

High-On-Off-Ratio Beam-Splitter Interaction for Gates on Bosonically Encoded Qubits

Benjamin J. Chapman^{1,2,*} Stijn J. de Graaf^{1,2,†} Sophia H. Xue^{1,2} Yaxing Zhang,^{1,2}
James Teoh,^{1,2} Jacob C. Curtis^{1,2} Takahiro Tsunoda,^{1,2} Alec Eickbusch^{1,2}
Alexander P. Read^{1,2} Akshay Koottandavida^{1,2} Shantanu O. Mundhada,^{1,2} Luigi Frunzio^{1,2}
M.H. Devoret,^{1,2} S.M. Girvin^{1,2} and R.J. Schoelkopf^{1,2}

¹*Departments of Physics and Applied Physics, Yale University, New Haven, Connecticut 06511, USA*

²*Yale Quantum Institute, Yale University, New Haven, Connecticut 06511, USA*



(Received 21 December 2022; accepted 5 May 2023; published 28 June 2023)

Encoding a qubit in a high-quality superconducting microwave cavity offers the opportunity to perform the first layer of error correction in a single device but presents a challenge: how can quantum oscillators be controlled while introducing a minimal number of additional error channels? We focus on the two-qubit portion of this control problem by using a three-wave-mixing coupling element to engineer a programmable beam-splitter interaction between two bosonic modes separated by more than an octave in frequency, without introducing major additional sources of decoherence. Combining this with single-oscillator control provided by a dispersively coupled transmon provides a framework for quantum control of multiple encoded qubits. The beam-splitter interaction g_{bs} is fast relative to the time scale of oscillator decoherence, enabling over 10^3 beam-splitter operations per coherence time and approaching the typical rate of the dispersive coupling χ used for individual oscillator control. Further, the programmable coupling is engineered without adding unwanted interactions between the oscillators, as evidenced by the high on-off ratio of the operations, which can exceed 10^5 . We then introduce a new protocol to realize a hybrid controlled-SWAP operation in the regime $g_{bs} \approx \chi$, in which a transmon provides the control bit for the SWAP of two bosonic modes. Finally, we use this gate in a SWAP test to project a pair of bosonic qubits into a Bell state with measurement-corrected fidelity of $95.5\% \pm 0.2\%$.

DOI: [10.1103/PRXQuantum.4.020355](https://doi.org/10.1103/PRXQuantum.4.020355)

I. INTRODUCTION

Bosonic codes [1–6]—quantum error-correction schemes in which a qubit is encoded in the multiple levels of an oscillator—offer the chance to perform error correction with a single quantum mode. Along with the long lifetimes of superconducting microwave cavities [7,8] and their simple error models, this hardware-efficiency suggests their use as a first layer of error correction for future quantum processors [9].

Storing quantum information in highly coherent oscillators presents a control challenge: the requisite nonlinearity must be introduced without adding significant error

channels or unwanted Hamiltonian terms. In circuit-QED implementations [10], universal control of a single bosonic qubit can be accomplished with resonant drives on an oscillator and a dispersively coupled transmon [11–15] but making this control first-order insensitive to errors on the ancilla [16,17], or developing alternative control methods based on noise-biased ancillas [18,19], is an area of active research. For example, when the ancilla is excited, ancilla decay can dephase the oscillator. In practice, this is frequently the dominant error in these systems [20–23].

To perform two-qubit operations, another ancilla can be dispersively coupled to a pair of oscillators. Manipulating the state of the ancilla provides an immediate path toward operations between the bosonic qubits [16,24] but introduces additional error pathways when it decays or dephases, due to its entanglement with the oscillators.

An alternative approach for generating two-qubit operations is to engineer a purely bilinear oscillator-oscillator coupling with a circuit in which the quantum degrees of freedom need not be explicitly excited or entangled with the oscillators [25]. While such a scheme only allows Gaussian interactions between the oscillators, it is largely

*benchapman@microsoft.com

†stijn.degraaf@yale.edu

‡These authors contributed equally to this work.

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

immune to decay or dephasing errors on the coupler, since ideally the coupler is only virtually excited. Heating of the coupler is then the only coupler error that causes a cavity error. The non-Gaussian resource needed for this approach can be provided by the ancillas used for single-oscillator control and therefore allows two-qubit gates to be performed without adding additional error-prone control hardware. This approach is analogous to well-known paradigms in linear optics in which universal control of continuous-variable systems is accomplished with Gaussian two-mode interactions and a non-Gaussian single-mode resource [26,27] but here it benefits from the comparatively strong nonlinearity of the dispersive shift.

Such a recipe for two-qubit operations on bosonically encoded qubits has been recently employed in the form of controlled-SWAP (CSWAP) and exponential-SWAP (ESWAP) gates [28]. The performance of these demonstrations, however, has been limited by the speed and fidelity of the engineered beam-splitter interaction [29]. Recent work suggests that spurious Hamiltonian interactions of the transmon coupling element have been responsible for these limitations [30].

At the same time, efforts throughout the superconducting qubit community have shown the promise of designing coupling elements (sometimes called tunable couplers) to mediate interqubit interactions. These devices enable strong coupling while suppressing spurious ZZ interactions between qubits and can achieve on-off ratios exceeding 10^3 [31,32]. Suppressing such unwanted interactions is critical for quantum error correction, which often relies on the assumption that errors are uncorrelated.

With this in mind, we improve on the efforts of Refs. [28,29,33] by designing a coupler based on a superconducting nonlinear asymmetric inductive element (SNAIL) dipole element [34]. The potential energy of the SNAIL can be made asymmetric with a static external flux to allow a pump-controlled beam-splitter interaction via three-wave mixing [35], similar to that achieved with superconducting quantum interference devices (SQUIDs) driven by a monochromatic flux pump [36,37]. Fine tuning the potential of the SNAIL coupler also allows *in situ* suppression of higher-order terms in the coupler and oscillator Hamiltonians, such as cross-Kerr (ZZ-like) interactions between the oscillators. SNAIL-based couplers constructed in this way are therefore complementary to SQUID-based couplers driven by bichromatic flux pumps, which offer less control of the static coupler Hamiltonian but suppress many unwanted pumped interactions with symmetry [38].

Here, we use a SNAIL-based coupler to realize a programmable beam-splitter interaction between two bosonic modes with over an octave separation in frequency [Fig. 1(a)]. The interaction is fast relative to the time scale of oscillator decoherence, enabling over 10^3 beam-splitter operations per coherence time. Further, the programmable

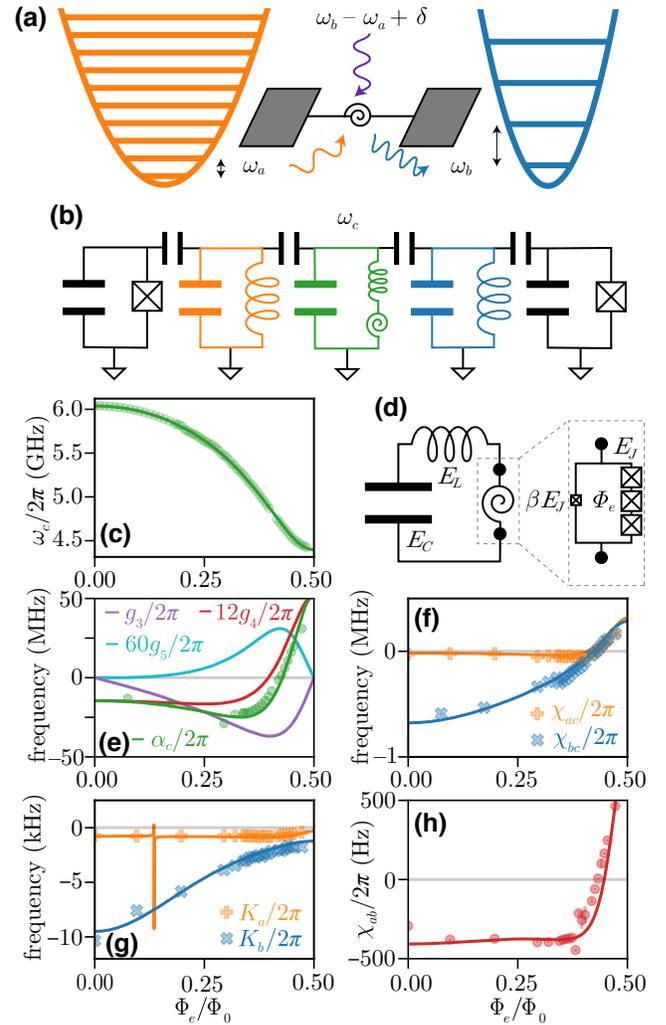


FIG. 1. Concept and coupler characterization. (a) A SNAIL-based [34] three-wave-mixing coupler exchanges photons between two high-quality superconducting microwave cavity modes with resonant frequencies of 3 and 7 GHz, at a rate proportional to the pump amplitude and with a phase set by the pump. The process is enabled by a pump photon (purple) with frequency equal to the 4-GHz detuning between the oscillators. (b) The effective circuit diagram. The cavity modes are each dispersively coupled to a transmon for readout and control. Readout resonators and Purcell filters for the transmons are not pictured. (c) Measurements of the coupler frequency ω_c (circles) as a function of the external flux bias in units of $\Phi_0 = h/2e$. The solid line is a fit to the circuit model in (d) [39]. The extracted parameters are $E_J/h = 90.0 \pm 0.3$ GHz, $E_L/h = 64 \pm 2$ GHz, $E_C/h = 177 \pm 2$ MHz, and SNAIL junction asymmetry $\beta = 0.147 \pm 0.001$. (e) The predicted Hamiltonian parameters, g_3 , g_4 , and g_5 , and the anharmonicity of the coupler α_c (lines) and the measured anharmonicity (circles). (f)–(h) Measurements (markers) and predictions (lines; see Appendix D) of (f) the cavity-coupler cross-Kerrs, (g) the cavity self-Kerrs, and (h) the cavity-cavity cross-Kerr. The resonance in (g) occurs when $\omega_c = 2\omega_a$.

coupling is engineered without adding unwanted interactions between the oscillators, as evidenced by the large

ratio between the beam-splitter rate and the magnitude of the strongest spurious cavity-cavity interactions during idle periods, here called the on-off ratio. We find that the on-off ratio is always above 2×10^3 and can exceed 10^5 at some operating points. To demonstrate the utility of the beam-splitter interaction as a primitive for constructing two-qubit gates, we combine it with a dispersively coupled transmon to realize a hybrid CSWAP gate. We use this to create a Bell state with measurement-corrected fidelity of $95.5\% \pm 0.2\%$.

II. COUPLER IMPLEMENTATION AND CHARACTERIZATION

The beam-splitter Hamiltonian in the interaction picture

$$\frac{\hat{\mathcal{H}}_{\text{bs}}}{\hbar} = g_{\text{bs}}(t) \left(e^{i\theta} \hat{a}^\dagger \hat{b} + e^{-i\theta} \hat{a} \hat{b}^\dagger \right), \quad (1)$$

generates the coherent exchange of excitations between two bosonic modes with annihilation operators \hat{a} and \hat{b} . We refer to these modes as Alice and Bob. To realize this interaction between bosonically encoded qubits, we link a pair of high-quality microwave cavities with a three-wave-mixing coupling element and control them by coupling each cavity to its own transmon ancilla. Figure 1(b) shows an effective circuit diagram of the system. For a simplified model of the superconducting package, see Fig. 16 in Appendix Q.

Drawing inspiration from the quantum amplifier community [39,40] and the insights of Refs. [29,30], we adopt a design philosophy that the coupler should act more like a switch than a qubit and strive to minimize all Hamiltonian interactions that are not needed to generate the desired beam-splitter interaction in Eq. (1). Constructing the coupler from a SNAIL dipole element [34] serves this goal, as the Hamiltonian of the coupler may be tuned *in situ* to suppress parasitic fourth-order interactions [39,41] such as self- and cross-Kerr interactions that distort multiphoton cavity states, while preserving the third-order nonlinearity g_3 that enables the pumped three-wave-mixing beam-splitter interaction $g_{\text{bs}} \propto g_3$ (see Appendix F). We emphasize that the amplitude g_{bs} and phase θ of the beam-splitter interaction are controlled by the amplitude and phase of the microwave pump, making the beam-splitter interaction in Eq. (1) fully programmable.

To demonstrate that the SNAIL can be used to suppress parasitic quartic interactions while preserving the desired cubic interaction, we characterize the coupler Hamiltonian following Ref. [39]. We first measure the resonant frequency of the coupler ω_c as a function of the external flux bias [Fig. 1(c)], which we deliver to the SNAIL via a superconducting flux transformer loop (see Appendix A). We then fit a circuit model [Fig. 1(d)] to those measurements [42]. Second quantization of the circuit Hamiltonian

and an expansion to fourth order then yields

$$\frac{\hat{\mathcal{H}}_c}{\hbar} \approx \omega_c \hat{c}^\dagger \hat{c} + g_3 (\hat{c} + \hat{c}^\dagger)^3 + g_4 (\hat{c} + \hat{c}^\dagger)^4, \quad (2)$$

where \hat{c} is the annihilation operator associated with the coupler mode. Here, g_3 and g_4 are the third- and fourth-order nonlinearities of the coupler, both of which are functions of the external flux and the fitted circuit parameters. Figure 1(e) shows the resulting prediction for their dependence on the external flux. (For a discussion on the design choices of the SNAIL circuit parameters, see Appendix B.)

Crucially, it is possible to choose the external flux to suppress g_4 while g_3 remains large. Suppressing the fourth-order interactions provides several advantages: it reduces the self- and cross-Kerr interactions of the cavities, which distort multiphoton cavity states and can therefore be a source of coherent errors; it reduces the anharmonicity of the coupler, which mitigates pump-induced decoherence effects [30]; and it reduces the cross-Kerr interaction between the cavities and the coupler itself, which mitigates cavity dephasing from thermal noise in the coupler.

To verify these predictions, we measure the anharmonicity α_c of the coupler, defined as the difference between its $|g\rangle \leftrightarrow |e\rangle$ and $|e\rangle \leftrightarrow |f\rangle$ transition frequencies. (Here, $|g\rangle$, $|e\rangle$, and $|f\rangle$ denote the first three energy eigenstates of the coupler mode.) Figure 1(e) shows that these measurements agree well with predictions made with perturbation theory in the small parameter $(p\phi_c)$ [43]:

$$\alpha_c = 12 \left(g_4 - 5 \frac{g_3^2}{\omega_c} \right) + \mathcal{O}(p\phi_c)^4, \quad (3)$$

where ϕ_c is the zero-point phase fluctuation of the coupler mode and p is the inductive participation of the SNAIL junctions in this mode (i.e., the fraction of the inductance arising from the junctions rather than the geometric inductance of the leads; see Appendix B).

Conveniently, the nonlinear parameters in the coupler Hamiltonian $\hat{\mathcal{H}}_c$ can also be used to predict the dependence of the joint cavity Hamiltonian $\hat{\mathcal{H}}$ on the external flux applied to the SNAIL, provided that the linear coupling rates of each oscillator to the coupler g_a and g_b are known. We infer these by measuring the dependence of the cavity frequencies on the external flux (see Appendix C). Expanded to fourth order, the joint cavity Hamiltonian takes the form

$$\begin{aligned} \hat{\mathcal{H}}/\hbar \approx & \omega_a \hat{a}^\dagger \hat{a} + \chi_a \hat{a}^\dagger \hat{a} (|e\rangle \langle e|)_a \\ & + \omega_b \hat{b}^\dagger \hat{b} + \chi_b \hat{b}^\dagger \hat{b} (|e\rangle \langle e|)_b + \hat{\mathcal{H}}_{\text{bs}}/\hbar \end{aligned}$$

$$\begin{aligned}
& + \frac{K_a}{2} \hat{a}^{\dagger 2} \hat{a}^2 + \frac{K_b}{2} \hat{b}^{\dagger 2} \hat{b}^2 + \hat{\mathcal{H}}_{\text{Ss}} \\
& + \chi_{ab} \hat{a}^{\dagger} \hat{a} \hat{b}^{\dagger} \hat{b} + \left(\chi_{ac} \hat{a}^{\dagger} \hat{a} + \chi_{bc} \hat{b}^{\dagger} \hat{b} \right) (|e\rangle \langle e|)_c.
\end{aligned} \tag{4}$$

Here, the first two lines contain the desired terms: a pair of oscillators dispersively coupled to transmon ancillas for individual cavity control and a beam-splitter interaction to couple the oscillators together. The cavity self-Kerr and Stark-shift interactions $\hat{\mathcal{H}}_{\text{Ss}}$ on the third line are parasitic, as are the cross-Kerr interactions on the fourth line.

To visualize the flux dependence of $\hat{\mathcal{H}}$, the predicted and measured self- and cross-Kerrs are shown in Figs. 1(f)–1(h). The cavity self-Kerr interactions [Fig. 1(g)] are determined by the nonlinearity they inherit from both the coupler and their transmon ancillas (see Appendix D). In our system, there is no value of external flux that nulls them but future designs could accomplish this by neutralizing the negative anharmonicity from the transmons with a positive anharmonicity from the coupler. To first order, the cavity Stark shifts are proportional to the self-Kerrs and could be nulled in the same way.

In contrast, the cavity cross-Kerr interactions [Figs. 1(f) and 1(h)] are dictated solely by the coupler Hamiltonian. To first order, they are proportional to g_4 , as is evident in the change of their sign when the external flux approaches approximately $0.4\Phi_0$, a feature that allows any of these interactions to be nulled via fine tuning of the external flux. Higher-order corrections, however, make the precise value of external flux that nulls these interactions slightly different from the external flux at which g_4 vanishes [39], as well as the external flux at which the other interactions vanish. These effects mean that many of the quartic interactions can be suppressed simultaneously but only one of the interactions can be finely nulled with a single value of the external flux.

In summary, the external flux bias of the SNAIL coupler serves a dual purpose. First, it breaks the symmetry of the confining potential of the SNAIL, enabling three-wave mixing and thus a switchable coupling between the cavities. At the same time, it provides a means of suppressing the undesired quartic interactions induced by the coupler, which give rise to the always-on cross-Kerr interaction χ_{ab} between the cavities, as well as the anharmonicity of the coupler.

III. THE BEAM-SPLITTER INTERACTION

Ideally, a tunable coupler creates a hierarchy of rates. When the coupler is “on,” the coupling should be strong, both in absolute terms (to reduce algorithm run times) and relative to the time scale of decoherence. Conversely, when the coupler is “off,” it should provide ample isolation; the interaction rates between the coupled modes should be

small relative to the time scale of decoherence. This last condition ensures that errors due to residual coupling are subdominant.

In our system, the cavities are strongly detuned, so the largest cavity-cavity Hamiltonian interaction when the coupler is “off” is the cross-Kerr interaction χ_{ab} , which describes a ZZ-like interaction between the cavities. The desired hierarchy of rates is then

$$\chi_{ab} \ll \tau_{\text{bs}}^{-1} \ll g_{\text{bs}}, \tag{5}$$

where τ_{bs} is a time scale for cavity decoherence in the presence of the beam-splitter interaction. We can quantify the degree to which this hierarchy is enforced with two dimensionless figures of merit. Anticipating its use in pulsed operations, we note that the 50:50 beam-splitter time (half the time required to swap a photon between the cavities) sets a representative time scale for the interaction: $t_{\text{bs}} \equiv \pi/4g_{\text{bs}}$. The first figure of merit is therefore the number of beam-splitter operations that can be performed in one coherence time, $\tau_{\text{bs}}/t_{\text{bs}}$. The second is the on-off ratio g_{bs}/χ_{ab} .

To measure these performance metrics, we turn on the beam-splitter coupling by applying a microwave pump tone with a frequency near to the detuning between the Stark-shifted resonance frequencies of the oscillators. As illustrated in Fig. 1(a), the pump supplies (or removes) the energy required to convert a photon from one oscillator to the other. When the pump frequency is precisely equal to the difference of the two Stark-shifted oscillator frequencies, we call it resonant.

We measure the rate of the beam-splitter interaction and the decoherence of the cavities while the beam splitter is “on” with the sequence in Fig. 2(a). To avoid coherent distortions of the cavity states caused by self- and cross-Kerr interactions, we characterize it in the single-photon manifold of the cavities. The cavities are first prepared in Fock state $|0, 1\rangle$ and a resonant pump is applied for a variable time t . The overlap of each cavity state with vacuum is then measured and the outcomes are correlated shot by shot to infer the photon-number probability distribution of the joint cavity state (for details, see Appendix G). Figure 2(b) shows the probability P_{10} of finding a photon in Alice and none in Bob (orange) or vice versa (P_{01} , blue). The two traces have equal amplitude and opposite phase, indicating that the prepared photon is exchanged between the two bosonic modes. The expected dynamics are [29]

$$\begin{aligned}
P_{01/10} &= \frac{1}{2} e^{-t/\tau} \left(1 \pm e^{-t/\tau_\phi} \cos(2g_{\text{bs}}t + \theta) \right) \\
&+ \mathcal{O}\left(\frac{\kappa_a - \kappa_b}{2g_{\text{bs}}}\right),
\end{aligned} \tag{6}$$

where κ_a and κ_b are the decay rates of Alice and Bob, respectively. Fitting to Eq. (6) (for details, see Appendix

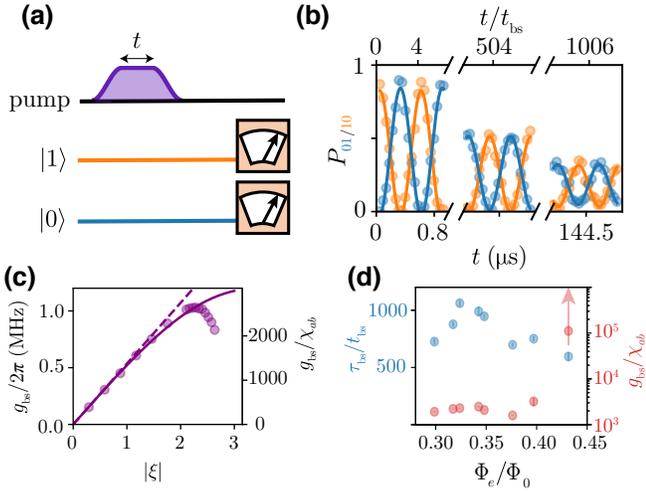


FIG. 2. A three-wave-mixing beam splitter. (a) The pulse sequence for measuring population exchange between the cavities. With an external flux applied to the coupler, the cavities are prepared in $|0, 1\rangle$ and the pump is applied resonantly for a variable time t . The cavities are then measured to determine if they are in vacuum and the outcomes are correlated shot by shot to infer the joint probability distribution of the photon population in the oscillators (see Appendix G). (b) The probability of finding one photon in Alice and none in Bob, P_{10} (orange), and vice versa, P_{01} (blue), as a function of time. The solid lines are fits to Eq. (6). See Fig. 8(c) for the full data set without axis breaks. (c) The fitted beam-splitter rate g_{bs} (left axis) and the on-off ratio (right axis), as a function of the dimensionless pump amplitude ξ , with external flux $\Phi_e/\Phi_0 = 0.35$. The circles show experimental data, the dashed line shows a prediction that is perturbative in the drive strength and treats the coupler Hamiltonian to third order, and the solid line shows a perturbation-theory prediction that accounts for resonant odd-order nonlinear interactions in the coupler Hamiltonian up to g_{17} [44] (for details, see Appendix F). (d) The maximum number of beam splitters that can be achieved per coherence time (left axis) and the corresponding on-off ratio (right axis) at each value of the external flux Φ_e .

H), we extract the beam-splitter rate g_{bs} , the time scale for decay from the single-photon manifold τ , and the time scale for dephasing τ_ϕ in the presence of the beam-splitter drive.

Figure 2(c) shows the beam-splitter rates extracted from a series of these measurements with different dimensionless pump amplitudes ξ , which can be interpreted as the classical displacement of the coupler by the pump. As expected for a three-wave-mixing process, initially the beam-splitter rate scales linearly with the pump amplitude, $g_{bs} \propto g_3 \xi$, in agreement with the perturbation-theory estimate (dashed line). As the amplitude increases, however, the expansion of the coupler Hamiltonian becomes less accurate and g_{bs} deviates from this linear scaling.

One possible explanation for the observed turnaround in g_{bs} is the presence of odd higher-order nonlinearities in

the Hamiltonian of the coupler that can give rise to resonant beam-splitter interactions that compete with g_3 . For example, the first correction of this kind adjusts the beam-splitter interaction by an amount proportional to $g_5 |\xi|^2 \xi$. This correction reduces the magnitude of the beam-splitter rate because g_3 and g_5 have opposite sign. The solid line in Fig. 2(c) shows a prediction of g_{bs} accounting for odd-order rotating-wave-approximation (RWA) processes up to g_{17} [see Eq. (F10)] and indicates that while this explanation is consistent with a slow-down, it fails to predict the sharp turnaround and suggests the importance of capturing non-RWA terms and treating the coupler Hamiltonian nonperturbatively [45]; e.g., in the charge basis.

Nevertheless, at its maximum, $g_{bs}/2\pi$ is greater than 1 MHz. This is approximately a factor of 50 faster than the beam-splitter rate used in Ref. [28] and it allows a 50:50 beam splitter (or half of a SWAP) to occur in $t_{bs} \approx 125$ ns.

To determine the on-off ratio, we compare this beam-splitter rate to the cross-Kerr interaction between the cavities, χ_{ab} . When the external flux is biased to $\Phi_e/\Phi_0 = 0.32$, the measured cavity-cavity cross-Kerr is $\chi_{ab}/2\pi = -390 \pm 10$ Hz (for measurement details, see Appendix E), allowing the on-off ratio to exceed 2×10^3 .

Fits to Eq. (6) also allow extraction of the decay times τ and dephasing times τ_ϕ (see Fig. 8). The fitted values of τ are relatively independent of the pump amplitude and near the inverse of the average cavity decay rate. The dephasing times τ_ϕ are also relatively independent of the pump amplitude and are much longer than the decay times τ . The decoherence in the presence of the beam splitter is therefore cavity-decay limited, which is significant because photon loss is typically the error that bosonic codes are designed to correct [1,3–5]. The dominant error during the beam-splitter gate is therefore correctable.

For times $t \ll \min(\tau, \tau_\phi)$, the fidelity of the time-evolved state diminishes at the composite time scale $\tau_{bs}^{-1} \equiv 1/\tau + 1/2\tau_\phi$ [38]. The unconventional weighting given here to decay and dephasing can be understood from the perspective that unlike a qubit, which cannot decay from its ground state, the single-photon manifold of the cavities is always subject to decay. We take τ_{bs} as a representative time scale for cavity decoherence and compare it to the beam-splitter time t_{bs} to quantify the number of 50:50 beam splitters that can be performed in a single coherence time. With $\Phi_e/\Phi_0 = 0.32$ and a strong drive, $\tau_{bs}/t_{bs} > 10^3$, indicating that high-fidelity gates could be achieved with this beam-splitter interaction.

To determine the flux bias that best enforces the hierarchy in Eq. (5), we repeat this procedure for many values of the external flux. Figure 2(d) shows the number of beam splitters per coherence time (left axis) and the on-off ratio (right axis) as a function of the external flux. For a given value of the external flux, we choose the pump amplitude that maximizes the beam splitters per coherence time and plot the on-off ratio achieved at that

pump amplitude. The quoted figures of merit are therefore achieved simultaneously.

Two clear operation points are visible. The first, near $\Phi_e/\Phi_0 = 0.32$, yields the performance presented in Fig. 2(b) and optimizes $\tau_{\text{bs}}/t_{\text{bs}}$. The second operation point occurs at the external flux $\Phi_e/\Phi_0 = 0.43$ that nulls χ_{ab} and optimizes the on-off ratio g_{bs}/χ_{ab} . To locate it, we resolve the cross-Kerr far below either of the individual line widths of the cavity by making a large displacement in the other cavity (for details, see Appendix E). With this technique, we null $|\chi_{ab}|/2\pi < 10$ Hz. Choosing $\Phi_e/\Phi_0 = 0.43$ still allows over 700 beam splitters per coherence time and due to the suppression of χ_{ab} , elevates the on-off ratio to over 10^5 . In particular, we expect such an operating point to find use in applications with special sensitivity to self- and cross-Kerrs; e.g., controlled-NOT (CNOT) gates between qubits encoded in grid states [2]. More broadly, this exceptional on-off ratio is highly attractive from a quantum error-correction standpoint, as uncorrelated errors are an assumption for many forms of error correction.

IV. HYBRID CONTROLLED SWAP OF BOSONICALLY ENCODED QUBITS

The fast beam-splitter rate engineered with this three-wave-mixing coupler brings the system into a regime where t_{bs} is short not only compared to the cavity coherence times but also to those of the transmon ancillas (see Appendix I). This regime, in which $T_2/t_{\text{bs}} \approx 400 \gg 1$, enables high-fidelity two-qubit gates constructed from the combination of the beam-splitter interaction and single-cavity dispersive control.

To demonstrate this, we perform a hybrid CSWAP gate on two bosonic qubits, in which the transmon coupled to Bob encodes the control bit:

$$\text{CSWAP} \equiv |g\rangle \langle g| \otimes \mathbb{1} + |e\rangle \langle e| \otimes \text{SWAP}. \quad (7)$$

(The beam-splitter interaction can also be used to realize a SWAP controlled on the state of a bosonically encoded qubit—for a proposed sequence for this gate, see Appendix O.) The CSWAP, also known as a quantum Fredkin gate [46], is a key operation for quantum implementations of random-access memory (QRAM) [47,48] and fundamental state-comparison and purification algorithms such as the SWAP test [49–52]. Previously, it has been demonstrated probabilistically in optical platforms [53] and deterministically in circuit QED [28] but performance there has been limited by a less favorable ratio between transmon coherence and the beam-splitter time $T_2/t_{\text{bs}} \approx 7$.

In our system, the control for the CSWAP arises from the dependence of the beam-splitter resonance condition on the frequency of each oscillator [see Fig. 1(a)], which in turn depends on the state of its transmon ancilla through dispersive coupling. For example, when Bob’s ancilla is

excited, the frequency at which the beam-splitter interaction is resonant shifts by χ_b (hereafter, simply χ). Pumping at $\omega_b - \omega_a + \chi$ for $t_{\text{gate}} = 2t_{\text{bs}}$ enacts a SWAP when the ancilla is in $|e\rangle$. The oscillator dynamics when the ancilla is in $|g\rangle$, however, depend on the relative strengths of g_{bs} and χ . A photon evolving under the detuned beam-splitter Hamiltonian $\hat{\mathcal{H}}_{\Delta\text{bs}} \equiv \hat{\mathcal{H}}_{\text{bs}} + \Delta\hat{a}^\dagger\hat{a}$ will swap between the oscillators at a rate Ω and with contrast $C \equiv \max(P_{10}) - \min(P_{10})$ given by

$$\begin{aligned} \Omega &= \sqrt{g_{\text{bs}}^2 + \frac{\Delta^2}{4}}, \\ C &= \frac{g_{\text{bs}}^2}{\Omega^2}, \end{aligned} \quad (8)$$

in analogy to a detuned Rabi drive on a qubit. Δ is the detuning of the beam-splitter drive from the ancilla-state-dependent resonance condition (for CSWAP, $\Delta = \chi$ when the ancilla is in $|g\rangle$). If $g_{\text{bs}} \ll |\chi|$, the contrast is approximately zero and the desired identity operation is achieved, albeit slowly [28]. A new approach is required, however, when g_{bs} approaches $|\chi|$, as now the cavity populations will exchange significantly when the control is in $|g\rangle$ [see Fig. 3(a)].

This problem can be partially resolved by fine tuning the beam-splitter rate to ensure that when the ancilla is in $|g\rangle$, the cavities undergo exactly two detuned SWAP operations during t_{gate} . To give a geometric interpretation of this process, we formalize the analogy to Rabi oscillations by moving to the Heisenberg picture. In that representation, evolution under the detuned beam-splitter Hamiltonian can be expressed as a unitary matrix acting on the vector of operators $\begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix}$. As the operator evolution obeys an SU(2) algebra [54], it can be visualized with trajectories on a Bloch sphere, like those shown in Fig. 3(b) [55]. Under $\hat{\mathcal{H}}_{\Delta\text{bs}}$, the operators rotate at a rate Ω about an axis

$$\vec{n} = \frac{1}{\Omega} \begin{pmatrix} g_{\text{bs}} \cos \theta \\ g_{\text{bs}} \sin \theta \\ \Delta/2 \end{pmatrix}. \quad (9)$$

To ensure that the operator executes precisely two detuned SWAP operations during t_{gate} —equivalent to the identity up to a phase-space rotation on both cavities—when the ancilla is in $|g\rangle$, we set the oscillation rate for a photon evolving under the detuned beam-splitter Hamiltonian to be twice the rate that it is under the resonant beam-splitter Hamiltonian, i.e., $\Omega = 2g_{\text{bs}}$. This condition is satisfied by setting $g_{\text{bs}} = |\chi|/(2\sqrt{3})$ [55]. This is illustrated in Fig. 3(a) when $t_{\text{gate}} = 750$ ns. The phase-space rotation, the angle of which corresponds to half the solid angle enclosed by the red trajectory in Fig. 3(b), is only imparted to the oscillators when the control is in $|g\rangle$. To account for this, we add a delay before and after the beam-splitter drive

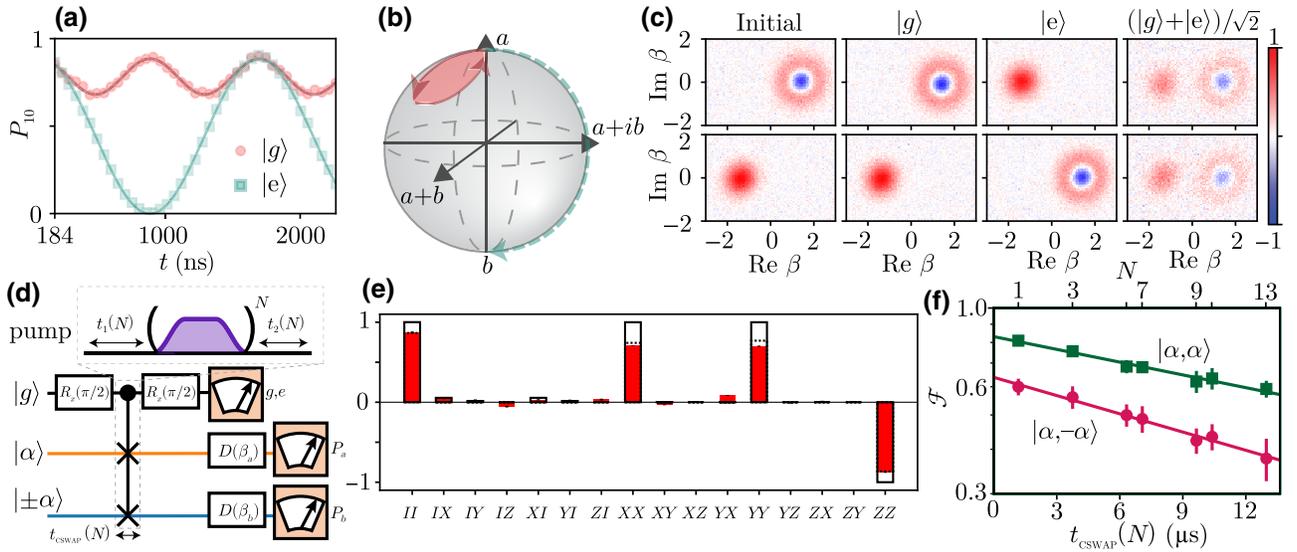


FIG. 3. Concept and performance of the conditional SWAP. (a) The measured probability P_{10} after applying the beam-splitter pump with $\delta = -|\chi|$ and $g_{bs} = \chi/2\sqrt{3}$ for time t , with the control transmon prepared in $|g\rangle$ (red circles) or $|e\rangle$ (turquoise squares). The lines are fits to Eq. (6). (b) An illustration of the evolution of the oscillator operator \hat{a} in the Heisenberg picture during the demonstrated CSWAP protocol. The SU(2) algebra generated by the detuned beam-splitter Hamiltonian $\mathcal{H}_{\Delta bs}$ allows these dynamics to be illustrated on a sphere. When the control is in $|g\rangle$, $\hat{a} \rightarrow \hat{a}$ (solid red). When the control is in $|e\rangle$, the cavity operators exchange and $\hat{a} \rightarrow \hat{b}$ (dashed turquoise). The evolution of the \hat{b} operator follows the same trajectories inverted through the origin. (c) The Wigner tomography of Alice (top row) and Bob (bottom row) initialized in $D(\alpha)|1\rangle$ and $|-\alpha\rangle$ and after performing a CSWAP with the control in $|g\rangle$, $|e\rangle$, and $(|g\rangle + |e\rangle)/\sqrt{2}$. (d) The gate sequence for a SWAP test followed by Wigner tomography, with a variable number N CSWAP gates. Appendix K details how to calibrate the delays $t_1(N)$ and $t_2(N)$ required to account for the control-state-dependent phase. (e) Joint-Pauli measurements on $|\Psi^+\rangle$ (red) with $N = 1$, compared to their ideal (solid) and expected (dashed) values. See Appendix M for details on the tomography and rescaling of YY bar and Appendix N for calculation of the expected values in the presence of decoherence. (f) The Bell fidelity \mathcal{F} (log scale) as a function of the odd number N of CSWAP gates used in the SWAP test (upper x axis) and the duration of the N CSWAP gates including pre- and postdelays (lower x axis), with the oscillators prepared in $|\alpha, -\alpha\rangle$ (circles). The upper x axis is unevenly spaced because the delays $t_1(N)$ and $t_2(N)$ depend on N . The squares show fidelity to $|\alpha, \alpha\rangle$ when oscillators are prepared in $|\alpha, \alpha\rangle$, as a control. The solid lines are linear fits; their slope is used to infer the SPAM-corrected infidelity of the CSWAP, in analogy with randomized benchmarking.

during which the dispersive shift acts, imparting an equal phase-space rotation to both oscillators when the control is in $|e\rangle$. This results in an overall phase-space rotation on the oscillators that is independent of the control state and can thus be removed with local operations (for details, see Appendix K).

In this experiment, the dispersive coupling $\chi/2\pi = -1.104$ MHz and so the required beam-splitter rate $g_{bs} = |\chi|/(2\sqrt{3}) = 2\pi \times 319$ kHz. Accounting for the pre- and postdelays, as well as pulse ramp times to ensure finite bandwidth, the total gate time is $1.3 \mu\text{s}$ and is limited by the dispersive shift χ .

To ensure that the phase correction is properly calibrated, we perform this fast CSWAP protocol on the displaced Fock states $D(\alpha)|1\rangle \otimes D(-\alpha)|0\rangle$ with $\alpha = \sqrt{2}$ [see the left column of Fig. 3(c)]. These multiphoton initial states are both sensitive to phase-space rotations and distinguishable under arbitrary rotation, such that a phase-space rotation on each oscillator cannot be mistaken for a SWAP (as would be the case for $|\alpha, -\alpha\rangle$).

The central columns of Fig. 3(c) show local Wigner tomography on each oscillator after enacting a CSWAP with the control initialized in $|g\rangle$ and $|e\rangle$ and serve as a truth table for the gate. The oscillator states emerge unchanged when the control is in $|g\rangle$ but are exchanged without distortion when the control is in $|e\rangle$.

Importantly, the quantum CSWAP gate is not merely a controlled SWAP on two quantum bits with a classical control; both targets and the control are quantum bits. To demonstrate this, we perform a SWAP test by preparing the control in a superposition state $(|g\rangle + |e\rangle)/\sqrt{2}$ and applying CSWAP followed by a second $\pi/2$ pulse and a measurement of the control [see Fig. 11(a)]. This protocol projects the oscillators into an eigenstate of SWAP, which is entangled with the state of the control. Measuring the control in $|g\rangle$ means that the oscillators are in the $+1$ eigenstate of SWAP; the unitary $(\mathbb{1} + \text{SWAP})/\sqrt{2}$ has acted on the joint oscillator state. The local Wigner tomography, unable to reveal entanglement between oscillators, shows mixed states [right column of Fig. 3(c)] composed equally of the

initial states of Alice and Bob, with half the contrast of the prior columns. In both oscillators, the two displaced Fock states appear on opposite sides of the tomogram, confirming that there is no residual control-state-dependent phase on the oscillator states.

To observe the interoscillator entanglement, we replace local Wigner tomography with a measurement of the joint Wigner function of the oscillators, in which the joint parity is obtained by correlating local parity measurements shot by shot:

$$W(\beta_A, \beta_B) = \langle D(\beta_A, \beta_B) P_A P_B D^\dagger(\beta_A, \beta_B) \rangle. \quad (10)$$

Here, $P_{A,B} \equiv e^{i\pi \hat{n}_{A,B}}$ is the photon-number parity operator in each oscillator. Having validated the phase calibration, we encode a qubit in the oscillator using the mapping $\{|\alpha\rangle, |-\alpha\rangle\}$ to represent logical states $\{|0_L\rangle, |1_L\rangle\}$. In this basis, it is possible to measure two-qubit Pauli operators without resorting to density-matrix reconstruction methods [24,56] (for details, see Appendix M). We prepare the oscillators in $|\alpha, -\alpha\rangle$, such that a SWAP test produces a Bell state in this basis $|\Psi^\pm\rangle \equiv \mathcal{N}(|\alpha, -\alpha\rangle \pm |-\alpha, \alpha\rangle)$, the parity of which is conditioned on the measurement of the control. (Here, $\mathcal{N} \approx 1/\sqrt{2}$ is a normalization factor). Postselecting on measuring $|g\rangle$ yields $|\Psi^+\rangle$.

Visualizing the full joint Wigner function of $|\Psi^+\rangle$ is difficult because it depends on two complex variables and is thus four-dimensional but the signatures of entanglement can be observed in slices in the real-real or imaginary-imaginary planes [see Figs. 11(b) and 11(c)]. These match qualitatively with the ideal simulated case [Figs. 11(d) and 11(e)]. The interference fringes in the imaginary-imaginary slice reveal the intercavity entanglement generated by the SWAP test.

To quantify this entanglement, we directly measure the expectation values of all 16 two-qubit Pauli operators $\langle \Psi^+ | \sigma_{i,A} \sigma_{j,B} | \Psi^+ \rangle$ ($i, j = 1, 2, 3$ or 4) in the coherent-state basis by sampling the joint Wigner function at 16 pairs of displacements (β_A, β_B) [24,56]. For an ideal Bell state, simultaneous measurements of the same single-qubit Pauli operator on each qubit are perfectly positively or negatively correlated, while all other pairs of Pauli measurements are uncorrelated: $|\langle \Psi^+ | \sigma_{i,A} \sigma_{j,B} | \Psi^+ \rangle| = \delta_{ij}$ (the Kronecker-delta). In this experiment, where the basis states for the Bell pair are quasiorthogonal coherent states with $\alpha = \sqrt{2}$, one caveat is that the measured value of $\langle YY \rangle$ is ideally only $|\langle i\pi/8\alpha | -i\pi/8\alpha \rangle|^2 \approx 0.73$ (see Appendix M). This nonideality arises because not all Pauli operators can be perfectly expressed as simple combinations of points in the measured Wigner function. It is the price paid for efficiently extracting the expectation values of the Pauli operators.

In Fig. 3(e), we compare these best-case outcomes for $|\Psi^+\rangle$ to our measurements. The fidelity to the ideal Bell state $\mathcal{F} \equiv 1/4(\langle II \rangle + \langle XX \rangle + \langle YY \rangle - \langle ZZ \rangle)$ is 74.1%

without correcting for state-preparation-and-measurement (SPAM) errors, which exceeds the classical limit of 50% [24].

To identify the origin of the infidelity, we estimate and simulate the effects of transmon and cavity decoherence using independently measured system parameters (for details of the error budget, see Appendix N). The expected values of the joint-Pauli measurements, indicated by the dashed lines in Fig. 3(e), yield an overall estimated fidelity within 2% of the measured value and suggest that the CSWAP gate is decoherence limited. Furthermore, the error budget suggests that the bulk of the errors arise from measurement of the joint oscillator state and not from the CSWAP gate itself.

To confirm this, we perform the Bell-state preparation sequence while replacing the single CSWAP with N CSWAP gates (with N an odd number), as shown in Fig. 3(d). Since CSWAP squares to the identity, this allows us to enhance the effect of gate errors while leaving the measurement unchanged. By fitting the exponential decay of the Bell fidelity as we increase the duration of the N CSWAP sequence, we extract a measurement-corrected infidelity of $4.5\% \pm 0.2\%$. This value, obtained by multiplying the fractional decrease in fidelity per unit time from the fitted slope by the total duration of one CSWAP (including delays), agrees with the predicted value of $4.2\% \pm 0.1\%$ from our error budget, which suggests that cavity and ancilla errors are approximately equally responsible for the infidelity.

As a control experiment, we run the same protocol with $|\alpha, \alpha\rangle$ as our initial cavity state. Since this state is already an eigenstate of SWAP and therefore ideally unchanged by a SWAP test, it is less prone to many of the error mechanisms that limit the fidelity of $|\Psi^+\rangle$ and it results in a lower measurement-corrected infidelity to the target state of $3.2\% \pm 0.2\%$, in agreement with our error budget.

V. CONCLUSIONS

SNAIL-based couplers can expand oscillator-control capabilities without adding significant sources of error because they need only be virtually excited to facilitate the coupling. Here, we show how such a coupler enables a fast and programmable three-wave-mixing beam-splitter interaction that does not degrade oscillator coherence and allows over 10^3 consecutive coherent beam-splitter operations.

Importantly, the strong oscillator-oscillator coupling g_{bs} is achieved without introducing unwanted interactions, as evidenced by its high on-off ratio, which can exceed 10^5 . The coupler therefore satisfies the desired hierarchy of rates $\chi_{ab} \ll \tau_{bs}^{-1} \ll g_{bs}$.

To demonstrate its utility as a primitive for two-qubit bosonic gates, we describe a new protocol that leverages

the beam splitter in a deterministic hybrid CSWAP gate in which a transmon serves as the control qubit. We use this to perform a SWAP test on coherent states with opposite phase, projecting the oscillators into a Bell state with a measurement-corrected infidelity of $4.5\% \pm 0.2\%$.

Looking forward, the hybrid CSWAP operation can be further optimized using established control techniques such as GRAPE [57] or REBURP [58] or implemented with the canonical construction of beam splitter, controlled parity, and inverse beam splitter (see Appendix P). Extending it to allow the control qubit to be provided by a bosonically encoded qubit (e.g., with the sequence proposed in Appendix O) will benefit QRAM implementations with bosonic codes. The beam-splitter interaction engineered in this work is also the first step in performing a universal gate set in circuit-QED dual-rail qubits, where dominant errors can be detected, leaving background Pauli errors that are orders of magnitude smaller [59]. Finally, the principles used to construct the hybrid CSWAP presented here can be generalized to design families of error-detectable two-qubit gates on bosonically encoded qubits [55]. As SNAIL-based couplers need not themselves add significant errors, such gates enable the long coherence times in superconducting cavities to translate to higher two-qubit gate fidelities with bosonic qubits, paving a path to hardware-efficient quantum error correction.

ACKNOWLEDGMENTS

We acknowledge N. Ofek, P. Reinhold, and Y. Liu for their work building the field-programmable gate array (FPGA) firmware and software, V. Sivak for providing the Josephson Array Mode Parametric Amplifier (JAMPA), and V. Joshi, G. Liu, and M. Malnou for their work on the design and assembly of the lumped-element SNAIL parametric amplifiers (LSPAs) used in the experiments. The LSPAs were fabricated by the National Institute of Standards and Technology (NIST) Quantum Sensors Group, with support from the NIST program on scalable superconducting computing. We thank A. Miano for help computing nonlinear SNAIL parameters at higher orders. We also thank N. Frattini, C. Wang, C. Hann, R. Andrews, L. Burkhardt, J. Venkatraman, X. Xiao, Y. Lu, A. Maiti, K. Chou, W. Kalfus, P. Lu, and C. Zhou for useful discussions. Use of facilities was supported by the Yale Institute for Nanoscience and Quantum Engineering (YINQE), Y. Sun, M. Rooks, and the Yale School of Engineering & Applied Science (SEAS) cleanroom. Research was sponsored by the Army Research Office (ARO) under Grant No. W911NF-18-1-0212. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official parties, either expressed or implied, of the Army Research Office (ARO) or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints

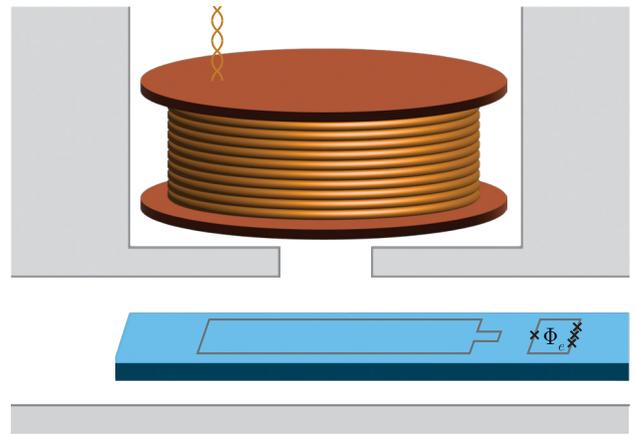


FIG. 4. Flux delivery with a superconducting pick-up loop. A sketched schematic of the superconducting pick-up loop used to transfer magnetic flux from the superconducting coil to the SNAIL loop. The diagram is not to scale. For simplicity, the pads and leads of the SNAIL-coupler circuit are omitted and only the SNAIL loop is shown.

for government purposes notwithstanding any copyright notation herein.

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

APPENDIX A: FLUX DELIVERY IN AN ALL-SUPERCONDUCTING PACKAGE

Flux biasing a SNAIL in an all-superconducting package is challenging because the Meissner effect [60] screens external magnetic fields and because oscillator coherence can be degraded if a magnetic coil wound on a spool of normal metal is located too near to the cavity. One solution to this problem is to use an all-superconducting pick-up loop [61] as an intermediary between the magnetic coil and the SNAIL loop, as illustrated in Fig. 4. To see how the magnetic coil integrates into the superconducting package, refer to Fig. 16. When a dc magnetic field from the coil threads through the pick-up loop, a screening current flows around the loop to counter this flux. The magnetic field from this screening current then threads through the loop of the SNAIL, delivering the flux bias. This process is analogous to the operation of a transformer, except that the properties of superconductors allow it to occur at dc.

Experimentally, we measure that between 20 and 40 mA are required to thread one flux quantum through the SNAIL using this method. To apply this large current without heating the mixing-chamber plate, we solder the superconducting magnet wire (SC-SW-M-0.127 mm Cu-Ni-clad Nb-Ti from Supercon) to a segment of copper cable and

heat sink their joint (and about 10 cm of wire on either side) to an oxygen-free high-conductivity (OFHC) copper bobbin secured to the 4 K plate. The other end of the copper wire is then routed out of the fridge and to a current source at room temperature. The absence of joints in the Nb-Ti wire below 4 K avoids placing a heat load on the dilution unit and allows us to send 200 mA through this line without any change in the mixing-chamber temperature.

APPENDIX B: SYSTEM DESIGN CONSIDERATIONS

The design of the experimental system is guided by enforcing the hierarchy of rates in Eq. (5). The parameter space of the design includes the bare-oscillator frequencies $\omega_{A,B}$, the linear coupling rates between the coupler and the cavities $g_{a,b}$, and the circuit parameters of the coupler: β , E_L , E_J , E_C , the number of large junctions in the shunting array M , and the number of SNAILs N arrayed in series in the coupler. For brevity, we omit discussion here on choices regarding the transmon and readout parameters.

1. Choice of coupler and bare-oscillator frequencies

When choosing the bare-oscillator frequencies, the primary consideration is that the difference of the oscillator frequencies sets the frequency of the beam-splitter pump. It is therefore convenient to parametrize the problem in terms of $\omega_{\pm} \equiv \omega_B \pm \omega_A$. (Without loss of generality, we take $\omega_B > \omega_A$.) To avoid low-order spurious interactions with waveguide modes in the three-dimensional (3D) package, $\omega_- + \omega_B$ should be much less than the waveguide cutoff ω_{wg} of the coupler or transmon tunnels. This consideration constrains ω_- from above. In our package, the waveguide cutoff frequencies are at $\omega_{\text{wg}}/2\pi \approx 20$ GHz, meaning that at least five pump photons are required to convert a photon in Bob into one of the waveguide modes. Ensuring that spurious interactions of this kind are of high order mitigates pump-induced decoherence on the cavities. As a general principle, reducing ω_- as much as possible helps to suppress unwanted pumped interactions.

The sum frequency ω_+ should be chosen to allow for acceptable thermal populations in both cavities—the Bose-Einstein distribution is exponential in frequency, so modest increases in oscillator frequency rapidly suppress residual thermal populations. Additionally, unwanted hybridization with auxiliary modes can be avoided by ensuring that they are not resonant with intermodulation products of the oscillator frequencies.

2. Choice of the linear coupling between the cavities and the coupler

The beam-splitter rate is proportional to the hybridization between the cavities and the coupler, as are the decay

and dephasing rates that the cavities inherit from the coupler. The linear coupling should therefore be chosen to maximize this coupling without meaningfully degrading the coherence of the cavities.

We parametrize the hybridization between each oscillator and the coupler in terms of their detuning $\Delta_{a,b} \equiv \omega_{a,b} - \omega_c$ and their linear coupling rate $g_{a,b}$. The decay rate that the oscillator inherits as a result of the hybridization is

$$\kappa_{a,b} \approx \left(\frac{g_{a,b}}{\Delta_{a,b}} \right)^2 \gamma_c, \quad (\text{B1})$$

where $\gamma_c \equiv 1/T_{1,c}$ is the decay rate of the coupler. (This can be seen by considering a complex bare coupler frequency $\omega_C \rightarrow \omega_C + i\gamma_c$ in Eq. (C3).)

By the same mechanism, the oscillator inherits a dephasing rate from the coupler if the coupler frequency fluctuates stochastically, e.g., because of flux noise in the loop of the SNAIL. How this inherited dephasing rate relates to the fluctuations of the coupler frequency will depend on its noise spectrum.

In the case of white noise with a constant power spectral density S_0 , we can identify the dephasing rate on the coupler as $\gamma_{C\phi} = \omega_C^2 S_0/4$ and the dephasing rate inherited by the oscillator as follows:

$$\kappa_{a,\phi} \approx \frac{g_a^4}{\Delta_a^4} \gamma_{C\phi}, \quad (\text{B2})$$

where we use the fact that $\omega_C - \omega_A \approx \Delta_a$.

It is more likely, however, that the noise spectrum is pink and of the form $S(f) = S_p/f$, in which case the dephasing rate of the coupler is given by $\gamma_{C\phi} = \omega_C \sqrt{S_p} \ln 0.401/f_m$. Here, f_m is a low-frequency cutoff that should be taken as the inverse of the longest time scale in the system and the value 0.401 arises from the Euler-Mascheroni constant [62]. The dephasing rate inherited by the oscillator is then

$$\kappa_{a,\phi} \approx \frac{g_a^2}{\Delta_a^2} \gamma_{C\phi}. \quad (\text{B3})$$

To be conservative, we use Eq. (B3) to model the oscillator dephasing inherited from the coupler.

With the inherited decay and dephasing rates calculated, the optimal degree of hybridization can be calculated with knowledge of the bare oscillator and coupler decoherence rates (and the assumption that all coupler dephasing arises from flux noise). Assuming bare-oscillator coherence at the level of 1 ms, we aim to make the inherited decay and dephasing rates much less than $(1 \text{ ms})^{-1}$. In our design, we assume coupler coherence times of $\gamma_C^{-1} = 10 \mu\text{s}$ and $\gamma_{C\phi}^{-1} = 20 \mu\text{s}$ based on previous measurements of similar devices [19]. With these assumptions and the detunings

in our system, choosing $g_{a,b}/2\pi \sim 100$ MHz results in inherited decay and dephasing rates around $(4 \text{ ms})^{-1}$.

As shown in Fig. 9(a), the measured T_1 time of the coupler exceeds this estimate. As a consequence, we see no flux dependence in the oscillator decay times. The dephasing time of the coupler is comparable to our estimate, resulting in a measurable variation in oscillator dephasing time with changing flux bias. Nevertheless, decay remains the dominant cavity error. In future designs, enhancing the coupling rates $g_{a,b}$ will allow a factor of several improvement in the beam-splitter rate without degrading oscillator coherence. The cost for this change will be a commensurate increase in χ_{ab} but as this rate can be heavily suppressed by tuning the external flux, this trade-off seems favorable.

3. Choice of the SNAIL coupler parameters

As discussed in the main text, the coupler parameters are chosen to suppress quartic interactions in its Hamiltonian and enhance cubic interactions. Realizing this goal requires relating those interactions to the circuit parameters of the coupler.

Following Refs. [34,39], we begin with general comments on the potential U_s of the SNAIL dipole, parametrized as pictured in the inset of Fig. 1(d). When the number of large junctions in the array $M > 1$ and the external flux is a dc bias [63], the potential may be written as

$$\frac{U_s(\hat{\varphi}_s)}{E_J} = -\beta \cos(\hat{\varphi}_s - \phi_e) - M \cos \frac{\hat{\varphi}_s}{M}. \quad (\text{B4})$$

Here, $\phi_e \equiv \Phi_e/\Phi_0$ is the phase difference imparted by the external flux and $\hat{\varphi}_s$ is the phase drop across the SNAIL; $\beta < 1$ is the ratio of Josephson energies for the primary and shunting Josephson junctions.

The position of the potential minimum φ_m at a given external flux is found by solving

$$\frac{dU_s/d\hat{\varphi}_s}{E_J} = \beta \sin(\hat{\varphi}_s - \phi_e) + \sin\left(\frac{\hat{\varphi}_s}{M}\right) = 0. \quad (\text{B5})$$

When $\beta < 1/M$, a single minimum exists and so to avoid multistability we work in this regime. The position of the potential minimum reaches its largest value $\varphi_m^{\max} = M \arcsin(\beta^{-1})$ when $\phi_e = \varphi_m^{\max} + \pi/2$.

Near the minimum φ_m the potential U_s is well approximated by fifth-order expansion in $\hat{\varphi}_s - \varphi_m$:

$$U_s(\hat{\varphi}_s) \approx \frac{c_2}{2!}(\hat{\varphi}_s - \varphi_m)^2 + \frac{c_3}{3!}(\hat{\varphi}_s - \varphi_m)^3 + \frac{c_4}{4!}(\hat{\varphi}_s - \varphi_m)^4 + \frac{c_5}{5!}(\hat{\varphi}_s - \varphi_m)^5, \quad (\text{B6})$$

with

$$\begin{aligned} c_2 &= \beta \cos(\varphi_m - \phi_e) + M^{-1} \cos \frac{\varphi_m}{M}, \\ c_3 &= \frac{M^2 - 1}{M^2} \sin \frac{\varphi_m}{M}, \\ c_4 &= -\beta \cos(\varphi_m - \phi_e) - M^{-3} \cos \frac{\varphi_m}{M}, \\ c_5 &= \frac{1 - M^4}{M^4} \sin \frac{\varphi_m}{M}, \end{aligned} \quad (\text{B7})$$

where Eq. (B5) is used to simplify the expression for c_3 and c_5 . The maximum magnitudes of c_3 and c_5 are obtained at φ_m^{\max} , at which point the ratio of their magnitudes is $|c_3/c_5| = M^2/(M^2 + 1)$. This suggests choosing $M > 2$ to attain a more favorable ratio of the desired c_3 to the unwanted c_5 , although the choice of M will also be guided by fabrication considerations. For our coupler, we use $M = 3$.

As discussed in Ref. [39], calculating the Hamiltonian of the full circuit in Fig. 1(d), including a linear inductance, rescales the potential coefficients $c_j \rightarrow \tilde{c}_j$ but does not change the location of the potential minima. The result is of the form given in Eq. (F2). Second quantization with $\hat{\varphi} - \varphi_m = \phi_c(\hat{c} + \hat{c}^\dagger)$, $\hat{N} = -i\phi_c^{-1}(\hat{c} - \hat{c}^\dagger)/2$ yields Eq. (2), with $\hat{\varphi}$ replaced by $\hat{\varphi}$. The coefficients in \mathcal{H}_c are

$$\begin{aligned} \omega_c &= \sqrt{8\tilde{c}_2 E_C E_J}, \\ \phi_c &= \left(\frac{2E_C}{\tilde{c}_2 E_J}\right)^{\frac{1}{4}}, \\ \hbar g_3 &= E_J \phi_c^3 \frac{\tilde{c}_3}{3!}, \\ \hbar g_4 &= E_J \phi_c^4 \frac{\tilde{c}_4}{4!}, \\ \hbar g_5 &= E_J \phi_c^5 \frac{\tilde{c}_5}{5!}, \end{aligned} \quad (\text{B8})$$

where

$$\begin{aligned} \tilde{c}_2 &= p c_2, \\ \tilde{c}_3 &= p^3 c_3, \\ \tilde{c}_4 &= p^4 \left(c_4 - 3 \frac{c_3^2}{c_2} (1-p) \right), \\ \tilde{c}_5 &= p^5 \left(c_5 - 10 \frac{c_4 c_3}{c_2} (1-p) + 15 \frac{c_3^2}{c_2^2} (1-p)^2 \right), \end{aligned} \quad (\text{B9})$$

and $p \equiv c_2 E_J / E_L + c_2 E_J$. A rotating wave approximation on \mathcal{H}_c yields a Kerr nonlinearity of the form $-\alpha_c / 2 \hat{c}^\dagger \hat{c}^2$ [see Eq. (3)]. We are now in a position to optimize the circuit parameters of the coupler for maximal g_3/g_5 and

minimal quartic interaction such as the anharmonicity and cross-Kerr χ_{ab} . As the frequency ω_c is constrained by other considerations, choosing the ϕ_c (equivalent to choosing the impedance of the coupler) fixes the linear circuit parameters E_L and E_C . Reducing ϕ_c is a straightforward way to increase the ratio g_3 relative to higher-order terms. Care should be taken, though, that the pump amplitude can be commensurately increased, to avoid reduction of the beam-splitter rate. In practice, this means ensuring that the drive can be delivered to the SNAIL coupler while sufficiently isolating other nonlinear components (such as the transmons). For more details, see Appendix J.

Replacing a single SNAIL with an array of N SNAILS accomplishes much the same goal (and at the same cost), provided that E_J is adjusted such that the total inductance of the array is equal to the inductance of the original SNAIL. We note that this is only possible when $E_J \ll E_L$ [64] but this requirement is well satisfied for our device.

We test multiple devices with $N = 3$ SNAILS in an array and find them to have excellent suppression of quartic interactions. Their performance, however, is limited by an undesired coupling of the strong beam-splitter pump to the transmons used for oscillator control. Reverting to a single SNAIL remedies this issue, as the same beam-splitter rate can be attained with a ninefold reduction in pump power. Future designs, though, could recover the enhanced suppression of higher-order terms by engineering better isolation between the coupler and the transmons, as in Ref. [35].

The only remaining parameter in the coupler circuit is the ratio of the Josephson energies β . Broadly speaking, larger values of β in the range $\beta < 1/M$ increase the nonlinearities g_3 and g_4 and cause the potential coefficients \tilde{c}_2 to be more sensitive to external flux. Assuming that undesirable quartic interactions have already been suppressed through other means, the trade-off in choosing β amounts to balancing the desire to increase g_3 against greater susceptibility to flux noise (via the increase of $d\omega_C/d\Phi_e$).

To quantify this, we perform numerical simulations sweeping β with E_J adjusted to hold ω_C fixed. They show that g_3 and $d\omega_C/d\Phi_e$ increase fractionally at the same rate when β is less than a critical value $\beta_c \approx 0.15$ (for our choices of the other parameters). Beyond this, i.e., when $\beta_c < \beta < 1/M$, the susceptibility to external flux increases more rapidly with β than g_3 .

Assuming that the flux noise on the coupler is pink, the decay and dephasing rates inherited by the cavities both scale as $(g_{a,b}/\Delta_{a,b})^2$. The optimal choice of β then depends on the expected coupler coherence and can be found with the following procedure. Beginning with $\beta = \beta_c$, calculate the dressed coherence of the cavities. If the cavities are dephasing limited, increase $(g_{a,b}/\Delta_{a,b})$ as much as possible until the inherited decay rate becomes non-negligible. Alternatively, if the cavities are decay limited, increase β

until the decay and dephasing times are balanced. Then adjust $(g_{a,b}/\Delta_{a,b})$ as needed.

Note that the above procedure requires knowledge of the bare-oscillator and coupler coherences, which are not always available in design. Nonetheless it illustrates the general considerations in the choice.

APPENDIX C: DETERMINING THE LINEAR OSCILLATOR-SNAIL COUPLING

An important parameter in predicting the nonlinear properties of the oscillators is the linear coupling strength g between each oscillator and the SNAIL coupler. This appears when considering their bare modes,

$$\hat{\mathcal{H}}_{AC}/\hbar = \omega_A \hat{A}^\dagger \hat{A} + \omega_C \hat{C}^\dagger \hat{C} + g_a (\hat{A}^\dagger \hat{C} + \hat{C}^\dagger \hat{A}), \quad (\text{C1})$$

where ω_A and ω_C are the bare frequencies of Alice and the coupler and g is taken to be real and independent of the external flux. The same expression holds for Bob. The Hamiltonian can be rewritten in matrix form,

$$\begin{aligned} \hat{\mathcal{H}}_{AC}/\hbar &= (\hat{A}^\dagger \quad \hat{C}^\dagger) \begin{pmatrix} \omega_A & g_a \\ g_a & \omega_C \end{pmatrix} \begin{pmatrix} \hat{A} \\ \hat{C} \end{pmatrix}, \\ &= (\hat{A}^\dagger \quad \hat{C}^\dagger) \mathbf{M} \begin{pmatrix} \hat{A} \\ \hat{C} \end{pmatrix}, \end{aligned} \quad (\text{C2})$$

and \mathbf{M} diagonalized to find the dressed frequencies

$$\begin{aligned} \omega_a &\approx \omega_A - \frac{g_a^2}{\omega_C - \omega_A}, \\ \omega_c &\approx \omega_C + \frac{g_a^2}{\omega_C - \omega_A}, \end{aligned} \quad (\text{C3})$$

where we use the fact that $g_a \ll \omega_C - \omega_A$. The same approximation relates the dressed frequencies to one another,

$$\omega_a \approx \omega_A - \frac{g_a^2}{\omega_c - \omega_A}. \quad (\text{C4})$$

By observing how the oscillator frequencies depend on the SNAIL frequency as we tune external flux, we can fit to obtain the linear coupling g and the bare-oscillator frequency, as shown in Fig. 5.

APPENDIX D: CALCULATION OF OSCILLATOR SELF- AND CROSS-KERRS

The analysis in Appendix C can be extended to account for two cavities (with dressed annihilation operators \hat{a} and

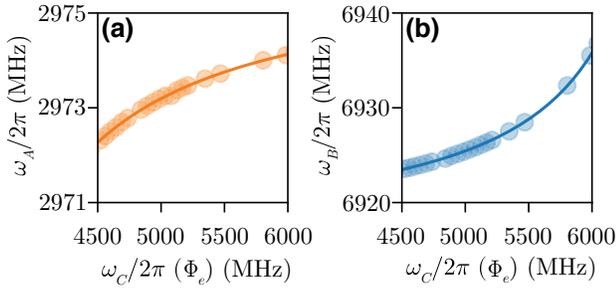


FIG. 5. The cavity frequencies. The circles show measurements of the dressed oscillator frequencies for (a) Alice and (b) Bob at a range of external flux biases, plotted as a function of the measured SNAIL frequency. The lines show fits to Eq. (C4), from which we extract the linear oscillator-SNAIL couplings of $g_a/2\pi = 75.6 \pm 0.2$ MHz and $g_b/2\pi = 134.9 \pm 0.1$ MHz and the bare-mode frequencies of $\omega_A/2\pi = 2976.018(16)$ MHz and $\omega_B/2\pi = 6915.945(17)$ MHz.

\hat{b}) coupled to the same SNAIL mode. In the dispersive limit, we approximate the bare SNAIL mode as

$$\hat{C} \approx \hat{c} + \frac{g_a}{\Delta_a} \hat{a} + \frac{g_b}{\Delta_b} \hat{b}, \quad (\text{D1})$$

where $\Delta_{a(b)} \equiv \omega_{a(b)} - \omega_c$. Insertion of Eq. (D1) into the Hamiltonian for the uncoupled SNAIL mode [Eq. (2)] allows us to calculate the energy of the combined system perturbatively.

The cross-Kerr between the cavities χ_{ab} can be defined as the frequency shift $\Delta E/\hbar$ proportional to $n_a n_b$. The g_4 nonlinearity yields a shift at first order in perturbation theory, while g_3 yields a shift at second order and so the cross-Kerr can be approximated as

$$\chi_{ab} \approx \left(24g_4 + 36\frac{g_3^2}{\tilde{\omega}} \right) \left(\frac{g_a}{\Delta_a} \right)^2 \left(\frac{g_b}{\Delta_b} \right)^2, \quad (\text{D2})$$

where

$$\tilde{\omega} \equiv \left(\frac{1}{\omega_a - \omega_b - \omega_c} + \frac{1}{-\omega_a + \omega_b - \omega_c} + \frac{1}{\omega_a + \omega_b - \omega_c} + \frac{1}{-\omega_a - \omega_b - \omega_c} \right)^{-1}.$$

Here, we assume that $g_a \ll \Delta_a$ and $g_b \ll \Delta_b$ we do not make any assumptions about the relative sizes of the mode frequencies.

The same procedure can be used to estimate the cavity self-Kerr, which is proportional to n^2 . To first order in g_4 and second order in g_3 ,

$$K_a \approx \left(12g_4 - 18g_3^2 \left(\frac{2\omega_c}{4\omega_a^2 - \omega_c^2} + \frac{4}{\omega_c} \right) \right) \left(\frac{g_a}{\Delta_a} \right)^4 + \frac{\chi_{at}^2}{4\alpha_{at}}, \quad (\text{D3})$$

where α_{at} is the anharmonicity of the transmon ancilla. The first term accounts for the coupler contribution to the cavity self-Kerr, while the second accounts for the ancilla contribution, which gives a fixed offset that is independent of external flux bias.

APPENDIX E: MEASUREMENT OF SELF- AND CROSS-KERR BETWEEN MODES

To sensitively measure the oscillator-oscillator cross-Kerr interaction $\chi_{ab} \hat{a}^\dagger \hat{a} \hat{b}^\dagger \hat{b}$, we find the linear frequency shift of one oscillator while populating the other with a large and variable number of photons. The cross-Kerr interaction is symmetric under exchange of \hat{a} and \hat{b} , so one can either displace Alice and detect a change in Bob's resonant frequency or vice versa. Moreover, the frequency shift can be detected either interferometrically or spectroscopically, giving four possible measurement configurations for χ_{ab} .

Figure 6(a) illustrates the pulse sequence used to interferometrically measure Alice's frequency. The protocol starts by displacing Bob by a variable amount α_d . We then apply a unit displacement to Alice with a detuned drive at frequency $\omega_a + \delta_i$ such that the resulting coherent state acquires phase at a rate $\delta_i/2\pi = 200$ kHz. In addition, the coherent state will pick up an extra phase due to the cross-Kerr at an average rate of $|\alpha_d|^2 \chi_{ab}$. After a variable delay t , a second detuned unit displacement returns the coherent state to the origin if and only if it has acquired π radians of phase during the delay. Figure 6(c) shows data measured in this way near the cross-Kerr free point at $\Phi_e/\Phi_0 = 0.43$. By detecting the probability P_0 of finding Alice in vacuum, we can fit the measured revivals with the model

$$P_0 = A_0 e^{1 + \cos((\delta_i + |\alpha_d|^2 \chi_{ab})t)} + C, \quad (\text{E1})$$

to infer χ_{ab} . Here, A_0 and C capture measurement infidelity and the detuning δ_i and displacement α_d are fixed by the protocol. The results of these fits are shown as circles in Fig. 6(e). The model neglects the effects of cavity decay but since measurements of Alice and Bob yield similar values for χ_{ab} despite their different decay rates, we deduce that corrections due to cavity decay are small.

Figure 6(b) illustrates the pulse sequences used to spectroscopically measure Alice's frequency. After displacing Bob by α_d , a long pulse with variable detuning δ_s is used to displace Alice, followed by a measurement to determine whether Alice is in vacuum. Figure 6(d) shows data measured in this way. Fitting a linear relation to the spectroscopy dip at each value of n , the slope gives a measurement of χ_{ab} .

Cavity self-Kerr ($K_a/2\hat{a}^\dagger \hat{a}^\dagger \hat{a} \hat{a}$) is measured very similarly to the interferometric measurement of cross-Kerr outlined above. Instead of displacing the other cavity, we vary the amplitude of the detuned displacement α , such

that the cavity acquires extra phase at a rate $K_a/2|\alpha|^2$. We can then extract the self-Kerr by measuring whether the cavity is in vacuum and fitting it to

$$P_0 = A_0 e^{1+\cos((\delta+K_a/2|\alpha|^2)t)} + C. \quad (\text{E2})$$

The cross-Kerr between the cavities and the SNAIL coupler ($\chi_{ac}\hat{a}^\dagger\hat{a}(|e\rangle\langle e|)_c$) is measured similarly to the spectroscopic measurement of the cavity-cavity cross-Kerr. The cavity is displaced by a variable displacement α_d before a long spectroscopy tone is applied to the coupler. We can then measure whether the coupler is in its ground state by performing a coupler-state-selective π pulse on the transmon not coupled to the cavity and reading out the transmon

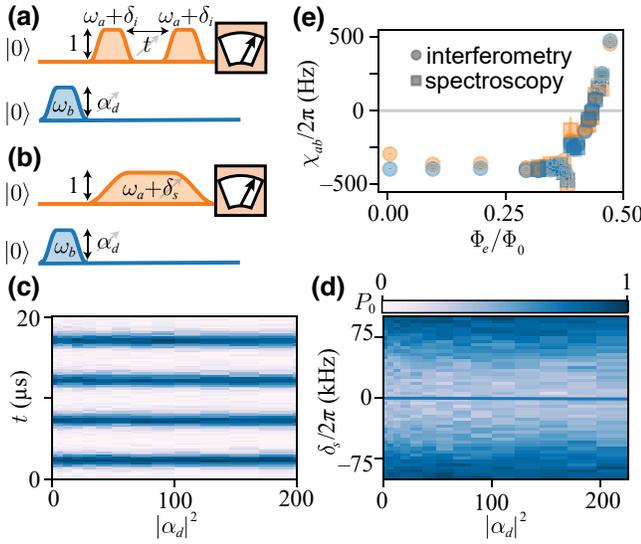


FIG. 6. Measurement of the oscillator-oscillator cross-Kerr. The pulse sequence for measuring the oscillator-oscillator cross-Kerr (a) interferometrically and (b) spectroscopically with Alice. Exchanging the cavity rails yields the complementary sequences used to measure with Bob. Both the interferometric and spectroscopic sequences start with a resonant and variable displacement of Bob, which populates that cavity with an average of $n = |\alpha_d|^2$ photons. For the interferometric measurement in (a), the resulting frequency shift is detected by off-resonantly displacing Alice by $\alpha = 1$, waiting a variable time t , and then making a second displacement of the same size before measuring whether Alice is in vacuum. For the spectroscopic measurement in (b), the frequency shift is detected by sweeping the frequency of a long displacement pulse on Alice and measuring whether it is in vacuum. (c),(d) The result of interferometric (c) and spectroscopic (d) measurements on Bob near the cross-Kerr free point $\Phi_e/\Phi_0 = 0.43$. The solid line in (d) is a linear fit to the centers of Gaussian fits on line cuts at each value of α_d . (e) The inferred cross-Kerr χ_{ab} as a function of the external flux, measured interferometrically (circles) and spectroscopically (squares), from Alice (orange) and Bob (blue). The measurements shown in Fig. 1(h) are the average of the two interferometric data sets.

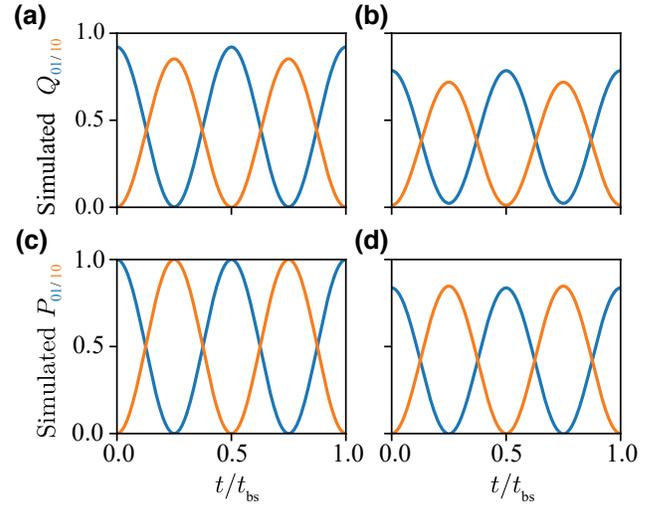


FIG. 7. Correcting for measurement infidelity. The top row shows mock data of the uncorrected probability $Q_{01/10}$ of measuring the oscillators in $|0, 1\rangle$ or $|1, 0\rangle$ during beam-splitter oscillations, assuming a fixed measurement infidelity, with either (a) perfect or (b) imperfect preparation of the initial state $|0, 1\rangle$. The bottom row shows the inferred probability $P_{01/10}$ that the oscillators are in $|0, 1\rangle$ or $|1, 0\rangle$ given the uncorrected data, with (c) perfect and (d) imperfect state preparation, after applying the correction in Eq. (G1). In practice, applying that correction requires populating the matrices $E_{(A)}$ and $E_{(B)}$, which can be done with measurements of Rabi oscillations. To generate the mock data, we take the conditional measurement probabilities to be $P_{(A)}(e|0) = 0.93$, $P_{(B)}(e|0) = 0.87$, $P_{(A)}(g|1) = 0.98$ and $P_{(B)}(g|1) = 0.99$. In the right column, we take the state-preparation fidelities to be $P_{(A)}(0|0) = 0.99$, $P_{(B)}(0|0) = 0.98$, $P_{(A)}(1|1) = 0.87$ and $P_{(B)}(1|1) = 0.84$, where $P_{(A/B)}(i|j)$ is the probability of actually preparing Fock state $|i\rangle$ when we try to prepare Fock state $|j\rangle$. We emphasize that these values are not taken from the experiment but are chosen to highlight the effect of the Rabi correction.

state. The spectroscopy data show a linear frequency shift proportional to $|\alpha_d|^2$, with the slope given by χ_{ac} .

APPENDIX F: DERIVATION OF THE BEAM-SPLITTER HAMILTONIAN FROM A DRIVEN SNAIL

We consider a system of two oscillators with associated annihilation operators \hat{a} and \hat{b} , which are jointly coupled to a conversion element with annihilation operator \hat{c} .

The system Hamiltonian is

$$\frac{\hat{\mathcal{H}}_s}{\hbar} = \omega_a \hat{a}^\dagger \hat{a} + \omega_b \hat{b}^\dagger \hat{b} + \omega_c \hat{c}^\dagger \hat{c} + \frac{\hat{\mathcal{H}}_d}{\hbar} + \frac{\hat{\mathcal{H}}_{\text{cnl}}}{\hbar}. \quad (\text{F1})$$

Here, ω_c is the g - e transition of the coupler, $\hat{\mathcal{H}}_d$ is the Hamiltonian of the drive, and $\hat{\mathcal{H}}_{\text{cnl}}$ is the nonquadratic part of the coupler Hamiltonian $\hat{\mathcal{H}}_c$.

The Hamiltonian of the coupler circuit is [39] (see also Appendix B):

$$\hat{\mathcal{H}}_c = 4E_C \hat{N}^2 + E_J \left(\frac{\tilde{c}_2}{2!} \hat{\phi}^2 + \frac{\tilde{c}_3}{3!} \hat{\phi}^3 + \frac{\tilde{c}_4}{4!} \hat{\phi}^4 + \frac{\tilde{c}_5}{5!} \hat{\phi}^5 \right) + \mathcal{O}(\hat{\phi})^6 \quad (\text{F2})$$

or, equivalently,

$$\omega_c \hat{c}^\dagger \hat{c} + \hat{\mathcal{H}}_{\text{cnl}} = E_J \left(\frac{\tilde{c}_3}{3!} \hat{\phi}^3 + \frac{\tilde{c}_4}{4!} \hat{\phi}^4 + \frac{\tilde{c}_5}{5!} \hat{\phi}^5 \right) + \mathcal{O}(\hat{\phi})^6,$$

where $\hat{\mathcal{H}}_{\text{cnl}}$ represents the terms in the Hamiltonian with order higher than two. Second quantization on Eq. (F2) yields Eq. (2). As the coupler is capacitively connected to the two oscillators [see Fig. 1(b)], the phase operator $\hat{\phi}$ depends on the displacement of all three hybridized modes, $\hat{\phi} \equiv \phi_a(\hat{a} + \hat{a}^\dagger) + \phi_b(\hat{b} + \hat{b}^\dagger) + \phi_c(\hat{c} + \hat{c}^\dagger)$. Here, the ϕ_i are the zero-point phases of the i th mode, with $i = a, b, c$.

To actuate the mixing process, we apply a pump with frequency ω_p and strengths ϵ . The Hamiltonian of the drive is

$$\hat{\mathcal{H}}_d / \hbar = -(\epsilon e^{-i\omega_p t} - \epsilon^* e^{i\omega_p t})(\hat{c} - \hat{c}^\dagger). \quad (\text{F3})$$

We now move to the frame displaced by the drive [65]. The unitary for the displacement transformation is

$$\hat{U}_d = e^{\tilde{\xi}(t)\hat{c}^\dagger - \tilde{\xi}^*(t)\hat{c}}, \quad (\text{F4})$$

where

$$\tilde{\xi} \equiv \xi e^{-i\omega_p t} \equiv \frac{\epsilon}{\omega_p - \omega_c} e^{-i\omega_p t}. \quad (\text{F5})$$

In the above, we neglect dissipation on the coupler mode and counter-rotating terms.

We then move into a frame corotating with the three modes via the unitary transformation

$$\hat{U}_r = e^{-i(\omega_a \hat{a}^\dagger \hat{a} + \omega_b \hat{b}^\dagger \hat{b} + \omega_c \hat{c}^\dagger \hat{c})t}. \quad (\text{F6})$$

The transformed phase operator is now

$$\begin{aligned} \hat{\phi} &= \hat{U}_r^\dagger \hat{U}_d^\dagger \hat{\phi} \hat{U}_d \hat{U}_r \\ &= \phi_a(\hat{a} e^{-i\omega_a t} + \hat{a}^\dagger e^{i\omega_a t}) + \phi_b(\hat{b} e^{-i\omega_b t} + \hat{b}^\dagger e^{i\omega_b t}) \\ &\quad + \phi_c(\hat{c} e^{-i\omega_c t} + \hat{c}^\dagger e^{i\omega_c t} + \xi e^{-i\omega_p t} + \xi^* e^{i\omega_p t}). \end{aligned} \quad (\text{F7})$$

When the frequency of the pump is Δ detuned from the difference $\omega_b - \omega_a$, a beam-splitter interaction of the form

$$g_{\text{bs}} e^{i\theta} \hat{a}^\dagger \hat{b} e^{i\Delta t} + \text{h.c.} \quad (\text{F8})$$

emerges, with (to lowest order)

$$\begin{aligned} g_{\text{bs}} &\approx E_J \phi_a \phi_b \phi_c \xi \tilde{c}_3, \\ &\approx 6 \frac{g_a}{\Delta_a} \frac{g_b}{\Delta_b} \xi g_3. \end{aligned} \quad (\text{F9})$$

Greater accuracy can be obtained by including the effect of higher-order nonlinear terms in the Hamiltonian (beyond g_3). To leading order within the RWA (but still treating the drive perturbatively) these modify the beam-splitter rate in Eq. (F9) to

$$g_{\text{bs}} \approx \frac{g_a}{\Delta_a} \frac{g_b}{\Delta_b} \xi^* \left(\sum_{m=1} (2m+1)! \frac{g_{2m+1}}{m!(m-1)!} |\xi|^{2m-2} \right). \quad (\text{F10})$$

Terms that are of higher order in g_i/Δ_i and those proportional to $g_{2m+1} \xi^* |\xi|^j$ with $j < 2m - 2$ are suppressed, since $(g_i/\Delta_i)^2 \ll 1$ and $(p\phi_c)^2 \ll 1$, respectively.

This expression predicts that at sufficient pump amplitude, the beam-splitter rate will deviate from a purely linear increase with amplitude. What physical considerations limit the pump amplitude itself?

The renormalized SNAIL potential (described in Appendix B) is 6π periodic in $\hat{\phi}$ with a single minimum within each period, provided that $\beta < 1/M$. Furthermore, adjacent maxima and minima are exactly 3π apart. If we approximate the maximum deviation from the potential minima due to the drive as $|\phi_{\text{max}}| = 2\phi_c |\xi|$, then the drive amplitude required to enter the adjacent well is approximately

$$|\xi_{\text{crit}}| \approx \frac{3\pi}{2\phi_c}. \quad (\text{F11})$$

For our system, this evaluates to $|\xi_{\text{crit}}| = 12.9$, a value that agrees closely with other estimates for the critical photon number given in Ref. [43]. This calculation suggests that further optimization of the coupler may allow even faster beam-splitter rates than those demonstrated in the present work, for which $\xi < 3$. Such work will likely require more exact treatments of the coupler Hamiltonian.

APPENDIX G: MEASUREMENT OF THE JOINT PROBABILITY DISTRIBUTION OF THE PHOTON POPULATION

The joint probability distribution of the photon population in the two cavities [shown in Fig. 2(b)] is obtained by measuring the state of both cavities simultaneously and correlating the outcomes shot by shot.

These measurements involve performing a 0-selective π pulse on the ancilla of each oscillator (which succeeds only if the oscillator is in vacuum) and reading out the ancilla state. Ideally, if an oscillator is in $|0\rangle$, we measure

its ancilla in $|e\rangle$ with unit probability. Likewise, if the oscillator is in a state $|\psi_0\rangle$ where $\langle 0|\psi_0\rangle = 0$, we measure the ancilla in $|g\rangle$. We multiply the results of the simultaneous measurements shot by shot to infer the joint oscillator state (e.g., the ancillas measured in $|e\rangle_A|g\rangle_B$ register as joint oscillator state $|0\rangle_A|\psi_0\rangle_B$). If we neglect oscillator heating, our cavities are confined to the $\{0, 1\}$ subspace and we can replace $|\psi_0\rangle$ with $|1\rangle$. We do so for ease of reading but the analysis works equally well with $|\psi_0\rangle$ undetermined.

Averaging the results yields the uncorrected probabilities $Q(i, j)$, where $i, j = \{e, g\}$. The uncorrected probabilities are the result of imperfect measurement arising from errors during our selective π pulses and readout. These imperfections lead to false positives, where the ancilla is measured in $|e\rangle$ despite the oscillator being in $|1\rangle$, and false negatives, where the ancilla is measured in $|g\rangle$ despite the oscillator being in $|0\rangle$.

For a single-oscillator measurement, we construct a matrix \mathbf{E} relating the probability of measuring the ancilla state to the true oscillator state:

$$\begin{pmatrix} Q(e) \\ Q(g) \end{pmatrix} = \begin{pmatrix} P(e|0) & P(e|1) \\ P(g|0) & P(g|1) \end{pmatrix} \begin{pmatrix} P(0) \\ P(1) \end{pmatrix} = \mathbf{E} \begin{pmatrix} P(0) \\ P(1) \end{pmatrix}.$$

Here, the matrix \mathbf{E} can be written in terms of the false positive and true positive rates, f and t . We can estimate these by taking the minimum and maximum points of our 0-selective Rabi oscillations, respectively:

$$\mathbf{E} = \begin{pmatrix} t & f \\ 1-t & 1-f \end{pmatrix}.$$

We can invert this matrix to obtain the oscillator state probabilities from our measurements:

$$\begin{pmatrix} P(0) \\ P(1) \end{pmatrix} = \mathbf{E}^{-1} \begin{pmatrix} Q(e) \\ Q(g) \end{pmatrix}.$$

The same procedure may be generalized for joint measurements, where we assume that measurement errors are uncorrelated between ancillas:

$$\begin{aligned} Q(i, j) &= E(i, j | n, m) P(n, m), \\ &= E_{(A)}(i | n) E_{(B)}(j | m) P(n, m), \\ \mathbf{Q} &= \mathbf{E}_{(A)} \mathbf{P} \mathbf{E}_{(B)}^T. \end{aligned}$$

Inverting this, we have

$$\mathbf{P} = \mathbf{E}_{(A)}^{-1} \mathbf{Q} (\mathbf{E}_{(B)}^T)^{-1}. \quad (\text{G1})$$

This method corrects the distortion in the measurement of the joint state of two oscillators when the measurement fidelity on each oscillator is different. We note that it does not normalize for imperfections in state preparation.

Figure 7 shows how the correction protocol transforms mock uncorrected data, with and without perfect state preparation. In the case of imperfect state preparation, the oscillations of P_{01} and P_{10} should have the same amplitude but not extend from 0 to 1.

To initialize the cavities in $|1\rangle_A|0\rangle_B$, we use a feedback routine whereby we displace Alice by $\alpha = 1$ and then simultaneously perform a 1-selective π pulse on Alice's ancilla and a 0-selective π pulse on Bob's ancilla. We then perform unselective π pulses on both ancillas and read out their state. If the cavities are in $|1\rangle_A|0\rangle_B$, we should find the ancillas in $|g\rangle_A|g\rangle_B$. If successful, we repeat the π pulses and read out once more, and continue if the result is once again $|g\rangle_A|g\rangle_B$. If either of these measurements fails, we wait long enough to let Alice relax to vacuum and restart from the beginning.

APPENDIX H: FITTING PROCEDURE FOR COHERENCE TIMES

We characterize the beam-splitter interaction by fitting the model in Eq. (6) to measurements like those shown in Figs. 2(a) and 2(b). The procedure has several steps. To begin, the duration and frequency of the pump is swept to produce a chevron pattern [as can be seen from Eq. (8)], from which we determine the precise pump frequency for a resonant beam-splitter interaction. Then, the pump frequency is fixed at resonance and measurements of $P_{01/10}$ are collected for a variety of unevenly spaced pump-duration times t . Readout imperfections in the data are corrected as described in Appendix G.

As illustrated in Fig. 2(b), we collect an oscillation of both curves by finely sampling 21 time points in a time window of duration approximately $6t_{\text{bs}}$. This process is repeated with the fine sampling centered around 41 coarsely spaced times between 0 and about $3\tau_{\text{bs}}$ (estimated from the unpumped cavity lifetimes). To extract g_{bs} and the upper and lower envelopes of the probability curves, we fit each of the finely sampled oscillations to a sinusoid, $A \sin(2g_{\text{bs}}t + \phi) + B$. Together, the upper and lower envelopes contain complete information about the decay and dephasing: their mean, B , decays with the time scale τ and their amplitude, A , decays at the time scale $(1/\tau + 1/\tau_\phi)^{-1}$. We fit the mean and amplitude of both P_{01} and P_{10} to extract τ and τ_ϕ .

The above procedure has several benefits relative to fitting Eq. (6) directly or fitting $P_\Sigma \equiv P_{01} + P_{10}$ and $P_{01/10}/P_\Sigma$ [29]. First, unlike fitting Eq. (6) directly, the decay and dephasing time can be extracted from separate traces, thereby reducing the number of decoherence time constants extracted from each fit from two to one. Second, unlike fitting to P_Σ and $P_{01/10}/P_\Sigma$, the two measured traces need not be combined. This allows decay and dephasing times to be inferred separately from P_{01} and P_{10} and compared as a consistency check.

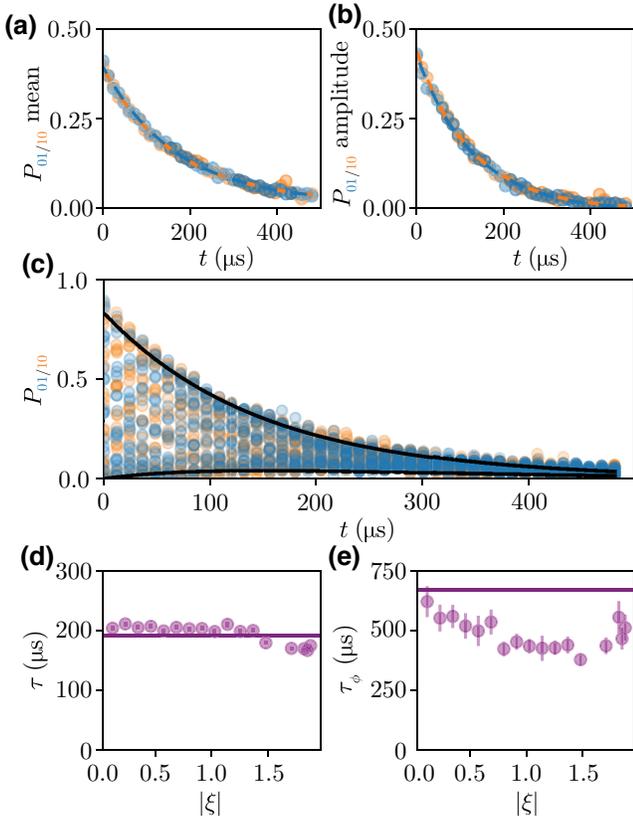


FIG. 8. Fitting pumped decay and dephasing times. (a) The decay of the mean of P_{01} (blue circles) and P_{10} (orange circles) is used to extract the decay time τ . (b) The decay of the amplitude of P_{01} (blue circles) and P_{10} (orange circles) is used to extract $(1/\tau + 1/\tau_\phi)^{-1}$, from which we obtain the dephasing time τ_ϕ . The dashed lines show the exponential decay fits. The decay constants for the mean of $P_{01/10}$ are $170 \pm 5 \mu\text{s}$ and $167 \pm 7 \mu\text{s}$, respectively, while the decay constants for the amplitude are $132 \pm 4 \mu\text{s}$ and $127 \pm 4 \mu\text{s}$. This yields an average fitted $\tau = 168 \pm 4 \mu\text{s}$ and $\tau_\phi = 560 \pm 70 \mu\text{s}$. (c) The complete trace of the measured $P_{01/10}$ curves as a function of the variable delay t , with envelopes (black lines) obtained from the fitted decay and dephasing times. (d),(e) Fitted decay times τ (d) and dephasing times τ_ϕ (e) as a function of the pump amplitude ξ . The solid lines shows the average unpumped decay and dephasing time of the cavities. The data are measured at the same flux point and pump amplitude as in Fig. 2(b).

To illustrate the intermediate steps in the fitting process, Fig. 8(a) shows the mean of P_{01} and P_{10} and Fig. 8(b) shows their amplitude, measured at $\Phi_e/\Phi_0 = 0.32$ and with $|\xi| = 1.83$. The time scales inferred from separate fits agree within their standard error; the quoted values for τ and τ_ϕ are the average of the two. To validate the results, we plot the readout-corrected traces for $P_{01/10}$ in Fig. 8(c), along with the envelopes inferred from the fitted time constants in Figs. 8(a) and 8(b).

Figures 8(d) and 8(e) show the time constants for decay from the single-photon manifold [Fig. 8(d)] and dephasing

[Fig. 8(e)] extracted in this way, as a function of the pump amplitude. The decay times are relatively independent of the pump amplitude. The dephasing times are slightly suppressed as pump amplitude increases but remain much longer than the decay time τ .

APPENDIX I: SYSTEM PARAMETERS

Tables I and II list the coherence times and thermal populations for the experimental system. The coherence rates of the coupler, which depend on the external flux bias, are plotted in Fig. 9(a). The decay times T_1 (red circles) are relatively independent of flux, but generally vary between 50 and 100 μs . The dephasing rates, however, exhibit a clear suppression near the flux sweet spot of $\Phi_0/2$, as expected from the flux dependence of the frequency $d\omega_c/d\Phi_e$ of the coupler, shown in Fig. 9(b). Table III lists the self- and cross-Kerr interactions of the transmons, which are not plotted in the main text.

In Table I, we see that the relaxation times T_{1c} of the two oscillators differ by a factor of 5. We first note the difference in oscillator resonant frequencies (2.976 GHz and 6.915 GHz), which would lead to a factor-of-2.32 difference in the oscillator lifetime given the same quality factor. Further, we measure the quality factor of the bare cavities (before any chips are inserted) to be 43×10^6 and 20×10^6 , respectively, which contributes another factor-of-2.15

TABLE I. The experimental time scales. All quoted coherence times are measured at the same time as the CSWAP experiment in Fig. 3(e) is performed. The coherence times for the repeated CSWAP experiment in Fig. 3(f) differ from these due to drifts over time. To compare the results of the repeated CSWAP experiment with theoretical estimates, we use coherence times measured at the same time as this experiment was performed. The following list details those coherence times that change by more than one standard deviation from the values in the table: $T_{1qA} = 112 \pm 8 \mu\text{s}$, $T_{1qB} = 45.6 \pm 9 \mu\text{s}$, $T_{2qA} = 97 \pm 10 \mu\text{s}$, and $T_{\phi cA} = 1510 \pm 90 \mu\text{s}$.

		Alice	Bob
Transmon $e \rightarrow g$ relaxation time	T_{1q}	$127.2 \pm 1.9 \mu\text{s}$	$57.1 \pm 0.6 \mu\text{s}$
Transmon T_2	T_{2q}	$114.4 \pm 2.9 \mu\text{s}$	$56.8 \pm 1.5 \mu\text{s}$
Transmon dephasing time	$T_{\phi q}$	$208 \pm 10 \mu\text{s}$	$113 \pm 6 \mu\text{s}$
Oscillator relaxation time	T_{1c}	$482 \pm 16 \mu\text{s}$	$91 \pm 4 \mu\text{s}$
Oscillator dephasing time	$T_{\phi c}$	$2010 \pm 220 \mu\text{s}$	$840 \pm 200 \mu\text{s}$
Parity-mapping time	τ_p	616 ns	432 ns
Transmon readout time	τ_{RO}		2.1 μs
CSWAP sequence length	τ_{CSWAP}		1.3 μs

TABLE II. The mode thermal populations. All values measured at $\Phi_e = 0.35\Phi_0$, the same external flux as for the CSWAP data in Fig. 3.

	Alice	Bob
Transmon P_e	$0.70 \pm 0.14\%$	$1.02 \pm 0.20\%$
Cavity n_{thermal}	$0.96 \pm 0.19\%$	$0.11 \pm 0.02\%$
SNAIL P_e	$2.42 \pm 0.47\%$	

difference in the oscillator relaxation time. Together, the differences in the oscillator resonant frequencies and bare quality factors are consistent with the factor-of-5 difference in the oscillator lifetime. However, these differences do not fully explain why the lifetime of either oscillator is lower than its bare lifetime.

This is not yet fully understood but there are a number of possible mechanisms that could degrade the quality factors. We estimate the contributions of Purcell decay due to the transmons and SNAIL from experimentally measured values of the T_1 values of these modes and their coupling to each oscillator and find them to be negligible. Further, since we see no change in the cavity T_1 as we tune the external flux bias and thus the detuning between the SNAIL and the cavity, it is unlikely that the cavity T_1 is Purcell limited by the SNAIL.

We also perform finite-element simulations of the Purcell decay due to the Purcell filters, readout resonators, and buffer mode, as well as direct radiative loss out of the qubit drive pins. These too only limit the cavity lifetimes to time scales that are an order of magnitude larger than the observed values.

The main contributions that we find, besides the still-dominant bare loss rate of the cavities, are the radiative loss out of the cavity drive pin, the coupling rate of which

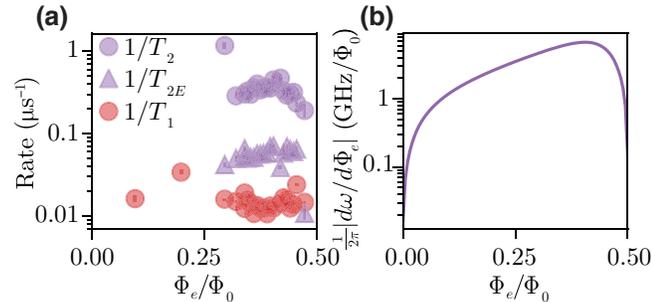


FIG. 9. The dependence of the SNAIL-coupler coherence on the external flux. (a) The SNAIL $1/T_1$ (red circles), $1/T_2$ (purple circles), and $1/T_{2E}$ (purple triangles) as a function of the external flux through the SNAIL loop. (b) The simulated gradient of the SNAIL frequency with respect to the flux, which determines its sensitivity to dephasing due to flux noise through the loop. T_{2E} is greatly increased near the flux sweet spot.

is designed to be slightly undercoupled, and the conductive losses in the magnet coil. Nevertheless, finite-element simulation suggests that these combined effects should not limit the oscillator lifetimes to less than 2.66 ms (Alice) and 0.81 ms (Bob).

Comparing T_{1c} across cool-downs in which the experimental setup is changed can provide a way to study the sources of loss in this complex system. One notable data point is a cool-down in which both transmon chips are inserted but the SNAIL chip and the magnet coil are removed. Here, we measure Bob's T_{1c} to be $333 \pm 18 \mu\text{s}$,

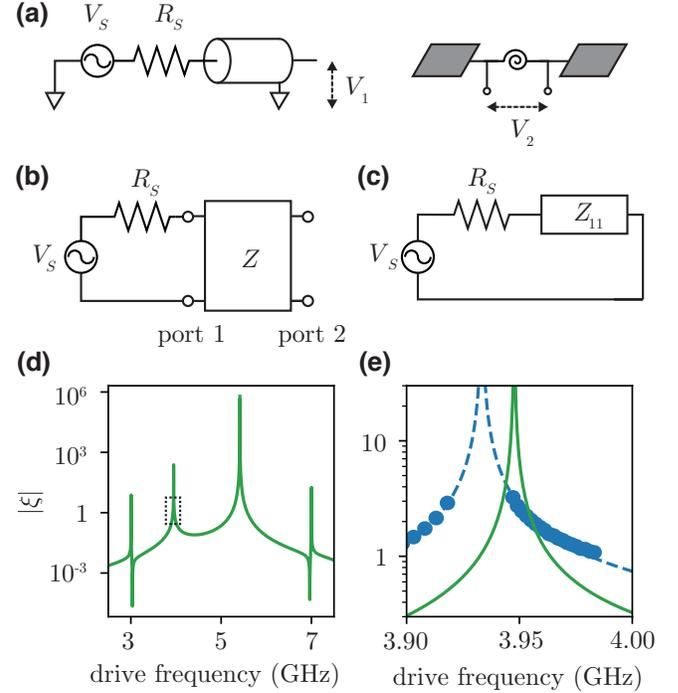


FIG. 10. The network picture of the driven system. (a) A microwave source with a voltage V_S and a resistance R_S connected via a transmission line to the drive pin of the package. It generates a voltage V_1 at the drive pin (port 1) and a voltage V_2 across the SNAIL (port 2). (b) The same circuit with the package, including the SNAIL itself, expressed as a two-port network with an impedance matrix \mathbf{Z} . (c) When port 2 is open circuited, we can treat the package as an impedance Z_{11} in series with the microwave source. (d) The simulated normalized pump amplitude $|\xi|$ with 100 pW delivered to the drive port, as a function of the drive frequency, when the external applied flux is $0.35\Phi_0$. The resonances at 3 and 7 GHz are due to Alice's and Bob's cavity modes, while the resonance at 5.2 GHz is due to the SNAIL mode. The peak near 4 GHz is due to the strip-line resonator that we introduce as a single-mode band-pass filter between the drive port and the SNAIL. (e) An enlargement of the dotted region in (d), showing the simulated prediction (green line), measured values extracted from measurements of the Stark shift (blue dots), and the simulated prediction shifted down in frequency by 14.3 MHz and up in amplitude by 2.84, to highlight the similar line shape to the data (blue dashed line).

lower than the bare cavity T_1 of $457 \pm 7 \mu\text{s}$ but higher than when the SNAIL and magnet are included ($91 \pm 4 \mu\text{s}$). However, large run-to-run variations in cavity T_1 on cool-downs with the same package and chips make it hard to precisely identify which changes are responsible. For example, on another cool-down with the SNAIL and magnet included, we measure Bob's T_{1c} to be $265 \pm 27 \mu\text{s}$.

APPENDIX J: DELIVERY OF STRONG MICROWAVE DRIVES

Delivering a strong microwave drive to the SNAIL is necessary to achieve a fast beam-splitter rate. However, it is important that the drive port does not Purcell limit the adjacent high- Q cavity modes. Furthermore, the drive power that we can deliver to the package is limited by the amount of power we can dissipate in the microwave attenuators before they start to heat the base temperature of the dilution refrigerator.

To satisfy these constraints, we filter the environment, as seen by the SNAIL and the adjacent cavity modes, by introducing a 3D superconducting strip-line resonator between the drive port and the SNAIL, with a resonance frequency ω_f close to $\omega_b - \omega_a$. This allows drive photons near 4 GHz to reach the SNAIL while shielding oscillator photons at 3 GHz and 7 GHz from the lossy drive line. (Care is taken that the resonance frequency of the strip line is sufficiently detuned from $\omega_b - \omega_a$ relative to its line width, to avoid unpumped conversion of photons in Bob.)

To confirm that the drive port does not Purcell limit the oscillators, we perform a finite-element eigenmode simulation of the package with the 3D strip-line resonator. From this, we extract the coupling Q of the oscillators and verify that they exceed 10^8 (limiting their T_1 times to no less than 5.3 ms and 2.3 ms for Alice and Bob, respectively).

The normalized pump amplitude ξ that we can deliver for a given input power can be calculated from the impedance matrix of the embedding network of the SNAIL. Figure 10(a) shows a simplified schematic of the microwave drive on the SNAIL coupler. A microwave generator sourcing a voltage V_S generates a voltage V_1 at the coupling pin inside the experimental package, which in turn generates a voltage V_2 across the SNAIL. Their ratio is determined by the properties of the embedding network of the SNAIL.

If we consider just the linear part of the circuit, we can treat the entire package including the SNAIL itself as an impedance matrix \mathbf{Z} , as shown in Fig. 10(b). This can be used to relate the voltage across the SNAIL to the current delivered from the source,

$$Z_{21} = \frac{V_2}{I_1}. \quad (\text{J1})$$

Meanwhile, the impedance presented to the source by the package is Z_{11} [Fig. 10(c)]. Current conservation indicates

that the source voltage V_S is related to the current I_1 by

$$V_S = (R_S + Z_{11})I_1. \quad (\text{J2})$$

We can eliminate I_1 from the equations to obtain

$$\frac{V_2}{V_S} = \frac{Z_{21}}{R_S + Z_{11}}. \quad (\text{J3})$$

Microwave generators typically display the rms power that they would deliver to a matched load,

$$P = \frac{|V_{\text{rms}}|^2}{2R_S} = \frac{|V_S|^2}{4R_S}, \quad (\text{J4})$$

allowing us to reach an expression for V_2 in terms of the source power:

$$|V_2| = \left| \frac{Z_{21}}{Z_{11} + R_S} \right| \sqrt{4PR_S}. \quad (\text{J5})$$

Assuming a monochromatic source at ω_p , the current passing through the SNAIL can be found by dividing V_2 by the impedance of the SNAIL at the pump frequency, $i\omega_p L_s$,

$$I(t) = \text{Re} \left(\frac{V_2}{i\omega_p L_s} e^{i\omega_p t} \right). \quad (\text{J6})$$

This couples to the flux of the coupler mode in the Hamiltonian

$$\hat{H}_d = I(t)\hat{\Phi} = \text{Re} \left(\frac{V_2}{i\omega_p L_s} e^{i\omega_p t} \right) \Phi_c^{\text{ZPF}} (\hat{c} + \hat{c}^\dagger). \quad (\text{J7})$$

By comparing this to Eq. (F3), we can write

$$|\epsilon| = \frac{\Phi_c^{\text{ZPF}} |V_2|}{2\hbar\omega_p L_s}. \quad (\text{J8})$$

From here, we can use the expression for ξ from Eq. (F5) to finally obtain

$$|\xi| = \left| \frac{Z_{21}}{Z_{11} + R_S} \right| \frac{\sqrt{PR_S} \Phi_c^{\text{ZPF}}}{\hbar\omega_p L_s (\omega_p - \omega_c)}. \quad (\text{J9})$$

The impedance matrix \mathbf{Z} can be obtained from a simulation of the package in an electromagnetic solver such as HFSS, while the zero-point flux of the SNAIL mode can be found by combining an electromagnetic simulation with further processing as, e.g., performed in the pyEPR package [66]. Figure 10(d) shows $|\xi|$ as a function of the drive frequency for an input power of 100 pW at the drive port of the package. The inset shows how the band-pass filter near 4 GHz allows us to obtain $|\xi| > 1$ over a 40-MHz bandwidth near the designed pump frequency. The simulation

matches measurements of ξ (found by measuring the coupler Stark shift and comparing to its anharmonicity) up to a shift of 14.3 MHz in frequency and a factor of 2.84 in power. We attribute the discrepancies to our uncertainty in the cold attenuation of the microwave lines and the precise location of the coupler chip after mounting.

APPENDIX K: CALIBRATING DELAYS DURING CSWAP

As described in Sec. IV, during the CSWAP pulse, the mode operators \hat{a} and \hat{b} will exchange if the ancilla is in $|e\rangle$ and will stay the same if the ancilla is in $|g\rangle$ but they will also acquire a dynamical phase that differs depending on the ancilla state. This phase can be obtained by calculating half the solid angle enclosed by the trajectory of the mode operators on the Bloch sphere. A phase ϕ applied to a mode operator ($\hat{a} \rightarrow e^{i\phi}\hat{a}$) corresponds to a rotation of the phase space of that mode by an angle ϕ . In order to perform just a CSWAP, we need to correct for this ancilla-state-dependent phase.

This can be done by adding delays to the sequence. Due to the dispersive interaction, Bob's state will rotate by an angle $\phi = \chi t$ during a delay of length t only when Bob's transmon is excited. Since the CSWAP pulse exchanges the oscillator states when the transmon is in $|e\rangle$, a delay after the pulse rotates Bob's final state, whereas a delay before the pulse rotates Alice's final state. The addition of two delays, t_1 before the pulse and t_2 after the pulse, is sufficient to deterministically erase the ancilla-state-dependent phase. There is a remaining phase on the mode operators that is independent of the ancilla state but this can be removed by applying a frame update to each cavity in software, by a predetermined phase.

The delays and the frame updates are calibrated by initializing the cavities in $D(\alpha)|1\rangle \otimes D(-\alpha)|0\rangle$, performing the CSWAP pulse with the ancilla in either $|g\rangle$ or $|e\rangle$ and measuring the Wigner functions of each mode, as shown in the center columns of Fig. 3(c). The angle in phase space between the fitted center of the displaced Fock state and its expected position on the x axis, for both initial ancilla states, allows us to determine the delays and frame updates required. We repeat this process with these delays and frame updates applied to confirm that we produce the expected Wigner function.

Empirically, the dynamical phase described above is not the only phase that needs to be accounted for. For example, the phase offset of the beam-splitter drive relative to the difference between the cavity reference signals (either due to different electrical lengths in the cables or due to a phase offset that we control in the DAC) can advance or retard the phase of the mode operators when a SWAP is performed. Additionally, when the beam-splitter drive is applied, the cavities are Stark shifted by differing amounts, again imparting differing phases to the mode operators.

We find that the magnitude of these Stark shifts is also dependent on the ancilla state.

The degrees of freedom under our control, namely the pre- and postdelays, the frame updates on each cavity, and the pump phase, are sufficient to account for the phases acquired during a single CSWAP. However, these same parameters will not correct the phases on a subsequent CSWAP—a new set of parameters is required.

To correct the phase for a series of N CSWAP gates, we could calibrate the delay required before and after every CSWAP pulse, as well as the frame rotations associated with each one. Alternatively, however, a single delay before and after the entire sequence, as well as a single frame rotation on each cavity, is sufficient. Since this allows us to perform a long sequence while calibrating fewer parameters, this is what we opt to do in the experiment, as shown in Fig. 3(d). In order to calibrate these delays and frame rotations, we use the same procedure as outlined above but replace the single CSWAP pulse with N of them. Since the duration of the single delay before and after depends on the number of CSWAP pulses performed, the total duration of the sequence is not linear in N , as evidenced by the uneven top x -axis scale in Fig. 3(f).

APPENDIX L: JOINT WIGNER TOMOGRAPHY ON A BELL STATE IN THE COHERENT-STATE BASIS

To demonstrate the quantum nature of the CSWAP gate, we perform a SWAP test followed by joint Wigner tomography on the cavities. The sequence is shown in Fig. 11(a). Line cuts of the measured displaced parities along the $\text{Im}\beta_a = \text{Im}\beta_b = 0$ plane are shown in Fig. 11(b) and line cuts along the $\text{Re}\beta_a = \text{Re}\beta_b = 0$ plane are shown in Fig. 11(c). The measurements compare well with simulations of the ideal distributions (i.e., simulations without decoherence) shown in Figs. 11(d) and 11(e).

APPENDIX M: PAULI TOMOGRAPHY

For a single-oscillator state, the results of Wigner tomography can be used to perform a maximum-likelihood reconstruction of the density matrix. For many-oscillator states, though, the cost of this technique scales exponentially and to fully characterize even a two-oscillator density matrix would take several days.

If, however, we choose a qubit basis within the Hilbert space of each oscillator, we can obtain expectation values of all the Pauli operators by probing the Wigner function in just a handful of points in phase space [24,56]. In the coherent-state basis $\{|\alpha\rangle, |-\alpha\rangle\}$, single-oscillator tomography can be simplified by conducting four displaced-parity measurements, the linear combinations of which approximate Pauli measurements.

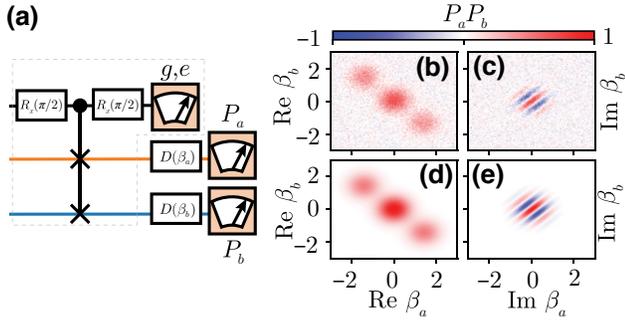


FIG. 11. Joint Wigner tomography on Ψ^+ . (a) A SWAP test sequence (in gray dashed lines) followed by a joint Wigner tomography sequence. The SWAP test is composed of a $\pi/2$ pulse on the control ancilla, followed by the CSWAP gate, the reverse $\pi/2$ pulse on the control ancilla, and its measurement. (b)–(e) Plane cuts from the (b),(c) measured and (d),(e) simulated joint Wigner distribution of $|\Psi^+\rangle$. Correlations between the real quadratures of the two cavities are shown in (b) and (d) in the plane $\text{Im}\beta_A = 0 = \text{Im}\beta_B$ and correlations between the imaginary quadratures are shown in (c) and (e) in the plane $\text{Re}\beta_A = 0 = \text{Re}\beta_B$.

To find the displacement β needed to perform a Pauli measurement, we first apply the projection operator

$$\hat{M} = (|\alpha\rangle\langle\alpha| + |-\alpha\rangle\langle-\alpha|) \quad (\text{M1})$$

to the displaced-parity operator $\hat{P}_\beta \equiv \hat{D}(\beta)\hat{P}\hat{D}(\beta)^\dagger$ to project it onto the coherent-state basis

$$\hat{M}\hat{P}_\beta\hat{M}^\dagger = \begin{pmatrix} \langle\alpha|P_\beta|\alpha\rangle & \langle\alpha|P_\beta|-\alpha\rangle \\ \langle-\alpha|P_\beta|\alpha\rangle & \langle-\alpha|P_\beta|-\alpha\rangle \end{pmatrix}. \quad (\text{M2})$$

We then find a set of displacements $\{\beta_i\}$ such that $\sum_i b_i \hat{M}\hat{P}_{\beta_i}\hat{M}^\dagger \in \{I, X, Y, Z\}$ for some coefficients b_i . (Here, I , X , Y , and Z are Pauli matrices; we omit the hats on these operators for brevity.)

When finding these displacements, the following relations prove useful:

$$\langle\beta|P_0|\beta\rangle = \langle\beta|-\beta\rangle = e^{-2|\beta|^2}, \quad (\text{M3})$$

$$\langle\beta|P_0|-\beta\rangle = \langle\beta|\beta\rangle = 1, \quad (\text{M4})$$

$$\langle\alpha|P_\beta|\alpha\rangle = \langle\alpha-\beta|-\alpha+\beta\rangle = e^{-2|\alpha-\beta|^2}, \quad (\text{M5})$$

$$\langle\alpha|P_\beta|-\alpha\rangle = \langle\alpha-\beta|\alpha+\beta\rangle = e^{2(\alpha^*\beta-\beta^*\alpha)}\langle\beta|-\beta\rangle. \quad (\text{M6})$$

We note that these properties are identical to those quoted in Ref. [56] except that in Eq. (M5) we do not conclude that $\langle\alpha|P_\beta|\alpha\rangle \ll 1$ (e.g., when $\beta = \alpha$, $\langle\alpha|P_\beta|\alpha\rangle = 1$).

We now choose to write the identity matrix I as the sum of $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ and $\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ and the Pauli- Z matrix as the difference. Setting $\hat{M}\hat{P}_{\beta_1}\hat{M}^\dagger = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, we must satisfy $\langle\alpha|\hat{P}_{\beta_1}|\alpha\rangle = 1$, which sets $\beta_1 = \alpha$. We verify that $\langle\alpha|\hat{P}_\alpha|-\alpha\rangle =$

$\langle-\alpha|\hat{P}_\alpha|\alpha\rangle \approx 0$ and $\langle-\alpha|\hat{P}_\alpha|-\alpha\rangle \approx 0$. Similarly, setting $\hat{M}\hat{P}_{\beta_2}\hat{M}^\dagger = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$, we obtain $\beta_2 = -\alpha$.

I can be constructed by summing the above two displaced-parity measurements,

$$I = \hat{M}\hat{P}_\alpha\hat{M}^\dagger + \hat{M}\hat{P}_{-\alpha}\hat{M}^\dagger, \quad (\text{M7})$$

while Z can be constructed by taking the difference,

$$Z = \hat{M}\hat{P}_\alpha\hat{M}^\dagger - \hat{M}\hat{P}_{-\alpha}\hat{M}^\dagger. \quad (\text{M8})$$

We then find the displacement β required to measure X by setting $\hat{M}\hat{P}_{\beta_3}\hat{M}^\dagger = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, which is satisfied by $\beta_3 = 0$. Thus measuring the photon-number parity without any displacement is equivalent to a Pauli- X measurement in this basis:

$$X = \hat{M}\hat{P}_0\hat{M}^\dagger. \quad (\text{M9})$$

Finally, we solve $\hat{M}\hat{P}_{\beta_4}\hat{M}^\dagger = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$ to obtain a Pauli- Y measurement. To ensure the correct relative phase between $\langle\alpha|P_{\beta_4}|-\alpha\rangle$ and $\langle-\alpha|P_{\beta_4}|\alpha\rangle$, we need $\beta_4 = ik\pi/8\alpha$, where k is any odd integer. However, choosing a larger $|\beta_4|$ results in a smaller magnitude for $\langle\alpha|P_{\beta_4}|-\alpha\rangle$ and $\langle-\alpha|P_{\beta_4}|\alpha\rangle$. To maximize the off-diagonal matrix elements, we thus choose $\beta_4 = i\pi/8\alpha$.

For $\alpha = \sqrt{2}$, this choice of β_4 yields off-diagonal elements with a magnitude of approximately 0.86, so the single-oscillator Pauli- Y measurement approximated by this displaced-parity measurement has a maximum amplitude of approximately 0.86. For general α ,

$$\hat{M}\hat{P}_{\frac{i\pi}{8\alpha}}\hat{M}^\dagger = \left\langle \frac{i\pi}{8\alpha} \middle| -\frac{i\pi}{8\alpha} \right\rangle Y = e^{-\frac{\pi^2}{32\alpha^2}} Y. \quad (\text{M10})$$

For a visual interpretation of how displaced-parity measurements can be used to approximate Pauli measurements, we mark the displacements of interest on Wigner functions of the states $|+Z\rangle$, $|-Z\rangle$, $|+X\rangle$, and $|+Y\rangle$ in Fig. 12.

To move from a single-oscillator measurement to a two-oscillator measurement, we need all 16 combinations of single-oscillator Pauli measurements $\{I, X, Y, Z\}$. The 16 displacements whose linear combination make up the two-oscillator Pauli measurements are labeled in Figs. 12(e)–12(h). Note that the two-oscillator measurement of YY has a maximum of amplitude of $e^{-\pi/16\alpha^2} \approx 0.73$ for $\alpha = \sqrt{2}$.

APPENDIX N: CSWAP ERROR BUDGET

For both the Bell-state preparation and the control experiment where we perform a SWAP test on $|\alpha, \alpha\rangle$ [see Figs. 3(c)–3(f)], we construct an error budget based on independent measurements of coherence times (summarized in Table I), the duration of each part of the experimental sequence, and the sensitivity of each measurement to various error channels.

The sequence in Fig. 3(d) can be divided into four parts: the controlled SWAP (including delays), the readout of the control transmon, the parity-mapping sequence for joint Wigner tomography, and the final transmon measurements.

The error channels considered are decay and dephasing on both ancillas and cavities. In this analysis, we neglect cavity no-jump evolution, errors during ancilla rotations, and coupler errors, since we expect them to have a negligible effect on fidelities.

The Bell-state fidelity is comprised of the expectation values of four two-qubit Pauli operators, each of which is affected differently by errors. For example, $\langle XX \rangle$ is the measured joint photon-number parity and so is only affected by errors that change the joint parity. Likewise, $\langle YY \rangle$ is measured near the origin of the joint Wigner function and is sensitive to parity jumps but less so to phase-space rotations. Any error that takes the even Bell state to the odd Bell state results in a flip of the measured $\langle XX \rangle$ or $\langle YY \rangle$ value.

In contrast, $\langle II \rangle$ and $\langle ZZ \rangle$ are identical for even and odd Bell states. They are, however, much more sensitive to rotations in phase space (coherent and incoherent).

Table IV summarizes the impact of errors on each Pauli measurement and how this contributes to the overall fidelity. Detailed justification for each error considered is provided below.

1. Errors during CSWAP

Following the CSWAP, the control-oscillator system will ideally be in the state

$$|\psi\rangle_{\text{post-CSWAP}} = \frac{|g\rangle |\alpha, -\alpha\rangle + |e\rangle |-\alpha, \alpha\rangle}{\sqrt{2}}. \quad (\text{N1})$$

The subsequent ancilla $-\pi/2$ rotation then generates a superposition of Bell states entangled with the ancilla,

$$|\psi\rangle_{\text{postrotation}} = \frac{|g\rangle |\Psi^+\rangle + |e\rangle |\Psi^-\rangle}{\sqrt{2}}, \quad (\text{N2})$$

at which point postselecting on the outcome $|g\rangle$ projects us onto the desired Bell state.

Ancilla dephasing errors during the CSWAP, however, can be mapped to ancilla bit-flip errors immediately after the $-\pi/2$ rotation. The result is that our postselection scheme will keep a shot in which we have prepared $|\Psi^-\rangle$, for which $\langle XX \rangle$ and $\langle YY \rangle$ have the opposite sign. As discussed previously, $\langle II \rangle$ and $\langle ZZ \rangle$ are identical for both states $|\Psi^+\rangle$ and $|\Psi^-\rangle$, and so will be unaffected.

Oscillator decay transforms simply under the beam-splitter Hamiltonian, such that an oscillator decay event at any point during a SWAP is equivalent to a superposition of decay on Alice and Bob occurring immediately after. In our case, we have the state $(\cos(gt_{\text{err}})\hat{a} +$

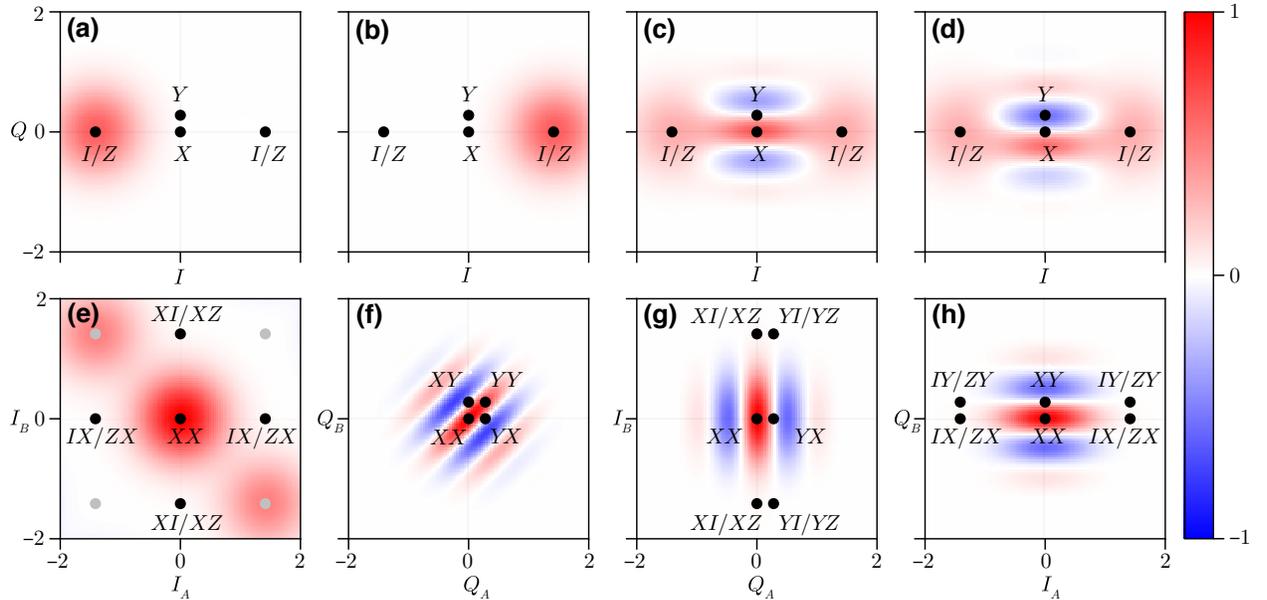


FIG. 12. Local and joint Wigner plots. The top row shows the displaced photon-number parity of single-oscillator states in the coherent-state bases (a) $|-Z\rangle = |-\alpha\rangle$, (b) $|+Z\rangle = |\alpha\rangle$, (c) $|+X\rangle = (|\alpha\rangle + |-\alpha\rangle)/N$, and (d) $|+Y\rangle = (|\alpha\rangle + i|-\alpha\rangle)/N$ for $\alpha = \sqrt{2}$. The black dots indicate the positions at which the displaced parity is sampled to obtain measurements of the single-oscillator Pauli operators. The bottom row shows slices of the displaced joint photon-number parity of the two-oscillator state $|\Psi^+\rangle$ along (e) real axes of both oscillators, (f) imaginary axes of both oscillators, (g) the imaginary axis of Alice and the real axis of Bob, and (h) vice versa. Here, the black dots indicate the positions at which the displaced joint parity is sampled to obtain measurements of the two-oscillator Pauli operators. The II , ZZ , IZ , and ZI measurements are obtained from linear combinations of measurements at the four gray dots.

TABLE III. The mode couplings.

		Alice	Bob
Transmon-oscillator dispersive shift	$\chi_{qc}/2\pi$	-766.4(5) kHz	-1104.0(5) kHz
Transmon anharmonicity and cavity dispersive shift	$K_q/4\pi$	-181.2537(4) MHz	-184.2860(3) MHz

$i \sin(gt_{\text{err}})\hat{b})|\psi\rangle_{\text{post-CSWAP}}$, where t_{err} describes when during the SWAP the error has occurred.

Ancilla decay during the CSWAP does not transform as simply. In the Heisenberg picture, where the evolution of cavity operators is visualized on a sphere, ancilla decay changes the rotation axis about which the Heisenberg operators rotate on the Bloch sphere partway through their trajectory. On average, this will scramble the cavity states; thus we expect this error to reduce all joint-Pauli measurements.

Finally, oscillator dephasing errors should primarily affect the II and ZZ measurements. This is because these are measured by probing points far from the origin of the joint Wigner function in phase space, where even small phase-space rotations have a large impact.

2. Errors during CSWAP readout

During the CSWAP readout, ancilla measurement (and postselection on $|g\rangle$) ideally projects the oscillators onto an even Bell state. However, an oscillator decay event in either cavity during this time will flip the joint photon-number parity of the state and thus the measurement outcomes for both XX and YY . Note that this happens regardless of when during the readout the photon loss occurs.

Ancilla decay during the readout leads not just to a projection on the incorrect Bell state but also to a random rotation of Bob's phase due to the dispersive shift to the ancilla, depending on when during the readout the error occurs. The XX -measurement outcome, which is insensitive to phase, is therefore flipped but the effect on $\langle YY \rangle$, $\langle II \rangle$ and $\langle ZZ \rangle$ is to reduce their amplitude uniformly.

Oscillator decay during the CSWAP readout will affect $\langle II \rangle$ and $\langle ZZ \rangle$ in the same manner as during the CSWAP itself.

3. Errors during parity measurement

Ancilla errors during the Wigner tomography, whether dephasing during the parity map or decay during the parity map and subsequent readout, affect all joint-Pauli measurements equally, flipping the sign of the result. The reduction in the measured value of each joint Pauli is thus reduced, on average, by twice the error probability.

Oscillator decay during the parity map affects the joint-Pauli measurements depending on the oscillator state after the Wigner displacement. For the XX measurement, this is (ideally) an eigenstate of joint parity and so decay will

change the joint parity of the oscillator state. The probability of measuring ± 1 depends on when during the parity map the error occurs and so we can treat it as depolarizing the measurement results.

For the II and ZZ measurements, however, the states after displacement are superpositions of vacuum and a coherent state, e.g., $|\psi\rangle = |0, 0\rangle + |2\alpha, -2\alpha\rangle$. Since $2|\alpha| \gg 1$, photon loss only negligibly affects the joint parity of this state.

4. Effects of Kerr and χ'

As the number of CSWAP operations N increases in Fig. 3(f), the effects of higher-order nonlinearities in the joint-cavity Hamiltonian become non-negligible. The most noticeable ones are cavity self-Kerrs and χ' .

Self-Kerrs of the cavities as shown in line three of Eq. (4) result in detunings of the cavity frequencies that depend on the cavity photon number quadratically. This accumulation of photon-number-dependent phase manifests as rotation and smearing of the cavity states in phase space. What further complicates the story is the sixth-order nonlinearity $\chi'_{bc}\hat{b}^{\dagger 2}\hat{b}^2(|e\rangle\langle e|)_e$. This parasitic term distorts Bob's cavity states in the same way as cavity self-Kerr but only when the control transmon is excited. At the operating point for the CSWAP experiment, away from the flux point where self-Kerr is minimized, we measure cavity self-Kerrs $K_a \approx -0.8$ kHz, $K_b \approx -3.0$ kHz, and $\chi'_{bc} \approx 0.9$ kHz.

The combined effect of the cavity self-Kerrs and χ' is a photon-number- and control-transmon-dependent rotation and smearing of cavity states in phase space. The rotation is accounted for when the pre- and postdelays are calibrated for different N but this cannot correct the smearing of the states, which is due to photon-number-dependent rotations. This smearing thus adds to the SPAM-corrected infidelities extracted from the slope in Fig. 3(f).

Comparing the simulated fidelity up to $N = 13$ rounds of CSWAP with and without self-Kerr terms and the χ' term in the Hamiltonian shows an increase in the slope of $0.19\% \pm 0.06\%$ per round for the Bell-state preparation and $0.10\% \pm 0.07\%$ per round for the control experiment. The error bars arise from uncertainty in the phase of the beam-splitter drive, which affects whether more photons are present in Alice (with less Kerr) or Bob (with more Kerr).

TABLE IV. The error budget for Bell-state preparation using CSWAP. The columns show the predicted decrease of the measured value of the nonzero joint-Pauli measurements for each error channel, the probability of occurrence of which is estimated by the expression in column 3. The values in parentheses are the predictions based, instead, on a Lindblad master equation simulation. $\bar{n} = 2$ is the average photon number in each oscillator. The time scales used are set out in Table I. The *YY* column has a scale factor of 73.46% applied, due to the quasiorthogonality of the basis states (see Appendix M).

Error channel	Probability of occurrence	Estimated (simulated) effect on <i>XX</i>	Estimated (simulated) effect on <i>YY</i>	Estimated (simulated) effect on <i>II</i>	Estimated (simulated) effect on <i>ZZ</i>
CSWAP					
Oscillator decay \hat{a}, \hat{b}	$\bar{n} (\tau_{\text{CSWAP}}/T_{1cA} + \tau_{\text{CSWAP}}/T_{1cB}) \approx 3.5\% \pm 0.1\%$	3.5(1)% (3.7%)	2.57(7)% (2.7%)	0% (0.2%)	0% (0.3%)
Oscillator dephasing $\hat{a}^\dagger \hat{a}, \hat{b}^\dagger \hat{b}$	$2\bar{n} (\tau_{\text{CSWAP}}/T_{\phi cA} + \tau_{\text{CSWAP}}/T_{\phi cB}) \approx 0.90\% \pm 0.15\%$	0% (0.3%)	0% (0.1%)	1.8(3)% (1.7%)	1.8(3)% (1.8%)
Ancilla decay $ e\rangle \rightarrow g\rangle$	$\frac{1}{2} \tau_{\text{CSWAP}}/T_{1qB} \approx 1.17\% \pm 0.01\%$	1.17(1)% (1.3%)	0.859(7)% (0.9%)	1.17(1)% (0.8%)	1.17(1)% (1.2%)
Ancilla dephasing $ e\rangle\langle e $	$\frac{\tau_{\text{CSWAP}}}{2T_{\phi qB}} \approx 0.59\% \pm 0.03\%$	1.18(6)% (1.3%)	0.87(4)% (0.9%)	0% (0.1%)	0% (0.2%)
CSWAP RO					
Oscillator decay \hat{a}, \hat{b}	$\bar{n} (\tau_{\text{RO}}/T_{1cA} + \tau_{\text{RO}}/T_{1cB}) \approx 4.7\% \pm 0.2\%$	9.4(3)% (9.0%)	6.9(2)% (6.5%)	0% (0%)	0% (0%)
Oscillator dephasing $\hat{a}^\dagger \hat{a}, \hat{b}^\dagger \hat{b}$	$2\bar{n} (\tau_{\text{RO}}/T_{\phi cA} + \tau_{\text{RO}}/T_{\phi cB}) \approx 1.2\% \pm 0.2\%$	0% (0%)	0% (0%)	2.4(4)% (1.8%)	2.4(4)% (1.9%)
Parity map					
Ancilla decay $ e\rangle \rightarrow g\rangle$	$\tau_{\text{RO}}/T_{1qB} \approx 3.15\% \pm 0.03\%$	6.30(7)% (6.0%)	3.15(3)% (2.9%)	3.15(3)% (2.1%)	3.15(3)% (3.1%)
Oscillator decay \hat{a}, \hat{b}	$XX/YY: \bar{n} (\tau_{pA}/T_{1cA} + \tau_{pB}/T_{1cB}) \approx 1.20\% \pm 0.04\%$	1.20(4)% (1.2%)	0.88(3)% (0.8%)	0% (0%)	0% (0%)
	$II/ZZ: 2\bar{n} (\tau_{pA}/T_{1cA} + \tau_{pB}/T_{1cB}) \approx 2.4\% \pm 0.08\%$				
Ancilla decay $ e\rangle \rightarrow g\rangle$	$\frac{1}{2} (\tau_{pA}/T_{1qA} + \tau_{pB}/T_{1qB}) \approx 0.620\% \pm 0.005\%$	0.620(5)% (0.6%)	0.455(4)% (0.4%)	0.620(5)% (0.5%)	0.620(5)% (0.6%)
Ancilla dephasing $ e\rangle\langle e $	$\frac{\tau_{pA}}{2T_{\phi qA}} + \tau_{pB}/2T_{\phi qB} \approx 0.34\% \pm 0.01\%$	0.68(3)% (0.6%)	0.50(2)% (0.4%)	0.68(3)% (0.5%)	0.68(3)% (0.7%)
Parity RO					
Ancilla decay $ e\rangle \rightarrow g\rangle$	$\frac{1}{2} (\tau_{\text{RO}}/T_{1qA} + \tau_{\text{RO}}/T_{1qB}) \approx 2.28\% \pm 0.02\%$	4.57(4)% (4.5%)	3.36(3)% (3.3%)	4.57(4)% (4.4%)	4.57(4)% (4.5%)
	Estimated gate error per round	5.7(1)%	4.2(1)%	2.9(3)%	2.9(3)%
			4.0% ± 0.1%		
	Estimated infidelity per round, including K, χ		4.2% ± 0.1%		
	Estimated (simulated) Bell fidelity	74.2(4)% (75.3%)	56.5(3)% (57.1%)	86.5(6)% (88.5%)	86.5(6)% (87.4%)
			75.9% ± 0.2%		
	Measured Bell fidelity	70.4(4)%	51.5(5)%	87(1)%	87(1)%
			74.1% ± 0.4%		

TABLE V. The error budget for the control experiment. The estimated reduction in the measured fidelity to the target state due to each error channel. The values in parentheses are predicted infidelities from master-equation simulations. The effects of Kerr and χ' are described in Appendix N 4.

	Error channel	Probability of occurrence	Estimated (simulated) effect on fidelity
CSWAP	Ancilla decay $ e\rangle\langle e $	$\tau_{\text{CSWAP}}/2T_{1qB} \approx 1.17\% \pm 0.02\%$	$1.17\% \pm 0.02\%$ (1.0%)
	Oscillator dephasing $\hat{a}^\dagger\hat{a}, \hat{b}^\dagger\hat{b}$	$2\bar{n}(\tau_{\text{CSWAP}}/T_{\phi cA} + \tau_{\text{CSWAP}}/T_{\phi cB}) \approx 0.9\% \pm 0.15\%$	$1.8\% \pm 0.3\%$ (1.3%)
CSWAP RO	Oscillator dephasing $\hat{a}^\dagger\hat{a}, \hat{b}^\dagger\hat{b}$	$2\bar{n}(\tau_{\text{RO}}/T_{\phi cA} + \tau_{\text{RO}}/T_{\phi cB}) \approx 1.2\% \pm 0.2\%$	$2.4\% \pm 0.4\%$ (1.9%)
Parity map	Ancilla decay $ e\rangle \rightarrow g\rangle$	$\frac{1}{2}(\tau_{pA}/T_{1qA} + \tau_{pB}/T_{1qB}) \approx 0.620\% \pm 0.005\%$	$0.620\% \pm 0.005\%$ (0.6%)
	Ancilla dephasing $ e\rangle\langle e $	$\tau_{pA}/2T_{\phi qA} + \tau_{pB}/2T_{\phi qB} \approx 0.34\% \pm 0.01\%$	$0.68\% \pm 0.03\%$ (0.6%)
Parity RO	Ancilla decay $ e\rangle \rightarrow g\rangle$	$\frac{1}{2}(\tau_{\text{RO}}/T_{1qA} + \tau_{\text{RO}}/T_{1qB}) \approx 2.28\% \pm 0.02\%$	$4.57\% \pm 0.04\%$ (4.5%)
	Estimated gate error per round		$3.0\% \pm 0.3\%$
	Estimated infidelity per round, including K and χ'		$3.1\% \pm 0.3\%$
	Estimated (simulated) fidelity		$89.1\% \pm 0.7\%$ (90.6%)
	Measured fidelity		89.5%

5. Combining errors

We estimate the combined fidelity as

$$\mathcal{F} \approx \prod_i (1 - p_i) + \sum_i q_i p_i \prod_{j \neq i} (1 - p_j), \quad (\text{N3})$$

where p_i is the probability of an error in the i^{th} channel and $q_i = 0$ or -1 is the expected measured value in the event of the error occurring. This expression neglects contributions

from multiple errors happening during the same run, which is a reasonable approximation given that $p_i \ll 1 \forall i$.

6. Control experiment on identical oscillator states

To validate our ability to predict errors, we conduct a control experiment where the cavities are initialized in $|\alpha, \alpha\rangle$ and perform a SWAP test on these nominally identical states. For this experiment, we measure fidelity to the target state $|\alpha, \alpha\rangle$ by probing the joint Wigner at $|\alpha, \alpha\rangle$.

This control experiment is insensitive to oscillator decay. Since the oscillator states remain eigenstates of the respective collapse operators during the CSWAP, they are only affected by no-jump shrinking, which is negligible for the time scales we are measuring. Meanwhile, during the Wigner tomography, the displacement brings the oscillator state back to the origin where oscillator decay has no

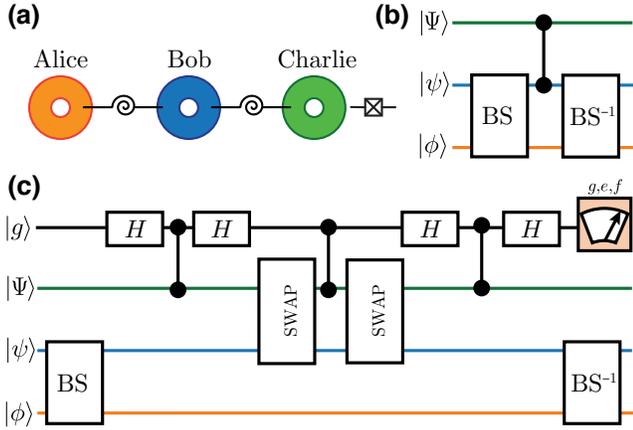


FIG. 13. A CSWAP with three bosonic qubits. (a) A schematic of the hardware required for the proposed CSWAP on three bosonically encoded qubits: Alice, Bob, and Charlie are bosonic modes hosted in high- Q stub cavities [67] and linked by SNAIL couplers of the kind presented in this work. A transmon ancilla is inserted into Charlie's cavity; it provides the requisite nonlinearity. (b) The canonical construction of a CSWAP from a beam splitter, a controlled-phase shift, and an inverse beam splitter (shown as “BS⁻¹”). (c) The full sequence for the proposed CSWAP. SWAP gates are unconditioned on the state of the transmon and only affect the bosonic modes—that is, they exchange the states of Alice and Charlie and do not affect the transmon rail.

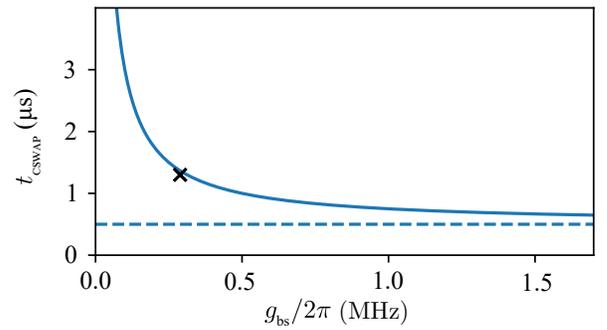


FIG. 14. A comparison of CSWAP approaches. The blue solid line shows the gate time as a function of the achievable beam-splitter rate for a CSWAP composed of a controlled phase shift sandwiched between two unconditional beam-splitter operations, with the dashed line showing the limit as the beam-splitter rate increases. The cross indicates the method implemented in this work. Both assume a dispersive shift to the transmon of $\chi/2\pi = -1.104$ MHz.

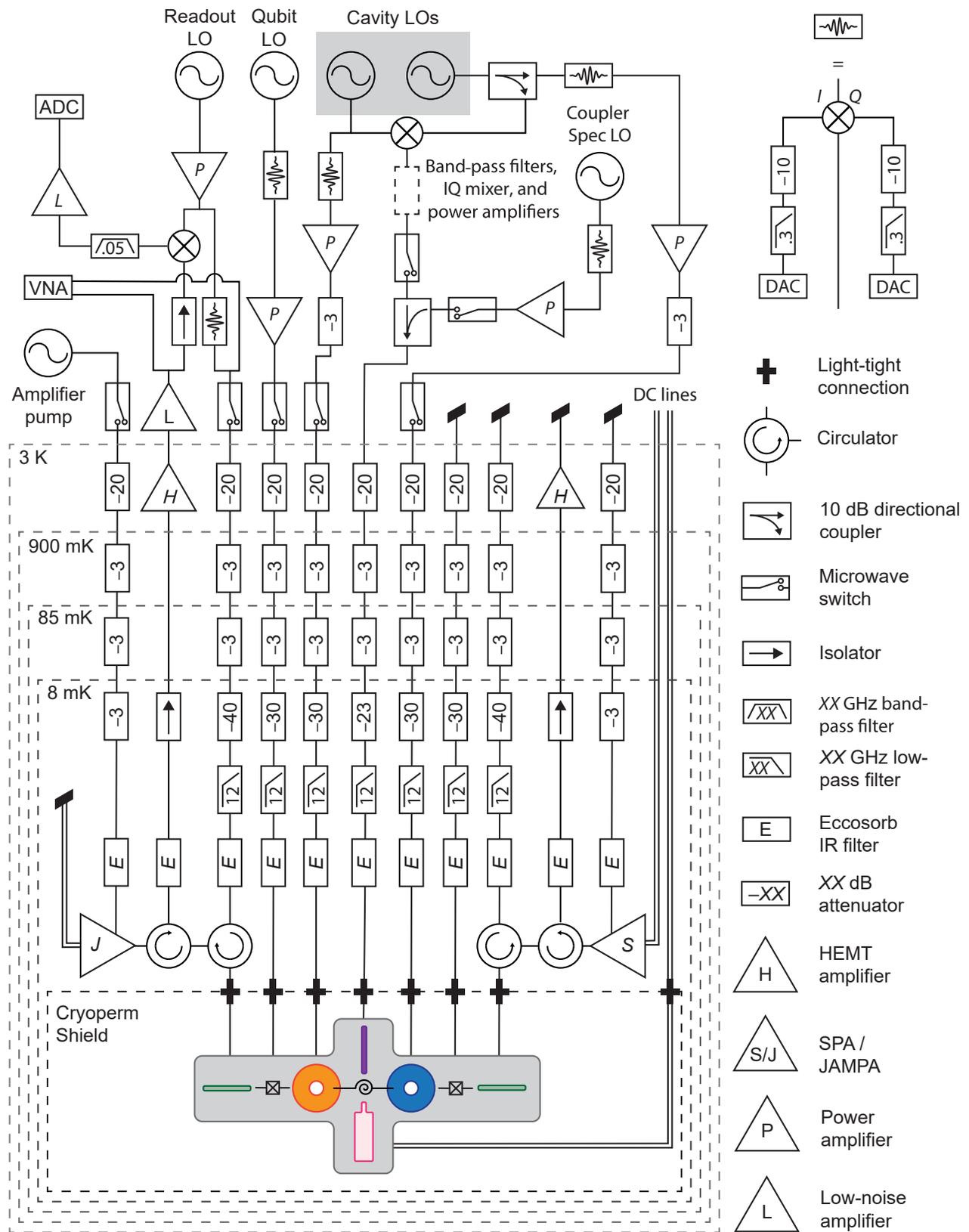


FIG. 15. The wiring diagram. The gray region within the Cryoperm shield shows the experimental package, containing two microwave cavities (orange and blue), a SNAIL coupler, two transmon ancillas, two readout resonators (green), a buffer mode (purple), and a flux transformer (pink). The on-chip Purcell filters (present in the experiment) are omitted from the diagram.

impact. For the same reason, oscillator dephasing has no impact during the Wigner tomography.

The joint-Pauli measurements are also impervious to transmon dephasing during the SWAP test because all $|e\rangle$ measurements resulting from transmon dephasing are discarded. So the only error channels we consider for this control experiment are ancilla decay and oscillator dephasing during the SWAP test, oscillator dephasing during the readout of the SWAP test, and transmon errors during the parity map and its readout. These estimated infidelity due to these error channels, shown in Table V, is consistent with the measured value.

APPENDIX O: CSWAP WITH THREE BOSONICALLY ENCODED QUBITS

A SWAP of two bosonically encoded qubits controlled on the state of a third bosonically encoded qubit can be realized with the hardware shown in Fig. 13(a). The proposed sequence utilizes the canonical construction of CSWAP from beam splitters and controlled phase shifts [Fig. 13(b)]. To realize the controlled phase shift between Charlie and Alice, the transmon ancilla serves as an intermediary. After the beam splitter, Charlie’s state is mapped to the transmon. A SWAP then exchanges Alice’s and Charlie’s states, such that the following controlled phase shift between the transmon (now storing Charlie’s original state) and Charlie (now storing Alice’s original state) accomplishes the desired operation. The sequence is then reversed to disentangle the transmon. Measurement of the transmon rail at the conclusion of the gate can provide a flag to signal errors.

We note that the efficacy of this sequence requires that the SWAP gates are unconditioned on the state of the transmon. One way to accomplish this is by pumping the beam-splitter interaction with $\delta = \chi/2$ while driving the transmon resonantly with Rabi-rate $\Omega \gg g_{bs}$ [55]. The fine tuning of Ω such that $2\Omega t_{bs} = \Omega t_{SWAP} = 2n\pi$ for some integer n ensures that the transmon state is also unchanged by this dynamical decoupling.

APPENDIX P: COMPARING CONTINUOUS AND TROTTERIZED CSWAP SEQUENCES

As illustrated in Fig. 13(b), a CSWAP can be constructed from a 50:50 beam splitter (unconditional on the ancilla state), a parity map (enacted by a delay of π/χ), and another unconditional 50:50 beam splitter, with opposite phase to the first. This differs from the approach taken in Sec. IV, in which we let the ancilla and both cavities interact simultaneously.

Analogously to the unconditional SWAP discussed in Appendix O, the unconditional 50:50 beam splitter can be performed by pumping the beam-splitter interaction

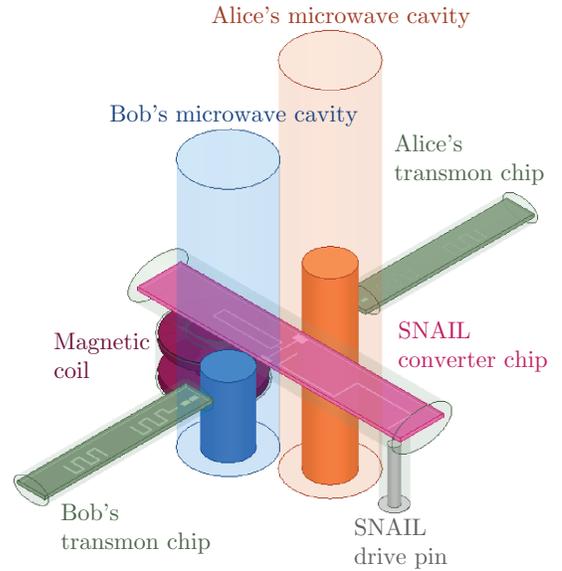


FIG. 16. The experimental device model. A schematic model of the superconducting device. The structure is machined out of high-purity aluminum. Each transmon chip houses a readout resonator and a Purcell filter, in addition to the transmon. Cavity drive pins, transmon drive pins, and transmon readout pins are omitted in this model for simplicity.

with $\delta = \chi/2$ while resonantly driving the control transmon with Rabi-rate $\Omega \gg g_{bs}$, in a time $\pi/4g_{bs}$. Assuming no restriction on how fast the beam-splitter pulse can be ramped on or off, this gives a duration of $\pi/2g_{bs} + \pi/\chi$ for the full CSWAP. (In practice, ensuring that the spectral content of the beam-splitter pulse is sufficiently narrow to avoid driving unwanted processes introduces some non-negligible ramp time.) In Fig. 14, this gate duration (solid blue line) is plotted as a function of g_{bs} for $\chi/2\pi = -1.104$ MHz, the value observed in our system. Comparing this “Trotterized” method to the performance of the “continuous” scheme used in this paper (black cross), we see that they give similar gate durations, given the same beam-splitter rate. The flexibility of the “Trotterized” method to further reduce the gate time as the achievable beam-splitter rate increases makes it a promising scheme for future CSWAP implementations.

APPENDIX Q: EXPERIMENTAL SETUP

To help minimize phase drift between our cavity drives, we source the local oscillators for Alice and Bob from channels 1 and 2 of the same SignalCore SC5510A generator. For quantum limited amplification of our readout signal at the base of the refrigerator, we use a Josephson-array-mode parametric amplifier (JAMPA) [64] and a lumped SNAIL parametric amplifier (LSPA) [34]. For a wiring diagram for this experiment, see Fig 16.

- [1] P. T. Cochrane, G. J. Milburn, and W. J. Munro, Macroscopically distinct quantum-superposition states as a bosonic code for amplitude damping, *Phys. Rev. A* **59**, 2631 (1999).
- [2] D. Gottesman, A. Kitaev, and J. Preskill, Encoding a qubit in an oscillator, *Phys. Rev. A* **64**, 012310 (2001).
- [3] M. Mirrahimi, Z. Leghtas, V. V. Albert, S. Touzard, R. J. Schoelkopf, L. Jiang, and M. H. Devoret, Dynamically protected cat-qubits: A new paradigm for universal quantum computation, *New J. Phys.* **16**, 045014 (2014).
- [4] M. H. Michael, M. Silveri, R. T. Brierley, V. V. Albert, J. Salmilehto, L. Jiang, and S. M. Girvin, New Class of Quantum Error-Correcting Codes for a Bosonic Mode, *Phys. Rev. X* **6**, 031006 (2016).
- [5] S. Puri, S. Boutin, and A. Blais, Engineering the quantum states of light in a Kerr-nonlinear resonator by two-photon driving, *npj Quantum Inf.* **3**, 1 (2017).
- [6] I. L. Chuang, D. W. Leung, and Y. Yamamoto, Bosonic quantum codes for amplitude damping, *Phys. Rev. A* **56**, 1114 (1997).
- [7] M. Reagor, H. Paik, G. Catelani, L. Sun, C. Axline, E. Holland, I. M. Pop, N. A. Masluk, T. Brecht, and L. Frunzio, *et al.*, Reaching 10 ms single photon lifetimes for superconducting aluminum cavities, *Appl. Phys. Lett.* **102**, 192604 (2013).
- [8] A. Romanenko, R. Pilipenko, S. Zorzetti, D. Frolov, M. Awida, S. Belomestnykh, S. Posen, and A. Grassellino, Three-Dimensional Superconducting Resonators at $T < 20$ mK with Photon Lifetimes up to $\tau = 2$ s, *Phys. Rev. Appl.* **13**, 034032 (2020).
- [9] W. Cai, Y. Ma, W. Wang, C.-L. Zou, and L. Sun, Bosonic quantum error correction codes in superconducting quantum circuits, *Fund. Res.* **1**, 50 (2021).
- [10] A. Blais, R.-S. Huang, A. Wallraff, S. M. Girvin, and R. J. Schoelkopf, Cavity quantum electrodynamics for superconducting electrical circuits: An architecture for quantum computation, *Phys. Rev. A* **69**, 062320 (2004).
- [11] Z. Leghtas, G. Kirchmair, B. Vlastakis, M. H. Devoret, R. J. Schoelkopf, and M. Mirrahimi, Deterministic protocol for mapping a qubit to coherent state superpositions in a cavity, *Phys. Rev. A* **87**, 042315 (2013).
- [12] R. W. Heeres, P. Reinhold, N. Ofek, L. Frunzio, L. Jiang, M. H. Devoret, and R. J. Schoelkopf, Implementing a universal gate set on a logical qubit encoded in an oscillator, *Nat. Commun.* **8**, 1 (2017).
- [13] T. Fösel, S. Krastanov, F. Marquardt, and L. Jiang, Efficient cavity control with SNAP gates (2020), arXiv preprint [ArXiv:2004.14256](https://arxiv.org/abs/2004.14256).
- [14] M. Kudra, M. Kervinen, I. Strandberg, S. Ahmed, M. Scigliuzzo, A. Osman, D. P. Lozano, M. O. Tholén, R. Borgani, D. B. Haviland, G. Ferrini, J. Bylander, A. F. Kockum, F. Quijandria, P. Delsing, and S. Gasparinetti, Robust Preparation of Wigner-Negative States with Optimized SNAP-Displacement Sequences, *PRX Quantum* **3**, 030301 (2022).
- [15] A. Eickbusch, V. Sivak, A. Z. Ding, S. S. Elder, S. R. Jha, J. Venkatraman, B. Royer, S. M. Girvin, R. J. Schoelkopf, and M. H. Devoret, Fast universal control of an oscillator with weak dispersive coupling to a qubit, *Nat. Phys.* **18**, 1464 (2022).
- [16] S. Rosenblum, Y. Y. Gao, P. Reinhold, C. Wang, C. J. Axline, L. Frunzio, S. M. Girvin, L. Jiang, M. Mirrahimi, and M. H. Devoret, *et al.*, A CNOT gate between multi-photon qubits encoded in two cavities, *Nat. Commun.* **9**, 1 (2018).
- [17] P. Reinhold, S. Rosenblum, W.-L. Ma, L. Frunzio, L. Jiang, and R. J. Schoelkopf, Error-corrected gates on an encoded qubit, *Nat. Phys.* **16**, 822 (2020).
- [18] S. Puri, A. Grimm, P. Campagne-Ibarcq, A. Eickbusch, K. Noh, G. Roberts, L. Jiang, M. Mirrahimi, M. H. Devoret, and S. M. Girvin, Stabilized cat in a driven nonlinear cavity: A fault-tolerant error syndrome detector, *PRX* **9**, 041009 (2019).
- [19] A. Grimm, N. E. Frattini, S. Puri, S. O. Mundhada, S. Touzard, M. Mirrahimi, S. M. Girvin, S. Shankar, and M. H. Devoret, Stabilization and operation of a Kerr-cat qubit, *Nature* **584**, 205 (2020).
- [20] L. Hu, Y. Ma, W. Cai, X. Mu, Y. Xu, W. Wang, Y. Wu, H. Wang, Y. P. Song, and C.-L. Zou, *et al.*, Quantum error correction and universal gate set operation on a binomial bosonic logical qubit, *Nat. Phys.* **15**, 503 (2019).
- [21] P. Campagne-Ibarcq, A. Eickbusch, S. Touzard, E. Zalusky, N. E. Frattini, V. V. Sivak, P. Reinhold, S. Puri, S. Shankar, and R. J. Schoelkopf, *et al.*, Quantum error correction of a qubit encoded in grid states of an oscillator, *Nature* **584**, 368 (2020).
- [22] V. V. Sivak, A. Eickbusch, B. Royer, S. Singh, I. Tsioutsios, S. Ganjam, A. Miano, B. L. Brock, A. Z. Ding, and L. Frunzio, *et al.*, Real-time quantum error correction beyond break-even (2022), arXiv preprint [ArXiv:2211.09116](https://arxiv.org/abs/2211.09116).
- [23] Z. Ni, S. Li, X. Deng, Y. Cai, L. Zhang, W. Wang, Z.-B. Yang, H. Yu, F. Yan, and S. Liu, *et al.*, Beating the break-even point with a discrete-variable-encoded logical qubit (2022), arXiv preprint [ArXiv:2211.09319](https://arxiv.org/abs/2211.09319).
- [24] C. Wang, Y. Y. Gao, P. Reinhold, R. W. Heeres, N. Ofek, K. Chou, C. Axline, M. Reagor, J. Blumoff, and K. M. Sliwa, *et al.*, A Schrödinger cat living in two boxes, *Science* **352**, 1087 (2016).
- [25] A third approach is to engineer a noise-biased coupler [68].
- [26] S. Lloyd and S. L. Braunstein, Quantum Computation over Continuous Variables, *Phys. Rev. Lett.* **82**, 1784 (1999).
- [27] E. Knill, R. Laflamme, and G. J. Milburn, A scheme for efficient quantum computation with linear optics, *Nature* **409**, 46 (2001).
- [28] Y. Y. Gao, B. J. Lester, K. S. Chou, L. Frunzio, M. H. Devoret, L. Jiang, S. M. Girvin, and R. J. Schoelkopf, Entanglement of bosonic modes through an engineered exchange interaction, *Nature* **566**, 509 (2019).
- [29] Y. Y. Gao, B. J. Lester, Y. Zhang, C. Wang, S. Rosenblum, L. Frunzio, L. Jiang, S. M. Girvin, and R. J. Schoelkopf, Programmable Interference between Two Microwave Quantum Memories, *Phys. Rev. X* **8**, 021073 (2018).
- [30] Y. Zhang, B. J. Lester, Y. Y. Gao, L. Jiang, R. J. Schoelkopf, and S. M. Girvin, Engineering bilinear mode coupling in circuit QED: Theory and experiment, *Phys. Rev. A* **99**, 012314 (2019).
- [31] Y. Chen, C. Neill, P. Roushan, N. Leung, M. Fang, R. Barends, J. Kelly, B. Campbell, Z. Chen, and B. Chiaro, *et al.*, Qubit Architecture with High Coherence and Fast Tunable Coupling, *Phys. Rev. Lett.* **113**, 220502 (2014).
- [32] F. Yan, P. Krantz, Y. Sung, M. Kjaergaard, D. L. Campbell, T. P. Orlando, S. Gustavsson, and W. D. Oliver,

- Tunable Coupling Scheme for Implementing High-Fidelity Two-Qubit Gates, *Phys. Rev. Appl.* **10**, 054062 (2018).
- [33] W. Pfaff, C. J. Axline, L. D. Burkhardt, U. Vool, P. Reinhold, L. Frunzio, L. Jiang, M. H. Devoret, and R. J. Schoelkopf, Controlled release of multiphoton quantum states from a microwave cavity memory, *Nat. Phys.* **13**, 882 (2017).
- [34] N. E. Frattini, U. Vool, S. Shankar, A. Narla, K. Sliwa, and M. H. Devoret, 3-wave mixing Josephson dipole element, *Appl. Phys. Lett.* **110**, 222603 (2017).
- [35] C. Zhou, P. Lu, M. Praquin, T.-C. Chien, R. Kaufman, X. Cao, M. Xia, R. Mong, W. Pfaff, and D. Pekker, *et al.*, A modular quantum computer based on a quantum state router (2021), arXiv preprint [ArXiv:2109.06848](https://arxiv.org/abs/2109.06848).
- [36] E. Zakka-Bajjani, F. Nguyen, M. Lee, L. R. Vale, R. W. Simmonds, and J. Aumentado, Quantum superposition of a single microwave photon in two different “colour” states, *Nat. Phys.* **7**, 599 (2011).
- [37] A. J. Sirois, M. A. Castellanos-Beltran, M. P. DeFeo, L. Ranzani, F. Lecocq, R. W. Simmonds, J. D. Teufel, and J. Aumentado, Coherent-state storage and retrieval between superconducting cavities using parametric frequency conversion, *Appl. Phys. Lett.* **106**, 172603 (2015).
- [38] Y. Lu, A. Maiti, J. W. Garmon, S. Ganjam, Y. Zhang, J. Claes, L. Frunzio, S. Girvin, and R. J. Schoelkopf, A high-fidelity microwave beamsplitter with a parity-protected converter (2023), arXiv preprint [ArXiv:2303.00959](https://arxiv.org/abs/2303.00959).
- [39] N. E. Frattini, V. V. Sivak, A. Lingenfelter, S. Shankar, and M. H. Devoret, Optimizing the Nonlinearity and Dissipation of a SNAIL Parametric Amplifier for Dynamic Range, *Phys. Rev. Appl.* **10**, 054020 (2018).
- [40] G. Liu, T. C. Chien, X. Cao, O. Lanes, E. Alpern, D. Pekker, and M. Hatridge, Josephson parametric converter saturation and higher order effects, *Appl. Phys. Lett.* **111**, 202603 (2017).
- [41] V. V. Sivak, N. E. Frattini, V. R. Joshi, A. Lingenfelter, S. Shankar, and M. H. Devoret, Kerr-Free Three-Wave Mixing in Superconducting Quantum Circuits, *Phys. Rev. Appl.* **11**, 054060 (2019).
- [42] We constrain that fit by probing the dc resistance of the coupler at room temperature and using the Ambegaokar-Baratoff relation [69] to infer the kinetic inductance of the SNAIL from that measurement.
- [43] N. E. Frattini, Ph.D. thesis, Graduate School of Arts and Sciences, Yale University (2021).
- [44] A. Miano, V. Joshi, G. Liu, W. Dai, P. Parakh, L. Frunzio, and M. Devoret, Hamiltonian extrema of an arbitrary flux-biased Josephson circuit (2023), arXiv preprint [ArXiv:2302.03155](https://arxiv.org/abs/2302.03155).
- [45] J. Venkatraman, X. Xiao, R. G. Cortiñas, A. Eickbusch, and M. H. Devoret, Static Effective Hamiltonian of a Rapidly Driven Nonlinear System, *Phys. Rev. Lett.* **129**, 100601 (2022).
- [46] G. J. Milburn, Quantum Optical Fredkin Gate, *Phys. Rev. Lett.* **62**, 2124 (1989).
- [47] V. Giovannetti, S. Lloyd, and L. Maccone, Quantum Random Access Memory, *Phys. Rev. Lett.* **100**, 160501 (2008).
- [48] V. Giovannetti, S. Lloyd, and L. Maccone, Architectures for a quantum random access memory, *Phys. Rev. A* **78**, 052310 (2008).
- [49] A. K. Ekert, C. M. Alves, D. K. Oi, M. Horodecki, P. Horodecki, and L. C. Kwak, Direct Estimations of Linear and Nonlinear Functionals of a Quantum State, *Phys. Rev. Lett.* **88**, 217901 (2002).
- [50] D. A. Abanin and E. Demler, Measuring Entanglement Entropy of a Generic Many-Body System with a Quantum Switch, *Phys. Rev. Lett.* **109**, 020504 (2012).
- [51] C.-H. Nguyen, K.-W. Tseng, G. Maslennikov, H. C. J. Gan, and D. Matuskevich, Experimental SWAP test of infinite dimensional quantum states (2021), arXiv preprint [ArXiv:2103.10219](https://arxiv.org/abs/2103.10219).
- [52] J. Carrasco, A. Elben, C. Kokail, B. Kraus, and P. Zoller, Theoretical and Experimental Perspectives of Quantum Verification, *PRX Quantum* **2**, 010102 (2021).
- [53] R. B. Patel, J. Ho, F. Ferreyrol, T. C. Ralph, and G. J. Pryde, A quantum Fredkin gate, *Sci. Adv.* **2**, e1501531 (2016).
- [54] J. Schwinger, On angular momentum (1952), reprinted in *Quantum Theory of Angular Momentum*, ed. L. C. Biedenharm and H. Van Dam (Academic Press, New York, 1965).
- [55] T. Tsunoda, J. D. Teoh, W. D. Kalfus, S. J. de Graaf, B. J. Chapman, J. C. Curtis, N. Thakur, S. M. Girvin, and R. J. Schoelkopf, Error-detectable bosonic entangling gates with a noisy ancilla (2022).
- [56] B. Vlastakis, A. Petrenko, N. Ofek, L. Sun, Z. Leghtas, K. Sliwa, Y. Liu, M. Hatridge, J. Blumoff, and L. Frunzio, *et al.*, Characterizing entanglement of an artificial atom and a cavity cat state with Bell’s inequality, *Nat. Commun.* **6**, 8970 (2015).
- [57] N. Khaneja, T. Reiss, C. Kehlet, T. Schulte-Herbrüggen, and S. J. Glaser, Optimal control of coupled spin dynamics: Design of NMR pulse sequences by gradient ascent algorithms, *J. Magn. Reson.* **172**, 296 (2005).
- [58] H. Geen and R. Freeman, Band-selective radiofrequency pulses, *J. Magn. Reson.* **93**, 93 (1991).
- [59] J. D. Teoh, P. Winkel, H. K. Babla, B. J. Chapman, J. Claes, S. J. de Graaf, J. W. O. Garmon, W. D. Kalfus, Y. Lu, A. Maiti, K. Sahay, N. Thakur, T. Tsunoda, S. H. Xue, L. Frunzio, S. M. Girvin, S. Puri, and R. J. Schoelkopf, Dual-rail encoding with superconducting cavities (2022), [ArXiv:2212.12077](https://arxiv.org/abs/2212.12077).
- [60] M. Tinkham, *Introduction to Superconductivity* (Courier Corporation, Chelmsford, MA, 2004).
- [61] J. E. Zimmerman, Sensitivity enhancement of superconducting quantum interference devices through the use of fractional-turn loops, *J. Appl. Phys.* **42**, 4483 (1971).
- [62] J. M. Martinis, S. Nam, J. Aumentado, K. M. Lang, and C. Urbina, Decoherence of a superconducting qubit due to bias noise, *Phys. Rev. B* **67**, 094510 (2003).
- [63] X. You, J. A. Sauls, and J. Koch, Circuit quantization in the presence of time-dependent external flux, *Phys. Rev. B* **99**, 174512 (2019).
- [64] V. V. Sivak, S. Shankar, G. Liu, J. Aumentado, and M. H. Devoret, Josephson Array-Mode Parametric Amplifier, *Phys. Rev. Appl.* **13**, 024014 (2020).
- [65] Z. Leghtas, S. Touzard, I. M. Pop, A. Kou, B. Vlastakis, A. Petrenko, K. M. Sliwa, A. Narla, S. Shankar, and M. J. Hatridge, *et al.*, Confining the state of light to a quantum manifold by engineered two-photon loss, *Science* **347**, 853 (2015).

- [66] Z. K. Mineev, Z. Leghtas, S. O. Mundhada, L. Christakis, I. M. Pop, and M. H. Devoret, Energy-participation quantization of Josephson circuits, *npj Quantum Inf.* **7**, 1 (2021).
- [67] M. Reagor, W. Pfaff, C. Axline, R. W. Heeres, N. Ofek, K. Sliwa, E. Holland, C. Wang, J. Blumoff, K. Chou, M. J. Hatridge, L. Frunzio, M. H. Devoret, L. Jiang, and R. J. Schoelkopf, Quantum memory with millisecond coherence in circuit QED, *Phys. Rev. B* **94**, 014506 (2016).
- [68] I. Pietikäinen, O. Černotík, S. Puri, R. Filip, and S. M. Girvin, Ancilla-error-transparent controlled beam splitter gate (2021), arXiv preprint [ArXiv:2112.04375](https://arxiv.org/abs/2112.04375).
- [69] V. Ambegaokar and A. Baratoff, Tunneling between Superconductors, *Phys. Rev. Lett.* **10**, 486 (1963).