Erasure Detection of a Dual-Rail Qubit Encoded in a Double-Post Superconducting Cavity

Akshay Koottandavida[*,†], Ioannis Tsioutsios[*,‡], Aikaterini Kargioti, Cassady R. Smith, Vidul R. Joshi, Wei Dai, James D. Teoh, Jacob C. Curtis, Luigi Frunzio, Robert J. Schoelkopf, and Michel H. Devoret[§]

Department of Applied Physics, Yale University, New Haven, Connecticut 06520, USA
and Yale Quantum Institute, Yale University, New Haven, Connecticut 06511, USA

Qubits with predominantly erasure errors present distinctive advantages for quantum error correction (QEC) and fault-tolerant quantum computing. Logical qubits based on dual-rail encoding that exploit erasure detection have been recently proposed in superconducting circuit architectures, with either coupled transmons or cavities. Here, we implement a dual-rail qubit encoded in a compact, double-post superconducting cavity. Using an auxiliary transmon, we perform erasure detection on the dual-rail subspace. We characterize the behavior of the code space by a novel method to perform joint-Wigner tomography. This is based on modifying the cross-Kerr interaction between the cavity modes and the transmon. We measure an erasure rate of 3.981 ± 0.003 $(\text{ms})^{-1}$ and a residual, postselected dephasing error rate up to 0.17 $(\text{ms})^{-1}$ within the code space. This strong hierarchy of error rates, together with the compact and hardware-efficient nature of this novel architecture, holds promise in realizing QEC schemes with enhanced thresholds and improved scaling.

Introduction.—Quantum error correction (QEC), the process of protecting quantum information from decoherence, is an essential ingredient toward fault-tolerant quantum computation (FTQC). QEC involves redundantly encoding logical qubits into an enlarged Hilbert space, targeting coherence that significantly exceeds that of its constituent components. QEC codes can be roughly categorized into discrete variable codes [1], such as the surface code [2–4], or continuous variable codes, such as the Gottesman-Kitaev-Preskill [5], binomial [6], or cat codes [7,8]. Experimentally, bosonic QEC codes have proven to be efficient in reducing error rates on the single logical-qubit level [9–11], while discrete variable codes have demonstrated suppression of logical error rates by increasing code distance [12]. Hence, combining the two approaches where error-corrected bosonic qubits form the base layer of surface code architectures may be a promising pathway toward FTQC.

Alternatively, recent studies have shown that systems with an erasure noise channel at the base layer can reduce the logical error rate by substantially increasing both the threshold and distance of the outer surface code [13,14]. Erasures are leakage errors to outside the computational subspace that are detected in real time and for which the physical qubits that were impacted are also located in space. Qubits predominantly subject to erasures, so-called "erasure qubits," have been proposed with metastable states of neutral atoms [14–16] and dual-rail qubits based on superconducting circuit quantum electrodynamics (cQED) architecture [17,18]. While the dual-rail encoding has been

a subject of extensive study in the quantum optics platform [19] and has been investigated in superconducting circuits [20–22], it has only recently been implemented in cQED systems in the context of erasure errors [23,24].

A successful incorporation of erasure qubits into a QEC architecture requires a system that exhibits strong hierarchy of errors [14]. This means that the dominant errors are converted to erasures and the remaining Pauli and leakage errors exhibit orders of magnitude lower rates. To this end, superconducting cavity modes are ideal candidates to encode a dual-rail erasure qubit, since they present natural noise bias, with photon loss being the dominant error mechanism [18,24]. In addition, superconducting cavities, especially those implemented in 3D geometries, exhibit longer lifetimes with lower intrinsic dephasing rates compared to transmons. Nonetheless, the nonlinearity of auxiliary transmons is still a necessary ingredient for the control of cavity modes. As a result, dual-rail qubits with cavity modes suffer from additional loss channels introduced by these nonlinear elements.

In this Letter, we present a dual-rail qubit implemented in a hardware-efficient, 3D cavity architecture—the symmetric and antisymmetric eigenmodes of a double-post coaxial superconducting aluminum cavity. The highly delocalized field distributions of the modes allow for a compact architecture in which a single, dispersively coupled auxiliary transmon provides the necessary nonlinearity for state preparation, erasure detection, and tomography. Hence, compared to other approaches [24], our architecture requires fewer resources. By modifying the dispersive
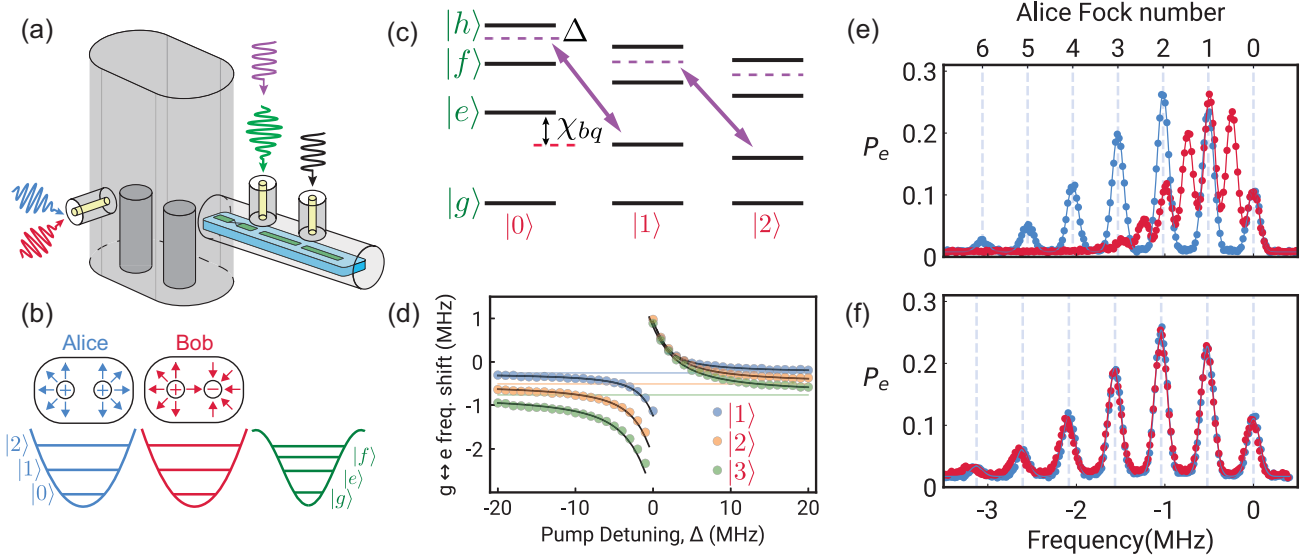
FIG. 1.   (a) Schematic of the double-post cavity, made of 5N aluminum with a transmon qubit, readout, and Purcell filter fabricated on a sapphire chip. (b) Field distributions of the symmetric (Alice) and antisymmetric (Bob) eigenmodes of the system, encoding the dual-rail qubit. (c) Energy-level diagram of the combined Bob-transmon system showing dispersive shifts. The purple arrows connect the pairs of levels $|n-1,h\rangle$ and $|n,e\rangle$ being coupled via the cross-Kerr tuning drive. (d) Avoided crossing observed when preparing $|n,e\rangle$ states in Bob-transmon, sweeping pump detuning Δ with fixed amplitude $\Omega/2\pi = 0.5$ MHz. No drive is applied to affect Alice-transmon coupling. Solid horizontal lines are a visual aid to the bare Fock state energies, and simulation results are overlaid in black lines. (e) Number-split peaks of the transmon when either Alice (blue) or Bob (red) is populated with a coherent state of amplitude $\alpha = 1.5$, without any pump. (f) Bob's peaks align with that of Alice in the presence of the cross-Kerr tuning pump with parameters $(\Omega/2\pi, \Delta/2\pi) = (0.5, -5.4)$ MHz.

interaction between the cavity modes and the transmon [25], we perform joint-Wigner tomography on the two modes to characterize the erasure detection circuit. We show that our quantum nondemolition (QND) erasure detection scheme converts cavity photon losses to erasures with a false-negative probability of only 0.28% per gate. We demonstrate a strong hierarchy of error rates with erasures occurring at a rate of 3.981 ± 0.003 (ms)$^{-1}$ and postselected Pauli $\hat{Z}$ errors at a rate of up to 0.170 (ms)$^{-1}$. Moreover, postselected bit-flip or Pauli $\hat{X}$ type errors occur at a negligible rate of $\sim 10^{-4}$ (ms)$^{-1}$. These results demonstrate the viability of incorporating such an erasure detection scheme in a circuit (so-called midcircuit erasure detection) and, consequently, use the dual-rail as an erasure qubit in concatenation codes to enhance QEC thresholds.

Experimental system.—Figure 1(a) depicts our experiment, which comprises a coaxial superconducting cavity made of high-purity (5N) aluminum [26] with two posts of equal length. The package hosts two modes, Alice ($\hat{a}$) and Bob ($\hat{b}$), which are the harmonic oscillators used to encode the dual-rail qubit. An auxiliary transmon fabricated on a sapphire chip is inserted into the package [27–30] (see Supplemental Material [31] for full system parameters). The delocalized electromagnetic field distribution of Alice and Bob [Fig. 1(b)] creates coupling to the transmon with similar strengths. This leads to a static dispersive interaction:

$$\hat{H}/\hbar = \chi_{aq}\hat{a}^\dagger\hat{a}|e\rangle\langle e| + \chi_{bq}\hat{b}^\dagger\hat{b}|e\rangle\langle e| \qquad (1)$$

with measured cross-Kerr rates $\chi_{aq}/2\pi = -0.514$ MHz and $\chi_{bq}/2\pi = -0.251$ MHz. This coupling enables our erasure detection scheme. However, the residual mismatch in cross-Kerr rates renders tomography in the combined Hilbert space a rather challenging task [32,33].

Cross-Kerr matching.—The motivation behind matching the cross-Kerr rates in Eq. (1) is to perform joint-Wigner tomography on the combined Hilbert space of Alice and Bob. The joint-Wigner function [34] given by

$$W(\alpha,\beta) = \frac{4}{\pi^2}\mathrm{Tr}[\hat{D}(-\alpha,-\beta)\rho\hat{D}(\alpha,\beta)\hat{\Pi}] \qquad (2)$$

requires the measuring the expectation value of the joint-parity operator:

$$\hat{\Pi} = (-1)^{\hat{a}^\dagger\hat{a}+\hat{b}^\dagger\hat{b}} \qquad (3)$$

of the displaced state. If $\chi_{aq} = \chi_{bq} = \chi$, this measurement is greatly simplified via a Ramsey-like sequence [Fig. 2(b)], similar to the single-mode Wigner tomography. By choosing the wait time such that $t_W = \pi/\chi$, the joint parity is mapped on to the transmon states (see Supplemental Material [31]). To match the cross-Kerr rates, we leverage the four-wave mixing property of the transmon [25] and
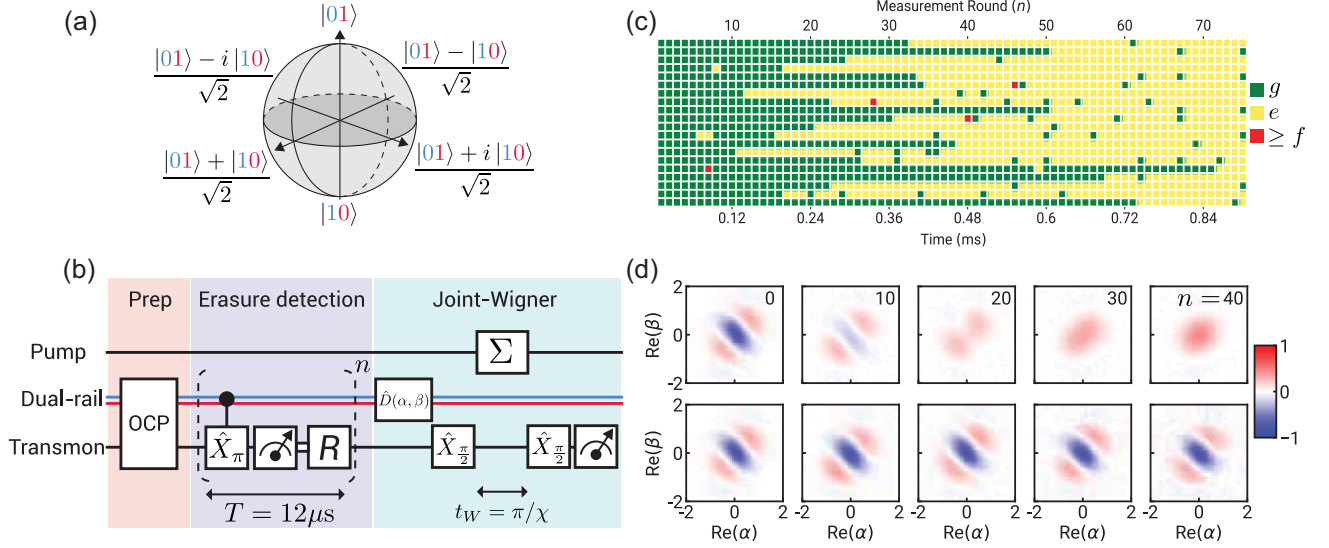
FIG. 2. (a) Dual-rail Bloch sphere where the code words are two-mode Fock states $|+Z\rangle_L = |01\rangle$ and $|-Z\rangle_L = |10\rangle$ (the first mode is Alice, and the second mode is Bob). (b) Erasure detection using a selective π pulse centered at the transmon frequency ($\hat{X}_\pi^0$). After n detection rounds, we perform joint-Wigner tomography on the system with the cross-Kerr matched Hamiltonian $\hat{H}_\Sigma$ represented as the Σ gate. (c) Experimental results of erasure detection for $n = 75$ rounds after preparing the $|+X\rangle_L$ state. Each row is an separate run of the circuit. Ideally, state transition from $|g\rangle \to |e\rangle$ heralds an erasure for the dual-rail qubit. But readout infidelities and transmon errors cause deviation from this behavior in the form of isolated "e" or "f" outcomes. (d) Experimental result of the full circuit shown in (b). Erasure detection is performed on the $|+X\rangle_L$ state for n rounds before measuring the $\mathrm{Re}(\alpha)$-$\mathrm{Re}(\beta)$ joint-Wigner cut of the state. The experiment is repeated for $n = 0, 10, 20, 30, 40$ rounds. The initial state clearly decays to the ground state due to cavity photon loss. Postselecting on a qubit being in "g" in the trajectory data, we can reconstruct the encoded information, visually proving faithful conversion of photon loss to erasures.

apply a microwave drive at frequency $\omega_\chi = \omega_{he} - \omega_b + \Delta$, with amplitude Ω. This process exchanges two photons from the transmon with one photon each from the drive and one of the cavities, coupling pairs of levels $|n, e\rangle \leftrightarrow |n-1, h\rangle$. Here, $|n\rangle$ denotes the nth Fock state of the relevant cavity mode, and $|e\rangle$ and $|h\rangle$ represent the first and third excited states, respectively, of the transmon. By detuning the pump by Δ, we shift the energy levels of the Fock states from their bare values. In the $\chi \ll \Omega \ll \Delta$ regime, this can be approximated as a change in the cross-Kerr rate between the modes [Fig. 1(c)].

For the purposes of this Letter, we choose to tune only χ_{bq} and keep χ_{aq} constant. We opted to match to the higher cross-Kerr rate of Alice in order to achieve faster gate times. This independent control of the cross-Kerr rates is possible due to the large detuning (≈ 200 MHz) between the cavity modes. The avoided crossings in Fig. 1(d) reveal the tuning of the energy levels of the Bob-transmon system due to the pump. We then set pump parameters (Ω, Δ) such that the cross-Kerr rates are matched. To confirm cross-Kerr matching, we perform spectroscopic measurements on the transmon after preparing a coherent state in either cavity mode. Figures 1(e) and 1(f) show the effect of the pump on the number split peaks [35] of the cavities, where Bob's peaks align with those of Alice. Hence, we can approximate the interaction Hamiltonian as

$$\hat{H}_\Sigma/\hbar \approx \chi(\hat{a}^\dagger \hat{a} + \hat{b}^\dagger \hat{b})|e\rangle\langle e| \tag{4}$$

up to six Fock states. The extracted pumped cross-Kerr rates were $\chi_{aq}/2\pi = -0.521$ MHz and $\chi_{bq}/2\pi = -0.527$ MHz. We note that by increasing the pump amplitude Ω we can potentially tune higher Fock states as well.

Erasure detection.—The dual-rail qubit is defined in the joint Hilbert space of Alice and Bob with $|+Z_L\rangle = |01\rangle$ and $|-Z_L\rangle = |10\rangle$ as the logical code words [Fig. 2(a)]. The dominant error channel in superconducting 3D cavity modes is single-photon loss. Photon losses in either Alice or Bob destroy the logical dual-rail encoding and leave the system in the error state $|00\rangle$. In our system, these errors occur at rates $\kappa_a = 4.454 \pm 0.044$ (ms)$^{-1}$ ($T_1^a = 224.5 \pm 2.2$ μs) and $\kappa_b = 3.339 \pm 0.018$ (ms)$^{-1}$ ($T_1^b = 299.4 \pm 1.6$ μs) for Alice and Bob, respectively. To detect photon losses, we leverage the dispersive interaction between the cavity modes and the transmon [Eq. (1)]. After initializing the system in the code space, we query the transmon state using a frequency selective pulse centered at its bare frequency (ω_{ge}). The transmon state flips only when the cavity modes are in $|00\rangle$ (error space). Our circuit to realize this is shown in Fig. 2(b). We prepare the states in the code space via optimal control pulses (OCP) [36–38], and the transmon is initialized in its

ground state ($|g\rangle$). The selective pulse to flip the transmon state is a long ($4\sigma = 8$ μs) Gaussian pulse ($\hat{X}_\pi$). This is followed by a readout and a fast reset of the transmon state to $|g\rangle$. The entire circuit takes exactly 12 μs to execute and should output readout result e if the cavity modes are in the error space and g otherwise.

Figure 2(c) displays the results of performing erasure detection for $n = 75$ rounds on the $|+X_L\rangle = (1/\sqrt{2})(|01\rangle + |10\rangle)$ state. Each row in the plot represents a different experimental shot. The first row depicts a near-ideal trajectory where the transmon state remains in $|g\rangle$ until measurement round 32. Because of a photon loss event in either cavity, the transmon frequency shifts to its bare value, causing the $\hat{X}_\pi$ gate to flip its state to $|e\rangle$. The transmon remains in this state for the remainder of the time due to the reset operation. Nonidealities arising from readout inefficiencies and transmon errors, however, cause many trajectories to deviate from the ideal behavior. To characterize the erasure detection circuit under these errors, we feed the raw trajectories to a hidden Markov model. The model learns a state transition matrix, describing the probability of transitions between code space and error space at each time step, and an emission matrix, which predicts the probability of an outcome (g/e) given a hidden state. From this model, we extract a false-negative probability, defined as the probability of misassigning an erasure as being within the code space, of 0.28% per measurement. Additionally, we extract a $> 99.9\%$ measure of QND on the cavity photon number, for our detection scheme. This means that the erasure detection itself induces minimal backaction compared to having no detection. This is a crucial feature for incorporating such erasure checks in a surface code. Finally, we note that each erasure check accrues a deterministic phase on superpositions of states in the dual-rail subspace. In the experiment, we calibrate this and apply software corrections to the tomography pulses to cancel it.

Joint tomography.—To determine the behavior of the code space during detection, we perform direct tomography on the cavities. Since the dual-rail code is defined in the joint Hilbert space of Alice and Bob, we measure the joint-Wigner function [32], using the circuit shown in Fig. 2(b). During the waiting time (t_W), we establish the cross-Kerr matched interaction Hamiltonian by applying the pump with the matching parameters, thereby measuring the expectation value of the joint-parity operator. Note that the joint-Wigner function is defined in a 4D space, since α and β are complex numbers, representing position and conjugate momentum variables of each mode. As a result, it is challenging to visualize and would require an exponential number of samples to characterize accurately.

Instead, to efficiently characterize the system, we perform partial tomography by sparsely sampling the values of (α, β). First, to visually verify the evolution of the states during erasure detection, we measure a 2D cut of the full

joint-Wigner space, specifically the Re(α)-Re(β) cut with Im(α) = Im(β) = 0. In the top row in Fig. 2(d), we observe the evolution of the $|+X_L\rangle$ state after $n = 0, 10, 20, 30, 40$ rounds of erasure detection. As expected, the state eventually decays to the error space $|00\rangle$. After discarding trajectories where erasures were detected, we are clearly able to recover the original information [Fig. 2(d), bottom panel]. This improvement comes at the expense of discarding more data as we track the system for longer times. Since we are measuring only a cut of the joint Wigner, we have only partial information about the system, and it is not enough to reconstruct the full density matrix.

Having visually confirmed that our erasure detection scheme enables faithful recovery of the encoded information, we proceed to quantify how good our detection scheme is. To achieve this, we measure the logical Pauli state vector components for the dual-rail subspace after preparing in the six cardinal states of the Bloch sphere. This measurement is performed using the same circuit as before [Fig. 2(b)]. By projecting the joint-Wigner function in Eq. (2) onto the dual-rail subspace, we extract the expectation values of the Pauli operators by sampling only 16 points (α_i, β_i) in phase space (see Supplemental Material [31]) [32,33].

Figure 3(a) illustrates the decay of the expectation values of the four Pauli operators, $(\hat{I}, \hat{X}, \hat{Y}, \hat{Z})$, for different states on the Bloch sphere, as a function of the number of detection rounds. The expectation values decay exponentially due to the single-photon loss channels. Crucially, we observe the decay of the identity operator ($\hat{I}$), indicating that the system decays outside the code space. Averaged over the six cardinal states on the Bloch sphere, we obtain an erasure rate of 3.981 ± 0.003 $(\text{ms})^{-1}$ or 4.8% per measurement.

Discarding the trajectories where an erasure was detected, we extract residual Pauli error rates, within the code space. The $\hat{I}$ and the $\hat{Z}$ operator expectation values are nearly constant, and we are able to provide an upper bound on their decay rates of only $\sim 10^{-4}$ $(\text{ms})^{-1}$. The postselected $\hat{X}$ and $\hat{Y}$ operators exhibit no-jump evolution, i.e., the conditional update of the density matrix of the system upon measuring no photon jumps [24]. For the dual-rail qubit, this effect causes any superposition of states to deterministically decay toward the cavity with lower decay rate (in our case, Bob). In principle, it is possible to exactly cancel this effect by designing modes with equal decay rates or by interleaving SWAP operations between the cavities such that the photon spends equal amounts of time in each mode [18]. Fitting this deterministic no-jump evolution, we extract residual dephasing rates up to 0.170 $(\text{ms})^{-1}$ ($\approx 0.2\%$ per measurement). Note that we are discarding exponentially many trajectories as we perform erasure detection for more rounds, as seen in the survival probability plot in Fig. 3(b), which shows the number of trajectories that survive as a function of detection rounds.
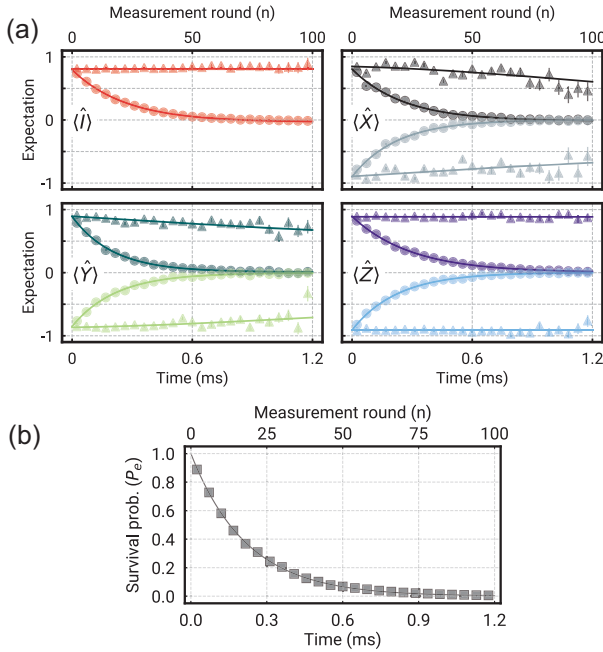
FIG. 3. (a) Expectation values of the logical Pauli operators $\hat{I}$, $\hat{X}$, $\hat{Y}$, and $\hat{Z}$ for the dual-rail subspace as a function of detection rounds, measured using the circuit in Fig. 2(b). Triangles indicate data after postselecting on no erasures from the dots. For $\langle\hat{X}\rangle$, the experiment is performed after preparing the $|+X\rangle$ (darker color at the top) and $|-X\rangle$ (lighter color at the bottom) states. The same is repeated for $\langle\hat{Y}\rangle$ and $\langle\hat{Z}\rangle$. Postselection preserves the expectation values for far longer. The fits (shown in solid lines) for the postselected data account for the no-jump evolution, due to the difference in lifetimes between Alice and Bob, and residual Pauli errors within the subspace. (b) Survival probability, i.e., the fraction of experimental shots that survive after n rounds.

Conclusions.—In this Letter, we demonstrated erasure detection for a dual-rail qubit implemented in a superconducting double-post coaxial cavity. As a result of the compact nature of this architecture, a single auxiliary transmon is sufficient for erasure detection and control of both cavity modes. We measure an average erasure rate of 3.981 ± 0.003 (ms)$^{-1}$, which corresponds to 4.8% per measurement, with a false-negative rate of 0.28%. In addition, we developed a protocol to perform joint-Wigner tomography which relies on matching the dispersive interaction rates between cavity modes and transmon. From 2D cuts of the joint-Wigner function space, we reconstruct the encoded information given the outcomes of the erasure detection and extract the residual Pauli errors within the dual-rail code space. We observe that dephasing type errors dominate at a rate up to 0.170 (ms)$^{-1}$, 0.2% per measurement, a result expected from the finite bit-flip rate of the transmon and the mismatch in the cross-Kerr rates during the erasure detection circuit. Finally, residual bit-flip and leakage errors are negligible with an upper bound of $\sim 10^{-4}$ (ms)$^{-1}$, highlighting a clear hierarchy of rates where erasure dominate over Pauli errors. These results, combined

with the high-fidelity beam splitters recently demonstrated [33,39], suggest a promising pathway toward concatenating superconducting cavity-based dual-rail qubits within a surface code and leverage the higher threshold and favorable scaling with code distance. Finally, it should be possible to further take advantage of the high noise bias [3,40–42], of the Pauli errors in the dual-rail subspace when designing the surface code architecture to further improve upon the advantage of erasure detection.

*These authors contributed equally to this work.
†akshay.koottandavida@yale.edu
‡ioannis.tsioutsios@yale.edu
§michel.devoret@yale.edu

[1] A. Steane, Multiple-particle interference and quantum error correction, Proc. R. Soc. A **452**, 2551 (1997).

[2] A. G. Fowler, M. Mariantoni, J. M. Martinis, and A. N. Cleland, Surface codes: Towards practical large-scale quantum computation, Phys. Rev. A **86**, 032324 (2012).

[3] J. P. Bonilla Ataides, D. K. Tuckett, S. D. Bartlett, S. T. Flammia, and B. J. Brown, The XZZX surface code, Nat. Commun. **12**, 2172 (2021).

[4] S. Krinner, N. Lacroix, A. Remm, A. Di Paolo, E. Genois, C. Leroux, C. Hellings, S. Lazar, F. Swiadek, J. Herrmann, G. J. Norris, C. K. Andersen, M. Müller, A. Blais, C. Eichler, and A. Wallraff, Realizing repeated quantum error correction in a distance-three surface code, Nature (London) **605**, 669 (2022).

[5] D. Gottesman, A. Kitaev, and J. Preskill, Encoding a qubit in an oscillator, Phys. Rev. A **64**, 012310 (2001).

[6] M. H. Michael, M. Silveri, R. T. Brierley, V. V. Albert, J. Salmilehto, L. Jiang, and S. M. Girvin, New class of

quantum error-correcting codes for a bosonic mode, Phys. Rev. X **6**, 031006 (2016).

[7] 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).

[8] 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).

[9] N. Ofek, A. Petrenko, R. Heeres, P. Reinhold, Z. Leghtas, B. Vlastakis, Y. Liu, L. Frunzio, S. M. Girvin, L. Jiang, M. Mirrahimi, M. H. Devoret, and R. J. Schoelkopf, Extending the lifetime of a quantum bit with error correction in superconducting circuits, Nature (London) **536**, 441 (2016).

[10] V. V. Sivak, A. Eickbusch, B. Royer, S. Singh, I. Tsioutsios, S. Ganjam, A. Miano, B. L. Brock, A. Z. Ding, L. Frunzio, S. M. Girvin, R. J. Schoelkopf, and M. H. Devoret, Real-time quantum error correction beyond break-even, Nature (London) **616**, 50 (2023).

[11] Z. Ni, S. Li, X. Deng, Y. Cai, L. Zhang, W. Wang, Z.-B. Yang, H. Yu, F. Yan, S. Liu, C.-L. Zou, L. Sun, S.-B. Zheng, Y. Xu, and D. Yu, Beating the break-even point with a discrete-variable-encoded logical qubit, Nature (London) **616**, 56 (2023).

[12] Google Quantum AI, Suppressing quantum errors by scaling a surface code logical qubit, Nature (London) **614**, 676 (2023).

[13] M. Grassl, T. Beth, and T. Pellizzari, Codes for the quantum erasure channel, Phys. Rev. A **56**, 33 (1997).

[14] Y. Wu, S. Kolkowitz, S. Puri, and J. D. Thompson, Erasure conversion for fault-tolerant quantum computing in alkaline earth Rydberg atom arrays, Nat. Commun. **13**, 4657 (2022).

[15] S. Ma, G. Liu, P. Peng, B. Zhang, S. Jandura, J. Claes, A. P. Burgers, G. Pupillo, S. Puri, and J. D. Thompson, High-fidelity gates and mid-circuit erasure conversion in an atomic qubit, Nature (London) **622**, 279 (2023).

[16] K. Sahay, J. Jin, J. Claes, J. D. Thompson, and S. Puri, High-threshold codes for neutral-atom qubits with biased erasure errors, Phys. Rev. X **13**, 041013 (2023).

[17] A. Kubica, A. Haim, Y. Vaknin, H. Levine, F. Brandão, and A. Retzker, Erasure qubits: Overcoming the ${T}{1}$ limit in superconducting circuits, Phys. Rev. X **13**, 041022 (2023).

[18] J. D. Teoh, P. Winkel, H. K. Babla, B. J. Chapman, J. Claes, S. J. de Graaf, J. W. O. Garmon, W. D. Kalfus, Y. Lu, A. Maiti, K. Sahay, N. Thakur, T. Tsunoda, S. H. Xue, L. Frunzio, S. M. Girvin, S. Puri, and R. J. Schoelkopf, Dual-rail encoding with superconducting cavities, Proc. Natl. Acad. Sci. U.S.A. **120**, e2221736120 (2023).

[19] I. L. Chuang and Y. Yamamoto, Quantum bit regeneration, Phys. Rev. Lett. **76**, 4281 (1996).

[20] E. Zakka-Bajjani, F. Nguyen, M. Lee, L. R. Vale, R. W. Simmonds, and J. Aumentado, Quantum superposition of a single microwave photon in two different 'colour' states, Nat. Phys. **7**, 599 (2011).

[21] Y.-P. Shim and C. Tahan, Semiconductor-inspired design principles for superconducting quantum computing, Nat. Commun. **7**, 11059 (2016).

[22] D. L. Campbell, Y.-P. Shim, B. Kannan, R. Winik, D. K. Kim, A. Melville, B. M. Niedzielski, J. L. Yoder, C. Tahan, S. Gustavsson, and W. D. Oliver, Universal nonadiabatic control of small-gap superconducting qubits, Phys. Rev. X **10**, 041051 (2020).

[23] H. Levine et al., Demonstrating a long-coherence dual-rail erasure qubit using tunable transmons, Phys. Rev. X **14**, 011051 (2024).

[24] K. S. Chou et al., Demonstrating a superconducting dual-rail cavity qubit with erasure-detected logical measurements, arXiv:2307.03169.

[25] S. Rosenblum, P. Reinhold, M. Mirrahimi, L. Jiang, L. Frunzio, and R. J. Schoelkopf, Fault-tolerant detection of a quantum error, Science **361**, 266 (2018).

[26] M. Reagor, H. Paik, G. Catelani, L. Sun, C. Axline, E. Holland, I. M. Pop, N. A. Masluk, T. Brecht, L. Frunzio, M. H. Devoret, L. Glazman, and R. J. Schoelkopf, Reaching 10 ms single photon lifetimes for superconducting aluminum cavities, Appl. Phys. Lett. **102**, 192604 (2013).

[27] J. Koch, T. M. Yu, J. Gambetta, A. A. Houck, D. I. Schuster, J. Majer, A. Blais, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, Charge-insensitive qubit design derived from the Cooper pair box, Phys. Rev. A **76**, 042319 (2007).

[28] E. M. Purcell, H. C. Torrey, and R. V. Pound, Resonance absorption by nuclear magnetic moments in a solid, Phys. Rev. **69**, 37 (1946).

[29] N. E. Frattini, U. Vool, S. Shankar, A. Narla, K. M. Sliwa, and M. H. Devoret, 3-wave mixing Josephson dipole element, Appl. Phys. Lett. **110**, 222603 (2017).

[30] V. V. Sivak, N. E. Frattini, V. R. Joshi, A. Lingenfelter, S. Shankar, and M. H. Devoret, Kerr-free three-wave mixing in superconducting quantum circuits, Phys. Rev. Appl. **11**, 054060 (2019).

[31] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.132.180601 for system parameters and details about tomography.

[32] C. Wang, Y. Y. Gao, P. Reinhold, R. W. Heeres, N. Ofek, K. Chou, C. Axline, M. Reagor, J. Blumoff, K. M. Sliwa, L. Frunzio, S. M. Girvin, L. Jiang, M. Mirrahimi, M. H. Devoret, and R. J. Schoelkopf, A Schrödinger cat living in two boxes, Science **352**, 1087 (2016).

[33] B. J. Chapman, S. J. de Graaf, S. H. Xue, Y. Zhang, J. Teoh, J. C. Curtis, T. Tsunoda, A. Eickbusch, A. P. Read, A. Koottandavida, S. O. Mundhada, L. Frunzio, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, High-on-off-ratio beam-splitter interaction for gates on bosonically encoded qubits, PRX Quantum **4**, 020355 (2023).

[34] S. Haroche and J.-M. Raimond, *Exploring the Quantum: Atoms, Cavities, and Photons* (Oxford University Press, New York, 2006).

[35] J. Gambetta, A. Blais, D. I. Schuster, A. Wallraff, L. Frunzio, J. Majer, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, Qubit-photon interactions in a cavity: Measurement-induced dephasing and number splitting, Phys. Rev. A **74**, 042318 (2006).

[36] N. Khaneja, T. Reiss, C. Kehlet, T. Schulte-Herbrüggen, and S. J. Glaser, Optimal control of coupled spin dynamics: Design of NMR pulse sequences by gradient ascent algorithms, J. Magn. Reson. **172**, 296 (2005).

[37] 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).

[38] P. Reinhold, Ph.D. thesis, Yale University, 2019.

[39] Y. Lu, A. Maiti, J. W. O. Garmon, S. Ganjam, Y. Zhang, J. Claes, L. Frunzio, S. M. Girvin, and R. J. Schoelkopf, High-fidelity parametric beamsplitting with a parity-protected converter, Nat. Commun. **14,** 5767 (2023).

[40] D. K. Tuckett, A. S. Darmawan, C. T. Chubb, S. Bravyi, S. D. Bartlett, and S. T. Flammia, Tailoring surface codes for highly biased noise, Phys. Rev. X **9,** 041031 (2019).

[41] D. K. Tuckett, S. D. Bartlett, and S. T. Flammia, Ultrahigh error threshold for surface codes with biased noise, Phys. Rev. Lett. **120,** 050505 (2018).

[42] S. Puri, L. St-Jean, J. A. Gross, A. Grimm, N. E. Frattini, P. S. Iyer, A. Krishna, S. Touzard, L. Jiang, A. Blais, S. T. Flammia, and S. M. Girvin, Bias-preserving gates with stabilized cat qubits, Sci. Adv. **6,** eaay5901 (2020).

Supplementary Information
"Erasure detection of a dual-rail qubit encoded in a double-post superconducting cavity"

Akshay Koottandavida,[1, *] Ioannis Tsioutsios,[1, *] Aikaterini Kargioti,[1] Cassady R. Smith,[1] Vidul R. Joshi,[1] Wei Dai,[1] James D. Teoh,[1] Jacob C. Curtis,[1] Luigi Frunzio,[1] Robert J. Schoelkopf,[1] and Michel H. Devoret[1, †]

[1]*Department of Applied Physics, Yale University, New Haven, CT 06520, USA*
Yale Quantum Institute, Yale University, New Haven, CT 06511, USA
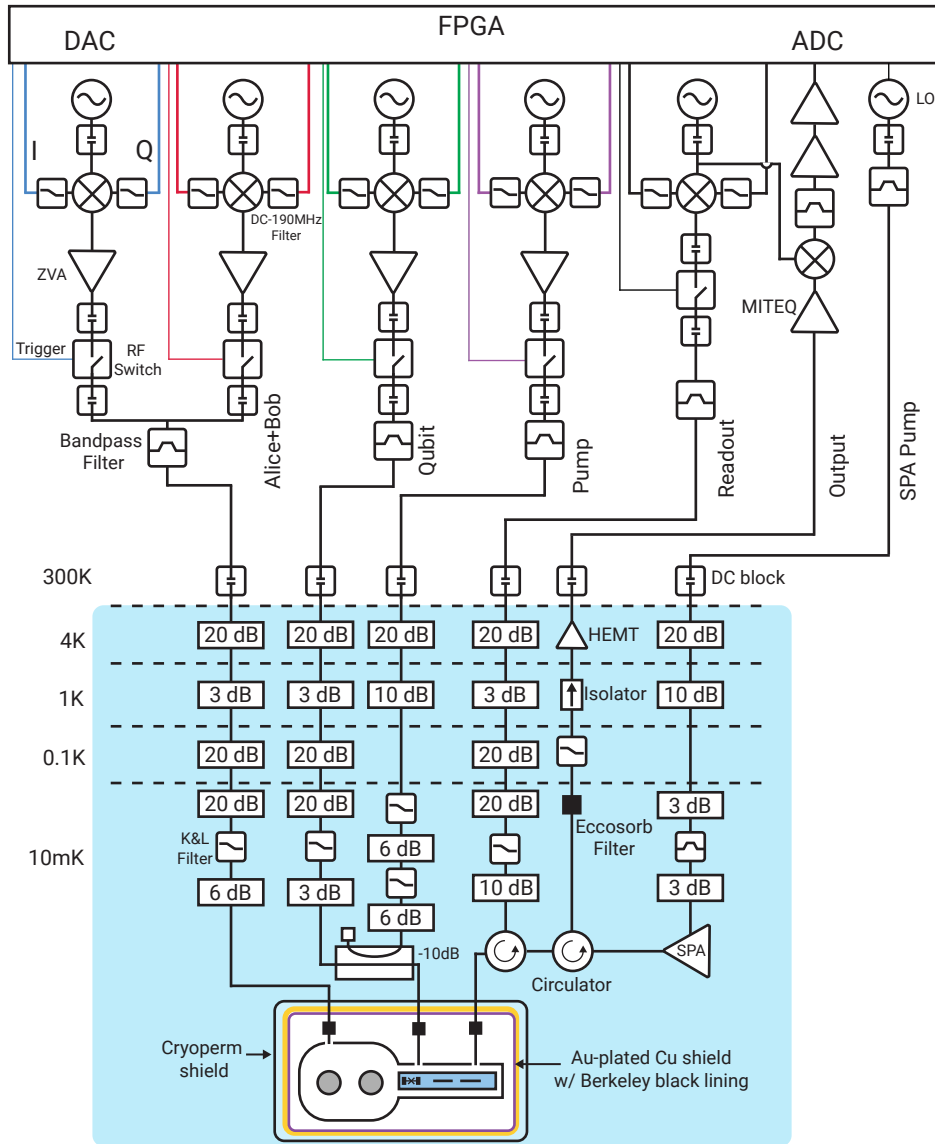(Dated: April 1, 2024)

CONTENTS

FIG. S 1. **Wiring diagram** Room temperature wiring for synthesizing control pulses and microwave wiring and shielding inside the dilution refrigerator.

EXPERIMENTAL SETUP AND SAMPLE PARAMETERS

Our experimental system consists of a coaxial stub cavity [1] with two posts, made of 5N Aluminium treated with a chemical etch to improve surface quality. The two harmonic oscillators are the two lowest frequency normal modes of the system obtained by the hybridisation of the $\lambda/4$ modes of the individual stubs. The auxiliary transmon [2] made of Aluminium is fabricated on a sapphire chip along with a readout resonator and a Purcell filter [3]. The chip is inserted into a tunnel waveguide that connects to the storage cavity modes and is held in place on one side using copper clamps. The entire package is rigidly attached to a gold-plated copper bracket that is mounted on to the base plate of a dilution refrigerator. The bracket is surrounded by a gold-plated copper can, coated with a thin layer of Berkeley black on the inside to absorb high frequency photons. The outer Cryoperm shield attenuates stray magnetic fields. The top of this can is sealed with a lid with SMA feedthroughs and each seam is sealed with Indium wires.

Control pulses for the relevant modes are synthesized via Digital-Analog Converters (DACs) from a Field Programmable Gate Array (FPGA) with a baseband of DC-250MHz [4]. These signals are up-converted to the required frequencies using IQ mixers and local oscillators (LO) which are then amplified and filtered accordingly. Fast RF switches are used to gate the signals in each line. These are then sent into different microwave lines in the dilution refrig-

erator each with different attenuation and filtering such that the noise temperature at relevant frequencies are around base plate temperature ($\sim$20mK). Readout signals are amplified via a quantum-limited SNAIL(Superconducting Nonlinear Asymmetric Inductive eLement) parametric amplifier(SPA) [5, 6], with a pump-port filter designed for efficient pump delivery. This amplified signal is further amplified by HEMT (High Electron Mobility Transistor) amplifier at the 4K stage and further by room temperature amplifier. This signal is then down-converted to a 50MHz signal using an IR mixer and the same LO used to up-convert the input readout pulse. After appropriate filtering and amplification, the Analog-Digital Converter(ADC) of the FPGA digitises, demodulates and integrates to obtain a readout value.

Alice	Frequency	$\omega_a = 2\pi \times 6.216\,\text{GHz}$
	Cross-Kerr shift	$\chi_{aq} = -2\pi \times 0.514\,\text{MHz}$
	Relaxation	$T_1 = 224.5 \pm 2.2\,\text{us}$ $\kappa_a = 4.454 \pm 0.044\,(\text{ms})^{-1}$
	Dephasing	$T_2^R = 452.4 \pm 8.6\,\text{us}$
	Thermal population	$\overline{n}_{th} = 0.0053 \pm 0.0002$
Bob	Frequency	$\omega_b = 2\pi \times 6.437\,\text{GHz}$
	Cross-Kerr shift	$\chi_{bq} = -2\pi \times 0.251\,\text{MHz}$
	Relaxation	$T_1 = 299.4 \pm 1.6\,\text{us}$ $\kappa_b = 3.339 \pm 0.018\,(\text{ms})^{-1}$
	Dephasing	$T_2^R = 604.6 \pm 5.2\,\text{us}$
	Thermal population	$\overline{n}_{th} = 0.0086 \pm 0.0003$
Transmon	Frequency	$\omega_q = 2\pi \times 4.681\,\text{GHz}$
	Anharmonicity	$\alpha = -2\pi \times 251\,\text{MHz}$
	Relaxation	$T_1 = 147.3 \pm 0.5\,\text{us}$
	Dephasing (Ramsey)	$T_{2R} = 87.2 \pm 1.3\,\text{us}$
	Dephasing (Echo)	$T_{2E} = 129.8 \pm 1.5\,\text{us}$
	Thermal population	$\overline{n}_{th} < 0.01$
Readout	Frequency	$\omega_r = 2\pi \times 8.159\,\text{GHz}$
	Cross-Kerr shift	$\chi_{qr} = -2\pi \times 0.432\,\text{MHz}$
	Coupling strength	$\kappa_{r(c)} = 3.943\,(\mu\text{s})^{-1}$
	Internal loss	$\kappa_{r(i)} = 1.030\,(\mu\text{s})^{-1}$

Supplementary Table I. Measured system parameters

CROSS-KERR TUNING

Deriving the effective Hamiltonian

We begin by writing the Hamiltonian of the Alice-Bob-Transmon system with a linear drive on the transmon delivered via a capacitively coupled port. Up to the leading order in the cosine potential of the Josephson junction, the Hamiltonian is

$$\hat{H}_0/\hbar = \omega_a \hat{a}^\dagger \hat{a} + \omega_b \hat{b}^\dagger \hat{b} + \omega_q \hat{q}^\dagger \hat{q} - g_4 \left[\phi_a(\hat{a} + \hat{a}^\dagger) + \phi_b(\hat{b} + \hat{b}^\dagger) + \phi_q(\hat{q} + \hat{q}^\dagger) \right]^4 - i\epsilon_D \cos \omega_d t (\hat{q} - \hat{q}^\dagger) \tag{1}$$

where ω_i is the frequency of the $i \in \{a, b, q, d\}$ mode representing Alice, Bob, transmon and the drive, respectively, and $g_4 = E_J/4!\hbar$ is the coefficient of the 4th order term representing the Josephson energy. First we perform a displaced frame transformation on the qubit to absorb the drive into the 4th order term. The Hamiltonion becomes

$$\hat{H}_D = \hat{D}(\alpha)\hat{H}_0\hat{D}^\dagger(\alpha) + i\dot{\hat{D}}(\alpha)\hat{D}^\dagger(\alpha) \tag{2}$$

where $\hat{D}(\alpha) = e^{\alpha\hat{q}^\dagger - \alpha^*\hat{q}}$ and $\hat{D}(\alpha)\hat{q}\hat{D}^\dagger(\alpha) \to \hat{q} - \alpha$. Choosing the displacement $\alpha = \xi e^{-i\omega_d t}$ such that $\xi = \frac{i\epsilon_d}{\omega_q - \omega_d}$, ignoring terms oscillating at $2\omega_d$, we can write the Hamiltonian in the displaced frame, up to constant terms, as

$$\frac{\hat{H}_D}{\hbar} = \omega_a\hat{a}^\dagger\hat{a} + \omega_b\hat{b}^\dagger\hat{b} + \omega_q\hat{q}^\dagger\hat{q} - g_4\left[\phi_a(\hat{a} + \hat{a}^\dagger) + \phi_b(\hat{b} + \hat{b}^\dagger) + \phi_q(\hat{q} + \hat{q}^\dagger + \xi e^{-i\omega_d t} + \xi^* e^{i\omega_d t})\right]^4 \tag{3}$$

Next we move into a frame where the modes are rotating at their bare frequencies. We add a slight detuning to Bob from its bare detuning (adding this detuning on the transmon has equivalent reuslt). The unitary for this transformation is then, $\hat{U}(t) = \exp\left[-i(\omega_a\hat{a}^\dagger\hat{a} + (\omega_b - \Delta)\hat{b}^\dagger\hat{b} + \omega_q\hat{q}^\dagger\hat{q})t\right]$ and we get the interaction Hamiltonian

$$\frac{\hat{H}_I}{\hbar} = \Delta\hat{b}^\dagger\hat{b} - g_4\left[\left(\phi_a\hat{a}e^{-i\omega_a t} + \phi_b\hat{b}e^{-i(\omega_b - \Delta)t} + \phi_q\hat{q}e^{-i\omega_q t} + \phi_q\xi e^{-i\omega_d t}\right) + h.c.\right]^4 \tag{4}$$

where $h.c.$ represents the hermitian conjugate of all the terms in the rounded brackets preceding it. For the cross-Kerr tuning process we drive a 2-photon transition at the frequency $\omega_d = 2\omega_q - \omega_b + \Delta$. Substituting this in to the above Hamiltonian and collecting all the static terms from the expansion of the 4th order term, with the rotating wave approximation(RWA) applied, we get

$$\begin{aligned}
\frac{\hat{H}_p}{\hbar} = {} & \Delta\hat{b}^\dagger\hat{b} + \frac{1}{2}\left(\Omega\hat{q}^2\hat{b}^\dagger + \Omega^*\hat{q}^{\dagger 2}\hat{b}\right) + \\
& \Delta_s^a\hat{a}^\dagger\hat{a} + \Delta_s^b\hat{b}^\dagger\hat{b} + \Delta_s^q\hat{q}^\dagger\hat{q} + \\
& \chi_{aq}\hat{a}^\dagger\hat{a}\hat{q}^\dagger\hat{q} + \chi_{bq}\hat{b}^\dagger\hat{b}\hat{q}^\dagger\hat{q} + \chi_{ab}\hat{a}^\dagger\hat{a}\hat{b}^\dagger\hat{b} + \\
& \frac{K_{aa}}{2}\hat{a}^{\dagger 2}\hat{a}^2 + \frac{K_{bb}}{2}\hat{b}^{\dagger 2}\hat{b}^2 + \frac{K_{qq}}{2}\hat{q}^{\dagger 2}\hat{q}^2
\end{aligned} \tag{5}$$

where, $\Delta_s^i = \frac{1}{2}K_{ii}|\xi|^2$ is the Stark-shift and $K_{ii} = -\frac{1}{2}E_J\phi_i^4$ is the self-Kerr of the i-th mode and $\chi_{ij} = -E_J\phi_i^2\phi_j^2$ is the cross-Kerr between modes i and j. In the first line in the above equation, $\Omega = \xi^*\sqrt{2\chi_{bq}K_{qq}}$ is the interaction rate of the cross-Kerr tuning process, detuned by Δ. Going into the interaction frame with respect to the transmon self-Kerr $U(t) = \exp\left(-i\frac{K_{qq}}{2}t\hat{q}^{\dagger 2}\hat{q}^2\right)$, we get

$$\begin{aligned}
\frac{\hat{H}_p}{\hbar} = {} & \Delta\hat{b}^\dagger\hat{b} + \left[\frac{\Omega}{2}\hat{b}^\dagger\left(\sqrt{2}\,|g\rangle\langle f|\,e^{-iK_{qq}t} + \sqrt{6}\,|e\rangle\langle h|\,e^{-3iK_{qq}t} + ..\right)\right] + h.c. + \\
& \chi_{aq}\hat{a}^\dagger\hat{a}\hat{q}^\dagger\hat{q} + \chi_{bq}\hat{b}^\dagger\hat{b}\hat{q}^\dagger\hat{q} + \chi_{ab}\hat{a}^\dagger\hat{a}\hat{b}^\dagger\hat{b} + \frac{K_{aa}}{2}\hat{a}^{\dagger 2}\hat{a}^2 + \frac{K_{bb}}{2}\hat{b}^{\dagger 2}\hat{b}^2
\end{aligned} \tag{6}$$

By choosing $\Delta \to \Delta - 3K_{qq}$ and moving into the appropriate rotating frame, we can selectively pick out the $|e\rangle \leftrightarrow |h\rangle$ transition to address and ignore the other pumped terms under the RWA($\Delta << K_{qq} << \omega_q$). Hence, we get,

$$\frac{\hat{H}_\chi}{\hbar} \approx (\Delta - 3K_{qq})\hat{b}^\dagger\hat{b} + \left(\Omega\,|e\rangle\langle h|\,\hat{b}^\dagger + \Omega^*\,|h\rangle\langle e|\,\hat{b}\right) + \chi_{aq}\hat{a}^\dagger\hat{a}\hat{q}^\dagger\hat{q} + \chi_{bq}\hat{b}^\dagger\hat{b}\hat{q}^\dagger\hat{q} \tag{7}$$

where we've redefined the drive amplitude to $\Omega \to \sqrt{\frac{3}{2}}\Omega$ and dropped the negligible terms like the self-Kerr rates of the cavities and the cross-Kerr between them. Note that the pump frequency is $\omega_d = 2\omega_q - \omega_b - 3K_{qq} + \Delta$. By diagonalizing the above pumped Hamiltonian, we obtain the behavior of the combined Fock states, as seen from the fits in Fig1.(c) of the main text.
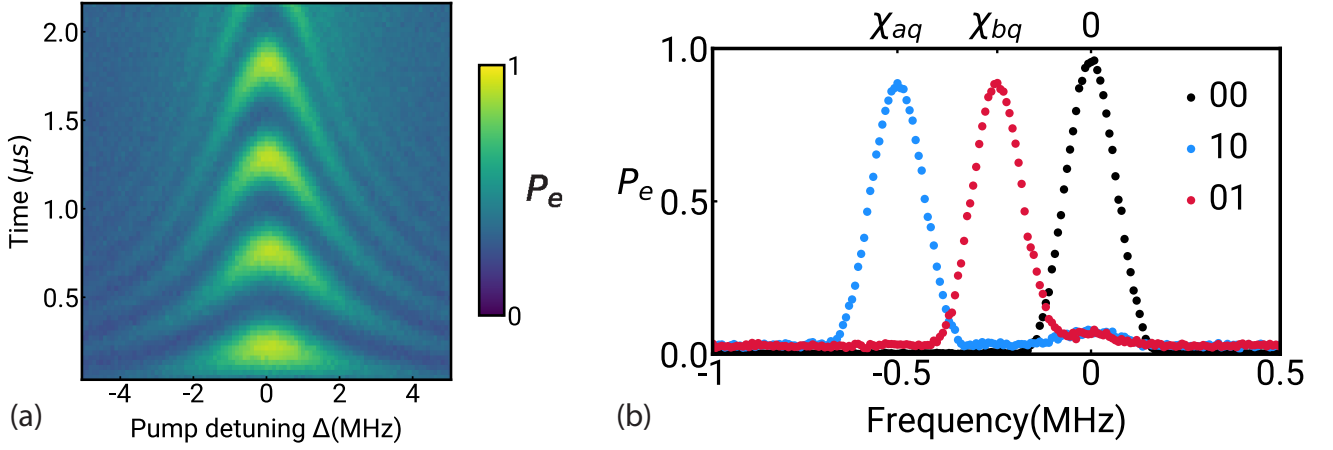
FIG. S 2. **a,** Chevron plot obtained after initialising the transmon in $|e\rangle$ and Bob cavity in $|1\rangle$ and sweeping time and frequency of the cross-Kerr tuning pump. Fit Rabi rate $\Omega = 2\pi \times 0.98$ MHz. Transmon spectroscopy after preparing Alice and Bob in states $|00\rangle$, $|01\rangle$ and $|10\rangle$. The Gaussian pulse used for spectroscopy is the same one used for erasure detection. The frequency axis referenced with respect to the bare transmon $g \to e$ frequency.

To get an intuition of how the above Hamiltonian modifies the cross-Kerr, let us consider the transmon coupled to just Bob, expand in the Fock basis and select only the term that couples $|h0\rangle \leftrightarrow |e1\rangle$. Here we write a simplified form of the above Hamiltonian for this term,

$$\frac{\hat{H}_p}{\hbar} = \Delta\,|e1\rangle\langle e1| + \Omega\,|e1\rangle\langle h0| + \Omega^*\,|h0\rangle\langle e1| \tag{8}$$

This describes two levels coupled to each other and we can diagonalize easily to find the eigenvalues and eigenstate. The eigenvalues are $\lambda_\pm = -\frac{\Delta}{2} \pm \frac{1}{2}\sqrt{\Delta^2 + 4|\Omega|^2}$. For the limit of $\Omega << \Delta$, the lowest eigenvalue can be approximated to the lowest order as $\lambda_+ \approx |\Omega|^2/\Delta$ and the corresponding eigenstate is $|\lambda_+\rangle = \big|\tilde{e1}\big\rangle \approx |e1\rangle + \frac{\Omega}{\Delta}|h0\rangle$. For large detunings, the hybridisation is weak and if the $|h\rangle$ level of the transmon is never occupied, then we can approximately rewrite the Hamiltonian as

$$\frac{\hat{H}_p}{\hbar} \approx \frac{|\Omega|^2}{\Delta}\left(\big|\tilde{e1}\big\rangle\big\langle\tilde{e1}\big| - \big|\tilde{h0}\big\rangle\big\langle\tilde{h0}\big|\right) \tag{9}$$

This corresponds to modified bare energy of the state. If we now perform the same for higher Fock states it will result in modified cross-Kerr due to the drive.

Characterizing the cross-Kerr tuning process

To characterize the cross-Kerr tuning, we prepare either of the cavities in the $|1\rangle$ state and the transmon in $|e\rangle$. By applying the cross-Kerr pump with varying frequency and time, we measure the probability of the transmon to be in $|e\rangle$. This reveals the chevron between the levels $|e1\rangle$ and $|h0\rangle$ with a Rabi rate Ω_{h0e1}. We use this method to calibrate the strength and the stark-shifted centre frequency of the process, for either cavity modes using Eq. (6). Fig.S 2a shows the chevron pattern of such an oscillation perform on the Bob cavity.

Gaussian pulse for erasure detection

Fig.S 2b shows transmon spectroscopy results obtained after preparing the cavities in $|00\rangle$, $|01\rangle$ and $|10\rangle$ states. The spectroscopy is performed using a Gaussian pulse with $\sigma = 2\mu$s, with a total duration of $4\sigma = 8\mu$s. This is the same pulse used in the erasure detection circuit.

DUAL-RAIL QUBIT

Encoding and error-channels

We define the dual-rail qubit in the joint-Hilbert space of Alice and Bob. The logical codewords are $|+Z_L\rangle = |01\rangle$ and $|-Z_L\rangle = |10\rangle$. Here the first mode in the ket is Alice and the second is Bob. The full error set for the code due to error channels of the cavities is $E = \{\hat{I}, \hat{a}, \hat{b}, \hat{a}^\dagger, \hat{b}^\dagger, \hat{n}_a, \hat{n}_b, e^{-\frac{1}{2}(\kappa_a \hat{n}_a + \kappa_b \hat{n}_b)t}\}$. These represent the identity, cavity photon loss, cavity photon gain, cavity dephasing and the no-jump error channels respectively. The i-th error channel $\hat{E}_i$ will act on an arbitrary state in the logical codespace, $|\psi\rangle = V|01\rangle + W|10\rangle$, and take the system into the error space as follows,

$$|E_i\rangle = \frac{\hat{E}_i |\psi\rangle}{\sqrt{\langle\psi| \hat{E}_i^\dagger \hat{E}_i |\psi\rangle}} \tag{10}$$

where E_i is the i-th error channel in the set E. From this, it becomes apparent that the error state for the photon loss channel, in either Alice or Bob, is the ground state of the system : $|E_1\rangle = |E_2\rangle = |00\rangle$. For the photon gain channels, we have

$$|E_3\rangle = \frac{V}{\sqrt{1+|W|^2}} |11\rangle + \frac{\sqrt{2}W}{\sqrt{1+|W|^2}} |20\rangle \tag{11}$$

$$|E_4\rangle = \frac{\sqrt{2}V}{\sqrt{1+|V|^2}} |02\rangle + \frac{W}{\sqrt{1+|V|^2}} |11\rangle \tag{12}$$

Note that for both the photon gain channels the total excitation in the cavities is 2. Dephasing error channel of the individual cavities will lead to dephasing within the dualrail codespace. Finally, there will be a backaction on the codespace due to the no-jump evolution of the states which will look like

$$|E_7\rangle = \frac{V|01\rangle + We^{-\frac{1}{2}\Delta\kappa t}|10\rangle}{\sqrt{|V|^2 + |W|^2 e^{-\Delta\kappa t}}} \tag{13}$$

where $\Delta\kappa = \kappa_a - \kappa_b$ is the difference in loss rates between the two modes. This no-jump evolution causes the state to be polarised to the longer lived mode, distorting the encoded information.

Erasure detection and trajectories

Hidden Markov decoding

To predict cavity photon loss and the most likely state of the system, we train a Categorical Hidden Markov Model (HMM) with the experimentally obtained trajectories. The HMM assumes 2 hidden states, Codespace (C) and Errorspace (E) and 2 measurement outcomes "g" and "e". The transition matrix element t_{ij} determines the probability of the system to make a transition from state $i \rightarrow j$ in the current time step. Similarly, the emission matrix element e_{mn} is the probability of observing the measurement outcome m given that the system is in state n at the current time step. In our experiment, we measure 10^4 trajectories for each of the 6 cardinal states in the dual-rail Bloch sphere and train an HMM for each state to learn the probabilities. Each trajectory is 167 measurement rounds long. Fig.S 3a shows a sample of these trajectories for the 6 cardinal states while Fig.S 3b shows the decoded trajectories as predicted by the HMM. Fig.S 3c and Fig.S 3d depicts the transition and emission matrices obtained after the training on the raw trajectories. Since each erasure detection round is $12\mu s$ long, we can convert these probabilities in to rates.

We expect the erasure rates to be κ_a and κ_b for the $|\pm Z\rangle$ states and $\frac{1}{2}(\kappa_a + \kappa_b)$ for the states on the equator. The HMM predicted values matches well with the rates extracted independently(from Table 1). The HMM also allows us to extract false positive and false negative probabilities.
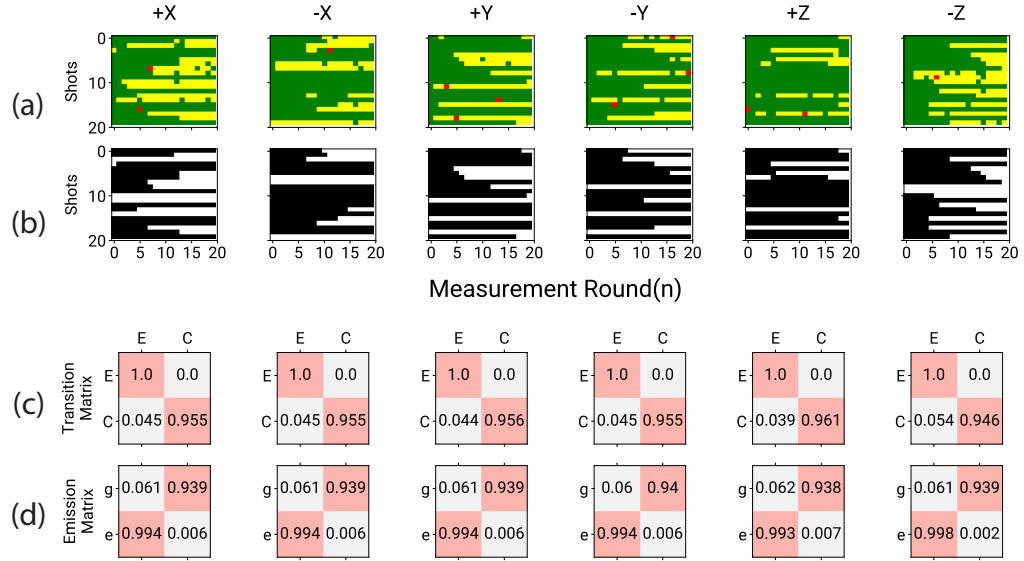
FIG. S 3. **a,** Raw trajectories measured after preparing the 6 cardinal states in the dual-rail subspace. Here we only show 20 samples trajectories with 20 erasure detection rounds in each trajectory. Each pixel represents a detection round with "g" as green, "e" as yellow and red represents states higher than "e". **b,** Corresponding decoded trajectories as predicted by the Hidden Markov Model (HMM). The model is trained for each states with 10^4 trajectories and more than 150 erasure detection rounds per trajectory. Black squares represents the system is within the dual-rail codespace and white squares represents the error space. **c,** Learned transition matrix for each of the 6 states. The off-diagonal elements represents the probability of the system to transition between codespace(C) and errorsapce(E). **d,** Learned emission matrix corresponding to the measurement outcomes "g" and "e".

State	Erasure prob.	Erasure rate
$|+Z\rangle$	0.039 per gate	$3.25(\mathrm{ms})^{-1}$
$|-Z\rangle$	0.054 per gate	$4.5(\mathrm{ms})^{-1}$
Equator states	0.045 per gate	$3.75(\mathrm{ms})^{-1}$

Supplementary Table II. Error probabilities calculated from the HMM model

$$\text{False negative prob.} = P(C \to E) \times P(g|E) = t_{CE} \times e_{gE} = 0.0028 \, \text{per gate}$$
$$\text{False positive prob.} = P(C \to C) \times P(e|C) = t_{CC} \times e_{eC} = 0.006 \, \text{per gate} \tag{14}$$

The above values are averaged over all 6 cardinal states.

TOMOGRAPHY

Parity mapping to measure single-mode Wigner

The single mode Wigner for a bosonic mode is defined as

$$W(\alpha) = \frac{2}{\pi} \text{Tr}\big[\rho \hat{P}_A(\alpha)\big] \tag{15}$$

where $\hat{P}_A(\alpha) = \hat{D}(-\alpha)\hat{\Pi}_A\hat{D}(\alpha)$ is the displaced parity operator with $\hat{\Pi}_A = (-1)^{\hat{a}^\dagger \hat{a}}$ being the parity operator for the mode. Assuming we have a single bosonic mode dispersively coupled to a transmon as $\hat{H} = \hbar\chi\hat{a}^\dagger\hat{a} |e\rangle\langle e|$, we can map the parity of the mode to the transmon state. To do so, after preparing the desired state, we initialise the transmon in the +X state. We then let the dispersive interaction act for exactly $t = \pi/\chi$ time. We then obtain,
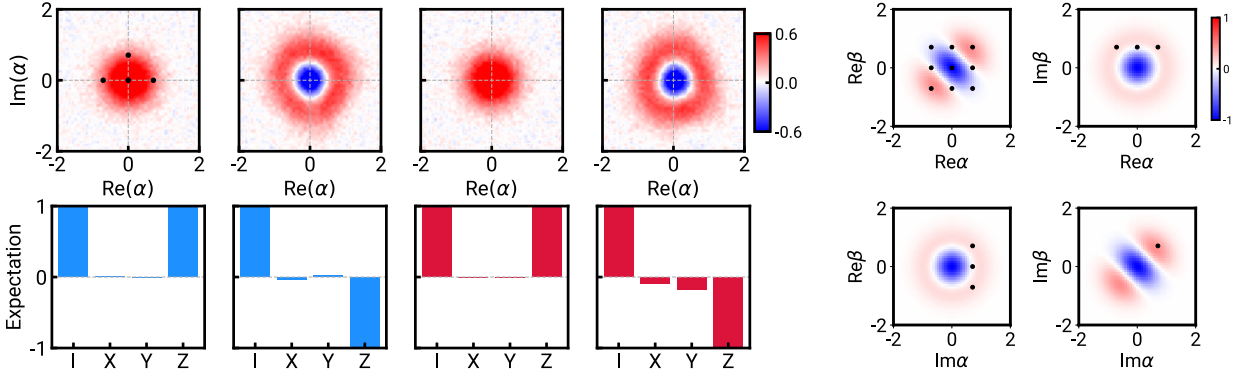
FIG. S 4. **a,** Measured single-mode Wigner functions for Alice and Bob after initialising in Fock states $|0\rangle$ and $|1\rangle$. The first plot shows the 4 points to be measured(black dots) in order to extract the expectation values of the Pauli operators within the $0-1$ Fock qubit. The bottom panel shows the expectation values of the Pauli operators using these 4 measurements for both Alice and Bob initialised in the same states(with $|+Z\rangle = |0\rangle, |-Z\rangle = |1\rangle$). We obtain the expected values up to SPAM errors. **b,** Simulated 2D cuts of joint-Wigner function for the $|+X\rangle_L$ state in the dual-rail subspace. In each of the plot, the other 2 values are set to 0. The black dots represents the 16 values in the joint phase space of Alice and Bob to be measured to extract the 16 2-qubit Pauli operator expectation values.

$$U(t)|\psi\rangle \otimes \frac{1}{\sqrt{2}}(|g\rangle + |e\rangle) = \frac{1}{\sqrt{2}}(|\psi,g\rangle + e^{-i\chi t \hat{a}^\dagger \hat{a}}|\psi,e\rangle) = \frac{1}{\sqrt{2}}(|\psi,g\rangle + e^{-i\pi \hat{a}^\dagger \hat{a}}|\psi,e\rangle) = \frac{1}{\sqrt{2}}(|\psi,g\rangle + \hat{\Pi}_A|\psi,e\rangle) \quad (16)$$

Hence, by measuring the transmon in the X-basis, we measure the parity of the bosonic mode.

Logical Pauli measurements

To extract the expectation values of the logical Pauli operators of the dual-rail subspace, we sample specific points in the joint-Wigner space of the cavities [7, 8]. To see this, we first define describe how to measure the expectation values of the logical Pauli operators for the Fock qubit using single-mode Wigner. We will then extend this method to measure the logical Pauli operators of the dual-rail qubit using Joint-Wigner.

Fock qubit

The logical codewords for the Fock qubit are $|+Z_L\rangle = |0\rangle$ and $|-Z_L\rangle = |1\rangle$. We then proceed to project the displaced parity onto the fock qubit subspace using the identity projector $\hat{I} = |0\rangle\langle 0| + |1\rangle\langle 1|$. We get

$$\hat{I}\hat{P}_A(\alpha)\hat{I} = \begin{bmatrix} \langle 0|\hat{P}_A|0\rangle & \langle 0|\hat{P}_A|1\rangle \\ \langle 1|\hat{P}_A|0\rangle & \langle 1|\hat{P}_A|1\rangle \end{bmatrix} = \begin{bmatrix} \langle \alpha|\hat{\Pi}_A|\alpha\rangle & \langle \alpha|\hat{\Pi}_A\hat{D}(\alpha)|1\rangle \\ \langle 1|\hat{D}(-\alpha)\hat{\Pi}_A|\alpha\rangle & \langle 1|\hat{D}(-\alpha)\hat{\Pi}_A\hat{D}(\alpha)|1\rangle \end{bmatrix} \quad (17)$$

Using Laguerre polynomials and the identity $(-1)^{\hat{a}^\dagger \hat{a}}|\alpha\rangle = |-\alpha\rangle$, we simplify the above equation to

$$\hat{I}_L\hat{P}_A(\alpha)\hat{I}_L = \begin{bmatrix} \langle \alpha|-\alpha\rangle & \langle -2\alpha|1\rangle \\ \langle 1|-2\alpha\rangle & \langle 1|\hat{P}(\alpha)|1\rangle \end{bmatrix} = e^{-2|\alpha|^2}\begin{bmatrix} 1 & 2\alpha^* \\ 2\alpha & 4|\alpha|^2 - 1 \end{bmatrix} \quad (18)$$

here we used $\langle -\alpha|\alpha\rangle = e^{-2|\alpha|^2}$ and $\langle 1|\hat{P}(\alpha)|1\rangle = e^{-2|\alpha|^2}(4|\alpha|^2 - 1)$. We then equate the above equation to the 4 Pauli operators $\{\hat{I}, \hat{X}, \hat{Y}, \hat{Z}\}$ and solving for α, we get $\alpha_i = \{0, \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, \frac{i}{\sqrt{2}}\}$. This tells us that if we restrict ourselves to the Fock qubit subspace then we only need to measure these four points in phase space to reconstruct the full density matrix. Fig.S 4a, shows the pictorially the four points in phase space and also provides intuition to why we expect it. Now, using these displacements, we can invert the above equation to get the Pauli operators as
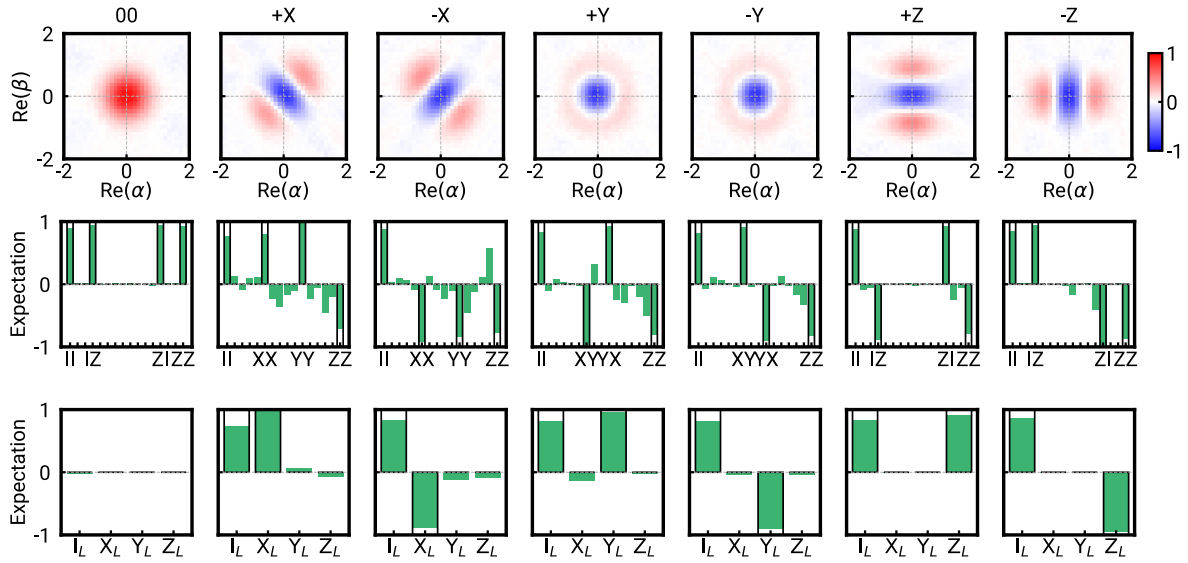
FIG. S 5. **a,** Experimentally measured 2d joint-Wigner cuts of the relevant states for the dual-rail code(6 cardinal states and the error state. The second row shows the 16 two-qubit Pauli expectation values extracted by measuring 16 points in phase space. Last row depicts the extracted logical Pauli expectation values of the dual-rail subspace. All the measured data shown includes SPAM errors.

$$
\begin{aligned}
\hat{I} &= \frac{e}{2}\Big[\hat{P}(\frac{1}{\sqrt{2}}) + \hat{P}(-\frac{1}{\sqrt{2}})\Big] \\
\hat{X} &= \frac{e}{2\sqrt{2}}\Big[\hat{P}(\frac{1}{\sqrt{2}}) - \hat{P}(-\frac{1}{\sqrt{2}})\Big] \\
\hat{Y} &= \frac{e}{2\sqrt{2}}\Big[2\hat{P}(\frac{i}{\sqrt{2}}) - \hat{P}(\frac{1}{\sqrt{2}}) - \hat{P}(-\frac{1}{\sqrt{2}})\Big] \\
\hat{Z} &= \hat{P}(0)
\end{aligned}
\tag{19}
$$

Joint-parity mapping to measure joint Wigner

We can easily extend this technique for our dual-rail qubit. Crucially, we will have to measure the joint-Wigner, defined as,

$$
W(\alpha, \beta) = \frac{4}{\pi^2}\mathrm{Tr}\Big[\rho\hat{P}_{\mathrm{J}}(\alpha, \beta)\Big]
\tag{20}
$$

where we define the displaced joint parity operator as $\hat{P}_J(\alpha, \beta) = \hat{P}_A(\alpha) \otimes \hat{P}_B(\beta)$. To measure the joint-parity, we again turn to the dispersively coupled transmon. In our system, the dispersive coupling is of the form $\hat{H} = \hbar\chi_{aq}\hat{a}^\dagger\hat{a}\,|e\rangle\langle e|\,\hbar\chi_{bq}\hat{b}^\dagger\hat{b}\,|e\rangle\langle e|$. Therefore, using the same parity mapping sequence as before, we would obtain,

$$
U(t)\,|\psi\rangle_{ab} \otimes \frac{1}{\sqrt{2}}(|g\rangle + |e\rangle) = \frac{1}{\sqrt{2}}(|\psi_{ab}, g\rangle + e^{-i\chi_{aq}t\hat{a}^\dagger\hat{a}+\chi_{bq}t\hat{b}^\dagger\hat{b}}\,|\psi_{ab}, e\rangle)
\tag{21}
$$

To obtain the joint-parity, we would need to choose t such that it is a multiple of both χ_{aq} and χ_{bq}, which in general is difficult. However, if $\chi_{aq} = \chi_{bq} = \chi$, we can set $t = \pi/\chi$ to again measure the joint-parity by measuring the transmon in the X-basis. Hence the need for cross-Kerr matching.
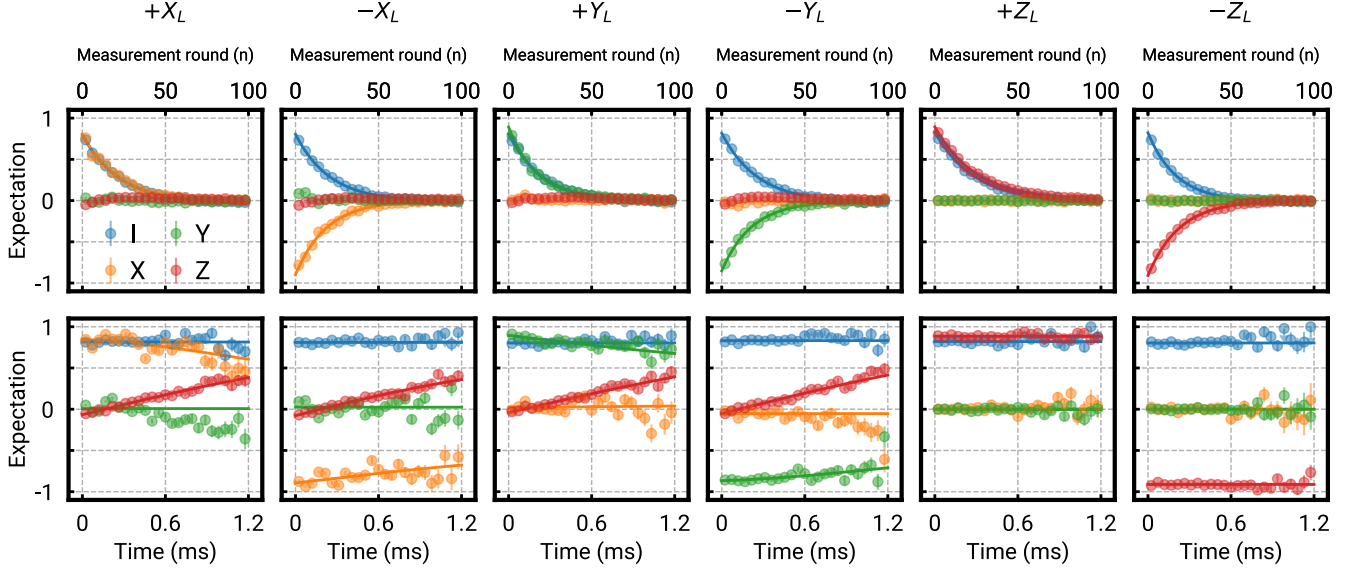
FIG. S 6. **a,** Measured expectation values of the dual-rail Pauli operators as a function of number of erasure detection rounds, with solid lines shows the fits to exponential decay. **b,** Expectation values postselected on no jumps in the trajectory. $\langle \hat{I}(t) \rangle$ and $\langle \hat{Z}(t) \rangle$ are fit to exponential decay. Since these plots are extremely flat, we can only extract an upper bound on the residual leakage and bit-flip rate. $\langle \hat{X}(t) \rangle$ and $\langle \hat{Y}(t) \rangle$ are fit to the analytical no-jump evolution with an additional exponential decay term to extract residual dephasing rate.

Dual-rail qubit

We first note that we cannot directly project on to the dual-rail subspace since, due to erasure errors, the system leaks out of this subspace. In order to incorporate the error state $|00\rangle$, we instead choose to restrict ourselves on to the two qubit subspace spanned by Fock qubit in both Alice and Bob : $\{|0\rangle, |1\rangle\}_A \otimes \{|0\rangle, |1\rangle\}_B$. The dual-rail subspace, including the error space, lives within this larger space, even though we have some states we are not concerned with. Now, projecting the displaced joint parity onto this subspace, we get,

$$(\hat{I}_A \otimes \hat{I}_B)\hat{P}_J(\alpha, \beta)(\hat{I}_A \otimes \hat{I}_B) = (\hat{I}_A\hat{P}_A(\alpha)\hat{I}_A) \otimes (\hat{I}_B\hat{P}_B(\beta)\hat{I}_B) \tag{22}$$

which we then equate to the 16 two-qubit Paulis $\{\hat{I}, \hat{X}, \hat{Y}, \hat{Z}\}_A \otimes \{\hat{I}, \hat{X}, \hat{Y}, \hat{Z}\}_B$. Its easy to see that this will yield 16 points, which are just combinations of the 4 points from single-mode displacements that we derived above. We write this set of 16 displacements explicitly below,

$$d = (\alpha, \beta) = \frac{1}{\sqrt{2}} \times \left\{ (0,0), (0,1), (0,-1), (0,i), (1,0), (1,1), (1,-1), (1,i), \right.$$
$$\left. (-1,0), (-1,1), (-1,-1), (-1,i), (i,0), (i,1), (i,-1), (i,i), \right\} \tag{23}$$

We will denote d_i as the i-th element in the above set, with $i = 0, 1, 2, ..., 15$. From these displacements, we can construct the 16 two-qubit Pauli matrices as follows,

$$\hat{I}_A\hat{I}_B = \frac{e^2}{4}\left[\hat{P}_J(d_5) + \hat{P}_J(d_6) + \hat{P}_J(d_9) + \hat{P}_J(d_{10})\right]$$

$$\hat{I}_A\hat{X}_B = \frac{e^2}{4\sqrt{2}}\left[\hat{P}_J(d_5) - \hat{P}_J(d_6) + \hat{P}_J(d_9) - \hat{P}_J(d_{10})\right]$$

$$\hat{I}_A\hat{Y}_B = \frac{e^2}{4\sqrt{2}}\left[2\hat{P}_J(d_7) - \hat{P}_J(d_5) - \hat{P}_J(d_6) + 2\hat{P}_J(d_{11}) - \hat{P}_J(d_9) - \hat{P}_J(d_{10})\right]$$

$$\hat{I}_A\hat{Z}_B = \frac{e}{2}\left[\hat{P}_J(d_4) + \hat{P}_J(d_8)\right]$$

$$\hat{X}_A\hat{I}_B = \frac{e^2}{4\sqrt{2}}\left[\hat{P}_J(d_5) + \hat{P}_J(d_6) - \hat{P}_J(d_9) - \hat{P}_J(d_{10})\right]$$

$$\hat{X}_A\hat{X}_B = \frac{e^2}{8}\left[\hat{P}_J(d_5) - \hat{P}_J(d_6) - \hat{P}_J(d_9) + \hat{P}_J(d_{10})\right]$$

$$\hat{X}_A\hat{Y}_B = \frac{e^2}{8}\left[2\hat{P}_J(d_7) - \hat{P}_J(d_5) - \hat{P}_J(d_6) - 2\hat{P}_J(d_{11}) + \hat{P}_J(d_9) + \hat{P}_J(d_{10})\right]$$

$$\hat{X}_A\hat{Z}_B = \frac{e}{2\sqrt{2}}\left[\hat{P}_J(d_4) - \hat{P}_J(d_8)\right]$$

$$\hat{Y}_A\hat{I}_B = \frac{e^2}{4\sqrt{2}}\left[2\hat{P}_J(d_{13}) + 2\hat{P}_J(d_{14}) - \hat{P}_J(d_5) - \hat{P}_J(d_6) - \hat{P}_J(d_9) - \hat{P}_J(d_{10})\right] \quad (24)$$

$$\hat{Y}_A\hat{X}_B = \frac{e^2}{8}\left[2\hat{P}_J(d_{13}) - 2\hat{P}_J(d_{14}) - \hat{P}_J(d_5) + \hat{P}_J(d_6) - \hat{P}_J(d_9) + \hat{P}_J(d_{10})\right]$$

$$\hat{Y}_A\hat{Y}_B = \frac{e^2}{8}\left[4\hat{P}_J(d_{15}) - 2\hat{P}_J(d_{13}) - 2\hat{P}_J(d_{14}) - 2\hat{P}_J(d_7) + \hat{P}_J(d_5) + \hat{P}_J(d_6) - 2\hat{P}_J(d_{11}) + \hat{P}_J(d_9) + \hat{P}_J(d_{10})\right]$$

$$\hat{Y}_A\hat{Z}_B = \frac{e^2}{2\sqrt{2}}\left[\hat{P}_J(d_{12}) - \hat{P}_J(d_4) - \hat{P}_J(d_8)\right]$$

$$\hat{Z}_A\hat{I}_B = \frac{e}{2}\left[\hat{P}_J(d_1) + \hat{P}_J(d_2)\right]$$

$$\hat{Z}_A\hat{X}_B = \frac{e}{2\sqrt{2}}\left[\hat{P}_J(d_1) - \hat{P}_J(d_2)\right]$$

$$\hat{Z}_A\hat{Y}_B = \frac{e}{2\sqrt{2}}\left[2\hat{P}_J(d_3) - \hat{P}_J(d_1) - \hat{P}_J(d_2)\right]$$

$$\hat{Z}_A\hat{Z}_B = \hat{P}_J(d_0)$$

For our purposes, we only care about the Pauli operators for the dual-rail, which can be obtained by linear combinations of the above matrices, since we work in the $\{|01\rangle, |10\rangle\}$ basis. Therefore, the logical Pauli operators for the dual-rail qubit are

$$\hat{I}_L = \frac{1}{2}\left[\hat{I}_A\hat{I}_B - \hat{Z}_A\hat{Z}_B\right]$$

$$\hat{X}_L = \frac{1}{2}\left[\hat{X}_A\hat{X}_B + \hat{Y}_A\hat{Y}_B\right]$$

$$\hat{Y}_L = \frac{1}{2}\left[\hat{Y}_A\hat{X}_B - \hat{X}_A\hat{Y}_B\right] \quad (25)$$

$$\hat{Z}_L = \frac{1}{2}\left[\hat{Z}_A\hat{I}_B - \hat{I}_A\hat{Z}_B\right]$$

Fig.S 4b shows the 16 points measured in phase space, depicted in different joint-Wigner cuts. Fig.S 4 shows the measurement of joint-Wigner cuts, the expectation values of the two-qubit Pauli operators and subsequently constructing the dual-rail Pauli expectation values from them.

Dual-rail Pauli transfer matrix with erasure detection

We measure the expectation values of the Pauli operators after preparing the 6 cardinal states as a function of number of erasure detection rounds. Top panel in Fig.S 6 shows the measured values with exponential fits. The decay of the identity operator clearly shows leakage outside of the codespace. The bottom panel shows postselected data. The postselected $\langle \hat{I} \rangle$ looks constant and we can only obtain an upper bound on residual leakage rate by fitting to an exponential decay ($< 10^{-4}(\text{ms})^{-1}$). We then fit the postselected $\langle \hat{X} \rangle$ and $\langle \hat{Y} \rangle$ to extract dephasing rates within the dualrail subspace. Taking into account the no-jump evolution, the analytical formula for the evolution of these operators will look like [9]

$$\langle \hat{X}(t) \rangle = \frac{(W^*V + V^*W)e^{-\frac{1}{2}\Delta\kappa t}}{|V|^2 + |W|^2 e^{-\Delta\kappa t}} \tag{26}$$

$$\langle \hat{Y}(t) \rangle = \frac{i(W^*V - V^*W)e^{-\frac{1}{2}\Delta\kappa t}}{|V|^2 + |W|^2 e^{-\Delta\kappa t}} \tag{27}$$

$$\langle \hat{Z}(t) \rangle = \frac{(|V|^2 - |W|^2)e^{-\frac{1}{2}\Delta\kappa t}}{|V|^2 + |W|^2 e^{-\Delta\kappa t}} \tag{28}$$

for an arbitrary state $|\psi\rangle = V|01\rangle + W|10\rangle$. On top of this, we add an exponentially decaying term to calculate residual dephasing and bit-flip rate within the codespace. We observe that the fits agree well with the measured data. We note that the increase in $\langle \hat{Z} \rangle$ for the equator states shows the polarisation of the Bloch vector towards the longer lived cavity (Bob) due to the no-jump backaction.

* These authors contributed equally to this work.
akshay.koottandavida@yale.edu
ioannis.tsioutsios@yale.edu
† michel.devoret@yale.edu

[1] M. Reagor, H. Paik, G. Catelani, L. Sun, C. Axline, E. Holland, I. M. Pop, N. A. Masluk, T. Brecht, L. Frunzio, M. H. Devoret, L. Glazman, and R. J. Schoelkopf, Reaching 10 ms single photon lifetimes for superconducting aluminum cavities, Applied Physics Letters **102**, 192604 (2013).

[2] J. Koch, T. M. Yu, J. Gambetta, A. A. Houck, D. I. Schuster, J. Majer, A. Blais, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, Charge-insensitive qubit design derived from the Cooper pair box, Physical Review A **76**, 042319 (2007).

[3] E. M. Purcell, H. C. Torrey, and R. V. Pound, Resonance Absorption by Nuclear Magnetic Moments in a Solid, Physical Review **69**, 37 (1946).

[4] P. Reinhold, , Ph.D. thesis (2019).

[5] N. E. Frattini, U. Vool, S. Shankar, A. Narla, K. M. Sliwa, and M. H. Devoret, 3-wave mixing Josephson dipole element, Applied Physics Letters **110**, 222603 (2017).

[6] V. Sivak, N. Frattini, V. Joshi, A. Lingenfelter, S. Shankar, and M. Devoret, Kerr-Free Three-Wave Mixing in Superconducting Quantum Circuits, Physical Review Applied **11**, 054060 (2019).

[7] C. Wang, Y. Y. Gao, P. Reinhold, R. W. Heeres, N. Ofek, K. Chou, C. Axline, M. Reagor, J. Blumoff, K. M. Sliwa, L. Frunzio, S. M. Girvin, L. Jiang, M. Mirrahimi, M. H. Devoret, and R. J. Schoelkopf, A Schrödinger cat living in two boxes, Science **352**, 1087 (2016).

[8] B. J. Chapman, S. J. de Graaf, S. H. Xue, Y. Zhang, J. Teoh, J. C. Curtis, T. Tsunoda, A. Eickbusch, A. P. Read, A. Koottandavida, S. O. Mundhada, L. Frunzio, M. Devoret, S. Girvin, and R. Schoelkopf, High-On-Off-Ratio Beam-Splitter Interaction for Gates on Bosonically Encoded Qubits, PRX Quantum **4**, 020355 (2023).

[9] J. D. Teoh, P. Winkel, H. K. Babla, B. J. Chapman, J. Claes, S. J. de Graaf, J. W. O. Garmon, W. D. Kalfus, Y. Lu, A. Maiti, K. Sahay, N. Thakur, T. Tsunoda, S. H. Xue, L. Frunzio, S. M. Girvin, S. Puri, and R. J. Schoelkopf, Dual-rail encoding with superconducting cavities, Proceedings of the National Academy of Sciences **120**, e2221736120 (2023).