb{1}$	96.70(29)%	—
$CX_{c=q_0}^{t=q_1}$	96.68(24)%	—
$CX_{c=q_1}^{t=q_0}$	98.84(17)%	—

numerical precision, which is over an order of magnitude lower than what was used for the pulse optimization (see Appendix D). We perform all simulations with the lowest $5 \times$ energy levels of the transmon to account for a leakage into higher levels. By taking the trace of the 4×4 submatrix which describes the ququart state, we verify for each simulation result that leakage into the fifth state is below 0.01%, justifying the truncation of the Hamiltonian at $\vec{d} = 5$ levels.

TABLE VI. Ququart gate infidelities extracted from simulation. DEC: decomposition, QOC: quantum optimal control.

Gate	Decoherence-free	Decoherence
H_4 DEC	0.42%	4.53%
$H \otimes H$ DEC	0.41%	4.78%
$CX_{c=q_0}^{t=q_1}$ DEC	0.38%	4.11%
$CX_{c=q_1}^{t=q_0}$ DEC	0.03%	2.18%
$H \otimes I$ DEC	0.62%	3.89%
$I \otimes H$ DEC	0.11%	1.82%
H_4 QOC	0.08%	4.28%
$H \otimes H$ QOC	0.26%	4.81%

We observe that the optimal control pulses can theoretically reduce the decoherence-free infidelities by more than 30%, and future work will be optimizing pulse duration to re-

duce the decoherence error further. The difference in infidelity between the ideal simulation here and the optimization target $1 - F_{\phi}^- \leq 10^{-4}$ arises from the increased time-step precision.

-
- [1] Y. Wang, Z. Hu, B. C. Sanders, and S. Kais, Qudits and high-dimensional quantum computing, *Front. Phys.* **8**, (2020).
- [2] A. S. Nikolaeva, E. O. Kiktenko, and A. K. Fedorov, Efficient realization of quantum algorithms with qudits, [arXiv:2111.04384](https://arxiv.org/abs/2111.04384).
- [3] Y. Deller, S. Schmitt, M. Lewenstein, S. Lenk, M. Federer, F. Jendrzejewski, P. Hauke, and V. Kasper, Quantum approximate optimization algorithm for qudit systems, *Phys. Rev. A* **107**, 062410 (2023).
- [4] A. Bocharov, M. Roetteler, and K. M. Svore, Factoring with qutrits: Shor’s algorithm on ternary and metaplectic quantum architectures, *Phys. Rev. A* **96**, 012306 (2017).
- [5] D. M. Nguyen and S. Kim, Quantum key distribution protocol based on modified generalization of Deutsch-Jozsa algorithm in d -level quantum system, *Int. J. Theor. Phys.* **58**, 71 (2019).
- [6] J. M. Baker, C. Duckering, and F. T. Chong, Efficient quantum circuit decompositions via intermediate qudits, in *2020 IEEE 50th International Symposium on Multiple-Valued Logic (ISMVL)* (IEEE, New York, 2020), pp. 303–308.
- [7] J. M. Baker, C. Duckering, P. Gokhale, N. C. Brown, K. R. Brown, and F. T. Chong, Improved quantum circuits via intermediate qutrits, *ACM Trans. Quantum Comput.* **1**, 1 (2020).
- [8] E. Kapit, Hardware-efficient and fully autonomous quantum error correction in superconducting circuits, *Phys. Rev. Lett.* **116**, 150501 (2016).
- [9] Z. Li, T. Roy, D. R. Perez, K.-H. Lee, E. Kapit, and D. I. Schuster, Autonomous error correction of a single logical qubit using two transmons, [arXiv:2302.06707](https://arxiv.org/abs/2302.06707).
- [10] S. Muralidharan, C.-L. Zou, L. Li, J. Wen, and L. Jiang, Overcoming erasure errors with multilevel systems, *New J. Phys.* **19**, 013026 (2017).
- [11] R. Majumdar, S. Basu, S. Ghosh, and S. Sur-Kolay, Quantum error-correcting code for ternary logic, *Phys. Rev. A* **97**, 052302 (2018).
- [12] M. S. Blok, V. V. Ramasesh, T. Schuster, K. O’Brien, J. M. Kreikebaum, D. Dahlen, A. Morvan, B. Yoshida, N. Y. Yao, and I. Siddiqi, Quantum information scrambling on a superconducting qutrit processor, *Phys. Rev. X* **11**, 021010 (2021).
- [13] M. Ringbauer, M. Meth, L. Postler, R. Stricker, R. Blatt, P. Schindler, and T. Monz, A universal qudit quantum processor with trapped ions, *Nat. Phys.* **18**, 1053 (2022).
- [14] P. Hrmo, B. Wilhelm, L. Gerster, M. W. van Mourik, M. Huber, R. Blatt, P. Schindler, T. Monz, and M. Ringbauer, Native qudit entanglement in a trapped ion quantum processor, *Nat. Commun.* **14**, 2242 (2023).
- [15] Y.-H. Luo, H.-S. Zhong, M. Erhard, X.-L. Wang, L.-C. Peng, M. Krenn, X. Jiang, L. Li, N.-L. Liu, C.-Y. Lu, A. Zeilinger, and J.-W. Pan, Quantum teleportation in high dimensions, *Phys. Rev. Lett.* **123**, 070505 (2019).
- [16] 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).
- [17] N. Goss, A. Morvan, B. Marinelli, B. K. Mitchell, L. B. Nguyen, R. K. Naik, L. Chen, C. Jünger, J. M. Kreikebaum, D. I. Santiago, J. J. Wallman, and I. Siddiqi, High-fidelity qutrit entangling gates for superconducting circuits, *Nat. Commun.* **13**, 7481 (2022).
- [18] T. Roy, Z. Li, E. Kapit, and D. I. Schuster, Realization of two-qutrit quantum algorithms on a programmable superconducting processor, [arXiv:2211.06523](https://arxiv.org/abs/2211.06523).
- [19] K. Luo, W. Huang, Z. Tao, L. Zhang, Y. Zhou, J. Chu, W. Liu, B. Wang, J. Cui, S. Liu, F. Yan, M.-H. Yung, Y. Chen, T. Yan, and D. Yu, Experimental realization of two qutrits gate with tunable coupling in superconducting circuits, *Phys. Rev. Lett.* **130**, 030603 (2023).
- [20] T. Roy, Z. Li, E. Kapit, and D. I. Schuster, Two-qutrit quantum algorithms on a programmable superconducting processor, *Phys. Rev. Appl.* **19**, 064024 (2023).
- [21] A. Litteken, L. M. Seifert, J. Chadwick, N. Nottingham, F. T. Chong, and J. M. Baker, Qompress: Efficient compilation for ququarts exploiting partial and mixed radix operations for communication reduction, in *Proceedings of the 28th ACM International Conference on Architectural Support for Programming Languages and Operating Systems, ASPLOS 2023* (Association for Computing Machinery, New York, NY, USA, 2023), Vol. 2, pp. 646–659.
- [22] A. Litteken, L. M. Seifert, J. D. Chadwick, N. Nottingham, T. Roy, Z. Li, D. Schuster, F. T. Chong, and J. M. Baker, Dancing the Quantum Waltz: Compiling Three-Qubit Gates on Four Level Architectures, in *Proceedings of the 50th Annual International Symposium on Computer Architecture* (Association for Computing Machinery, New York, NY, USA, 2023), pp. 1–14.
- [23] S. Cao, M. Bakr, G. Campanaro, S. D. Fasciati, J. Wills, D. Lall, B. Shteynas, V. Chidambaram, I. Rungger, and P. Leek, Emulating two qubits with a four-level transmon qudit for variational quantum algorithms, [arXiv:2303.04796](https://arxiv.org/abs/2303.04796).
- [24] Y. Dong, Q. Liu, J. Wang, Q. Li, X. Yu, W. Zheng, Y. Li, D. Lan, X. Tan, and Y. Yu, Simulation of two-qubit gates with a superconducting qudit, *Phys. Status Solidi B* **259**, 2100500 (2022).
- [25] N. Anders Petersson and F. Garcia, Optimal control of closed quantum systems via B-splines with carrier waves, *SIAM J. Sci. Comput.* **44**, A3592 (2022).
- [26] S. Günther, N. A. Petersson, and J. L. Dubois, *Quandary: An open-source C++ package for high-performance optimal control of open quantum systems* (IEEE Computer Society, 2021), pp. 88–98.
- [27] M. H. Goerz, D. Basilewitsch, F. Gago-Encinas, M. G. Krauss, K. P. Horn, D. M. Reich, and C. P. Koch, Krotov: A Python implementation of Krotov’s method for quantum optimal control, *SciPost Phys.* **7**, 080 (2019).
- [28] N. Khaneja, T. Reiss, C. Kehlet, T. Schulte-Herbrüggen, and S. J. Glaser, Optimal control of coupled spin dynamics: Design of NMR pulse sequences by gradient ascent algorithms, *J. Magn. Reson.* **172**, 296 (2005).

- [29] T. Propson, B. E. Jackson, J. Koch, Z. Manchester, and D. I. Schuster, Robust quantum optimal control with trajectory optimization, *Phys. Rev. Appl.* **17**, 014036 (2022).
- [30] Q-CTRL, Boulder Opal (2023).
- [31] H. Ball, M. J. Biercuk, A. R. R. Carvalho, J. Chen, M. Hush, L. A. D. Castro, L. Li, P. J. Liebermann, H. J. Slatyer, C. Edmunds, V. Frey, C. Hempel, and A. Milne, Software tools for quantum control: Improving quantum computer performance through noise and error suppression, *Quantum Sci. Technol.* **6**, 044011 (2021).
- [32] R.-B. Wu, B. Chu, D. H. Owens, and H. Rabitz, Data-driven gradient algorithm for high-precision quantum control, *Phys. Rev. A* **97**, 042122 (2018).
- [33] D. Pizzuoli, C. J. Wood, D. J. Egger, B. Rosand, and K. Ueda, Qiskit dynamics: A python package for simulating the time dynamics of quantum systems, *J. Open Source Softw.* **8**, 5853 (2023).
- [34] S. Günther, N. A. Petersson, and J. L. DuBois, Quantum optimal control for pure-state preparation using one initial state, *AVS Quantum Sci.* **3**, 043801 (2021).
- [35] A. J. Goldschmidt, J. L. DuBois, S. L. Brunton, and J. N. Kutz, Model predictive control for robust quantum state preparation, *Quantum* **6**, 837 (2022).
- [36] O. R. Meitei, B. T. Gard, G. S. Barron, D. P. Pappas, S. E. Economou, E. Barnes, and N. J. Mayhall, Gate-free state preparation for fast variational quantum eigensolver simulations, *npj Quantum Inf.* **7**, 155 (2021).
- [37] X. Wu, S. L. Tomarken, N. A. Petersson, L. A. Martinez, Y. J. Rosen, and J. L. DuBois, High-fidelity software-defined quantum logic on a superconducting qudit, *Phys. Rev. Lett.* **125**, 170502 (2020).
- [38] L. M. Seifert, J. Chadwick, A. Litteken, F. T. Chong, and J. M. Baker, Time-efficient qudit gates through incremental pulse re-seeding, in *2022 IEEE International Conference on Quantum Computing and Engineering (QCE)* (IEEE Computer Society, Washington, DC, 2022), pp. 304–313.
- [39] J. Chadwick and F. T. Chong, Efficient control pulses for continuous quantum gate families through coordinated re-optimization, [arXiv:2302.01553](https://arxiv.org/abs/2302.01553).
- [40] A. R. R. Carvalho, H. Ball, M. J. Biercuk, M. R. Hush, and F. Thomsen, Error-robust quantum logic optimization using a cloud quantum computer interface, *Phys. Rev. Appl.* **15**, 064054 (2021).
- [41] D. Gottesman, Fault-tolerant quantum computation with higher-dimensional systems, in *Quantum Computing and Quantum Communications*, Lecture Notes in Computer Science, edited by C. P. Williams (Springer, Berlin, Heidelberg, 1999), pp. 302–313.
- [42] J. Koch, T. M. Yu, J. Gambetta, A. A. Houck, D. I. Schuster, J. Majer, A. Blais, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, Charge-insensitive qubit design derived from the Cooper pair box, *Phys. Rev. A* **76**, 042319 (2007).
- [43] 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).
- [44] P. Dita, Factorization of unitary matrices, *J. Phys. A: Math. Gen.* **36**, 2781 (2003).
- [45] N. A. Petersson, F. M. Garcia, A. E. Copeland, Y. L. Rydin, and J. L. DuBois, Discrete adjoints for accurate numerical optimization with application to quantum control, [arXiv:2001.01013](https://arxiv.org/abs/2001.01013).
- [46] 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).
- [47] E. Magesan, J. M. Gambetta, and J. Emerson, Scalable and robust randomized benchmarking of quantum processes, *Phys. Rev. Lett.* **106**, 180504 (2011).
- [48] P. Selinger, Generators and relations for n-qubit Clifford operators, *Log. Methods Comput. Sci.* **11**, 1570 (2015).
- [49] E. Magesan, J. M. Gambetta, B. R. Johnson, C. A. Ryan, J. M. Chow, S. T. Merkel, M. P. da Silva, G. A. Keefe, M. B. Rothwell, T. A. Ohki, M. B. Ketchen, and M. Steffen, Efficient measurement of quantum gate error by interleaved randomized benchmarking, *Phys. Rev. Lett.* **109**, 080505 (2012).
- [50] M. Mohseni, A. T. Rezakhani, and D. A. Lidar, Quantum-process tomography: Resource analysis of different strategies, *Phys. Rev. A* **77**, 032322 (2008).
- [51] D. Gottesman, The Heisenberg representation of quantum computers, *Group22: Proceedings of the XXII International Colloquium on Group Theoretical Methods in Physics*, edited by S. P. Corney, R. Delbourgo, and P. D. Jarvis (International Press, Cambridge, MA, 1999), pp. 32–43. [arXiv:quant-ph/9807006](https://arxiv.org/abs/quant-ph/9807006).
- [52] E. Hostens, J. Dehaene, and B. De Moor, Stabilizer states and Clifford operations for systems of arbitrary dimensions and modular arithmetic, *Phys. Rev. A* **71**, 042315 (2005).
- [53] J. Tolar, On Clifford groups in quantum computing, *J. Phys.: Conf. Ser.* **1071**, 012022 (2018).
- [54] QISKIT contributors, QISKIT: An open-source framework for quantum computing (2023).