

Passive Quantum Error Correction of Photon Loss at Breakeven

Shruti Shirol¹, Sean van Geldern¹, Hanzhe Xi, and Chen Wang^{1*}

Department of Physics, University of Massachusetts-Amherst, Amherst, Massachusetts 01002, USA



(Received 29 October 2025; accepted 3 March 2026; published 22 May 2026)

Physical qubits in a quantum computer are often represented by superposition states of single particles or excitations. Decay of the excitation itself is a fundamental error channel that is difficult to overcome via external drive or control techniques. Quantum error correcting codes, which encode information in superpositions involving multiple excitations, provide a path to preserve information beyond the capacity of individual excitations, but typically require exquisite active operations on the system. Here, we demonstrate a steady-state driven-dissipative quantum system, composed of a superconducting cavity and a transmon ancilla, that preserves a logical qubit beyond the photon-lifetime limit by about 5% using a binomial encoding. This realization of continuous quantum error correction at the breakeven point highlights the quantitative competitiveness of passive correction strategies, while circumventing some demanding hardware requirements of its active counterparts.

DOI: [10.1103/nvbm-97vs](https://doi.org/10.1103/nvbm-97vs)

Subject Areas: Quantum Physics, Quantum Information

I. INTRODUCTION

Environmental noise in quantum computing devices leads to stochastic evolution of quantum states, i.e., decoherence, the effects of which must be suppressed to avoid logical errors in the computation. Strategies to counter decoherence have been jointly pursued on a stack of three levels: construction of the physical device, choices of the encoding and external drive fields, and implementations of quantum error correction (QEC) circuits. Notably, in the middle level, suppressing errors passively by continuously driving a quantum system [1–5] offers strong appeal in delivering performance gains with minimal overhead. The possibility of such passive protection depends on the nature of the physical noise channels. A common class of noise manifests as quasistatic fluctuations of Hamiltonian parameters, e.g., $1/f$ dephasing [6], which can be suppressed by continuous dynamical decoupling without incurring quantum resource overhead [7–10]. Suppression of dephasing up to higher order has further been accomplished by incorporating QEC encodings in the expanded Hilbert space of driven and/or dissipative oscillators [3,4,11–13] and high-spin systems [14,15].

However, some decoherence processes are practically broadband and necessitate the use of QEC codes to protect the information even at the lowest order. Stochastic loss of

elementary excitations, such as T_1 decay of qubit excitations or cavity photons, belongs to this category and is the current limiting factor of superconducting quantum processors. To overcome this limit, active QEC has been a focus of pursuit using, e.g., the surface code for transmon qubits [16,17], the repetition code for dissipative cat-state qubits [18], and a variety of bosonic codes [19–24]. Because of the quality requirements of the system components and control infrastructure to implement active QEC, only a few state-of-the-art experiments on T_1 -limited quantum hardware are able to overcome the QEC overhead to reach the so-called breakeven point [16,21–24], where the error-corrected logical qubit retains information longer than any individual physical component of the system.

Is it possible to engineer continuous quantum dynamics to protect logical information beyond the lifetime limit of the information-carrying excitations? The foundations of autonomous QEC (AQEC) have been explored since the early 2000s [25,26], where continuously acting Markovian dissipation is considered for reverting the action of discrete errors as they stochastically occur. Several practical implementations of continuous AQEC in superconducting devices have been proposed [27–30] and implemented [31–33]. However, no continuous AQEC experiments so far are able to tailor the dissipation sufficiently well to reach the breakeven point.

In this work, we demonstrate continuous and autonomous quantum error correction at the breakeven point for a superconducting cavity qubit. The logical qubit is encoded with a binomial bosonic code and passively protected from single-photon losses by an efficient parity-recovery dissipation enabled by a three-level superconducting transmon ancilla. We report a time-averaged logical qubit coherence

*Contact author: wangc@umass.edu

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

time of $196 \pm 1 \mu\text{s}$ that narrowly exceeds the T_1 -limited coherence time of individual photons in the cavity, $186 \pm 1 \mu\text{s}$. We further discuss the resemblance of the continuous AQEC actions to active QEC cycles, as well as their corresponding remaining error channels. Together with a growing frontier of discrete-cycle-based AQEC strategies [34,35] (which have also reported breakeven performance during the preparation of this manuscript [36,37] using similar binomial encodings), our results add to an emerging array of strategies to extend logical coherence times at the interface of quantum hardware engineering and error correction.

II. PASSIVE QEC TECHNIQUE

Our AQEC scheme targets a class of rotation-symmetric bosonic codes [38] that encode a logical qubit in an eigenspace of the generalized photon-number parity operator of a cavity, $\hat{\Pi}_m = e^{i\frac{2\pi}{m}\hat{a}^\dagger\hat{a}}$, using code words with equal mean photon number $\bar{n}$. For $m = 2$, this construction divides the Hilbert space into even and odd photon-number subspaces, sufficient to correct a single-photon-loss event. Prominent examples of experimental implementation include the four-legged cat-state code [21] and the kitten code [20,24], where repeated measurements of the photon-number parity lead to real-time detection of photon-loss errors and recovery of the logical information.

In contrast, our scheme passively restores the correct parity by engineering a continuous dissipative process, eliminating the need for real-time feedback. Specifically, we implement parity recovery by selective photon addition (PReSPA) [31], which adds a photon only when the storage mode is in the even-parity error subspace. It can be described by an effective Lindbladian dissipator $\mathcal{D}[\hat{\Pi}_{\text{cor}}]$, where $\hat{\Pi}_{\text{cor}} = \sqrt{\kappa_{\text{cor}}} \sum_n |2n+1\rangle\langle 2n|$. The PReSPA operator realizes an approximate recovery map that fully preserves the logical information in the limit of $\bar{n} \rightarrow \infty$. At finite photon numbers, distortion of the code word coefficients over time leads to a secondary logical dephasing [31], which may be corrected with an engineered four-photon dissipation in the future [39,40].

In this demonstration, we employ a binomial code in an odd-parity subspace with $\bar{n} = 3$,

$$|0_L\rangle = \frac{|1\rangle + |5\rangle}{\sqrt{2}}, \quad |1_L\rangle = |3\rangle, \quad (1)$$

which is illustrated in Fig. 1(a) with its logical Bloch sphere representations. The choice is a result of balancing the code word distortion effect at small $\bar{n}$ and the increased rate of photon loss at larger $\bar{n}$.

To experimentally implement the dissipator operator $\hat{\Pi}_{\text{cor}}$ in a storage cavity mode, we use two auxiliary quantum modes, as represented in Fig. 1(b). The nonlinear transmon ancilla allows photon-number-selective coherent

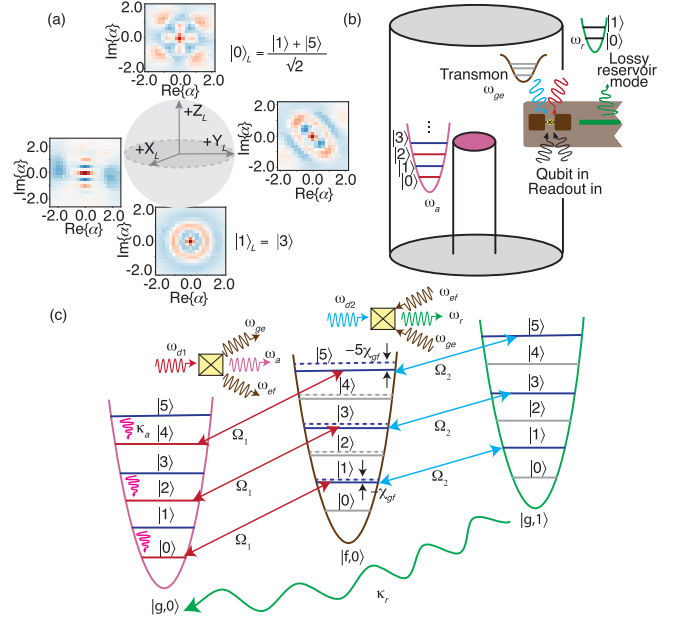


FIG. 1. Experimental setup and correction protocol. (a) The Bloch sphere representation of the logical qubit and the measured Wigner function of the pole ($|0_L\rangle$, $|1_L\rangle$) and the equator ($|+X_L\rangle$, $|+Y_L\rangle$) states. (b) Cartoon depiction of the experimental device, showing the three-dimensional (3D) post cavity storage resonator, transmon ancilla, and quasiplanar readout mode that is also used as a reservoir mode. One input port is used for the transmon drive, transmon readout, and the AQEC drives as shown. (c) Level diagram depicting the transition paths of the AQEC scheme, implemented with two combs of simultaneous continuous drives (red arrows and blue arrows) and dissipation from the reservoir mode (green squiggly arrows). Above the drive-tone arrows, four-wave-mixing diagrams denote the conversion of excitations for each drive.

control of the storage mode through the dispersive shift ($\chi_{ge} = 1.12$ and $\chi_{gf} = 2.07$ MHz). The quasiplanar reservoir mode [41] with a short lifetime ($\kappa_r/2\pi = 0.58$ MHz) extracts entropy by dissipating photons. The relevant terms of the system Hamiltonian in the laboratory frame are

$$\begin{aligned} \frac{\hat{H}_0}{\hbar} = & \omega_a \hat{a}^\dagger \hat{a} + \omega_q \hat{q}^\dagger \hat{q} + \omega_r \hat{r}^\dagger \hat{r} - \frac{\alpha_q}{2} \hat{q}^{\dagger 2} \hat{q}^2 - \chi_{ge} |e\rangle\langle e| \hat{a}^\dagger \hat{a} \\ & - \chi_{gf} |f\rangle\langle f| \hat{a}^\dagger \hat{a} - \chi_{qr} |e\rangle\langle e| \hat{r}^\dagger \hat{r}, \end{aligned} \quad (2)$$

where $\hat{a}$, $\hat{q}$, $\hat{r}$ are the bosonic annihilation operators of the storage, transmon, and reservoir modes, respectively. $|e\rangle$ and $|f\rangle$ represent the first and second excited state of the transmon, and α_q is the transmon anharmonicity. Other higher-order terms (such as cavity anharmonicity) have negligible effects on the AQEC process and are not included in Eq. (2). More details of the Hamiltonian parameters and the coherence times of the components are discussed in Table I in Appendix A.

Two combs of continuous drives at frequency $\omega_{d1} = \omega_a + (2\omega_q - \alpha - n\chi_{gf})$ and $\omega_{d2} = \omega_r - (2\omega_q - \alpha - n\chi_{gf})$

for $n = 1, 3, 5$ are applied simultaneously that implement an effective two-stage cascaded dissipative process, shown in Fig. 1(c). The drive at ω_{d1} activates the parity-selective photon addition and the drive at ω_{d2} implements ancilla reset by converting the transmon excitation to a reservoir-mode photon that is spontaneously emitted. These drive tones realize a rotating-frame (rotating with $\hat{H}_0$) Hamiltonian under the rotating-wave approximation,

$$\frac{\hat{H}_d}{\hbar} = \sum_{n=1,3,5} \Omega_1 |n, f, 0\rangle \langle n-1, g, 0| + \Omega_2 |n, g, 1\rangle \langle n, f, 0| + \text{H.c.} \quad (3)$$

The matched values of $\omega_{d1} + \omega_{d2}$, Ω_1 , and Ω_2 across all three even-to-odd Fock-state conversion processes ensure that the frequency and the temporal profile of the emitted photon from the reservoir mode are identical, which conceals the “which-path” information and preserves the coherence of PReSPA. Similar comb-driving techniques targeting a specific set of photon-number states have also been recently applied in the context of parity mapping and unitary cavity control [24,42].

To obtain fast error correction with minimal spurious actions on the code states, there are two key considerations in our implementation of PReSPA. First, we exploit the intrinsically strongest four-wave-mixing (FWM) sideband transitions available in the transmon-cavity system, which use the $|f\rangle$ state of the transmon as the intermediate state [43]. Compared to the previous study that transits through the $|e\rangle$ state of the transmon [31], this strategy allows us to obtain

large drive rate Ω_2 at much weaker drive fields and hence poses much lower risk of spurious transitions. The use of the $|f\rangle$ state also benefits from its larger dispersive shift, allowing for the use of larger Ω_1 , thus faster overall correction, while satisfying the parity selectivity requirement $\Omega_1 \ll \chi_{gf}$.

Second, we choose a rate hierarchy near the critical damping condition that yields the fastest nonoscillatory relaxation of the logical state back to the code space. We find this condition to be $\Omega_1 : \Omega_2 : \kappa_r \approx \frac{1}{13} : \frac{1}{4} : 1$ in a two-stage cascaded “optical pumping” model of PReSPA, which is a rather generic extension of the well-known Λ -system model. As we discuss in Appendix B, the effective decay rate of the initial state can be inferred from the imaginary component of the eigenvalues of a 3×3 effective non-Hermitian Hamiltonian, as shown in Fig. 2(b) (which shows some surprisingly rich behavior despite the simplicity of the model). Compared to operating in the adiabatic regime with strong κ_r , this strategy allows us to obtain the fastest possible effective dissipation given any practical constraints on Ω_1 or Ω_2 .

III. EXPERIMENTAL RESULTS

To experimentally tune up the optimal drive rates Ω_1 , Ω_2 and to characterize the effectiveness of the dissipative photon-addition process, we measure the time-domain dynamics of the storage photon numbers under the combined drives, starting in different initial even-parity Fock states [Fig. 2(a)]. The photon-number-dependent dispersive shift of the transmon allows for tracking of the $|n\rangle$ -state probability of the storage cavity. This is done by measuring

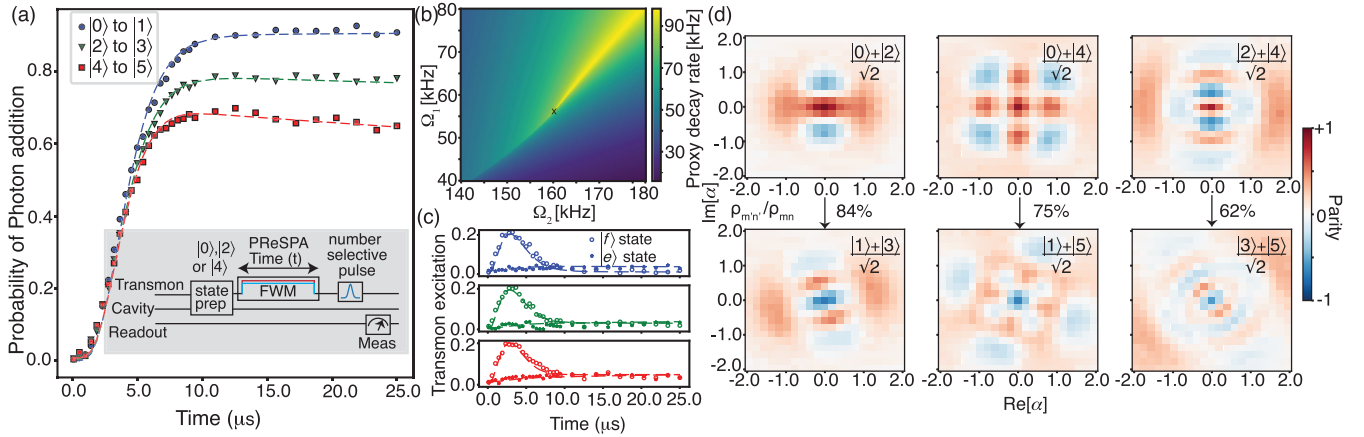


FIG. 2. Characterization of PReSPA performance on even-odd conversion. (a) Starting in even-parity Fock states, we show the cavity populations in converted odd-parity Fock states over pump times. The cavity population in $|n\rangle$ presented here corresponds to $P_{|n,g\rangle} - P_{|n,e\rangle}$, hence a conservative estimate of the success probability of photon addition. (See Appendix A. 6 for discussions.) Dashed lines show the fit to the master-equation simulations, with $\Omega_1 = 55$ and $\Omega_2 = 160$ kHz. (b) Effective rate of the two-stage cascaded dissipation process, evaluated from the smallest imaginary component of the eigenvalues of matrix [Eq. (B2)]. The $\times$ denotes the drive rates used in the experiment. (c) The transmon $|e\rangle$, $|f\rangle$ population for the three different cavity initial states as in (a) over time, using the same master equation fit as (a). (d) Wigner tomography for prepared initial even-parity superposition states (top) and the odd-parity states after 15 μ s of PReSPA drives (bottom). There is a deterministic phase-space rotation due to the self-Kerr of the cavity. The coherence preservation factor, defined by the ratio of the corresponding off-diagonal density matrix elements, is stated next to the arrows between the Wigner functions. All initial states are prepared using numerically optimized control pulses [44].

the transmon after exciting its $|g\rangle \rightarrow |e\rangle$ transition with a spectrally selective microwave pulse centered at frequency $\omega_q - n\chi_{ge}$. The population data over time are fit to the numerical solution of the master equation with Hamiltonian $\hat{H}_d$ and dissipators $\mathcal{D}[\hat{f}], \mathcal{D}[\hat{q}], \mathcal{D}[\hat{a}]$ (with corresponding jump rates), for each even-to-odd-parity Fock-state conversion, with Ω_1 and Ω_2 as fit parameters. In this scheme, the transmon is partially excited to the $|f\rangle$ state that might decay to $|e\rangle$ during the correction cycle, which is also tracked and included in the fit [Fig. 2(c)]. The combs of FWM drives induce an ac Stark shift in the transmon frequency, which is characterized in separate spectroscopy measurements and taken into account in the frequencies of the applied PReSPA drives.

The ability to achieve relatively fast continuous QEC rates at low-drive power is a notable technical advance of this AQEC demonstration. From Fig. 2(a), we achieve a dissipation halftime $\tau_{\text{cor}} \approx 4 \mu\text{s}$, which is sufficiently fast in comparison to the optimal active QEC cycle time in most previous studies of bosonic QEC [20–22]. This PReSPA operating condition fits to drive rates of $\Omega_1/2\pi \approx 55$ and $\Omega_2/2\pi \approx 160$ kHz for each of the three conversion paths, while the total ac Stark shift of the transmon (which is often quoted as a proxy for parametric drive strengths) is less than 100 kHz for all drives combined. This Stark shift is about 2 orders of magnitude lower than the onset of prominent spurious transitions revealed in recent studies, Refs. [45,46] and places our driven system deep in the lowest-order perturbative regime of Josephson nonlinearity. Indeed, we observe no multitone-mixing effects, the required

amplitude of each drive tone to yield the same FWM Rabi rates simply scales as the bosonic factor, and each FWM tone contributes an additional Stark shift of $\Delta_{\text{ss1}} = \Omega_1^2/n\chi_{ge}$ and $\Delta_{\text{ss2}} = \Omega_2^2/\chi_{qr}$ as predicted by perturbation theory (see Appendix B). Minimal Stark shift also makes the experiment insensitive to temporal drifts of microwave power and eliminates the need for frequent calibration. Nevertheless, we still observe a very small transmon heating effect, which we will revisit in the Discussion section. Because of the relatively low-cavity T_1 time of 136 μs (attributed to unexpectedly high surface losses of the aluminum cavity) in this experiment, the conversion fidelity of the photon addition decreases substantially from $n = 1$ to 3 and 5, which limits the overall performance of this AQEC demonstration.

In another set of measurements, we apply PReSPA drives on a pairwise superposition of even-parity Fock states to show preservation of coherence after the photon addition. We perform direct Wigner function tomography using the transmon ancilla to characterize both the initial state and the converted state after applying the drives for 15 μs [Fig. 2(d)] for comparison. Density matrix reconstruction informs that PReSPA produces the corresponding odd-parity states with well-preserved coherence. As the mean photon number increases for the initial even-parity superposition states, the preserved coherence decreases primarily from competing photon-loss events.

We use the process fidelity to quantify the overlap between the experimentally realized quantum map (χ_{exp}) and the identity map ($\mathcal{I}$) on the logical qubit, thereby providing a comprehensive measure of how accurately

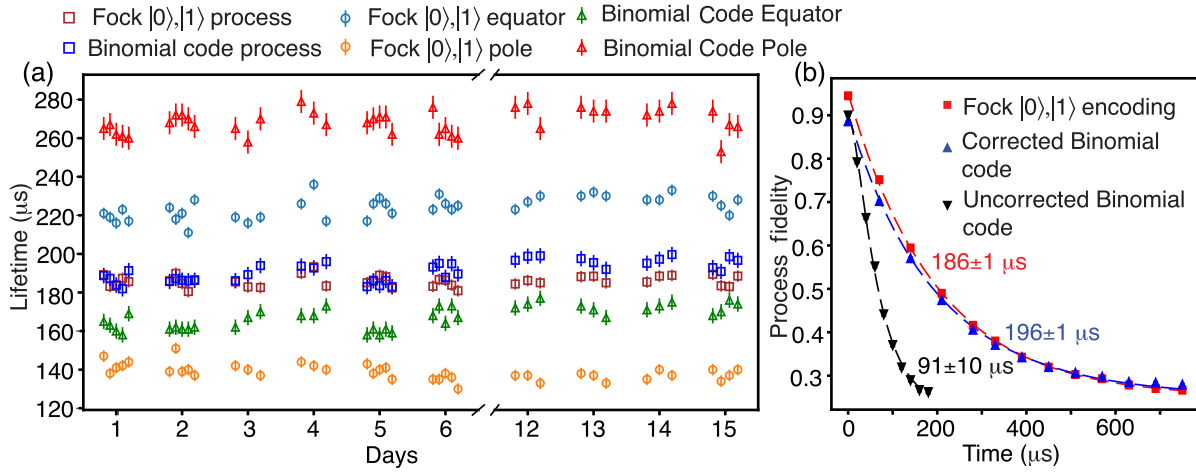


FIG. 3. Logical qubit lifetime under AQEC. (a) The state and process fidelity lifetimes for the corrected binomial code using PReSPA and the $|0\rangle, |1\rangle$ -Fock encoding, as measured over multiple days. Pole- and equator-state fidelities are fit to an exponential decay lifetime with error bars determined from fit uncertainties. The process fidelity lifetimes for both the encodings are computed as the inverse of its decay slope at $t \rightarrow 0$, or $\tau_{\text{process}} = 1/(\frac{2}{3}T_{\text{eq}}^{-1} + \frac{1}{3}T_p^{-1})$. (b) Process fidelity over time of the binomial code with and without correction along with the $|0\rangle, |1\rangle$ -Fock encoding, averaged over all data acquired in the last four days of measurement in (a). The dashed red and blue curves are the sum of two exponential functions with two timescales T_{eq} and T_p , computed from independent exponential fits of the state fidelity of pole and equator states. The uncorrected binomial code is fit to a sum of two exponential functions, taking into account partial information recovery from the decoding process when the cavity has lost only one photon, with the $1/e$ decay time shown in the plot.

PRESPA preserves the encoded information over time. The process fidelity, formally defined as $\mathcal{F}_{\text{process}} = \text{Tr}[\chi_{\text{exp}}\mathcal{I}]$, is linearly related to the experimentally measured average state fidelity $\mathcal{F}_{\text{av}}$ of the six cardinal states of the logical Bloch sphere by $\mathcal{F}_{\text{process}} = 0.25 + 1.5(\mathcal{F}_{\text{av}} - 0.5)$. Because of the difference in exponential decay times of the pole states T_p and equator states T_{eq} , the process fidelity decays as a sum of two exponential functions. Since the decay rate of the process fidelity at the short-time limit is the most relevant metric for the QEC performance, we will quote its inverse as the logical lifetime $\tau_{\text{process}} = 1/(\frac{2}{3}T_{\text{eq}}^{-1} + \frac{1}{3}T_p^{-1})$.

We tracked the fidelity lifetimes (T_p , T_{eq} , τ_{process}) of the corrected binomial code and the best physical qubit of system, a Fock $\{|0\rangle, |1\rangle\}$ encoding, for multiple days, as shown in Fig. 3(a). The AQEC remains effective with no recalibration of PRESPA drives throughout the experiment. Averaged over the whole dataset, the binomial code shows a mean process fidelity lifetime of $\tau_{\text{process}} = 191 \pm 1 \mu\text{s}$ (standard error), which exceeds the Fock $|0\rangle, |1\rangle$ encoding, $\tau_{\text{process}}^{|0\rangle, |1\rangle} = 186 \pm 1 \mu\text{s}$, by 3%. Close inspection of Fig. 3(a) reveals small but statistically significant changes of both the raw physical coherence properties and the logical qubit lifetime over days. Notably, over the last four days, the binomial code under AQEC showed a mean lifetime of $\tau_{\text{process}} = 196 \pm 1 \mu\text{s}$, about 5% beyond its breakeven reference. This improvement may be attributed to a slightly lower excited-state population and better T_1 time of the transmon.

To further illustrate the performance of the continuous AQEC, we plot in Fig. 3(b) the time decay of the process fidelity of the corrected binomial code against the breakeven reference and the uncorrected binomial code, for data averaged over the last four days. Compared to the binomial encoding under free evolution (black), the AQEC provides an improvement of about 2.2 times. The infidelity of the state preparation and decoding leads to lower initial process fidelity for the corrected binomial code (blue) when compared to the Fock encoding (red). The two decay curves cross at $t \sim 400 \mu\text{s}$, showing that we can indeed recover slightly more information from the protected logical qubit than from the physical qubit at long times after accounting for the overhead of encoding and decoding.

IV. DISCUSSION

The continuous AQEC demonstrated here is distinct from the effect of driven dynamical protection. The latter is responsible for protecting spin-locked or Floquet qubits [8–10] and plays a major role in dissipative- or Kerr-cat-state qubits [4, 11–13]. Dynamical protection exploits the (typically low-frequency) spectral feature of the noise, using an engineered effective energy gap or the quantum Zeno effect to prevent errors that cumulate from noisy continuous quantum dynamics over time. The noise spectrum of cavity photon decay, on the other hand, is mostly white over the practically relevant spectral range. Its

dynamics is better understood as comprising intrinsically discrete quantum jumps as a result of Fermi's golden rule coupling to a high density of environmental states, which is unaffected by the PRESPA drives.

Despite the continuous nature of the AQEC, its underlying steps of actions have a direct correspondence with the active QEC paradigm based on measurement cycles, as illustrated in Fig. 4. Continuous AQEC acts as repeated applications of probabilistic correction cycles over infinitesimal time steps [25, 26]. Similar to active QEC, each cycle is composed of conditional excitation of the ancilla for syndrome detection, a correction unitary conditioned on the ancilla state, and an ancilla reset to extract the entropy. The probabilistic nature of the correction lies in an effective weak measurement of the syndrome via the ancilla, as

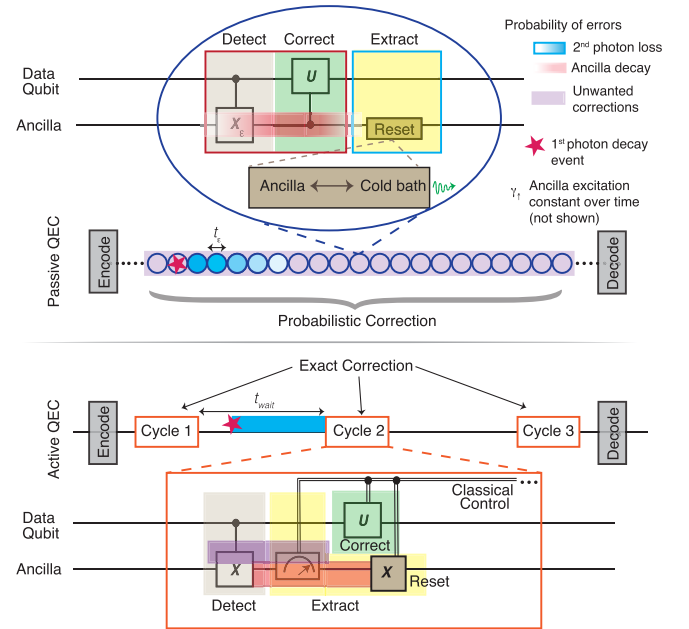


FIG. 4. Comparison of the actions and residual error channels in continuous AQEC versus active QEC. Illustration showing an overview of the passive (top) and active (bottom) QEC schemes. The small circles in the passive case show the correction happening continuously after the error occurs, where each circle can be viewed as a finite probability in the small time step (t_c) of the correction process, outlined in the circuit above, occurring. The discrete error correction cycles in the active case are shown, outlined by the orange rectangles. An enlarged circuit diagram for these measurement-based correction cycles is shown below. The uncorrectable error pathways are shown as shaded regions, with specific errors outlined in the legend in the upper right corner, where the more opaque regions correspond to larger instantaneous error rates. For example, after an initial error, the probability of a second error decreases over time in the passive case thanks to the increasing chance of having completed the correction, shown as a progressively lighter blue shade over time. In the active case, the probability is a constant until reaching the next active QEC operation, thus shown by the constant blue shading.

opposed to the strong measurement in active QEC cycles. Furthermore, the AQEC applies an ancilla-controlled unitary on the encoded qubit (controlled U) for the recovery operation, as opposed a direct U activated by classical logic after measurement. It is worth noting that a hybrid approach of the two paradigms has employed discrete cycles of unitary correction and dissipative reset without measurement feedback [14,15,34–37].

Figure 4 also facilitates an intuitive comparison of main sources of imperfections in our passive AQEC protocol versus active QEC. The susceptibility to uncorrectable second-order photon-loss errors depends on the average time the system spends in the error space after a first photon loss, which is τ_{cor} and $\tau_{\text{wait}}/2$, respectively. Speeding up the correction incurs a trade-off of increased probabilities of injecting uncorrectable errors. A main type of injected errors can be understood as spurious “unwarranted corrections” that drives the code states into error states, which in the case of AQEC is caused by reduced parity-subspace selectivity under stronger dissipation and in the case of active QEC is caused by increased gate and readout errors under more frequent QEC cycles. Assuming a static dispersive coupling between the data and ancilla qubits, both protocols are susceptible to forward propagation of ancilla decay error during correction. Similarly, both types of QEC are susceptible to any spontaneous excitation of the ancilla at all times.

In Appendix E, we present a detailed analysis of these remaining error channels in our binomial qubit in comparison with its closest actively corrected counterpart as studied in Ref. [24]. Assuming the transmon readout performance therein and equal level of hardware coherence, we find that the two protocols provide overall similar performances in corrected logical qubit lifetimes with slight differences in error-budget trade-offs. The main performance advantages of AQEC are lower impact from unwanted corrections from parity-measurement errors and recovery unitary control errors. The main advantages of active QEC are more flexibility to reduce transmon decay errors and the ability to correct for the distortion of the bosonic code word coefficients. Given the hardware coherence in the present device, AQEC slightly outperforms active QEC in our analysis, which shows the competitiveness of this passive strategy in addition to its hardware efficiency in eliminating the technically demanding high-fidelity readout and fast feedback.

Moving forward, the appealing convenience of the passive QEC strategy for superconducting circuit QED comes at the cost of managing always-on parametric drives in the background. It is well known that strong off-resonance drives in Josephson circuits can lead to drive-induced unwanted state transitions (DUSTs) [45–47], which limits fast readout and parametric gate operations [46] and is also a major bottleneck for engineering driven qubits [10,13]. Despite working in the weak-drive regime,

we observe a slow increase of the excited-state occupancy of the transmon over long timescales, translating to an excess transmon heating rate over equilibrium of $\Delta\gamma_{\uparrow} \approx 0.4 \text{ ms}^{-1}$ under our AQEC conditions. Its impact on our reported logical qubit lifetime is insignificant, but will become a major trade-off factor in devices with much better cavity T_1 times. In Appendix D, we present a systematic study of this millisecond-scale DUST effect at low-drive power, which shows an unexpected near-linear dependence on drive amplitude and no apparent memory effect (which should exclude extrinsic factors such as heating of off-chip filters or attenuators). The mechanism of the heating effect remains elusive and is an important topic of future study in order to advance continuous AQEC techniques for cavity bosonic qubits. Future directions further include phase-space stabilization by engineering four-photon dissipation [39,40], increasing code distances by employing higher-photon-number code states [38], and incorporation of logical gate operations [48,49].

In conclusion, we have experimentally demonstrated a passively and continuously corrected quantum memory with logical lifetimes reaching breakeven. Although there is no obvious thermodynamic description of this driven system as in the true sense of self-correction [50], we note that our logical qubit is protected under a time-independent Hamiltonian and dissipation processes in a predefined rotating frame. Our demonstration is, to the best of our knowledge, the first example that quantum information stored in the collective state of elementary excitations under a quasiequilibrium environment persists beyond the limit imposed by the excitation lifetime.

ACKNOWLEDGMENTS

We thank J. M. Gertler for contributions to the device design and B.-J. Liu and Y.-Y. Wang for helpful discussions. This research was supported by the U.S. Department of Energy (DE-SC0021099). We also acknowledge partial support from Defense Advanced Research Projects Agency (HR00112490360).

DATA AVAILABILITY

The data that support the findings of this article are openly available [51].

APPENDIX A: MATERIALS AND METHODS

1. Qubit and package setup

We use a 3D-planar hybrid circuit QED architecture as introduced in Ref. [41]. The 3D package is made with high-purity (99.9995% pure) aluminum that undergoes chemical etching to remove surface contaminants, thus increasing the internal Q for modes. The package is built to have a vertical cylindrical waveguide with a post in its center to make a 3D high- Q $\lambda/4$ -resonator mode in the system (Fig. 2) and a

horizontal cylindrical tunnel where a sapphire chip containing the transmon and a stripline $\lambda/2$ resonator can be inserted. Compared to typical lifetimes of similar cavity modes of approximately 1 ms, in this experiment, we measured an unexpectedly low-cavity T_{1a} of 136 μ s, which we attribute to a degraded quality of the cavity surface over the initial setup phase of the experiment. In a later troubleshooting test (after the completion of all data acquisition of the experiment), we found that a 20-s chemical etching of the aluminum surface (corresponding to an estimated etch of 50 nm) restored the cavity to an earlier-observed T_{1a} of 1.1 ms. The fixed-frequency transmon qubit is fabricated on a double-side-polished sapphire chip using the Dolan bridge method. In a double-angle electron beam evaporator, we deposit layers of Al/ AlO_x /Al that makes the Josephson junction. The first and second layer of aluminum have 40- and 60-nm thickness, respectively. The oxide layer is about 1–2 nm, resulting from 10 min of oxidation at 100 mbar. The 2D (low- Q) stripline resonator that is used for readout and as a cold bath in the AQEC is fabricated on the same chip. The corners of the metal strip are rounded to reduce the surface dielectric loss. The transmon capacitively couples with high- Q ($\chi_{ge}/2\pi = 1.15$, $\kappa_a/2\pi = 1.1$ kHz) and low- Q ($\chi_{qr}/2\pi = 1.13$, $\kappa_r/2\pi = 0.58$ MHz) resonator modes. The coupling between the storage and reservoir mode is undesired and $\chi_{ar}/2\pi \approx 9.3$ kHz. The chip is held in place using a clamp that contains deformable indium. In order to increase (decrease) the χ_{ge} , we add (reduce) more indium in the clamp that pushes (pulls) the chip toward (away) the post cavity.

2. Fridge and measurement setup

The package is mounted on the a bracket inside an Amuneal can on the mixing chamber (MXC). The fridge setup is shown in Appendix E, Fig. 7. With this configuration, we achieved the qubit and cavity coherence values stated in Table I. We have seen an improvement in the transmon relaxation and spurious excitation rates by adding Eccosorb filters, a 30-dB attenuator on the mixing chamber, and an Eccosorb foam on the walls of the Amuneal can. The Eccosorb filters improve qubit T_{1q} by lowering the incident IR radiation [52]. The room-temperature microwave control setup we have is typical for circuit QED architectures where the transmon and cavity drives are generated using sideband mixing of a local oscillator (LO) tone and intermediate frequency (IF) tone from an arbitrary waveform generator (AWG). We transmit the readout drive from the transmon input port to readout output port to probe the transmon state. The transmon input port is also used to drive the transmon. The readout signal exiting the device is amplified at 4 K with a HEMT that has 40-dB gain, followed by a second amplifier with 35-dB gain at room temperature. The signal is demodulated to 50 MHz and amplified again before going to the analog-to-digital converter (ADC). The readout fidelity is 88%, which would be insufficient for high-

TABLE I. System parameters. This table lists all the mode frequencies, coupling strengths between modes, and coherence properties of each mode. We state only the most relevant and significant fourth-order coupling terms of the Hamiltonian. The higher-order (sixth-order and beyond) terms are negligible except for $\chi'_{q'}$, which becomes important at higher storage photon populations and thus added to the Hamiltonian for numerical pulse optimization used for initial-state preparation.

Parameter	Symbol	Value
Transmon frequency	$\omega_q/2\pi$	3482.9 MHz
Storage cavity frequency	$\omega_a/2\pi$	4657.9 MHz
Readout frequency	$\omega_r/2\pi$	8725 MHz
Transmon anharmonicity	$\alpha/2\pi$	134.28 MHz
Storage cavity anharmonicity	$K/2\pi$	3.3 kHz
Storage-transmon $g-e$ coupling	$\chi_{ge}/2\pi$	1.12 MHz
Storage-transmon $e-f$ coupling	$\chi_{ef}/2\pi$	0.95 MHz
Storage-transmon 6 th -order coupling	$\chi'_{q'}/2\pi$	1.9 kHz
Readout-transmon coupling	$\chi_r/2\pi$	1.13 MHz
Readout decay	$\kappa_r/2\pi$	0.58 MHz
Transmon $ e\rangle \rightarrow g\rangle$ decay	T_{1q}	50 μ s
Transmon $ f\rangle \rightarrow e\rangle$ decay	T_{1ef}	31 μ s
Transmon T_2 Ramsey	T_{2R}^*	53 μ s
Transmon T_2 echo	T_{2E}	70 μ s
Transmon $g-f$ T_2 Ramsey	T_{2gf}^*	30 μ s
Transmon $ e\rangle$ population		1.7%
Storage cavity $ 1\rangle \rightarrow 0\rangle$	T_{1a}	136 μ s
Storage cavity T_2	T_{2a}	235 μ s
Storage cavity $ 1\rangle$ -state population		0.6%

fidelity single-shot readout but suffices for the AQEC experiment. Future experiments could improve on readout fidelity with a parametric amplifier and incorporate the midcircuit measurements for erasure detection. The difference in the measured voltages for the transmon in state $|g\rangle$ versus $|e\rangle$ is the raw contrast and is used to scale all the measurements thereafter to determine the cavity and transmon excitation probabilities. Since we have a nonzero excited-state population (1.7%) of the transmon in the equilibrium, we appropriately scale the raw contrast value to approximately equal full $|g\rangle - |e\rangle$ contrast. Details of how the measurements in Fig. 2(a) are scaled are discussed in the following sections.

3. Parameter choices

The frequency placements and coupling strengths are shown in the parameter list presented in Table I. The transmon, storage cavity, and readout mode frequencies are chosen so that the required four-wave-mixing combs are in a suitable range for the microwave setup. The first and second stages of AQEC use pumping tones at frequencies $\omega_{d1} = \omega_a + 2\omega_{ge} - \alpha - n\chi_{gf} = 11.49$ and $\omega_{d2} = \omega_r - (2\omega_{ge} - \alpha - n\chi_{gf}) = 1.89$ GHz, respectively for the mode frequencies as listed in Table I. We place the mode

frequencies such that the FWM frequencies required for PReSPA do not overlap with other modes in the system or induce multiple-photon processes. During the two-stage cascaded PReSPA process, the transmon is driven between $|g\rangle$ and $|f\rangle$ states, which enables an intrinsically stronger interaction owing to the increased participation of the nonlinear transmon mode. We can drive at low enough strengths that produce Stark shifts $\ll \Omega_1, \Omega_2$. For these weak drives, we do not expect multiphoton transitions. The drives also produce consistent Rabi rates without the need for frequent tune-up procedures since the temporal drifts in amplitude or frequency of the drives do not affect the experiment conditions. There is also no parametric mixing effects from multiple tones (e.g., as in Ref. [31]).

The PReSPA drive rates for first and second stage AQEC follow the hierarchy $\Omega_1 < \Omega_2 \ll \chi$ and thus selectively drive when the logical qubit is in the error space manifold. A stronger storage-transmon dispersive coupling strength allows faster correction rates since it can accommodate larger Ω_1 . However, increasing χ_{ge} also increases the other higher-order terms of the Hamiltonian, such as self-Kerr K and χ'_q which contribute to error budget of the correction scheme. By using the state $|f\rangle$ of the transmon in the PReSPA, since now $\Omega_1 \ll \chi_{gf}$ where $\chi_{gf} = \chi_{ge} + \chi_{ef}$, a larger Ω_1 can be accommodated without increasing the coupling between storage and transmon. With the reservoir loss rate $\kappa_r/2\pi = 0.58$ MHz, a critically damped driving is observed at $\Omega_1 = 55$, $\Omega_2 = 160$ kHz. A sweep of pumping rates to find optimal dissipation on the error state is shown in Fig. 2(b) and has already been discussed in the main text.

4. Four-wave-mixing tones generation and calibration

The two four-wave-mixing combs are sideband modulated using two separate LOs and IF channels from an AWG for in-phase and quadrature (IQ) mixing. The three square pulses with 100 ns ramp up and ramp down times with frequency $\chi_{gf}, 3\chi_{gf}, 5\chi_{gf}$ are digitally generated from the AWG IQ pair and mixed with the LO drives to produce $\omega_{d1} - \chi_{gf}, \omega_{d1} - 3\chi_{gf}, \omega_{d1} - 5\chi_{gf}$ simultaneously. The IF drives from another pair of AWG channels mix with the second LO drive to produce $\omega_{d2} + \chi_{gf}, \omega_{d2} + 3\chi_{gf}, \omega_{d2} + 5\chi_{gf}$ simultaneously. These signals are combined using power dividers, amplified, and driven through the qubit input port.

To find the transmon Stark shift in the presence of the AQEC drives, we perform a spectroscopy sweeping the drive frequencies such that $\omega_{d1} + \omega_{d2}$ is constant and measure the transmon conditional on $|1\rangle$ -state cavity population. The applied drives have frequency $\omega_{d1} + \Delta f$ and $\omega_{d2} - \Delta f$ with varying Δf . The Stark shift is predicted from finding this detuning Δf at which the $|1\rangle$ -state population maximizes. As discussed in the main text, the tones in the first comb are scaled by $1/\sqrt{n}$ (where $n = 1, 3, 5$) to compensate for the bosonic factors. The second comb has constant amplitude for all the tones across photon numbers since there is no bosonic enhancement factor to compensate for.

5. State preparation

For various PReSPA calibration and characterization experiments [e.g., Figs. 2(a), 2(c), and 2(d)], we prepare even-parity Fock states and a superposition of even-parity Fock states. An optimal control theory (OCT) pulse, optimized using a gradient-based optimizer from Q-CTRL [53], is used to generate these arbitrary states in the storage cavity. We also used OCT pulses to prepare the six Bloch sphere cardinal states and a decoding pulse that maps the cavity state to the transmon state, which is then characterized to get logical state fidelity and eventually the PReSPA process fidelity in Figs. 3(a) and 3(b).

6. Cavity and transmon population measurement

We probe the photon number in our storage cavity using the transmon ancilla. We use a Gaussian pulse 2.4 μ s long (4σ) that is selective in frequency, allowing us to drive the ancilla conditioned on different Fock states in the storage cavity. In order to get the data in Fig. 2(a), we read the transmon state after PReSPA pumping in two distinct measurements. In the first one, we do not use any transmon drive pulse and read the background right after applying the pump tone. In the consecutive measurement, we use a selective Rabi π_{ge} pulse centered at $\omega_q - n\chi_{ge}$ after the PReSPA pumps that excites the transmon to $|g\rangle \leftrightarrow |e\rangle$ conditional on population in cavity. We can represent the readout values from these experiments in terms of the equations $P_{|n,g\rangle}A + P_{|n,e\rangle}B + P_{|n,f\rangle}C = \text{background}$ (readout value without transmon pulse) and $P_{|n,e\rangle}A + P_{|n,g\rangle}B + P_{|n,f\rangle}C = \text{data}$ (readout value with transmon pulse), where $P_{|n,g\rangle}, P_{|n,e\rangle}, P_{|n,f\rangle}$ are the populations of the respective states and A, B, C are the raw readout values measured when the transmon is initialized in the state $|g\rangle, |e\rangle, |f\rangle$. Subtracting the two equations and dividing by $A - B$ gives the plotted data, $P_{|n,g\rangle} - P_{|n,e\rangle} = (\text{background} - \text{data})/(A - B)$.

This measurement is repeated for all the even-parity Fock initial states and is presented in Fig. 2(a). The population difference $P_{|n,g\rangle} - P_{|n,e\rangle}$ from the numerical solution of the master equation with Hamiltonian $\hat{H}_d$ [Eq. (3)] and dissipators $\mathcal{D}[\hat{r}], \mathcal{D}[\hat{a}], \mathcal{D}[\hat{q}]$ is used to fit these data for all $n = 1, 3, 5$ to extract the drive rates Ω_1, Ω_2 . The fit yields the populations after $t = 15 \mu$ s of PReSPA drives (near convergence) to be $P_{|1,g\rangle} = 0.92, P_{|1,e\rangle} = 0.04$ for the initial state of $|0, g\rangle$. The populations for other odd-parity Fock states after $t = 15 \mu$ s of correction drives for respective even-parity Fock states are $P_{|3,g\rangle} = 0.78, P_{|3,e\rangle} = 0.04$, and $P_{|5,g\rangle} = 0.67, P_{|5,e\rangle} = 0.045$. These $|e\rangle$ -state populations result from a transmon $|f\rangle \rightarrow |e\rangle$ decay in the intermediary step of PReSPA, which terminates the correction prematurely but nonetheless accomplishes a photon-addition process but with a random phase. Based on this analysis, the data in Fig. 2(a) may have underestimated the success probability of PReSPA (approximately $P_{|n,g\rangle}$) by up to 4% and the actual cavity

photon population $P_{|n\rangle}$ by up to 8%. We presented the conservative estimates since we did not measure $P_{|n,e\rangle}$ at the time (which is possible by using number-selective $|e\rangle - |f\rangle$ pulses).

We also measure the intermediary transmon states using a protocol similar to transmon thermal population measurement, which assumes higher states are not occupied [54,55]. A sequence of measurements with Rabi π pulses driving $|f\rangle - |h\rangle$ and $|e\rangle - |h\rangle$ measures the population in $|f\rangle$ and $|e\rangle$.

7. Measuring the fidelity of logical state

The performance of PReSPA on the logical code words is gauged by measuring the fidelity of logical states against the ideal Bloch sphere cardinal states, which can then be used to compute the process fidelity of the correction map. As discussed in the main text, the lifetime τ_{process} is calculated from the pole (T_p) and equator (T_{eq}) logical state lifetimes. When $t \ll T_{1a}$, the process fidelity can be estimated to have a single exponential decay, with a rate that is a weighted average of pole and equator logical state decay rates. The inverse of this average gives the lifetime $\tau_{\text{process}} = [6/(4/T_{\text{eq}} + 2/T_p)]$.

To determine the fidelity of each Bloch sphere cardinal state, the logical state is decoded to the ancilla, which can be characterized with tomography. We use a decoding OCT pulse that performs the following map:

$$|u_0, g\rangle \rightarrow |0, g\rangle, |v_0, g\rangle \rightarrow |0, e\rangle, |u_1, g\rangle \rightarrow |1, g\rangle, \quad (\text{A1})$$

where $u_0 = [(|1\rangle + |5\rangle)/\sqrt{2}]$, $u_1 = [(|1\rangle - |5\rangle)/\sqrt{2}]$, and $v_0 = |3\rangle$. The decoding was also designed to recover states in the error space after single-photon losses, even though it was not as perfect as for the code-space states.

The pole-state fidelity is a straightforward measure of the transmon population but equator states require a frame change where the superposition of $|0_L\rangle$ and $|1_L\rangle$ are stationary. Using Ramsey-like measurement, we find the rotation frequency of $|0_L\rangle$ and that of $|1_L\rangle$ with respect to the state $|0_L\rangle$. A correct phase on the OCT pulse used for decoding changes the frame to that where $|0_L\rangle$ is stationary. After decoding, in the transmon-state tomography step, we drive the transmon around the axis adjusted by the rotation angle between $|0_L\rangle$ and $|1_L\rangle$ at the time of decoding, which is now the angle between $|g\rangle$ and $|e\rangle$. This is repeated for all four equator states, which are then averaged to quote the equator lifetime T_{eq} .

APPENDIX B: MODEL AND OPTIMIZATION OF TWO-STAGE CASCADED DISSIPATION

For efficient pumping, we choose the drive rates that satisfy the critical damping condition for a two-step dissipation scheme, yielding the fastest nonoscillatory relaxation of the logical state back to the code space.

This condition can be determined by studying a simplified version of the full master equation governing the system, which is effectively a non-Hermitian Hamiltonian [56]. We can treat $|n-1, g, 0\rangle, |n, f, 0\rangle, |n, g, 1\rangle$ as a manifold of excited states and $|n, g, 0\rangle$ as ground state for all $n = 1, 3, 5$. Since there are no interactions in the ground-state manifold to consider when optimizing the correction process, we can write an effective non-Hermitian Hamiltonian ($\hat{H}_{\text{eff}}$) as

$$\hat{H}_{\text{eff}} = \hat{H}_d - \sum_{n=1,3,5} \frac{i\kappa_r}{2} |n, g, 0\rangle \langle n, g, 1|. \quad (\text{B1})$$

The Ω_1, Ω_2 are required to match for all the n ; thus, the dynamics for all transition paths are identical. We may use a single 3×3 non-Hermitian Hamiltonian matrix to describe each of the three dissipative processes,

$$\hat{H}_{\text{diss}} = \begin{pmatrix} 0 & \Omega_1 & 0 \\ \Omega_1^* & 0 & \Omega_2 \\ 0 & \Omega_2^* & -i\kappa_r/2 \end{pmatrix}. \quad (\text{B2})$$

The imaginary component of the eigenvalue of the eigenstate that has the largest overlap with $|n-1, g, 0\rangle$ approximately gives the effective dissipation of error state back to code state. In the absence of the Ω_1 drive, this reduces to the familiar three-level λ system, which has a critical damping at $\Omega_2 = \kappa_r/4$. Fixing Ω_2 to this critically damped point of $\kappa_r/4$ maximizes the imaginary component of the eigenvalue corresponding to the largest eigenvector state overlap with $|n-1, g, 0\rangle$. Since this eigenvector will have the largest effective coupling to our initial error state, maximizing the imaginary component of this eigenvector will maximize the amount of dissipation that can be inherited by our initial error state, $|n-1, g, 0\rangle$. Increasing Ω_1 until the eigenvalue associated with our initial error state bifurcates maximizes the dissipation of our error state to our code state with an optimal value of $\Omega_1 \approx \kappa_r/13$. The rates used in the experiment are close to the analyzed optimal drive rates of $\Omega_1 = \kappa_r/13, \Omega_2 = \kappa_r/4$. A numerical plot of the imaginary component of this eigenvalue as a function of drive strengths is shown in Fig. 2(b). For the drive rates higher than what was used in the experiment (denoted by $\times$), the eigenvalue gains a real component, making it oscillatory and hence suboptimal.

APPENDIX C: DRIVEN TRANSMON HAMILTONIAN

In this section, we derive the transmon Hamiltonian under combs of off-resonant drive in the perturbative regime. We show how four-wave-mixing processes occur under frequency-matching conditions and connect the Rabi rate of the FWM transitions Ω_1 and Ω_2 with circuit parameters, especially its simple relationship with Stark

shift, Eq. (C9). The derivation follows the standard fourth-order expansion of the Josephson potential in the displaced frame (as in Refs. [31,57]), which was first introduced in Ref. [3].

We have a system containing two linear resonators and one nonlinear resonator (transmon) that are capacitively coupled with each other. The linear modes are detuned from and weakly couple only to the transmon with strengths $g_{aq}, g_{rq} \ll \omega_a - \omega_q, \omega_r - \omega_q$, where g_{aq}, g_{rq} are storage- and reservoir-transmon coupling strengths. This weak participation of the transmon mode in the high- and low- Q cavity modes induces a nonlinearity that allows for dispersive readout and control of the storage cavity. The Hamiltonian for this three-mode system can be separated into linear and nonlinear parts, such as $\hat{H}_{\text{tot}} = \hat{H}_l + \hat{H}_{\text{nl}}$, where

$$\frac{\hat{H}_l}{\hbar} = \omega_a \hat{a}^\dagger \hat{a} + \omega_q \hat{q}^\dagger \hat{q} + \omega_r \hat{r}^\dagger \hat{r}, \quad (\text{C1})$$

$$\frac{\hat{H}_{\text{nl}}}{\hbar} = -E_J \left(\cos \hat{\phi} + \frac{\hat{\phi}^2}{2} \right). \quad (\text{C2})$$

In this hybridized Hamiltonian, $\hat{\phi}$ is written in terms of the zero-point fluctuation of each mode, $\hat{\phi} = \phi_a(\hat{a}^\dagger + \hat{a}) + \phi_q(\hat{q}^\dagger + \hat{q}) + \phi_r(\hat{r}^\dagger + \hat{r})$. Expanding the cosine to fourth order and keeping only dominant photon-number-preserving nonlinear interaction terms, we arrive Eq. (2) of the main text. The other nonlinear terms such as storage anharmonicity given by self-Kerr $= -(K/2)\hat{a}^{\dagger 2}\hat{a}^2$ and higher-order (sixth-order) dispersive shifts $= (\chi'_q/2)\hat{q}^\dagger \hat{q} \hat{a}^{\dagger 2} \hat{a}^2$ do not affect the correction scheme significantly and are thus omitted from the equation. For corrections on code words with more than three Fock components, the χ'_q leads to slightly off-resonant pumping drives as in Ref. [31].

The FWM drive combs are detuned from the transmon frequency by $\Delta_{d1-\omega_{ge}} \approx 8$ and $\Delta_{d2-\omega_{ge}} \approx 2.7$ GHz. Because of the larger input coupling to the reservoir mode compared to the transmon, the Stark shift on the transmon from these drives is largely dominated by the displacement of the reservoir rather than direct driving. Nevertheless, we provide a general derivation assuming FWM drives couple to both the modes,

$$\frac{\hat{H}_{\text{drive}}}{\hbar} = \sum_{n=0}^2 (\epsilon_{1q,n} \hat{q}^\dagger + \epsilon_{1r,n} \hat{r}^\dagger) e^{i[\omega_{d1} - (2n+1)\chi_{gf}]t} + (\epsilon_{2q} \hat{q}^\dagger + \epsilon_{2r} \hat{r}^\dagger) e^{i[\omega_{d2} + (2n+1)\chi_{gf}]t} + \text{H.c.}, \quad (\text{C3})$$

where $\epsilon_{1q/1r,n}$ is the drive strength for the first step that is constant in time and scales as $\epsilon_{1q/1r,n} = (\epsilon_{1q/1r}/\sqrt{2n+1})$ for $n = 0, 1, 2$ to compensate for the bosonic enhancement factor. However, $\epsilon_{2q/r}$ does not depend on n since

there is no change in the photon number. We are working in the regime where $\omega_q, \omega_a, \omega_r, \omega_{d1}, \omega_{d2}, \omega_{d1/d2} - \omega_{q/r} \gg \epsilon_{1q/r,n}, \epsilon_{2q/r}$, which allows us to consider the drives to be only a perturbation on the system. A unitary transformation to a displaced slowly rotating drive frame simplifies the Hamiltonian. For example, with $\hat{U}_{q,k} = e^{i\omega_k \hat{q}^\dagger \hat{q}} \hat{D}_q(\xi_k) e^{-i\omega_k \hat{q}^\dagger \hat{q}}$ for a drive k with frequency ω_k and drive strength ϵ_k , the Hamiltonian can be transformed to eliminate transmon drive terms from $\hat{H}_{\text{drive}}$, when $\xi_k = [-\epsilon_{q,k}/(\omega_k - \omega_q)]$. The transmon operator transforms in this frame to gain time dependence $\hat{q} \rightarrow \hat{U}_{q,k}^\dagger \hat{q} \hat{U}_{q,k} \rightarrow \hat{q} + \xi_k e^{-i\omega_k t}$. Similarly, the unitary $\hat{U}_{r,k} = e^{i\omega_k \hat{r}^\dagger \hat{r}} \hat{D}_r(\zeta_k) e^{-i\omega_k \hat{r}^\dagger \hat{r}}$, where $\zeta_k = [-\epsilon_{r,k}/(\omega_k - \omega_r)]$ eliminates the drive terms coupled to the reservoir from $\hat{H}_{\text{drive}}$ and the operator gains time dependence $\hat{r} \rightarrow \hat{r} + \zeta_k e^{-i\omega_k t}$. This displacement can be done for each of the six drive tones without loss of generality, leading to $\hat{q} \rightarrow \hat{q} + \sum_{n=0}^2 (\xi_{1,n} e^{-i\omega_{d1,n} t} + \xi_{2,n} e^{-i\omega_{d2,n} t})$ and the reservoir operator $\hat{r} \rightarrow \hat{r} + \sum_{n=0}^2 (\zeta_{1,n} e^{-i\omega_{d1,n} t} + \zeta_{2,n} e^{-i\omega_{d2,n} t})$ in the final displaced reference frame. The operators are then substituted in the $\hat{\phi}$ of the original Hamiltonian,

$$\hat{\phi} = \phi_a \hat{a} + \phi_q \hat{q} + \phi_r \hat{r} + \sum_{n=0}^2 [(\phi_q \xi_{1,n} + \phi_r \zeta_{1,n}) e^{-i\omega_{d1,n} t} + (\phi_q \xi_{2,n} + \phi_r \zeta_{2,n}) e^{-i\omega_{d2,n} t}] + \text{H.c.} \quad (\text{C4})$$

This can be simplified by replacing the weighted sum of displacements with constants $\phi_q \beta_{1,n} = \phi_q \xi_{1,n} + \phi_r \zeta_{1,n}$ and $\phi_q \beta_{2,n} = \phi_q \xi_{2,n} + \phi_r \zeta_{2,n}$.

Finally, we perform another unitary transformation on the Hamiltonian,

$$\hat{U} = e^{i\omega_q \hat{q}^\dagger \hat{q}} e^{i\omega_a \hat{a}^\dagger \hat{a}} e^{i\omega_r \hat{r}^\dagger \hat{r}}. \quad (\text{C5})$$

The operators in this frame are $\hat{a} \rightarrow e^{-i\omega_a t} \hat{a}$, $\hat{q} \rightarrow e^{-i\omega_q t} \hat{q}$, $\hat{r} \rightarrow e^{-i\omega_r t} \hat{r}$ and the Hamiltonian changes as $\hat{H}_{\text{rot}} \rightarrow \hat{U} \hat{H} \hat{U}^\dagger + i\hbar \dot{\hat{U}} \hat{U}^\dagger$, which cancels $\hat{H}_l$. We can then do a fourth-order expansion of the cosine to

$$\begin{aligned} \frac{\hat{H}_{\text{rot}}}{\hbar} = & -\frac{\alpha_q}{2} \hat{q}^\dagger \hat{q}^\dagger \hat{q} \hat{q} - \chi_{ge} |e\rangle \langle e| \hat{a}^\dagger \hat{a} \\ & - \chi_{gf} |f\rangle \langle f| \hat{a}^\dagger \hat{a} - \frac{K}{2} \hat{a}^{\dagger 2} \hat{a}^2 - \chi_{qr} |e\rangle \langle e| \hat{r}^\dagger \hat{r} \\ & - \sum_{n=0}^2 \left[\frac{E_J}{2} \phi_q^3 \phi_a \beta_{1,n} e^{i[\alpha_q + (2n+1)\chi_{gf}]t} \hat{q}^{\dagger 2} \hat{a}^\dagger \right. \\ & \left. + \frac{E_J}{2} \phi_q^3 \phi_r \beta_{2,n} e^{-i[\alpha_q + (2n+1)\chi_{gf}]t} \hat{q}^2 \hat{r}^\dagger \right] + \text{H.c.} \quad (\text{C6}) \end{aligned}$$

The drives also produce a Stark shift on the transmon, which was absorbed in its frequency before changing the

TABLE II. Parameters of the two FWM drive transitions. This table shows the measured Stark shifts (Δ_{ss}) and drive strengths (Ω_n measured) along with the predicted β_n and drive strengths (Ω_n estimated), inferred from the Stark shift value, for the two stages of FWM tones. Note that these values are all only for the tones used in the $|0\rangle \rightarrow |1\rangle$ storage addition process.

FWM drive	Δ_{ss} (kHz)	β_n	Ω_n estimated (kHz)	Ω_n measured (kHz)
$ 0, g, 0\rangle \rightarrow 1, f, 0\rangle$	3	0.0033	57	55
$ 1, f, 0\rangle \rightarrow 1, g, 1\rangle$	32	0.011	190	160

frame in the above equation. These Stark shift terms can be calculated from the fourth-order terms of the cosine expansion. The following are the nonrotating and slowly rotating terms from the expansion that produce the Stark shift on the transmon:

$$\begin{aligned} \frac{\hat{H}_{\text{Stark}}}{\hbar} = & -E_J \phi_q^4 \sum_{n=0}^2 (|\beta_{1,n}|^2 + |\beta_{2,n}|^2) \hat{q}^\dagger \hat{q} \\ & - E_J \phi_q^4 \sum_{k=0}^2 \sum_{\substack{l=0 \\ l \neq k, l > k}}^2 [\beta_{1,k} \beta_{1,l}^* e^{i2(k-l)\chi_{gf}t} \\ & + \beta_{2,k} \beta_{2,l}^* e^{i2(l-k)\chi_{gf}t}] \hat{q}^\dagger \hat{q}. \end{aligned} \quad (\text{C7})$$

Ignoring the slow-rotating cross terms and writing only the static shift in transmon frequency in terms of the known parameters $\chi_{qr}, \chi_{ge}, \alpha_q$, we get the following:

$$\Delta_{ss} = 2\alpha_q \sum_{n=0}^2 (|\beta_{1,n}|^2 + |\beta_{2,n}|^2). \quad (\text{C8})$$

The drive rates to first order are just coefficients of $\hat{q}^{\dagger 2} \hat{a}^\dagger$ and $\hat{q}^2 \hat{r}^\dagger$ from Eq. (C6) with the bosonic factor $\sqrt{2}$ for adding two excitations to the transmon $|g\rangle \rightarrow |f\rangle$ and $\sqrt{2n+1}$, where $n = 0, 1, 2$ for cavity excitation,

$$\begin{aligned} \Omega_{1,n} &= \frac{E_J}{2} \phi_q^3 \phi_a \beta_{1,n} \sqrt{2(2n+1)} = \sqrt{(2n+1)\chi_{ge}\Delta_{ss1}}, \\ \Omega_{2,n} &= \frac{E_J}{2} \phi_q^3 \phi_r \beta_{2,n} \sqrt{2(2n+1)} = \sqrt{(2n+1)\chi_{qr}\Delta_{ss2}}. \end{aligned} \quad (\text{C9})$$

The $\beta_{1,n} = \sqrt{[\Delta_{ss1}/(2\alpha_q)]}$ and $\beta_{2,n} = \sqrt{[\Delta_{ss2}/(2\alpha_q)]}$ are substituted above to get the drive rates in terms of Stark shift on transmon from each tone of the two drive pumps. These calculated drive rates $\Omega_{1,n}, \Omega_{2,n}$ for one transition $n = 0$ are listed in Table II.

APPENDIX D: SPURIOUS QUBIT EXCITATIONS UNDER WEAK OFF-RESONANT DRIVES

A main concern of continuous AQEC with driven dissipation is that the always-on off-resonant drive tones may cause spurious excitation of the transmon, which, in turn, will dephase the logical qubit. To investigate whether this process may occur for the drive tones used in the AQEC experiment reported in the main text, we apply intentionally detuned PReSPA drives to the initial equilibrium state of the system (very close to $|0g0\rangle$) for a long period of time and measure the excited-state population P_e of the transmon. Because of the lack of high-fidelity shot-single measurements, we use the $|e\rangle - |f\rangle$ Rabi oscillation to measure P_e [54,55]. Detuning of tens to hundreds of megahertz was introduced to avoid any resonant excitations instigating the FWM processes. We find steady-state $P_e = 3.8\%$ under the continuous pumps at experimental settings during the same periods when the data presented in the main text were acquired, which is higher than the $P_e = 1.7\%$ when the system is in equilibrium. The transmon T_1 was not degraded under these continuous detuned drives. This additional transmon population can be converted to an effective additional heating rate $\Delta\gamma_\uparrow$ of 0.42 ms^{-1} .

We performed a more systematic measurement of this drive-induced qubit excitation as a function of the drive strength ϵ_1 and ϵ_2 (which can be converted to FWM Rabi rates Ω_1 and Ω_2) in a subsequent cooldown of the same device with slightly different coupling parameters. The transmon $T_{1ge} \approx 170$, $T_{1ef} \approx 75 \mu\text{s}$ in this cooldown is longer than previously measured, allowing for greater resolution in detecting a small heating effect. We further used a driven-dissipative reset sequence to initialize the transmon in $|g\rangle$ by a driving the system from $|0e0\rangle$ or $|0f0\rangle$ to $|0g1\rangle$ from where it spontaneously decays, which increased the throughput of the experiment. We further applied shot-to-shot sweeping of the drive amplitude to distinguish between instantaneous driven response of the system and any long-term cumulative impact on the environment (such as heating up off-the-chip microwave components of the system, which are expected to react in timescales of seconds or longer).

In Fig. 5, we present the measured transmon population in $|g\rangle, |e\rangle, |f\rangle$, or higher states following three different pumping times $t_p = 75, 175, 275 \mu\text{s}$. One single-frequency pumping tone is applied at various detunings from $\omega_{d1} = 11.5 \text{ GHz}$. In the top panels of Figs. 5(a)–5(c), each curve corresponds to one independent measurement where the drive amplitude ϵ_1 is swept along the x axis from left to right from shot to shot, averaged over many rounds of the sweep.

In the experiment, after pumping for time t_p for each ϵ_1 , we obtain readout values (D_1, D_2, D_3, D_4) from four different iterations of the measurements using the following transmon pulse sequences:

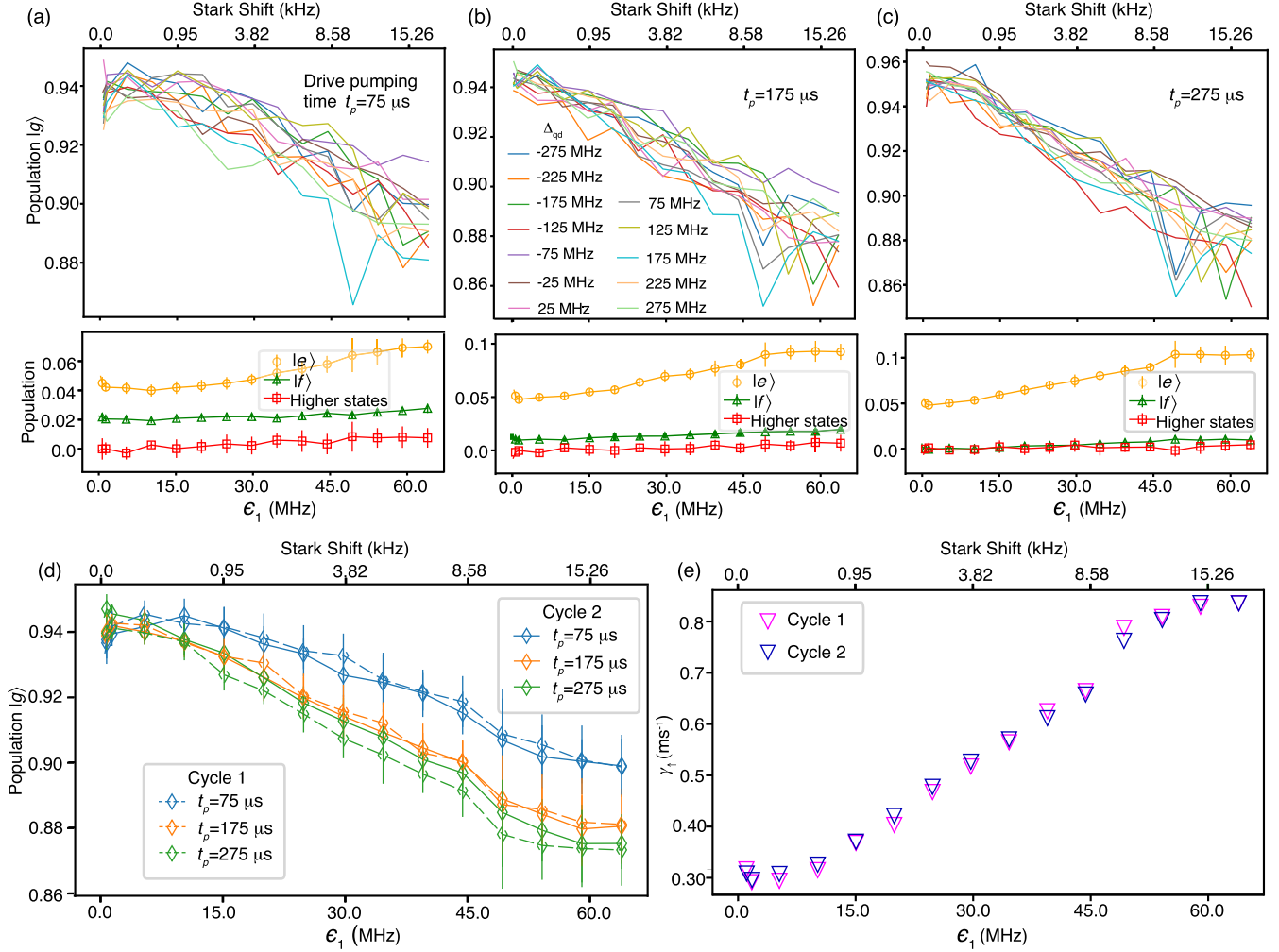


FIG. 5. Transmon excitations by a single-tone off-resonant drive. (a) The $|g\rangle$ -state populations of the transmon driven with a single-parameter (ϵ_1) sweep of an off-resonant tone that is applied for $t_p = 75 \mu\text{s}$. Each curve represents a measurement repeated for a drive frequency detuned from the FWM transition, ω_{d1} . The subplots below show the $|e\rangle$ -, $|f\rangle$ -, and higher excited-state populations, averaged across measurements with different drive frequencies. (b),(c) The measurement is repeated for different drive pumping times $t_p = 175, t_p = 275 \mu\text{s}$, respectively. The shared legend for all the figures is shown in (b). (d) The $|g\rangle$ -state population is averaged across drive frequencies for the three pump times. The error bars represent the spread of the data across drive frequencies. Cycles 1 and 2 are measurements from two iterations of the same experiment. (e) The excitation rate $\gamma_\uparrow$ is analyzed from population data for different drive amplitudes ϵ_1 . The Rabi rate $\Omega_1 = 55 \text{ kHz}$ requires a relatively weak drive, $\epsilon_1 \approx 26 \text{ MHz}$, that yields a Stark shift $\Delta_{\text{ss}} = 3 \text{ kHz}$ and corresponds to $\gamma_\uparrow \approx 0.5 \text{ ms}^{-1}$.

(1) No transmon pulse;
 (2) $|e\rangle \leftrightarrow |f\rangle$ pulse followed by $|g\rangle \leftrightarrow |e\rangle$ pulse;
 (3) $|g\rangle \leftrightarrow |e\rangle$ pulse followed by another $|e\rangle \leftrightarrow |f\rangle$ followed by $|g\rangle \leftrightarrow |e\rangle$ pulse;
 (4) $|g\rangle \leftrightarrow |e\rangle$,
 which give four equations, $gA + eB + fC + R = D_1$, $fA + gB + eC + R = D_2$, $fA + eB + gC + R = D_3$, and $eA + gB + fC + R = D_4$, where g, e, f are the populations in respective states. A, B, C are the readout values (in arbitrary units) for transmon in $|g\rangle, |e\rangle, |f\rangle$ states, and R is a small residual signal from potential higher-excited-state occupancy. We do not generally impose constraints of $g + e + f = 1$ since it is possible that the drive may excite the

transmon to higher-excited states. At small $\epsilon_1 \approx 0$, after sufficient t_p after reset, the system is assumed to be in equilibrium with population in only $|g\rangle, |e\rangle, |f\rangle$ states, which we use for a self-consistent calibration of A, B, C throughout the duration of the experiment. Using the calibrated A, B, C values, the data (D_1, D_2, D_3, D_4) are used to solve for g, e, f , and R giving the population values as plotted in Fig. 5 (where the “higher-state” population is computed as $1 - g - e - f$).

As seen in Figs. 5(a)–5(c), with increasing ϵ_1 the transmon is driven out of the $|g\rangle$ state. In the subpanels, we plot the $|e\rangle$ -, $|f\rangle$ -, and higher- $|h\rangle$ -state populations, as calculated from the above analysis. The populations mostly

do not depend on the detuning of the drives from the FWM frequency, allowing us to plot the average of all of them, with error bars representing the standard deviation of the spread for different frequencies in Fig. 5(d). The pump time of $t_p = 275 \mu\text{s}$ is the longest we have measured, at which the transmon approaches a steady state.

It should be noted that our readout produces some high-excited state beyond $|f\rangle$ due to measurement-induced state transition [47], which cannot be efficiently reset to $|g\rangle$ and is responsible for our $|f\rangle$ -state population of approximately 2% at a wait time $t_p = 75 \mu\text{s}$ after the transmon reset even at negligible drive power. From the data, we observe that the $|f\rangle$ -state population consistently goes down as t_p increases from 75 to 275 μs even at the highest drive power. The dominant effect of increasing drive power appears to be populating the $|e\rangle$ state rather than producing high-excited states which then gradually cascade down the transmon levels over time.

We believe the higher-excited-state population at higher drive amplitude is caused by relatively instantaneous effects within each reset-measurement cycle (of 100–300 μs) rather than any cumulated heating over much longer timescales. This is ensured by our measurement mode of sweeping up ϵ_1 from shot to shot. We have also carried out measurements where we sweep ϵ_1 down, yielding very similar results. (Some minor differences are observed for the first points after a “loop over” which we attribute to a slightly different reset fidelity. Even if it indicates a cumulative heating effort, it poses an upper bound to the memory timescale to $\lesssim 1 \text{ ms}$.)

Motivated by the above observation, we analyze the spurious excitation based on a simple instantaneous heating model. We assume that the ratio between the excitation rates and relaxation rates, $r = \gamma_{\uparrow}/\gamma_{\downarrow}$, between the neighboring transmon states, is the same for $|g\rangle - |e\rangle$ and $|e\rangle - |f\rangle$ transitions (neglecting the minor transition

frequency differences in the detailed balance relation), but is a function of the drive amplitude, i.e., $r = r(\epsilon_1)$. We did not observe any appreciable degradation to the transmon T_1 time in the presence of the drives; hence, we assume the down-transition rates do not change under drives. Then we can model the evolution of the transmon state based on a rate matrix with a single parameter r , which we use to fit the measured qubit population data at four different pumping times ($t_p = 75, 175$, and $275 \mu\text{s}$, plus the initial state $t_p = 0$), independently for each ϵ_1 . The fit result for r is converted to the $|g\rangle \rightarrow |e\rangle$ transition rate $\gamma_{\uparrow} = rT_{1ge}$ and presented in Fig. 5(e). We see that this rate increases approximately linearly with the pump amplitude, except for perhaps a very small curved region near $\epsilon_1 = 0$ where the Stark shift is well below 1 kHz. We acquire data in two cycles, which show similar results over varying amplitudes. The baseline here at equilibrium is 0.3 ms^{-1} .

In Fig. 6, we plot the analyzed population data for a comb of three drive tones near ω_{d1} with fixed amplitude and a comb of three drive tones near ω_{d2} with varying amplitude. This measurement is also carried out at different detunings and averaged over different detuning frequency values. The amplitude of the first drive comb is calibrated to match the AQEC condition presented in the main text. The AQEC condition used in the main text corresponds to $\epsilon_2 = 11.5 \text{ MHz}$, which from the figure gives $\gamma_{\uparrow} = 0.7 \text{ ms}^{-1}$, that is about $\Delta\gamma_{\uparrow} = 0.4 \text{ ms}^{-1}$ above the baseline.

Despite the very different qubit T_1 time and a different attenuator and filter configuration in this cooldown, we find a very similar drive-induced excitation rate to what was coarsely measured during the AQEC experiment. This coincidence, together with the lack of memory effect, suggests that this apparent heating rate may be intrinsic to the driven transmon rather than extrinsic. However, the

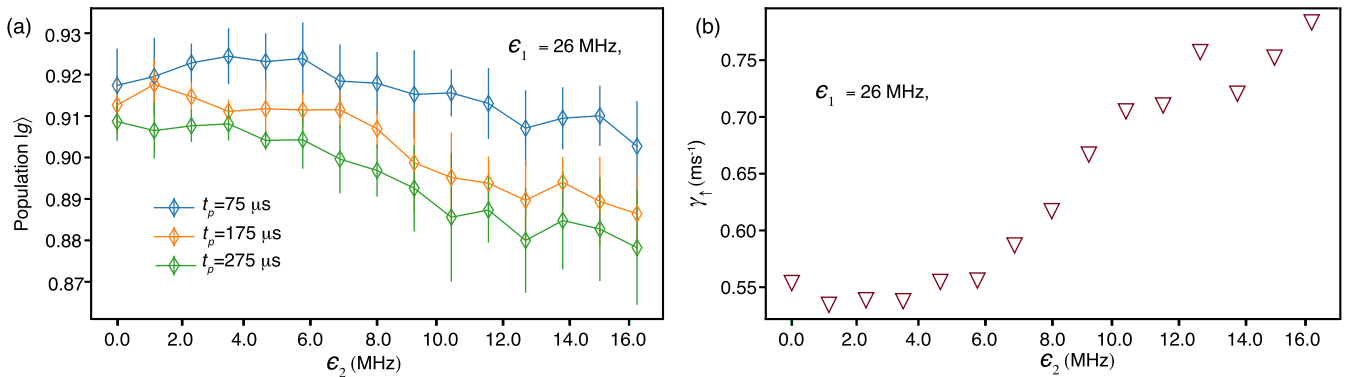


FIG. 6. Transmon excitations by two combs of off-resonant drives. The two combs, each with three tones spaced by $2\chi_{gf}$, are placed at 100s of megahertz detuned from ω_{d1} and ω_{d2} to mimic the FWM pumping strength under the AQEC condition. (a) $|g\rangle$ -state population as a function of E_2 , which is averaged over 12 different drive frequencies detuned from ω_{d2} and is presented for three different pump times. The measurement has one drive comb at fixed strength $\epsilon_1 = 26 \text{ MHz}$ and fixed detuning from ω_{d1} . (b) The $\gamma_{\uparrow}$ rate obtained from analyzing the population data of (a). The drive strengths that produce the Rabi rates Ω_1, Ω_2 used in the experiment are $\epsilon_1 \approx 26$, $\epsilon_2 \approx 11.5 \text{ MHz}$, respectively, which would produce a $\gamma_{\uparrow} \approx 0.7 \text{ ms}^{-1}$.

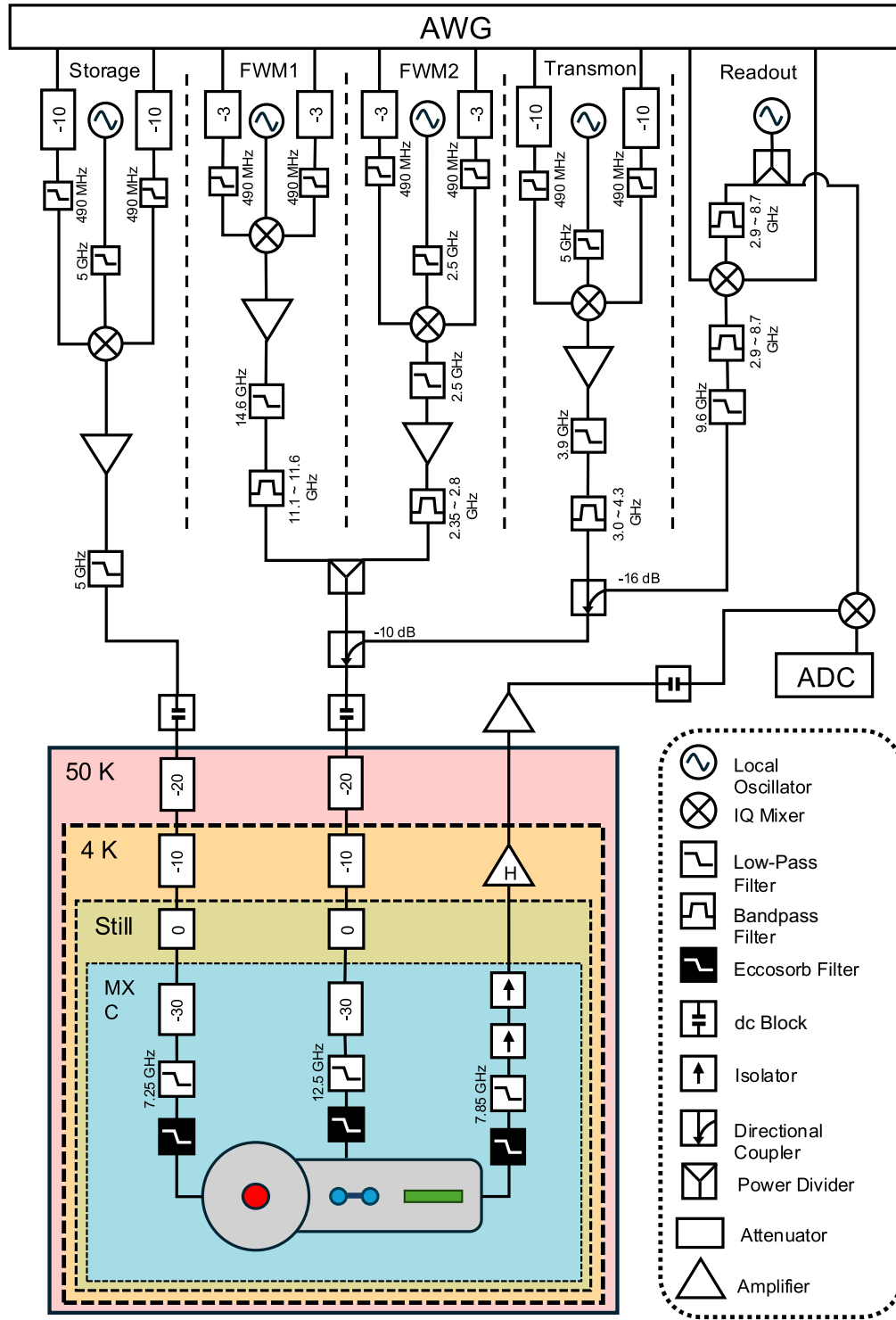


FIG. 7. Dilution fridge diagram and room-temperature microwave setup. Filtering and attenuation and amplification setup for the experiment. There are five independent microwave channels to drive the three modes in the system and two FWM processes. All the coherent control signals are generated by IQ modulation. The storage mode has a dedicated drive port. The transmon and readout modes are driven using the same input port along with the two FWM drives that are combined using power dividers. The output port carries signal that is amplified with low-noise HEMT at 4 K and room-temperature amplifier at the top of the fridge. We use a combination of filters and amplifiers from minicircuits at room temperature and Marki filters suited for cryogenic purposes inside the dilution fridge. The aluminum package containing the sapphire chip and storage cavity is housed inside a magnetic-field-protecting (Amuneal) can that is placed on the mixing chamber of the dilution fridge.

near-linear relationship is particular surprising as any multiphoton off-resonance excitation mechanism should have a quadratic or higher power-law dependence on drive amplitude. In fact, current understanding in DUST [47] suggests the existence of turn-on thresholds that is much higher than the regime explored here. The mechanism of the unexpected DUST observation in this low-power pumping regime with multimillisecond transition rates will therefore require a theoretical understanding of weaker and more subtle effects in the driven circuit QED systems.

APPENDIX E: COMPARING CONTINUOUS PASSIVE QEC WITH ACTIVE QEC IN CIRCUIT QED

In this appendix, we do a comparative analysis of the errors in the two approaches, i.e., active measurement based and passive QEC for binomial qubits with a transmon ancilla. The dominant error channel, single-photon loss at rate κ_a , induces both jump and no-jump evolution that need to be corrected. In active measurement-based corrections, such as Ref. [24], which we will extensively use as our comparison, a correction unitary was applied based on a measurement outcome that corrected for jump or no-jump errors accumulated over a finite time t_{wait} . In our protocol, however, we engineer a dissipative scheme to correct for the parity-flip error from the jumps. There remains a distortion effect in the code word coefficients from the jump or no-jump evolution at finite $\bar{n}$ that would ultimately need a four-photon dissipative process [40] to fully stabilize. To mitigate this uncorrected effect, we used a larger photon number, $\bar{n} = 3$, compared to $\bar{n} = 2$ in

Ref. [24]. Apart from the jump and no-jump errors, the bosonic logical states are also affected by ancilla errors and the induced nonlinearities in the harmonic oscillator modes, which are not corrected for in either of the correction schemes. As a pedagogical comparison, we will address each error avenue and compare its mechanism and magnitude of error for the two different schemes, with a full summary listed in Table III.

1. Double-photon-loss errors

Incurring a second photon-loss error before the first is corrected is the canonical second-order error that is uncorrected in a first-order QEC implementation. Its rate is the result of direct competition between the raw error rate $\bar{n}\kappa$ and the rate of error correction κ_{cor} . For a binomial code, a double-photon-loss error causes an effective bit-flip error, while leaving the parity unchanged and is thus uncorrectable in either passive or active QEC. The double-photon-loss rate is $\bar{n}^2(\kappa_a^2/\kappa_{\text{cor}})$ for both correction schemes. This expression can be derived from Poisson statistics in the limit of $\kappa_{\text{cor}} \gg \kappa_a$. We have defined the correction rate (κ_{cor}) as the inverse of one half an active QEC cycle time τ_{cyc} (which is often chosen to be slower than the speed limit of parity tracking), $\kappa_{\text{cor}} = 2/\tau_{\text{cyc}}$. In the quantitative example comparison of Table III, we choose $\tau_{\text{cyc}} = 2/\kappa_{\text{cor}} = 2\tau_{\text{cor}} = 8 \mu\text{s}$ for the active QEC, which is twice the characteristic time of our experimental AQEC ($\tau_{\text{cor}} = 1/\kappa_{\text{cor}} = 4 \mu\text{s}$, which we will also refer to as the effective passive “cycle time”) so that the baseline second-order error rate is held equal. The heuristic picture for this factor of 2 difference is that the continuous correction starts only

TABLE III. Error-budget comparison of active versus passive QEC. Here we present how each error contributes to the loss of fidelity in the logical code space in the measurement-based versus passive QEC in the binomial code. We show an approximate analytic scaling on the left side of the columns and an associated calculated error rate using the device parameters from this AQEC experiment on the right side of the columns. For computing the error budget in active QEC, we consider two cases: (1) A base case with the same $\bar{n} = 3$ binomial code and correction rate $\kappa_{\text{cor}} = (2/\tau_{\text{cyc}}) = (4 \mu\text{s})^{-1}$ as in our AQEC experiment for a more direct comparison and (2) an example case with the $\bar{n} = 2$ kitten code and $\kappa_{\text{cor}} = (8 \mu\text{s})^{-1}$ optimized for overall logical error rate. In either case, we have applied the performance metrics of the ancilla readout, reset, and recovery unitary from Ref. [24] for calculating error rates: $\tau_{\text{ex}} = 1.3 \mu\text{s}$ is the total ancilla excitation time after a positive syndrome measurement; $\epsilon_{\text{meas}} = 1.4\%$, $\epsilon_j = 3.9\%$, and $\epsilon_{\text{nj}} = 1.4\%$ are the readout, jump correction, and no-jump correction infidelities, respectively. The notation $f(x)$ denotes a functional dependence whose detail is beyond the scope of the discussion.

Error sources	Active measurement-based QEC		Continuous AQEC	
	Error rate (ms^{-1})		Analytic scaling	Error rate (ms^{-1}) ($\bar{n} = 3$)
Double-photon loss	$\bar{n}^2\kappa_a^2/\kappa_{\text{cor}}$	1.8	$\bar{n}^2\kappa_a^2/\kappa_{\text{cor}}$	1.8
Ancilla relaxation $ e\rangle \rightarrow g\rangle$ or $ f\rangle \rightarrow e\rangle$	$\bar{n}\kappa_a(\tau_{\text{ex}}/T_{1ge})$	0.5	$(\bar{n}\kappa_a)/(5\kappa_{\text{cor}}T_{1ef})$	0.6
Unwanted corrections	$\frac{1}{2}\epsilon_{\text{meas}}\kappa_{\text{cor}}$	1.3	$4\kappa_{\text{cor}}^3/\chi_{gf}^2$	0.4
Imperfect recovery unitary	$\frac{1}{2}\epsilon_{\text{nj}}\kappa_{\text{cor}} + \bar{n}\kappa_a\epsilon_j$	2.6	$\kappa_a f(\bar{n})$	1.3
Storage-mode nonlinearity	$\frac{1}{6}(K/\kappa_{\text{cor}})^2\bar{n}\kappa_a$	0.1	$\frac{1}{6}(K/\kappa_{\text{cor}})^2\bar{n}\kappa_a$	0.1
Ancilla spurious excitation $ g\rangle \rightarrow e\rangle$	$\gamma_{\uparrow 0}$	0.3	$\gamma_{\uparrow 0} + f(\Omega_1, \Omega_2)$	0.7
Ancilla dephasing	$(\pi\bar{n}\kappa_a)/(\chi_{ge}T_\phi)$	0.2	$(\bar{n}^2\kappa_a^2)/(\kappa_{\text{cor}}\Omega_2^2T_\phi^2)$	<0.01

after a photon-loss event, thus the system is susceptible to the second photon-loss error for the duration of the characteristic correction time $1/\kappa_{\text{cor}}$. The active QEC cycles run on a schedule uncorrelated with the loss events. Thus, on average, the time interval between the first loss event and the next correction operation is $\tau_{\text{cyc}}/2$, motivating our definition of $\tau_{\text{cyc}}/2 = \tau_{\text{cor}}$.

2. Unwanted corrections

The interactions in both schemes leading to correction come at a cost, since they also introduce associated errors. In the passive scheme, the combs of four-wave-mixing tones can cause detuned transitions out of the logical code space. In the measurement-based scheme, incorrect readout assignment can lead to an incorrect unitary correction being applied, thus scrambling the information and driving the system out of the logical code space.

For the error rate of a detuned transition in the continuous scheme, we can refer to the 3×3 matrix describing the cascaded dissipative dynamics, Eq. (B2), with a detuning term of χ ,

$$\hat{H}_{\text{detune}} = \begin{pmatrix} 0 & \Omega_1 & 0 \\ \Omega_1 & \chi & \Omega_2 \\ 0 & \Omega_2 & -i\kappa/2 \end{pmatrix}. \quad (\text{E1})$$

In the limit of $\kappa \gg \Omega_2 \gg \Omega_1$, the imaginary component of the eigenvalue of interest yields a detuned transition rate of $[(\Omega_1^2 \Omega_2^2)/\chi^2 \kappa]$, which approaches approximately Ω_1^3/χ^2 as the drives approach the critical damping regime. This shows the clear benefit that an increase in χ provides to this error. It should be noted that this detuned transition can be further mitigated to a large degree in a hybrid active-passive QEC scheme [34,35] by using cascaded pulses with spectrally narrow envelopes instead of continuous ones. If the first pump tone uses a Gaussian envelope with a temporal width of σ_q , the spectral width will also follow a Gaussian dependence, where the amplitude at our detuned transition will be $\propto e^{-\chi^2 \sigma_q^2/2}$. By choosing a Gaussian pulse with appropriate width, one can minimize destabilization of the code state from these detuned transitions to a great extent.

In the active scheme, this error comes from a measurement infidelity (ϵ_{meas}). Reference [24], using the measurement-based approach, had on average $\epsilon_{\text{meas}} \sim 1.4\%$ measurement infidelity. Therefore, frequent application of active QEC cycles are often undesirable. There is substantial room for improvement on readout fidelity, but it faces strong limitations in measurement-induced state transitions, which is beyond the scope of this discussion. The passive case has a lower error rate for this pathway with our current experimental parameters (0.4 versus 1.3 ms^{-1} , as in Table III). Since the measurement-based scheme error rate is $\propto \kappa_{\text{cor}}$, while the continuous approach is $\propto \kappa_{\text{cor}}^3$, the passive scheme

benefits greater from a reduced κ_{cor} as storage lifetimes increase.

3. Ancilla errors

The ancilla is an integral part of a QEC protocol. The errors on the ancilla, such as $|e\rangle \rightarrow |g\rangle$ decay, pure dephasing, or spurious $|g\rangle \rightarrow |e\rangle$ excitation (with rate $\gamma_{\uparrow}$) may lead to a loss of logical information. Since the ancilla is utilized in both schemes a bit differently, the rate at which it affects the logical states differ.

The errors from the heating rate $\gamma_{\uparrow}$ give a constant error rate, irrespective of the correction rate, in both protocols. However, in the passive case, due to the external drives, the effective $\gamma_{\uparrow}$ is higher than the bare undriven rate $\gamma_{0\uparrow}$, giving a benefit to the active case. Furthermore, when the corrections are done in cycles rather than continuously, it is, in principle, possible to decouple the transmon during wait times [58] using a tunable ancilla to mitigate the dephasing effect of excitations during the idle wait times. In the continuous AQEC, the correction happens stochastically; thus, we cannot completely decouple the transmon from the oscillator mode. A possible approach for the passive scheme to mitigate this issue is to use strong drives that cancel χ_{ge} , as done in Ref. [59], to protect the bosonic state from dephasing from the transmon $|g\rangle - |e\rangle$ transitions (while our PReSPA uses the $|f\rangle$ state). As already discussed, we observe that with strong drives the $\gamma_{\uparrow}$ rate increases; thus, any drives in addition to the four-wave-mixing tones used for correction are going to increase it more and this could pose experimental challenges.

The second ancilla error channel is the $|e\rangle \rightarrow |g\rangle$ or $|f\rangle \rightarrow |e\rangle$ decay, which causes cavity logical dephasing. It is relevant only for the time in which the ancilla is in the excited state during the correction process. In the passive case, the ancilla is in the $|f\rangle$ state for approximately 20% of the correction time t_{cor} under critical damping dynamics. Thus, the probability of ancilla decay for each correction is roughly $[t_{\text{cor}}/(5T_1)]$. The active protocol is susceptible to ancilla decay starting from parity mapping and ending at ancilla reset. As the cavity lifetimes become longer, we could perform fewer correction cycles over time in the active case and thus reduce the ancilla- T_1 -induced errors in the active scheme. The situation with continuous AQEC is more complicated. If we assume the use of critically damped cascaded dissipation (which minimizes unwarranted corrections and heating), a slower correction implies the use of slower reservoir decay rate κ_r . This leaves the ancilla in the excited state for longer and increases the probability of ancilla decay per correction, leaving ancilla- T_1 -induced logical error rates a constant. However, if we assume fixed κ_r and go deeper into the overdamped regime as cavity lifetime improves (thus accepting a steeper reduction of correction rate than the reduction of FWM pumping amplitude), slower correction reduces the

excitation probability of the ancilla during the correction, yielding the same benefit as in the case of active QEC.

The third ancilla error is dephasing, with timescale T_ϕ . The parity mapping in the active scheme needs to be fast to avoid dephasing errors that can lead to incorrect mapping followed by an incorrect correction unitary. In the passive scheme, the decoherence in the $|g\rangle - |f\rangle$ subspace merely slows the first step of correction without causing logical error. To a good approximation, the ancilla frequency changes by approximately Γ_ϕ over time, acting as an effective detuning. In the passive case, this results in a slower correction rate by a factor of about $[1 - (\Gamma_\phi^2/\Omega_2^2)]$, which will increase the double-photon-loss rate by about $\bar{n}^2(\kappa_a^2/\kappa_{\text{cor}})(\Gamma_\phi^2/\Omega_2^2)$ in the limit where $\Gamma_\phi \ll \Omega_2$. For most reasonably long T_ϕ values ($\gtrsim 10 \mu\text{s}$), this will be a negligible effect. In the active scheme, ancilla dephasing errors will occur during the error mapping period ($t_{\text{map}} \approx \pi/\chi_{ge}$), resulting in a flipped ancilla state and thus an incorrect assignment, resulting in an error rate of approximately $[(\pi\bar{n}\kappa_a)/(\chi_{ge}T_\phi)]$. This is a relatively small error for reasonably long T_ϕ values. It should be noted that this error can be mitigated by performing multiple syndrome measurements and majority voting. However, since the error rate is small, the error overhead incurred by three or more measurements is larger than the mitigated error from this dephasing effect in most cases, so it is usually not practical. Overall, both schemes suffer little from ancilla dephasing errors, but the passive scheme has a built-in advantage.

4. Imperfect unitary recovery

Both the active and passive schemes have imperfections in the unitary operation performed for recovery. For the passive case, this imperfection stems from the jump and no-jump evolution terms in the Lindblad master equation. The stochastic nature of these terms will cause a superposition of $|0_L\rangle = (1/\sqrt{2})(|1\rangle + |5\rangle)$ to lose coherence and trend toward a mixed state even with a perfect correction operator. This is because the no-jump and jump evolutions drive the state into different components of $|1\rangle + |5\rangle$ and $|1\rangle - |5\rangle$ in a stochastic way, while they act differently on the other code state, $|1_L\rangle = |3\rangle$, since there is no complement state. This causes a loss of phase information, as one code word is driven to a complementary state, while the other is not. It should be noted that the logical phase information is stored in the relative phase between $|0_L\rangle$ and $|1_L\rangle$, not within the code word of $|0_L\rangle$, making the logical dephasing not as severe as one may naively expect. Nonetheless, some logical information is lost. The loss in fidelity from this code-state distortion effect in our experiment is approximately 0.5% per QEC cycle (τ_{cor}) based on numerical calculations.

In the active correction scheme, as in Ref. [24], since the binomial state is fully corrected with an applied unitary

accounting for jump and no-jump evolution after a correction cycle, errors incurred from the correction cycle are due to the correction infidelity. There are two correction cases, correcting the no-jump evolution from no error or correcting the jump error depending on the measurement outcome. Each one of these has an infidelity associated with it, ϵ_{nj} and ϵ_j , respectively. We use the values of $\epsilon_{\text{nj}} = 1.4\%$ and $\epsilon_j = 3.9\%$ from Ref. [24] to calculate the error rate in Table III.

5. Error rate comparison conclusion

The error rates tabulated in Table III show a performance advantage for the passive approach for the $\bar{n} = 3$ case. In the active case, the QEC performance can be improved by using the smallest binomial encoding (the $\bar{n} = 2$ kitten code) and reducing the cadence of QEC cycles, which then yields overall performance very similar to the passive QEC scheme demonstrated in our experiment.

-
- [1] N. Timoney, I. Baumgart, M. Johanning, A. Varón, A. Retzker, and C. Wunderlich, *Quantum gates and memory using microwave-dressed states*, *Nature (London)* **476**, 185 (2011).
 - [2] X. Xu, Z. Wang, C. Duan, P. Huang, P. Wang, Y. Wang, N. Xu, X. Kong, F. Shi, X. Rong, and J. Du, *Coherence-protected quantum gate by continuous dynamical decoupling in diamond*, *Phys. Rev. Lett.* **109**, 070502 (2012).
 - [3] Z. Leghtas, S. Touzard, I. M. Pop, A. Kou, B. Vlastakis, A. Petrenko, K. M. Sliwa, A. Narla, S. Shankar, M. J. Hatridge, M. Reagor, L. Frunzio, R. J. Schoelkopf, M. Mirrahimi, and M. H. Devoret, *Confining the state of light to a quantum manifold by engineered two-photon loss*, *Science* **347**, 853 (2015).
 - [4] A. Grimm, N. E. Frattini, S. Puri, S. O. Mundhada, S. Touzard, M. Mirrahimi, S. M. Girvin, S. Shankar, and M. H. Devoret, *Stabilization and operation of a Kerr-cat qubit*, *Nature (London)* **584**, 205 (2020).
 - [5] A. Laucht, R. Kalra, S. Simmons, J. P. Dehollain, J. T. Muhonen, F. A. Mohiyaddin, S. Freer, F. E. Hudson, K. M. Itoh, D. N. Jamieson *et al.*, *A dressed spin qubit in silicon*, *Nat. Nanotechnol.* **12**, 61 (2017).
 - [6] E. Paladino, Y. M. Galperin, G. Falci, and B. L. Altshuler, *1/f noise: Implications for solid-state quantum information*, *Rev. Mod. Phys.* **86**, 361 (2014).
 - [7] L. Viola and E. Knill, *Robust dynamical decoupling of quantum systems with bounded controls*, *Phys. Rev. Lett.* **90**, 037901 (2003).
 - [8] F. Yan, S. Gustavsson, J. Bylander, X. Jin, F. Yoshihara, D. G. Cory, Y. Nakamura, T. P. Orlando, and W. D. Oliver, *Rotating-frame relaxation as a noise spectrum analyser of a superconducting qubit under drive*, *Nat. Commun.* **4**, 2337 (2013).
 - [9] Z. Huang, P. S. Mundhada, A. Gienis, D. I. Schuster, A. A. Houck, and J. Koch, *Engineering dynamical sweet spots to protect superconducting qubits from 1/f noise*, *Phys. Rev. Appl.* **15**, 034065 (2021).

- [10] L. B. Nguyen, Y. Kim, A. Hashim, N. Goss, B. Marinelli, B. Bhandari, D. Das, R. K. Naik, J. M. Kreikebaum, A. N. Jordan *et al.*, *Programmable Heisenberg interactions between Floquet qubits*, *Nat. Phys.* **20**, 240 (2024).
- [11] A. Hajr, B. Qing, K. Wang, G. Koolstra, Z. Pedramrazi, Z. Kang, L. Chen, L. B. Nguyen, C. Jünger, N. Goss *et al.*, *High-coherence Kerr-Cat qubit in 2D architecture*, *Phys. Rev. X* **14**, 041049 (2024).
- [12] U. Réglade *et al.*, *Quantum control of a cat qubit with bit-flip times exceeding ten seconds*, *Nature (London)* **629**, 778 (2024).
- [13] N. E. Frattini, R. G. Cortiñas, J. Venkatraman, X. Xiao, Q. Su, C. U. Lei, B. J. Chapman, V. R. Joshi, S. M. Girvin, R. J. Schoelkopf, S. Puri, and M. H. Devoret, *Observation of pairwise level degeneracies and the quantum regime of the Arrhenius law in a double-well parametric oscillator*, *Phys. Rev. X* **14**, 031040 (2024).
- [14] Y. Li, Q. Mei, Q.-X. Jie, W. Cai, Y. Li, Z. Liu, Z.-J. Chen, Z. Xie, X. Cheng, X. Zhao *et al.*, *Beating the break-even point with autonomous quantum error correction*, [arXiv:2504.16746](#).
- [15] K. DeBry, N. Meister, A. V. Martinez, C. D. Bruzewicz, X. Shi, D. Reens, R. McConnell, I. L. Chuang, and J. Chiaverini, *Error correction of a logical qubit encoded in a single atomic ion*, [arXiv:2503.13908](#).
- [16] R. Acharya, D. A. Abanin, L. Aghababaie-Beni, I. Aleiner, T. I. Andersen, M. Ansmann, F. Arute, K. Arya, A. Asfaw, N. Astrakhantsev *et al.*, *Quantum error correction below the surface code threshold*, *Nature (London)* **638**, 920 (2024).
- [17] S. Krinner, N. Lacroix, A. Remm, A. Di Paolo, E. Genois, C. Leroux, C. Hellings, S. Lazar, F. Swiadek, J. Herrmann *et al.*, *Realizing repeated quantum error correction in a distance-three surface code*, *Nature (London)* **605**, 669 (2022).
- [18] H. Putterman, K. Noh, C. T. Hann, G. S. MacCabe, S. Aghaieibodi, R. N. Patel, M. Lee, W. M. Jones, H. Moradinejad, R. Rodriguez *et al.*, *Hardware-efficient quantum error correction via concatenated bosonic qubits*, *Nature (London)* **638**, 927 (2025).
- [19] P. Campagne-Ibarcq, A. Eickbusch, S. Touzard, E. Zalys-Geller, N. E. Frattini, V. V. Sivak, P. Reinhold, S. Puri, S. Shankar, R. J. Schoelkopf *et al.*, *Quantum error correction of a qubit encoded in grid states of an oscillator*, *Nature (London)* **584**, 368 (2020).
- [20] L. Hu, Y. Ma, W. Cai, X. Mu, Y. Xu, W. Wang, Y. Wu, H. Wang, Y. Song, C.-L. Zou *et al.*, *Quantum error correction and universal gate set operation on a binomial bosonic logical qubit*, *Nat. Phys.* **15**, 503 (2019).
- [21] N. Ofek, A. Petrenko, R. Heeres, P. Reinhold, Z. Leghtas, B. Vlastakis, Y. Liu, L. Frunzio, S. M. Girvin, L. Jiang *et al.*, *Extending the lifetime of a quantum bit with error correction in superconducting circuits*, *Nature (London)* **536**, 441 (2016).
- [22] V. V. Sivak, A. Eickbusch, B. Royer, S. Singh, I. Tsioutsios, S. Ganjam, A. Miano, B. Brock, A. Ding, L. Frunzio *et al.*, *Real-time quantum error correction beyond break-even*, *Nature (London)* **616**, 50 (2023).
- [23] B. L. Brock, S. Singh, A. Eickbusch, V. V. Sivak, A. Z. Ding, L. Frunzio, S. M. Girvin, and M. H. Devoret, *Quantum error correction of qubits beyond break-even*, *Nature (London)* **641**, 612 (2025).
- [24] Z. Ni, S. Li, X. Deng, Y. Cai, L. Zhang, W. Wang, Z.-B. Yang, H. Yu, F. Yan, S. Liu *et al.*, *Beating the break-even point with a discrete-variable-encoded logical qubit*, *Nature (London)* **616**, 56 (2023).
- [25] O. Oreshkov, in *Quantum Error Correction*, edited by D. A. Lidar and T. A. Brun (Cambridge University Press, Cambridge, England, 2013) Chap. 8, pp. 237–264.
- [26] C. Ahn, A. C. Doherty, and A. J. Landahl, *Continuous quantum error correction via quantum feedback control*, *Phys. Rev. A* **65**, 042301 (2002).
- [27] J. Cohen and M. Mirrahimi, *Dissipation-induced continuous quantum error correction for superconducting circuits*, *Phys. Rev. A* **90**, 062344 (2014).
- [28] V. V. Albert, S. O. Mundhada, A. Grimm, S. Touzard, M. H. Devoret, and L. Jiang, *Pair-cat codes: Autonomous error-correction with low-order nonlinearity*, *Quantum Sci. Technol.* **4**, 035007 (2019).
- [29] Q. Xu, G. Zheng, Y.-X. Wang, P. Zoller, A. A. Clerk, and L. Jiang, *Autonomous quantum error correction and fault-tolerant quantum computation with squeezed cat qubits*, *npj Quantum Inf.* **9**, 78 (2023).
- [30] E. Kapit, *Hardware-efficient and fully autonomous quantum error correction in superconducting circuits*, *Phys. Rev. Lett.* **116**, 150501 (2016).
- [31] J. M. Gertler, B. Baker, J. Li, S. Shirol, J. Koch, and C. Wang, *Protecting a bosonic qubit with autonomous quantum error correction*, *Nature (London)* **590**, 243 (2021).
- [32] Z. Li, T. Roy, D. Rodríguez Pérez, K.-H. Lee, E. Kapit, and D. I. Schuster, *Autonomous error correction of a single logical qubit using two transmons*, *Nat. Commun.* **15**, 1681 (2024).
- [33] W. P. Livingston, M. S. Blok, E. Flurin, J. Dressel, A. N. Jordan, and I. Siddiqi, *Experimental demonstration of continuous quantum error correction*, *Nat. Commun.* **13**, 2307 (2022).
- [34] D. Lachance-Quirion, M.-A. Lemonde, J. O. Simoneau, L. St-Jean, P. Lemieux, S. Turcotte, W. Wright, A. Lacroix, J. Fréchette-Viens, R. Shillito *et al.*, *Autonomous quantum error correction of Gottesman-Kitaev-Preskill states*, *Phys. Rev. Lett.* **132**, 150607 (2024).
- [35] B. De Neeve, T.-L. Nguyen, T. Behrle, and J. P. Home, *Error correction of a logical grid state qubit by dissipative pumping*, *Nat. Phys.* **18**, 296 (2022).
- [36] L. Sun, Y. Xu, Y. Zhou, Z. Hua, W. Wang, J. Zhou, Z. Jie Chen, L. Z. de Paula, Q.-X. Jie, G. Xue, H. Yu, W. Cai, C.-L. Zou, and L. Sun, *Extending coherence time beyond break-even point using only drives and dissipation*, [arXiv:2509.22191](#).
- [37] Z. Ni, L. Hu, Y. Cai, L. Zhang, J. Mai, X. Deng, P. Zheng, S. Liu, S.-B. Zheng, Y. Xu, and D. Yu, *Autonomous quantum error correction beyond break-even and its metrological application*, [arXiv:2509.26042](#).
- [38] A. L. Grimsmo, J. Combes, and B. Q. Baragiola, *Quantum computing with rotation-symmetric bosonic codes*, *Phys. Rev. X* **10**, 011058 (2020).
- [39] M. Mirrahimi, Z. Leghtas, V. V. Albert, S. Touzard, R. J. Schoelkopf, L. Jiang, and M. H. Devoret, *Dynamically*

- protected cat-qubits: A new paradigm for universal quantum computation*, *New J. Phys.* **16**, 045014 (2014).
- [40] A. Vanselow, B. Beaupre, L. Lattier, M. Villiers, A. Denis, P. Morfin, Z. Leghtas, and P. Campagne-Ibarcq, *Dissipating quartets of excitations in a superconducting circuit*, [arXiv:2501.05960](https://arxiv.org/abs/2501.05960).
- [41] C. Axline, M. Reagor, R. Heeres, P. Reinhold, C. Wang, K. Shain, W. Pfaff, Y. Chu, L. Frunzio, and R. J. Schoelkopf, *An architecture for integrating planar and 3D cQED devices*, *Appl. Phys. Lett.* **109**, 042601 (2016).
- [42] M. Kudra, M. Jirlow, M. Kervinen, A. M. Eriksson, F. Quijandria, P. Delsing, T. Abad, and S. Gasparinetti, *Experimental realization of deterministic and selective photon addition in a bosonic mode assisted by an ancillary qubit*, *Quantum Sci. Technol.* **10**, 045037 (2025).
- [43] J. Huang, T. J. DiNapoli, G. Rockwood, M. Yuan, P. Narasimhan, E. Gupta, M. Bal, F. Crisa, S. Garattoni, Y. Lu *et al.*, *Fast sideband control of a weakly coupled multimode bosonic memory*, [arXiv:2503.10623](https://arxiv.org/abs/2503.10623).
- [44] R. W. Heeres, P. Reinhold, N. Ofek, L. Frunzio, L. Jiang, M. H. Devoret, and R. J. Schoelkopf, *Implementing a universal gate set on a logical qubit encoded in an oscillator*, *Nat. Commun.* **8**, 94 (2017).
- [45] W. Dai, S. Hazra, D. K. Weiss, P. D. Kurilovich, T. Connolly, H. K. Babla, S. Singh, V. R. Joshi, A. Z. Ding, P. D. Parakh, J. Venkatraman, X. Xiao, L. Frunzio, and M. H. Devoret, *Characterization of drive-induced unwanted state transitions in superconducting circuits*, *Phys. Rev. X* **16**, 011011 (2026).
- [46] M. Xia, C. Lledó, M. Capocci, J. Repicky, B. D'Anjou, I. Mondragon-Shem, R. Kaufman, J. Koch, A. Blais, and M. Hatridge, *Exceeding the parametric drive strength threshold in nonlinear circuits*, [arXiv:2506.03456](https://arxiv.org/abs/2506.03456).
- [47] M. F. Dumas, B. Groleau-Paré, A. McDonald, M. H. Muñoz Arias, C. Lledó, B. D'Anjou, and A. Blais, *Measurement-induced transmon ionization*, *Phys. Rev. X* **14**, 041023 (2024).
- [48] Y. Ma, Y. Xu, X. Mu, W. Cai, L. Hu, W. Wang, X. Pan, H. Wang, Y. P. Song, C.-L. Zou, and L. Sun, *Error-transparent operations on a logical qubit protected by quantum error correction*, *Nat. Phys.* **16**, 827 (2020).
- [49] P. Reinhold, S. Rosenblum, W.-L. Ma, L. Frunzio, L. Jiang, and R. J. Schoelkopf, *Error-corrected gates on an encoded qubit*, *Nat. Phys.* **16**, 822 (2020).
- [50] B. J. Brown, D. Loss, J. K. Pachos, C. N. Self, and J. R. Wootton, *Quantum memories at finite temperature*, *Rev. Mod. Phys.* **88**, 045005 (2016).
- [51] S. Shirol, [10.5281/zenodo.19190491](https://zenodo.org/record/19190491) (2026).
- [52] T. Connolly, P. D. Kurilovich, S. Diamond, H. Nho, Charlotte G. L. Böttcher, L. I. Glazman, V. Fatemi, and M. H. Devoret, *Coexistence of nonequilibrium density and equilibrium energy distribution of quasiparticles in a superconducting qubit*, *Phys. Rev. Lett.* **132**, 217001 (2024).
- [53] Q-CTRL, Boulder Opal (2025), <https://q-ctrl.com/boulder-opal>.
- [54] X. Y. Jin, A. Kamal, A. P. Sears, T. Gudmundsen, D. Hover, J. Miloshi, R. Slattery, F. Yan, J. Yoder, T. P. Orlando, S. Gustavsson, and W. D. Oliver, *Thermal and residual excited-state population in a 3D transmon qubit*, *Phys. Rev. Lett.* **114**, 240501 (2015).
- [55] 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).
- [56] F. Reiter and A. S. Sørensen, *Effective operator formalism for open quantum systems*, *Phys. Rev. A* **85**, 032111 (2012).
- [57] J. M. Gertler, S. van Geldern, S. Shirol, L. Jiang, and C. Wang, *Experimental realization and characterization of stabilized pair-coherent states*, *PRX Quantum* **4**, 020319 (2023).
- [58] F. Valadares, N.-N. Huang, K. T. N. Chu, A. Dorogov, W. Chua, L. Kong, P. Song, and Y. Y. Gao, *On-demand transposition across light-matter interaction regimes in bosonic cQED*, *Nat. Commun.* **15**, 5816 (2024).
- [59] S. Rosenblum, P. Reinhold, M. Mirrahimi, L. Jiang, L. Frunzio, and R. Schoelkopf, *A fault-tolerant detection of a quantum error*, *Science* **361**, 266 (2018).