

Fast unconditional reset and leakage reduction of a tunable superconducting qubit via an engineered dissipative bath

Gihwan Kim^{1,2,*}, Andreas Butler^{1,2,*}, Vinicius S. Ferreira^{1,2,†}, Xueyue (Sherry) Zhang^{1,2,‡}, Alex Hadley¹, Eunjong Kim^{3,4} and Oskar Painter^{1,2,§}

¹*Kavli Nanoscience Institute and Thomas J. Watson, Sr., Laboratory of Applied Physics, California Institute of Technology, Pasadena, California 91125, USA*

²*Institute for Quantum Information and Matter, California Institute of Technology, Pasadena, California 91125, USA*

³*Department of Physics & Astronomy, Seoul National University, Seoul 08826, South Korea*

⁴*Institute of Applied Physics, Seoul National University, Seoul 08826, South Korea*

 (Received 26 November 2024; revised 14 February 2025; accepted 5 June 2025; published 7 July 2025)

Rapid and accurate initialization of qubits, known as reset, is a crucial building block for various tasks in quantum information processing, such as quantum error correction and estimation of the statistics of noisy quantum devices with many qubits. We demonstrate unconditional reset of a frequency-tunable transmon qubit that simultaneously resets multiple excited states using a metamaterial waveguide engineered to provide a cold bath over a wide spectral range, while providing strong protection against Purcell decay of the qubit. We report a reset error below 0.13% (0.16%) when prepared in the first (second) excited state of the transmon within 88 ns. In addition, through the sharp roll-off in the density of states of the metamaterial waveguide, we implement a leakage-reduction unit that selectively resets the transmon's second excited state to 0.285(3)% residual population within 44 ns while acting trivially in the computational subspace as an identity operation that preserves encoded information with an infidelity of 0.72(1)%.

DOI: [10.1103/6wc6-78y3](https://doi.org/10.1103/6wc6-78y3)

I. INTRODUCTION

Fast and high-fidelity reset of qubits into fiducial states is a necessary capability of quantum processors designed for demanding quantum information and computation tasks [1]. In the context of quantum error correction, high-fidelity qubit reset minimizes faulty state preparation of auxiliary qubits, and rapid qubit reset reduces the idling error experienced by spectator qubits [2–7]. In addition, operations capable of resetting leakage states into the computational subspace offer the benefit of converting uncorrectable leakage errors into correctable errors [8–12]. More generally, fast and accurate qubit reset improves experimental repetition rates and state-preparation fidelities,

bringing the measurement times of higher-order statistics of noisy many-qubit devices within a feasible range.

In superconducting qubit systems, reset schemes that use projective single-shot measurement followed by conditional single-qubit control can realize reset within a few hundred nanoseconds [13–16]; however, measurement-based conditional reset is limited in speed by inevitable temporal overhead in postprocessing, and it is limited in accuracy by readout infidelity, for example due to measurement-induced state transitions [17,18].

More recently, state-of-the-art unconditional transmon qubit reset has been achieved by activating resonant exchange coupling between a transmon and a damped auxiliary mode such as a readout resonator and allowing the system to decay to its collective ground state [8,19–24]. Such reset protocols prepare the ground state independent of the prior state of the qubit, and they are thus not limited by readout time and fidelity. In addition, unconditional reset schemes help reduce the impact of state-misclassification errors on fault-tolerant logical operations [7]. Nonetheless, the speed of these reset schemes is limited by the decay rate of the auxiliary mode, which is itself constrained by the need to suppress Purcell decay and, in the case in which the mode belongs to the qubit's readout resonator, to optimize readout fidelity [8,22,25,26].

*These authors contributed equally.

†Current address: Google Quantum AI, Santa Barbara, CA.

‡Current address: Department of Applied Physics and Applied Mathematics, Columbia University, New York, NY 10027.

§Contact author: opainter@caltech.edu

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

Furthermore, additional driving or complicated pulse shaping is required for each additional excited state to be reset due to the narrow-banded nature of the auxiliary modes.

In this work, we go beyond the state of the art by implementing fast, unconditional reset of the first two excited states of a frequency-tunable transmon qubit by coupling it to a broadband metamaterial waveguide (MMWG) strongly damped to a cold environment [27–31]. The MMWG provides stronger protection against Purcell decay than single-pole Purcell filters, thanks to the sharp extinction of its density of states (DOS) outside its passband. This enables significantly faster reset than in previous works by alleviating the trade-off between large achievable damping and idling qubit Purcell relaxation. The bandwidth of the MMWG can be made large relative to the transmon anharmonicity, enabling simultaneous reset of multiple excited states. In addition, band edges with roll-off steeper than the transmon anharmonicity enable implementation of so-called leakage-reduction units (LRUs) [10], which selectively reset the second excited state of the transmon without disturbing the information encoded in the computational subspace. Leveraging these capabilities of the MMWG, in combination with dynamical activation of the interaction via parametric flux modulation [32–34], we demonstrate simultaneous reset of a transmon's three lowest-energy eigenstates and an LRU that selectively resets a transmon's second excited state.

II. RESULTS

We perform our measurements on a frequency-tunable transmon qubit (Q_1) on the same device used in Ref. [30], biased at its upper sweet spot (USS) with a g - e transition frequency $\omega_{ge}/2\pi = 7.63$ GHz and an anharmonicity $\eta/2\pi \equiv (\omega_{ef} - \omega_{ge})/2\pi = -179$ MHz. The qubit has a dedicated microwave drive line for XY control and a flux line for slow and fast frequency control.

In this system, Q_1 is capacitively coupled to a single unit cell of an MMWG comprising a 1D coupled array of 52 lumped-element microwave resonators, as schematically shown in Fig. 1(a). Such a structure can be modeled as a set of harmonic modes with a finite passband with sharp band edges, and this allows rapid photon emission from coupled modes [29,35,36]. At either end of the MMWG is a tapering section consisting of four resonators that impedance-match the MMWG modes to $50\ \Omega$ transmission lines serving as input and outputs to the waveguide, providing strong damping to its environment across a wide and flat passband from approximately 5 to 7 GHz. We measure Q_1 using standard dispersive readout techniques, with the MMWG serving as a Purcell filter waveguide for readout. Because of the need to estimate a small residual excitation population for characterizing qubit reset on the order of the thermal population of the MMWG, achieving sufficiently high sensitivity in single-shot readout is crucial. More

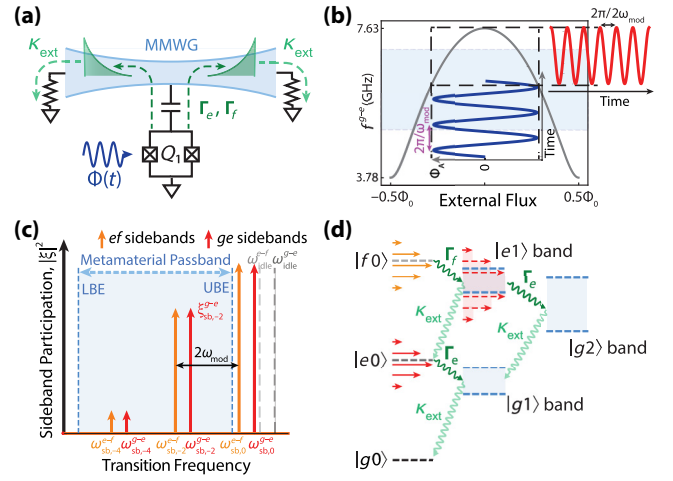


FIG. 1. Schematic of the reset protocol. (a) Illustration of reset of Q_1 . An external flux pulse $\Phi(t)$ applied to the SQUID loop of Q_1 causes bidirectional emission of excitation from Q_1 into the MMWG, which is damped to its external environment. (b) Tuning curve of the flux-tunable transmon qubit. Here, Φ_0 is the magnetic flux quantum. Parking the qubit at the upper sweet spot, where the qubit is first-order insensitive to the external flux, a flux modulation tone of frequency ω_{mod} (dark blue curve) modulates the qubit frequency at $2\omega_{\text{mod}}$ (red curve). The region shaded in pale blue denotes the extent of the MMWG passband. (c) Sideband picture of the Q_1 during reset operation. Gray dashed lines represent idling g - e and e - f transition frequencies of Q_1 , while colored lines indicate sideband frequencies with parametric modulation at ω_{mod} turned on. (d) Energy-level diagram showing the flow of reset for different states. The notation $|sn\rangle$ is shorthand for Q_1 in state $|s\rangle$ and n photons being in the waveguide. The red and gold arrows denote modulation sidebands. Note that states of nonzero MMWG photon number live in a band of closely spaced energy levels, represented by the semiopaque blue bands. The transition frequencies of the states in these bands can also be modulated when they include excitations in the qubit, as illustrated by the semiopaque red rectangles representing the set of sidebands of the $|e1\rangle$ band because they depend on qubit frequency.

details of device parameters, readout, and thermal population of the MMWG can be found in Appendixes B–D.

Our qubit reset protocol is based on dynamically activating an exchange interaction between Q_1 and the MMWG modes. When Q_1 is idling, detuned above the MMWG passband, the rapid extinction of the MMWG DOS outside its band edges strongly protects Q_1 from Purcell decay through the $50\ \Omega$ ports of the MMWG. To turn on an exchange interaction, we apply a single-tone flux modulation pulse to Q_1 's superconducting quantum interference device (SQUID) loop, which activates a resonance condition between Q_1 and the MMWG passband, rapidly converting this Purcell decay protection into Purcell decay enhancement and inducing qubit reset. More information on relaxation times (T_1) as a function of qubit frequency

can be found in Appendix D and Supplementary Material III of Ref. [30].

As illustrated in Figs. 1(b) and 1(c), when Q_1 is biased at its USS, a flux modulation pulse of frequency ω_{mod} and amplitude Φ_m grows sidebands separated by even integer multiples ($\pm n2\omega_{\text{mod}}$) from the time-averaged transition frequencies. Note that the negative convexity of the qubit's tuning curve at its bias point shifts the time-averaged transition frequencies of Q_1 downward.

Sidebands of the transmon that fall within the passband of the MMWG are coupled to an effective continuum of modes and mediate radiative emission into the MMWG, as illustrated in the energy-level diagrams of Figs. 1(c) and 1(d). The rate of emission of the transmon is roughly set by Fermi's golden rule, involving the DOS of the MMWG and the transmon-MMWG coupling matrix element, which is determined by the bare coupling between Q_1 and its unit cell, and the participation of the sidebands. Multiple excited states of the transmon can be simultaneously reset by the same mechanism, thanks to the large bandwidth of the MMWG, as shown in Figs. 1(c) and 1(d), where the emission processes for states up to the second excited state are illustrated.

Emitted photons leave the waveguide at an effective damping rate κ_{ext} determined by the finite reflectivity of the MMWG tapers and a round-trip time of approximately 14 ns, estimated from the group velocity and the number of unit cells in the MMWG [29]. Note that full reset of both the qubit and the waveguide is important to avoid subsequent qubit control errors caused by photons remaining in the waveguide, potentially inducing population revival or dephasing [29,31].

We have chosen parametric modulation to avoid shelving and retrieval of excitation in and out of long-lived resonances whose frequencies collide with Q_1 's frequency during flux tuning [37]. We find multiple such resonances, including two-level system (TLS) defects and MMWG modes near the band edges that are not sufficiently damped to output lines, one of which we measure to have a relaxation time of approximately 1.7 μs . See Appendix E for details of reset with direct flux tuning and Landau-Zener-Stückelberg interference with long-lived resonances.

To calibrate a reset operation for at least the first two excited states of Q_1 , we first define a flux pulse $\Phi(t; \Phi_A, \omega_{\text{mod}}, \tau_{\text{pulse}}) = \Phi_A \mathcal{E}(t; \tau_{\text{pulse}}) \sin(\omega_{\text{mod}} t)$, parameterized with Φ_A (modulation amplitude), ω_{mod} (modulation frequency), and τ_{pulse} (total flux pulse duration). The envelope function $\mathcal{E}(t; \tau_{\text{pulse}})$ is a square pulse of duration τ_{pulse} with zero-amplitude buffers $\tau_B = 2$ ns at the start and end, filtered with a Gaussian kernel of width $\sigma = 1$ ns [12]. We perform a 3D sweep of the parameters Φ_A , ω_{mod} , and τ_{pulse} to characterize the reset at each point.

The pulse sequence for characterizing a reset operation is provided in Fig. 2(a). We begin by preresetting Q_1 with a precalibrated reset pulse and heralding Q_1 in its

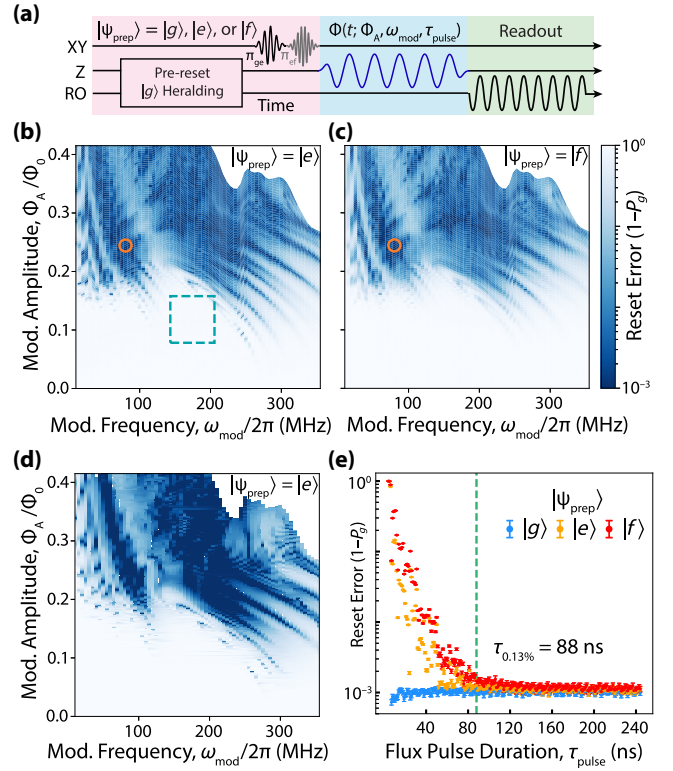


FIG. 2. Calibration and characterization of unconditional reset. (a) Pulse sequence for characterizing reset operation. (b),(c) Reset errors obtained from scan of reset pulse parameters ω_{mod} and Φ_A , for Q_1 prepared in (b) $|e\rangle$ and (c) $|f\rangle$. The dashed teal box indicates the scanned region for Fig. 3(b). (d) Simulation of the measurement of (b) up to single-excitation subspace. (e) Reset errors obtained by sweeping τ_{pulse} with 1-ns resolution, for $\Phi_A = 0.244\Phi_0$ and $\omega_{\text{mod}} = 80$ MHz [orange hollow circles in (b) and (c)]. Error bars indicate 95% confidence interval (t -test) with 30 samples, where each sample is obtained from 50 000 single-shot measurements.

ground state with high probability using repeated rounds of projective measurement with a strict discrimination boundary [21]. Our ground-state preparation has measured infidelity $1 - P_{g,\text{HERALD}} \approx 0.02\%$, which is significantly smaller than the thermal population of approximately 0.1% in the MMWG. We then apply a combination of π_{ge} and π_{ef} pulses to prepare Q_1 in its $|g\rangle$, $|e\rangle$, or $|f\rangle$ state, after which we apply the appropriate reset pulse, followed by readout of the qubit. The measured state populations are inverted using the calibrated confusion matrix. We define the reset error (infidelity) as $1 - P_g$, where P_g is the measured probability of the qubit being in the ground state after reset.

Figures 2(b) and 2(c) show the reset errors obtained from scanning the frequency and amplitude of a reset pulse of fixed length $\tau_{\text{pulse}} = 104$ ns for prepared states $|e\rangle$ and $|f\rangle$. Thanks to the broadband nature of the MMWG, there are many frequency-amplitude pairs that

could parameterize viable reset pulses with reset errors $< 1\%$ for both prepared states. The results from simulating the experiment of Fig. 2(b) agree very well with experimental data, replicating most of the significant qualitative features of the data, as shown in Fig. 2(d) (see Appendix F for further details). Sets of curved features found in Figs. 2(b)–2(d) show where sidebands are dominantly interacting with the longer-lived band-edge modes of the MMWG passband. In addition, the interference patterns found at the upper-leftmost regions ($\omega_{\text{mod}}/2\pi < 40$ MHz and $\Phi_A > 0.17\Phi_0$) of Figs. 2(b)–2(d) show evidence of coherent dynamics due to the relatively slow tuning of the qubit at lower modulation frequencies through the more coherent band-edge modes [37].

Figure 2(e) shows the reset quality at a fixed frequency and amplitude as τ_{pulse} is varied. For a given frequency and amplitude, we define the time τ_{ϵ}^i for reset of error ϵ such that for $\tau_{\text{pulse}} > \tau_{\epsilon}^i$, $1 - P_g < \epsilon$, for initial state $|i\rangle \in \{|g\rangle, |e\rangle, |f\rangle, \dots\}$. We consider this definition to take into account the time required for the depletion of photons in the MMWG after emission, which can re-excite the qubit upon spurious reflection from the MMWG taper. Figure 2(e) shows a frequency-amplitude pair that gives one of the fastest resets, as characterized by $\tau_{1\%}^e = 44$ ns, $\tau_{0.3\%}^e = 51$ ns, and $\tau_{0.13\%}^e = 88$ ns when prepared in $|e\rangle$, with a steady-state population $P_e^{\text{s.s.}} = 0.11\%$ (2) found from a fit. This residual excitation population is close to the qubit's excited-state population when thermalized to the MMWG with its g - e transition frequency near the center of the passband. In addition, it is clearly demonstrated that the reset operation achieves simultaneous reset of multiple excited states, characterized by $\tau_{1\%}^f = 51$ ns, $\tau_{0.3\%}^f = 70$ ns, and $\tau_{0.16\%}^f = 88$ ns.

Similarly, rapid reset of multiple excited states with a small residual excited-state population of approximately 0.2% within approximately 97 ns is also achieved when Q_1 is biased at its lower sweet spot (LSS) with a g - e transition frequency $\omega_{ge}^{\text{LSS}}/2\pi \approx 3.78$ GHz. This demonstrates flexibility in choosing the idling frequency relative to the strongly damped auxiliary modes used for unconditional reset, while maintaining both speed and the ability to reset multiple excited states simultaneously. For the results of unconditional reset from the LSS, refer to Appendix F.

The sharp roll-off of the DOS at the upper band edge of the MMWG relative to the transmon anharmonicity enables implementation of LRUs by selectively activating dissipation of $|f\rangle$ and higher excited states into the MMWG while maintaining the coherence of the g - e subspace. Figure 3(a) shows the desired sideband configuration for implementing an LRU. An e - f sideband is placed within the passband of the MMWG, inducing radiative emission from $|f\rangle$ to $|e\rangle$, while the corresponding g - e sideband, which is greater in frequency by the transmon anharmonicity, remains outside the passband.

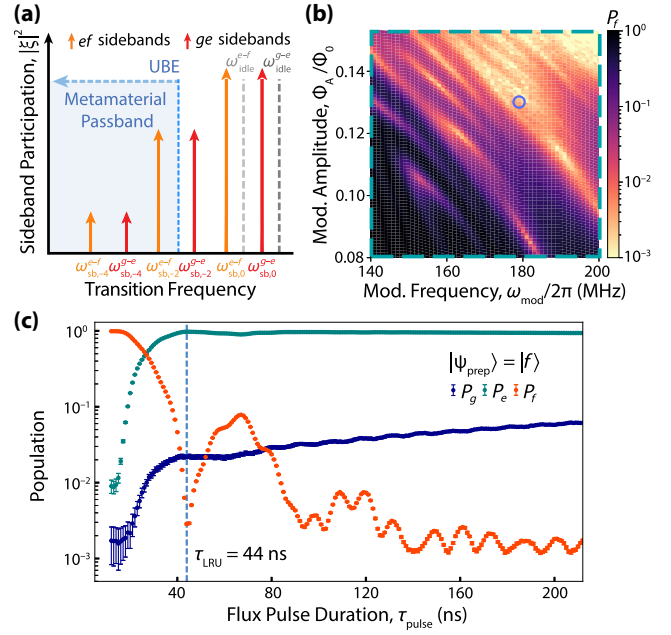


FIG. 3. Calibration and benchmarking of LRU operation. (a) Sideband picture of the parametric modulation of Q_1 during LRU operation. (b) The $|f\rangle$ population after applying the LRU operation implemented by parametric flux modulation, with two-dimensional scan of pulse parameters over the regions indicated in Figs. 2(b) and 2(c) (dashed teal box). (c) Resulting $|g\rangle$, $|e\rangle$, and $|f\rangle$ populations obtained after applying the LRU operation to a qubit initially prepared in $|f\rangle$, where τ_{pulse} is swept with 1-ns resolution, for $\Phi_A = 0.13\Phi_0$ and $\omega_{\text{mod}}/2\pi = 179$ MHz [blue hollow circle in (b)]. Error bars indicate 95% confidence interval (t -test) with 30 samples, where each sample is obtained from 100 000 single-shot measurements.

It is desirable to minimize the time taken for the LRU so as to minimize the decoherence suffered by the g - e subspace due to relaxation and dephasing processes during leakage reset. This generally requires increasing the coupling of the e - f transition to the MMWG modes, which is accomplished in our parametric reset scheme by increasing the modulation amplitude for a given modulation frequency; however, depending on the amplitude of modulation, higher-order sidebands of the g - e transition lying inside the MMWG passband may have non-negligible participation, such as the fourth-order sideband of the g - e transition, as illustrated in Fig. 3(a), and will limit the fidelity of the LRU.

To determine viable parameter candidates for the LRU, we perform a more granular frequency-amplitude sweep with total flux pulse duration $\tau_{\text{pulse}} = 132$ ns. The explored parameter region is that indicated by the dashed teal box in Fig. 2(b), where we observe insignificant $|e\rangle$ decay and expect the second-order e - f sideband to just be entering the passband of the MMWG. In this measurement, we chose a smoother Gaussian filter parameter $\sigma = 3$ ns and

buffer duration $\tau_B = 6$ ns to suppress spectral broadening of the second-order sideband of the g - e transition, which can induce unwanted additional radiative decay of the $|e\rangle$ state [12]. The pulse sequence is identical to that of Fig. 2(a). As can be seen in Fig. 3(b), multiple prominent regions of $|f\rangle$ relaxation are visible, which we believe correspond to interaction of the e - f transition with different modes near the upper band edge of the MMWG.

The set of optimal LRU parameters is found by trying to both maximize the $|f\rangle$ removal fraction and minimize the impact on remaining $|e\rangle$ population following LRU application (see Appendix G for further details regarding the calibration of the LRU). For optimal modulation frequency $\omega_{\text{mod}}/2\pi = 179$ MHz and modulation amplitude $\Phi_A = 0.13\Phi_0$ [blue circle in Fig. 3(b)], we plot in Fig. 3(c) the measured final-state g , e , and f probabilities as a function of LRU pulse length τ_{pulse} for an initial state of $|f\rangle$. We find the fastest LRU operation of 44 ns with a residual $|f\rangle$ population of 0.285(3)%. Note that we define the time needed for the LRU as the time to the first local minimum of the $|f\rangle$ population—in contrast to the definition of the unconditional reset time, which includes the time required for the depletion of MMWG—for the following reasons. First, leakage errors are expected to occur infrequently [9]; hence, the vast majority of the time, we expect no emitted photons after the LRU operation. Any time spent waiting for waveguide ringdown in those cases serves no purpose and only incurs additional decoherence on spectator qubits. Second, in the event that leakage has occurred and the LRU has successfully returned the leaked population back to the computational subspace, a detected error in the next syndrome measurement is essentially inevitable. In this case, waiting for waveguide depletion to avoid additional computational errors is less valuable.

With the calibrated LRU, we study its efficacy in removing leakage errors over long sequences and its decoherence impacts on the g - e subspace by using randomized benchmarking (RB). For this analysis, we consider four different RB primitive sequences: reference, interleaved LRU (“LRU”), inject 3% leakage after each Clifford gate (“leakage injection”), and leakage injection followed by LRU (“leakage injection + LRU”), shown schematically in Fig. 4(a) [12,40]. Figures 4(b) and 4(c) compare the measured ground-state survival population and leakage accumulation, respectively, for each of these RB sequences. The RB curves are fitted to the leakage-accumulation and survival-population models provided in Ref. [38], which allows us to extract the steady-state leakage population of the RB sequences. When the leakage error rate per Clifford operation is significantly higher than the relaxation rates of the leakage states, the survival population P_g is expected to drop below 50% due to the buildup of a higher steady-state leakage population, as clearly observed in the case of “leakage injection,” where P_f exceeds 20%. In

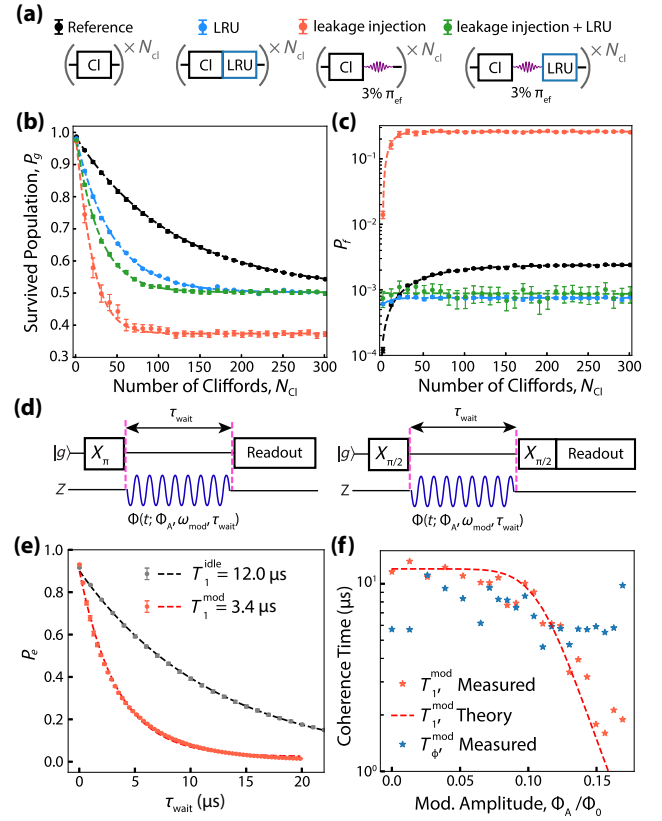


FIG. 4. Randomized benchmarking (RB) with LRU operation. (a) Primitive pulse sequences used for RB to assess the infidelity of the LRU as an idling gate (labeled “LRU”) and the efficacy of the LRU against realistic leakage errors (labeled “leakage injection” and “leakage injection + LRU”). Each sequence in parentheses is repeated N_{Cl} times to run an RB experiment. A π_{ef} pulse with a reduced amplitude is used to artificially inject a leakage error of 3%. Every LRU operation is followed by a pre-calibrated virtual-Z operation that cancels the phase accumulated during the LRU. (b) Survival population (P_g) and (c) accumulated leakage (P_f) after applying RB sequences with different numbers of Clifford gates (N_{Cl}). Dashed lines indicate fits to the model from Ref. [38]. Error bars indicate 95% confidence intervals (t -test) across 95 RB traces obtained over approximately 25 h. Each RB experiment was performed with 100 random Clifford sequences and averaged over 1000 repetitions. (d) Pulse sequences used for measuring the relaxation coherence time, T_1^{mod} (left), and the Ramsey coherence time, T_2^{mod} (right), under continuous flux modulation during the waiting times τ_{wait} . (e) Comparison between relaxation coherence time without flux modulation (T_1^{idle}) and with flux modulation, using the parameters $\omega_{\text{mod}}/2\pi = 179$ MHz and $\Phi_A = 0.13\Phi_0$, as used in the LRU. Dashed lines represent fits to an exponential decay profile. (f) Plot of T_1^{mod} and pure dephasing coherence time T_ϕ^{mod} measured at different Φ_A , with $\omega_{\text{mod}}/2\pi = 179$ MHz. Here, T_ϕ^{mod} is obtained from its relationship to the Ramsey coherence time $(T_2^{\text{mod}})^{-1} = (2T_1)^{-1} + T_\phi^{-1}$ [39]. The red dashed curve represents the calculated T_1^{mod} based on the sideband spectrum of the g - e transition during modulation and a simplified decay-rate model within the MMWG passband provided in Appendix G.

contrast, we find that the use of the LRU effectively limits the accumulation of leakage significantly below that of the reference RB, resulting in a steady-state leakage population of approximately 0.08%, which is close to the readout floor. Furthermore, the use of the LRU maintains a small steady-state leakage population of approximately 0.09%, even when preceded by the injection of 3% leakage. These results clearly demonstrate that the LRU is sufficiently effective in suppressing typical leakage-error rates for single- and two-qubit gates in transmon qubits.

Noting the sufficiently small leakage accumulation in both the “Reference” and “LRU” cases, we conduct a conventional interleaved RB (iRB) analysis [41], revealing a gate infidelity of 0.72(1)%. To better understand the sources of this infidelity, we quantify the impact of flux modulation on the g - e subspace by measuring coherence with the pulse sequence shown in Fig. 4(d). We observe a reduced relaxation coherence time $T_1^{\text{mod}} = 3.3 \mu\text{s}$ during the flux modulation with the same parameters as used in the LRU operation, as shown in Fig. 4(e). In addition, we observe a clear reduction in T_1^{mod} as the modulation amplitude is increased at a fixed modulation frequency, as illustrated in Fig. 4(f). We attribute this reduction in relaxation coherence time to the activation of higher-order g - e sidebands within the MMWG passband, which serves as an unintended channel for photon emission into the MMWG. This interpretation is supported by the strong agreement with the numerically estimated T_1^{mod} (red dashed curve) obtained from numerical calculation of sideband strengths and assuming constant emission rates within the MMWG passband. We also perform a Ramsey experiment with flux modulation activated to measure the decoherence time during the LRU operation and use this along with T_1^{mod} to estimate the pure dephasing rate during modulation, T_ϕ^{mod} . We find $T_\phi^{\text{mod}} = 3.7 \mu\text{s}$, roughly uncorrelated with the modulation amplitude, as shown in Fig. 4(f).

From these independent measurements, we estimate that such decay through unwanted sidebands contributes an infidelity of approximately 0.23% for the pulse parameters used for the LRU. Including the contribution from background (idling) decoherence rates, we estimate an infidelity of 0.72(1)% for the LRU restricted to the g - e subspace, and this is in good agreement with the measured infidelity. Further improvements in infidelity are expected for qubits with longer idling coherence times, since they allow the use of a weaker modulation amplitude and longer pulse duration for the LRU, suppressing decay through unwanted g - e sidebands. For further discussion and analysis of LRU fidelity, refer to Appendix G.

III. SUMMARY AND OUTLOOK

We have demonstrated simultaneous unconditional reset of a transmon with a residual excitation population below 0.13% when prepared in $|e\rangle$ and below 0.16% when

prepared in $|f\rangle$, in 88 ns. We additionally implemented an LRU operation that resets a transmon’s second excited state to 0.285(3)% residual excitation in 44 ns and has an overlap infidelity of 0.72(1)% with the identity operation in the computational subspace.

Although the reset and LRU operations presented here are remarkably fast, additional avenues for improvement are clear. The qubit-waveguide coupling could be increased further, as the qubit relaxation times are not yet Purcell limited ($T_1^{\text{Purcell-limit}} > 1 \text{ s}$). Narrowing the passband of the waveguide would increase the emission rate by increasing the photonic DOS in the passband. Such a narrowing of the passband would also help suppress spurious reflections at the input and output ports by improving the achievable impedance matching and provide further Purcell protection during the LRU operation by steepening the roll-off of the MMWG band edges. Further reduction in the reset time could be achieved by decreasing the number of unit cells in the MMWG. This would allow the MMWG to ring down more quickly by reducing the round-trip time, compensating for the increased group delay from a narrower passband. State-of-the-art fabrication techniques and loss-optimized MMWG design choices can suppress relaxation at idle qubit frequency due to internal loss in the MMWG resonators, without compromising the reset rate (see Appendix D for more details). Regarding leakage reduction, the use of an MMWG with a narrower passband could also prevent emission via higher-order g - e sidebands through the MMWG passband [31], providing better protection of the g - e subspace and yielding lower infidelity for the LRU. Further discussion and analysis of LRU performance is presented in Appendix G.

The reported reset error level is understood to be limited by the thermal population of the MMWG. We believe that the temperature of the MMWG (approximately 45 mK) was elevated above the mixing-plate temperature in our system due to an incidental overlap of one of our traveling-wave parametric amplifier (TWPA) pumps with the MMWG passband, and insufficient isolation of our device from the pump tone. Ameliorating this systematic issue, as well as improving the MMWG tapering to its environment, could help better reduce the residual excitation population after reset (see Appendix D for further details).

While the unconditional nature of the reset obviates the need for measurement of the transmon state, information—partial or otherwise—about the state of the transmon prior to reset can still be extracted by collecting the outgoing photonic field from any damped ends of the MMWG [31,42,43]. This could, in the most ideal case, allow for conversion of leakage errors into erasure errors by detecting when leakage events occur. It is also worth noting that the demonstrated unconditional reset and LRU, in principle, simultaneously reset a few higher excited states including the third ($|h\rangle$) and fourth excited states of the

transmon, as their transition sidebands are expected to reside within the broad passband of the MMWG. Although this has not been characterized in this work due to insufficient coherence in these states caused by their spectral proximity to the MMWG passband (bare ω_{fh} is expected to be approximately 7.12 GHz, upper band edge approximately 7.1 GHz), a device in which the MMWG possesses a lower-frequency passband would allow the reset of such states to be characterized.

Finally, the reset and leakage-reduction method presented here can be extended to multiqubit systems using a single MMWG, with a few technical considerations. To prevent unwanted excitation swapping between qubits instead of efficient resetting, modulation parameters should be carefully chosen to ensure that qubit sidebands do not overlap. To extend leakage reduction to multiqubit systems, it is important to address interactions mediated by the MMWG due to overlap between qubit-photon bound states [30,35,44]. One possible approach is to leverage these interactions to implement parametric multiqubit gates or continuous multiqubit interactions with built-in leakage reduction. This can be achieved by selecting modulation parameters that enable controlled interaction within the g - e subspace while placing higher-order sidebands of transitions involving leakage states inside the MMWG passband. Alternatively, such interactions can be minimized by increasing the detuning from the MMWG band edge, which reduces localization lengths. In addition, since bound states exhibit an exponentially decaying wave function with distance, selectively spacing qubits by skipping a few resonators can further reduce unintended coupling.

ACKNOWLEDGMENTS

We thank Mo Chen, Lucia De Rose, Piero Chiappina, Steven Wood, Srujan Meesala, David Lake, Jamison Stevens, and Matt Davidson for helpful discussions. We appreciate MIT Lincoln Laboratories for providing the traveling-wave parametric amplifier used in the microwave readout chain in our experimental setup. This work was supported through a grant from the Department of Energy (Grant No. DE-SC0020152), the Institute for Quantum Information and Matter, an NSF Physics Frontiers Center (Grant No. PHY-2317110), and through a sponsored research agreement with Amazon Web Services. E.K. was supported by the New Faculty Startup Fund from Seoul National University, the POSCO Science Fellowship of POSCO TJ Park Foundation, the National Research Foundation of Korea (NRF) grant funded by the Ministry of Science and ICT (MSIT) (Grant No. RS-2024-00334169), and the LAMP Program of the National Research Foundation of Korea (NRF), funded by the Ministry of Education (Grant No. RS-2023-00301976).

DATA AVAILABILITY

The data are not publicly available. The data are available from the authors upon reasonable request.

APPENDIX A: MEASUREMENT SETUP

Our fridge wiring is identical to that of Fig. S6 of the Supplementary Material in Ref. [30], and a wiring diagram of our external control electronics can be seen in Fig. S1. It is similar to Fig. S5 of the Supplementary Material in Ref. [30], except we have removed the 10 dB of attenuation on the fast flux lines of subfigure (a) to increase the dynamic range of our reset pulses, and the up-conversion of our microwave XY signals is now handled by two Quantum Machines Octave units, obviating the need for some of the in-house filters, mixers, attenuators, amplifiers, and fast RF switches on those XY lines. Note that the readout up-conversion and down-conversion chain continue to use our previous in-house mixing configuration.

APPENDIX B: DEVICE CHARACTERIZATION

The circuit model of our device was identical to that shown in Supplementary Material Sec. I of Ref. [30], including fit values for the inductances and capacitances, except that we also included the readout resonators of each qubit as lumped-element LC resonators with center frequencies ω_{ri} obtained using a fit to spectroscopy data from the device, inductances $L_r = 4.5$ nH, and capacitances $C_{ri} = 1/(\omega_{ri}^2 L_r)$. The readout resonators are capacitively coupled to each qubit with capacitance C_{qr} and the unit cell of the MMWG with capacitance C_{wr} . We did not include parasitics for the readout resonators.

APPENDIX C: QUBIT READOUT CALIBRATION AND CHARACTERIZATION

We perform three-level state discrimination of Q_1 with the standard transmon dispersive readout technique [26,45]. Our input and output signal paths and amplification chain for readout are identical to those of Figs. S5–S6 in the Supplementary Material of Ref. [30].

In this work, we send a readout input microwave pulse at around the frequency of R_1 (the readout resonator of Q_1), $\omega_{\text{drive}} \approx \omega_{R1}$, with a 248-ns-long square pulse envelope. We start the square envelope with a 20-ns kick at twice the amplitude for fast charging of the readout resonator [46] and convolve the envelope with a Gaussian kernel of 10-ns standard deviation.

We measure temporal down-converted signals of both the microwave signals reflected (superscripted R) and transmitted (superscripted T) from the readout resonator and the MMWG Purcell filter using the OPX+, improving the signal-to-noise ratio by approximately 3 dB. For each single-shot measurement (subscripted k), we collect the signals for 348 ns, which is 100 ns longer than the

input readout pulse, accounting for the transient response of R_1 . The measured reflected and transmitted signals are then each mapped to pairs of real numbers (I_k^R, Q_k^R) and (I_k^T, Q_k^T) by integrating them with precalibrated sets of integration weights, as defined in the Supplementary Material of Ref. [5].

The four real numbers $u_{kj} \equiv (I_k^R, Q_k^R, I_k^T, Q_k^T)_j$ characterize the k th single-shot readout of Q_1 prepared in state $|j\rangle$, with $j \in \{g, e, f\}$. Here, $\{u_{kj}\}$ form three Gaussian clusters in a four-dimensional real space. Linear discrimination boundaries for real-time state discrimination are obtained by performing linear discriminant analysis (LDA) [47,48]. We first collect a total of 3×10^6 single-shot readout data (10^6 for each prepared state). Instead of directly feeding the measurement data into LDA, we cluster the data using the Gaussian-mixture clustering method, an unsupervised clustering technique assuming mixtures of Gaussian distributions. During this process, we use the GAUSSIANMIXTURE API from the PYTHON package SCIKIT-LEARN [21,48], identifying the corresponding state by selecting the state with the largest assignment probability. Next, we generate an artificially created ideal dataset $\{u_{kj, \text{ideal}}\}$, numerically drawn from the individual Gaussian distributions of each state. Finally, this ideal dataset is fed into the LDA clustering algorithm using the LINEARDISCRIMINANTANALYSIS API from SCIKIT-LEARN, yielding linear discrimination boundaries. Using such artificially generated ideal samples instead of measured data makes the process of determining discrimination boundaries significantly more resilient against various error sources including state-preparation errors, T_1 relaxation, and measurement-induced state transitions.

We calibrate and characterize the readout using the readout confusion matrix $M \equiv \{M_{ij}\}$, where M_{ij} is the probability of assigning state $|i\rangle$ when Q_1 is prepared in state $|j\rangle$, giving the following relationship:

$$\mathbf{P}^A = M\mathbf{P}, \quad (\text{C1})$$

where $\mathbf{P} = (P_g, P_e, P_f)^T$ is the column vector of the transmon state population and $\mathbf{P}^A = (P_g^A, P_e^A, P_f^A)^T$ is the column vector of the assigned state population. The measured populations present in this work, unless otherwise noted, are obtained by multiplying the inverse of the confusion matrix to account for the measurement errors:

$$\mathbf{P} = M^{-1}\mathbf{P}^A. \quad (\text{C2})$$

We aim to perform readout that is capable of accurately estimating residual excited-state populations while maintaining sufficiently large three-state assignment fidelity $F_{\text{RO}} \equiv (M_{11} + M_{22} + M_{33})/3$. We first run a two-dimensional scan over readout pulse frequency and readout pulse amplitude and optimize for F_{RO} . Then, we increase the readout pulse amplitude until the probability of assigning $|e\rangle$ or $|f\rangle$ when a sample is drawn from the ideal

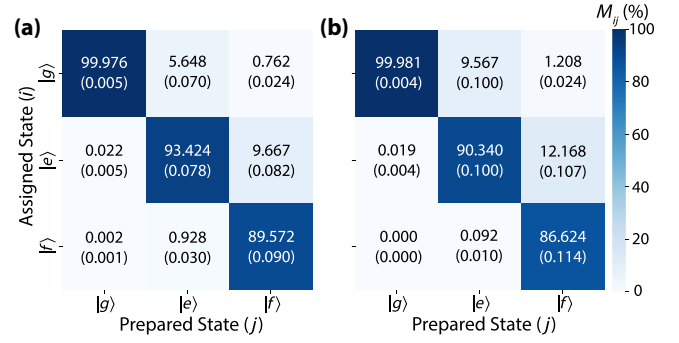


FIG. 5. Estimated confusion matrix of single-shot readout. Experimentally characterized readout confusion matrix M_{ij} used for estimating Q_1 's population after applying the (a) reset and (b) LRU operation including the three lowest-energy eigenstates of Q_1 . The mean and standard error (in parentheses) of each matrix element are obtained over 100 repetitions, with each repetition consisting of 100 000 single-shot readouts.

distribution of $|g\rangle$ is sufficiently small, approximately 10^{-4} .

We start every measurement by heralding the $|g\rangle$ state using a stricter discrimination boundary with a target probability $P_{g, \text{HERALD}}^{\text{target}} = 0.9999$. This minimizes the effect of state-preparation errors, enabling more accurate characterization of M . Note that negative populations often found in similar works may appear if contributions from state-preparation errors are not sufficiently excluded.

Figure 5(a) shows the readout confusion matrices used for characterizing unconditional reset operations. We achieve a readout floor $1 - M_{gg} \approx 2 \times 10^{-4} \approx 1 - P_{g, \text{HERALD}}$, which is sufficiently small for estimating typical residual excited-state populations of approximately 0.1%, while maintaining a high three-state readout fidelity of $F_{\text{RO}} \approx 94.3\%$. Figure 5(b) shows the readout confusion matrices used for characterizing LRU operations. For this readout, we use a different pulse with higher amplitude to better estimate the residual population of $|f\rangle$ states as characterized by $M_{fe} \approx 0.09\%$, while compromising F_{RO} down to approximately 92.3% and maintaining a similar readout floor of approximately 2×10^{-4} .

APPENDIX D: CHARACTERIZATION OF METAMATERIAL WAVEGUIDE AS A COLD BATH

1. Thermal population of the metamaterial waveguide

The thermal population of the MMWG is responsible for the residual excited-state population observed after applying the unconditional reset operations. We estimate the thermal population of the MMWG by thermalizing Q_1 to the MMWG for a sufficiently long time and reading out the excited-state population of Q_1 . Figures 6(a) and 6(b) show the pulse sequence used for this process. We first prepare Q_1 in its $|g\rangle$ or $|e\rangle$ state using an appropriate combination

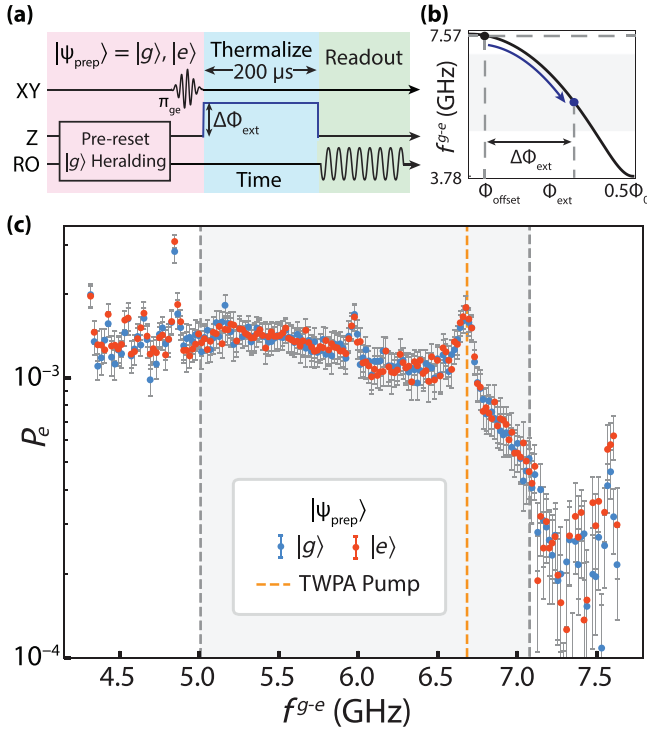


FIG. 6. Characterization of the thermal-excitation population at various g - e transition frequencies of Q_1 . (a),(b) Pulse sequence for characterizing the thermal population. At each measurement, Q_1 is thermalized to its environment while a 200- μ s-long square flux pulse is applied, which brings g - e transition frequencies f^{g-e} to various points. (c) Residual Q_1 excited-state population P_e after thermalization for prepared states $|g\rangle$ (blue) and $|e\rangle$ (red). The range over which f^{g-e} is swept includes the MMWG passband (gray shaded region) and USS frequency (rightmost point). The orange dashed line indicates the frequency of one the TWPA pumps (6.684 GHz) used for readout.

of heralding into $|g\rangle$ and a π_{ge} . Then, the g - e transition frequency of Q_1 is tuned from 7.57 GHz to another frequency f^{g-e} by sending a 200- μ s-long square flux pulse. Finally, we record Q_1 's residual excited-state population via dispersive readout, as explained in Appendix C.

The measured residual excited-state population P_e over different f^{g-e} at which Q_1 idles is shown in Fig. 6. The residual populations are almost the same for the two prepared states over most of the range, suggesting that Q_1 reached a steady state at each point, and the residual excited-state population can be understood as the thermal population of the MMWG. The thermal population of approximately 0.1% at 6 GHz corresponds to an effective temperature $T_{\text{eff}} \approx 43$ mK, which is consistent with values reported in various works. There is an unexpected peak at 6.684 GHz due to insufficient isolation of one of the TWPA pumps driven at 6.684 GHz. This suggests that the thermal populations could be further reduced by better isolating signals and noise flowing from the TWPA to the output line (note that we use a 40-dB isolator) or by

avoiding overlap between the MMWG passband and the frequencies of the TWPA pumps.

2. Relaxation of qubits induced by metamaterial waveguide

To achieve effective reset and leakage reduction while maintaining the coherence of the qubit when idle, it is necessary to ensure a high emission rate at the sideband frequencies and a sufficiently long relaxation rate (T_1) at the idle qubit frequency. This requires a sharp roll-off in the relaxation rate near the band edge of the MMWG. Here, T_1 can be modeled as T_1^{model} in

$$1/T_1^{\text{model}} = 1/T_1^{\text{background}} + 1/T_1^{\text{Purcell}} + 1/T_1^{\text{internal}}, \quad (\text{D1})$$

where the total relaxation rate is decomposed into the background relaxation rate $1/T_1^{\text{background}}$, which is independent of the MMWG, the Purcell-limited relaxation rate $1/T_1^{\text{Purcell}}$, and the internal-loss-induced relaxation rate $1/T_1^{\text{internal}}$ from the MMWG resonators.

A plot of T_1^{Purcell} is shown in Fig. 7. This is numerically calculated based on the admittance seen from the qubit node toward the MMWG output ports. Due to the multipole nature of the MMWG, T_1^{Purcell} changes rapidly near the band edge. For instance, T_1^{Purcell} decreases from over 1 ms at 7.2 GHz to below 10 ns at 7.0 GHz. For further details on the estimation of T_1^{Purcell} , refer to Ref. [30].

Another key factor determining the sharpness of the qubit relaxation rate is material loss in the lumped-element resonators in the MMWG. Due to strong qubit-resonator

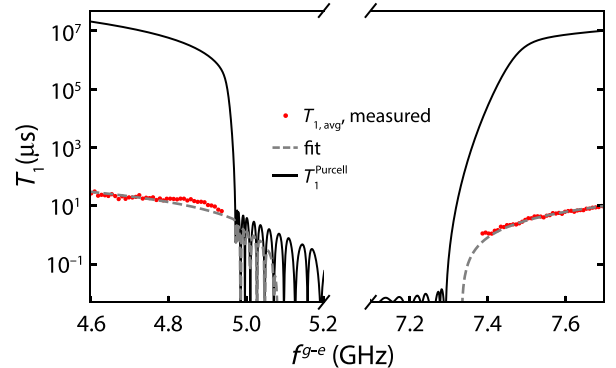


FIG. 7. Plot of T_1 as a function of qubit frequency and a comparison with the model. Average relaxation time (T_1) of Q_1 - Q_{10} as a function of operating frequency f^{g-e} . The T_1 values were measured in a separate cooldown. Data points are selected for frequency ranges where at least three qubits were measured. For better visualization, the frequency range [5.2, 7.1] GHz, where no measured data points are available, has been omitted from the plot, as indicated by diagonal break lines on the x axis. The gray dashed line represents a fit to the modeled relaxation time T_1^{model} in Eq. (D1). The black solid line indicates the Purcell-limited relaxation time T_1^{Purcell} , as estimated in Ref. [30].

coupling and resulting hybridization, the qubit-like eigenstates inherit losses from the resonators that are not suppressed by the bandpass filter response of the MMWG. The $1/T_1^{\text{internal}}$ value of a qubit coupled to one of the MMWG resonators can be approximated as follows:

$$1/T_1^{\text{internal}} = \sum_k p_k \kappa_k^i, \quad (D2)$$

$$p_k \approx \frac{g_k^2}{(\omega_q - \omega_k)^2},$$

where $k = 2\pi m/N$ (for integer m satisfying $k \in [-\pi, \pi]$) is the discrete wave vector of an MMWG mode with frequency ω_k , internal decay rate κ_k , and coupling g_k to the qubit. Here, p_k represents the participation of mode k in the hybridized qubit-like eigenstate.

The dispersion relation of the MMWG and the coupling between the qubit and the MMWG modes can be approximated as follows [29,35]:

$$\omega_k = \omega_c - 2J \cos(k), \quad (D3)$$

$$g_k \approx \frac{g}{\sqrt{N}},$$

where ω_c is the center frequency of the MMWG, J is the hopping rate between adjacent resonators, g is the coupling strength, and we assume a sufficiently long MMWG with N resonators and periodic boundary conditions.

In the continuum limit $N \rightarrow \infty$, where $\Delta k = 2\pi/N \rightarrow dk$, assuming all MMWG modes have the same internal quality factor Q_i , T_1^{internal} simplifies to

$$\begin{aligned} 1/T_1^{\text{internal}} &= \frac{g^2}{Q_i} \sum_k \frac{1}{N} \frac{\omega_c - 2J \cos(k)}{(\omega_q - \omega_c + 2J \cos(k))^2} \\ &\approx \frac{g^2}{2\pi Q_i} \int_{-\pi}^{\pi} \frac{\omega_c - 2J \cos(k)}{(\omega_q - \omega_c + 2J \cos(k))^2} dk \\ &= \frac{g^2}{Q_i} \frac{(\omega_q - \omega_c)\omega_c + 4J^2}{(\omega_q - \omega_c)^2 - 4J^2}^{3/2}, \end{aligned}$$

where we use the relationship $\kappa_k = \omega_k/Q_i$.

By incorporating this analytical expression into Eq. (D1), we compare T_1^{model} with the measured T_1 values, as shown in Fig. 7. To mitigate site-specific variations, such as resonances with TLS defects that could obscure the comparison, we use the average relaxation rate $T_{1, \text{avg}}$ of Q_1 - Q_{10} . A fit to the data, which provides good agreement, as shown in Fig. 7, suggests an estimated Q_i of 2.3×10^4 for the lumped-element resonators in the MMWG.

This analysis suggests that decreasing qubit-MMWG coupling enhances T_1^{internal} by reducing overall hybridization. Nonetheless, the reset rate, which is primarily determined by the emission rate $\Gamma_{\text{ID}} = g^2/J$ at the center of

the passband, can be preserved by reducing J proportionally to g^2 . This implies that a device with a narrower MMWG passband and reduced qubit-MMWG coupling could achieve longer T_1^{internal} without compromising the reset speed. For example, reducing the MMWG bandwidth from 2 GHz to 500 MHz while halving the coupling can improve T_1^{internal} by a factor of 4 without degrading reset performance.

Furthermore, significant improvements in quality factors are possible with recent advances in material engineering and by choosing designs optimized for loss. State-of-the-art superconducting microwave resonators have demonstrated quality factors exceeding 10^6 for both lumped-element resonators and coplanar waveguide resonators [49–51]. Incorporating such high-quality resonators into the MMWG could further enhance T_1^{internal} while maintaining efficient reset and leakage reduction.

APPENDIX E: RESET WITH DIRECT FLUX TUNING AND LONG-LIVED BAND-EDGE RESONANCES

Theoretically, direct tuning of the transmon's transition frequencies within the MMWG provides an even larger emission rate $\Gamma_{\text{ID}} = \lambda_i^2 g_{\text{uc}}^2/J$ in the Markovian limit at the center of the passband [35,36], where λ_i is the relevant transition-matrix element for the i th state in the transmon annihilation operator. Consequently, it is expected that Q_1 would achieve a reset error of 0.1% within 20 ns when prepared in $|e\rangle$, significantly faster than the reset demonstrated in the main text; however, as noted in the main text, we have chosen parametric flux modulation to avoid adiabatic shelving of excitation into other modes coupled to the qubit, which sets a lower bound on achievable reset error and introduces non-Markovianity into the environment [52].

Figures 8(a)–8(c) show the quality of the reset by direct flux tuning. Despite the rapid initial decay of excited states, the reset error stops decreasing and reveals coherent dynamics dominating at longer pulse durations, over a wide range of g - e transition frequencies during the flux pulse, including the entire MMWG passband.

Such phenomena can be understood by Landau-Zener-Stückelberg (LZS) interference with two passages: a fraction of the excitation is transferred to long-lived resonances coupled to Q_1 during the frequency collision at the rising edge of the flux pulse, and it is retrieved by the same mechanism at the falling edge of the flux pulse [37].

To test this hypothesis, we first probe the presence of a strongly coupled long-lived resonance by activating a sufficiently weak sideband of the g - e transition of Q_1 around the upper band edge of the MMWG. We indeed observe damped Rabi oscillation when there is a weak sideband at

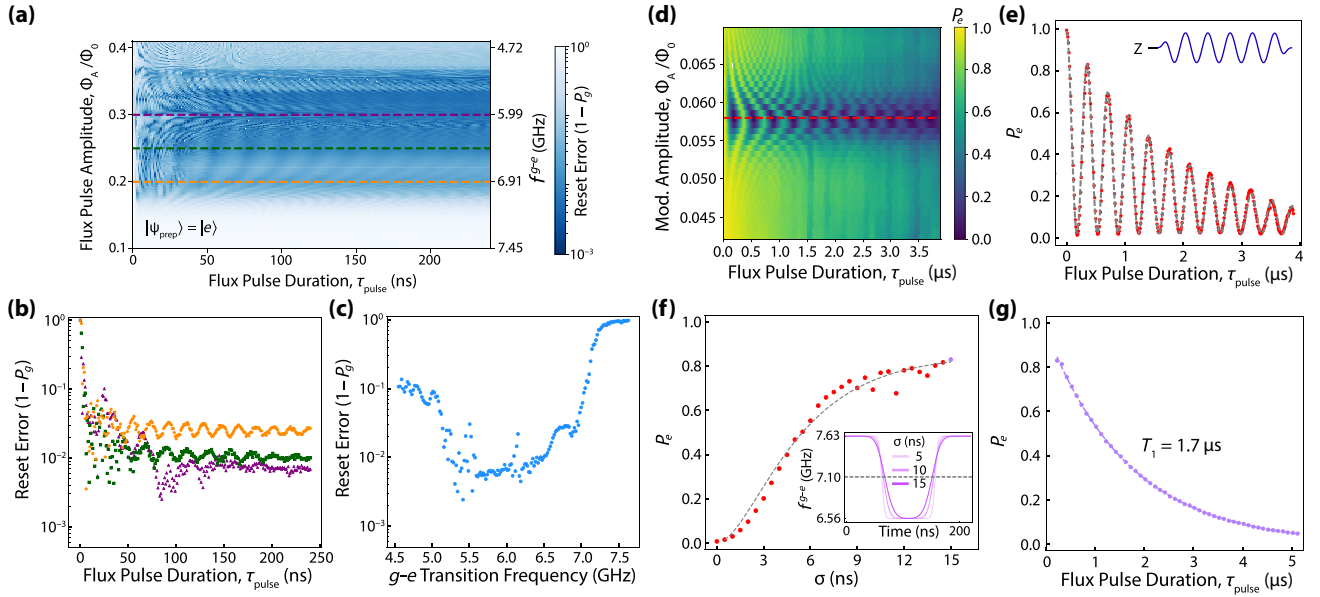


FIG. 8. Reset with direct flux tuning and probing the presence and characterization of the long-lived resonances. (a) Quality of reset with direct flux tuning without ac modulation. The flux pulse amplitude Φ_A and flux pulse duration τ_{pulse} are swept. (b) Cross-sectional view of (a) at various flux pulse amplitudes ($\Phi_A \approx 0.2\Phi_0, 0.25\Phi_0, 0.3\Phi_0$), showing the reset error over different pulse durations. (c) Reset errors at $\tau_{\text{pulse}} = 240$ ns over different flux pulse amplitudes. [Cross-sectional view of (a) at $\tau_{\text{pulse}} = 240$ ns]. (d) Probing the presence of one of the long-lived modes by coherent swap interaction. Here, Q_1 is prepared in $|e\rangle$, followed by a weak parametric flux modulation with $\omega_{\text{mod}} = 250$ MHz. The flux pulse amplitude and flux pulse duration are swept. (e) Cross-sectional view of (d) (red dashed line). The measured populations (red dots) are fitted with a damped Rabi-oscillation model (gray dashed line) [53]. (f) Characterization of the long-lived mode by Landau-Zener tunneling experiments. The tuning rate of f^{g-e} of Q_1 when it crosses the frequency of the long-lived mode is varied by sweeping the filter width σ of the Gaussian kernel used for filtering the square flux pulses. Inset: the measured residual excited-state populations (red dots) are fitted to a model (gray dashed line). (g) Measured residual excited-state population (purple dots) obtained by sweeping τ_{pulse} , with $\sigma = 15$ ns. The fit (gray dashed line) characterizes the relaxation lifetime $T_1 = 1.7 \mu\text{s}$ of the long-lived mode. See the text of this appendix for details of the Landau-Zener tunneling experiment.

approximately 7.1 GHz, which is a signature of two coupled coherent modes, as evidenced by a chevron pattern and perfect fit to a model found in Figs. 8(d) and 8(e) [53].

We then proceed to conduct a quantitative comparison with predictions from the theory of LZS interference. We perform a two-passage experiment, in which the time-dependent g - e transition frequency $\omega_{ge}(t)$ of Q_1 prepared in $|e\rangle$ crosses the frequency of the long-lived resonance (7.1 GHz) twice, by applying a square flux pulse (without ac modulation). Here, the tuning rate of $\omega_{ge}(t)$ when it crosses 7.1 GHz is controlled by the filter width σ of the Gaussian kernel used for filtering the square flux pulse, as illustrated in the inset of Fig. 8(f). The remaining $|e\rangle$ population P_R is measured over different σ from 0 ns (no filtering) to 15 ns, for a fixed duration of $\tau_{\text{pulse}} = 220$ ns, with buffers $\tau_B = 60$ ns at the start and the end. This choice of temporal parameters for flux pulses provides sufficient duration for the tested range of σ , during which Q_1 dumps most of its excitation into the MMWG when it is inside the MMWG passband, such that P_R is dominated by the population retrieved from the long-lived resonance during the second

passage. Therefore, the theoretical prediction for this measurement is given by the following set of equations, rather than the equations assuming maximum coherence for the resonances provided in Ref. [37]:

$$\begin{aligned}
 P_R &= P_0(1 - P_{\text{dia}})^2, \\
 P_{\text{dia}} &= e^{-2\pi\Gamma}, \\
 \Gamma &= \frac{g_{Q_1-M}^2}{d\omega_{ge}/dt|_{\omega_{ge}=\omega_M}},
 \end{aligned} \tag{E1}$$

where P_0 is the revival population in the limit of infinitely fast frequency tuning, M denotes the aforementioned long-lived resonance with a 0-1 transition frequency $\omega_M = 7.1$ GHz, and g_{Q_1-M} is the coupling between Q_1 and M . We find excellent agreement between the measured data and the fit to Eq. (E1), with a strong coupling $g_{Q_1-M} \approx 23.6$ MHz, as shown in Fig. 8(f). In addition, we measure a relaxation lifetime $T_1 = 1.7 \mu\text{s}$ for the resonance M , by sweeping the flux pulse duration with $\sigma = 15$ ns, as provided in Fig. 8(g).

It is worth noting that the rate of adiabatic exchange of population $P_{\text{adia}} = 1 - P_{\text{dia}}$ provided in Eq. (E1) is much greater for higher excited states of the transmons due to their larger transition-matrix elements. Consequently, this imposes challenges on resetting higher excited states using flux tuning with the degree of adiabaticity required for unconditional reset of multiple excited states via narrow-banded auxiliary modes [8]. To overcome this issue, we have chosen parametric flux modulation, which provides continuous dissipation through sidebands while interleaving multiple avoided-crossing passages with long-lived resonances coupled to transmons, and it does not require knowledge or careful characterization of the relevant long-lived resonances.

APPENDIX F: UNCONDITIONAL RESET

1. Unconditional reset from the lower sweet spot

In this appendix, we demonstrate the unconditional reset of a qubit when its idling frequency is below the passband of the MMWG. By positioning the MMWG passband above the qubit's idling frequency, we can achieve a lower residual population in the qubit after reset. This leverages the reduced thermal population at the higher frequency band of the MMWG, which is difficult to attain at the qubit's idling frequency [21,22].

We begin by biasing Q_1 at its LSS, where we measure a g - e transition frequency of $\omega_{ge}^{\text{LSS}} = 3.78$ GHz and relaxation coherence time of $T_1 \approx 30$ μ s. As with the unconditional reset demonstrated in the main text for Q_1 biased at its USS, as illustrated in Fig. 9(a), we apply a parameterized flux modulation pulse to activate sidebands within the passband of the MMWG. Figures 9(b) and 9(c) display the reset errors ($1 - P_g$) as a function of the frequency (ω_{mod}) and amplitude (Φ_A) for a reset pulse with a fixed duration of $\tau_{\text{pulse}} = 104$ ns, for prepared states $|e\rangle$ and $|f\rangle$.

Similar to the reset achieved from the USS, we find numerous frequency-amplitude pairs that enable unconditional reset with errors below 1% for both prepared states. Figure 9(d) shows the reset errors obtained by varying the total pulse duration τ_{pulse} for the specific frequency-amplitude pair indicated by the orange circles in Figs. 9(b) and 9(c). Using the reset-time definition from the main text, we find $\tau_{0.2\%}^e = 97$ ns and $\tau_{0.3\%}^f = 105$ ns, which are faster and as accurate as previously reported unconditional resets. This demonstrates the MMWG's capability for rapid and simultaneous reset of multiple excited states. In addition, we observe a steady-state population of 0.15% from an exponential fit, corresponding to an effective temperature of approximately 28 mK at the transition frequency of 3.78 GHz. This indicates that the MMWG effectively provides a cold bath for reset.

Note that the population estimates shown in Fig. 9(d) are acquired differently from other results in this work due

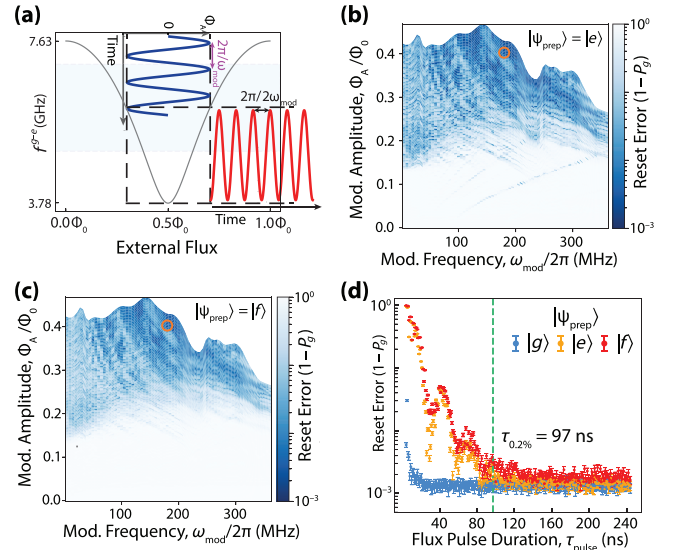


FIG. 9. Unconditional reset from the LSS. (a) Tuning curve of the flux-tunable transmon qubit centered at the LSS. Here, Φ_0 is the magnetic flux quantum. Parking the qubit at the LSS, where it is first-order insensitive to external flux, a flux modulation tone of frequency ω_{mod} (dark-blue curve) modulates the qubit frequency at $2\omega_{\text{mod}}$ (red curve), similar to the case when the qubit is biased at its USS. The region shaded in pale blue denotes the extent of the MMWG passband. (b),(c) Reset errors obtained from scans of reset-pulse parameters ω_{mod} and Φ_A , for Q_1 prepared in (b) $|e\rangle$ and (c) $|f\rangle$. (d) Reset errors obtained by sweeping τ_{pulse} with 1-ns resolution, for $\Phi_A = 0.4\Phi_0$ and $\omega_{\text{mod}} = 180$ MHz [orange hollow circles in (b),(c)]. Error bars indicate 95% confidence intervals obtained from bootstrapping.

to limited ground-state-preparation fidelity, which compromised the accuracy of the confusion-matrix estimation. Instead, we determined the state populations by fitting IQ points from 100 000 repeated measurements to a mixture of Gaussian distributions using the GAUSSIANMIXTURE API from the PYTHON package SCIKIT-LEARN. The confidence interval for the estimated reset error was calculated through bootstrapping with 1000 resamples.

2. Simulation of reset dynamics

We performed a Lindblad master-equation simulation of the reset dynamics using the QUTIP [54] PYTHON package to confirm the reset dynamics observed in measurements. Because of the overhead of quantifying the transfer function of our full flux lines for all the frequencies of interest in the reset measurements, we did not attempt to use these simulations to obtain a quantitative fit to the measured data but only to verify that the measured qualitative behavior was expected in our best theoretical model of the device.

We performed a full dissipationless circuit-quantization procedure using the capacitance and inductance matrices of the device to define its Hamiltonian. Our circuit model was identical to that of Supplementary Material Sec. I in

Ref. [30], including the fit values for the inductances and capacitances, except that we also included the readout resonators of each qubit as lumped-element LC resonators with center frequencies ω_{ri} fit from spectroscopy data of the device, inductances $L_r = 4.5$ nH, and capacitances $C_{ri} = 1/(\omega_{ri}^2 L_r)$. The readout resonators are capacitively coupled to each qubit with capacitance C_{qr} and the unit cell of the MMWG with capacitance C_{wr} . We did not fit the parasitics for the readout resonators.

We modeled dissipation at either end of the MMWG by defining two collapse operators on the modes corresponding to the final taper cells at either end of the waveguide. The associated decay rates for these collapse operators were defined to be the decay rate of the charge on the taper cell capacitor found by modeling the last taper cells as independent (that is, not coupled to the rest of the MMWG) LC resonators coupled capacitively to a $50\ \Omega$ resistor standing in for the environment and solving the equations of motion for the charge. More accurate modeling would incorporate more unit cells of the metamaterial waveguide.

Because we were not considering any drive on the system, and because of the excitation-number-preserving nature of the Jaynes-Cummings Hamiltonian and the fact that the defined collapse operators only remove excitations from the system, we could restrict the total number of excitations in our system to be less than some finite value N . We chose $N = 1$ because, given the total number of degrees of freedom in the MMWG-qubit-RO resonator system, $N = 2$ was prohibitively expensive to simulate.

We simulated the reset of Q_1 for different flux pulses defined by different frequency-amplitude pairs with Q_1 beginning at its USS and a single excitation in the eigenstate of the full Hamiltonian with the largest overlap with the bare Q_1 eigenstate. Reset was simulated using a time-dependent term in the Hamiltonian that varied the frequency of Q_1 over time according to the frequency and amplitude of the flux pulse, assuming the same Gaussian $\sigma = 1$ ns filtering of the flux pulse in simulation as was used in the actual measurement, with reset pulses of total length 104 ns. The results of the simulation can be seen in Fig. 2(d).

3. Evaluation of the demonstrated reset protocol

In Fig. 10, we present a comparison of experimentally demonstrated reset protocols for superconducting qubits, evaluated by reset errors ($1 - P_g$) and the time required to achieve these reset errors (T_{reset}) [8,9,13,16,19–24]. For conditional reset schemes, T_{reset} includes the time needed for readout and the application of feedback on the qubits. For unconditional reset schemes, we quote T_{reset} as either the pulse duration after which reset errors remain below the quoted reset errors or the total reset time including depletion of auxiliary modes.

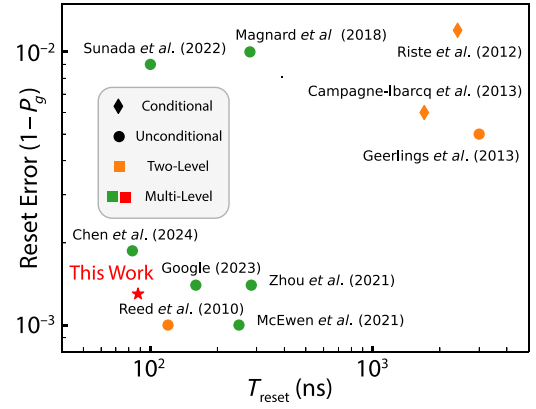


FIG. 10. Comparison of experimentally demonstrated reset protocols. Experimentally characterized reset errors ($1 - P_g$) and the time to reach the quoted reset errors (T_{reset}) of various reset schemes reported in Refs. [8,9,13,16,19–23].

4. Protocol for resetting multiple qubits

The unconditional reset protocol can be extended to support the simultaneous reset of multiple qubits, as illustrated in 11(a). Due to the broad passband of the MMWG, multiple qubit transition sidebands can be accommodated within the passband while suppressing unwanted crosstalk and frequency crowding. A practical strategy is to space the sidebands of qubit transitions sufficiently far apart, as illustrated in 11(b). To minimize reset errors arising from the formation of dark states between qubits or reabsorption of emitted excitations by other qubits in the MMWG, the sideband spacing should be sufficiently larger than the spectral linewidth of each sideband, as indicated by the green shaded regions around the arrows. Such spacing of transition sidebands can be realized by carefully choice of parametric modulation parameters.

APPENDIX G: LEAKAGE-REDUCTION UNITS

1. Calibration of LRU parameters

To characterize LRUs at each combination of parameters, we conduct a three-dimensional sweep of Φ_A , ω_{mod} , and τ_{pulse} . For each frequency-amplitude pair, we determine the LRU duration $\tau_{\text{LRU}}(\omega_{\text{mod}}, \Phi_A)$ by identifying the point at which the population P_f of Q_1 , initially prepared in $|f\rangle$, reaches a local minimum, as shown in Fig. 12(a). This definition of the LRU operation duration is the same as the definition provided in the main text. Figure 12(b) depicts the first-excited-state population P_e after applying the LRU with the calibrated τ_{LRU} , when Q_1 is initially prepared in $|e\rangle$. We find multiple LRU operation candidates, highlighted as the red triangles, where both a short τ_{LRU} and insignificant decay of P_e are simultaneously achieved. Finally, using the iRB technique [41], we estimate the gate fidelity of the LRU operation, followed by virtual-Z correction, when acting as an idling gate on the g - e subspace

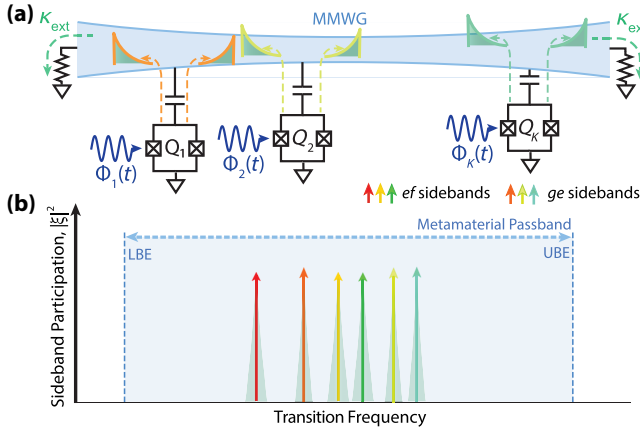


FIG. 11. Schematic of the reset protocol for multiple qubits with a single MMWG. (a) Illustration of reset of multiple qubits $Q_1, Q_2, \dots, Q_K$. External flux pulses $\Phi_i(t)$ cause emission of excitation from Q_i into the MMWG, where $i \in \{1, 2, \dots, K\}$. (b) Strategy for placing transition sidebands for simultaneous reset of multiple qubits. The green shaded regions around the arrows indicate the spectral lineshapes of the sidebands.

of Q_1 . We have only included the points where the leakage population ($P_f < 0.02$) is sufficiently low to ensure effective leakage removal, and where the population revival after reaching the first minimum is less than 20%, excluding parameters where sufficient depletion of the MMWG modes is not achieved.

At each candidate parameter, coherence limits of the gate infidelities (IFs) are calculated using the following formula derived from Ref. [55]:

$$\text{IF}(t_{\text{gate}}, T_\phi, T_1) \equiv 1 - F_g \approx \frac{t_{\text{gate}}}{3} \left(\frac{1}{T_\phi} + \frac{1}{T_1} \right), \quad (\text{G1})$$

where F_g represents the gate fidelity and t_{gate} denotes the duration of the gate. In estimating infidelities, we approximate the LRU pulse as comprising two components: a buffer duration (12 ns) for the rising and falling edge of the envelope, during which the qubit is assumed to experience idling decoherence times ($T_1^{\text{idling}} \approx 12 \mu\text{s}$ and $T_\phi^{\text{idling}} \approx 7.3 \mu\text{s}$), and the on time ($\tau_{\text{LRU}} - 12 \text{ ns}$), during which we assume that the qubit experiences enhanced decoherence times measured under a continuous flux modulation tone, as detailed in the main text.

Figure 12(c) shows the distribution of LRU infidelities for the selected candidate parameters, along with calculated coherence limits and T_1 limits, which represent the contribution of T_1 to the overall coherence limit, as a function of modulation amplitude Φ_A . The LRU infidelities align well with the coherence limit, suggesting that no other error channels play a significant role. At low Φ_A , the LRU infidelities decrease as Φ_A increases, likely due to a reduced contribution from dephasing errors as τ_{LRU} shortens; however, at high Φ_A , despite the shortening of τ_{LRU} ,

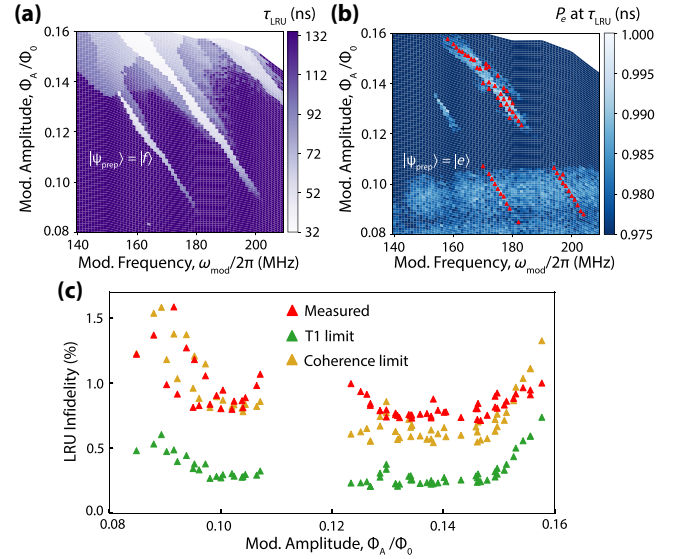


FIG. 12. Calibration of LRU and impact of decoherence on infidelity in the g - e subspace. (a) Plot of τ_{LRU} over a range of $(\omega_{\text{mod}}, \Phi_A)$ combinations. Here, τ_{LRU} is defined as the τ_{pulse} at which P_f reaches the first local minimum when Q_1 is initially prepared in $|f\rangle$, obtained from a three-dimensional sweep of $(\omega_{\text{mod}}, \Phi_A, \tau_{\text{pulse}})$. Note that the points without any local minimum are colored purple. (b) Remaining P_e population with a flux pulse of duration τ_{LRU} , when Q_1 is initially prepared in $|e\rangle$. Red triangles indicate the parameter combinations for LRUs tested with iRB. (c) Measured iRB infidelity of LRUs with different parameter combinations, along with T_1 and coherence limits. The coherence times (T_1 and T_2^*) are obtained from coherence measurements with the corresponding flux modulation tone applied continuously.

the LRU infidelities begin to increase significantly due to increased T_1 decay rates during the sideband interaction. These two competing contributions result in a plateau of minimal infidelities in the range $\Phi_A \in [0.13\Phi_0, 0.15\Phi_0]$, from which we select the LRU parameters used in the LRU operation discussed in the main text. The absence of candidates in the region $\Phi_A \in [0.11\Phi_0, 0.12\Phi_0]$ is likely caused by a collision between the average qubit frequency of Q_1 (zeroth sideband) and a nearby resonance, such as TLS defects. This hypothesis is supported by the observed enhanced decay of P_e across the entire modulation frequency ω_{mod} , which minimally contributes to the average qubit frequency.

2. Infidelity analysis of the LRU operation and ideal MMWG design for LRU

During the flux modulation with amplitude Φ_A and modulation frequency ω_{mod} , the qubit-metamaterial interaction Hamiltonian up to double-excitation subspace can

be written as follows:

$$\begin{aligned} \hat{H}_{Q-MM}^{\text{int}} &\approx \sum_k \sum_{j \in \{1,2\}} \sum_m \lambda_j g_k \xi_m e^{i(\omega_k - \omega_m^{j-1,j})t} |j-1\rangle \langle j| \hat{a}_k^\dagger + \text{H.c.}, \end{aligned}$$

where $|0\rangle$, $|1\rangle$, and $|2\rangle$ indicate the transmon's three lowest-energy eigenstates $|g\rangle$, $|e\rangle$, and $|f\rangle$; $\lambda_j \approx \sqrt{j}$ is the matrix element of the ladder operator for the transition between $|j-1\rangle$ and $|j\rangle$; $\omega_m^{j-1,j} = \overline{\omega^{j-1,j}}(t) + m\omega_{\text{mod}}$ is the frequency of the m th-order sideband of the $|j-1\rangle$ - $|j\rangle$ transition; ξ_m is the amplitude of the m th sideband; k denotes MMWG mode index; ω_k and $\hat{a}_k$ are the corresponding frequency and ladder operator; and g_k is the coupling between MMWG mode k and the qubit.

During an LRU operation, the coherent part of the dynamics of the e - f transition is dominated by the interaction with a single lossy waveguide mode located near the upper band edge of the MMWG. In Fig. 13(a), we show the dynamics of the leakage population P_f with varied LRU durations τ_{LRU} , which is accurately fitted to Eq. (G2) describing a qubit interacting with a lossy waveguide mode denoted with index l :

$$\begin{aligned} H_{\text{eff}} &= \begin{pmatrix} |f0\rangle & |e1\rangle \\ \Delta_{\text{sb}} - i\frac{\Gamma_{ef}}{2} & g_{\text{sb}} \\ g_{\text{sb}} & -i\frac{\kappa_l}{2} \end{pmatrix} \begin{pmatrix} |f0\rangle \\ |e1\rangle \end{pmatrix}, \\ P_f^{\text{sol}}(t) &= (1 - P_{\text{s.s.}}) \left| \frac{\alpha_- e^{\alpha_+ t} - \alpha_+ e^{\alpha_- t}}{\alpha_- - \alpha_+} \right|^2 + P_{\text{s.s.}}, \\ \alpha_{\pm} &= \frac{1}{2} \left(-b \pm \sqrt{b^2 - 4c} \right), \\ b &= \frac{\kappa_l + \Gamma_{ef}}{2} + i\Delta_{\text{sb}}, \quad c = g_{\text{sb}}^2 + i\frac{\Delta_{\text{sb}}\Gamma_{ef}}{2} + \frac{\kappa_l\Gamma_{ef}}{4}, \end{aligned} \quad (\text{G2})$$

where H_{eff} is the effective Hamiltonian for the transmon-lossy MMWG mode's $\{|f0\rangle, |e1\rangle\}$ subspace via sideband interaction, $P_f^{\text{sol}}(t)$ is the analytical solution for Q_1 's $|f\rangle$ state population during the sideband interaction with the initial condition $P_f^{\text{sol}}(0) = 1$, $\Delta_{\text{sb}} \equiv \omega_{-2}^{1,2} - \omega_l$ is the detuning between the lossy mode l and the closest sideband ($m = -2$) of the e - f transition, $g_{\text{sb}} \equiv \sqrt{2}g_l\xi_{-2}$ is the effective coupling through the sideband, κ_l is the damping rate of the lossy mode, $\Gamma_{ef} = 1/T_1^{\text{ef}}$ is the bare relaxation rate of the second excited state ($T_1^{\text{ef}} \approx 4.7 \mu\text{s}$), and $P_{\text{s.s.}}$ is the steady-state population.

On the other hand, the information encoded in the g - e subspace is affected by added decoherence by activating sidebands inside the passband of the MMWG. The total relaxation rates of the transmon's states, $\Gamma_{\text{tot}}^{i-1,i}$, for the $|i-1\rangle$ - $|i\rangle$ transition can be approximated by summing the decay rates of each sideband. Assuming

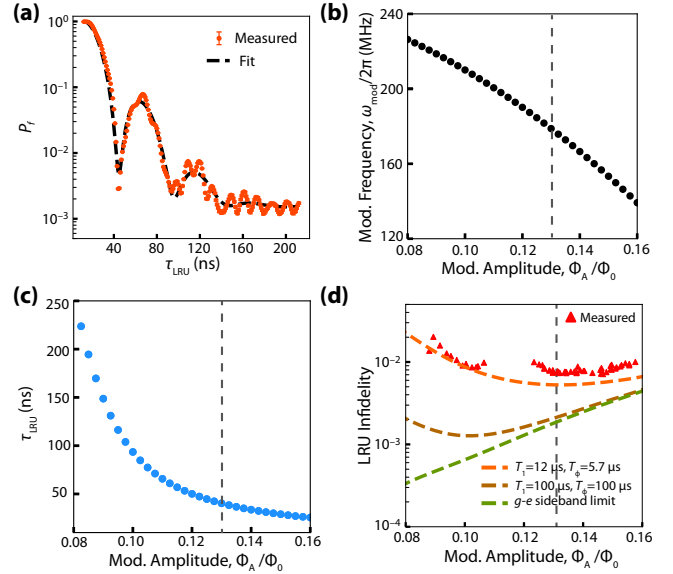


FIG. 13. Approximation to single lossy-mode model and infidelity analysis for the LRU operation. (a) Resulting $|f\rangle$ population obtained after applying the LRU operation, with the same red curve as in Fig. 3(c). The black dashed line is the fit to the single lossy-mode model given by Eq. (G2), with $g_l/2\pi = 46.7$ MHz, $\omega_l/2\pi = 6.93$ GHz, $\kappa_l/2\pi = 17.6$ MHz, and $P_{\text{s.s.}} = 0.15\%$. (b) Modulation frequencies ω_{mod} needed for different modulation amplitudes Φ_A to locate the second lower ($m = -2$) sideband frequency of Q_1 at the lossy mode's frequency ω_l to induce resonant interaction. (c) Plot of τ_{LRU} values estimated from the single lossy-mode model, at which P_f values (i.e., full-swap) reach their first local minima, for different modulation amplitudes Φ_A with modulation frequencies ω_{mod} provided in (b). (d) Measured infidelities and infidelities estimated from Eq. (G4), over different Φ_A values. To understand the interplay between the optimal operating point of the LRU operation and the background decoherence rates of the qubits, we compare infidelity estimates with the measured decoherence rates (dashed red line), a case with improved coherence assuming $T_1 = 100 \mu\text{s}$ and $T_\phi = 100 \mu\text{s}$ (dashed brown line), and the case without any background decoherence, labeled as the g - e sideband limit (dashed green line). Dashed gray lines in (b),(c),(d) indicate the modulation amplitude $\Phi_A = 0.13\Phi_0$ used for the LRU operation shown in the main text, located at the plateau around the local minimum condition predicted by the model.

a constant emission rate inside the MMWG passband, $\Gamma_{1D}^{i-1,i}(\omega) \approx (\lambda_i^2 g_{uc}^2)/J$, and background relaxation rate, $\Gamma_0^{i-1,i} \equiv 1/(T_1^{i-1,i})$, outside the passband, we can use the following approximation:

$$\Gamma_{\text{tot}}^{i-1,i} \approx \Gamma_0^{i-1,i} + \Gamma_{1D}^{i-1,i} \left(\sum_{\omega_m^{i-1,i} \in \text{Passband}} |\xi_m|^2 \right). \quad (\text{G3})$$

The dynamics and infidelity of LRU operations can be modeled using Eqs. (G2) and (G3). Figure 13(a)

shows a fit to the measured P_f for the model of Eq. (G2), from which we obtain $g_l/2\pi = g_{sb}/(\sqrt{2}|\xi_{-2}|) = 46.7$ MHz, $\omega_l/2\pi = 6.93$ GHz, $\kappa_l/2\pi = 17.6$ MHz, and $P_{s.s.} = 0.15\%$. Based on this model, we find the LRU parameters ω_{mod} [Fig. 13(b)] and τ_{LRU} [Fig. 13(c)] at each modulation amplitude Φ_A , such that $\Delta_{sb} = 0$ during the flux modulation and $P_f < 0.02$ after the LRU operation.

With this model and using Eqs. (G3) and (G1), we estimate infidelity of LRU operations for different flux modulation amplitudes as shown in Fig. 13(d) (dashed orange line), from which good agreement with the measured LRU infidelity (red triangles) is found. In addition, we find that the modulation amplitude used in the main text (dashed gray line) is close to the optimal condition predicted by the model.

Further, we analyze the interplay between the background coherence of the qubit and the optimal condition for LRU operation by comparing the infidelity estimates with different coherence times assumed: $T_1 = 12$ μs , $T_\phi = 5.7$ μs (this device, dashed orange line), $T_1 = 100$ μs , $T_\phi = 100$ μs (dashed brown line), and no background decoherence (dashed green line, labeled *g-e* sideband limit). As longer background coherence tolerates longer LRU operation, further improvements in infidelity are achieved by using weaker modulation, which suppresses higher-order sidebands inside the MMWG passband. Therefore, we expect that an LRU infidelity of approximately 0.1% is achievable for qubits with improved background coherence (approximately 100 μs).

APPENDIX H: MICROWAVE-INDUCED LEAKAGE RESET

It is possible to perform a leakage reset of a transmon dispersively coupled to a damped harmonic resonator by driving the transmon with a microwave charge drive at the difference frequency between the $|f\ 0\rangle$ and $|g\ 1\rangle$ states of the coupled system [21,56]. This has the advantage of obviating the need for modulating the transmon frequency, thus avoiding the possibility of tuning the transmons through resonance with any other coupled modes in its spectrum (TLSs, other qubits, etc.) and even allowing a leakage reset to be performed on fixed-frequency transmons.

Microwave leakage reset of a transmon capacitively coupled to a single harmonic mode is possible because the dispersive exchange coupling weakly dresses the bare $|f\ 0\rangle$ and $|g\ 1\rangle$ states in part with the $|e\ 1\rangle$ and $|e\ 0\rangle$ states, respectively. The dressed states can then be coupled to one another via a charge drive on the transmon, as each dressed state contains some component of a state that differs from some component of the other only by a single excitation difference in the transmon part. If this driving is resonant with the energy splitting of the two states, the dressed $|f\ 0\rangle$ and $|g\ 1\rangle$ states will strongly hybridize and undergo Rabi

oscillations, and if the harmonic mode is strongly damped, the hybridized state will then decay to the global $|g\ 0\rangle$ ground state of the system. Because such a drive will in general not be resonant with any transitions out of the $|g\ 0\rangle$ or $|e\ 0\rangle$ states, the reset will only occur for the leakage population in the second excited state, and the logical subspace of the transmon will be preserved.

By replacing the single harmonic mode with a metamaterial waveguide and coupling the transmon to a single unit cell of the waveguide, the qubit couples to the many different harmonic eigenmodes of the metamaterial and can be reset through any of them. The Hamiltonian for such a system can be modeled as

$$H = \omega_q a^\dagger a + \frac{\eta}{2} a^\dagger a^\dagger a a \quad (\text{H1})$$

$$+ \sum_{x=1}^{x=N} \omega_r b_x^\dagger b_x + J(b_x b_{x+1}^\dagger + b_x^\dagger b_{x+1}) \quad (\text{H2})$$

$$+ g(a^\dagger b_{x_0} + a b_{x_0}^\dagger) \quad (\text{H3})$$

$$+ \Omega \cos(\omega_d t)(a + a^\dagger), \quad (\text{H4})$$

where ω_q is the transmon frequency; η is the transmon anharmonicity; x indexes a periodic linear chain of N identical nearest-neighbor coupled resonators comprising the metamaterial waveguide, each with frequency ω_r and nearest-neighbor coupling J ; g is the coupling between the transmon and the resonator indexed by x_0 ; and Ω is the strength of a drive at frequency ω_d on the transmon. The second line gives the bare metamaterial Hamiltonian, which can itself be diagonalized to give the bare eigenmodes of the waveguide in terms of the unit-cell modes, and it will also give the expression for b_{x_0} in terms of the eigenmodes. When the microwave drive frequency ω_d is resonant with the energy splitting between the $|f\ 0\rangle$ state and the $|g\ 1\rangle$ state, where the 0 and 1 indices refer to the 0 and 1 photon states of the single unit cell to which the transmon is coupled, it will induce an effective resonant coupling between the states $\tilde{g} = \frac{1}{\sqrt{2}} \frac{\eta \Omega g}{\Delta(\Delta + \eta)}$, where $\Delta \equiv \omega_q - \omega_r$ [21,56].

To give a simple example of microwave reset using such a waveguide, it is possible to design a waveguide in a regime where g/J is small while g is on the order of at least a few times the free spectral range (FSR) of the metamaterial, around $(4J/N)$. In this regime, a microwave drive resonant with the $|f\ 0\rangle$ state and the waveguide passband will turn on a resonant exchange interaction between the transmon and a bath of multiple waveguide modes, and the transmon's second excited state will roughly undergo exponential decay via photon emission into the waveguide at a rate $\Gamma \sim \tilde{g}^2/J$ [35]. If the ends of the waveguide are well damped to the environment, as with our tapered metamaterial waveguides, the photonic excitation

will propagate away, and the transmon-metamaterial system will decay to its global ground state. For $J/(2\pi) = 100$ MHz, $g/(2\pi) = 250$ MHz, $\eta/(2\pi) = -250$ MHz, and $\Delta_{q-uc}/(2\pi) = -400$ MHz. With these parameters, we get an FSR of around 8 MHz, $\tilde{g}/(2\pi) \sim 25$ MHz, and $\Gamma_{1D}/(2\pi) \sim 6.5$ MHz. It is not necessary to operate in this specific regime; the transmon leakage population should still be theoretically resettable even if the $|f, 0\rangle$ state is so weakly coupled that it interacts with well-defined modes of the waveguide, or so strongly coupled to individual modes that it undergoes nonexponential coherent dynamics with them as the coupled system decays.

-
- [1] D. P. DiVincenzo, The physical implementation of quantum computation, *Fortschr. Phys.: Prog. Phys.* **48**, 771 (2000).
 - [2] H. K. Ng, D. A. Lidar, and J. Preskill, Combining dynamical decoupling with fault-tolerant quantum computation, *Phys. Rev. A* **84**, 012305 (2011).
 - [3] M. D. Reed, L. DiCarlo, S. E. Nigg, L. Sun, L. Frunzio, S. M. Girvin, and R. J. Schoelkopf, Realization of three-qubit quantum error correction with superconducting circuits, *Nature* **482**, 382 (2012).
 - [4] Z. Chen *et al.*, Exponential suppression of bit or phase errors with cyclic error correction, *Nature* **595**, 383 (2021).
 - [5] S. Krinner, N. Lacroix, A. Remm, A. Di Paolo, E. Genois, C. Leroux, C. Hellings, S. Lazar, F. Swiadek, J. Herrmann, G. J. Norris, C. K. Andersen, M. Müller, A. Blais, C. Eichler, and A. Wallraff, Realizing repeated quantum error correction in a distance-three surface code, *Nature* **605**, 669 (2022).
 - [6] R. Acharya *et al.*, Suppressing quantum errors by scaling a surface code logical qubit, *Nature* **614**, 676 (2023).
 - [7] G. P. Gehér, M. Jastrzebski, E. T. Campbell, and O. Crawford, To reset, or not to reset—that is the question, *npj Quantum Inf.* **11**, 39 (2025).
 - [8] M. McEwen *et al.*, Removing leakage-induced correlated errors in superconducting quantum error correction, *Nat. Commun.* **12**, 1761 (2021).
 - [9] K. C. Miao *et al.*, Overcoming leakage in quantum error correction, *Nat. Phys.* **19**, 1780 (2023).
 - [10] F. Battistel, B. Varbanov, and B. Terhal, Hardware-efficient leakage-reduction scheme for quantum error correction with superconducting transmon qubits, *PRX Quantum* **2**, 030314 (2021).
 - [11] J. F. Marques, H. Ali, B. M. Varbanov, M. Finkel, H. M. Veen, S. L. M. van der Meer, S. Valles-Sanclemente, N. Muthusubramanian, M. Beekman, N. Haider, B. M. Terhal, and L. DiCarlo, All-microwave leakage reduction units for quantum error correction with superconducting transmon qubits, *Phys. Rev. Lett.* **130**, 250602 (2023).
 - [12] N. Lacroix, L. Hofele, A. Remm, O. Benhayoune-Khadraoui, A. McDonald, R. Shillito, S. Lazar, C. Hellings, F. m. c. Swiadek, D. Colao-Zanuz, A. Flasby, M. B. Panah, M. Kerschbaum, G. J. Norris, A. Blais, A. Wallraff, and S. Krinner, Fast flux-activated leakage reduction for superconducting quantum circuits, *Phys. Rev. Lett.* **134**, 120601 (2025).
 - [13] D. Ristè, J. G. van Leeuwen, H.-S. Ku, K. W. Lehnert, and L. DiCarlo, Initialization by measurement of a superconducting quantum bit circuit, *Phys. Rev. Lett.* **109**, 050507 (2012).
 - [14] D. Ristè, C. C. Bultink, K. W. Lehnert, and L. DiCarlo, Feedback control of a solid-state qubit using high-fidelity projective measurement, *Phys. Rev. Lett.* **109**, 240502 (2012).
 - [15] J. E. Johnson, C. Macklin, D. H. Slichter, R. Vijay, E. B. Weingarten, J. Clarke, and I. Siddiqi, Heralded state preparation in a superconducting qubit, *Phys. Rev. Lett.* **109**, 050506 (2012).
 - [16] P. Campagne-Ibarcq, E. Flurin, N. Roch, D. Darson, P. Morfin, M. Mirrahimi, M. H. Devoret, F. Mallet, and B. Huard, Persistent control of a superconducting qubit by stroboscopic measurement feedback, *Phys. Rev. X* **3**, 021008 (2013).
 - [17] D. Sank *et al.*, Measurement-induced state transitions in a superconducting qubit: Beyond the rotating wave approximation, *Phys. Rev. Lett.* **117**, 190503 (2016).
 - [18] M. Khezri, A. Opremcak, Z. Chen, K. C. Miao, M. McEwen, A. Bengtsson, T. White, O. Naaman, D. Sank, A. N. Korotkov, Y. Chen, and V. Smelyanskiy, Measurement-induced state transitions in a superconducting qubit: Within the rotating-wave approximation, *Phys. Rev. Appl.* **20**, 054008 (2023).
 - [19] M. D. Reed, B. R. Johnson, A. A. Houck, L. DiCarlo, J. M. Chow, D. I. Schuster, L. Frunzio, and R. J. Schoelkopf, Fast reset and suppressing spontaneous emission of a superconducting qubit, *Appl. Phys. Lett.* **96**, 203110 (2010).
 - [20] K. Geerlings, Z. Leghtas, I. M. Pop, S. Shankar, L. Frunzio, R. J. Schoelkopf, M. Mirrahimi, and M. H. Devoret, Demonstrating a driven reset protocol for a superconducting qubit, *Phys. Rev. Lett.* **110**, 120501 (2013).
 - [21] P. Magnard, P. Kurpiers, B. Royer, T. Walter, J.-C. Besse, S. Gasparinetti, M. Pechal, J. Heinsoo, S. Storz, A. Blais, and A. Wallraff, Fast and unconditional all-microwave reset of a superconducting qubit, *Phys. Rev. Lett.* **121**, 060502 (2018).
 - [22] Y. Zhou, Z. Zhang, Z. Yin, S. Huai, X. Gu, X. Xu, J. Allcock, F. Liu, G. Xi, Q. Yu, H. Zhang, M. Zhang, H. Li, X. Song, Z. Wang, D. Zheng, S. An, Y. Zheng, and S. Zhang, Rapid and unconditional parametric reset protocol for tunable superconducting qubits, *Nat. Commun.* **12**, 5924 (2021).
 - [23] Y. Sunada, S. Kono, J. Ilves, S. Tamate, T. Sugiyama, Y. Tabuchi, and Y. Nakamura, Fast readout and reset of a superconducting qubit coupled to a resonator with an intrinsic Purcell filter, *Phys. Rev. Appl.* **17**, 044016 (2022).
 - [24] L. Chen, S. P. Fors, Z. Yan, A. Ali, T. Abad, A. Osman, E. Moschandrea, B. Lienhard, S. Kosen, H.-X. Li *et al.*, Fast unconditional reset and leakage reduction in fixed-frequency transmon qubits, *ArXiv:2409.16748*.
 - [25] E. Purcell, Spontaneous emission probabilities at radio frequency, *Phys. Rev.* **69**, 681 (1946).
 - [26] E. Jeffrey, D. Sank, J. Y. Mutus, T. C. White, J. Kelly, R. Barends, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, A. Megrant, P. J. J. O'Malley, C. Neill, P. Roushan, A. Vainsencher, J. Wenner, A. N. Cleland, and J. M. Martinis,

- Fast accurate state measurement with superconducting qubits, *Phys. Rev. Lett.* **112**, 190504 (2014).
- [27] M. Mirhosseini, E. Kim, V. S. Ferreira, M. Kalaei, A. Sipahigil, A. J. Keller, and O. Painter, Superconducting metamaterials for waveguide quantum electrodynamics, *Nat. Commun.* **9**, 3706 (2018).
- [28] E. Kim, X. Zhang, V. S. Ferreira, J. Banker, J. K. Iverson, A. Sipahigil, M. Bello, A. González-Tudela, M. Mirhosseini, and O. Painter, Quantum electrodynamics in a topological waveguide, *Phys. Rev. X* **11**, 011015 (2021).
- [29] V. S. Ferreira, J. Banker, A. Sipahigil, M. H. Matheny, A. J. Keller, E. Kim, M. Mirhosseini, and O. Painter, Collapse and revival of an artificial atom coupled to a structured photonic reservoir, *Phys. Rev. X* **11**, 041043 (2021).
- [30] X. Zhang, E. Kim, D. K. Mark, S. Choi, and O. Painter, A superconducting quantum simulator based on a photonic-bandgap metamaterial, *Science* **379**, 278 (2023).
- [31] V. S. Ferreira, G. Kim, A. Butler, H. Pichler, and O. Painter, Deterministic generation of multidimensional photonic cluster states with a single quantum emitter, *Nat. Phys.* **20**, 865 (2024).
- [32] F. Beaudoin, M. P. da Silva, Z. Dutton, and A. Blais, First-order sidebands in circuit QED using qubit frequency modulation, *Phys. Rev. A* **86**, 022305 (2012).
- [33] J. D. Strand, M. Ware, F. Beaudoin, T. A. Ohki, B. R. Johnson, A. Blais, and B. L. T. Plourde, First-order sideband transitions with flux-driven asymmetric transmon qubits, *Phys. Rev. B* **87**, 220505 (2013).
- [34] M. Silveri, J. Tuorila, E. Thuneberg, and G. Paraoanu, Quantum systems under frequency modulation, *Rep. Prog. Phys.* **80**, 056002 (2017).
- [35] G. Calajó, F. Ciccarello, D. Chang, and P. Rabl, Atom-field dressed states in slow-light waveguide QED, *Phys. Rev. A* **93**, 033833 (2016).
- [36] A. González-Tudela and J. I. Cirac, Markovian and non-Markovian dynamics of quantum emitters coupled to two-dimensional structured reservoirs, *Phys. Rev. A* **96**, 043811 (2017).
- [37] S. N. Shevchenko, S. Ashhab, and F. Nori, Landau–Zener–Stückelberg interferometry, *Phys. Rep.* **492**, 1 (2010).
- [38] C. J. Wood and J. M. Gambetta, Quantification and characterization of leakage errors, *Phys. Rev. A* **97**, 032306 (2018).
- [39] G. Ithier, E. Collin, P. Joyez, P. J. Meeson, D. Vion, D. Esteve, F. Chiarello, A. Shnirman, Y. Makhlin, J. Schrieffer, and G. Schön, Decoherence in a superconducting quantum bit circuit, *Phys. Rev. B* **72**, 134519 (2005).
- [40] G. Huber, F. Roy, L. Koch, I. Tsitsilin, J. Schirck, N. Glaser, N. Bruckmoser, C. Schweizer, J. Romeiro, G. Krylov *et al.*, Parametric multi-element coupling architecture for coherent and dissipative control of superconducting qubits, *ArXiv:2403.02203*.
- [41] 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).
- [42] J.-C. Besse, S. Gasparinetti, M. C. Collodo, T. Walter, P. Kurpiers, M. Pechal, C. Eichler, and A. Wallraff, Single-shot quantum nondemolition detection of individual itinerant microwave photons, *Phys. Rev. X* **8**, 021003 (2018).
- [43] K. Reuer, J.-C. Besse, L. Wernli, P. Magnard, P. Kurpiers, G. J. Norris, A. Wallraff, and C. Eichler, Realization of a universal quantum gate set for itinerant microwave photons, *Phys. Rev. X* **12**, 011008 (2022).
- [44] N. M. Sundaresan, R. Lundgren, G. Zhu, A. V. Gorshkov, and A. A. Houck, Interacting qubit-photon bound states with superconducting circuits, *Phys. Rev. X* **9**, 011021 (2019).
- [45] P. Krantz, M. Kjaergaard, F. Yan, T. P. Orlando, S. Gustavsson, and W. D. Oliver, A quantum engineer’s guide to superconducting qubits, *Appl. Phys. Rev.* **6**, 021318 (2019).
- [46] D. T. McClure, H. Paik, L. S. Bishop, M. Steffen, J. M. Chow, and J. M. Gambetta, Rapid driven reset of a qubit readout resonator, *Phys. Rev. Appl.* **5**, 011001 (2016).
- [47] R. A. FISHER, The use of multiple measurements in taxonomic problems, *Ann. Eugen.* **7**, 179 (1936).
- [48] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, Scikit-Learn: Machine learning in Python, *J. Mach. Learn. Res.* **12**, 2825 (2011).
- [49] G. Calusine, A. Melville, W. Woods, R. Das, C. Stull, V. Bolkhovskiy, D. Braje, D. Hover, D. K. Kim, X. Miloshi, D. Rosenberg, A. Sevi, J. L. Yoder, E. Dauler, and W. D. Oliver, Analysis and mitigation of interface losses in trenched superconducting coplanar waveguide resonators, *Appl. Phys. Lett.* **112**, 062601 (2018).
- [50] W. Woods, G. Calusine, A. Melville, A. Sevi, E. Golden, D. Kim, D. Rosenberg, J. Yoder, and W. Oliver, Determining interface dielectric losses in superconducting coplanar-waveguide resonators, *Phys. Rev. Appl.* **12**, 014012 (2019).
- [51] L. Shi, T. Guo, R. Su, T. Chi, Y. Sheng, J. Jiang, C. Cao, J. Wu, X. Tu, G. Sun, J. Chen, and P. Wu, Tantalum microwave resonators with ultra-high intrinsic quality factors, *Appl. Phys. Lett.* **121**, 242601 (2022).
- [52] M. Spiecker, P. Paluch, N. Gosling, N. Drucker, S. Matityahu, D. Gusenkova, S. Günzler, D. Rieger, I. Takmakov, F. Valenti, P. Winkel, R. Gebauer, O. Sander, G. Catelani, A. Shnirman, A. V. Ustinov, W. Wernsdorfer, Y. Cohen, and I. M. Pop, Two-level system hyperpolarization using a quantum Szilard engine, *Nat. Phys.* **19**, 1320 (2023).
- [53] N. Kosugi, S. Matsuo, K. Konno, and N. Hatakenaka, Theory of damped Rabi oscillations, *Phys. Rev. B* **72**, 172509 (2005).
- [54] J. Johansson, P. Nation, and F. Nori, QuTiP: An open-source Python framework for the dynamics of open quantum systems, *Comput. Phys. Commun.* **183**, 1760 (2012).
- [55] L. H. Pedersen, N. M. Møller, and K. Mølmer, Fidelity of quantum operations, *Phys. Lett. A* **367**, 47 (2007).
- [56] M. Pechal, L. Huthmacher, C. Eichler, S. Zeytinoğlu, A. A. Abdumalikov, S. Berger, A. Wallraff, and S. Filipp, Microwave-controlled generation of shaped single photons in circuit quantum electrodynamics, *Phys. Rev. X* **4**, 041010 (2014).