A mid-circuit erasure check on a dual-rail cavity qubit using the joint-photon number-splitting regime of circuit QED

Stijn J. de Graaf,[1,2,*] Sophia H. Xue,[1,2,*] Benjamin J. Chapman,[1,2] James D. Teoh,[2,3,4] Takahiro Tsunoda,[1,2] Patrick Winkel,[1,2] John W. O. Garmon,[2,3] Kathleen M. Chang,[1,2] Luigi Frunzio,[1,2,4] Shruti Puri,[1,2] and Robert J. Schoelkopf[1,2,3,4,†]

[1]*Department of Applied Physics, Yale University, New Haven, Connecticut 06511, USA*
[2]*Yale Quantum Institute, Yale University, New Haven, Connecticut 06511, USA*
[3]*Department of Physics, Yale University, New Haven, Connecticut 06511, USA*
[4]*Quantum Circuits, Inc., 25 Science Park, New Haven, Connecticut 06511, USA*

Quantum control of a linear oscillator using a static dispersive coupling to a nonlinear ancilla underpins a wide variety of experiments in circuit QED. Extending this control to more than one oscillator while minimizing the required connectivity to the ancilla would enable hardware-efficient multi-mode entanglement and measurements. We show that the spectrum of an ancilla statically coupled to a single mode can be made to depend on the joint photon number in two modes by applying a strong parametric beamsplitter coupling between them. This 'joint-photon number-splitting' regime extends single-oscillator techniques to two-oscillator control, which we use to realize a hardware-efficient erasure check for a dual-rail qubit encoded in two superconducting cavities. By leveraging the beamsplitter coupling already required for single-qubit gates, this scheme permits minimal connectivity between circuit elements. Furthermore, the flexibility to choose the pulse shape allows us to limit the susceptibility to different error channels. We use this scheme to detect leakage errors with a missed erasure fraction of $(9.0 \pm 0.5) \times 10^{-4}$, while incurring an erasure rate of $2.92 \pm 0.01\%$ and a Pauli error rate of $0.31 \pm 0.01\%$, both of which are dominated by cavity errors.

INTRODUCTION

Controlling the state of an oscillator is a powerful resource, enabling the implementation of hardware efficient error-correcting codes for quantum computing [1–7], simulations of bosonic systems [8–12] and the generation of metrologically useful states for quantum-enhanced sensing [13–15]. In circuit quantum electrodynamics (cQED) [16, 17], where these oscillators take the form of standing modes in microwave resonators, most of the techniques developed for single-oscillator control [18–25] rely on a static dispersive coupling between the oscillator and a nonlinear ancilla qubit [26].

Moving beyond control of a single linear mode affords new capabilities, including the generation of multi-mode entanglement [27–29] and measurements of joint properties of multiple modes [30–33]. In the context of quantum error correction, it both enables gates between qubits encoded in individual modes [34, 35], as well as implementations of natively multi-mode error correcting codes, such as the pair-cat [36], Chuang-Leung-Yamamoto (CLY) [37], or dual-rail codes [38]. This may be done by complementing the dispersive control with tunable beamsplitter interactions [39], which swap states between oscillators, allowing the nonlinear ancilla to interact with each oscillator in turn.

Recent progress in generating stronger tunable beamsplitter interactions between high-Q cavities without compromising their long coherence times or introducing unwanted nonlinearity [40, 41] provides access to a regime where the inter-oscillator coupling strength exceeds the typical coupling strength to the nonlinear ancilla. This presents the opportunity to treat the coupled oscillators collectively. Thus a single ancilla, statically coupled to only one of the modes, can be used to measure either joint or individual properties of the combined system. Such operations include the recently proposed joint-parity measurement or erasure check of a dual-rail cavity qubit [42, 43].

In this article, we observe a new regime that emerges in the presence of an increasingly strong beamsplitter drive between two bosonic modes, which is an analog of dispersive number splitting [26], but for the total excitation number, N, shared between two modes. In this regime the spectra of the combined multi-mode system are well predicted by treating the parametrically coupled oscillators as a single spin with $S = N/2$ that has a dispersive coupling to the ancilla. This provides an intuitive picture for arbitrary N. We find a range of operating points that are now accessible where we can measure the total photon number in the coupled oscillators, without measuring the photon number in either cavity.

An important example of an operation enabled in this regime is a mid-circuit erasure check [33, 44], an essential ingredient for a dual-rail qubit [38] spanned by the single excitation manifold $|0, 1\rangle$ and $|1, 0\rangle$ of two superconducting cavity modes. Erasure qubits [33, 43–51] rely on detecting dominant leakage errors and resetting these states back into the codespace. Since the time and location of these (erasure) errors are known, erasure qubits yield high thresholds when embedded in a higher-level error-correcting code [48, 52, 53]. The dual-rail cavity

arXiv:2406.14621v1 [quant-ph] 20 Jun 2024

qubit is a prototypical example of this as its errors are dominated by detectable leakage to $|0,0\rangle$, with Pauli errors within the codespace much less likely [43].

We implement a minimally-invasive erasure check by only exciting the less-coherent ancilla transmon when the dual-rail qubit has already leaked from the codespace. This limits the probability of declaring an erasure due to errors in the measurement itself to $2.92\pm0.01\%$ per check, of which only $0.51\pm0.02\%$ are false positives due to transmon errors. With a missed-erasure probability of only $(9.0 \pm 0.5) \times 10^{-4}$, dual-rail logical states are also well-preserved, with a Pauli error rate of $0.31\pm0.01\%$, of which only $0.12 \pm 0.01\%$ is induced by transmon decoherence. As in the previous demonstration of mid-circuit erasure detection for a dual-rail cavity qubit by Koottandavida *et al.* [33] this scheme engineers an ancilla spectrum that distinguishes between computational states and the joint vacuum state, $|0,0\rangle$. However, it does so by making use of the tunable beamsplitter already required for single- and two-qubit gates and minimizes the connectivity between circuit elements. Realizing such a high-performance and minimally-invasive mid-circuit erasure check is a crucial capability for improving error correction [48] via erasure conversion.

In the following sections, we begin by describing the emergence of parametrically-activated joint-photon number-splitting, before showing its application in a dual-rail erasure check. An extension of this technique to a two-qubit gate is described in Appendix A.

RESULTS

Spectroscopic observation of joint-photon number-splitting

Combining two common elements of oscillator control, a tunable beamsplitter interaction between two oscillators and a fixed dispersive interaction between a single oscillator and an ancillary qubit, enables measurements and control conditioned on the total photon number of the system, via an extension of the strong dispersive regime of circuit QED [26]. The Hamiltonian of this system in the interaction picture is

$$\frac{\hat{\mathcal{H}}}{\hbar} = \frac{g_{\mathrm{bs}}(t)}{2} \left(e^{i\varphi}\hat{a}\hat{b}^\dagger + e^{-i\varphi}\hat{a}^\dagger\hat{b} \right) - \Delta\hat{b}^\dagger\hat{b} + \chi\hat{b}^\dagger\hat{b}\,|e\rangle\langle e|, \quad (1)$$

where the oscillator Alice and Bob modes are represented by lowering operators $\hat{a}$ and $\hat{b}$, and $|e\rangle$ is the excited state of the ancillary qubit (see Fig. 1a). The ancilla is coupled to only one of the modes with a dispersive interaction strength χ that is fixed, whereas the amplitude g_{bs}, phase φ and frequency detuning Δ of the beamsplitter drive are all controllable in time. The dispersive term (activated by intentionally exciting the ancilla out of its ground state) and the beamsplitter term are often used alternately: either narrow-bandwidth pulses on the ancilla enact photon-number-selective operations on a single mode or a beamsplitter routes states between modes [28]. However, when both terms are activated simultaneously and $g_{\mathrm{bs}} \sim \chi$, the number-split spectrum of the ancilla is modified to depend on the joint photon number in both oscillators.

The Hamiltonian in Eq. 1 is realized with a pair of superconducting microwave stub cavities as the oscillators [54] and a fixed-frequency transmon as the ancilla. The use of a SNAIL coupler [40, 55–57] situated between the cavities allows us to engineer a microwave-activated beamsplitter interaction with amplitude up to $|g_{\mathrm{bs}}|/2\pi = 2.05$ MHz, in excess of the dispersive coupling $|\chi|/2\pi = 1.07$ MHz, while preserving the coherence and linearity of the cavity modes.

The emergence of joint-photon number-splitting is revealed by probing the transmon spectrum in the presence of a variable-amplitude beamsplitter drive, as illustrated in Fig. 1b. Importantly, the beamsplitter drive is applied with a 'symmetric' frequency detuning, $\Delta = \chi/2$ — since Bob's frequency shifts by χ when the transmon is in $|e\rangle$, this condition ensures the beamsplitter drive is equally detuned from resonance when the transmon is in $|g\rangle$ or in $|e\rangle$. Because the Hamiltonian conserves the total oscillator photon number N and the average cavity T_1 is much longer than the pulse duration T_{p}, the dynamics may be considered separately for different values of N. Fig. 1c shows the resulting spectra for different initial states with $N = 0, 1$ and 2.

In the absence of a coupling between the cavities ($g_{\mathrm{bs}} = 0$) the transmon spectrum displays the familiar photon number-splitting regime, with transitions separated in frequency by χ per photon in the Bob mode but independent of the photon number in the uncoupled Alice mode. This can be seen from the linecuts at $g_{\mathrm{bs}} = 0$ (in the left panels of Fig. 1c) when initializing the cavities in different two-oscillator Fock states. However, in the presence of a strong beamsplitter drive ($|g_{\mathrm{bs}}| \sim |\chi|$) we enter the joint-photon number-splitting regime. Here, the transmon spectrum exhibits $2N + 1$ prominent transition lines, with the central transition at a frequency detuning $\delta\omega = N\chi/2$ becoming the dominant one at large values of g_{bs}.

We can compare these features to analytical predictions obtained using Schwinger's angular momentum operator formalism [42, 58], which considers the two cavities containing N total photons as a single spin with $S = N/2$ (see Fig. 1a). This formalism provides a way to map oscillator operators onto spin operators:

$$\hat{S}_x \equiv \frac{\hat{a}^\dagger\hat{b} + \hat{b}^\dagger\hat{a}}{2}, \qquad \hat{S}_y \equiv \frac{\hat{a}^\dagger\hat{b} - \hat{b}^\dagger\hat{a}}{2i},$$
$$\hat{S}_z \equiv \frac{\hat{a}^\dagger\hat{a} - \hat{b}^\dagger\hat{b}}{2}, \qquad \frac{N}{2}\hat{\mathbb{I}} \equiv \frac{\hat{a}^\dagger\hat{a} + \hat{b}^\dagger\hat{b}}{2}. \quad (2)$$

For a symmetrically-detuned beamsplitter drive ($\Delta = \chi/2$) this gives a spin Hamiltonian for this system:

$$\frac{\hat{\mathcal{H}}}{\hbar} = g_{\mathrm{bs}}\hat{S}_x + \frac{\chi}{2}\hat{S}_z\hat{\sigma}_z - \frac{N\chi}{4}\hat{\sigma}_z, \quad (3)$$
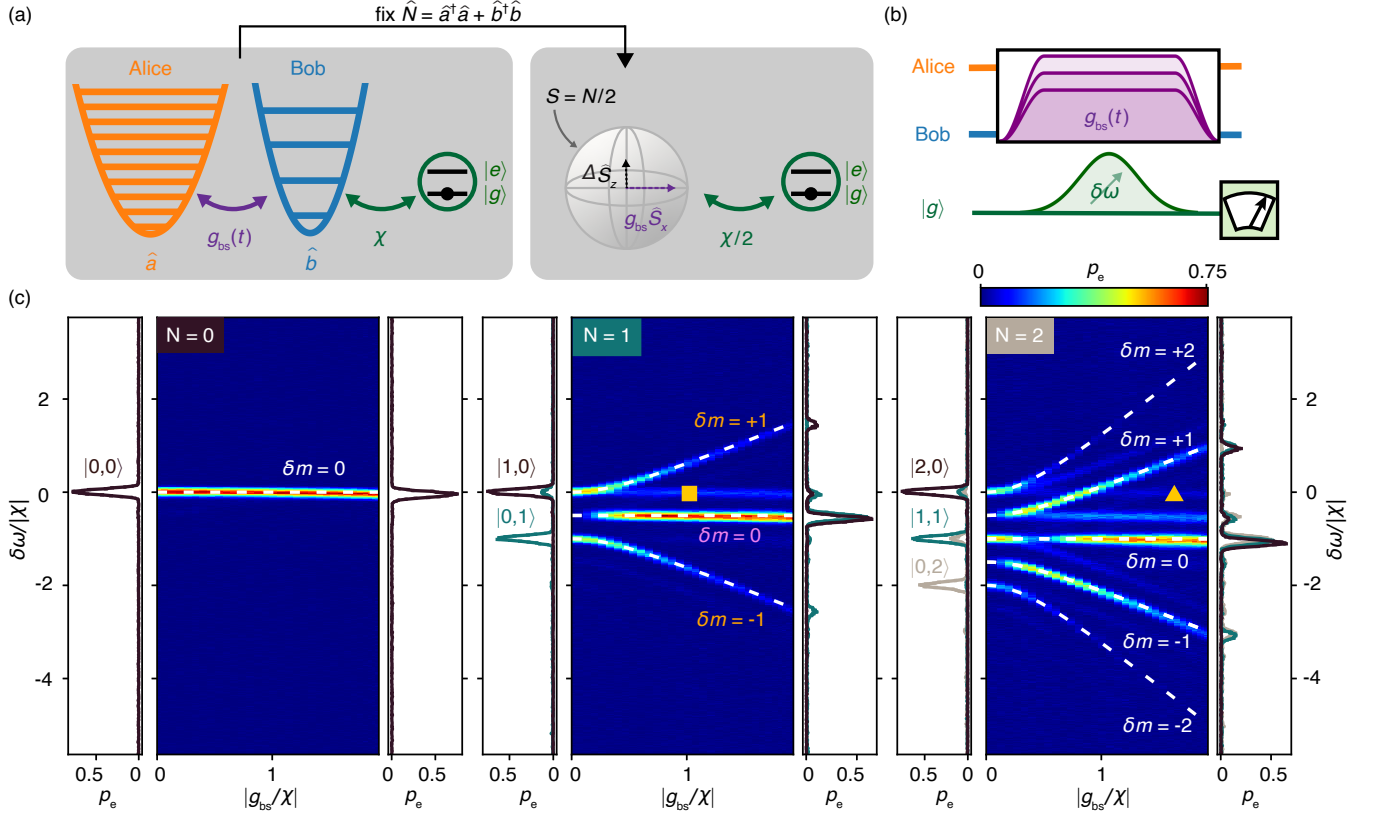
FIG. 1. **Ancilla spectroscopy in the presence of a beamsplitter drive.** (a) System schematic showing two oscillators coupled by a tunable beamsplitter interaction and an ancilla qubit statically coupled to one oscillator via a fixed dispersive interaction. Two coupled oscillators with total photon number N are equivalent to a spin with $S = N/2$. (b) Pulse sequence used for spectroscopy experiment. (c) Transmon spectra in the presence of increasing beamsplitter drive amplitude for input oscillator states with fixed total photon number N. Colorplots show initial states $|0,0\rangle$ ($N = 0$), $\frac{|0,1\rangle+|1,0\rangle}{\sqrt{2}}$ ($N = 1$) and $\frac{|0,2\rangle + \sqrt{2}|1,1\rangle + |2,0\rangle}{2}$ ($N = 2$). Predicted transitions (dashed white lines) are labeled by the change in the magnetic quantum number δm in the oscillator-spin model (see Fig. 2). Left (right) panels show spectra for all initial two-oscillator Fock states in each N-photon manifold at the lowest (highest) value of $|g_{bs}/\chi|$ shown in the colorplots. The yellow square (triangle) in the middle (right) panel indicates operating point for the erasure check (CPHASE gate described in Appendix A).

where $\hat{\sigma}_z = |g\rangle\langle g| - |e\rangle\langle e|$ is the Pauli-Z operator for the ancilla and the beamsplitter drive phase is chosen to be $\varphi = 0$. When the amplitude of the spectroscopy drive on the qubit is small relative to $\max(|g_{bs}|, |\chi|)$, its effect in the Hamiltonian may be treated as a perturbation. $\hat{\mathcal{H}}$ is then block-diagonal with respect to the ancilla subspace, allowing us to separately consider Hamiltonians conditioned on the ancilla state, $\hat{\mathcal{H}}_g \equiv \langle g|\hat{\mathcal{H}}|g\rangle$ and $\hat{\mathcal{H}}_e \equiv \langle e|\hat{\mathcal{H}}|e\rangle$:

$$\hat{\mathcal{H}} = \hat{\mathcal{H}}_g |g\rangle\langle g| + \hat{\mathcal{H}}_e |e\rangle\langle e|. \qquad (4)$$

Diagonalizing $\hat{\mathcal{H}}_g$ and $\hat{\mathcal{H}}_e$ gives spin eigenstates (representing the state of the coupled oscillators) before and after the ancilla transition from $|g\rangle$ to $|e\rangle$, with each Hamiltonian defining a spin quantization axis along which these eigenstates are aligned.

In this spin picture, the application of the beamsplitter drive is analogous to a transverse magnetic field. As the beamsplitter strength increases, the energy eigenstates

are no longer oscillator Fock states ($\hat{S}_z$ eigenstates) but more closely resemble symmetric and antisymmetric superpositions of these Fock states ($\hat{S}_x$ eigenstates). Meanwhile the eigenenergies, shown for total photon number $N = 1$ in Fig. 2a, display a zero-field splitting by $\chi/2$ and a high-field Zeeman-like splitting that approaches g_{bs} as the transverse field is increased (i.e. by increasing the beamsplitter strength). The eigenstates are labelled by their projection along the quantization axis:

$$\hat{\mathcal{H}}_g |m_g\rangle = \left(-\frac{N\chi}{4} + m_g \hbar\Omega\right)|m_g\rangle, \qquad (5)$$

$$\hat{\mathcal{H}}_e |m_e\rangle = \left(+\frac{N\chi}{4} + m_e \hbar\Omega\right)|m_e\rangle, \qquad (6)$$

where the magnetic quantum numbers can take the values,

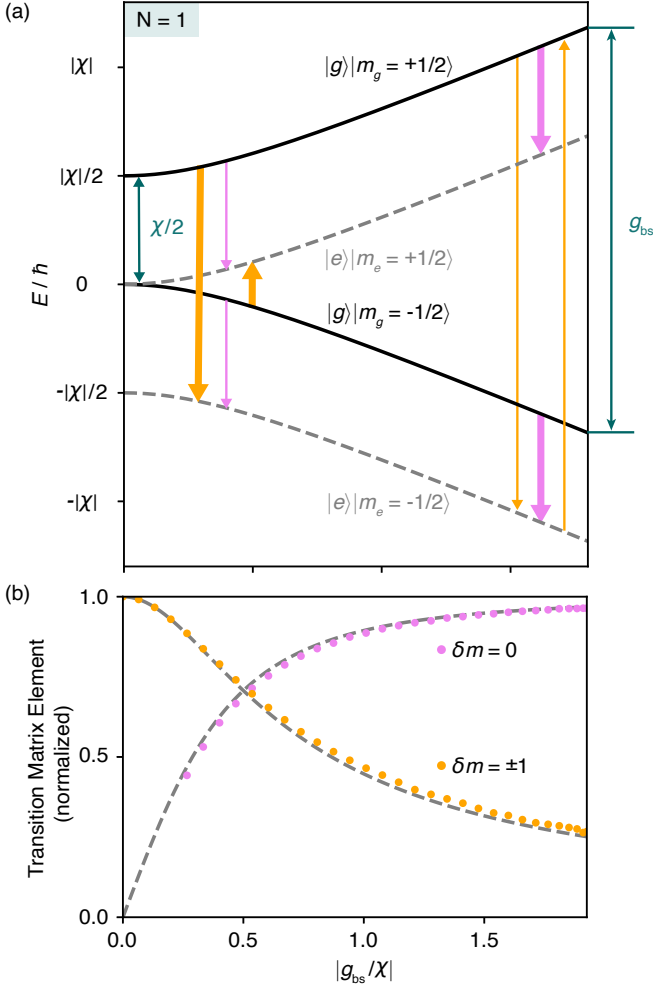$$m_{g/e} = -S, \ldots, S = -\frac{N}{2}, \ldots, \frac{N}{2},$$

FIG. 2. **Ancilla transition matrix elements in the presence of beamsplitter.** (a) Energy level diagram for the ancilla-oscillator system with $N = 1$ photon shared between the two oscillators. States are labeled by their transmon state ($|g\rangle$ or $|e\rangle$) and oscillator-spin state (magnetic quantum number $m_{g/e} = \pm 1/2$). Arrow thicknesses illustrate the strength of transition matrix elements at low and high beamsplitter amplitudes. (b) Measured transition matrix elements along central ($\delta m = 0$, pink) and lower ($\delta m = -1$, orange) transitions of $N = 1$ spectrum, obtained from ancilla Rabi oscillation rates. Dashed lines show predictions for normalized transition matrix elements obtained from the spin model.

and Ω can be interpreted as a Larmor frequency:

$$\Omega = \sqrt{g_{bs}^2 + \left(\frac{\chi}{2}\right)^2}. \quad (7)$$

The energy difference between each possible pair of $\hat{\mathcal{H}}_g$ and $\hat{\mathcal{H}}_e$ eigenstates allows us to accurately predict the $2N + 1$ unique transition frequencies observed in the spectra:

$$\omega_{\delta m} = \frac{N\chi}{2} + \delta m\,\Omega, \quad \delta m \equiv m_e - m_g = -N, \ldots, N. \quad (8)$$

These frequencies are shown as white dashed lines in Fig. 1c, and show good agreement with the observed spectra [59]. Faint features corresponding to transitions in the $N - 1$ photon manifold can also be seen due to photon loss during the spectroscopy pulse. The symmetric beamsplitter detuning condition $\Delta = \chi/2$ provides a unique operating point where each transition frequency has a degeneracy $(N + 1) - |\delta m|$. (In Appendix D, we show how shifting away from this point breaks the degeneracies between these transitions, yielding $(N + 1)^2$ unique frequencies.)

Furthermore, the strength of each transition is proportional to the overlap between the initial ($\hat{\mathcal{H}}_g$) and final ($\hat{\mathcal{H}}_e$) eigenstates. As g_{bs} is increased, both sets of eigenstates start to align along the x-axis, with increasing overlap between states where $m_g = m_e$. As a result, off-central transitions with $\delta m \neq 0$ become suppressed in favor of transitions with $\delta m = 0$, as can be seen from their changing brightness in Fig. 1c. We verify this quantitatively for transitions in the $N = 1$ manifold by measuring the rate of ancilla Rabi oscillations across a range of $|g_{bs}/\chi|$ values (described in Appendix C). The measured rates for the $\delta m = 0$ and $\delta m = \pm 1$ transitions (proportional to the transition matrix element) are shown in Fig. 2b and agree well with the model predictions.

The predictions from this spin model explain the emergence of a joint-photon number-splitting regime when $g_{bs} > \chi$, where the qubit spectrum depends on the total number of photons in both oscillators. This is most clear in linecuts of the spectra at the largest values of g_{bs}/χ in Fig. 1c, where the central $\delta m = 0$ transition line dominates. Independent of which state we initialize in the oscillators, we see a single dominant transition frequency at $\omega_0 = N\chi/2$ for each value of N. This can be interpreted as the indistinguishability of photons in Alice and Bob from the perspective of the qubit when the beamsplitter interaction is stronger than the dispersive shift. The extent to which the qubit can still distinguish the two oscillators can be seen in the much smaller satellite $\delta m = \pm 1$ peaks, at a detuning of $\pm g_{bs}$, which would be further suppressed at higher beamsplitter amplitudes.

In the following section we show how understanding these joint-photon number-split spectra provides a way to extend established single-oscillator control techniques to operations on multiple oscillators.

Mid-circuit erasure check of a dual-rail qubit

We use the $|g_{bs}| \gtrsim |\chi|$ regime to construct a mid-circuit erasure check for a dual-rail qubit encoded in two superconducting cavities. The basis states for this encoding are $|0_{\mathrm{L}}\rangle \equiv |1, 0\rangle$ and $|1_{\mathrm{L}}\rangle \equiv |0, 1\rangle$ and their dominant error channel is photon loss to the common $|0, 0\rangle$ leakage state, which we hope to detect (see Fig. 3a). Detecting these leakage errors (in order to convert them to erasures) while preserving logical information in states that have not suffered such an error is an essential task in this
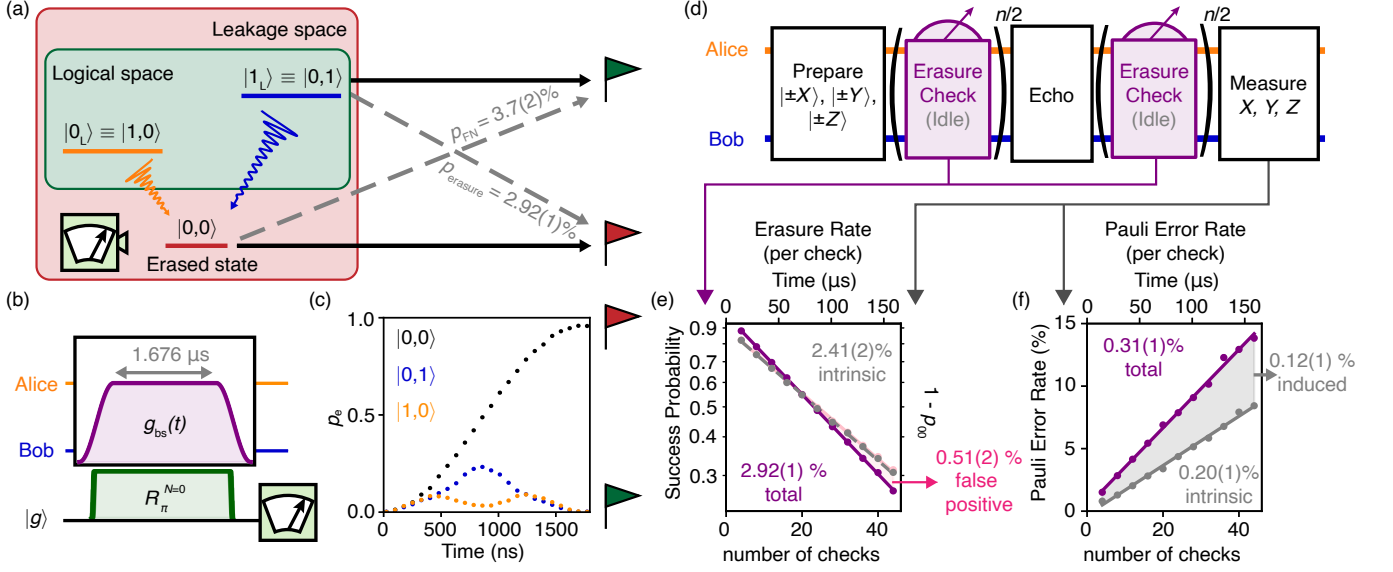
FIG. 3. **Characterizing a dual-rail mid-circuit erasure check.** (a) Logical code space of a dual-rail encoded qubit indicating decay to the leakage state $|0,0\rangle$, which we seek to detect and convert to erasure errors. Black arrows show the ideal mapping processes shown while gray arrows indicate misassignment processes that lead to false negative (FN) errors and erasure errors respectively. We measure a false negative error rate of $p_{\text{FN}} = 3.7(2)\%$ per erasure check and a total erasure error rate of $p_{\text{erasure}} = 2.92(1)\%$ per erasure check. Green (red) flag indicates erasure check reporting "no erasure" ("erasure detected"). (b) Pulse sequence used to implement the erasure check, consisting of simultaneous flat-top pulses on both the transmon and the beamsplitter drive, followed by transmon readout. (c) Measured transmon trajectories during the erasure check for logical and leakage states. (d) Pulse sequence used to measure state fidelity after n repeated erasure checks (or n periods of idling for the duration of an erasure check) on the encoded dual-rail qubit. A single echo pulse (a cavity SWAP operation) is performed halfway through the sequence in order to suppress the effect of no-jump backaction resulting from different cavity decay rates. (e) Success probability for passing n consecutive checks (purple), and probability of remaining in logical subspace ($1 - p_{00}$) after n checks, unconditioned on erasure check results, with (gray) and without (pink) the beamsplitter pump applied. (f) Total Pauli error rate when performing erasure checks (purple) or idling (gray). Lines show linear fits.

architecture.

A key requirement of the erasure check is that it does not introduce additional errors. Firstly, since the check may be performed multiple times per round of syndrome measurements, erasure errors during the check must be minimized for the code to operate below the error-correction threshold of the higher-level code, such as the surface code [60–62]. Secondly, Pauli errors must remain much less likely than erasure errors to preserve the bias that enables a high threshold [48]. Both of these criteria require a way of limiting errors induced by the less-coherent ancilla transmon.

A powerful and established technique for measuring bosonic modes while preventing the ancilla from polluting the logical state is to use a three-level ancilla and to apply a microwave drive to ensure that the dispersive shift is unchanged when the ancilla is in $|e\rangle$ or in $|f\rangle$ (often known as χ-matching) [23]. This erasure check, however, bypasses the need for χ-matching drives since it only needs to catch events when the logical information has already been lost.

The modification of the transmon spectrum in the presence of a beamsplitter drive allows us to perform an erasure check with low susceptibility to transmon errors.

In the number-splitting regime for a single oscillator, a narrow-bandwidth pulse on the ancilla enables a π-pulse conditioned on zero photons in the oscillator. Likewise, the joint-photon number-splitting regime allows us to excite the transmon if and only if there are zero total photons in two oscillators (i.e. when the dual-rail qubit has leaked to $|0,0\rangle$).

For transmon pulses selective on $N = 0$, we do not require $g_{\text{bs}} \gg \chi$ but only that the $N = 0$ transition is sufficiently detuned from all other transitions we would like to avoid. In the context of the dual-rail encoding, these are the transitions in the $N = 1$ manifold (corresponding to the logical subspace), for which the detuning saturates at $\chi/2$ for $|g_{\text{bs}}/\chi| \geq \sqrt{3}/2$ (see Fig. 1c). This establishes a wide range of values for $|g_{\text{bs}}/\chi|$ that can enact an erasure check and allows for a more flexible implementation than the erasure check proposed in [42, 43], based on measuring the joint photon number parity, where $|g_{\text{bs}}/\chi|$ is fixed (see Appendix J for a detailed comparison of the two approaches).

Furthermore, a variety of pulse shapes on the transmon drive may be used to perform the erasure check while trading off susceptibility to different error sources. For example, a long, highly frequency-selective Gaussian

pulse will limit transmon excitation when in the $N = 1$ manifold, at the expense of more idling errors during the check. To minimize the combined rate of transmon errors and idling errors, we use a shorter square pulse with a duration $T_p = 1.820$ μs while applying a beamsplitter drive with amplitude $|g_{bs}/\chi| = 1.04$ (as indicated by the square symbol in Fig. 1c). These values of g_{bs} and T_p are calibrated to ensure that when starting in the dual-rail logical subspace, both the transmon and oscillators return to their initial states at the end of the sequence despite the finite selectivity of the transmon pulse (see Fig. 3b,c; for detailed calibration procedures, see Appendix E). While this scheme places no special requirement on the adiabaticity of the beamsplitter drive ramp, we use a 120 ns cosine-shaped ramp to stay well within the bandwidth limit imposed by the on-chip Purcell filter through which the beamsplitter drive is applied. We align the center of this ramp in time with the center of the 24 ns cosine ramp on the transmon pulse, which we find from simulation ensures optimal performance.

The trajectories of the transmon state for different input oscillator states in Fig. 3c showcase the operation of the erasure check: while the transmon ends in $|e\rangle$ when the oscillators are in $|0, 0\rangle$ (thus flagging an erasure), it returns to $|g\rangle$ for input states $|0, 1\rangle$ and $|1, 0\rangle$. The relatively small area under the $|0, 1\rangle$ and $|1, 0\rangle$ curves indicates that the transmon is less likely to be excited in these cases and so transmon decay and dephasing errors are less likely to induce false positives or logical Pauli errors. Meanwhile the relatively small difference between the $|0, 1\rangle$ and $|1, 0\rangle$ curves indicates the limited extent to which transmon dephasing errors allow the environment to distinguish between different logical states, thereby inducing Pauli errors when they occur [63].

We test the fraction of leakage errors caught by the check by preparing $|0, 0\rangle$ and performing a single mid-circuit erasure check followed by destructive photon-number-selective measurements of each cavity [44]. When post-selecting on the final state remaining in $|0, 0\rangle$, we find a false negative rate $p_{FN} = 3.7 \pm 0.2\%$, consistent with results of master equation simulations using the physical transmon error rates in our system. While the false negative rate *is* sensitive to transmon errors, in the operation of an error-correcting code the actual probability of missing an erasure ($p_{miss} = p_{erasure} \times p_{FN} \sim 10^{-3}$) is also multiplied by the small probability that an erasure has been suffered since the previous check. We find that this rate of false negatives is still admissible for high fault-tolerant thresholds in the surface code. In fact, when performing erasure checks after every two qubit gate, the surface code threshold (per step consisting of a gate plus an erasure check) with such a level of false negatives is $p_{th} = 3.71 \pm 0.02\%$, which is well in excess of the 1% Pauli noise threshold. As a point of comparison, if the false negative rate were 0, our threshold would be $p_{th} = 3.79 \pm 0.02\%$, signifying that the threshold is minimally affected by the rate of false negatives. In both cases, we have assumed that 90% of errors are

erasures. For more details on these simulations, refer to Appendix F. As such, false negatives, which have a very small impact on p_{th}, are less costly than false positives, which contribute to the overall physical error rate that should be kept small with respect to p_{th}. This informs the design of the erasure check where the transmon is excited only if $N = 0$.

A single experiment is used to evaluate the performance of the mid-circuit erasure check when acting on the code space, in terms of both the erasure error rate and the Pauli error rate induced by the check. This tomography sequence, shown in Fig. 3d, consists of preparing the six dual-rail cardinal states $|\pm X\rangle, |\pm Y\rangle$, and $|\pm Z\rangle$, repeating the erasure check n times and then measuring the logical operators $\hat{X}_L, \hat{Y}_L$ and $\hat{Z}_L$ using photon-number-selective readout of each cavity (see Appendix G). To separate the contribution from idling errors, we perform the same sequence but replacing each erasure check with a delay of the same duration. Beamsplitter pulses are used to enact single-qubit gates on the dual-rail qubit, allowing us to prepare states on the equator of the Bloch sphere and perform logical measurements. An echo pulse is added to remove the effect of no-jump back-action at long times. When the decay rate in each cavity is different (as is especially the case here, with $T_{1,a} = 347$ μs and $T_{1,b} = 109$ μs), post-selecting on no photon loss leads to a deterministic polarization towards one pole of the Bloch sphere and results in an approximately-Gaussian envelope on both the idling and erasure check data on a timescale set by the difference in the T_1 times [43]. Introducing the echo therefore allows us to better resolve errors induced by the check itself.

The erasure rate is extracted by looking at the success probability (i.e. the likelihood of passing n successive checks) as a function of n, averaged over all input states (see Fig. 3e). The slope of the exponential decay shows a total erasure rate per check $p_{erasure} = 2.92 \pm 0.01\%$. To determine what fraction of these flagged events are 'intrinsic' erasures due to photon loss to $|0, 0\rangle$, as opposed to false positive events predominantly due to transmon errors, we may instead ignore the results of the mid-circuit checks and only ask how often the end-of-line measurement yields $|0, 0\rangle$. This gives a photon loss rate of $2.41 \pm 0.02\%$ per check, consistent with the value obtained when idling for the same duration, $2.43 \pm 0.02\%$, indicating that the erasure check does not induce additional photon loss in the cavities. This value is the 'intrinsic' erasure rate, from which we infer that the remaining $0.51 \pm 0.02\%$ of erasures are false positive errors. That the majority of detected erasures are due to photon loss in the high-Q cavities, rather than transmon errors, shows the ability of the scheme to tolerate decoherence in the transmon ancilla. Furthermore, the flexibility to choose different transmon pulse shapes allows us to trade off between 'intrinsic erasures' (suppressed with a shorter pulse) and false positives (suppressed with a longer, more selective pulse), which can therefore be tailored to the relative decoherence rates in the system. This flexibility,

not present in the joint-parity-based approach [42], also allows us to trade off false positives ($p_{\text{FP}} = 0.51\%$) for less-costly false negatives ($p_{\text{FN}} = 3.7\%$).

To verify that this mid-circuit erasure check preserves the logical qubit state, we probe the fidelity of all 6 dual-rail cardinal states conditioned on passing n successive erasure checks. Fitting the slope of the post-selected Pauli error probabilities (obtained from the average state fidelity; see Appendix H) as a function of measurement rounds n allows us to precisely resolve the error rate for a single round (Fig. 3f). We find that the logical information is well-preserved during the course of the measurement, with an overall Pauli error per check $p_{\text{Pauli}} = 0.31 \pm 0.01\%$, compared to $0.20 \pm 0.01\%$ when idling. This indicates that background cavity errors dominate, with transmon-induced Pauli errors contributing at most the remaining $0.12 \pm 0.01\%$, highlighting the robustness of the scheme against transmon decoherence. We note that while the echo pulse used to mitigate against no-jump backaction will also reduce the effect of low-frequency dephasing noise on the intrinsic error rate, errors induced by transmon decoherence should not be affected.

Taken together, these results demonstrate the efficacy of a flexible hardware-efficient mid-circuit erasure check, making use of only the beamsplitter interaction (used for gates) and dispersive transmon coupling to a single mode (used for state preparation). This check preserves a large ratio of the erasure error rate to the Pauli error rate, with both quantities remaining dominated by 'intrinsic' errors associated with the hardware. With improvement of the mode coherences closer to state-of-the-art values [64–66], we expect this scheme to yield significantly below-threshold performance. Indeed, with the same χ, g_{bs} and readout duration τ_{RO} but with a transmon T_1 and T_ϕ of 200 μs and an average cavity T_1 of 1000 μs, we predict from master equation simulations a total erasure rate $p_{\text{erasure}} = 0.49\%$ and a transmon-induced Pauli error rate of 0.035%, at which point shot-noise dephasing from photons in the readout resonator starts playing a larger role (see Appendix K). Increasing χ and reducing τ_{RO} to reduce the check duration provides another way to improve performance. Separately, increasing g_{bs} would help reduce Pauli errors from transmon dephasing by making the dual-rail logical states more indistinguishable during the check [63] (see Appendix I).

DISCUSSION

These results show that a strong beamsplitter drive, when paired with dispersive coupling to a single mode, offers a versatile means of measuring and controlling multiple oscillators. The tunable beamsplitter allows for switching between single-cavity operations (accessible when $g_{\text{bs}} = 0$) and their equivalent multi-cavity operations on the total photon number (accessible when $g_{\text{bs}} \gtrsim \chi$). As an important example of this, we have

demonstrated how photon-number selective measurements can be extended to joint-photon-number selective measurements to enable a mid-circuit erasure check for dual-rail qubits. Further examples of this principle include the extension of the "selective number arbitrary phase" (SNAP) gate to a joint-SNAP-like gate selective on the photon number in two cavities, enabling a tunable CPHASE(θ) gate for two dual-rail qubits (see Appendix A).

The mid-circuit erasure check for dual-rail qubits enabled by joint-photon-number selective measurements represents a new, hardware-efficient way of performing the essential ingredient for an error-correcting surface code with dual-rail cavity qubits. The design of this check, which asks whether the joint photon number is zero (as opposed to a joint-parity check which asks whether it is even), leverages two important features of erasure errors: that they are rare, and that the state need not be preserved once an erasure is detected. Therefore by only minimally exciting the ancilla when in the dual-rail code space, it ensures that the contribution of transmon-induced erasures and Pauli errors is subdominant, at the cost of more false negatives, to which the code is more tolerant. One caveat is that this check does not catch rare but damaging heating events, although this could be mitigated with the addition of a selective pulse acting on the two-photon manifold.

Improvement of the mode coherences to state-of-the-art values should enable performance substantially below the erasure threshold for the surface code. With high-fidelity state preparation, logical erasure-detected measurements, single-qubit gates and hardware-efficient mid-circuit erasure checks demonstrated, key next steps will include a fully error-detectable two-qubit gate, as well as fast qubit reset to turn leakage detection into erasure conversion.

METHODS

Device

The device used for this experiment was previously used in Chapman et al. [40] and includes two superconducting $\lambda/4$ post cavities machined from 99.999%-purity Aluminum as the two oscillators. Into each of these cavities we insert an EFG sapphire chip supporting a transmon qubit, a readout resonator and a Purcell filter. These are used to prepare and readout states in the two oscillators. The transmon coupled to Bob's cavity is used to determine the spectrum in Fig. 1 and to perform the erasure check in Fig. 3. Detailed characterization of system parameters are shown in Table II and discussed in Appendix M.

To provide the beamsplitter interaction between these cavities, we add a capacitively-shunted superconducting nonlinear asymmetric inductive element (SNAIL) [56], a superconducting loop with three nominally-identical

Josephson junctions (each with Josephson energy $E_{\text{J}}/h = 90.0 \pm 0.3$ GHz) in series on one arm of the loop and a single junction (with Josephson energy $\alpha = 0.147 \pm 0.001$ times smaller) on the other. The capacitive shunt is implemented by adding leads to the two ends of the SNAIL connecting to large capacitor pads. This gives the SNAIL a charging energy $E_{\text{C}}/h = 177 \pm 2$ MHz, while the leads also contribute a series linear inductance with energy $E_{\text{L}}/h = 64 \pm 2$ GHz. This circuit is patterned onto a sapphire chip which is inserted into a tunnel that intersects both cavities, thereby generating a linear coupling between the SNAIL mode and each of the oscillator modes. Passing a DC magnetic flux through the SNAIL loop (delivered via a superconducting flux transformer) modifies the potential of the SNAIL, allowing us to generate a third-order nonlinearity. By applying a single microwave drive that couples to the charge operator of the SNAIL mode, at a frequency equal to the difference between the two cavity frequencies, we can generate the beamsplitter interaction via a three-wave-mixing process. The frequency detuning from this resonance condition is the detuning Δ indicated in Eq. 1, and the amplitude and phase of the drive determine g_{bs} and φ.

Any magnetic flux bias Φ_{ext} besides a half integer multiple of Φ_0 gives the SNAIL a third-order nonlinearity and at low drive powers, the magnitude of this nonlinearity determines the ratio between the applied drive amplitude and g_{bs}. However, as was seen previously [40], at larger drive powers, this linear relationship breaks down and there exists a maximum value of g_{bs} at each Φ_{ext}. The flux bias used in this experiment ($\Phi_{\text{ext}} = 0.334\Phi_0$) was chosen to give the highest maximum value of g_{bs} while avoiding any unwanted resonances at any value of g_{bs} as it is increased to this maximum value. This therefore provides access to the widest range of values for $|g_{\text{bs}}/\chi|$. Appendix B provides details on how we measure and calibrate the value of g_{bs}.

The aluminum package is mounted to the mixing chamber plate of a Bluefors XLD400sl dilution refrigerator, at a temperature of 8 mK.

Oscillator-spin model

As in Tsunoda et al. [42], we can express the system Hamiltonian in Eq. 1 in terms of the Schwinger angular momentum operators [58] (defined in Eq. 2):

$$\frac{\hat{\mathcal{H}}}{\hbar} = g_{\text{bs}} \cos \varphi \hat{S}_x - g_{\text{bs}} \sin \varphi \hat{S}_y - \left(\Delta - \chi |e\rangle \langle e| \right) \left(\frac{\hat{N}}{2} - \hat{S}_z \right). \quad (9)$$

As indicated in Eq. 4, $\hat{\mathcal{H}}$ is block-diagonal in the ancilla subspace, allowing us to separately consider the ancilla-state-dependent Hamiltonians $\hat{\mathcal{H}}_g$ and $\hat{\mathcal{H}}_e$.

This Hamiltonian also conserves total photon number $\hat{N}$ and so if we restrict ourselves to the N-photon man-

ifold, we may make the replacement $\hat{N} \to N$. In the N-photon manifold, the Schwinger angular momentum operators defined above are the standard spin-$N/2$ operators, with eigenvalues running from $-N/2$ to $N/2$. We may combine these spin operators into a single spin operator $\hat{S}_{\hat{\Omega}}$ pointing along a quantization axis $\hat{\Omega}$ such that the Hamiltonian is diagonal. Writing the Hamiltonian separately when the qubit is in $|g\rangle$ and in $|e\rangle$ gives:

$$\frac{\hat{\mathcal{H}}_g}{\hbar} = \frac{-N\Delta}{2} + |\vec{\Omega}_g| \hat{S}_{\hat{\Omega}_g}, \quad (10)$$

$$\frac{\hat{\mathcal{H}}_e}{\hbar} = \frac{N(\chi - \Delta)}{2} + |\vec{\Omega}_e| \hat{S}_{\hat{\Omega}_e}, \quad (11)$$

where

$$\vec{\Omega}_g = \begin{pmatrix} g_{\text{bs}} \cos \varphi \\ -g_{\text{bs}} \sin \varphi \\ \Delta \end{pmatrix}, \quad (12)$$

$$\vec{\Omega}_e = \begin{pmatrix} g_{\text{bs}} \cos \varphi \\ -g_{\text{bs}} \sin \varphi \\ \Delta - \chi \end{pmatrix}. \quad (13)$$

Transition frequencies

Since the Hamiltonian is diagonal in each case, we may simply read off the eigenenergies as

$$\left\{ \frac{E_g}{\hbar} \right\} = \left\{ -\frac{N\Delta}{2} + m_g \sqrt{g_{\text{bs}}^2 + \Delta^2} \right\}, \quad (14)$$

$$\left\{ \frac{E_e}{\hbar} \right\} = \left\{ \frac{N(\chi - \Delta)}{2} + m_e \sqrt{g_{\text{bs}}^2 + (\Delta - \chi)^2} \right\}, \quad (15)$$

where $m_g, m_e = -N/2, ..., N/2$. If we take all the differences between these energies, we obtain $(N+1)^2$ transition frequencies,

$$\{\omega\} = \left\{ \frac{N\chi}{2} + m_e \sqrt{g_{\text{bs}}^2 + (\Delta - \chi)^2} \right. \\ \left. -m_g \sqrt{g_{\text{bs}}^2 + \Delta^2} \right\}. \quad (16)$$

The case where $\Delta = \chi/2$, the condition required to observe joint-photon number-splitting, is special as many of the transition frequencies become degenerate, leaving $2N + 1$ unique transitions:

$$\{\omega\}_{\Delta = \chi/2} = \left\{ \frac{N\chi}{2} + \delta m \sqrt{g_{\text{bs}}^2 + \left(\frac{\chi}{2}\right)^2} \right\} \quad (17)$$

where $\delta m \equiv m_e - m_g = -N, ..., N$.

Transition matrix elements

From the eigenstates of $\mathcal{H}_g$ and $\mathcal{H}_e$, which we denote as $|m_g\rangle$ and $|m_e\rangle$, we may also find the matrix elements

for each transition:

$$|M|_{m_g \to m_e} = \left| \langle m_g| \langle g| \left(\frac{\epsilon}{2} |g\rangle \langle e| + \frac{\epsilon^*}{2} |e\rangle \langle g| \right) |e\rangle |m_e\rangle \right|$$

$$= \frac{|\epsilon|}{2} |\langle m_g|m_e\rangle| \tag{18}$$

The overlap between the initial and final oscillator states therefore tells us the strength of each transition. The overlap between a spin-$N/2$ eigenstate with magnetic quantum number m_g along one axis and another eigenstate with magnetic quantum number m_e along another axis at an angle $\delta\theta$ from the first is given by the Wigner (small) d-matrix [67]:

$$|\langle m_g|m_e\rangle| = \left| d^{N/2}_{m_g,m_e}(\delta\theta) \right|. \tag{19}$$

In our case, the angle between the spin axes $\hat{\Omega}_g$ and $\hat{\Omega}_e$ is

$$\delta\theta = \arctan\left(\frac{\Delta}{g_{\text{bs}}}\right) - \arctan\left(\frac{\Delta - \chi}{g_{\text{bs}}}\right). \tag{20}$$

For the specific case shown in Fig. 2b, where $N = 1$ and $\Delta = \chi/2$, the angle difference becomes $\delta\theta = 2\arctan\left(\chi/2g_{\text{bs}}\right)$. The matrix element for the central transition is therefore

$$|M_{\delta m=0}| \propto \left| d^{\frac{1}{2}}_{\frac{1}{2},\frac{1}{2}}(\delta\theta) \right| = \left| d^{\frac{1}{2}}_{-\frac{1}{2},-\frac{1}{2}}(\delta\theta) \right| \tag{21}$$

$$= \cos\left(\frac{\delta\theta}{2}\right) \tag{22}$$

$$= \cos\left(\arctan\left(\frac{\chi}{2g_{\text{bs}}}\right)\right) \tag{23}$$

$$= \frac{g_{\text{bs}}}{\sqrt{g_{\text{bs}}^2 + \left(\frac{\chi}{2}\right)^2}}. \tag{24}$$

Likewise for the off-central transitions,

$$|M_{\delta m=\pm 1}| \propto \left| d^{\frac{1}{2}}_{\frac{1}{2},-\frac{1}{2}}(\delta\theta) \right| = \left| d^{\frac{1}{2}}_{-\frac{1}{2},\frac{1}{2}}(\delta\theta) \right| \tag{25}$$

$$= \sin\left(\frac{\delta\theta}{2}\right) \tag{26}$$

$$= \frac{\left(\frac{\chi}{2}\right)}{\sqrt{g_{\text{bs}}^2 + \left(\frac{\chi}{2}\right)^2}}. \tag{27}$$

DATA AND CODE AVAILABILITY

The data that support the findings of this study, and the code used to perform simulations are available from the corresponding author upon reasonable request.

AUTHOR CONTRIBUTIONS

S.J.d.G., S.H.X. and R.J.S. conceived the experiment. S.J.d.G, S.H.X. and B.J.C. set up the experiment hardware. B.J.C. fabricated the devices with help from L.F.. J.T., T.T., P.W. and J.W.O.G. advised on characterizing error sources in the dual-rail encoding and the comparison of this erasure check to prior schemes. K.M.C. and S.P. ran simulations of the surface code threshold in the presence of false negatives. S.J.d.G. and S.H.X. performed the measurements, analyzed the data, and wrote the manuscript with input from all authors.

COMPETING INTERESTS

L.F. and R.J.S. are founders and shareholders of Quantum Circuits Inc. (QCI)

[1] Ofek, N. *et al.* Extending the lifetime of a quantum bit with error correction in superconducting circuits. *Nature* **536**, 441–445 (2016).

[2] Flühmann, C. *et al.* Encoding a qubit in a trapped-ion mechanical oscillator. *Nature* **566**, 513–517 (2019).

[3] Hu, L. *et al.* Quantum error correction and universal gate set operation on a binomial bosonic logical qubit. *Nature Physics* **15**, 503–508 (2019).

[4] Campagne-Ibarcq, P. *et al.* Quantum error correction of a qubit encoded in grid states of an oscillator. *Nature* **584**, 368–372 (2020).

[5] Gertler, J. M. *et al.* Protecting a bosonic qubit with autonomous quantum error correction. *Nature* **590**, 243–248 (2021).

[6] Ni, Z. *et al.* Beating the break-even point with a discrete-variable-encoded logical qubit. *Nature* **616**, 56–60 (2023).

[7] Sivak, V. V. *et al.* Real-time quantum error correction beyond break-even. *Nature* **616**, 50–55 (2023).

[8] Braumüller, J. *et al.* Analog quantum simulation of the Rabi model in the ultra-strong coupling regime. *Nature Communications* **8**, 779 (2017).

[9] Hu, L. *et al.* Simulation of molecular spectroscopy with circuit quantum electrodynamics. *Science Bulletin* **63**, 293–299 (2018).

[10] Owens, C. *et al.* Quarter-flux Hofstadter lattice in a qubit-compatible microwave cavity array. *Physical Review A* **97**, 013818 (2018).

[11] Wang, C. S. *et al.* Efficient Multiphoton Sampling of Molecular Vibronic Spectra on a Superconducting Bosonic Processor. *Physical Review X* **10**, 021060 (2020).

[12] Wang, C. S. *et al.* Observation of wave-packet branching through an engineered conical intersection. *Phys. Rev. X* **13**, 011008 (2023).

[13] McCormick, K. C. *et al.* Quantum-enhanced sensing of a single-ion mechanical oscillator. *Nature* **572**, 86–90 (2019).

[14] Wang, W. *et al.* Heisenberg-limited single-mode quantum metrology in a superconducting circuit. *Nature Communications* **10**, 4382 (2019).

[15] Dixit, A. V. *et al.* Searching for dark matter with a superconducting qubit. *Phys. Rev. Lett.* **126**, 141302 (2021).

[16] Wallraff, A. *et al.* Strong coupling of a single photon to a superconducting qubit using circuit quantum electrodynamics. *Nature* **431** (2004).

[17] Blais, A., Huang, R. S., Wallraff, A., Girvin, S. M. & Schoelkopf, R. J. Cavity quantum electrodynamics for superconducting electrical circuits: An architecture for quantum computation. *Physical Review A* **69**, 1–14 (2004).

[18] Heeres, R. W. *et al.* Cavity State Manipulation Using Photon-Number Selective Phase Gates. *Physical Review Letters* **115**, 137002 (2015).

[19] Heeres, R. W. *et al.* Implementing a universal gate set on a logical qubit encoded in an oscillator. *Nature Communications* **8**, 94 (2017).

[20] Eickbusch, A. *et al.* Fast universal control of an oscillator with weak dispersive coupling to a qubit. *Nature Physics* **18**, 1464–1469 (2022).

[21] Diringer, A. A., Blumenthal, E., Grinberg, A., Jiang, L. & Hacohen-Gourgy, S. Conditional-not Displacement: Fast Multioscillator Control with a Single Qubit. *Physical Review X* **14**, 011055 (2024).

[22] Reinhold, P. *et al.* Error-corrected gates on an encoded qubit. *Nature Physics* **16**, 822–826 (2020).

[23] Rosenblum, S. *et al.* Fault-tolerant detection of a quantum error. *Science* **361**, 266–270 (2018).

[24] Sun, L. *et al.* Tracking Photon Jumps with Repeated Quantum Non-Demolition Parity Measurements. *Nature* **511**, 444–448 (2014).

[25] Roy, S. *et al.* Synthetic high angular momentum spin dynamics in a microwave oscillator (2024). arXiv:2405.15695 [quant-ph].

[26] Schuster, D. I. *et al.* Resolving photon number states in a superconducting circuit. *Nature* **445**, 515–518 (2007).

[27] Wang, H. *et al.* Deterministic entanglement of photons in two superconducting microwave resonators. *Phys. Rev. Lett.* **106**, 060401 (2011).

[28] Gao, Y. Y. *et al.* Entanglement of bosonic modes through an engineered exchange interaction. *Nature* **566**, 509–512 (2019).

[29] Chakram, S. *et al.* Multimode photon blockade. *Nature Physics* **18**, 879–884 (2022).

[30] Wang, C. *et al.* A schrödinger cat living in two boxes. *Science* **352**, 1087–1091 (2016).

[31] Wollack, E. A. *et al.* Quantum state preparation and tomography of entangled mechanical resonators. *Nature* **604**, 463–467 (2022).

[32] Gertler, J. M., van Geldern, S., Shirol, S., Jiang, L. & Wang, C. Experimental realization and characterization of stabilized pair-coherent states. *PRX Quantum* **4**, 020319 (2023).

[33] Koottandavida, A. *et al.* Erasure Detection of a Dual-Rail Qubit Encoded in a Double-Post Superconducting Cavity. *Physical Review Letters* **132**, 180601 (2024).

[34] Rosenblum, S. *et al.* A CNOT gate between multiphoton qubits encoded in two cavities. *Nature Communications* **9**, 652 (2018).

[35] Xu, Y. *et al.* Demonstration of Controlled-Phase Gates between Two Error-Correctable Photonic Qubits. *Physical Review Letters* **124**, 120501 (2020).

[36] Albert, V. V. *et al.* Pair-cat codes: autonomous error-correction with low-order nonlinearity. *Quantum Science and Technology* **4**, 035007 (2019).

[37] Chuang, I. L., Leung, D. W. & Yamamoto, Y. Bosonic quantum codes for amplitude damping. *Physical Review A* **56**, 1114–1125 (1997).

[38] Chuang, I. L. & Yamamoto, Y. Simple quantum computer. *Physical Review A* **52**, 3489–3496 (1995).

[39] Gao, Y. Y. *et al.* Programmable Interference between Two Microwave Quantum Memories. *Physical Review X* **8**, 021073 (2018).

[40] Chapman, B. J. *et al.* High-On-Off-Ratio Beam-Splitter Interaction for Gates on Bosonically Encoded Qubits. *PRX Quantum* **4**, 020355 (2023).

[41] Lu, Y. *et al.* High-fidelity parametric beamsplitting with a parity-protected converter. *Nature Communications* **14**, 5767 (2023).

[42] Tsunoda, T. *et al.* Error-Detectable Bosonic Entangling Gates with a Noisy Ancilla. *PRX Quantum* **4**, 020354 (2023).

[43] Teoh, J. D. *et al.* Dual-rail encoding with superconducting cavities. *Proceedings of the National Academy of Sci-*

ences **120**, e2221736120 (2023).

[44] Levine, H. *et al.* Demonstrating a Long-Coherence Dual-Rail Erasure Qubit Using Tunable Transmons. *Physical Review X* **14**, 011051 (2024).

[45] Grassl, M., Beth, T. & Pellizzari, T. Codes for the quantum erasure channel. *Physical Review A* **56**, 33–38 (1997).

[46] Kubica, A. *et al.* Erasure Qubits: Overcoming the ${T}_{1}$ Limit in Superconducting Circuits. *Physical Review X* **13**, 041022 (2023).

[47] Scholl, P. *et al.* Erasure conversion in a high-fidelity Rydberg quantum simulator. *Nature* **622**, 273–278 (2023).

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

[49] Kang, M., Campbell, W. C. & Brown, K. R. Quantum Error Correction with Metastable States of Trapped Ions Using Erasure Conversion. *PRX Quantum* **4**, 020358 (2023).

[50] Ma, S. *et al.* High-fidelity gates and mid-circuit erasure conversion in an atomic qubit. *Nature* **622**, 279–284 (2023).

[51] Chou, K. S. *et al.* Demonstrating a superconducting dual-rail cavity qubit with erasure-detected logical measurements (2023). arXiv:2307.03169 [quant-ph].

[52] Barrett, S. D. & Stace, T. M. Fault Tolerant Quantum Computation with Very High Threshold for Loss Errors. *Physical Review Letters* **105**, 200502 (2010).

[53] Delfosse, N. & Zémor, G. Linear-time maximum likelihood decoding of surface codes over the quantum erasure channel. *Physical Review Research* **2**, 033042 (2020).

[54] Reagor, M. *et al.* Quantum memory with millisecond coherence in circuit QED. *Physical Review B* **94**, 014506 (2016).

[55] Frattini, N. E. *et al.* 3-wave mixing Josephson dipole element. *Applied Physics Letters* **110** (2017).

[56] Frattini, N. E. *Three-wave Mixing in Superconducting Circuits: Stabilizing Cats with SNAILs.* Ph.D. thesis, Yale University (2021).

[57] Zhou, C. *et al.* Realizing all-to-all couplings among detachable quantum modules using a microwave quantum state router. *npj Quantum Information* **9**, 54 (2023).

[58] Schwinger, J. On angular momentum. Tech. Rep. NYO-3071, Harvard Univ., Cambridge, MA (United States); Nuclear Development Associates, Inc. (US) (1952).

[59] We note that the spectrum for $N = 1$ is reminiscent of the Mollow triplet [68, 69] seen when probing the frequency of a strongly-driven two-level-system. Here we are instead probing the frequency of a two-level-system that is dispersively coupled to the strongly driven mode.

[60] Kitaev, A. Fault-tolerant quantum computation by anyons. *Annals of Physics* **303**, 2–30 (2003).

[61] Dennis, E., Kitaev, A., Landahl, A. & Preskill, J. Topological quantum memory. *Journal of Mathematical Physics* **43**, 4452–4505 (2002).

[62] Fowler, A. G., Stephens, A. M. & Groszkowski, P. High-threshold universal quantum computation on the surface code. *Physical Review A* **80**, 052312 (2009).

[63] Ma, W.-L. *et al.* Path-independent quantum gates with noisy ancilla. *Phys. Rev. Lett.* **125**, 110503 (2020).

[64] Place, A. P. M. *et al.* New material platform for superconducting transmon qubits with coherence times exceeding 0.3 milliseconds. *Nature Communications* **12**,

[65] Ganjam, S. *et al.* Surpassing millisecond coherence in on chip superconducting quantum memories by optimizing materials and circuit design. *Nature Communications* **15**, 3687 (2024).

[66] Oriani, A. E. *et al.* Niobium coaxial cavities with internal quality factors exceeding 1.5 billion for circuit quantum electrodynamics (2024). arXiv:2403.00286 [quant-ph].

[67] Wigner, E. *Gruppentheorie und ihre Anwendung auf die Quantenmechanik der Atomspektren* (Vieweg+Tuebner Verlag Wiesbaden, 1931), 1 edn.

[68] Mollow, B. R. Power Spectrum of Light Scattered by Two-Level Systems. *Physical Review* **188**, 1969–1975 (1969).

[69] Baur, M. *et al.* Measurement of Autler-Townes and Mollow Transitions in a Strongly Driven Superconducting Qubit. *Physical Review Letters* **102**, 243602 (2009).

[70] Fowler, A. G. Coping with qubit leakage in topological codes. *Phys. Rev. A* **88**, 042308 (2013).

[71] Ghosh, J., Fowler, A. G., Martinis, J. M. & Geller, M. R. Understanding the effects of leakage in superconducting quantum-error-detection circuits. *Phys. Rev. A* **88**, 062329 (2013).

[72] Suchara, M., Cross, A. W. & Gambetta, J. M. Leakage suppression in the toric code. *Quantum Info. Comput.* **15**, 997–1016 (2015).

[73] Brown, N. C. & Brown, K. R. Comparing zeeman qubits to hyperfine qubits in the context of the surface code: $^{174}\mathrm{Yb}^{+}$ and $^{171}\mathrm{Yb}^{+}$. *Phys. Rev. A* **97**, 052301 (2018).

[74] Brown, N. C. & Brown, K. R. Leakage mitigation for quantum error correction using a mixed qubit scheme. *Phys. Rev. A* **100**, 032325 (2019).

[75] Miao, K. C. *et al.* Overcoming leakage in quantum error correction. *Nature Physics* **19**, 1780–1786 (2023).

[76] Chang, K. *et al.* Surface code with imperfect erasure checks (In prep.).

[77] Horodecki, M., Horodecki, P. & Horodecki, R. General teleportation channel, singlet fraction, and quasidistillation. *Physical Review A* **60**, 1888–1898 (1999).

[78] Johansson, J. R., Nation, P. D. & Nori, F. QuTiP: An open-source Python framework for the dynamics of open quantum systems. *Computer Physics Communications* **183**, 1760–1772 (2012).

[79] Gambetta, J. *et al.* Qubit-photon interactions in a cavity: Measurement-induced dephasing and number splitting. *Phys. Rev. A* **74**, 042318 (2006).

Appendix A: Constructing a tunable CPHASE(θ) gate using joint-photon number-splitting

Just as the joint-photon-number-splitting regime enables joint-photon-number selective measurements, it can also be used to extend the established selective-number-arbitrary-phase (SNAP) gate for a single oscillator [18] to a multi-oscillator 'joint-SNAP' (with some caveats). SNAP relies on the number-splitting regime to perform two consecutive photon-number-selective π-pulses with different phases, such that they enclose a geometric phase on the Bloch sphere of the transmon, thereby applying photon-number selective phases to the oscillator state.

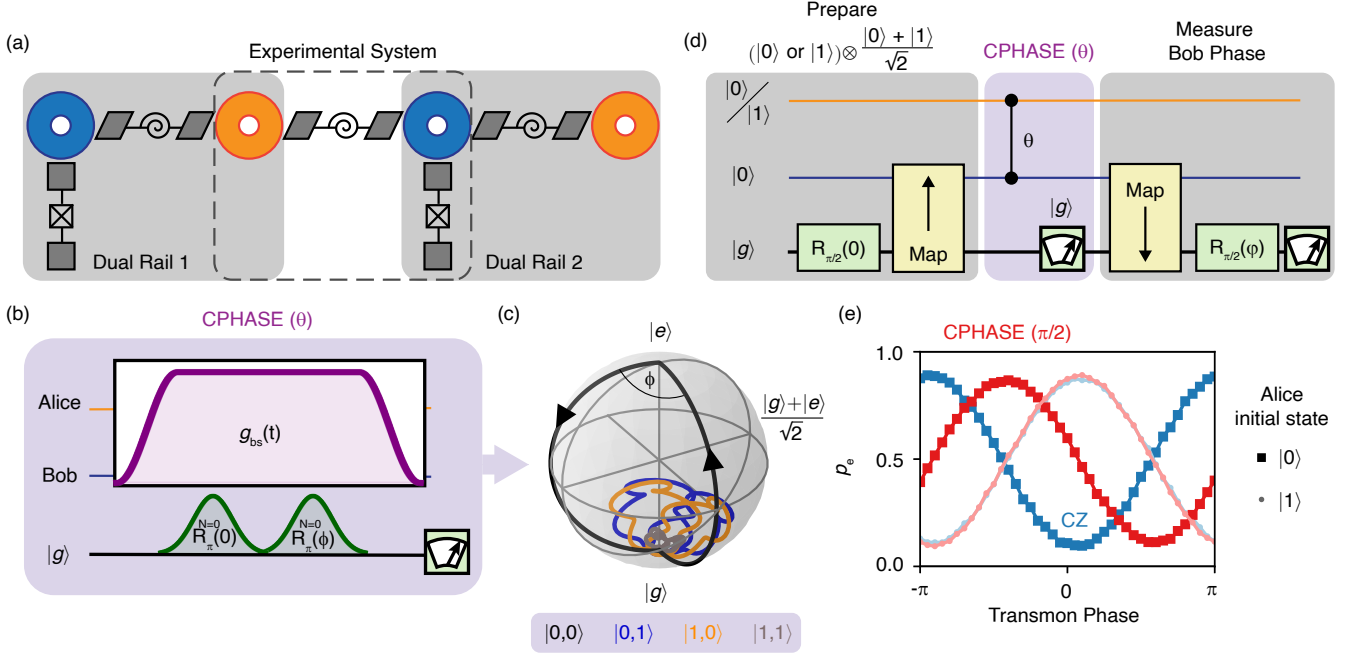Similarly, the application of a strong beamsplitter drive

FIG. 4. **CPHASE(θ) gate.** (a) Illustration of two adjacent dual-rail cavity qubits, with beamsplitter couplings enabled by SNAILs. The orange and blue annuli represent lambda/4 post cavities viewed from above. The setup in this experiment (within the dashed lines) can be used to operate a gate between these two hypothetical encoded dual-rail qubits. (b) Pulse sequence to obtain transmon trajectories during joint-SNAP CPHASE(θ) gate. (c) Bloch sphere trajectories of the transmon state conditioned on starting in every valid state for two oscillators shared between two dual-rail qubits. (d) Ramsey sequence used to probe the phase acquired by Bob's oscillator, conditioned on the state in Alice. (e) Results of Ramsey experiment to probe the phase acquired by Bob's oscillator during CPHASE(θ) gate with $\theta = \pi/2$ (red) and $\theta = \pi$ (blue, CZ), when Alice oscillator state is $|0\rangle$ (connected squares) and $|1\rangle$ (solid lines).

during the same SNAP sequence makes the constituent transmon π-pulses selective on the joint photon number in two oscillators, thereby allowing one to apply an arbitrary phase to each joint-photon-number component of the two-oscillator state:

$$\hat{U}_{\text{Joint-SNAP}} = \sum_{N=0}^{N_{\max}} e^{i\phi_N} \hat{P}_N, \qquad \text{(A1)}$$

where $\hat{P}_N$ is the projector onto two-oscillator states with total photon number N and $N_{\max}$ is the highest total photon number state we wish to address. However, as is described in App. L, for $N_{\max} > 1$, an increasingly large ratio $|g_{\text{bs}}/\chi|$ is required to ensure that different states with the same N evolve together. This construction is therefore most useful for applying a phase conditioned on $N = 0$ total photons.

The application of an $N = 0$ - selective phase implements a tunable CPHASE(θ) gate on two Fock-encoded qubits or equivalently, if the two oscillators are each part of a different dual-rail qubit (as in Fig. 4a), a logical CPHASE(θ) gate on dual-rail qubits. This gate is a maximally entangling gate between qubits, a key ingredient for the dual-rail architecture, and consists of applying a tunable phase only to the $|1_{\text{L}}\rangle|1_{\text{L}}\rangle$ two-qubit dual-rail state while leaving other dual-rail states untouched.

Since the logical information for a single dual-rail qubit is redundantly encoded in two oscillators, we can effect this by applying a phase to the $|0,0\rangle$ Fock state of the two addressed oscillators. The application of a geometric phase to the joint-vacuum state of two oscillators has previously been shown in the context of Schrödinger cat qubits using a transmon with a static dispersive coupling to both modes [35] but is here achieved with a transmon statically coupled to one mode and a strong beamsplitter interaction that temporarily activates the coupling to the other. (In principle, applying the phase to any *one* of $|0,0\rangle$, $|0,0\rangle$, $|0,0\rangle$ or $|0,0\rangle$, combined with local rotations on the cavities, can effect a CPHASE gate.)

The constraints on the value of $|g_{\text{bs}}/\chi|$ are slightly different in the case of the CPHASE(θ) gate than for the erasure check. Whereas previously we only needed to avoid transitions in the $N = 1$ total photon number manifold, the $|1,1\rangle$ Fock state (with $N = 2$) is a valid initial state for two oscillators belonging to two different dual-rail qubits. As such, the need to drive an $N = 0$ transition while avoiding both $N = 1$ and $N = 2$ transitions leads us to choose a value of $|g_{\text{bs}}/\chi| \approx 1.6$ (see triangle Fig. 1c). Further increasing $|g_{\text{bs}}/\chi|$ would increase the separation to unwanted transitions at the potential cost of additional decoherence as we approach the maximum value accessible in this device.

The simulated Bloch sphere trajectories of the transmon state for all valid input oscillator states (shown in Fig. 4c) show the principle of this operation for $\text{CPHASE}(\pi) = \text{CZ}$. As shown in Fig. 4b, $|0,0\rangle$, $|0,1\rangle$, $|1,0\rangle$ and $|1,1\rangle$ are prepared, before performing back-to-back selective π-pulses with a phase difference that is chosen to ensure that $\theta = \pi$. Whereas for initial states $|0,1\rangle$, $|1,0\rangle$ and $|1,1\rangle$ the transmon trajectories stay near the ground state and return to $|g\rangle$, the $|0,0\rangle$ trajectory passes through $|e\rangle$ and encloses a large geometric phase before returning to $|g\rangle$.

To verify that this protocol performs a $\text{CPHASE}(\theta)$ gate on two Fock-encoded qubits, we can perform the Ramsey sequence shown in Fig. 4d. This starts by preparing either $|0\rangle|+\rangle$ or $|1\rangle|+\rangle$ in the oscillators, where $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$. The superposition states are loaded into the oscillator by preparing $(|g\rangle + |e\rangle)/\sqrt{2}$ in the transmon and performing an 'encoding' optimal control pulse (OCP) which maps $|\psi\rangle_{\text{qubit}}|0\rangle_{\text{oscillator}} \rightarrow |g\rangle_{\text{qubit}}|\psi\rangle_{\text{oscillator}}$. We then enact the $\text{CPHASE}(\theta)$ gate for variable θ, including a transmon readout to catch errors that leave it out of the ground state, before mapping the oscillator state onto the transmon with a 'decoding' OCP (which performs the inverse of the 'encoding' OCP) and measuring the phase of the transmon using a variable-phase $\pi/2$-pulse. In this way, the transmon measurement allows us to probe the phase acquired by the Bob oscillator state during the gate.

The results for $\theta = \pi/2$ and $\theta = \pi$, shown in Fig. 4e, show that a tunable phase can be imparted on Bob's superposition state dependent on the state in Alice's cavity. The measured transmon population oscillations encode the phase of Bob's cavity superposition. When Alice is initialized in $|1\rangle$ (connected dots) these oscillations lie on top of one another, indicating that the two-oscillator states $|1,0\rangle$ and $|1,1\rangle$ have the same relative phase regardless of the choice of θ. Meanwhile the relative position of the oscillations when Alice is initialized in $|0\rangle$ (squares), can be tuned depending on the enclosed phase, such that we are free to choose the relative phase between the two-oscillator states $|0,0\rangle$ and $|0,1\rangle$. Together with arbitrary rotations on either cavity, which can be performed virtually, this provides the necessary control for a $\text{CPHASE}(\theta)$ gate. The loss in contrast of the oscillations is dominated by SPAM errors during the OCT encode and decode pulses, as well as cavity photon loss. Full characterization of the dual-rail CPHASE gate performance would require process tomography and would ideally be performed on a system with four modes comprising two dual-rails (as shown in Fig. 4a). This would allow for both high-fidelity SPAM and the detection of photon loss during the gate.

This gate construction has some, but not all of the desired features for a multi-cavity encoding scheme such as the dual-rail. By appending the $\text{CPHASE}(\theta)$ operation with a transmon readout, some, but not all, transmon errors can be caught and converted to erasures. Fig. 5 shows the simulated erasure rate and post-selected state
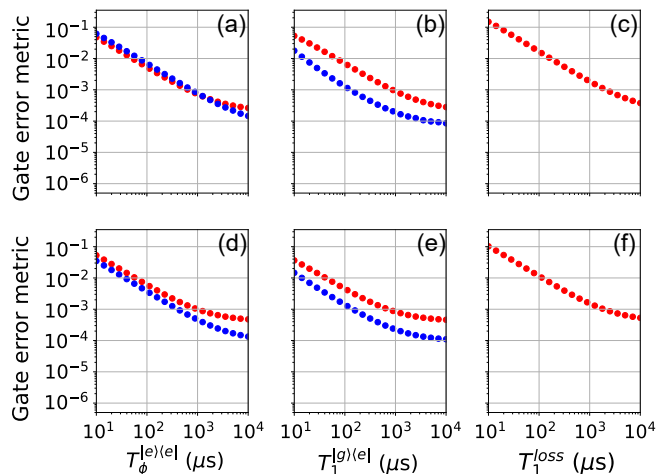

FIG. 5. **Projected dual-rail CPHASE performance.** Simulated erasure rate (red) and post-selected infidelity (blue) for the dual-rail $\text{CPHASE}(\pi)$ gate, with chopped Gaussian transmon pulse of total duration 3.194 µs (a-c) and 'shaped-square' transmon pulse shape of total duration 1.994 µs (d-f). In each column, either the transmon dephasing rate $T_\phi^{|e\rangle\langle e|}$, transmon decay rate $T_1^{|g\rangle\langle e|}$ or cavity photon loss rates T_1^{loss} are varied with no other loss channels included. All simulation results are averaged over all 36 two-qubit states formed from every combination of the 6 dual-rail cardinal states in either qubit. For all simulations, we use a dispersive shift of $\chi/2\pi = 1$ MHz, and an infinite-bandwidth square beamsplitter pulse whose amplitude is optimized to minimize coherent errors while restricting it to obey $g_{\text{bs}}/2\pi < 2$ MHz. Impact of Stark shifts due to the beamsplitter drive is neglected.

infidelity, which depend linearly on the underlying physical error rates but with different offsets. (The plateau at long coherence times is due to residual coherent errors that could be removed by using a pulse shape with more degrees of freedom than a simple square pulse.) Whereas the post-selected infidelity due to transmon decay errors is distinguishably lower than the corresponding erasure rate, indicating that most transmon decay errors can be caught by the check, for transmon dephasing errors they are comparable.

This construction therefore does not achieve the full error-detecting and fault-tolerant properties of the $ZZ(\theta)$ gate for dual-rail qubits proposed by Tsunoda et al. [42]. Adapting this $\text{CPHASE}(\theta)$ gate to be fully error-detectable to transmon errors would require the use of three transmon energy levels (to protect against transmon relaxation) and simultaneous transmon drives on all joint-photon-number-selective peaks (to protect against transmon dephasing) [22].

Appendix B: Calibrating beamsplitter strength and Stark shifts

The beamsplitter amplitude g_{bs} and resonance frequency $\omega_{\Delta=0}$ (i.e. the drive frequency required to achieve a detuning $\Delta = 0$) are measured by initializing the state $|0, 1\rangle \otimes |g\rangle$, turning on the beamsplitter drive and measuring the photon number in Bob's cavity after some time t. By varying the frequency and duration of the beamsplitter drive, we observe the chevron pattern seen in Fig. 6a, which we can fit to:

$$P_1(\omega, t) = A\left[\cos^2\left(\frac{\Omega t}{2} + \phi\right) \right.$$
$$\left. + \left(\frac{\omega - \omega_{\Delta=0}}{\Omega}\right)^2 \sin^2\left(\frac{\Omega t}{2} + \phi\right) \right] + c,$$

(B1)

where the detuned oscillation rate is given by

$$\Omega = \sqrt{g_{\mathrm{bs}}^2 + (\omega - \omega_{\Delta=0})^2},$$

(B2)

ϕ captures the phase accumulation while the beamsplitter drive is ramping up, and A and c capture the contrast and offset due to SPAM.

Fig. 6b shows the fitted values of g_{bs} over the range of microwave drive amplitudes used in this experiment. As described in previous work [40], the measured beamsplitter amplitude is linearly proportional to drive amplitude at low amplitudes but deviates at high amplitudes. In order to interpolate between the amplitudes used for the calibration experiment, we fit the g_{bs} curve to a degree-5 polynomial (with the intercept forced to pass through the origin).

The fits to the chevrons also provide an estimate of the beamsplitter resonance frequency, which at low pump powers is equal to the difference in the undriven cavity frequencies, $\omega_b^0 - \omega_a^0$. This in turn allows us to estimate the frequency $\omega_{\Delta=\chi/2}$ at which to drive the beamsplitter to achieve a 'symmetric' detuning $\Delta = \chi/2$. This frequency shifts downwards due to Stark shifts of the cavity frequencies induced by the beamsplitter drive, with Bob's cavity frequency shifting downward by more than Alice's:

$$\omega_{\Delta=\chi/2} = \omega_{\Delta=0} + \chi/2$$
$$= \left(\omega_b^0 + \delta_b^{\mathrm{Stark}}\right) - \left(\omega_a^0 + \delta_a^{\mathrm{Stark}}\right) + \chi/2$$
$$= \left(\omega_b^0 - \omega_a^0 + \chi/2\right) +$$
$$\left(\delta_b^{\mathrm{Stark}} - \delta_a^{\mathrm{Stark}}\right).$$

(B3)

We perform a finer calibration of this frequency by measuring the oscillations of the cavities' photon number induced by the beamsplitter drive when the transmon is in $|g\rangle$ or in $|e\rangle$. The point at which the frequencies of these oscillations are matched marks the symmetric $\Delta = \chi/2$ point for which the data in Fig. 1 is taken. These values, relative to $\omega_b^0 - \omega_a^0$, are plotted in Fig. 6c.
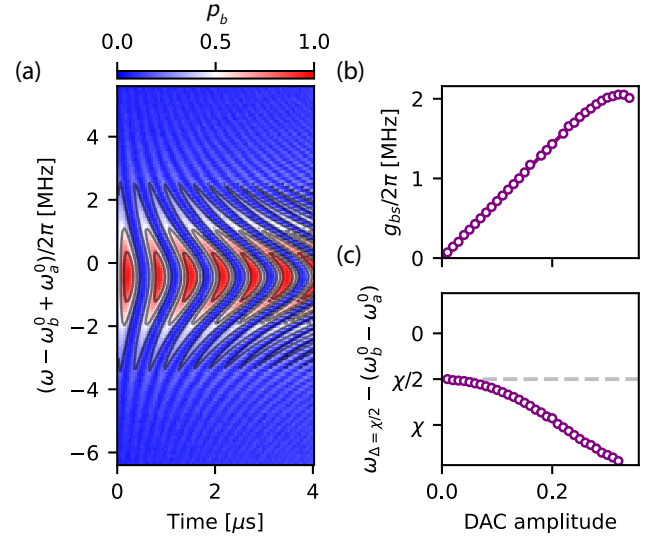

FIG. 6. **Beamsplitter amplitude and resonance frequency.** (a) Measured oscillations in the population of Bob's cavity, p_b, as a function of applied beamsplitter drive frequency. The data is shown for a DAC amplitude of 0.25 for the beamsplitter drive. Contours show the fit to Eq. B1, from which g_{bs} and a first estimate for the resonance frequency are obtained. (b) Fitted beamsplitter amplitude g_{bs} as a function of DAC amplitude. (c) Beamsplitter drive frequency required to ensure a symmetric detuning $\Delta = \chi/2$ as a function of DAC amplitude, relative to the bare resonance frequency.

Appendix C: Measuring transition matrix elements

To measure the transition matrix elements in Fig. 2b, we perform a power-Rabi experiment which determines the amplitude of the spectroscopy tone required to perform a π-pulse on the ancilla. This quantity is inversely proportional to the transition matrix element. We run this sequence at each value of $|g_{\mathrm{bs}}/\chi|$ for both the central $\delta m = 0$ transition and the lower $\delta m = -1$ transition. The $\delta m = -1$ transition is chosen over the $\delta m = +1$ transition, owing to its distance from other transitions, allowing for a cleaner measurement, but both $\delta m = \pm 1$ transitions are expected to share the same matrix element.

This measurement requires us to ensure that the frequency of the spectroscopy tone is on resonance with the transition. If the spectroscopy tone has some detuning, the rate of Rabi oscillations would be overestimated. To mitigate this, we precisely find the ancilla resonance frequency for each value of $|g_{\mathrm{bs}}/\chi|$ by taking a vertical slice of the $N = 1$ colorplot in Fig. 1c in the vicinity of the desired transition, fitting it to a Gaussian and extracting the fitted center frequency.

The power-Rabi sequence consists of initializing the oscillators in the state $(|0, 1\rangle + i |1, 0\rangle)/\sqrt{2}$, playing a Gaussian-shaped pulse of total duration 14.8 μs with varying amplitude and measuring the state of the transmon at the end. The rate of the oscillations (see Fig. 7 for
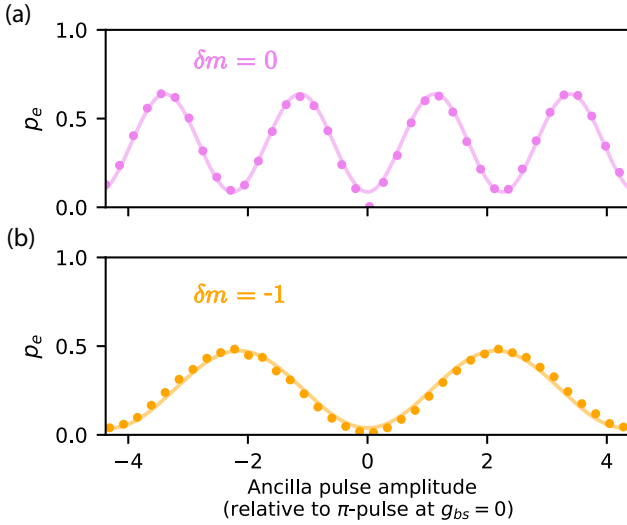
FIG. 7. **Power Rabi oscillations used to extract matrix elements.** Measured probability of finding the ancilla in $|e\rangle$ after a selective ancilla pulse of variable amplitude for the $\delta m = 0$ (a) and $\delta m = -1$ (b) transitions in the $N = 1$ photon manifold of Fig. 1c. A symmetric beamsplitter drive is applied during the pulse, in this case with an amplitude $g_{\text{bs}} \approx |\chi|$. The ancilla pulse amplitude is normalized to the value required for a π-pulse in the absence of a beamsplitter drive. The frequency of these oscillations is used to extract the transition matrix elements in Fig. 2.
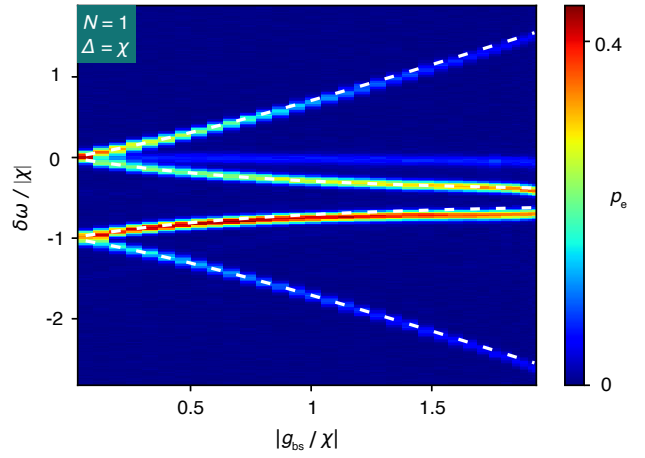
FIG. 8. **Driven transmon spectrum with non-symmetric beamsplitter detuning.** Measured probability of exciting the transmon when initializing the cavities in $|\psi\rangle = (|0,1\rangle + i\,|1,0\rangle)/\sqrt{2}$ and applying a selective transmon pulse in the presence of a beamsplitter drive with amplitude g_{bs} and detuning $\Delta = \chi$. The white dashed lines show predicted transition frequencies from spin-oscillator model.

an example) is inversely proportional to the amplitude required for a single π-pulse, and so is proportional to the transition matrix element. We avoid the need to precisely calibrate the drive power delivered to the package by normalizing all the power-Rabi oscillation frequencies to the value at $g_{\text{bs}} = 0$ for $\delta m = -1$. Importantly, this technique is not sensitive to infidelity in state preparation and measurement, as this just affects the amplitude of the oscillations, not their rate. The experimental $\delta m = 0$ curve in Fig. 2b does not extend to $g_{\text{bs}}/\chi = 0$. The reason for this is that as the transition matrix element decreases, the transmon drive amplitude $|\epsilon|$ required to perform a π-pulse increases. As $|\epsilon|$ becomes comparable to $\max(g_{\text{bs}}, \chi)$, the assumption that the transmon drive can be treated as a perturbation in the Hamiltonian breaks down. Specifically, we find that there is no longer a single degenerate $\delta m = 0$ transition at $\delta\omega = 0$.

Appendix D: Degeneracy breaking for non-symmetric beamsplitter detuning

In the main text, we primarily consider the case where a beamsplitter drive is applied with a symmetric detuning, $\Delta = \chi/2$, so-called because it lies exactly halfway between the beamsplitter resonance when the ancilla qubit is in $|g\rangle$ and the resonance when the ancilla is in $|e\rangle$. In this case, many of the transition frequencies are degenerate. However, as can be seen in Eq. 16, we expect this

degeneracy to break when we move to a non-symmetric beamsplitter detuning, $\Delta \neq \chi/2$. For example, in the case that the total photon number $N = 1$, we expect four (rather than three) separate transition frequencies.

We can verify this by measuring the transmon spectrum when the cavities are initialized in $|\psi\rangle = (|0,1\rangle + i\,|1,0\rangle)/\sqrt{2}$ and we apply a beamsplitter drive with detuning $\Delta = \chi$, shown in Fig. 8. The white dashed lines show the theory predictions which agree well with the measured spectrum, with the exception of a slight shift downwards in frequency of the experimental curves with respect to theory at large beamsplitter amplitudes. This is at least in part due to a slight negative Stark shift of the transmon frequency when a strong beamsplitter drive is applied.

Appendix E: Tune-up of joint-photon number-selective pulses

In general, a variety of transmon pulse shapes may be chosen for the joint-photon number-selective pulse. The choice of pulse shape provides a way to trade off between idling and false negative errors (reduced with a shorter, less frequency selective pulse) and false positive and induced Pauli errors (reduced with a longer, more frequency selective pulse). In the erasure check demonstrated in this work, we considered a cosine-ramped square pulse. Later in this section we also provide a method for tuning up a longer chopped Gaussian pulse, as used for the CPHASE gate in Appendix A.

Erasure check with square pulse

We first consider cosine-ramped square transmon pulses of the form

$$\frac{\hat{\mathcal{H}}_{\text{transmon}}}{\hbar} = f(t)\,\hat{\sigma}_x, \tag{E1}$$

with

$$f(t) = \begin{cases} \frac{A}{2}\left(1 - \cos\left(\frac{\pi t}{t_r}\right)\right), & 0 < t < t_r \\ A, & t_r \le t \le T_{\text{p}} - t_r \\ \frac{A}{2}\left(1 + \cos\left(\frac{\pi(t - t_r + T_{\text{p}})}{t_r}\right)\right), & T_{\text{p}} - t_r < t < T_{\text{p}}. \end{cases} \tag{E2}$$

where A is the pulse amplitude, T_p is the pulse duration and t_r is the duration of the cosine ramp. The ramp time t_r should be kept sufficiently long to prevent coherent errors from the $|f\rangle$ level of the transmon.

When using a square transmon pulse, there are discrete operating points that ensure that dual-rail and ancilla states both return to where they began. As mentioned in the main text, we find from simulation that aligning the center (in time) of the cosine ramps of the transmon pulse with those of the beamsplitter drive ensures optimal performance.

The tune-up procedure consists of:

1. Calibrate the following quantities as a function of DAC amplitude:

   - beamsplitter amplitude g_{bs},
   - beamsplitter detuning $\omega_{\Delta=\chi/2}$,
   - ancilla resonance frequency ω_{ancilla}.

   As described in Appendix B, the frequencies may shift as the drive amplitude is increased.

2. Identify the maximum beamsplitter amplitude $g_{\text{bs}}^{\text{max}}$ that can be accessed without introducing significantly more errors than when idling. Above certain drive amplitudes, one may see enhanced photon loss, dephasing or heating.

3. Find an initial starting point for the operating parameters: T_p, g_{bs}, A, the beamsplitter drive frequency ω, and the detuning of the ancilla drive from its undriven resonance frequency $\delta\omega$. For infinite bandwidth pulses ($t_r = 0$) and no Stark shifts, the following analytical expressions approximate the ideal values:

$$T_p = \frac{2\pi\sqrt{3}}{|\chi|} \qquad g_{\text{bs}} = |\chi|\sqrt{\frac{m^2}{3} - \frac{1}{4}} \tag{E3}$$

$$A = \frac{|\chi|}{4\sqrt{3}} \qquad \omega = \omega_b - \omega_a + \chi/2, \tag{E4}$$

with $\delta\omega = 0$ and where the integer $m > 1$ can be varied in order to operate at different values of

g_{bs}. We choose values close to $m = 2$ to ensure $g_{\text{bs}} < g_{\text{bs}}^{\text{max}}$.

A more precise starting point can be obtained by performing a Schrödinger equation simulation which includes the ramp times as well as Stark shifts (i.e. how the ancilla resonance frequency and $\omega_{\Delta=\chi/2}$ change with pump amplitude) and running gradient descent optimization using the infidelity of the state transfers

$$|0, 0, g\rangle \to |0, 0, e\rangle, \tag{E5}$$
$$|0, 1, g\rangle \to |0, 1, g\rangle, \tag{E6}$$
$$|1, 0, g\rangle \to |1, 0, g\rangle \tag{E7}$$

as a cost function.

4. Starting with the parameters found above, fine-tune the ancilla drive amplitude A and detuning $\delta\omega$ experimentally by initializing the cavities in $|0, 0\rangle$ and performing the ancilla pulse with the beamsplitter drive applied. Choose the parameters that maximize the probability of exciting the transmon, thereby minimizing false negatives. Since the $|0, 0\rangle$ state is unaffected by the beamsplitter drive, this calibration depends very weakly on changes in g_{bs} in the following step.

5. With T_p and A fixed, perform the following experiment while sweeping the values of g_{bs} and ω:

   - Initialize the system in $|0, 1, g\rangle$ or $|1, 0, g\rangle$,
   - Perform N successive erasure checks,
   - Perform an erasure-detected logical measurement.

   From this single experiment, we obtain three key metrics to maximize:

   - Probability of passing N checks when initializing in $|0, 1, g\rangle$,
   - Probability of passing N checks when initializing in $|1, 0, g\rangle$,
   - Probability of returning to $|0, 1, g\rangle$ when initializing in $|0, 1, g\rangle$.

   With the optimal choice of T_p, there should be a point in (g_{bs}, ω) space that maximizes all three, thereby minimizing both false positives and coherent Pauli errors.

6. If there is not optimal choice for g_{bs} and ω, T_p can be adjusted. If the operating points that optimized for false positives lie at a lower g_{bs} than is required to ensure the dual-rail cavity states return to where they began, the pulse duration T_p can be increased (and vice-versa). The previous step can then be repeated.

   We can probe whether the pulse duration is set correctly via a spectroscopy experiment. This involves
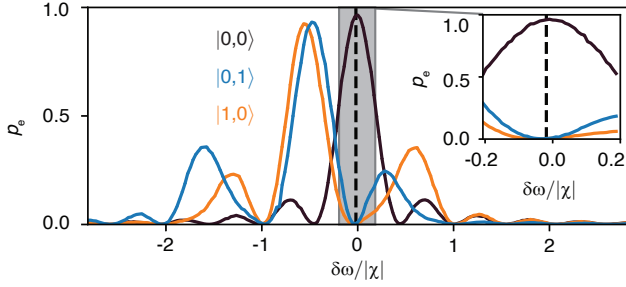
FIG. 9. **Square Pulse Spectroscopy.** Probability of erasure check reading out $|e\rangle$ as a function of transmon pulse detuning during the check when initializing the dual-rail qubit in $|0,0\rangle$ (black), $|0,1\rangle$ (blue) or $|1,0\rangle$ (orange) state. The results are post-selected on the total photon number in the oscillators remaining the same after the check.

preparing the cavities in $|0,0\rangle$, $|0,1\rangle$ and $|1,0\rangle$, and performing a single erasure check with a variable detuning $\delta\omega$ on the ancilla pulse. We also post-select on the total photon number in the two cavities remaining the same after the check to remove the effect of photon loss or gain. Fig. 9 shows the result for the parameters optimized in this experiment, with the maximum of the $|0,0\rangle$ peak lining up with the minimum of the $|0,1\rangle$ and $|1,0\rangle$ traces. The frequency at which they line up (indicated with the dashed vertical line) is the ancilla frequency detuning used for the check. The slight gap between the peaks of the $|0,1\rangle$ and $|1,0\rangle$ curves is expected, even at $\Delta = \chi/2$, and is due to the ancilla drive amplitude no longer being much smaller than g_{bs}. If the extrema are not aligned at zero detuning, increasing (decreasing) the duration of the square pulse allows us to narrow (widen) the spectrum of the curves such that they align. Doing so will then require re-optimizing g_{bs} and ω in this iterative tune-up scheme.

7. Finally, double-check the calibration of A and $\delta\omega$ now that g_{bs} and ω have been adjusted.

Erasure Check with Gaussian pulse

Alternatively, one can consider (amongst other narrow bandwidth pulse shapes) chopped-Gaussian pulses of the form:

$$f(t) = A\left[\exp\left(-\frac{(t - n_{chop}\sigma)^2}{2\sigma^2}\right) - \exp\left(-\frac{n_{chop}^2}{2}\right)\right],$$

(E8)

where σ is the RMS width of the Gaussian and the pulse duration is $T_p = 2n_{chop}\sigma$.

1-2. Same as for the square pulse sequence.

3. Set the beamsplitter amplitude to a value in the range $\sqrt{3}|\chi|/2 < g_{bs} < g_{bs}^{max}$. Larger values within this range will suppress the impact of ancilla decoherence on dual-rail Pauli errors. Set the beamsplitter frequency to $\omega_{\Delta=\chi/2}$. Tune up a π-pulse on the ancilla with the cavities in $|0,0\rangle$ for these beamsplitter parameters.

4. Prepare $|0,1\rangle$ and $|1,0\rangle$ in the cavities and sweep the selectivity (given by σ) of the ancilla pulse while keeping the product of A and σ fixed. By observing the probability of exciting the ancilla, we can determine a minimum selectivity in order to ensure we do not flag $|0,1\rangle$ and $|1,0\rangle$ as erasures. Set σ to a value above this threshold.

5. Finally, we must ensure that the cavity states return to their initial states at the end of the sequence (to avoid coherent Pauli errors). Prepare $|1,0\rangle$ and $|0,1\rangle$, perform the erasure check with a variable beamsplitter amplitude, and measure the dual-rail state at the end. Depending on how close we are to performing a full revolution, there are a few options:

   - If the beamsplitter amplitude can be increased to perform an integer number of revolutions without exceeding g_{bs}^{max}, do this. This ensures the pulse is as short as it can be.

   - If not, increase σ until an integer number of revolutions is performed (at the same g_{bs}).

6. Having made this adjustment, other parameters may need to be changed:

   - If g_{bs} has been adjusted, also adjust the beamsplitter detuning (to ensure $\Delta = \chi/2$), and the transmon detuning (if it is Stark shifted by the drive).

   - If σ has been adjusted, fine-tune the amplitude of the transmon pulse such that a complete π-pulse is performed.

7. Return to Step 5 to verify whether the dual-rail computational states return to their initial states. If not, iterate over Steps 5 and 6 until this is true.

Adapting tune-up procedure for CPHASE(θ)

When performing the CPHASE(θ) gate in Appendix A we also care about the input two-cavity state $|1,1\rangle$. In this experiment we adapt the Gaussian pulse tune-up procedure by (a) ensuring the starting value of g_{bs} is sufficiently large to avoid driving transitions in the $N = 2$ manifold, and (b) adding the state $|1,1\rangle$ to the σ and g_{bs} calibration sequences in steps 4 and 5.

The square pulse tune-up procedure may also be adapted for use with CPHASE(θ) but more degrees of

Appendix F: Fault-tolerant threshold simulations

We find fault-tolerant thresholds of the CSS surface code, assuming noisy two-qubit gates and mid-circuit erasure checks applied to every qubit after each gate. These noisy two-qubit gates have error probability p in which its qubits experience an erasure or Pauli error. We define a parameter called the erasure fraction, $R_e = p_{eras}/(p_{eras} + p_{Pauli})$. Although we anticipate an erasure fraction of $R_e = 0.9$, extracted from the erasure check demonstrated in this work, we do not expect R_e to change significantly for two-qubit gates. Namely, the beamsplitter and dispersive interactions required for the erasure checks are also utilized in proposals for entangling gates [42, 43].

With probability $p(1-R_e)$, both qubits involved in the gate experience a Pauli error uniformly drawn from the set $\{I, X, Y, Z\}^{\otimes 2}$. Furthermore, with probability pR_e, one qubit from the two-qubit gate leaves the computational space. To describe what happens to other unleaked qubits, we assume that a leaked qubit induces computational Pauli errors on any qubit it interacts with [70–74]. Thus, the unleaked qubit in the gate undergoes a Pauli error that is dependent on which two-qubit gate is applied [74]. Specifically, in a CZ gate, the unleaked qubit experiences an I or Z error with equal probability. If the control (target) qubit is leaked in a CX gate, then the unleaked target (control) qubit experiences an I or X (I or Z) error. The proposed method for two-qubit gates in the dual-rail cavity architecture, the $ZZ(\theta)$ gate, follows the specific model of leakage-induced Pauli errors studied here [42, 43]. For this gate proposal, one can show that a leaked qubit induces rotational errors on the other qubit along one Pauli axis.

We model false negatives of erasure checks as the following. When erasure occurs during a two-qubit gate, the erasure check after this gate returns no flag with the false-negative probability 3.7%. Given a false negative, we assume the detectable leakage will be detected during the next round of mid-circuit erasure checks. We therefore ignore the effect of doubly-missed erasures, which are expected to occur with a probability on the order of 10^{-5}. These events, while damaging in the context of error-correction and a subject for future work, are rare relative to leakage rates in state-of-the-art transmon-based architectures (on the order of 10^{-3}) [75]. As before, the leaked qubit is assumed to induce computational errors on any qubit it interacts with before it is detected. To account for these correlated errors, we use a modified, weighted union-find decoder. For details on this decoder, we refer to [76].

Appendix G: Erasure-detected end-of-line logical measurement

Logical X, Y and Z measurements at the end of the sequence in Fig. 3d consist of a 50:50 beamsplitter pulse that rotates the dual-rail state onto the desired measurement axis, followed by an erasure-detected end-of-line (EOL) measurement, illustrated in Fig. 10a [51]. In this section, we describe the measurement sequence we use and show that the post-selection probability does not depend on the number of preceding successful mid-circuit checks.

The logical measurement starts with a transmon reset sequence that conditionally resets both transmons and double checks that they are indeed in $|g\rangle$ after the reset. Resetting the transmons to $|g\rangle$ is crucial because starting the EOL measurement with one or more transmon in $|e\rangle$ could lead to an error in state assignment.

Subsequently, cavity-photon-number-selective π pulses excite the measurement transmons conditioned on 1 photon in the adjacent cavity. We record this result and conditionally reset the transmons to $|g\rangle$ before swapping the cavity states with two consecutive 50:50 beamsplitter pulses. We then perform a second photon-number-selective measurement. This echoed sequence, where the state of each cavity is measured twice (once with each measurement transmon), removes bias due to any differences between the measurement fidelities of the transmons or between the coherences of the cavities. We only keep results in which (a) these two rounds of measurement are consistent and (b) they indicate a total of 1 photon in both cavities (i.e. $|g\rangle_A |e\rangle_B$ followed by $|e\rangle_A |g\rangle_B$, or $|e\rangle_A |g\rangle_B$ followed by $|g\rangle_A |e\rangle_B$).

Fig. 10b shows the probability of passing the EOL measurement, conditioned on passing all previous mid-circuit erasure checks, for the data presented in Fig. 3f. This probability is consistent regardless of the number of mid-circuit checks performed thus confirming that the EOL measurement does not introduce any bias in our measurements. This 'strict' EOL measurement sequence discards a relatively large fraction of the data, with an average success probability of $79.4 \pm 0.4\%$, but in doing so, ensures a faithful measurement of the final dual-rail states.

Appendix H: Dual-rail state fidelity after repeated erasure checks

To characterize the Pauli error rate per erasure check (post-selected on passing every check), we perform the sequence shown in Fig. 3d, in which we prepare all six dual-rail cardinal states, perform n successive erasure checks (with an echo pulse inserted halfway) and then perform a logical measurement along the axis in which the initial state was prepared. Ideally the erasure check should leave the initial state alone, and this measurement sequence allows us to obtain the state fidelity for
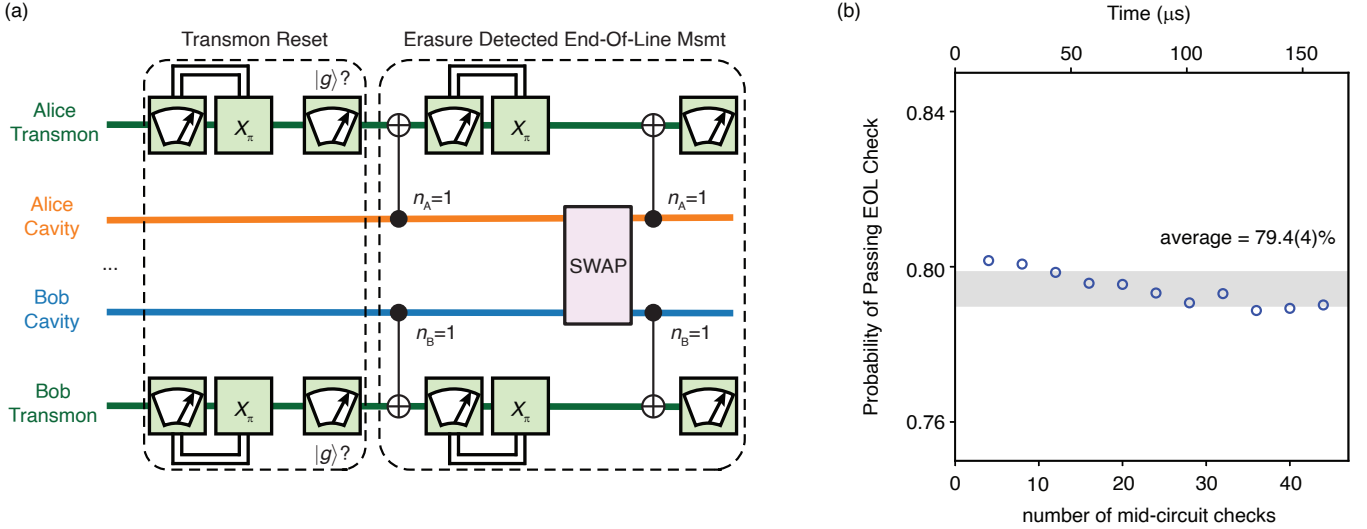
FIG. 10. **Erasure-detected end-of-line (EOL) logical measurement.** (a) Pulse sequence for the erasure detected EOL check comprises transmon reset and erasure-detected EOL measurement.(b) Probability of passing the EOL measurement plotted against the number of preceding mid-circuit erasure checks before the EOL measurement. Gray horizontal line shows average probability.

each cardinal state. This fidelity is post-selected on passing all the erasure checks, and on the end-of-line logical measurement not finding the cavities in $|0,0\rangle$.

The state fidelities for initial states $|\pm X\rangle_L$, $|\pm Y\rangle_L$ and $|\pm Z\rangle_L$ are shown in purple in Fig. 11. The grey lines show the same fidelities when we instead idle for the duration of the beamsplitter and transmon pulses. Whereas the idling dual-rail state is strongly biased against having bit-flip errors (indicated by the high $\pm |Z\rangle_L$ state fidelity), the errors induced on the state during the erasure check are closer to a depolarizing channel. Just as with single-qubit gates, this removal of the bias is an inevitable consequence of applying the beamsplitter drive.

From the average state infidelities per erasure check, we extract the post-selected Pauli error rate via [77]

$$p_{\mathrm{Pauli}} = p_x + p_y + p_z = \frac{3}{2}\left(1 - \bar{F}\right). \qquad \text{(H1)}$$

The factor of $3/2$ comes from the fact that only two out of the three types of errors (X, Y or Z) will flip any of the cardinal states (e.g. the $|+Z\rangle$ state is unaffected by Z-type errors).

We also measure the expectation values of the logical operators perpendicular to the axis along which the initial state was prepared (Fig. 12). Ideally these values should be zero. In Fig. 12, the logical measurement results display small biases for both performing erasure check (top row) and when idling (bottom row). These biases are largest for X_L and Y_L measurements but they do not grow or noticeably oscillate in time. We speculate that this indicates that the sequence is not perfectly symmetric about the echo pulse.
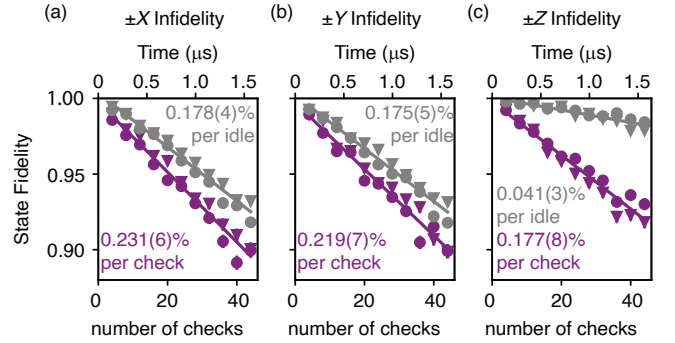

FIG. 11. **Dual-rail state fidelities after repeated erasure checks.** (a-c) Purple markers show state fidelity of logical plus ($|+X\rangle$, $|+Y\rangle$ or $|+Z\rangle$) (dots) or logical minus ($|-X\rangle$, $|-Y\rangle$ or $|-Z\rangle$)(triangles) states after passing n erasure checks. Gray markers show same fidelities when replacing each erasure check with a delay of the same duration. Lines show linear fits.

Appendix I: Error scaling for erasure check

To help decide which parameters to use for the erasure check, we would like to determine how the erasure and Pauli error rates scale as we vary the decoherence rates of the system, the duration of the transmon pulse and the ratio of the beamsplitter amplitude to the dispersive shift. Here we do so with the aid of QuTiP [78] master equation simulations.

The key takeaways are that (a) increasing the pulse duration and commensurately reducing the transmon pulse amplitude reduces the probability of suffering transmon-induced erasure and Pauli errors, (b) increasing the beamsplitter amplitude reduces the probability
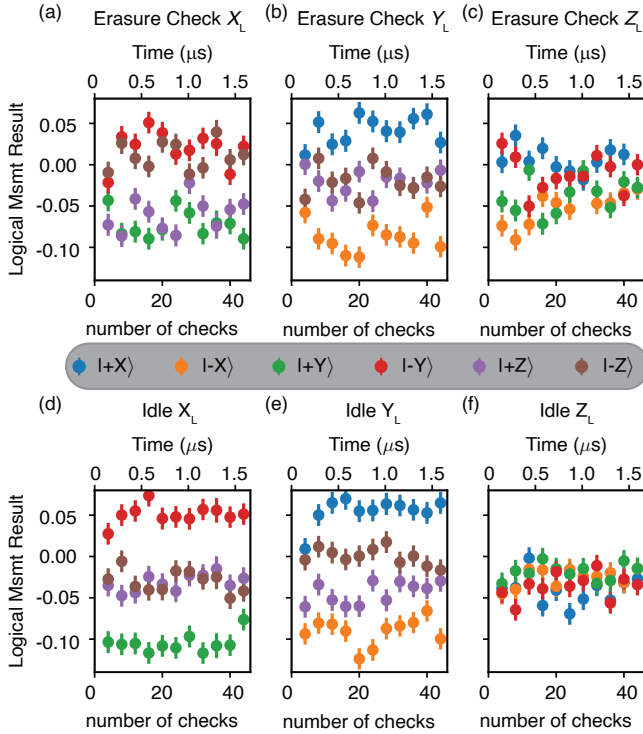
FIG. 12. **Logical measurement results breakdown.** (a-c) Residual X, Y and Z logical measurement results for different initial states after passing n mid-circuit erasure checks. (d-e) Same data when we instead idle for the duration of the beamsplitter and transmon pulses.

of suffering transmon-induced Pauli errors, and (c) the likelihood of transmon-induced erasures and transmon-induced Pauli errors scale differently for transmon dephasing and transmon relaxation.

We can show this using a family of square transmon pulses that has consistent ideal performance in the absence of collapse operators while giving us the freedom to sweep the pulse duration. As a starting point, we center the pulse on the $N = 0$ transition frequency, with a duration T_p chosen to put a notch in the spectrum at the central $N = 1$ transition

$$T_\mathrm{p} = \frac{2\pi}{|\chi|}\sqrt{4n^2 - 1}, \tag{I1}$$

where the integer n indicates the number of detuned Rabi oscillations the transmon undergoes when the cavities are in $|0, 1\rangle$ or $|1, 0\rangle$. Meanwhile we can also choose the beamsplitter amplitude to ensure that the oscillator states approximately return to their starting positions at the end of T_p:

$$g_\mathrm{bs} = |\chi|\sqrt{\frac{m^2}{4n^2 - 1} - \frac{1}{4}}, \tag{I2}$$

where the integer m denotes the number of detuned beamsplitter oscillations the oscillator states undergo during the pulse duration. We then locally optimize the
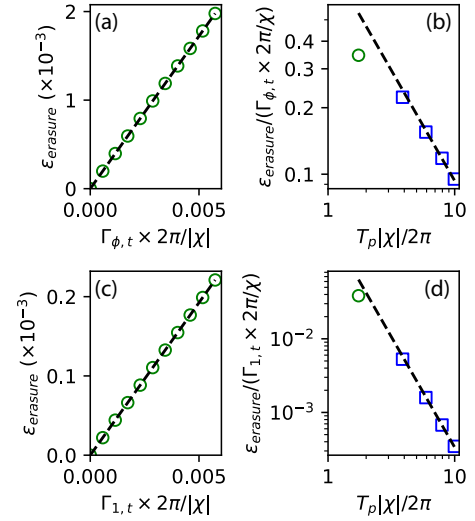

FIG. 13. **Scaling of erasures arising from transmon errors.** (a) Simulated average Pauli error rate during erasure check as a function of the normalized transmon dephasing rate. Data is shown for a pulse duration $T_\mathrm{p} = 1.7 \times 2\pi/|\chi|$ and a beamsplitter rate $g_\mathrm{bs} = 2.8|\chi|$ with the dashed line indicating a quadratic fit. (b) Linear component extracted from these quadratic fits (i.e. sensitivity of the erasure rate to changes in the normalized transmon dephasing rate) for pulses of varying duration. Black dashed line shows $1/x$ fit to last three points. The green point corresponds to the specific choice of pulse duration shown in the left column. (c)-(d) Same analysis for transmon relaxation-induced errors, where the fit in (d) is now to a $1/x^3$ curve.

beamsplitter amplitude and pulse duration to suppress coherent errors from the above approximations. The pulses found are very close to these guessed parameters with negligible coherent errors.

Erasures due to transmon errors

We would like to know how often (due to transmon errors) the transmon will end up in $|e\rangle$ given that the state was initialized in $|\psi\rangle_L \otimes |g\rangle$, averaged over dual-rail logical cardinal states $|\psi\rangle_L$. These are the false positives we see due to transmon dephasing and relaxation.

We choose to set $m = 5n$, corresponding to $|g_\mathrm{bs}/\chi| \approx 2.5$, and sweep the value of n from 1 to 5. For each choice of the pulse duration, we can sweep the transmon dephasing (relaxation) rate $\Gamma_{\phi,t}$ ($\Gamma_{1,t}$) and simulate how the erasure rate varies. We fit the resulting curve to a quadratic (see an example in Fig. 13a,c) and extract the linear component. The resulting slopes are then plotted as a function of pulse duration in Fig. 13b,d.

For transmon dephasing we find that

$$\epsilon_\mathrm{erasure} \propto \frac{\Gamma_{\phi,t}}{\chi}\left(T_\mathrm{p}\chi\right)^{-1}, \tag{I3}$$
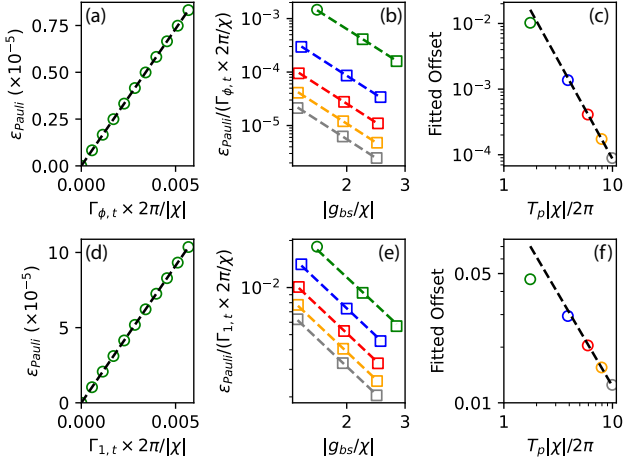
FIG. 14. **Scaling of Pauli errors during erasure check arising from transmon errors.** (a) Simulated average Pauli error rate during erasure check as a function of transmon dephasing rate. Data is shown for a pulse duration $T_\mathrm{p} = 1.7 \times 2\pi/|\chi|$ and a beamsplitter rate $g_\mathrm{bs} = 1.6|\chi|$ with dashed line indicating a quadratic fit. (b) Linear slope extracted from quadratic fits for pulses of varying $|g_\mathrm{bs}/\chi|$ and pulse duration (indicated by the colors). Within each pulse duration, the curves are fit to $A \times (g_\mathrm{bs}/\chi)^{-4}$ where the offset A is the free parameter. (c) Fitted offsets A as a function of pulse duration. Dashed black line shows a cubic fit. (d)-(f) Same analysis for transmon relaxation errors. The fit in (e) is now to $A \times (g_\mathrm{bs}/\chi)^{-2}$ and the fit in (f) is a linear fit.

whereas for transmon relaxation errors we find good agreement to

$$\epsilon_\mathrm{erasure} \propto \frac{\Gamma_{1,t}}{\chi} \left(T_\mathrm{p}\chi\right)^{-3}. \tag{I4}$$

Provided that $g_\mathrm{bs} \gtrsim |\chi|$, we find that the beamsplitter rate has a negligible impact on the erasure rate. However, it will have a significant impact on the Pauli error rates.

Logical Pauli errors due to transmon errors

We can perform the same analysis for Pauli errors, averaging the induced Pauli error rate over all dual-rail cardinal state as we sweep the transmon error rates (Fig 14a,d). To highlight the effect of the beamsplitter rate, we sweep both the pulse duration (via n) and g_bs (by using three different values set by $m = 3n, 4n, 5n$). Fig. 14b,e show the slope of Pauli errors versus transmon error rate (obtained from the left panels) as a function of $|g_\mathrm{bs}/\chi|$.

Increasing the beamsplitter amplitude quickly suppresses Pauli errors, with fits showing excellent agreement to a quartic suppression with $|g_\mathrm{bs}/\chi|$ for transmon dephasing, and a quadratic suppression for transmon decay.

The colored bands indicate pulses with the same pulse duration. Increasing the pulse duration also clearly has a beneficial impact on transmon-induced Pauli errors. To look at this scaling, we plot the offset of the previous polynomial fits versus $|g_\mathrm{bs}/\chi|$ and see how they vary as a function of $T_\mathrm{p}\chi$. These offsets are shown in Fig. 14c,f. The dashed lines show fits of the last three points to a cubic (transmon dephasing) or linear (transmon relaxation). While there is slight deviation from the fit at short pulse durations, we find that to good approximation:

$$\epsilon_\mathrm{Pauli} \propto \frac{\Gamma_{\phi,t}}{\chi} \left(\frac{g_\mathrm{bs}}{\chi}\right)^{-4} (T_\mathrm{p}\chi)^{-3} \propto \frac{\Gamma_{\phi,t}}{g_\mathrm{bs}^4 T_\mathrm{p}^3}, \tag{I5}$$

in the case of transmon dephasing, and

$$\epsilon_\mathrm{Pauli} \propto \frac{\Gamma_{1,t}}{\chi} \left(\frac{g_\mathrm{bs}}{\chi}\right)^{-2} (T_\mathrm{p}\chi)^{-1} \propto \frac{\Gamma_{1,t}}{g_\mathrm{bs}^2 T_\mathrm{p}}, \tag{I6}$$

in the case of transmon relaxation.

Appendix J: Comparison to joint-parity-based erasure check proposal

The erasure check presented in this work differs from approach initially proposed in [42, 43], which relies on measuring the parity of the combined photon number in the two oscillators. Here we will show how to interpret this joint-parity scheme in the context of the driven spectroscopy shown in this paper, and compare the expected performance of this scheme to the erasure check presented here.

Spectroscopic interpretation of joint-parity measurement

The joint-parity measurement consists of two $\pi/2$ pulses on the qubit, separated by a beamsplitter drive with amplitude $|g_\mathrm{bs}/\chi| = \sqrt{3}/2$ applied for a duration of $2\pi/\chi$, followed by a transmon readout. This scheme is analogous to the parity-map protocol for a single oscillator, where the addition of the beamsplitter drive allows the measurement qubit to accumulate an equal phase from photons in either oscillator.

When the beamsplitter amplitude is set to $|g_\mathrm{bs}/\chi| = \sqrt{3}/2$, the expression for the transition frequencies in Eq. 8 simplifies considerably:

$$\omega_{\delta m|g_\mathrm{bs}/\chi=\sqrt{3}/2} = \left(\frac{N}{2} + \delta m\right)\chi, \quad \delta m = -N, \dots, N. \tag{J1}$$

From this, we can see that for odd (even) N, all the transition frequencies lie at odd (even) integer multiples of $\chi/2$. This is higlighted in Fig. 15.

If we assume that the $\pi/2$-pulses are instantaneous, then the qubit pulse takes the form

$$f(t) \propto \delta\left(t + \frac{\pi}{\chi}\right) + \delta\left(t - \frac{\pi}{\chi}\right), \tag{J2}$$
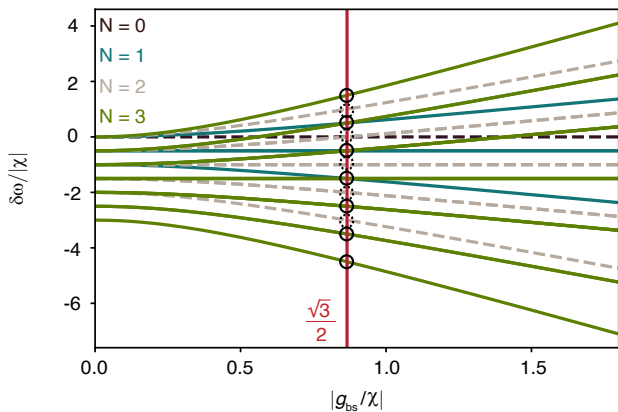
FIG. 15. **Spectroscopic interpretation of joint-parity measurement.** The ancilla transition frequencies for a 'symmetric' beamsplitter detuning $\Delta = \chi/2$ and total photon number in the two coupled oscillators of $N = 0$ (black), $N = 1$ (green), $N = 2$ (beige, dashed) or $N = 3$ (olive). The vertical red line highlights the beamsplitter amplitude $|g_{\mathrm{bs}}/\chi| = \sqrt{3}/2$, which is used for the joint parity measurement. At this value, all of the even N transitions lie at a detuning equal to an even integer multiple of $\chi/2$ (dashed circles) and all of the odd N transitions lie at odd integer multiples of $\chi/2$ (solid circles). The ancilla pulse ideally consists of two delta functions separated by $2\pi/\chi$, whose frequency spectrum has nodes at the solid circles and antinodes at the dashed circles.

which has the Fourier transform:

$$F(t) \propto \cos\left(\frac{\pi\omega}{\chi}\right). \tag{J3}$$

The frequency spectrum of the transmon pulse therefore has antinodes (nodes) at all of the transition frequencies with even (odd) joint-photon-number-parity. This provides an alternative interpretation of why the previously proposed scheme indeed maps the joint-photon-number parity of the two oscillators onto the transmon ancilla state.

Performance comparison

The joint-parity measurement may be used with a two-level ($|g\rangle - |e\rangle$) or three-level ($|g\rangle - |f\rangle$) qubit. In both cases, transmon dephasing errors are fully error-detectable and do not contribute to the post-selected Pauli error rate. In the $|g\rangle - |f\rangle$ case, the $|e\rangle$ level is reserved as a flag state, allowing error detection of transmon relaxation errors as well. However, these flagged transmon errors do contribute to the overall erasure rate.

To compare the performance of the joint-parity scheme with the photon-number-selective scheme demonstrated in this work, we use master equation simulations considering only transmon relaxation and dephasing errors. In both cases we assume no limitation on pulse bandwidth,

	Joint-photon-number		Joint-parity	
	$|g\rangle - |e\rangle$	$|g\rangle - |f\rangle$	$|g\rangle - |e\rangle$	$|g\rangle - |f\rangle$
p_{erasure}	0.61%	2.92%	1.40%	5.43%
p_{Pauli}	0.16%	0.10%	0.23%	0.0048%

TABLE I. **Simulated error rates during erasure check due to transmon relaxation and dephasing only.** The transmon-induced erasure probability p_{erasure} and the transmon-induced Pauli error probability p_{Pauli} (post-selected on not detecting an erasure) obtained from master equation simulations of the joint-photon-number measurement scheme demonstrated in this work, and the previously proposed joint-photon-number-parity measurement [42]. Both schemes are simulated using a $|g\rangle$-$|e\rangle$ ancilla and with a $|g\rangle$-$|f\rangle$ ancilla where the $|e\rangle$ level is reserved as a flag state.

and use the system parameters in Table II, assuming $T_1^{fe} = T_1^{eg}/2$ and that $T_\phi^{gf} = T_\phi^{ge}/4$ (as would be the case for a harmonic oscillator). While typically $\chi_{gf} \approx 2\chi_{ge}$, this larger χ can only be accessed by commensurately increasing gbs. Supposing that χ can be optimized based on the available g_{bs}, we assume these values to be equal in each case. For the joint-photon-number-selective scheme we use a square pulse shape with $g_{\mathrm{bs}} = 1.04$ MHz and $T_p = 1.699$ µs.

The results shown in Table I show that when using a $|g\rangle - |e\rangle$ transmon, the joint-photon-number measurement demonstrates fewer transmon-induced false positive erasures and fewer transmon-induced Pauli errors than the joint-parity measurement.

The benefit of using a $|g\rangle - |f\rangle$ transmon to suppress Pauli errors due to transmon decoherence is substantially more pronounced for the joint-parity measurement since the scheme already protects against transmon dephasing errors. This results in a very low transmon-induced Pauli error rate. However it does so at the expense of a large additional erasure rate.

As discussed in [42], in reality the finite bandwidth of the beamsplitter pulse during the joint-parity sequence would necessitate the use of optimal control pulses to ensure that the pulse completes within $2\pi/|\chi|$. Alternatively, without pulse shaping, delays can be added to make the total pulse duration $4\pi/|\chi|$. However this comes at the cost of twice as many transmon-induced errors and more idling errors.

When the error rates are dominated by transmon decoherence (as opposed to intrinsic errors), the greater suppression of Pauli errors offered by the joint-parity scheme ensures a greater bias towards erasure errors during the check. This could therefore enable a higher error-correcting threshold (closer to 7% than the approximately 4% in this work). However, this improvement comes with a very large increase in the overall error rate. Since the performance of an erasure code depends on the ratio of the error rate to the threshold, the joint-photon-number-selective scheme is preferred.

One caveat is that the joint-photon-number-parity

measurement also catches heating to dual-rail states with $N = 2$, unlike the version of the check presented in this work. These uncaught events are rare but are especially damaging for error correction. However, one could adapt the photon-number-selective scheme by adding a 2-photon-selective transmon pulse to prevent this.

Appendix K: Predicting erasure check performance with longer coherences

We simulate the estimated the performance improvement with longer transmon coherences ($T_1 = T_\phi = 200$ μs) and cavity lifetimes ($T_{1,A} = T_{1,B} = 1000$ μs). We use the same dispersive shift χ and readout duration τ_{RO} used in this experiment. We choose a square pulse shape (without ramps) with pulse duration $T_p = 1.699$ μs and beamsplitter rate $g_{bs}/2\pi = 1.038$ MHz, the closest ideal pulse parameters to the experimentally optimized parameters we use.

For each dual-rail input state, we initialize the state, wait τ_{RO} (to capture the effect of photon loss during readout), and then perform the pulse. From this final state we extract the probability of the cavities being in $|0, 0\rangle$ (the intrinsic erasure rate, $p_{erasure}^{intrinsic} = 0.334\%$) and the probability of suffering a Pauli error given that the transmon was in $|g\rangle$ (the induced Pauli error rate, $p_{Pauli}^{induced} = 0.035\%$). To obtain the false positive rate, we run the same simulation without the cavity collapse operators, and measure the probability of the transmon ending in $|e\rangle$ ($p_{FP} = 0.164\%$).

An effect not included in these simulations that will start to become important is measurement-induced dephasing on the dual-rail qubit due to photons in the readout resonator. This dephasing rate is given by

$$\Gamma_\phi = \frac{\bar{n}\kappa\chi^2}{\kappa^2 + \chi^2}, \tag{K1}$$

where $\bar{n}$ is the average photon number in the readout resonator, $\kappa = 2\pi \times 1.77$ MHz is its linewidth, and χ is the dispersive shift between the resonator and Bob's cavity [79]. We estimate χ from the anharmonicity of Bob's transmon $\alpha = -2\pi \times 185$ MHz, its dispersive shift on Bob's cavity $\chi_{ct} = -2\pi \times 1.066$ MHz and its dispersive shift on Bob's readout resonator $\chi_{tr} = -2\pi \times 0.86$ MHz, as

$$\chi \approx \frac{\chi_{tr}\chi_{ct}}{\alpha} \approx 2.5 \text{ kHz.} \tag{K2}$$

We then take $\bar{n} \approx 10$ photons for 1 μs during the readout. Putting these values into Eq. K1, and using $p_{Pauli} = \Gamma_\phi t/2$, we obtain an approximate Pauli error rate per erasure check due to measurement-induced dephasing of $p_{Pauli} \approx 0.01\%$.

Appendix L: Requirements on $|g_{bs}/\chi|$ at larger photon numbers

As discussed in the Methods, the state of the two coupled oscillators with N photons between them is analogous to the state of a single spin with $S = N/2$. In this picture, we saw that the matrix element for a transition from the eigenstate of $\hat{H}_g$ (the Hamiltonian when the qubit is in $|g\rangle$) with magnetic quantum number m_g to the m_e eigenstate of $\hat{H}_e$ (the Hamiltonian when the qubit is in $|e\rangle$) can be expressed in terms of the small Wigner d-matrix, $|d_{m_g,m_e}^{N/2}(\delta\theta)|$. Furthermore, we saw that for the choice of beamsplitter detuning $\Delta = \chi/2$, all the transition frequencies for the same total photon number N and change in magnetic quantum number $\delta m \equiv m_e - m_g$ become degenerate, allowing us to see joint-photon number-splitting.

However, by looking at the elements of the Wigner d-matrix, we find that these transitions with the same frequency do *not* always have the same transition matrix element. The transition matrix elements for $N = 1$ and $N = 2$ are shown in Fig. 16. Since the d-matrix obeys the symmetry $|\delta_{i,j}^{N/2}(\beta)| = |\delta_{j,i}^{N/2}(\beta)|$, we find that $|M_{\frac{1}{2}\to\frac{1}{2}}| = |M_{-\frac{1}{2}\to-\frac{1}{2}}|$ and so all the degenerate transitions for $N = 1$ also have the same matrix element. By contrast, in the 2-photon manifold we find that the transition matrix element from $m_g = 0$ to $m_e = 0$ is given by

$$\left| M_{0\to0}^{(N=2)} \right| = \frac{|\epsilon|}{2} \left| d_{0,0}^1(\delta\theta) \right| = \frac{|\epsilon|}{2} \left| \cos(\delta\theta) \right|, \tag{L1}$$

whereas the matrix element from $m_g = \pm 1$ to $m_e = \pm 1$ is

$$\left| M_{\pm1\to\pm1}^{(N=2)} \right| = \frac{|\epsilon|}{2} \left| d_{\pm1,\pm1}^1(\delta\theta) \right| = \frac{|\epsilon|}{2} \frac{|1 + \cos\delta\theta|}{2}, \tag{L2}$$

where $\delta\theta = 2\arctan(\chi/2g_{bs})$.

The implication is that joint-photon-number-selective qubit pulses for $N \geq 2$ will proceed at different rates for different oscillator states within the N-photon manifold, thereby introducing coherent errors. However, at large beamsplitter rates, the difference in rates becomes suppressed as $(\chi/g_{bs})^2$.

Appendix M: System parameters

Table II lists the coherences and thermal populations of the modes in the system, as well as the dispersive coupling strength between each cavity and its ancilla transmon.

The SNAIL parameters are measured at the external flux bias used to obtain the data ($\Phi_{ext} = 0.334\Phi_0$). Its low T_{2R} is due to low-frequency flux noise away from its flux sweet spot, but introducing an echo increases its T_{2E} coherence substantially.

As mentioned in the Methods, this device was previously used in Chapman et al. [40]. One change that

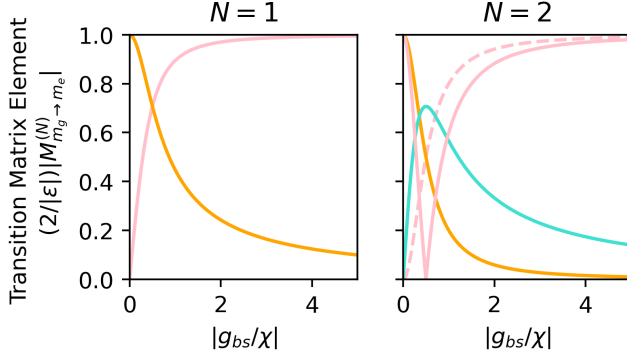| Parameter	Unit	Alice		SNAIL	Bob	
		Transmon	Cavity		Cavity	Transmon
ω	GHz	4.783 720(1)	2.973 421(1)	5.192 429(1)	6.926 368(2)	5.402 311(3)
T_1	μs	123.4 ± 0.9	347 ± 2	69 ± 2	108.5 ± 0.6	42.4 ± 0.4
T_{2R}	μs	31.6 ± 0.9	441 ± 23	3.0 ± 0.2	165 ± 7	33.1 ± 0.9
T_{2E}	μs	40.5 ± 0.7	-	14.4 ± 0.6	-	63.2 ± 0.8
n_{th}	%	0.23 ± 0.06	3.82 ± 0.10	4.1 ± 0.6	0.53 ± 0.03	0.52 ± 0.07
χ	MHz	0.7773(4)			1.0660(9)	

TABLE II. **Measured system parameters.**

FIG. 16. **Theoretical transition matrix elements.** Predicted transition matrix elements (normalized for the amplitude of the spectroscopy tone) for a) $N = 1$ and b) $N = 2$ total photons in the two oscillators, when the beamsplitter detuning is symmetric ($\Delta = \chi/2$). Colors distinguish transitions that are degenerate in frequency with $\delta m = 0$ (pink), $\delta m = \pm 1$ (orange) and $\delta m = \pm 2$ (turquoise). For $N = 2$ and $\delta m = 0$, the matrix element for the transition $0 \to 0$ (solid) differs from those for $\pm 1 \to \pm 1$ (dashed).

was made to the device was an increase the coupling to

Bob's readout resonator, increasing its linewidth from $\kappa = 2\pi \times (0.89 \pm .05)$ MHz to $\kappa = 2\pi \times (1.77 \pm .04)$ MHz. This allowed us to reduce the duration of the readout and therefore reduce idling errors.

As discussed in Chapman et al. [40], the oscillator coherence times in this system are lower than is typical for $\lambda/4$ superconducting stub cavities. The bare cavities, without any chips inserted, were measured to have lifetimes of 2.3 ms and 460 μs, respectively. These lifetimes therefore do not explain the measured coherences. Meanwhile, finite element simulations including Purcell loss out of the system ports and conductor loss due to the magnetic coil suggest that these effects should not limit the oscillator lifetimes to less than 2.66 ms and 810 μs. We often find large cooldown-to-cooldown variations in T_1 when changes are made inside the package (e.g. changing a coupling pin length, or re-inserting a chip), even when these changes are not expected to be related. For example, the slight increase in *Bob's* readout pin length described earlier coincided with a decrease in *Alice's* cavity T_1 from 485 ± 14 μs to 347 ± 2 μs. Meanwhile, on another cooldown with all of the chips inserted, Bob's cavity T_1 was measured to be 265 ± 27 μs. This difficult-to-predict behavior could point to unexpected package modes as the culprit.