nature portfolio

https://doi.org/10.1038/s41567-022-01630-y

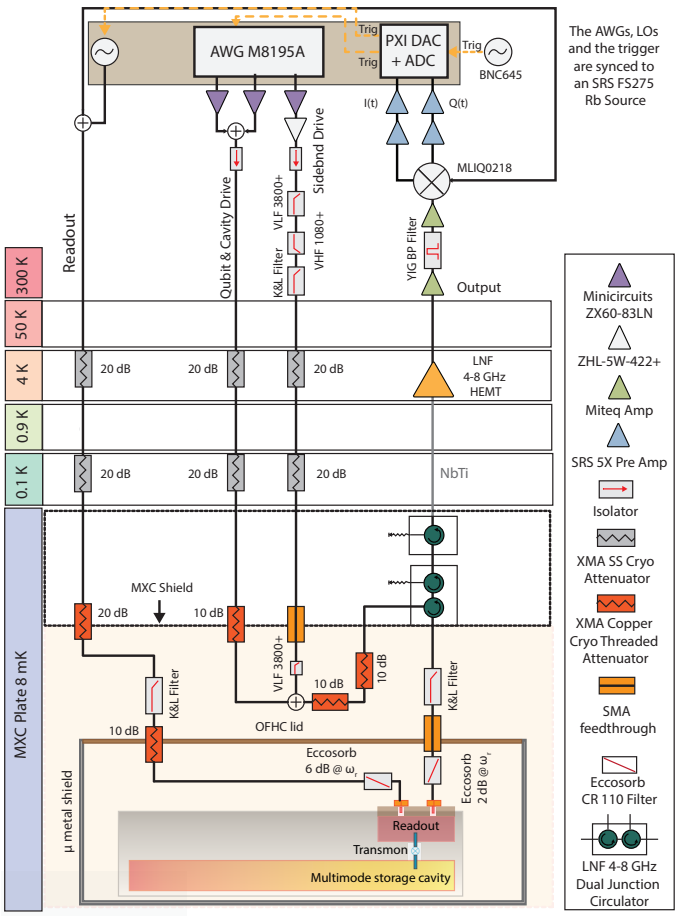
Supplementary information

Multimode photon blockade

In the format provided by the authors and unedited

Supplementary Information: Multimode photon blockade

I. CRYOGENIC SETUP AND CONTROL INSTRUMENTATION



Supplementary Figure 1 | Schematic of the cryogenic setup, microwave wiring and filtering, and control instrumentation.

The multimode cavity device is heat sunk to an OFHC copper plate connected to the base stage of a Bluefors LD-400 dilution refrigerator (7-8 mK). The sample is surrounded by a can containing two layers of μ-metal shielding, with the inside of the inner layer connected to a can made out of copper shim that is painted on the inside with Berkeley black and attached to the copper can lid. A schematic of the cryogenic setup, control instrumentation, and device wiring is shown in SFig. 1. The device is machined from a single piece of 5N5 aluminium and consists of a readout cavity and a multimode storage cavity fabricated using the flute method described in [1]. The cavities are bridged by a 3D transmon circuit whose fabrication is detailed in the next section. All controls are performed through the readout cavity, by driving at the qubit and storage mode frequencies. The pulses are directly digitally synthesized using a four-channel, 64 GSa/s arbitrary waveform generator (Keysight M8195A). The combined signals are sent to the device after being attenuated at each of the thermal stages, as shown in SFig. 1. The transmitted signal from the readout resonator passes through three cryogenic circulators (thermalized at the base stage) and is amplified using a HEMT amplifier (anchored at 4 K). Outside the fridge, the signal is filtered (tunable narrow band YIG filter with a bandwidth of 80 MHz) and further amplified. The amplitude and phase of the resonator transmission signal are obtained through a homodyne measurement, with the transmitted signal demodulated using an IQ mixer and a local oscillator at the readout resonator frequency. The homodyne signal is amplified (SRS preamplifier) and recorded using a fast ADC card (Keysight M3102A PXIe 500 MSa/s digitizer).

II. FABRICATION OF THE TRANSMON CIRUIT

The transmon qubit was fabricated on a 430 µm thick C-plane (0001) sapphire wafer with a diameter of 50.8 mm. The wafer was cleaned with organic solvents (Toluene, Acetone, Methanol, Isopropanol, and DI water) in an ultrasonic bath to remove contamination, then annealed at 1200 °C for 1.5 hours. Prior to film deposition, the wafer underwent a second clean with organic solvents (Toluene, Acetone, Methanol, Isopropanol, and DI water) in an ultrasonic bath. The junction was made out of aluminum using a combination of optical and electron-beam lithography. The base layer of the device, which includes the capacitor pads for the transmon, consists of 120 nm of Al deposited via electron-beam evaporation at 1Å/s. The features were defined via optical lithography using AZ MiR 703 photoresist and exposure by a Heidelberg MLA150 Direct Writer. The resist was developed for 1 minute in AZ MIF 300 1:1. The features were etched in a Plasma-Therm inductively coupled plasma (ICP) etcher using chlorine based etch chemistry (30 sccm Cl_2, 30 sccm BCl_2, 10 sccm Ar). This was followed by a second layer of optical patterning and thermal evaporation of 50 nm of Au for the alignment marks used for ebeam lithography. The resist was subsequently removed by leaving the wafer in 80°C N-Methyl-2-pyrrolidone (NMP) for 4 hours. The junction mask was defined through electron-beam lithography of a bi-layer resist (MMA-PMMA) in the Manhattan pattern using a Raith EBPG5000 Plus E-Beam Writer, with overlap pads that allow for direct galvanic contact to the optically defined capacitors. The resist stack was developed for 1.5 minutes in a solution of 3 parts IPA and 1 part DI water. Before deposition, the overlap regions on the pre-deposited capacitors were milled *in-situ* with

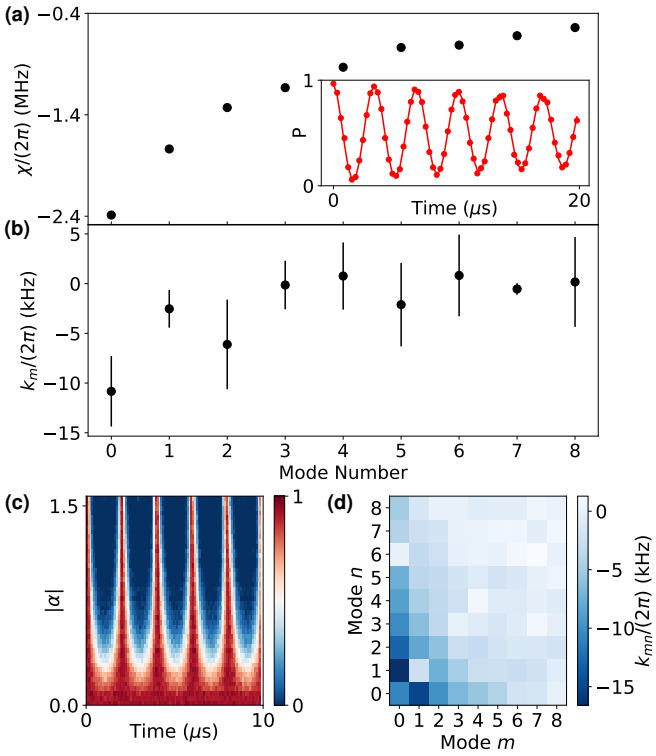
an argon ion mill to remove the native oxide. The junction was then deposited with a three step electron-beam evaporation and oxidation process. First, an initial 35 nm layer of Al was deposited at 1 nm/s at an angle of 29° relative to the normal of the substrate, azimuthally parallel to one of the fingers in the Manhattan pattern. Next, the junction was exposed to 20 mBar of a high-purity mixture of Ar and O_2 (80:20 ratio) for 12 minutes to allow the first layer to grow a native oxide. Finally, a second 120 nm layer of Al was deposited at 1 nm/s at the same 29° angle relative to the normal of the substrate, but azimuthally orthogonal to the first layer of Al. After evaporation, the remaining resist was removed via liftoff in 80°C NMP for 3 hours, leaving only the junction directly connected to the base layer. After both the evaporation and liftoff, the device was exposed to an ion-producing fan for 30 minutes to avoid electrostatic discharge of the junction.

III. CALIBRATION OF THE MULTIMODE HAMILTONIAN

The Hamiltonian of the multimode cavity QED system realized by the transmon and the storage modes is:

$$H = \omega_q |e\rangle\langle e| + \sum_{m=0}^{N-1} \{\omega_m a_m^\dagger a_m + \chi_m a_m^\dagger a_m |e\rangle\langle e| + \frac{k_m}{2} a_m^\dagger a_m (a_m^\dagger a_m - 1)\} + \sum_{n\neq m} k_{mn} a_m^\dagger a_m a_n^\dagger a_n, \quad (1)$$

where ω_q is the frequency of the transmon $|g\rangle$-$|e\rangle$ transition, ω_m the memory mode frequencies, χ_m the dispersive shifts, k_m the self-Kerr shift of each mode, and k_{mn} the cross-Kerr interactions between the modes. The value of ω_q is accurately obtained through a standard Ramsey measurement on the transmon. The χ_m are initially calibrated by performing qubit spectroscopy with a resolved pulse swept near the qubit frequency, following a coherent drive at the cavity frequency. The χ_m are then determined more precisely with a Ramsey experiment on the transmon $|g\rangle$-$|e\rangle$ transition after the addition of a photon in the cavity mode, as shown in SFig. 2(a). The photon is added to the cavity either by initializing the transmon in $|f\rangle$ and then driving the $|f0\rangle - |g1_m\rangle$ transition, or by performing a Rabi oscillation on the cavity in the presence of a blockade at $|2\rangle$, as described in the main text. The self-Kerr shifts k_m of the cavity modes are obtained by performing a cavity Ramsey experiment, with the measured values shown in SFig 2(b). This experiment is conducted by varying the time (τ) between two coherent cavity pulses (with the the phase of the second cavity pulse advanced by $2\pi\nu_R\tau$) and subsequently measuring the population in $|0\rangle$ using a resolved transmon π pulse. The magnitude of the coherent state α injected in the cavity is also swept, and the resulting data is fit to the expression:



Supplementary Figure 2 | Calibrations of multimode cavity dispersive shift and self-Kerr interactions. (a) Dispersive shift calibration for each of the manipulable modes. The measurement is performed by placing a photon in a mode, followed by a qubit Ramsey and fitting to the resulting oscillation frequency. An example is shown in the inset. (b) Self-Kerr calibration for each of the modes. The measurement is performed through cavity Ramsey and fitting to the resulting spectrum vs. time and the magnitude $|\alpha|$ of the cavity displacement. (c) Self-kerr data for cavity mode number 3. (d) Cross-shift between mode pairs. The measurement is performed by placing a photon in mode m, then sweeping the cavity frequency when probing the 0 to 1 photon peak of mode n.

$P_0(t) = |\exp(-\alpha^2)\sum_n \frac{1}{n!}\alpha^{2n}\exp(-itn(\omega_m + k_m n/2))|^2$, as shown in SFig. 2(c) for cavity mode 3. The cross-Kerrs k_{mn} are obtained by adding a photon to mode m and performing cavity spectroscopy on a different mode n. This procedure is also repeated, on the same mode to verify the consistency of the self-Kerr shifts. The values of k_{mn} are shown in SFig. 2(d). A summary of all the measured parameters of the Hamiltonian, as well as Liouvillian terms corresponding to transmon and cavity decoherence and decay, is provided in STable 1.

The minimal description of the dynamics during the blockade of a single mode includes the dispersive coupling between the transmon and the cavity, a Rabi drive on the transmon $|g\rangle$-$|e\rangle$ transition, the self-Kerr of the mode, and the cavity drive. The corresponding Hamiltonian in the frame rotating at the dressed mode and transmon

Parameter	Hamiltonian/Liouvillian Term	Quantity	Value
Transmon frequency	$\omega_q \left\|e\right\rangle \left\langle e\right\|$	$\omega_q/(2\pi)$	4.99 GHz
Storage cavity frequencies	$\omega_m a_m^\dagger a_m$	$\omega_m/(2\pi)$	see SFig. 1
Readout frequency	$\omega_r a_r^\dagger a_r$	$\omega_r/(2\pi)$	7.79 GHz
Readout dispersive shift	$\chi_r a_r^\dagger a_r \left\|e\right\rangle \left\langle e\right\|$	$\chi_r/(2\pi)$	1 MHz
Storage mode dispersive shifts	$\chi_m a_m^\dagger a_m \left\|e\right\rangle \left\langle e\right\|$	$\chi_m/(2\pi)$	see SFig. 2
Storage mode self-Kerrs	$\frac{k_m}{2} a_m^\dagger a_m (a_m^\dagger a_m - 1)$	$k_m/(2\pi)$	"
Storage mode cross-Kerrs	$k_{mn} a_m^\dagger a_m a_n^\dagger a_n$	$k_{mn}/(2\pi)$	"
Transmon $\left\|e\right\rangle \rightarrow \left\|g\right\rangle$ relaxation	$\frac{1}{T_1^q}(1+\bar{n})\mathcal{D}\left[\left\|g\right\rangle \left\langle e\right\|\right]$	T_1^q	86 ± 6 μs
Transmon $\left\|g\right\rangle - \left\|e\right\rangle$ dephasing	$(\frac{1}{T_2^q} - \frac{1}{2T_1^q})\mathcal{D}\left[\left\|e\right\rangle \left\langle e\right\|\right]$	T_2^q	58 ± 4 μs
Readout linewidth	$\kappa_r \mathcal{D}[a_r]$	$\kappa_r/(2\pi)$	0.52 MHz
Storage mode relaxation	$\frac{1}{T_1^m}\mathcal{D}[a]$	T_1^m	~ 2 ms, see [1]
Transmon thermal population	$\frac{\bar{n}}{T_1^q}\mathcal{D}\left[\left\|e\right\rangle \left\langle g\right\|\right]$	$\bar{n}$	1.2 ± 0.5 %
Storage mode dephasing	"	T_2^m	$\sim 2-3$ ms, see [1]

Supplementary Table 1 | Multimode cQED system parameters

frequencies is:

$$\hat{H} = \chi \hat{a}^\dagger \hat{a} \left|e\right\rangle \left\langle e\right| + \frac{\kappa}{2}\hat{a}^\dagger \hat{a}\left(\hat{a}^\dagger \hat{a} - 1\right) + \left\{\Omega(t)\left|g\right\rangle \left\langle e\right| + \epsilon(t)\hat{a} + \text{c.c.}\right\}. \tag{2}$$

To blockade the $\left|n_0\right\rangle$ Fock state of a single mode, the transmon is driven at frequency $\omega_q + \chi n_0$. The blockade drive can thus be expressed as $\Omega(t) = \tilde{\Omega} e^{-i\chi n_0 t}$. We make the blockade drive term static through the frame transformation $\hat{U} = e^{-i\chi \left|e\right\rangle \left\langle e\right| n_0 t}$, resulting in:

$$\tilde{\hat{H}} = \chi\left(\hat{a}^\dagger \hat{a} - n_0\right)\left|e\right\rangle \left\langle e\right| + \frac{\kappa}{2}\hat{a}^\dagger \hat{a}\left(\hat{a}^\dagger \hat{a} - 1\right) + \left\{\Omega \left|g\right\rangle \left\langle e\right| + \epsilon(t)\hat{a} + \text{c.c.}\right\}. \tag{3}$$

The blockade is valid in the regime that $\epsilon\sqrt{n_0} < \Omega < \chi$. The first of these conditions prevents leakage to $\left|\widetilde{g, n_0}\right\rangle, \left|\widetilde{e, n_0}\right\rangle$, while the second selectively blockades only the $\left|g, n_0\right\rangle \leftrightarrow \left|e, n_0\right\rangle$ transition and minimally affects transitions corresponding to other photon numbers. This Hamiltonian can be simplified by individually diagonalizing each photon number subspace ($\left|g, n\right\rangle$, $\left|e, n\right\rangle$). The blockade drive is resonant with $\left|g, n_0\right\rangle \rightarrow \left|e, n_0\right\rangle$, splitting those levels by 2Ω and mixing them equally. For levels on either side of n_0, the dressing between the ground and excited states is proportional to $\Omega/\left(\chi\left(n - n_0\right)\right)$ to leading order in Ω/χ. The Hamiltonian can be rewritten in terms of these dressed states as:

$$\tilde{\hat{H}} - \chi\left(\hat{a}^\dagger \hat{a} - n_0\right)\left|e\right\rangle \left\langle e\right| - \frac{\kappa}{2}\hat{a}^\dagger \hat{a}\left(\hat{a}^\dagger \hat{a} - 1\right) = \sum_n \xi(t)\left(\sqrt{n+1}\left|\widetilde{g, n}\right\rangle \left\langle \widetilde{g, n+1}\right| + \text{c.c} + \ldots\right) + \sum_n \sqrt{\chi^2\left(n - n_0\right)^2/4 + \Omega^2}\left(\left|\widetilde{e, n}\right\rangle \left\langle \widetilde{e, n}\right| - \left|\widetilde{g, n}\right\rangle \left\langle \widetilde{g, n}\right|\right) \tag{4}$$

In the above, we have dropped the drive terms that couple the dressed ground and excited states, which are off-resonant and suppressed by Ω/χ. The physics of the blockade can be approximated within a truncated Hilbert space that involves only the dressed transmon ground state, described by the following Hamiltonian:

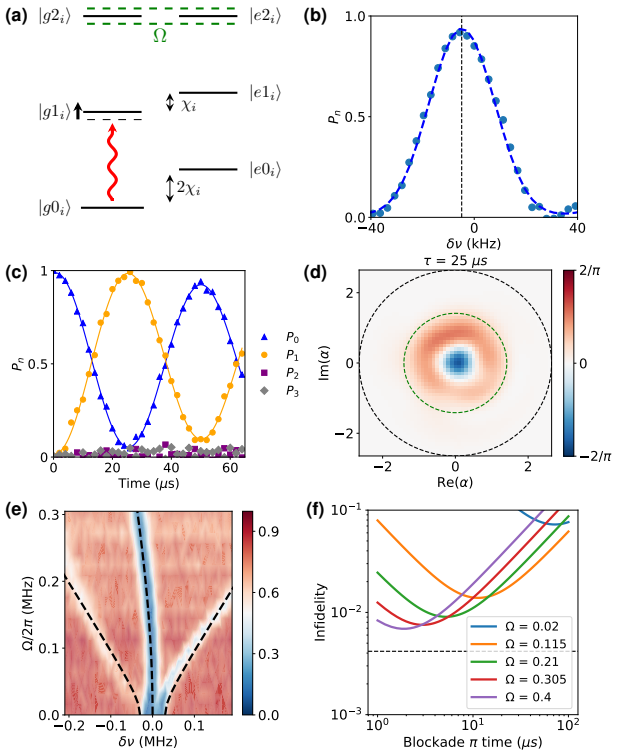
$$\begin{aligned} H &\approx \sum_n \left(\left[\frac{\kappa}{2}n(n-1) + \chi(n - n_0) - \sqrt{\chi^2\left(n - n_0\right)^2/4 + \Omega^2}\right]\left|\widetilde{g, n}\right\rangle \left\langle \widetilde{g, n}\right| + \xi(t)\left(\sqrt{n+1}\left|\widetilde{g, n}\right\rangle \left\langle \widetilde{g, n+1}\right| + \text{c.c}\right)\right) \\ &\approx \sum_n \left(\left[\frac{\kappa}{2}n(n-1) - \frac{\Omega^2}{\chi\left(n - n_0\right)}\right]\left|\widetilde{g, n}\right\rangle \left\langle \widetilde{g, n}\right| + \xi(t)\left(\sqrt{n+1}\left|\widetilde{g, n}\right\rangle \left\langle \widetilde{g, n+1}\right| + \text{c.c}\right)\right). \end{aligned} \tag{5}$$

We note that the dressed cavity Fock state energies are corrected by the self-Kerr interaction of the cavity, as well

as the photon number dependent Stark shift induced by the blockade drive.

We characterize the blockade of a single mode using the experiments shown in SFig. 3. SFig. 3(a) shows the energy level diagram describing the blockade of the $|2\rangle$ state of mode i by driving the the transmon $|g2_i\rangle - |e2_i\rangle$ transition with a Rabi strength Ω, which shifts each level by $\pm\Omega$. As is evident from the Hamiltonian SEqn. (5), the blockade drive also Stark shifts the other cavity states. We find the optimal Stark-shifted cavity drive frequency via spectroscopy in the presence of the blockade with a weak cavity drive near mode i with strength $\epsilon \ll \Omega$ as shown in SFig. 3(b). Driving weakly at this frequency in conjunction with blockade of $|g2_i\rangle$ generates high-contrast Rabi oscillations between $|g0_i\rangle$ and $|g1_i\rangle$ as shown in SFig. 3(c), which is used to add a single photon in that mode in SFig. 3(d). The complete blockade spectrum is probed using a spectroscopy experiment in which we first prepare cavity mode i in Fock-state $|1\rangle$, sweep the frequency of a weak cavity drive tone (over a larger range than in (b)), and monitor the population of $|g1_i\rangle$ as a function of the Rabi amplitude (Ω) of the blockade drive as shown in SFig. 3(e) . In addition to the peak corresponding to the Stark shifted $|0\rangle$ - $|1\rangle$ transition, we also see two Rabi split $|1\rangle$ - $|2\pm\rangle$ transitions. The Rabi amplitude Ω is calibrated separately using transmon Rabi oscillations.

As described in the main text, despite no direct occupation, the infidelity of state preparation using photon blockade is limited by decay and dephasing of the transmon in the limit of no intrinsic cavity loss. This infidelity is due to a combination of leakage (in this case to the dressed $|2\rangle$ state) and subsequent decay via the transmon ($\epsilon/(\Omega^2 T^q)$), and Purcell decay of cavity states from additional dressing due to the off-resonant blockade drive ($\Omega^2/(\epsilon\chi_m^2 T^q)$ for the $|1\rangle$ state). Here, T^q is the qubit decoherence time that is the minimum of T_2^q and T_1^q. Optimizing the cavity drive strength results in a minimum infidelity of $\sim 1/(\chi_m T^q)$ for preparing $|1\rangle$ (blockade π pulse). This infidelity is comparable to that from qubit decay and dephasing for SNAP gates and is indicated by the dashed black line in SFig. 3(f). We also show the calculated infidelity with the inclusion of intrinsic cavity decay, which is larger for state preparation via photon blockade due to the weaker cavity drives and correspondingly longer pulse times.



Supplementary Figure 3 | Blockade calibrations. (a) Energy level diagram demonstrating the physics of single-mode photon blockade in a frame rotating at the applied blockade drive frequency. (b) Cavity spectroscopy experiment in the presence of a blockade drive resonant with $|g2_i\rangle - |e2_i\rangle$ transition (depicted by the red and black arrows in (a)). The fitted center frequency is marked by the vertical dashed black line, and indicates the Stark shift of the $|g0_i\rangle - |g1_i\rangle$ transition from the bare cavity resonance frequency due to the blockade. The population of $|g1_i\rangle$ is monitored with a resolved $|g1_i\rangle \rightarrow |e1_i\rangle$ π pulse. (c) Populations over time produced by a constant cavity drive ($\epsilon/(2\pi) = 10$ kHz) at the Stark-shifted cavity resonance, in the presence of a blockade drive ($\Omega/(2\pi) = 107$ kHz) resonant with $|g2_i\rangle - |e2_i\rangle$. (d) Wigner tomography of $|1\rangle$ prepared using blockaded cavity Rabi oscillation π pulse, with a fidelity of $\mathcal{F} = 0.967 \pm 0.024$. The dashed green circle (with radius $\sqrt{2}$) indicates the boundary in phase space enforced by the blockade, and as expected all population is contained within it. (e) Cavity spectrum showing the energy splitting of the $|2\pm\rangle$ due to the blockade as well as the Stark shift of the $|0\rangle$ - $|1\rangle$ transition (central vertical blue line). The theoretical curves for the Rabi split $|1\rangle$ - $|2\pm\rangle$ transition are indicated by the dashed black lines. (f) Theoretical blockade infidelity as a function of the the time required for a blockade π pulse ($1/(2\epsilon)$), for different blockade drive strengths Ω, as a result of both transmon decoherence and cavity decay.

IV. GENERATION OF OPTIMAL CONTROL PULSES IN THE PRESENCE OF THE BLOCKADE DRIVE

Optimal control pulses were generated with the GrAPE algorithm using the package developed in [2] using two methods. The first approach used the Hamiltonian in SEqn. (3), which is in the frame rotating at the blockade frequency, using a Hilbert space with 2 transmon levels and 5-7 cavity levels,. It includes a fixed transmon blockade drive included in the drift Hamiltonian and a cavity drive term in the control Hamiltonian. The cavity drives were written as real time-dependent fields $(x(t), y(t))$ acting on the quadratures $\hat{x} = a + a^\dagger$ and $\hat{y} = -i(a - a^\dagger)$. We impose amplitude constraints on the optimal control pulses to satisfy the blockade cri-

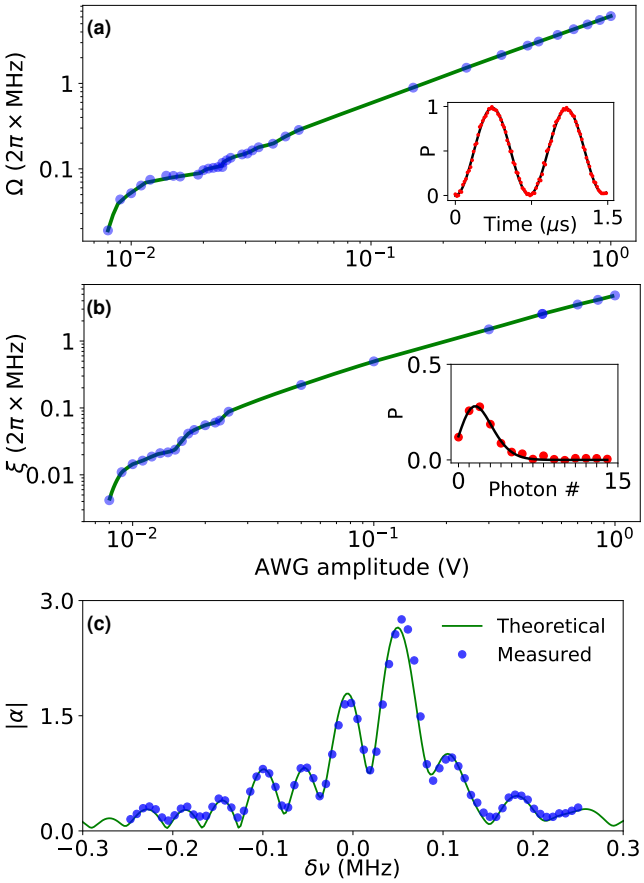
terion $\epsilon\sqrt{n_0} < \Omega < \chi$, setting a maximum allowed cavity drive amplitude of $2\pi \times (10 - 15)$ kHz $\approx \Omega/10$. We also explicitly forbid population of the dressed eigenstates at and above the blockaded level to reduce unwanted leakage.

The initial pulses produced had a bandwidth much greater than χ. While this could be decreased by adding bandwidth constraints during the optimal control pulse generation, here we filtered the pulses to a bandwidth of $\pm\chi/2$ about the cavity frequency after generation, with no detriment to the simulated or experimentally measured fidelities.

The second approach used the simplified Hamiltonian given by SEqn. (5), which includes only the cavity photon number states below the blockade level (n_0) and the cavity drive (ϵ). It correctly incorporates the Stark shifts of the cavity levels, but approximates the blockade as perfect, with leakage minimized solely by constraining the cavity drive strength. This simpler optimal control problem resulted in faster pulse convergence and had similar experimental performance for the state preparation sequences of $|1\rangle, |2\rangle$ while blockading $|3\rangle$. This improved convergence arises from not needing to manage interference effects to cancel leakage through the blockaded level. It also allowed for the implementation of a qutrit shift gate operation that simultaneously takes $|0\rangle \rightarrow |1\rangle$, $|1\rangle \rightarrow |2\rangle$, and $|2\rangle \rightarrow |0\rangle$. The pulses heuristically resulted in a gate fidelity of ~ 0.8, which was 5% worse than the simulated fidelity when including mode and transmon decoherence and decay in accordance with STable 1. The decay and decoherence times of all of the manipulable cavity modes are $\gtrsim 2$ ms [1].

A. Qubit and cavity drive calibrations

To convert between the optimal control pulse amplitudes (in frequency units) and the control voltages output by the arbitrary waveform generator (AWG), we measured transfer functions for the blockade and cavity drives. The cavity transfer function was determined by driving the target cavity mode for varying times and drive amplitudes and measuring the photon number distribution of the resulting coherent state ($|\alpha\rangle$) using resolved qubit spectroscopy. For a given cavity drive amplitude, we measured $|\alpha|$ as a function of the drive duration (τ) and extract the cavity drive strength from the slope ($|\xi| = 2|\alpha|/\tau$). This process was repeated for different cavity drive amplitudes to obtain the transfer function for the cavity drive strength versus the AWG control amplitude in SFig. 4(a). The qubit transfer function was obtained by driving the transmon $|g\rangle$-$|e\rangle$ transition at a fixed amplitude and fitting the resulting Rabi oscillation. The blockade Rabi drive strength Ω extracted as a function of the control voltage is shown in SFig. 4(b). While the transfer functions are linear at higher amplitudes, they become nonlinear at amplitudes < 25 mV due to rounding/digitization artifacts from the AWG (8 bit). When input into the experiment, the transfer function data was linearly interpolated with an odd copy reflected about the origin to handle negative drive fields produced by the optimal control. The optimal control pulses are finally shifted back to the lab frame according to $f(t) = x(t)\cos(\omega_m t) - y(t)\sin(\omega_m t)$, where ω_m is the frequency of the target mode.



Supplementary Figure 4 | Cavity and qubit drive calibrations. (a) Qubit calibration performed by fitting to Rabi oscillations as a function of AWG amplitude, with an example oscillation shown in the inset. For both the qubit and cavity, the calibration function is not strictly linear at all amplitudes due to digitization effects from the 8-bit control electronics. (b) The cavity drive calibration is performed by populating the cavity with coherent states and fitting the resulting photon number Poisson distribution, as a function of AWG amplitude. An example of a prepared state and fit is shown in the inset. The function is calibrated individually for each mode involved in our experiments. (c) Fourier transform of the optimal control pulse used to generate Fock state $|1\rangle$ in the cavity. The theoretical and experimental spectrums are in good agreement, particularly around the mode frequency where most of the contribution resides.

1. *Measuring the FFT of the optimal control pulses*

The Fourier transform of the optimal control pulse can be measured in-situ by using the cavity as a narrow band ($\sim$ 20 Hz) spectral filter. We apply the optimal control pulse to the cavity mode (with the blockade drive off) while varying the central carrier frequency, and measure the resulting photon number distribution via resolved qubit spectroscopy. Since the mode only responds on resonance, the magnitude of the resulting coherent state ($|\alpha(\omega_c)|$) as a function of the carrier frequency (ω_c) allows us to determine the FFT of the optimal control pulse. The experimentally measured $|\alpha(\omega_c)|$ and the theoretical FFT of the AWG output pulse with the calibrated transfer function are shown in SFig. 4(c).

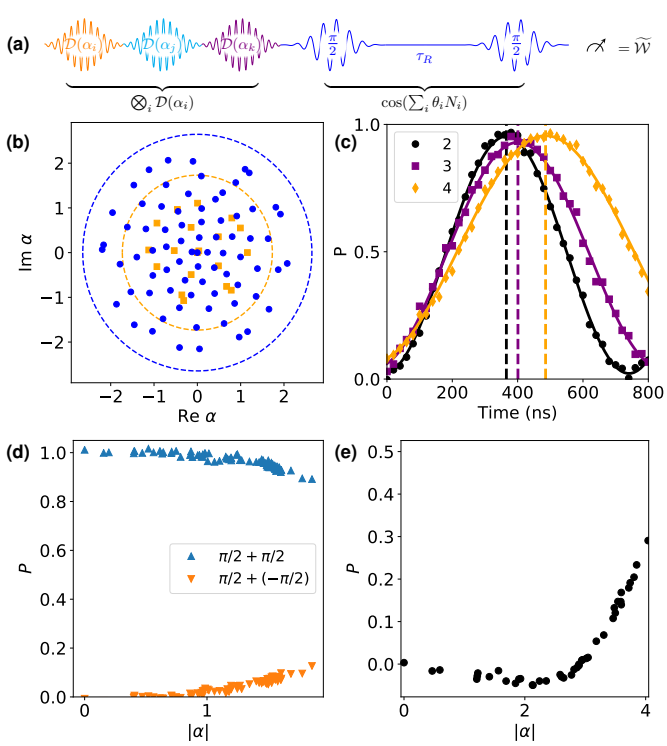
V. WIGNER TOMOGRAPHY

Single mode Wigner tomography is performed via a measurement of the photon number parity ($\hat{\Pi}$) following a series of displacements of the cavity mode. This effectively measures the Wigner operator, $\hat{\mathcal{W}}(\alpha) = \hat{\mathcal{D}}_\alpha \hat{\Pi} \hat{\mathcal{D}}_{-\alpha}$. The measurements of the Wigner operator for a set of mode displacements α_i, $x_i = \mathrm{Tr}[\hat{\mathcal{W}}(\alpha_i)\rho]$ were inverted to reconstruct the density matrix ρ following the procedure described in [3]. By converting the Wigner operator and density matrix to vectors, we express $x_i = \langle\langle \mathcal{W}(\alpha_i)|\rho\rangle\rangle$, and construct a matrix $\mathcal{M}$ with $\mathcal{M}_{ij} = \langle\langle \mathcal{W}(\alpha_i)|_j$ that represents measurements of the Wigner operator at all the displacements. The number of columns of $\mathcal{M}$ is d^2, where d is the truncated dimension of the Hilbert space of the cavity up to which the tomography is valid, while the rows correspond to the points in phase space where measurements are sampled ($> d^2$). Since $\mathcal{M}$ is a non-square matrix, we calculate the density matrix by acting the Moore-Penrose pseudoinverse of $\mathcal{M}$ on the vector of measurements $\vec{x}$, i.e., $|\rho\rangle\rangle = (\mathcal{M}^T\mathcal{M})^{-1}(\mathcal{M}^T\vec{x})$. The density matrix extracted from this inversion is made physical by forcing it to have unit trace and imposing positive semi-definiteness. We constrain the trace of the density matrix through the use a Lagrange multiplier (λ) and perform the inversion as below:

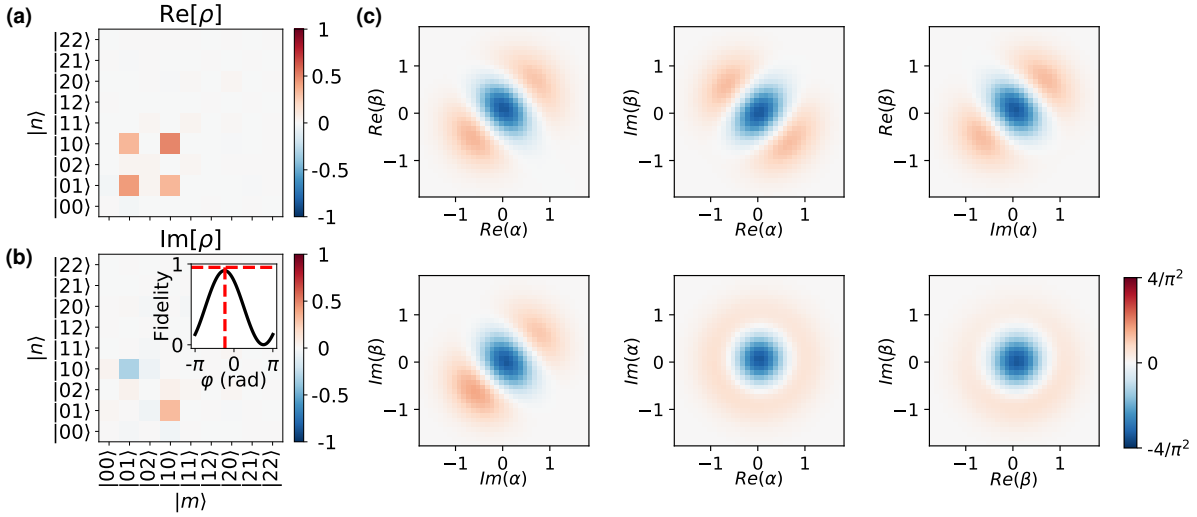
$$\begin{bmatrix} |\rho\rangle\rangle \\ \lambda \end{bmatrix} = \begin{bmatrix} \mathcal{M}^T\mathcal{M} & \langle\langle I|^T \\ \langle\langle I| & 0 \end{bmatrix}^{-1} \begin{bmatrix} \mathcal{M}^T\vec{x} \\ 1 \end{bmatrix}. \tag{6}$$

Positive semi-definiteness is usually imposed with Cholesky decomposition or other methods. We instead impose that condition using the algorithm presented in [4], involving diagonalizing $\mathcal{M}$, iteratively redistributing any negative eigenvalues equally across the remaining positive ones, and subsequently reconstructing ρ.

The displacements we perform are to an optimized set of points in phase space that are chosen with the method described in [3], and are shown in SFig. 5(a). In particular, the set of points minimizes the *condition number* κ—the absolute value of the ratio of the maximum to the minimum eigenvalue of $\mathcal{M}$. Minimizing κ increases the likelihood of $\mathcal{M}$ being invertible, and reduces amplification of error from the inverted ρ—measurement noise of magnitude ϵ results in an error no greater than $\kappa\epsilon$ in the reconstructed density matrix. For our single mode experiments, we used a total of 75 Wigner points with



Supplementary Figure 5 | Multimode Wigner tomography calibrations. (a) Pulse sequence for Wigner tomography of 3 modes. Sequential cavity displacements are followed by a qubit Ramsey measurement with wait time τ_R. (b) Set of cavity displacements used to reconstruct the cavity state with Wigner tomography. The 75 points used in the single mode case (blue) and the 18 points used for each mode in the three-mode tomography case (orange) are shown. Dashed circles indicate the square root of the maximum photon number to which the reconstruction is accurate. (c) Ramsey sequence after the addition of a single photon in each mode. Dashed vertical lines indicate the times that correspond to a perfect parity measurement for each mode. This is subsequently used to calibrate the angles θ_j that correspond to a wait time τ_R. (d) Calibrating the bandwidth of the parity measurement. The pulse sequence is a cavity displacement followed by two $\pi/2$ pulses, where the second $\pi/2$ pulse either has phase 0 (blue) or π (orange). The finite bandwidth of the qubit $\pi/2$ pulses and the increasing dispersive shift at larger $|\alpha|$ result in imperfect $\pi/2$ pulses, reducing the range of possible measurement results. (e) Calibration of the cross-Kerr between the readout resonator and a cavity mode. The pulse sequence is a cavity displacement of magnitude $|\alpha|$ followed by readout. This does not significantly affect the measurements presented in this work, since the values of $|\alpha|$ used in the Wigner tomography experiments were < 2.



Supplementary Figure 6 | Density matrix reconstruction and multimode Wigner tomography data for a two-mode W-state. (a) Real and (b) Imaginary part of the density matrix of the reconstructed two-mode state. The value of $\phi = -0.730$ for the state $(|10\rangle + e^{i\phi}|01\rangle)/\sqrt{2}$ is determined by the maximum fidelity projection, and is shown in the inset. The corresponding state fidelity is 0.918 ± 0.012. c) The 6 orthogonal 2D slices of the two-mode Wigner function through the origin of phase space, involving all combinations of the real and imaginary quadratures of both modes.

a maximum photon number of 7, with $\kappa = 1.6$. For the multimode measurements of the single-photon W-state, we used 18 Wigner points and a maximum photon number of 3 per cavity mode.

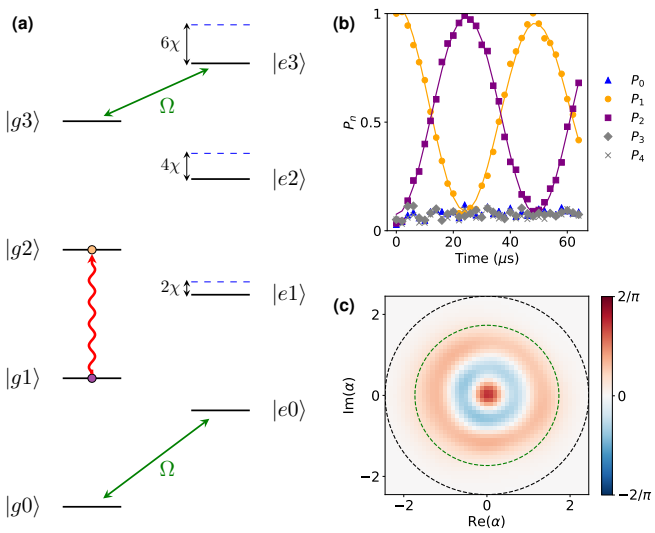
A. Generalized Wigner tomography

The parity measurement in Wigner tomography is performed using a qubit Ramsey sequence that is composed of two broadband $\pi/2$ pulses with opposite phases, separated by a wait time $\tau = 1/(2\chi_m)$ during which the qubit acquires a $\theta = \pi$ phase if a single photon is in cavity mode i, followed by qubit readout. Since rotations are only distinguishable modulo 2π, all odd (even) photon numbers will place the qubit in the excited (ground) state, resulting in a measurement of photon number parity $\hat{\Pi} = \cos(\pi\hat{N})$. A similar qubit Ramsey sequence that idles for an arbitrary τ corresponds to a phase shift of $\theta = 2\pi\chi_m\tau$ for a single photon in mode m, and a measurement of $\hat{\Theta} = \cos(\theta\hat{N})$. Displacing the cavity mode prior to this general qubit Ramsey sequence allows for the measurement of a generalized Wigner operator $\hat{\mathcal{W}}(\alpha,\theta) = \hat{\mathcal{D}}_\alpha\hat{\Theta}\hat{\mathcal{D}}_{-\alpha}$. As long as θ is known, we can invert the measurements of the expectation value of this operator for a series of known displacements to obtain the density matrix, like in the case of $\theta = \pi$. The error in reconstruction fidelity is dependent on θ, with the smallest error occurring around $\theta = \pi$—where the generalized Wigner function has maximum contrast, and is also the least sensitive to errors in the calibration of θ.

B. Multimode Wigner tomography

Multimode Wigner tomography has previously been performed via measurements of the joint photon number parity following displacements of each of the cavity modes, effectively corresponding to a measurement of a joint Wigner operator. However, joint photon number parity measurements become challenging when the modes do not have the same χ_m, requiring the use of higher transmon levels [5], or additional transmons [6]. The generalized Wigner tomography protocol described in the previous section provides a workaround, allowing us to replace the joint photon number parity operator $\bigotimes_m \hat{\Pi}_m$ with a generalized operator $\bigotimes_m \hat{\Theta}_m$. Since θ_m need not be identical between modes, a single qubit Ramsey time τ that corresponds to different θ_m for each mode m can be utilized to perform the measurement. We can then characterize our state without engineering χ_m.

In the case of two modes, we used cavity modes 3 and 4 and combined the measurements at three different Ramsey times τ_1, τ_2, and τ_3 to reconstruct the state density matrix. While only one τ is necessary, additional times improve the accuracy of the final state reconstruction. The set of τ's at which we measured was $\{\tau_j\} = [419.8, 483.3, 454.0]$ ns, which corresponds to $\{\theta_3\} = [\pi, 3.63, 3.39]$ and $\{\theta_4\} = [2.74, \pi, 2.94]$. For these sets of angles, the condition numbers are $\kappa_{3,4} = 1.6$. For the three mode case, we used modes 2,3, and 4 of our cavity, and chose the Ramsey time that corresponds to $\theta_3 = \pi$. We made this choice because χ_3 is between χ_2 and χ_4, resulting in a set of $\{\theta\}$ that are as close to π as possible. This is desirable for reasons as discussed in Sec-



Supplementary Figure 7 | State preparation with multiple blockade drives. (a) Energy level diagram showing two drives resonant with $|g0\rangle - |e0\rangle$ and $|g3\rangle - |e3\rangle$ that have Rabi strengths Ω and used to simultaneously blockade $|0\rangle$ and $|3\rangle$, respectively. (b) Populations over time produced by a constant cavity drive ($\epsilon/(2\pi) = 10$ kHz) with blockade drives ($\Omega/(2\pi) = 107$ kHz) at $|0\rangle$ and $|3\rangle$ after preparing the cavity in $|1\rangle$. (c) Wigner tomography of $|2\rangle$ as prepared in (b) corresponding to a fidelity of $\mathcal{F} = 0.925 \pm 0.008$. The dashed green circle indicates the boundary of the blockaded subspace (circle of radius $\sqrt{3}$).

tion V A. The resulting tomography angles for this single value of τ are (in radians) $\{\theta_2, \theta_3, \theta_4\} = [3.44, \pi, 2.64]$. The corresponding values of κ are $\kappa_{2,3,4} = 1.6$.

In addition to the density matrices of the multimode W-states presented in the main text, here we provide 2D slices of their multimode Wigner functions. These can be seen for the two and three mode states in SFig. 6 and 8, respectively. The slices correspond to all pairwise combinations of real and imaginary quadratures of each of the modes, leading to 6 slices in the two-mode case and 15 in the three-mode case.

C. Multimode state phase determination and gauge freedom

We determine the phases ϕ_j of our multimode W-states, $(|10\rangle + e^{i\phi}|01\rangle)/\sqrt{2}$ in the two-mode case and $(|100\rangle + e^{i\phi_1}|010\rangle + e^{i\phi_2}|001\rangle)/\sqrt{3}$ in the three-mode case, by maximizing the fidelity of the projection onto those states as a function of the ϕ_j's. That is, we map our prepared state onto the appropriate (two- or three-mode) expected W-state while sweeping the phase parameters, and pick the angles that give us the projected value closest to 1. This is shown in the insets of SFig. 6 and SFig. 8.

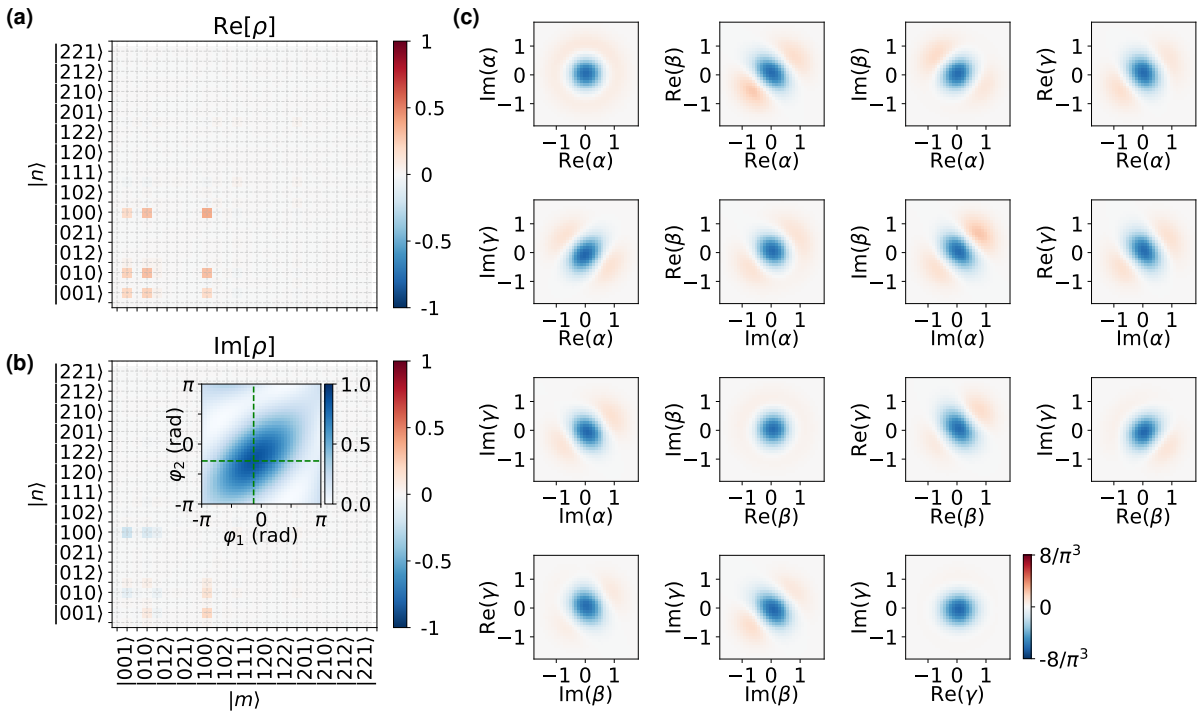
We are able to prepare states with different phases by varying the relative phases of our cavity drives. There is a 2π gauge freedom in the definition of the phase of each cavity mode. For a given choice of these phases, we determine the phase of the prepared states using Wigner tomography. These phases can be can be modified by a gauge transformation, allowing us to make the reconstructed density matrices real, as in SFig. 5 of the main text.

D. Systematic errors in Wigner tomography

In addition to experimental noise, the Wigner tomography reconstruction has systematic errors that appear in the parity measurement and come from two main sources. The first source is the limited bandwidth of the parity measurement. We mitigate this by using DRAG pulse shaping to maximize the bandwidth of the pulses (Gaussian pulse with $\sigma = 5$ ns). The second source is readout error arising from the cross-Kerr interaction between the storage and readout modes. This results in a systematic shift in the readout voltage of transmon states that depends on the number of photons in the storage modes. We calibrate both these errors using the protocols described below, and use them to correct the Wigner tomography. The correction to the bandwidth of the parity measurement is calibrated by displacing the cavity to each of the phase space points used in the Wigner tomography and subsequently applying two $\pi/2$ pulses with either the same or opposite phase, with no wait time in between. This would ideally place the qubit in either the excited or ground state, corresponding to $P_e = 1$ or 0, respectively. However, despite the large bandwidth of the qubit $\pi/2$ pulses, the dispersive shift takes the pulses off resonance for larger $|\alpha|$. As shown in SFig. 5(d), this reduces the contrast of the parity measurement and therefore the Wigner operator measurement. We compensate for this effect by scaling the Wigner operator measurement for a given state and displacement using a linear transformation $(\mathcal{W}(\alpha, \rho) \to a\mathcal{W}(\alpha, \rho) + b)$, that takes the upper and lower bounds for the parity measurement (c_1, c_2) to their ideal values $(1, 0)$, i.e. $a, b = 1/(c_1 - c_2), -c_2/(c_1 - c_2)$. This correction is performed for each Wigner point used in the tomography, with varying calibrated values of a and b. Wigner points with larger values of $|\alpha|$ deviate more from the ideal 0 to 1 range, as the magnitude of the cross-Kerr and dispersive shift effects scales with $|\alpha|$.

VI. EXTENSIONS TO MORE MODES (SITES) AND PHOTONS (BODIES)

We present experiments where we extend photon blockade beyond the discussion presented in the main text. The experiments presented there all involve using a single drive to blockade one or more levels. Using

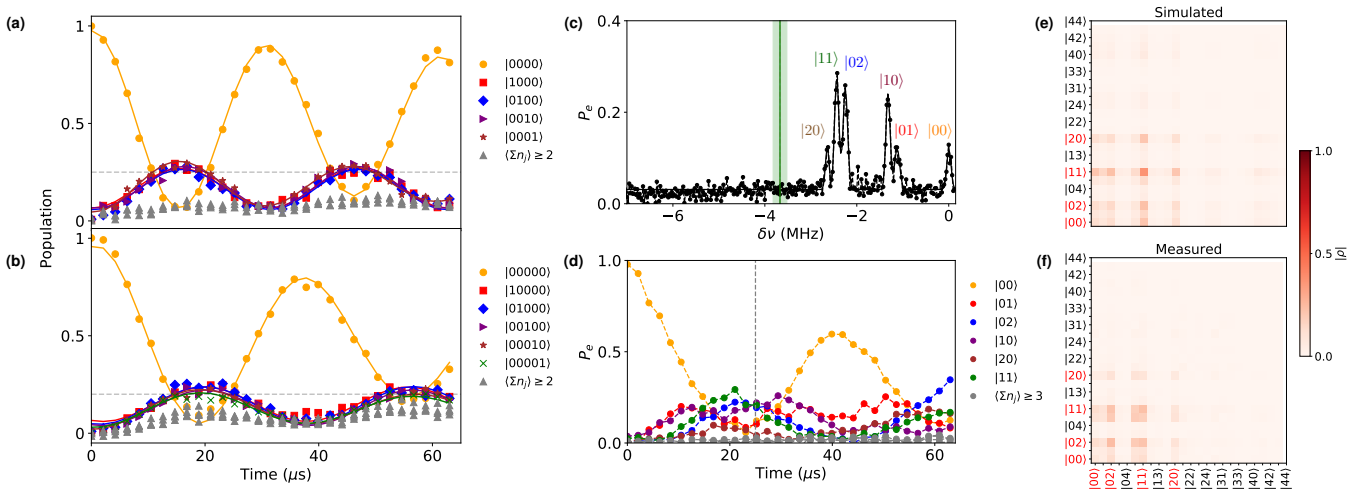


Supplementary Figure 8 | Density matrix reconstruction and multimode Wigner tomography data for a three-mode W-state. (a) Real and (b) Imaginary part of the density matrix of the reconstructed three mode state. The values of $\phi_1 = -0.403, \phi_2 = -0.866$ of the prepared state $(|100\rangle + e^{i\phi_1}|010\rangle + e^{i\phi_2}|001\rangle)/\sqrt{3}$ are again determined by the maximum fidelity projection, and shown in the inset. The corresponding state fidelity is 0.864 ± 0.014. (c) The 15 orthogonal 2D slices of the three-mode Wigner function through the origin of phase space, involving all combinations of the real and imaginary quadratures of the three modes.

multiple drive tones to simultaneously target different transitions allows for more flexibility with blockade and preparing more complex quantum states. We illustrate this on a single mode by simultaneously driving $|g0\rangle - |e0\rangle$ and $|g3\rangle - |e3\rangle$, as illustrated in the energy level diagram shown in SFig. 7(a). This results in the simultaneous blockade of $|0\rangle$ and $|3\rangle$ by shifting both of those levels off resonance from the cavity frequency, isolating $|1\rangle$ and $|2\rangle$ from the rest of the Hilbert space. As a result, initializing Fock state $|1\rangle$ and driving at the cavity frequency produces Rabi oscillations between those two levels, as shown in SFig. 7(b), which can be used to prepare the target mode in $|2\rangle$, as shown in SFig.7(c).

We also extend multimode photon blockade to more modes (sites) by applying the protocol described in the main text for preparing two- and three-mode W states (where we simultaneously blockade all multimode states with a total of 2 photons), to four and five modes. Starting from the vacuum state and driving all the cavity modes simultaneously in the presence of the blockade drive results in the correlated population oscillations shown in SFig. 9(a) and (b). As before, this corresponds to oscillations between the vacuum and four- and five-mode W states which is consistent with the measured populations. Completely characterizing the four- and five-mode W states with our multimode Wigner tomography protocol required a prohibitive number of measurements and is omitted in this work. This may be overcome in future experiments by increasing the experiment duty cycle through active reset or by implementing efficient protocols for state certification without full tomography. The phase of the W-state can also be inferred by interferometry experiments that can be implemented with beamsplitter operations between modes. From master equation simulations, we expect the fidelities of the four- and five-mode W states to be $\mathcal{F} = 0.8$ and $\mathcal{F} = 0.69$, respectively. The fidelities decrease with the addition of more modes (sites) due to the mismatch in the dispersive shifts of different modes, which leads to imperfect blockade when using a single tone. Achieving multimode photon blockade therefore requires larger blockade drive strengths, which in turn results in larger participation of the transmon in cavity states and enhanced decay. We note that these state preparation fidelities for W-states with more modes may be improved by using multiple drives to blockade different multimode states, or potentially by using off-resonant charge-sideband drives [7] to match the dispersive shifts across the modes.

Finally, we also implement a photon blockade interaction between $\mathrm{N} = 3$ photons distributed across two modes. Here, the blockade drive shifts the energy of all two-mode 3-photon states and is applied at the mean



Supplementary Figure 9 | Extensions of the blockade. Populations over time for simultaneous photon blockade of (a) four modes ($\Omega/(2\pi) = 206$ kHz) and (b) five modes ($\Omega/(2\pi) = 178$ kHz) with constant, equal cavity drive strengths ($\epsilon/(2\pi) = 9$ kHz and $\epsilon/(2\pi) = 7$ kHz, respectively). The results are oscillations between the ground state and entangled W states. Solid lines indicate the fit to exponentially decaying sinusoidal oscillations and the dashed grey lines correspond to mode populations for an ideal W state (0.25 and 0.2, respectively). The increase in population of the blockaded states is due to leakage to the two-photon manifold and subsequent decay arising from participation of the transmon in those states due to the blockade drive. (c) Spectroscopy of a state prepared with constant cavity drives while blockading all two-mode states with total of $n = 3$ photons. The x-axis is defined relative to the transmon qubit frequency. The blockade drive frequency and drive strength are indicated by the green vertical line and rectangle, respectively, and the relevant two-mode Fock state peaks are labeled. As expected, there are no visible peaks corresponding to Fock states with of 3 or more total photons. (d) Populations over time produced by constant cavity drives in the presence of a two-mode blockade of a total of N = 3 photons. The dashed gray vertical line indicates the time corresponding to the spectrum shown in (c). Solid circles indicate points at which data was collected, and the dashed lines serve as a guide to the eye and do not correspond to fits. (e) Simulated density matrix and (f) density matrix reconstruction from Wigner tomography measurements of the state produced in (c). Multimode Fock states within the blockaded subspace are indicated in red (only every other state is labeled). In both cases, the density matrix is nonzero primarily only for the Fock states in that subspace.

transmon frequency corresponding to 3 photons ($(3\chi_i + 3\chi_j)/2$), indicated by the green line in SFig. 9(c). Driving both cavity modes uniformly in the presence of the blockade results in restricted dynamics in the Hilbert space of two-mode Fock states with ≤ 2 photons, with very little leakage outside the blockaded subspace. This is evident in the qubit spectroscopy plot shown in SFig. 9(c), performed after uniform drives on both modes for 25 μs (dashed vertical line in (d)), in the presence of the 3-photon blockade drive. The dynamics are more complicated compared to the blockade of N = 2 photons, and there is no longer a time when the population fully returns to the vacuum state, as shown in SFig. 9(d). The restriction of population and coherences to within the blockaded subspace is also evident in the density matrix, shown in (f), reconstructed from Wigner tomography measurements. These results are qualitatively similar to those found in the corresponding master-equation simulations, shown in (e). These experiments show that multimode photon blockade can be implemented at higher photon numbers, and demonstrate a 3-photon (3-body) interaction between photons in two modes (sites).

VII. MASTER EQUATION SIMULATIONS OF BLOCKADE DYNAMICS

We use QuTip [8] to simulate multimode blockade dynamics using a master equation that includes the decay of the cavity modes ($\kappa_m = 1/T_1^m$), as well as the decay ($\gamma_q = 1/T_1^q$) and dephasing ($\gamma_\phi^q = 1/T_2^q - 1/(2T_1^q)$) of the transmon:

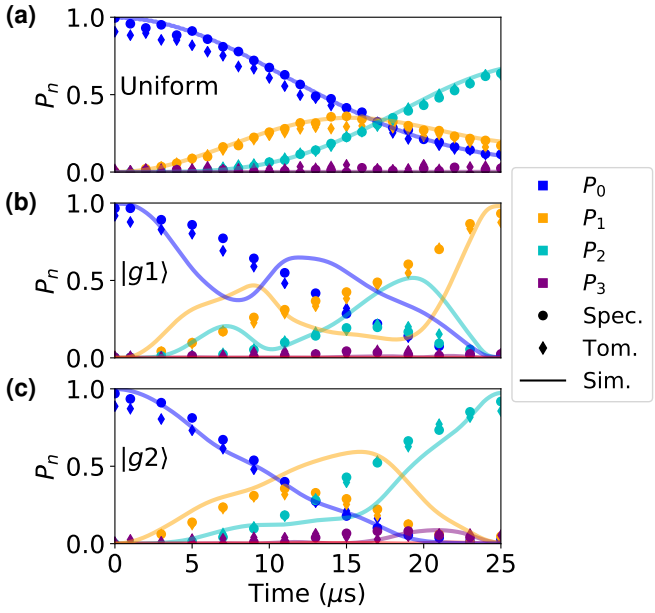
$$\dot{\hat{\rho}} = -i[\hat{H}, \hat{\rho}] + \sum_m \kappa_m \mathcal{D}[\hat{a}_m]\hat{\rho} + \gamma_q n_q^{\text{th}} \mathcal{D}[|e\rangle\langle g|]\hat{\rho} + \gamma_q(1 + n_q^{\text{th}})\mathcal{D}[|g\rangle\langle e|]\hat{\rho} + \gamma_q^\phi \mathcal{D}[|e\rangle\langle e|]\hat{\rho}. \quad (7)$$

Here, $\mathcal{D}$ is the Lindblad dissipator, and $\hat{H}$ is the blockade Hamiltonian given by SEqn. (3) for the single-mode case. We include the thermal occupation of the transmon ($n_q^{\text{th}} = 1.2 \pm 0.5\%$), but ignore the thermal population of the storage cavity modes ($n_m^{\text{th}} \leq 0.01\%$). For the case of multiple cavity modes and drives, this Hamiltonian—written in a frame co-rotating with the blockade ($\nu_b = \nu_q + \delta\nu_b$) and cavity mode frequencies

(ν_m), generalizes to:

$$\hat{H} = \left\{ \sum_m \chi_m \hat{a}_m^\dagger \hat{a}_m - \delta\nu_b \right\} |e\rangle\langle e| + \sum_m \frac{k_m}{2} \hat{a}_m^\dagger \hat{a}_m \left(\hat{a}_m^\dagger \hat{a}_m - 1 \right) + \sum_{m \neq n} \frac{k_{mn}}{2} \hat{a}_m^\dagger \hat{a}_m \hat{a}_n^\dagger \hat{a}_n + \{ \Omega |g\rangle\langle e| + \sum_m \epsilon_m(t) \hat{a}_m + \text{c.c.} \}, \tag{8}$$

where Ω is the blockade Rabi frequency, k_m, k_{mn} the self and cross-Kerr interactions, and $\epsilon_m(t)$ the time dependent cavity drive amplitudes. All the drive tones are sent through the readout port and are directly coupled only to the readout resonator ($\hat{H}_d = \epsilon_r \cos(\omega_d t)(\hat{a}_r + \hat{a}_r^\dagger)$). However, the coupling between the transmon and the modes and their resulting dressing leads to effective transmon and storage mode drives when the readout is driven on resonance with either of them. To lowest order in the dispersive approximation, the resulting transmon and storage mode drives are $\Omega \approx \epsilon_r g_r/(2\Delta_r)$ and $\epsilon_m \approx \epsilon_r g_r g_m/(2\Delta_r \Delta_m)$, respectively. For generating the optimal control pulses, we treat the drives as being directly on the transmon and the storage modes. This approximation is valid because the detuning between the transmon and the readout (Δ_r) and storage modes (Δ_m) is large compared to the coupling strengths (g_r, g_m), which we additionally verify using master equation simulations that include the readout cavity/drive.



Supplementary Figure 10 | Simulations of single-mode optimal control. State evolution vs. time produced by a uniform cavity drive (a) and optimal control pulses that prepare $|g1\rangle$ (b) or $|g2\rangle$ (c), in the presence of a blockade drive at $n = 3$. Populations were measured directly using qubit spectroscopy (circles), and were also extracted from the density matrices obtained from Wigner tomography (diamonds). The results of the master equation simulations are represented by the solid lines.

A. Single-mode optimal control pulses

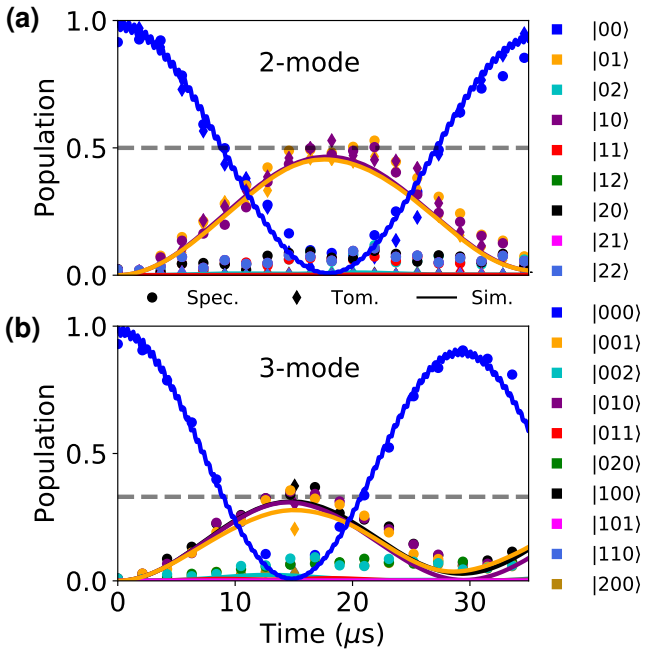
We first use the master equation simulations to obtain the expected fidelity in experiments involving the blockade of a single cavity mode, corresponding to the data presented in Fig. 3 of the main text. An analysis of the population evolution generated by a uniform cavity drive, and the optimal control pulses used to prepare $|1\rangle$ and $|2\rangle$ while blockading $|3\rangle$, is shown in SFig. 10.

The populations are extracted from the density matrices reconstructed from Wigner tomography (diamonds), as well as through number resolved qubit spectroscopy (circles), from the raw data presented in Fig. 3 of the main text. We note that we account for the measurement error arising from the decay during the resolved qubit pulse used for the spectroscopy—a Gaussian pulse with $\sigma = 0.9$ μs (duration $= 4\sigma$), by normalizing by the height of the $|0\rangle$ peak obtained from spectroscopy of the vacuum state. No such normalization is performed for the density matrices obtained from the Wigner tomography, which accounts for the slightly lower extracted populations. Apart from this difference ($\sim 5\%$), the populations extracted using both these methods are consistent.

The master equation simulations of a uniform cavity drive in the presence of a $|g3_i\rangle \leftrightarrow |e3_i\rangle$ blockade drive match well with the experimentally measured populations. We varied the Hamiltonian parameters relevant for the blockade $(\chi, k, \Omega, \delta)$ in the simulations, and found that the independently calibrated parameter values also produced the best overlap with the measured state. Despite this, the population trajectories measured by applying slices of the optimal control pulses that prepared $|1\rangle$ and $|2\rangle$ at varying times did not perfectly match with the simulations. Given the close match between the experimental and simulated trajectories for a uniform pulse, the discrepancy in the trajectory is believed to be due to distortions of the optimal control pulse from impedance mismatches along the drive line before reaching the device. While the trajectories themselves deviate, the final states prepared by the optimal control pulses still result in fidelities of $\mathcal{F} = 0.953 \pm 0.022$ ($|1_4\rangle$) and 0.965 ± 0.022 ($|2_4\rangle$), compared to simulated fidelities of 0.981 and 0.974, respectively.

B. Simulations of multimode N-body interactions

We simulate the dynamics arising from the multimode blockade interactions by using the Hamiltonian in SEqn. 8 in the master equation in SEqn. 7). The



Supplementary Figure 11 | Simulations of two- and three-mode W-states. Populations of different multimode states in the presence of uniform cavity drives and a blockade drive at the mean of the frequencies corresponding to two photons for (a) two (ν_3, ν_4; $\Omega/(2\pi) = 207$ kHz) and (b) three cavity modes (ν_2, ν_3, ν_4; $\Omega/(2\pi) = 227$ kHz). For two cavity modes, in addition to monitoring populations through qubit spectroscopy (circles), we extract populations from the reconstructed density matrices obtained from Wigner tomography (diamonds). For the 3 mode case, this comparison to tomography is only made for a pulse corresponding to the preparation of the three-mode W-state ($\tau = 15\mu s$). In both cases, the cavity drive strength on each mode was $\epsilon/(2\pi) = 10$ kHz. The populations obtained through master equation simulations in the presence of transmon decoherence and cavity decay are represented by the solid lines.

Hamiltonian—co-rotating at the blockade and cavity frequency, is valid for a single blockade drive frequency, as is used in all the experiments presented in this work.

We prepare two- and three-mode W-states by using a blockade drive detuned from the qubit frequency by the average of the dispersive shifts from adding 2 photons in any combination of modes. For the two-mode case, we study the temporal evolution arising from uniform (and equal strength) cavity drives on both modes (3, 4) in the presence of the blockade drive using photon number resolved qubit spectroscopy and two-mode Wigner tomography. The extracted populations in the different multimode Fock states are shown as a function of the drive duration in SFig. 11(a). At a time $\tau = 18.7\ \mu s$, this produces an entangled two-mode W-state, $|\psi\rangle = (|10\rangle + e^{i\phi}|01\rangle)/\sqrt{2}$. The measured fidelity from Wigner tomography ($\mathcal{F} = \mathrm{Tr}[\rho_W \rho] = 0.918 \pm 0.012$), and that obtained from master equation simulations (0.919) were consistent.

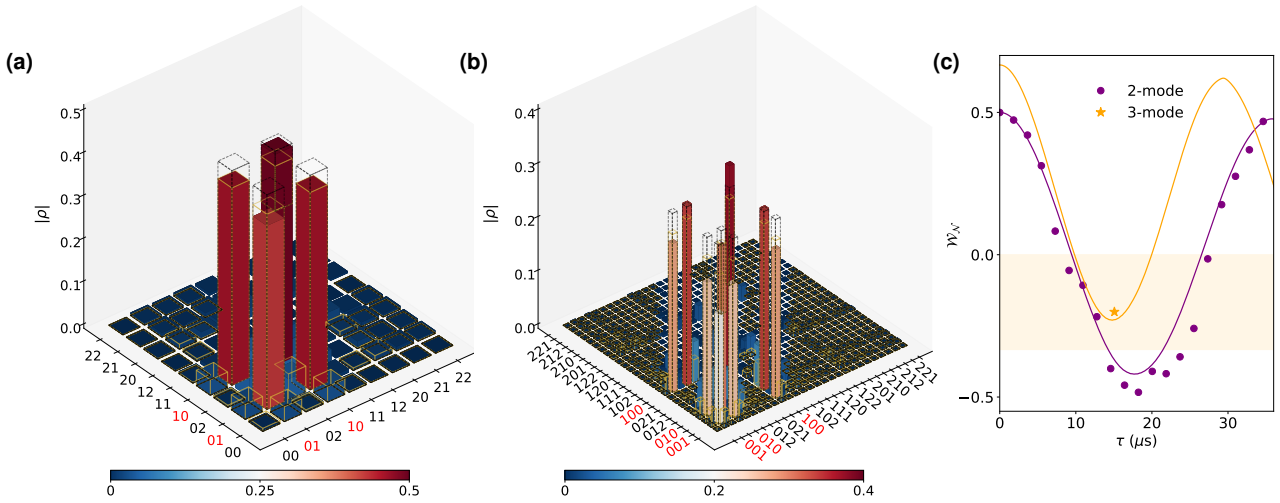
A similar comparison between the experiment and master equation simulations for the three-mode W-state preparation sequence is shown in SFig. 11(b). Here, the populations are measured as a function of the cavity drive duration using photon number resolved qubit spectroscopy. For both this experiment and the two-mode case, we used a longer resolved qubit π pulse (Gaussian pulse with $\sigma \sim 3\ \mu s$) than in the single mode case, in order to accurately resolve the differences between the dispersive shifts.

At the cavity drive duration that corresponds to the W-state, we reconstruct the state using three-mode Wigner tomography, resulting in the density matrices and Wigner functions shown in SFig. 8. The simulated fidelity of the three-mode W-state was $\mathcal{F} = 0.896$, compared to the experimental measured value of 0.864 ± 0.014. A comparison of the simulated, measured, and ideal state populations that produce these fidelities is shown in SFig. 12 for both the two- and three-mode W-states. The simulations (yellow edge boxes) include loss and decoherence effects from the transmon and cavity, as well as the qubit temperature. The simulated and measured data are generally in good agreement. The ideal W-states (black dashed edge boxes) are included to serve as a guide. In order for the blockade drive to simultaneously address the dispersively shifted peaks corresponding to two photons in any combination of modes, we pick a blockade Rabi strength ($\Omega/(2\pi) = 227$ kHz), which is roughly twice that used in the single-mode blockade experiments, resulting in 5% higher participation of the transmon in the cavity levels from off-resonant dressing.

From the reconstructed density matrices for two- and three-mode W-state preparation shown in SFig. 12(a) and (b), we extract the W-state entanglement witness [9, 10],

$$\hat{\mathcal{W}}_{\mathcal{N}} = \frac{N-1}{N} - |W_N\rangle\langle W_N| \Rightarrow \left\langle \hat{\mathcal{W}}_N \right\rangle = \frac{N-1}{N} - \mathcal{F}. \quad (9)$$

where N is the number of entangled modes. For each measured density matrix ρ, we extract the witness by sweeping the free phases that characterize the W-state to maximize the state fidelity $\mathcal{F} = \mathrm{Tr}[\rho\rho_W]$. The results of these measurements are presented in SFig. 12(c). In the two-mode case, we evaluate the witness as a function of the duration of the cavity drive, while in the three-mode case, we evaluate it at the single time corresponding to the creation of the W-state. We note that $-1/3 < \left\langle \hat{\mathcal{W}}_3 \right\rangle = -0.2 < 0$ indicates genuine tripartite entanglement for the three-mode W-state (the orange band in SFig. 12).



Supplementary Figure 12 | Comparison of simulated, prepared, and ideal multimode W-states. (a) Absolute value of the two-mode W-state density matrix. Populations are represented with colors ranging from red to blue. Dashed black boxes indicate the ideal W-state populations, while yellow boxes show the simulated populations after accounting for the effects of cavity and transmon decoherence and decay. (b) Same as (a), but for the three-mode case. (c) W-state entanglement witness for 2 and 3 modes. The purple circles correspond to measurements of the two-mode witness as function of the cavity pulse duration in the presence of the blockade drive. The orange ⋆ corresponds to the measured witness for the three-mode state at the time corresponding to the W-state. The solid lines represent the result of master equation simulations of the witness. The orange band between -1/3 and 0 are the witness values that indicate W-state-like tripartite entanglement.

[1] S. Chakram, A. E. Oriani, R. K. Naik, A. V. Dixit, K. He, A. Agrawal, H. Kwon, and D. I. Schuster, Physical Review Letters **127**, 107701 (2021).

[2] N. Leung, M. Abdelhafez, J. Koch, and D. Schuster, Physical Review A **95**, 042318 (2017).

[3] P. Reinhold, *Controlling Error-Correctable Bosonic Qubits*, Ph.D. thesis, Yale University (2019).

[4] J. A. Smolin, J. M. Gambetta, and G. Smith, Physical review letters **108**, 070502 (2012).

[5] C. Wang, Y. Y. Gao, P. Reinhold, R. W. Heeres, N. Ofek, K. Chou, C. Axline, M. Reagor, J. Blumoff, K. Sliwa, *et al.*, Science **352**, 1087 (2016).

[6] Y. Ma, X. Pan, W. Cai, X. Mu, Y. Xu, L. Hu, W. Wang, H. Wang, Y. P. Song, Z.-B. Yang, *et al.*, arXiv:2005.09849 (2020).

[7] S. Rosenblum, P. Reinhold, M. Mirrahimi, L. Jiang, L. Frunzio, and R. J. Schoelkopf, Science **361**, 266 (2018).

[8] J. R. Johansson, P. D. Nation, and F. Nori, Computer Physics Communications **183**, 1760 (2012).

[9] O. Gühne and G. Tóth, Physics Reports **474**, 1 (2009).

[10] J. A. Mlynek, A. A. Abdumalikov Jr, J. M. Fink, L. Steffen, M. Baur, C. Lang, A. F. van Loo, and A. Wallraff, Physical Review A **86**, 053838 (2012).