Efficient multimode Wigner tomography

Kevin He,[1,2] Ming Yuan,[3] Yat Wong,[3] Srivatsan Chakram,[4] Alireza Seif,[3] Liang Jiang,[3] and David I. Schuster[1,2,3]

[1]*James Franck Institute, University of Chicago, Chicago, Illinois 60637, USA*
[2]*Department of Physics, University of Chicago, Chicago, Illinois 60637, USA*
[3]*Pritzker School of Molecular Engineering, University of Chicago, Chicago, Illinois 60637, USA*
[4]*Department of Physics and Astronomy, Rutgers University, Piscataway, NJ 08854, USA*

Advancements in quantum system lifetimes and control have enabled the creation of increasingly complex quantum states, such as those on multiple bosonic cavity modes. When characterizing these states, traditional tomography scales exponentially in both computational and experimental measurement requirement, which becomes prohibitive as the state size increases. Here, we implement a state reconstruction method whose sampling requirement instead scales polynomially with subspace size, and thus mode number, for states that can be expressed within such a subspace. We demonstrate this improved scaling with Wigner tomography of multimode entangled W states of up to 4 modes on a 3D circuit quantum electrodynamics (cQED) system. This approach performs similarly in efficiency to existing matrix inversion methods for 2 modes, and demonstrates a noticeable improvement for 3 and 4 modes, with even greater theoretical gains at higher mode numbers.

Quantum state tomography (QST) is the process of determining the quantum state of a system, and is a fundamental part of quantum information processing. The exponentially greater complexity of quantum systems compared to classical ones makes efficient QST a challenging task. In its conventional formulation, obtaining full state information has a processing and measurement requirement that scales exponentially with the size of the Hilbert space [1, 2]. However, physical states of interest typically have some structure that we can exploit to simplify the measurement complexity. Direct fidelity estimation (DFE) is a technique that utilizes this, and has been applied to matrix product states or stabilizer states in many-qubit systems [2–5] to efficiently produce partial information about the system state. In the remainder of this work, we refer to such states as DFE-efficient.

Efficient QST is especially relevant in continuous variable systems with bosonic cavity modes, whose Hilbert spaces are arbitrarily large. These systems have applications in error correction codes [6–8], quantum optics [9], quantum simulation [10], and quantum information processing [11]. For a single mode, full state information is obtained by measuring operators like the Wigner operator [12] or Q function operator at different mode displacements [13]. Efficient QST in the multimode case is much more challenging. Several efforts use multiple cavities to propose or produce increasingly complex states like multimode cat states [14], W states [15], multimode GKP states [16], GHZ states [17], and other multimode Fock state superpositions [18, 19] that have a variety of applications in quantum error correction and logical encodings, as well as quantum simulation. In particular, W states have unique multipartite entanglement and protection against photon loss that gives them applications in quantum communication. Some proposed theoretical methods are able to extract multimode state information while circumventing the exponential scaling of observation number with the number of

modes. These include techniques that apply additional unitaries between modes as part of the measurement process, perform targeted measurements with polynomial post-processing [20], make use of ancillary modes and a known excitation number [21], or apply operators based on excitation counting [13]. To obtain more full state information, we instead directly measure the density matrices of potentially mixed states confined in a subspace of interest.

In this work, we use the Direct Extraction of (Density) Matrix Elements from Subspace Sampling Tomography (DEMESST) method to *demystify* and reconstruct quantum state density matrices. DEMESST applies when an unknown state lies in a polynomial dimensional subspace. It only requires local operations if the subspace is spanned by a set of finite local operations acting on a DFE-efficient state [22]. Under these conditions, DEMESST has a polynomially scaling sampling requirement, and applies to both discrete qubit and continuous cavity systems. For certain multimode cavity states, the total measurement number will therefore depend polynomially on the number of modes, rather than exponentially. This is especially advantageous when the subspace that an expected state lives in is much smaller than the full space. With DEMESST, we individually sample measurements for each basis operator in a polynomial subspace, and subsequently reconstruct a density matrix by combining them. Additionally, we implement this method with Wigner operators and Wigner tomography, thus performing the measurements with purely local operations on the modes and eliminating errors that may be associated with multimode unitaries or beamsplitters. With this approach, we measure fidelities of W states prepared in up to 4 bosonic modes, which is beyond existing demonstrations and advances the state of the art.

The structure of this paper is as follows: we first describe the details of DEMESST and elaborate on how it operates. We then discuss existing QST methods for

continuous variable systems and how they have been optimized [15, 23]. We compare DEMESST to these previous methods by testing the observation number required to accurately reconstruct the density matrices of approximate entangled multimode W states of 2, 3, and 4 modes on a superconducting cQED system, without making prior assumptions about the populations or phases of the state components. Finally, we experimentally confirm the advantage of DEMESST and verify its expected properties.

DEMESST scales polynomially with mode number for states that have support in a polynomial subspace of DFE-efficient basis operators [3] (for example, states with a known maximum excitation number in the Fock basis). This is accomplished by leveraging that information, and rather than sampling displacements corresponding to all basis operators, only sampling for the ones that are expected to support the state. This is illustrated schematically in Fig. 1(a). We reconstruct a density matrix ρ in a polynomial subspace by independently measuring basis operators and the corresponding matrix elements for each projection of ρ within the subspace, through methods similar to DFE [3, 4], without introducing bias (see supplementary information). The reconstructed state is given by

$$\rho = \begin{bmatrix} \mathrm{Tr}[\rho O_{\vec{n}_1, \vec{n}_1}] & \mathrm{Tr}[\rho O_{\vec{n}_2, \vec{n}_1}] & \cdots \\ & \mathrm{Tr}[\rho O_{\vec{n}_2, \vec{n}_2}] & \\ \vdots & & \ddots \end{bmatrix} \quad (1)$$

With this approach, we avoid extracting irrelevant information about states outside the subspace of interest, thereby lowering the number of measurements required for an accurate result (see supplementary information). For example, if we know a 3-mode state has a maximum of 2 photons ($M = 3, N = 3$ for 0, 1, or 2 photon population), we reconstruct it with DEMESST by measuring the matrix elements associated with that subspace, namely the one formed by $|000\rangle, |001\rangle, |011\rangle, |002\rangle$, and permutations. We eliminate unnecessary sampling of unnecessarily high photon number states like $|111\rangle$ or $|012\rangle$ and beyond. Additionally, we build upon existing work [2, 3] by developing and applying this approach to QST of bosonic modes and arbitrary mixed states rather than qubits and pure states. Here, we implement DEMESST on a multimode cavity system with Wigner tomography.

Wigner tomography uses measurements of the Wigner operator $\mathcal{W}(\vec{\alpha}) = \mathcal{D}(\vec{\alpha})\Pi\mathcal{D}(-\vec{\alpha})$ acting on a bosonic state ρ to reconstruct it. Here, $\mathcal{D}(\vec{\alpha}) = \bigotimes_i \mathcal{D}(\alpha_i)$ is the displacement operator and Π is a parity measurement. Existing inversion-based Wigner tomography methods operate by taking the Wigner functions of a set of displacements $\{\vec{\alpha}\}$ to construct a measurement matrix $\mathcal{M}$ that maps to states as $\vec{x} = \mathcal{M}|\rho\rangle\rangle$, where $|\rho\rangle\rangle$ is the vectorized form of ρ. Inverting $\mathcal{M}$ to minimize$_\rho \|\mathcal{M}|\rho\rangle\rangle - \vec{x}\|$
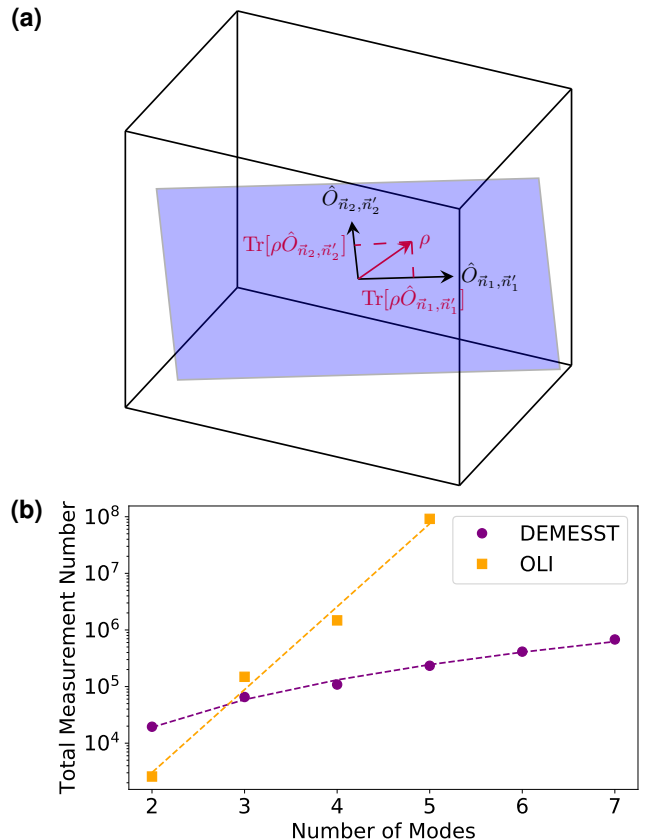

(a)

(b)

Figure 1 | Theoretical comparison and schematic representation of tomography methods. (a) Schematic representing the DEMESST method. Rather than sampling an entire multimode operator space (3D space), if a state lives in some number of subspaces (blue 2D plane), we restrict the sampling to each of those instead. The $\{O\}$ basis operators are of the form $O_{\vec{n}, \vec{n}'} = |\vec{n}\rangle\langle\vec{n}'|$ for generic basis states $|\vec{n}\rangle, |\vec{n}'\rangle$ (see supplementary information). Assuming an orthonormal basis, the state ρ is given by $\rho = \sum_{\vec{n}_1, \vec{n}_2} \mathrm{Tr}[\rho O_{\vec{n}_2, \vec{n}_1}]O_{\vec{n}_1, \vec{n}_2}$. This improves the overall efficiency of the sampling, especially for states with support across large numbers of modes. In practice, we use Hermitian $\{O\}$ that are accessible through experiment. (b) Number of measurements required for the DEMESST (purple, circles) and OLI (orange, squares) methods to reach a 90% state reconstruction fidelity on W states of up to 7 modes, assuming perfect state preparation. OLI scales exponentially with the size of the Hilbert space (and therefore the number of modes M), while DEMESST scales only polynomially.

allows us to determine the physical (unit trace and positive semidefinite) ρ that was most likely to have produced $\vec{x}$. The set $\{\vec{\alpha}\}$ is optimized by minimizing the condition number (the ratio of largest to smallest eigenvalue) of $\mathcal{M}$ and thus the error magnification, using the techniques presented in [15, 23]. We refer to this method as Optimized Linear Inversion (OLI). In this approach, to make
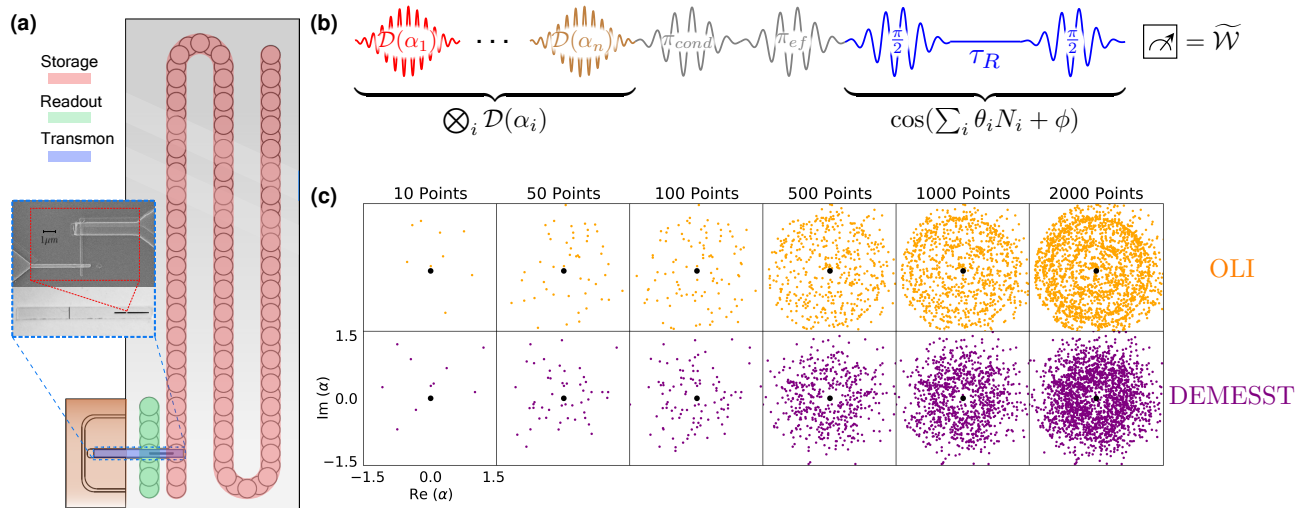
Figure 2 | **Experimental system and scheme**. (a) Diagram of the multimode flute cavity with a storage cavity and readout cavity that both couple to a transmon chip. Experimental drives are input through the readout cavity, or through a direct drive on the storage cavity. (b) Wigner tomography pulse sequence. Initial cavity displacements and a final generalized multimode parity measurement implement the tomography, while optional conditional pi pulses are used for projecting targeted modes out of the transmon ge subspace for the DEMESST approach (see supplementary information). An additional angle ϕ is applied between the $\pi/2$ pulses of the parity measurement to project it onto the real axis. (c) Cavity displacement plots for the OLI (orange) and DEMESST (purple) sampling methods. The OLI has ring features corresponding to measurement of all Fock states up to a cutoff, while DEMESST has points more densely located in the phase space based on the basis state being measured.

the problem tractable, we choose a cutoff dimension N to truncate the Hilbert space. Reconstructing ρ for a single mode therefore requires at least N^2 measurements to determine each density matrix parameter. For multiple modes, the size of the Hilbert space and thus the number of required measurements will scale exponentially, requiring at least N^{2M} observations for M modes. We compare OLI with DEMESST by testing their performance on experimentally prepared W states.

W states are excellent candidates for testing our Wigner tomography sampling methods. For 2 modes, this forms a dual rail encoding, $|W_2\rangle = (|10\rangle + e^{i\phi} |01\rangle)/\sqrt{2}$. For 3 modes, $|W_3\rangle = (|100\rangle + e^{i\phi_1} |010\rangle + e^{i\phi_2} |001\rangle)/\sqrt{3}$, and similarly for 4 modes, $|W_4\rangle = (|1000\rangle + e^{i\phi_1} |0100\rangle + e^{i\phi_2} |0010\rangle + e^{i\phi_3} |0001\rangle)/\sqrt{4}$. Here, the ϕ_j's are *a priori* unknown phases on each of the state components, and are determined through measurement. Additionally, due to imperfect state preparation, we make no assumptions about the component populations and precisely measure each one separately, making our measurement task more difficult. We further generalize this: rather than restricting ourselves to the pure states $W_j = |W_j\rangle \langle W_j|$, we measure the full, possibly mixed density matrices. W states are great representative states because they are irreducible multimode states that generalize straightforwardly to arbitrary numbers of modes, and have a well-defined photon number. We prepare them easily using photon blockade [15, 24–26].

We first investigate the simulated theoretical performance of DEMESST and OLI on M-mode W states. Assuming perfect state preparation, we compare the number of observations required to accurately reconstruct the W state with 90% fidelity. This is shown in Fig. 1(b). As discussed previously, inversion-based methods like OLI have a sampling requirement that scales exponentially with mode number M. In contrast, the DEMESST method scales polynomially with the subspace dimension and thus M, demonstrating an advantage that increases with M. For two modes, OLI performs better due to overhead required for the DEMESST approach (see supplementary information). However, for larger M, DEMESST requires fewer measurements to converge to the same level of fidelity, and scales much more efficiently than OLI. We proceed to demonstrating this expected behavior in experiment.

We generate W states and implement DEMESST and OLI on a superconducting 3D cQED platform. The system consists of a transmon qubit coupled to a 3D readout cavity and a 3D multimode storage cavity. A schematic of this hardware setup is shown in Fig. 2(a). The single storage cavity supports many bosonic cavity modes at roughly equally spaced microwave frequencies. The transmon allows for universal control of the cavity modes and also mediates interactions like photon blockade [24, 26] between the storage modes. We use four of the modes to prepare our multimode W states. We also

	Simulated Fidelity	OLI Fidelity	DEMESST Fidelity	OLI:DEMESST Distance
2-mode	0.971	0.966(5)	0.96(1)	0.96(2)
3-mode	0.956	0.949(4)	0.955(4)	1.55(10)
4-mode	0.912	0.912(7)	0.911(7)	2.3(2)

Table 1 | OLI and DEMESST Tomography results. Simulations and Wigner tomography measurements are performed on M-mode W states of varying size. The fidelities are in good agreement, and the Frobenius norm matrix distance ratio demonstrates exponential improvement as M increases.

use the transmon to implement the parity measurements necessary for the multimode Wigner tomography.

The tomography sampling methods use the same generalized Wigner tomography sequence, where each cavity mode is displaced before performing a generalized parity measurement on the transmon [15]. This procedure allows us to perform multimode tomography measurements despite having unequal dispersive shifts χ_m for different modes, without requiring χ engineering techniques [27, 28] or additional control pulses. This pulse sequence is shown in Fig. 2(b). For the DEMESST method, we may also include a π_{ge} pulse conditioned on certain cavity populations followed by an π_{ef} pulse on the transmon. These pulses project one or more of the cavity modes to the transmon $|f\rangle$ level and outside the $|g\rangle - |e\rangle$ qubit space of the Wigner tomography. This effectively removes those modes from the measurement and allows us to perform it more easily. We apply this when the basis operator being measured has one or more cavity modes in vacuum. For example, to sample the 3-mode matrix element $|001\rangle \langle 010|$, we project the first mode to $|f\rangle$ and reduce the sampling to the 2-mode $|01\rangle \langle 10|$. This projection allows us to reduce the size of the sampling problem for that element to a lower number of modes (see supplementary information).

We first compare the state reconstructed from the OLI method with simulation, which we use as a baseline for later comparison to DEMESST. With OLI, we find prepared W state fidelities that are in good agreement with the simulated fidelities, as shown in Table 1. The simulations include error from transmon and cavity decoherence and decay, and state preparation errors such as leakage outside of the blockaded subspace. We now continue to the DEMESST performance.

For the DEMESST approach, we reconstruct the 2–4 mode density matrices by measuring Wigner operators corresponding to multimode Fock basis states with up to 2 photons. Even though W states have at most one photon, we measure two-photon operators to capture imperfect state preparation errors. These observations directly provide the values of each density matrix element. From the density matrix, we obtain the component populations and phase angles ϕ_j of our prepared approximate W states by calculating the phase angle value that best matches the resulting data. These angles are then verified

to match with the ones obtained from the OLI approach.

We quantify the performance of the DEMESST and OLI sampling methods versus total measurement number with two metrics: reconstructed state fidelity and Frobenius norm matrix distance. The fidelities are computed with respect to an ideal W state, while the matrix distances are calculated with respect to the experimentally prepared state reconstructed at the maximum measurement number. These results are shown in Fig. 3. The final fidelities obtained from the DEMESST approach are presented in Table 1, and are consistent with the OLI results. Error bars are obtained from the results of 10 independent sets of 10 repetitions of tomography measurements for each sampled displacement. The number of distinct displacements for the OLI method therefore equals the Total Measurement Number shown on the x-axis in Fig. 3 divided by 10. For the DEMESST method, the number of distinct displacements for each basis element is further divided by the number of elements. For the 2-mode W state, the two methods perform similarly, while for 3 and 4 modes, DEMESST performs better than OLI with faster convergence to the final state fidelity.

This improvement is most evident in the matrix distance comparisons. The distances are computed using the Frobenius norm, with error bars again computed from 10 independent sets of 10 repetitions for each displacement. The final state density matrix against which the distances are computed is obtained by considering all 100 repetitions, rather than sets of 10, which is why the final distances do not completely vanish. The behavior of both sampling methods is nearly identical for the 2-mode W state. However, for the 3-mode case, the DEMESST has noticeably faster convergence versus total measurement number x, as the matrix distance d to the final state is smaller, as seen by the fit coefficient to $d = ax^b$ ($a = 1.1 \pm 0.1$ for DEMESST versus 1.71 ± 0.02 for OLI). This effect is further enhanced in the 4-mode case. The ratio of these fit values scales roughly geometrically, as shown in Table 1, reflecting the fact that OLI scales exponentially while DEMESST only scales polynomially versus total measurement number. In all cases, the distances fall off roughly as $x^{-1/2}$, as expected.

An advantage of the DEMESST sampling method compared to OLI is its self-consistency. Individual density matrix elements for any multimode state are mea-
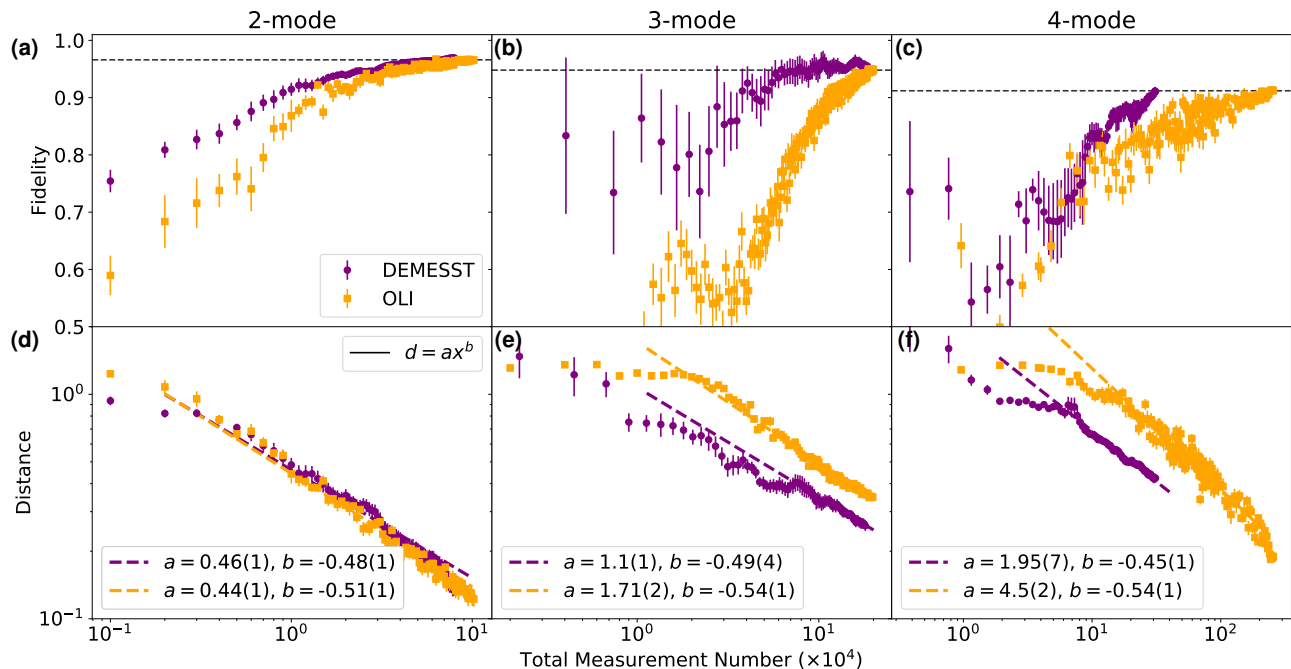
Figure 3 | Tomography fidelity and matrix distances for DEMESST and OLI sampling methods. Sampling was performed on approximate entangled W states of 2–4 modes. The top panels (a)–(c) show fidelities versus an ideal (exactly equal population coefficients) W state for 2–4 modes, with dashed horizontal lines indicating the final converged fidelity obtained from the OLI method. These final fidelities are, for 2–4 modes, $0.966 \pm 0.005, 0.949 \pm 0.004$, and 0.912 ± 0.007 for OLI and $0.96 \pm 0.01, 0.954 \pm 0.004$, and 0.911 ± 0.007 for DEMESST and are in good agreement. The bottom panels (d)–(f) show Frobenius norm matrix distances between the state at a given measurement number versus the final measured state. The rates of convergence are close to $1/\sqrt{x}$ or a power of -0.5, as expected. As the mode number increases, the DEMESST method performs increasingly more efficiently by requiring fewer measurements to reach a given level of convergence or error threshold.

sured independently, without needing to choose a cutoff maximum photon number or Hilbert space size that could subject the reconstructed state to inversion errors. This eliminates the risk of obtaining an inaccurate tomography result if, for example, the prepared state contains population beyond the space spanned by our chosen basis during OLI sampling.

We verify that the DEMESST tomography sampling method leads to self-consistent measurement results. We check the traces of our prepared W states and compare them with unity, the expected value for physical states. This allows us to confirm that our prepared state indeed lives in our chosen, measured Hilbert space. The results are shown in Fig. 4 as average trace versus observation number. Like before, the averages are taken over 10 independent sets of 10 measurement repetitions for each sampled displacement. We find that in all cases ((a)–(c)), the observed traces are near one. We attribute large deviations from unity at low measurement numbers to noise and statistics, and attribute the final traces being slightly less than one to imperfect state preparation that produces population outside the measured subspace. We perform a further check by considering only a 2-mode

subspace of a prepared 4-mode W state. This is shown in Fig. 4(d). As expected, the measured trace converges to a value near 0.5, as we are effectively only observing half of the total state population. This demonstrates that the DEMESST method provides accurate results for each basis operator independently. In particular, we can identify when we have measured insufficient basis elements to fully characterize a state, such as when the state lives partially (or entirely) outside the corresponding space, which is a useful capability in itself.

In summary, we have applied the DEMESST sampling method to characterize multimode cavity states with Wigner tomography. DEMESST is most appropriate for multimode states that have population contained in a small subspace of DFE-efficient elements of an overall Hilbert space, and outperforms traditional optimized inversion-based methods by scaling polynomially rather than exponentially with mode number. We observe this improvement for W states on 3 and 4 modes. Here, we have presented comparisons using the multimode Fock basis on multimode W states, but DEMESST also applies to different bases that more readily support other states; this tomography method can even be used for
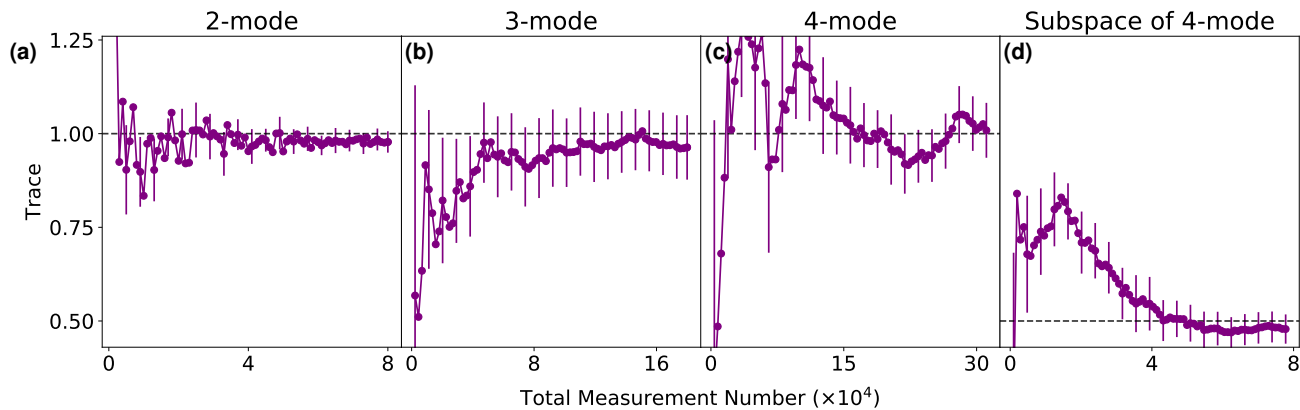
Figure 4 | Trace verification for the DEMESST sampling method. Trace versus point number for prepared (a) 2-mode, (b) 3-mode, and (c) 4-mode W states. Error bars are shown for every fourth point, as well as the final one. They are obtained from comparing the traces from individual sets of measurements. As expected, the traces converge to values near unity. (d) Trace versus point number when measuring only a 2-mode subspace of a prepared 4-mode W state. Due to only measuring half of the populated space, the trace converges to 0.5. This demonstrates that the DEMESST sampling method is self-consistent and does not depend on the chosen measurement subspace.

DFE by choosing as a basis the intended target state. While Wigner tomography was presented in this work, the method also operates beyond the bosonic Wigner function, and works for both continuous and discrete systems. This approach operates without coupling gates between modes, and would be useful for calibrating entangled states over distributed quantum networks. Ultimately, the DEMESST sampling method enables efficient measurement of large states, which will be crucial as the size of quantum hardware systems increases and more complicated states are generated and applied for quantum simulation, bosonic logical state encoding, and error correction.

[1] Mohseni, M., Rezakhani, A. & Aspuru-Guzik, A. Direct estimation of single- and two-qubit hamiltonians and relaxation rates. *Physical Review A* **77**, 042320 (2008).

[2] da Silva, M., Landon-Cardinal, O. & Poulin, D. Practical characterization of quantum devices without tomography. *Physical Review Letters* **107**, 210404 (2011).

[3] Cao, S. *et al.* Generation of genuine entanglement up to 51 superconducting qubits. *Nature* **1**, 738–742 (2023).

[4] Flammia, S. & Liu, Y. Direct fidelity estimation from few pauli measurements. *Physical Review Letters* **106**, 230501 (2011).

[5] Zhu, D. *et al.* Cross-platform comparison of arbitrary quantum states. *Nature Communications* **13** (2022).

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

[7] Ni, Z. *et al.* Beating the break-even point with a discrete-variable-encoded logical qubit. *Nature* **1**, 56–60 (2023).

[8] Sivak, V. *et al.* Real-time quantum error correction beyond break-even. *Nature* **1**, 50–55 (2023).

[9] Sundaresan, N. M. *et al.* Beyond strong coupling in a

multimode cavity. *Physical Review X* **5**, 021035 (2015).

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

[11] Naik, R. *et al.* Random access quantum information processors using multimode circuit quantum electrodynamics. *Nature Communications* **8**, 1–7 (2017).

[12] Cahill, K. & Glauber, R. Density operators and quasiprobability distributions. *Physical Review Journals Archive* **177**, 1882 (1969).

[13] Shen, C. *et al.* Optimized tomography of continuous variable systems using excitation counting. *Physical Review A* **94**, 052327 (2016).

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

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

[16] Royer, B., Singh, S. & Girvin, S. Encoding qubits in multimode grid states. *Physical Review X Quantum* **3**, 010335 (2022).

[17] Bäumer, E. *et al.* Efficient long-range entanglement using dynamic circuits. *Arxiv preprint* (2023).

[18] Ouyang, Y. & Chao, R. Permutation-invariant constant-excitation quantum codes for amplitude damping. *IEEE Transactions on Information Theory* **66**, 2921–2933 (2019).

[19] Niu, M., Chuang, I. & JH, S. Hardware-efficient bosonic quantum error-correcting codes based on symmetry operators. *Physical Review A* **97**, 032323 (2018).

[20] Cramer, M. *et al.* Efficient quantum state tomography. *Nature Communications* **1** (2010).

[21] Chen, W. *et al.* Scalable and programmable phononic network with trapped ions. *Nature Physics* **19**, 877–883 (2023).

[22] Wong, Y. *et al.* Efficient tomography of a polynomial subspace. *in preparation* .

[23] Reinhold, P. *Controlling Error-Correctable Bosonic Qubits*. Ph.D. thesis, Yale University (2019). URL `https://rsl.yale.edu/sites/default/files/files/RSL_Theses/Reinhold-Thesis%20(1).pdf`.

[24] Heeres, R. W. *et al.* Cavity state manipulation using photon-number selective phase gates. *Physical review letters* **115**, 137002 (2015).

[25] Heeres, R. W. *et al.* Implementing a universal gate set on a logical qubit encoded in an oscillator. *Nature Communications* **8**, 1–7 (2017).

[26] Bretheau, L., Campagne-Ibarcq, P., Flurin, E., Mallet, F. & Huard, B. Quantum dynamics of an electromagnetic mode that cannot contain n photons. *Science* **348**, 776–779 (2015).

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

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

Supplementary Information: Efficient multimode Wigner tomography

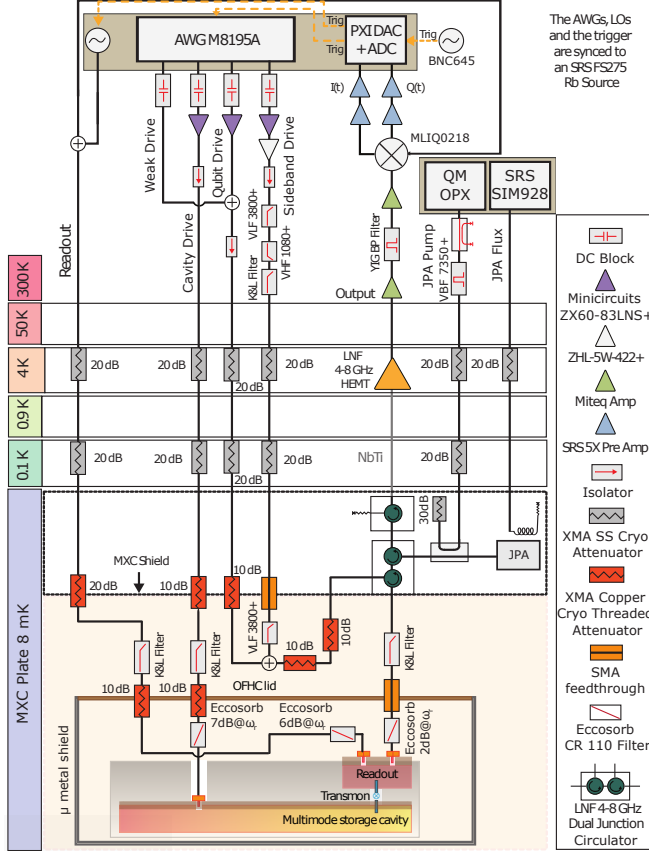## I. CRYOGENIC SETUP AND CONTROL INSTRUMENTATION

Figure S1 | 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 (10-12 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. S1. 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 [S 1] and also used in [S 2]. The cavities are bridged by a 3D transmon circuit. Controls are performed through the readout cavity or the direct cavity drive line, 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. S1. 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. CALIBRATION OF THE MULTIMODE HAMILTONIAN

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

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

where $\omega_q/(2\pi) = 4.96$ GHz is the frequency of the transmon $|g\rangle$ - $|e\rangle$ transition, $\omega_m/(2\pi)$ ranging from $5.72 - 6.47$ GHz are the cavity mode frequencies, $\chi_m/(2\pi)$ ranging from -1.64 to -0.91 MHz are the dispersive shifts, k_m the self-Kerr shift of each mode, and k_{mn} the cross-Kerr interactions between the modes. The Kerr nonlinearities range from $-6-7$ kHz. Parameter values are determined with the processes described in the supplement of [S 2]. A summary of the system parameters used in the experiment, as well as Liouvillian terms corresponding to transmon and cavity decoherence and decay, is provided in STable 2. The 2-mode experiments used the last two modes in the table, the 3-mode experiments used the last three modes, and the 4-mode experiments used all four of the modes.

III. THE WIGNER FUNCTION AND ITS GENERALIZATION

The Wigner function is the quasiprobability distribution of a state in phase space, and is one of the most important functions in the field of quantum optics. For a single-mode state ρ, the Wigner function is defined as [S 3]

$$
W_\rho(\alpha) = 2\text{Tr}[\rho D(\alpha) e^{i\pi a^\dagger a} D^\dagger(\alpha)],
\tag{2}
$$

Parameter	Hamiltonian/Liouvillian Term	Quantity	Value(s)
Transmon frequency	$\omega_q \lvert e \rangle \langle e \rvert$	$\omega_q/(2\pi)$	4.97 GHz
Storage cavity frequencies	$\omega_m a_m^\dagger a_m$	$\omega_m/(2\pi)$	5.716, 5.965, 6.223, 6.469 GHz
Readout frequency	$\omega_r a_r^\dagger a_r$	$\omega_r/(2\pi)$	7.79 GHz
Readout dispersive shift	$\chi_r a_r^\dagger a_r \lvert e \rangle \langle e \rvert$	$\chi_r/(2\pi)$	1 MHz
Storage mode dispersive shifts	$\chi_m a_m^\dagger a_m \lvert e \rangle \langle e \rvert$	$\chi_m/(2\pi)$	-1.636, -1.269, -1.093, -0.906 MHz
Storage mode self-Kerrs	$\frac{k_m}{2} a_m^\dagger a_m (a_m^\dagger a_m - 1)$	$k_m/(2\pi)$	9.0, 5.2, 4.2, 3.2 kHz
Storage mode cross-Kerrs	$k_{mn} a_m^\dagger a_m a_n^\dagger a_n$	$k_{mn}/(2\pi)$	$-6 - 0$ kHz
Transmon $\lvert e \rangle \to \lvert g \rangle$ relaxation	$\frac{1}{T_1^q}(1+\bar{n})\mathcal{D}\big[\lvert g \rangle \langle e \rvert \big]$	T_1^q	$108 \pm 7\ \mu s$
Transmon $\lvert g \rangle - \lvert e \rangle$ dephasing	$\left(\frac{1}{T_2^q} - \frac{1}{2T_1^q}\right)\mathcal{D}\big[\lvert e \rangle \langle e \rvert \big]$	T_2^q	$165 \pm 14\ \mu s$
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	$\sim 1-2$ ms, see [S 1]
Transmon thermal population	$\frac{\bar{n}}{T_1^q}\mathcal{D}\big[\lvert e \rangle \langle g \rvert \big]$	$\bar{n}$	1.5 ± 0.5 %
Storage mode dephasing	$\left(\frac{1}{T_2^m} - \frac{1}{2T_1^m}\right)\mathcal{D}\big[\lvert 1 \rangle \langle 1 \rvert \big]$	T_2^m	$\sim 1.5 - 3$ ms, see [S 1]

Table 2 | Multimode cQED system parameters

where $D(\alpha) = e^{\alpha a^\dagger - \alpha^* a}$ is the displacement operator. We can see that $W_\rho(\alpha)$ is proportional to the expectation value of the parity operator with the state ρ displaced by complex amplitude $-\alpha$. Similarly, we can introduce the Wigner function for an operator O with finite Frobenius norm (F-norm) ($\lVert O \rVert_F = \sqrt{\mathrm{Tr}[O^\dagger O]} < +\infty$) by substituting the state ρ with the operator O in Eqn. (2).

The Wigner function of a multimode M-mode state ρ can be expressed as

$$W_\rho(\vec{\alpha}) = 2^M \mathrm{Tr}[\rho D(\vec{\alpha}) e^{i\pi \sum_{m=1}^{M} a_m^\dagger a_m} D^\dagger(\vec{\alpha})], \quad (3)$$

where $D(\vec{\alpha}) = \prod_{m=1}^{M} e^{\alpha_m a_m^\dagger - \alpha_m^* a_m}$. As described in the main text, due to the different dispersive shifts of our cavity modes, we instead implement a generalized parity operator [S 2]. We can express the corresponding generalized version of the Wigner function, as introduced in [S 4], as

$$W_\rho(\vec{\alpha}, \vec{\theta}) = \frac{2^M \mathrm{Tr}[\rho D(\vec{\alpha}) e^{i \sum_{m=1}^{M} \theta_m a_m^\dagger a_m} D^\dagger(\vec{\alpha})]}{\prod_{m=1}^{M}[1 + i \cot(\theta_m/2)]}, \quad (4)$$

where θ_m can be different and need not equal π for the M modes. We also denote

$$\tilde{W}_\rho(\vec{\alpha}, \vec{\theta}) = \mathrm{Tr}[\rho D(\vec{\alpha}) e^{i \sum_{m=1}^{M} \theta_m a_m^\dagger a_m} D^\dagger(\vec{\alpha})], \quad (5)$$

where $\tilde{W}_\rho(\vec{\alpha}, \vec{\theta})$ can in general be a complex number, since $e^{i \sum_{m=1}^{M} \theta_m a_m^\dagger a_m}$ is no longer Hermitian. Also, $\lvert \tilde{W}_\rho(\vec{\alpha}, \vec{\theta}) \rvert \leq 1$. We can see from the form of Eqn. (5) that $\tilde{W}_\rho(\vec{\alpha}, \vec{\theta}) = \tilde{W}_{\rho^\dagger}^*(\vec{\alpha}, -\vec{\theta})$. This generalized Wigner function $\tilde{W}_\rho(\vec{\alpha}, \vec{\theta})$ is what we experimentally measure. In the next section, we show that $\tilde{W}_\rho(\vec{\alpha}, \vec{\theta})$ for quantum states plays a similar role as the usual Wigner function (Eqn. (3)) when calculating the expectation values with other operators.

IV. ESTIMATING EXPECTATION VALUES WITH WIGNER SAMPLING

In this section, we discuss the sampling method for estimating the expectation value of the Wigner function with an operator with finite F-norm O. We then analyze the overhead required to reach a certain accuracy threshold.

We start with the identity that reflects the relationship between expectation values and the Wigner function:

$$\mathrm{Tr}[\rho O] = \int \frac{\mathrm{d}^{2M}\vec{\alpha}}{\pi^M} W_\rho(\vec{\alpha}) W_O(\vec{\alpha}). \quad (6)$$

For the generalized Wigner function, we have a similar expression:

$$\begin{aligned} \mathrm{Tr}[\rho O] &= \int \frac{\mathrm{d}^{2M}\vec{\alpha}}{\pi^M} W_\rho(\vec{\alpha}, -\vec{\theta}) W_O(\vec{\alpha}, \vec{\theta}) \\ &= C_M \int \mathrm{d}^{2M}\vec{\alpha}\ \tilde{W}_\rho(\vec{\alpha}, -\vec{\theta}) \tilde{W}_O(\vec{\alpha}, \vec{\theta}), \end{aligned} \quad (7)$$

where $C_M = \prod_{m=1}^{M}[2(1 - \cos \theta_m)/\pi]$.

The equations above have been applied for direct fidelity estimation between an experimentally prepared state ρ and a target pure state σ [S 5]. For example, we can rewrite Eqn. (7) as

$$\begin{aligned} F(\rho, \sigma) = \mathrm{Tr}[\rho\sigma] &= C_M \int \mathrm{d}^{2M}\vec{\alpha}\ \lvert \tilde{W}_\sigma(\vec{\alpha}, \vec{\theta}) \rvert^2 \frac{\tilde{W}_\rho(\vec{\alpha}, -\vec{\theta})}{\tilde{W}_\sigma(\vec{\alpha}, -\vec{\theta})} \\ &= \int \mathrm{d}^{2M}\vec{\alpha}\ p_{W^2}(\vec{\alpha}) \frac{\mathrm{Re}[e^{i\phi(\vec{\alpha})} \tilde{W}_\rho(\vec{\alpha}, -\vec{\theta})]}{\lvert \tilde{W}_\sigma(\vec{\alpha}, -\vec{\theta}) \rvert}, \end{aligned}$$
$$(8)$$

where $p_{W^2}(\vec{\alpha}) = C_M|\tilde{W}_\sigma(\vec{\alpha}, \vec{\theta})|^2$ satisfies $p_{W^2}(\vec{\alpha}) \geq 0$ and $\int \mathrm{d}^{2M}\vec{\alpha} \ p_{W^2}(\vec{\alpha}) = \mathrm{Tr}[\sigma^2] = 1$, such that we can treat $p_{W^2}(\vec{\alpha})$ as a probability distribution function that we can then use to sample a set of displacement vectors $\{\vec{\alpha}^{(k)}\}$. Given a set $\{\vec{\alpha}^{(k)}\}$, we can measure $\mathrm{Re}[e^{i\phi(\vec{\alpha}^{(k)})}\tilde{W}_\rho(\vec{\alpha}^{(k)}, -\vec{\theta})]$ and calculate the average $\frac{\mathrm{Re}[e^{i\phi(\vec{\alpha}^{(k)})}\tilde{W}_\rho(\vec{\alpha}^{(k)}, -\vec{\theta})]}{|\tilde{W}_\sigma(\vec{\alpha}^{(k)}, -\vec{\theta})|}$ to obtain an estimate of $F(\rho, \sigma)$. In Sec. VI, we show explicitly that by repeatedly measuring whether the qubit is in the $|g\rangle$ level, we can obtain a series of binomial outcomes $A_j^{(k)} \in \{1, -1\}$ whose expectation values lead to exactly $\mathrm{Re}[e^{i\phi(\vec{\alpha}^{(k)})}\tilde{W}_\rho(\vec{\alpha}^{(k)}, -\vec{\theta})]$. We call the above sampling method the W^2 tomography sampling method.

In general, we have the freedom to choose other probability distribution functions $p(\vec{\alpha})$ to generate sampling points $\{\vec{\alpha}^{(k)}\}$, since

$$\mathrm{Tr}[\rho O]$$
$$= C_M \int \mathrm{d}^{2M}\vec{\alpha} \ p(\vec{\alpha}) \frac{|\tilde{W}_O(\vec{\alpha}, \vec{\theta})|}{p(\vec{\alpha})} \mathrm{Re}[e^{i\phi(\vec{\alpha})}\tilde{W}_\rho(\vec{\alpha}, -\vec{\theta})]. \tag{9}$$

We can calculate the average of $C_M \frac{|\tilde{W}_O(\vec{\alpha}^{(k)}, \vec{\theta})|}{p(\vec{\alpha}^{(k)})} A_j^{(k)}$ to obtain an estimate for $\mathrm{Tr}[\rho O]$. However, in certain situations, there will be an optimal choice for $p(\vec{\alpha})$. For example, if we perform single shot measurements where we only measure each sampling vector $\vec{\alpha}^{(k)}$ once, the variance of the estimator $C_M \frac{|\tilde{W}_O(\vec{\alpha}^{(k)}, \vec{\theta})|}{p(\vec{\alpha}^{(k)})} A^{(k)}$ will be limited by

$$C_M^2 \int \mathrm{d}^{2M}\vec{\alpha} \ p(\vec{\alpha}) \frac{|\tilde{W}_O(\vec{\alpha}, \vec{\theta})|^2}{p^2(\vec{\alpha})} - (\mathrm{Tr}[\rho O])^2$$
$$\geq \frac{[C_M \int \mathrm{d}^{2M}\vec{\alpha} \ |\tilde{W}_O(\vec{\alpha}, \vec{\theta})|]^2}{\int \mathrm{d}^{2M}\vec{\alpha} \ p(\vec{\alpha})} - (\mathrm{Tr}[\rho O])^2 \tag{10}$$
$$= \left[C_M \int \mathrm{d}^{2M}\vec{\alpha} \ |\tilde{W}_O(\vec{\alpha}, \vec{\theta})|\right]^2 - (\mathrm{Tr}[\rho O])^2.$$

Here, we have used the Cauchy-Schwarz inequality and the fact that $\int \mathrm{d}^{2M}\vec{\alpha} \ p(\vec{\alpha}) = 1$. The minimum is achieved when $p(\vec{\alpha}) \propto |\tilde{W}_O(\vec{\alpha}, \vec{\theta})|$, which we call the DEMESST sampling method. It is also worth noting that, for the W^2 method where $p_{W^2}(\vec{\alpha}) = C_M|\tilde{W}_O(\vec{\alpha}, \vec{\theta})|^2$, the integral shown in the first line of Eqn. (10) will be divergent. To avoid this, Ref. [S 5, 6] have proposed choosing a threshold cutoff value to discard some sampling vectors $\vec{\alpha}^{(k)}$ that make the denominator of $\frac{A_j^{(k)}}{|\tilde{W}_O(\vec{\alpha}^{(k)}, \vec{\theta})|}$ too small. This cutoff procedure can lead to bias when estimating $\mathrm{Tr}[\rho O]$ and makes the error analysis more complicated. A detailed analysis of the effect of this cutoff in a multimode setting is beyond the scope of this work.

Instead, we focus on the DEMESST method. The probability distribution function is given by

$$p_\mathrm{D}(\vec{\alpha}) = \frac{|\tilde{W}_O(\vec{\alpha}, \vec{\theta})|}{Z_O}, \tag{11}$$

where

$$Z_O = \int \mathrm{d}^{2M}\vec{\alpha} \ |\tilde{W}_O(\vec{\alpha}, \vec{\theta})|. \tag{12}$$

In the DEMESST method, we must average $C_M Z_O A_j^{(k)}$ over all sampling vectors $\vec{\alpha}^{(k)}$ and all possible binomial outcomes from the qubit measurement per sampling vector. In the limit where we do one qubit measurement per $\vec{\alpha}^{(k)}$, we can use Hoeffding's inequality to estimate the number of samples N_spl required to reach an accuracy ϵ_1 with probability $1 - \delta_1$ to be

$$P\left(\left|\frac{C_M Z_O}{N_\mathrm{spl}} \sum_{k=1}^{N_\mathrm{spl}} A^{(k)} - \mathrm{Tr}[\rho O]\right| \geq \epsilon_1\right) \leq \delta_1 \tag{13}$$

when

$$N_\mathrm{spl} \geq \lceil \frac{2C_M^2 Z_O^2}{\epsilon_1^2} \ln(2/\delta_1) \rceil. \tag{14}$$

We can see that, in general, $N_\mathrm{spl} \propto (C_M Z_O)^2$. In the next section, we analyze the properties of Z_O for our operators of interest.

V. DENSITY MATRIX RECONSTRUCTION PROCEDURE

In this section, we discuss our scheme for reconstructing an unknown state using the DEMESST sampling method to estimate each element of its density operator in the Fock basis. We consider a system with M modes and maximum total photon number N between those modes. This restricts the dimension of the Hilbert space to $\binom{M+N}{N}$. We focus on the scaling of the sampling overhead (number of samples required) vs. mode number M, in the limit where M is much larger than $2N$, and show that this overhead scales polynomially vs. M with bounded photon number N, demonstrating the efficiency of the DEMESST approach as M increases.

We assume that the operator O (in $\mathrm{Tr}[\rho O]$) is in one of the following forms:

$$O_{\vec{n}, \vec{n}} = |\vec{n}\rangle\langle\vec{n}|,$$
$$O_{\vec{n}, \vec{n}'}^R = \frac{|\vec{n}\rangle\langle\vec{n}'| + |\vec{n}'\rangle\langle\vec{n}|}{\sqrt{2}} \quad (\vec{n} \neq \vec{n}'), \tag{15}$$
$$O_{\vec{n}, \vec{n}'}^I = i\frac{|\vec{n}\rangle\langle\vec{n}'| - |\vec{n}'\rangle\langle\vec{n}|}{\sqrt{2}} \quad (\vec{n} \neq \vec{n}').$$

Here $|\vec{n}\rangle = \bigotimes_{m=1}^M |n_m\rangle$ and $|\vec{n}'\rangle = \bigotimes_{m=1}^M |n'_m\rangle$ are Fock basis states that satisfy $\sum_{m=1}^M n_m \leq N$ and $\sum_{m=1}^M n'_m \leq$
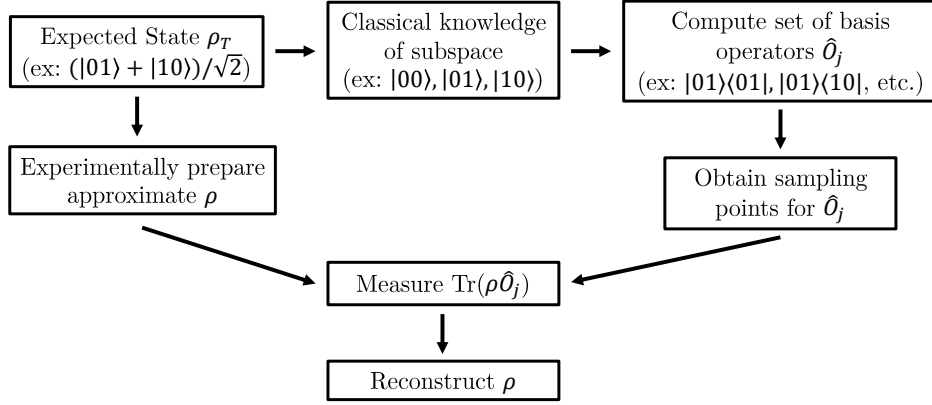
Figure S2 | Flowchart depicting the steps involved in our density matrix reconstruction procedure, of which DEMESST is an example. Starting with a target expected state ρ_T, we can use our knowledge of ρ_T to identify a target subspace and compute the operators $\hat{O}_j$ that span that basis. We then use those operators to generate sampling points, which may involve projections to be done efficiently, as described in Sec. IV. By experimentally implementing the operators on our prepared state, we can reconstruct the state ρ and compare it with the ideal target ρ_T. Even if the prepared state ρ is not close to ρ_T, as long as it lives in the same subspace, we can effectively reconstruct it.

N. The operators O are chosen such that $O = O^\dagger$ and $\text{Tr}[O^\dagger O] = 1$. For a system with M modes and maximum total photon number N, we have $\binom{M+N}{N}^2$ of these operators.

One essential observation is that, when $M > 2N$, for any $(\vec{n}, \vec{n}')$ pair there are at least $(M - 2N)$ elements in the set $S \equiv S_{(\vec{n}, \vec{n}')} = \{m | n_m = n'_m = 0\}$. We also denote $\bar{S} = \{1, 2, \ldots, M\} \backslash S$. Because of this, the corresponding operators shown in Eqn. (15) can always be decomposed as

$$O = \Big(\bigotimes_{m \in S} |0\rangle\langle 0|_m \Big) \otimes O_{\bar{S}}, \qquad (16)$$

where $O_{\bar{S}}$ has support on the modes with index $m \in \bar{S}$. We can see that the number of elements in $\bar{S}$ is no greater than $2N$, and independent of M. Similarly, the generalized Wigner function of such an operator O satisfies

$$\tilde{W}_O(\vec{\alpha}, \vec{\theta}) = \Big(\prod_{m \in S} \tilde{W}_{|0\rangle\langle 0|}(\alpha_m, \theta_m) \Big) \cdot \tilde{W}_{O_{\bar{S}}}(\vec{\alpha}_{\bar{S}}, \vec{\theta}_{\bar{S}}), \qquad (17)$$

where $\vec{\alpha}_{\bar{S}}$, $\vec{\theta}_{\bar{S}}$ contain those elements in $\vec{\alpha}$, $\vec{\theta}$ whose mode index $m \in \bar{S}$. Again, $\tilde{W}_{O_{\bar{S}}}(\vec{\alpha}_{\bar{S}}, \vec{\theta}_{\bar{S}})$ is independent of M.

Now, we consider the sampling overhead to obtain a precise estimate of $\text{Tr}[\rho O]$ when O satisfies the properties described above. One approach is to sample directly according to the M-mode function $|\tilde{W}_O(\vec{\alpha}, \vec{\theta})|/Z_O$. Based on Eqn. (14), the key quantity we should focus on is

$C_M Z_O$, which satisfies

$$C_M Z_O = \Big(\prod_{m=1}^{M} \frac{2(1 - \cos\theta_m)}{\pi} \Big) \int \text{d}^{2M}\vec{\alpha} \ |\tilde{W}_O(\vec{\alpha}, \vec{\theta})|$$
$$= \prod_{m \in S} \Big(\frac{2(1 - \cos\theta_m)}{\pi} \int \text{d}^2\vec{\alpha} \ |\tilde{W}_{|0\rangle\langle 0|}(\alpha_m, \theta_m)| \Big)$$
$$\cdot C_{\bar{S}} \int \text{d}^{2|\bar{S}|}\vec{\alpha}_{\bar{S}} \ |\tilde{W}_{O_{\bar{S}}}(\vec{\alpha}_{\bar{S}}, \vec{\theta}_{\bar{S}})|$$
$$= 2^{M - |\bar{S}|} C_{\bar{S}} Z_{O_{\bar{S}}}. \qquad (18)$$

Here, $C_{\bar{S}} = \prod_{m \in \bar{S}}[2(1 - \cos\theta_m)/\pi]$. Since $O_{\bar{S}}$ is supported on at most $2N$ modes, which is independent of M when $M > 2N$, the only M-dependence in $C_M Z_O$ comes from the 2^M factor. Unfortunately, this is still unfavorable, as it grows exponentially with mode number M.

To resolve this issue, consider $O_{\bar{S}}$, which is nontrivially supported on the modes contained in $\bar{S}$, rather than the full operator O itself. We introduce the projection operator $P_S = \bigotimes_{m \in S} |0\rangle\langle 0|_m$ and denote $\rho_{\bar{S}} = \text{Tr}_S[\rho P_S]$. Here $\text{Tr}_S[\bullet]$ means the partial trace over all modes with $m \in S$. We can see that

$$\text{Tr}[\rho O] = \text{Tr}[\rho_{\bar{S}} O_{\bar{S}}], \qquad (19)$$

where $\rho_{\bar{S}}$ and $O_{\bar{S}}$ are wholly supported on modes in $\bar{S}$, which contains at most $2N$ elements. We can perform DEMESST sampling according to $O_{\bar{S}}$ as follows:

$$\text{Tr}[\rho_{\bar{S}} O_{\bar{S}}]$$
$$= C_{\bar{S}} Z_{O_{\bar{S}}} \int \text{d}^{2|\bar{S}|}\vec{\alpha}_{\bar{S}} \ p_\text{D}(\vec{\alpha}_{\bar{S}}) \ \text{Re}[e^{i\phi(\vec{\alpha}_{\bar{S}})} \tilde{W}_{\rho_{\bar{S}}}(\vec{\alpha}_{\bar{S}}, -\vec{\theta}_{\bar{S}})]. \qquad (20)$$

In the experiment, we utilize the $|f\rangle$ level of the transmon to effectively restrict the cavity state to $\rho_{\bar{S}}$. From the measurement, we obtain binomial outcomes $A_j^{(k)} \in \{1, -1\}$ with expectation values $\text{Re}[e^{i\phi(\vec{\alpha}_{\bar{S}}^{(k)})}\tilde{W}_{\rho_{\bar{S}}}(\vec{\alpha}_{\bar{S}}^{(k)}, -\vec{\theta}_{\bar{S}})]$. More details about this experimental protocol is presented in Sec. VI. Like before, we can use Hoeffding's inequality to estimate the sampling overhead. If we only measure once per sampling vector $\vec{\alpha}_{\bar{S}}^{(k)}$, we will have

$$P\left(\left|\frac{C_{\bar{S}}Z_{O_{\bar{S}}}}{N_{\text{spl}}}\sum_{k=1}^{N_{\text{spl}}} A^{(k)} - \text{Tr}[\rho_{\bar{S}}O_{\bar{S}}]\right| \geq \epsilon_2\right) \leq \delta_2 \quad (21)$$

when

$$N_{\text{spl}} \geq \lceil \frac{2C_{\bar{S}}^2 Z_{O_{\bar{S}}}^2}{\epsilon_2^2}\ln(2/\delta_2)\rceil. \quad (22)$$

When N is bounded, $C_{\bar{S}}Z_{O_{\bar{S}}}$ is independent of mode number M when $M > 2N$. Therefore, N_{spl} in Eqn. (22) scales as

$$N_{\text{spl}} \sim \mathcal{O}_M\left(\frac{f(N)}{\epsilon_2^2}\ln(2/\delta_2)\right), \quad (23)$$

where $\mathcal{O}_M$ indicates that we focus only on the scaling over M in the large M limit, and $f(N) = 2C_{\bar{S}}^2 Z_{O_{\bar{S}}}^2$ is a function that depend solely on N and the specific form of O from Eqn. (15). We also introduce $f_{\max}(N)$ to represent the maximum value of $f(N)$ from those $\binom{M+N}{N}^2$ operators.

We can now consider our reconstructed density matrix $\hat{\rho}$. By performing expectation value estimation on the unknown state ρ with all $\binom{M+N}{N}^2$ operators with form in Eqn. (15), we can achieve

$$P(B) \geq 1 - \binom{M+N}{N}^2\delta_2 \quad (24)$$

with total sample number

$$N_{\text{tot}} \sim \mathcal{O}_M\left[\binom{M+N}{N}^2\frac{f_{\max}(N)}{\epsilon_2^2}\ln(2/\delta_2)\right], \quad (25)$$

where B requires all the conditions below:

$$|\langle\vec{n}|(\hat{\rho} - \rho)|\vec{n}\rangle| \leq \epsilon_2,$$
$$|\text{Re}[\langle\vec{n}|(\hat{\rho} - \rho)|\vec{n}'\rangle]| \leq \epsilon_2/\sqrt{2} \quad (\vec{n} \neq \vec{n}'), \quad (26)$$
$$|\text{Im}[\langle\vec{n}|(\hat{\rho} - \rho)|\vec{n}'\rangle]| \leq \epsilon_2/\sqrt{2} \quad (\vec{n} \neq \vec{n}').$$

If B is satisfied, we will have the Frobenius norm distance between our reconstructed state and the true state satisfy

$$||\hat{\rho} - \rho||_F = \sqrt{\sum_{\vec{n},\vec{n}'}|\langle\vec{n}|(\hat{\rho} - \rho)|\vec{n}'\rangle|^2} \leq \binom{M+N}{N}\epsilon_2. \quad (27)$$

Also, if ρ is a pure state, we will have

$$|F(\hat{\rho}, \rho) - 1| = |\text{Tr}[(\hat{\rho} - \rho)\rho]| \leq \epsilon_2 \cdot \sum_{\vec{n},\vec{n}'}|\langle n|\rho|\vec{n}'\rangle|$$
$$\leq \epsilon_2 \cdot \sqrt{\sum_{\vec{n},\vec{n}'}|\langle n|\rho|\vec{n}'\rangle|^2} \cdot \binom{M+N}{N}$$
$$= \epsilon_2 \cdot \sqrt{\text{Tr}[\rho^2]} \cdot \binom{M+N}{N} = \binom{M+N}{N}\epsilon_2. \quad (28)$$

In summary, by choosing $\epsilon_2 = \epsilon/\binom{M+N}{N}$ and $\delta_2 = \delta/\binom{M+N}{N}^2$, our sampling method will require the total sample number

$$N_{\text{tot}} \sim$$
$$\mathcal{O}_M\left[\binom{M+N}{N}^4\frac{f_{\max}(N)}{\epsilon^2}\ln\left(2\binom{M+N}{N}^2/\delta\right)\right] \quad (29)$$

to achieve

$$P(||\hat{\rho} - \rho||_F \leq \epsilon) \geq 1 - \delta, \quad (30)$$

even if we only perform a single measurement for each sampling instance. With the same amount of sampling, we can also achieve

$$P(|F(\hat{\rho}, \rho) - 1| \leq \epsilon) \geq 1 - \delta \quad (31)$$

when ρ is a pure state.

The procedure for DEMESST is as follows: first, the Wigner function corresponding to a chosen basis operator is normalized to a probability distribution based on its absolute value. Then, displacement vectors are sampled from the resulting inverse cumulative distribution function (CDF) by randomly selecting a value between 0 and 1 to find the corresponding angular and radial values of the displacements (see supplementary information). This calculation is performed efficiently by utilizing Laguerre functions and their inverses. During measurement, each displacement vector will have the original sign of its Wigner function value preserved, so that if the Wigner function was negative at that point, the final measured value will be multiplied by -1. Observing a set of these will provide an estimate of the chosen density matrix element. Repeating this for multiple elements will thus produce the density matrix of the prepared state.

VI. DEMESST EXPERIMENTAL PROTOCOL

For simplicity, in this section we assume that the cavity is initialized as a pure state $|\psi\rangle$. However, the same arguments will apply for a generic density operator ρ, which can always be decomposed as $\rho = \sum_i c_i|\psi_i\rangle\langle\psi_i|$ and understood as an ensemble average of a set of pure states $\{|\psi_i\rangle\langle\psi_i|\}$ with probability c_i.

First, we consider the generalized Wigner function of an M-mode state. We assume that the cavity-qubit state is initialized as $|\psi\rangle|g\rangle$. To perform the Wigner tomography measurement, we apply a short (large-bandwidth) drive to each cavity mode, then apply a large-bandwidth $\pi/2$ pulse on the qubit to begin the parity measurement. After these operations, the qubit part becomes $\exp(-i\frac{\pi}{4}\sigma_y)|g\rangle = \frac{|g\rangle+|e\rangle}{\sqrt{2}}$, and the cavity part becomes $|\psi_D\rangle = D(-\vec{\alpha})|\psi\rangle$. Then, as part of the parity measurement, we wait for a time t. Due to the dispersive interaction Hamiltonian $H_{\text{int}} = \sum_m \chi_m a_m^\dagger a_m |e\rangle\langle e|$, the cavity modes will be entangled with the qubit as $\frac{1}{\sqrt{2}}[|\psi_D\rangle|g\rangle + e^{-i\sum_m \theta_m a_m^\dagger a_m}|\psi_D\rangle|e\rangle]$, where $\theta_m = \chi_m t$. In principle, we could choose any time t, as long as none of the θ_m are integer multiples of 2π. However, in practice, we select t to make each of the θ_m as close to π modulo 2π as possible. This choice provides the maximum contrast and is closest to the ideal multimode parity operator. To complete our generalized parity measurement, we apply another $\pi/2$ pulse on the qubit, but with different phase from the initial one. Specifically, we consider the qubit rotation along $\vec{r} = -\sin\phi\,\vec{e}_x - \cos\phi\,\vec{e}_y$. By applying this $\exp(-i\frac{\pi}{4}\vec{r}\cdot\vec{\sigma})$ operation, the final cavity-qubit entangled state $|\Psi\rangle$ will be

$$|\Psi\rangle = \frac{|\psi_D\rangle + e^{i\phi}e^{-i\sum_m \theta_m a_m^\dagger a_m}|\psi_D\rangle}{2}|g\rangle$$
$$+ \frac{-e^{-i\phi}|\psi_D\rangle + e^{-i\sum_m \theta_m a_m^\dagger a_m}|\psi_D\rangle}{2}|e\rangle. \qquad (32)$$

Thus, when performing readout on the qubit, the final probability of achieving $|g\rangle$ will be

$$P_g = \frac{1 + \text{Re}\{e^{i\phi}\text{Tr}[D^\dagger(\vec{\alpha})|\psi\rangle\langle\psi|D(\vec{\alpha})e^{-i\sum_m \theta_m a_m^\dagger a_m}]\}}{2}$$
$$= \frac{1}{2}\{1 + \text{Re}[e^{i\phi}\tilde{W}_{|\psi\rangle\langle\psi|}(\vec{\alpha}, -\vec{\theta})]\}. \qquad (33)$$

Therefore, if we record $A = 1$ upon measuring $|g\rangle$ and $A = -1$ otherwise, the expectation value of A will be exactly $\text{Re}[e^{i\phi}\tilde{W}_\rho(\vec{\alpha}, -\vec{\theta})]$. This derivation applies for any ϕ, but as mentioned before, the choice of ϕ depends on the operator O and the sampling vector $\vec{\alpha}$.

We must modify the experimental protocol above to measure the generalized Wigner function for the projected state $\rho_{\bar{S}}$, which is defined in Sec. V. In particular, we utilize the second excited state $|f\rangle$ of the transmon to implement subsystem tomography and measure the Wigner values for only the projected states $\rho_{\bar{S}}$, which is similar to the idea used in [S 7]. Our W states are generated using multimode photon blockade as described in [S 8] to ensure that our maximum total photon number is $N = 1$. Consequently, for the density matrix reconstruction, the Hilbert space will be spanned by $\{|\vec{n}\rangle\,|\,\sum_{m=1}^M n_m \leq 1\}$. In this case, the projected operator $O_{\bar{S}}$ introduced in Eqn. (16) will be supported on

at most 2 modes. Because of the different dispersive couplings between the qubit and distinct cavity modes (STable 2), we can selectively target each of the modes with sufficiently narrow-bandwidth qubit pulses to help perform the necessary projections. Therefore, before the parity measurement, we first apply several long (narrow-bandwidth) qubit π pulses with frequencies $\omega_q + \chi_m$ for $m \in S$, such that the qubit that coupled with the $(I - P_S)|\psi\rangle$ component of the multimode cavity state will transfer from $|g\rangle$ to $|e\rangle$, while the component that coupled with $P_S|\psi\rangle$ will stay in $|g\rangle$. Then we give the transmon a short π pulse on the $|e\rangle-|f\rangle$ transition. After those steps, the cavity-transmon state becomes

$$|\Psi\rangle = P_S|\psi\rangle|g\rangle + (I - P_S)|\psi\rangle|f\rangle. \qquad (34)$$

Finally, we can use the procedure that we described before when focusing on the Wigner value measurement of a generic M-mode state ρ. We only need to drive those modes with index $m \in \bar{S}$ such that those modes are displaced by $D(-\vec{\alpha}_{\bar{S}})$. The probability to measure $|g\rangle$ from the final qubit readout is

$$P_{g,\phi} = \frac{\text{Tr}[\rho_{\bar{S}}] + \text{Re}[e^{i\phi}\tilde{W}_{\rho_{\bar{S}}}(\vec{\alpha}_{\bar{S}}, -\vec{\theta}_{\bar{S}})]}{2}, \qquad (35)$$

where $\rho_{\bar{S}} = \text{Tr}_S[P_S|\psi\rangle\langle\psi|]$. In practice, the result above is unaffected by the order in which we perform the qubit π pulses and cavity displacements, assuming that we add additional π pulses to target the displaced state. We found this order to work better in the experiment, despite the need for a larger comb of π pulse frequencies.

Now, if we proceed similarly to before and assign $A = 1$ for $|g\rangle$ and $A = -1$ otherwise, the expectation value of A will not give us our desired result. To solve this issue, we utilize the freedom we have in choosing the phase of the second qubit $\pi/2$ pulse in the parity measurement. If we choose the second qubit $\pi/2$ rotation to be along $-\vec{r}$ instead of $\vec{r}$ (or equivalently choose $(\phi + \pi)$ instead of ϕ, then the probability of measuring $|g\rangle$ will be

$$P_{g,(\phi+\pi)} = \frac{\text{Tr}[\rho_{\bar{S}}] - \text{Re}[e^{i\phi}\tilde{W}_{\rho_{\bar{S}}}(\vec{\alpha}_{\bar{S}}, -\vec{\theta}_{\bar{S}})]}{2}. \qquad (36)$$

Therefore, comparing with Eqn. (35), one solution to recovering our desired quantity is to first generate a random binomial number s. There is a 50% probability that $s = 1$ and 50% probability that $s = -1$. If we get $s = 1$, we choose the second qubit $\pi/2$ rotation to be along $\vec{r}$, and otherwise choose $-\vec{r}$ instead. In both cases, we assign $A = 1$ if the qubit measurement outcome is $|g\rangle$, and assign $A = -1$ if it is not $|g\rangle$. The expectation value of sA will be exactly $\text{Re}[e^{i\phi}\tilde{W}_{\rho_{\bar{S}}}(\vec{\alpha}_{\bar{S}}, -\vec{\theta}_{\bar{S}})]$. An advantage of this procedure is that it will work even when it is difficult to distinguish $|e\rangle$ and $|f\rangle$ levels in qubit readout, as long as we can distinguish $|g\rangle$ outcomes from others. The same applies for other permutations of these three
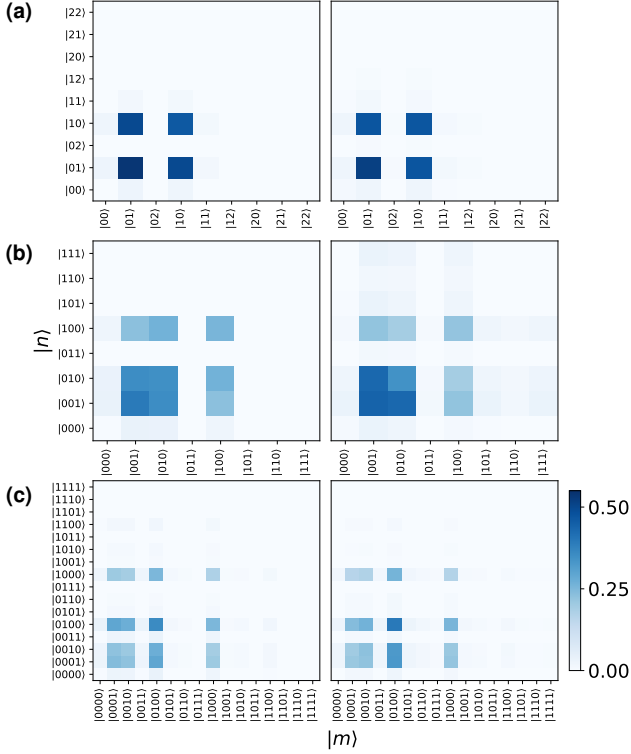
Figure S3 | Absolute value of the final density matrices determined using the DEMESST tomography sampling method (left column) and the OLI method (right column) for W states of (a) 2, (b) 3, and (c) 4 modes. The results for DEMESST and OLI are in good agreement.

states, as long as the experimental protocol is adjusted accordingly. In the actual experiment, we did not use this trick of random number s generation, since we can perform more than a single measurement per sampling vector $\vec{\alpha}_{\bar{S}}$. Instead, we repeated the experiment 10 times for rotation of the second $\pi/2$ along $\vec{r}$ and 10 times along $-\vec{r}$. Finally, we subtracted the averaged probability of measuring $|g\rangle$ between the two cases to obtain an estimate for $\mathrm{Re}[e^{i\phi}\tilde{W}_{\rho_{\bar{S}}}(\vec{\alpha}_{\bar{S}}, -\vec{\theta}_{\bar{S}})]$.

VII.   FINAL RECONSTRUCTED DENSITY MATRICES

Using the DEMESST and OLI tomography methods, we can reconstruct the final density matrices of our prepared W states. These matrices are final in the sense that they are the results obtained from the entire set of measurements performed in the experiment, i.e. 100 averages for each distinct displacement, so that the total measurement number is 10 times the maximum measurement number shown in Fig. 3 of the main text. Given a fixed total measurement number that equals the product of a number of distinct sets of cavity displacements

times the number of averages that each displacement is repeated, we would gain the most information from maximizing the number of distinct displacements and measuring each a single time. We choose instead to average each measurement 10 times due to our imperfect readout fidelity, to minimize our average number but still be able to obtain accurate measurement results. We then repeat this process 10 times to obtain statistics, resulting in a total of 100 averages for each distinct displacement. This choice lets us balance this theoretical maximal information of singleshot measurements with our measurement errors.

The final reconstructed density matrices for 2, 3, and 4 modes are shown in SFig. S3. We plot the absolute values of the density matrix elements so that we have a single matrix grid for each combination of tomography method and mode number. We can see that the two methods are in good agreement, with the largest visible deviation being in the 3-mode case for Fock basis elements with nonzero population in the second (middle) cavity mode. Nevertheless, the distances between the two final matrices as determined by the two methods is still low, and is 0.05 for the 2-mode case, 0.22 for the 3-mode case, and 0.30 for the 4-mode case. These distances are all below the corresponding minimum distances at the maximum total measurement number presented in the main text, and so this difference should not have significantly affected those results.

Besides some slight deviations in the measured populations of individual density matrix elements, the remaining distance between the final reconstructed matrices can be explained by small differences in the fit phase angles of the W states. For the 2-mode W state with form $|W_2\rangle = (|10\rangle + e^{i\phi}|01\rangle)/\sqrt{2}$, using the DEMESST method we obtain a fit $\phi_D = 0.04$, and using the OLI method we measure $\phi_O = 0.03$. In the 3-mode case, for W state with phase angles defined as $|W_3\rangle = (|100\rangle + e^{i\phi_1}|010\rangle + e^{i\phi_2}|001\rangle)/\sqrt{3}$, we measure $\phi_{1D} = -0.19$ and $\phi_{2D} = 1.57$, while $\phi_{1O} = -0.12$ and $\phi_{2O} = 1.57$. Similarly to what we see in the populations, the deviation is primarily in the middle mode. Finally, for 4 modes and W state $|W_4\rangle = (|1000\rangle + e^{i\phi_1}|0100\rangle + e^{i\phi_2}|0010\rangle + e^{i\phi_3}|0001\rangle)/\sqrt{4}$, we find $\phi_{1D} = -1.36, \phi_{2D} = -2.90$, and $\phi_{3D} = 0.60$, and $\phi_{1O} = -1.38, \phi_{2O} = -3.02$, and $\phi_{3O} = 0.63$. These angles are obtained by discretely sweeping the ϕ_j values over the full 2π range for each of the modes and determining which set of ϕ_j gives the largest fidelity when compared to an ideal W state with those phases.

VIII.   INFIDELITY AND MATRIX DISTANCE SIMULATIONS

In this section, we present simulations of the infidelity and Frobenius norm matrix distance vs. point number
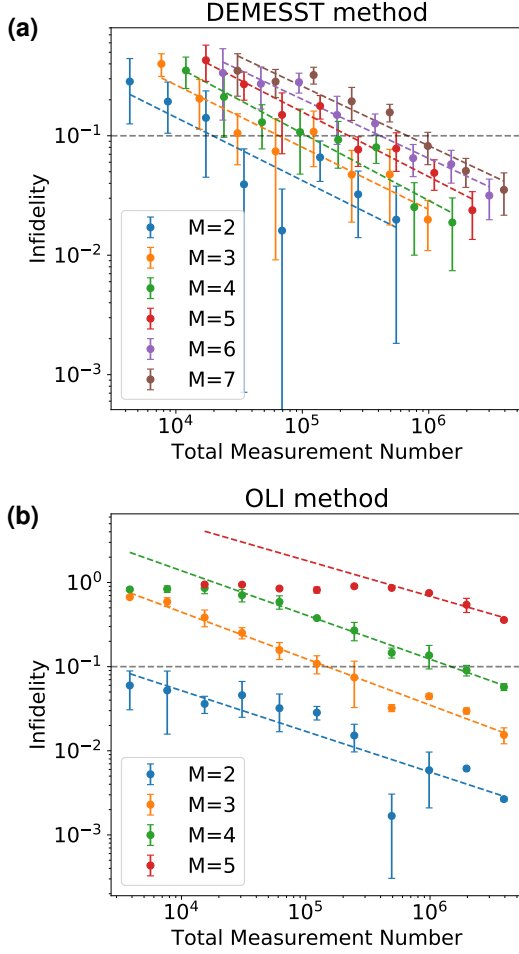
Figure S4 | Simulated infidelity vs. total measurement number for W states of different mode numbers for the two tomography methods, (a) DEMESST and (b) OLI. The infidelity is computed as the fidelity difference between the reconstructed state at a given total measurement number vs. an ideal M-mode W state. Error bars are obtained from repeating the simulation multiple times while including readout bit flip errors. The infidelities decrease to lower values more quickly for the DEMESST approach, especially for larger mode numbers. The dashed horizontal lines indicates an infidelity of 0.1 (90% fidelity).

Figure S5 | Simulated matrix distance vs. total measurement number for W states of different mode numbers for the two tomography methods, (a) DEMESST and (b) OLI. The distance is computed as the Frobenius norm between the reconstructed state at a given measurement number and a final simulated W state with state preparation errors from photon blockade and decoherences. Error bars are obtained from repeating the simulation multiple times while including readout bit flip errors. The distances decrease to lower values more quickly for the DEMESST approach, especially for larger mode numbers.

for our two tomography methods, DEMESST and OLI, following the same procedure described in the main text. Wigner tomography measurements are sampled assuming perfect state preparation, and the infidelities are computed with respect to an ideal M-mode W state, and the results are shown in SFig. S4. Error bars are obtained by repeating the simulations while modeling bit flip readout errors. The resulting infidelity vs. total measurement number for each M is fit to a power law, and the intersection with 0.1 (90% fidelity) is used to generate the values plotted in Fig. 1(a) in the main text. We can see that the OLI method requires fewer measurements than DEMESST for 2 modes, but DEMESST has a lower sam-
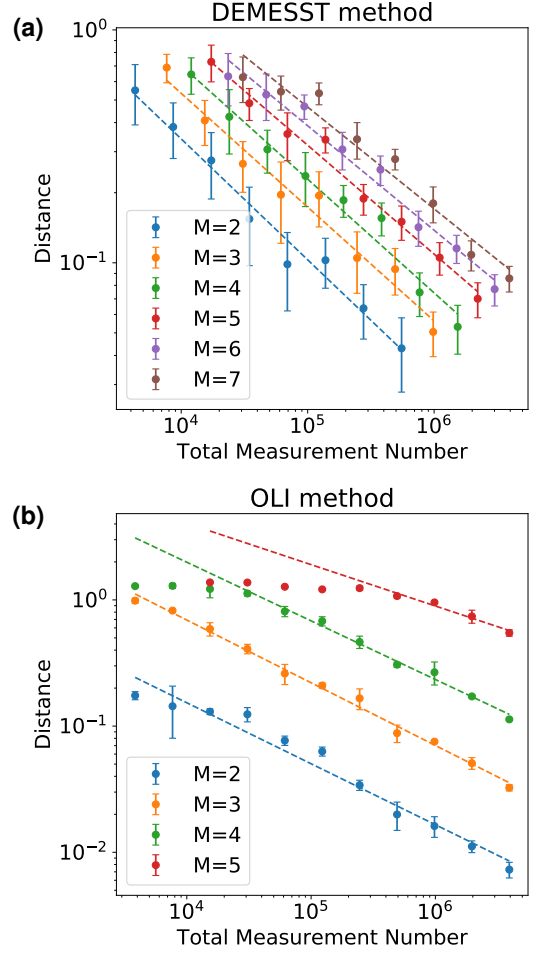
pling requirement for 3 or more modes. This effect becomes increasingly apparent for larger M. DEMESST scales polynomially with M, while OLI scales exponentially with M.

To compute the Frobenius norm matrix distances, we obtain the density matrix of an imperfect W state prepared by photon blockade, with errors from transmon and cavity decoherence and leakage through the blockade. We then simulate Wigner tomography measurements on that state for the total measurement numbers and using the same cavity displacement sampling points as used in our experiment, then reconstruct the density matrix while

including readout and bit flip error. These simulated results are shown in SFig. S5, with error bars obtained from repeating the simulations with the readout errors. As expected, the DEMESST method performs increasingly more efficiently as the mode number increases.

Comparing to the experimental data presented in the main text, the matrix distance results are similar, albeit with some differences. For example, we can see that the measurement number at which the simulated distance reaches roughly 0.1 for 2 modes is slightly less than 10^5, while we observe a distance slightly above 0.1 at that point number in our data. For the 3-mode case, in the experiment we observe a distance of roughly 0.3 at 10^5 measurements for DEMESST and 0.4 for OLI, which is close to the simulated distances of roughly 0.25 and 0.3, respectively. For the 4-mode case, we measure a distance of 0.4 at roughly 2×10^5 measurements for DEMESST, compared to a simulated distance of roughly 0.2, and a distance of roughly 1.0 vs 0.8 for OLI at that measurement number. We attribute the discrepancies to fluctuations over time in the readout error that may affect the accuracy of the simulation. This effect is particularly pronounced for the 4-mode case, where more measurements are required.

IX. W^2 STATE RECONSTRUCTION

Another Wigner tomography sampling method that we implement is the W^2 method, which was first introduced in [S 5] and that we briefly discussed in Sec. IV. In this approach, sets of coherent cavity displacements α_i are chosen using rejection sampling. This approach computes the overlap between a prepared state and a desired target state. For M modes, a cutoff c and a displacement vector $(\beta_1, ..., \beta_M)$ is randomly sampled from a uniform distribution between 0 and a maximum value of $|\mathcal{W}(\alpha_1, ..., \alpha_M)|^2 \prod_{i=1}^{M} |\alpha_i|$, where $\mathcal{W}$ corresponds to the target state. If $|\mathcal{W}(\beta_1, ..., \beta_M)|^2 \prod_{i=1}^{M} |\beta_i| > c$, the vector is kept. This ensures that we measure cavity displacements that provide the most information about the state, while also avoiding displacements with large magnitude or Wigner values near zero, which are more susceptible to experimental errors. After measuring a set of n of these vectors, the final overlap fidelity is computed as $\frac{1}{n} \sum_{i=1}^{n} \mathcal{W}_{\exp}(\vec{\alpha_i})/\mathcal{W}_{\text{ideal}}(\vec{\alpha_i})$. This approach allows for direct fidelity estimation of a prepared state with an ideal state. In particular, the W^2 method can be used in a similar manner to the DEMESST, where the fidelity estimation is performed with respect to multimode Fock state basis elements. Repeating for multiple elements can thus provide a reconstructed density matrix.

Experimentally, we use the W^2 method as an additional check on our prepared W states. We set the target state to be the multimode W state with ϕ's determined from the DEMESST and OLI methods. The W^2
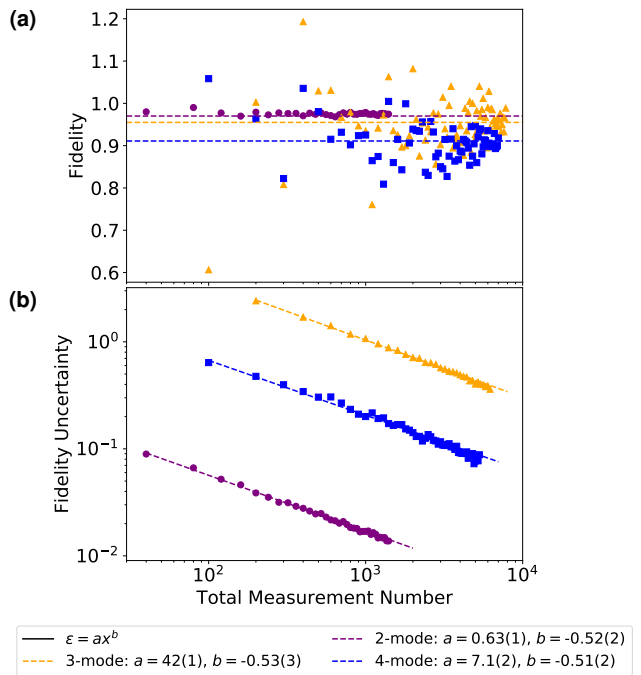

Figure S6 | Experimental results for the W^2 tomography sampling method. (a) State reconstruction fidelities for 2- (purple circles), 3- (orange triangles), and 4-mode (blue squares) W states. Horizontal lines indicate the fidelities obtained from the OLI method, which are consistent with the W^2. (b) Error magnitudes for the fidelities shown in (a). All results follow a roughly $1/\sqrt{x}$ relationship vs. total measurement number, as expected.

measurements then provide a direct fidelity estimation of the prepared state with the expected target. Since the sampling uses these angles, we present the W^2 results independently from the DEMESST and OLI, which do not utilize that information. The results are shown in SFig. S6. The fidelities for the maximum provided observation number are 0.972 ± 0.013, 0.95 ± 0.35, and 0.90 ± 0.08 for the 2-, 3-, and 4-mode W states respectively. These averages are consistent with the results of the previous DEMESST and OLI methods, with the OLI fidelities indicated by the horizontal lines in SFig. S6(a), and the data converges quickly to the expected fidelity obtained from those two approaches, although with large uncertainties, as shown in SFig. S6(b). One reason for these errors is the relatively low total measurement number compared with the other methods. However, an odd behavior is that the 3-mode data has much greater uncertainties than even the 4-mode case, when we would expect the uncertainties to increase monotonically with mode number. Some possible explanations for this behavior could be a particularly low readout fidelity during data collection or fluctuations in drive strength during the measurement sequence that modify the effective Wigner operator differently for distinct sets of cavity

mode displacements. This could also be caused by the choice in cutoff, as derived in Sec. IV. All the uncertainties have the expected $1/\sqrt{x}$ with total observation number.

------

[1] Chakram, S. *et al.* Seamless high-q microwave cavities for multimode circuit quantum electrodynamics. *Physical Review Letters* **127**, 107701 (2021).

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

[3] Cahill, K. E. & Glauber, R. J. Density Operators and Quasiprobability Distributions. *Physical Review* **177**, 1882–1902 (1969). URL https://link.aps.org/doi/10.1103/PhysRev.177.1882.

[4] Cahill, K. E. & Glauber, R. J. Ordered Expansions in Boson Amplitude Operators. *Physical Review* **177**, 1857–1881 (1969). URL https://link.aps.org/doi/10.1103/PhysRev.177.1857.

[5] da Silva, M. P., Landon-Cardinal, O. & Poulin, D. Practical Characterization of Quantum Devices without Tomography. *Physical Review Letters* **107**, 210404 (2011). URL https://link.aps.org/doi/10.1103/PhysRevLett.107.210404.

[6] Flammia, S. T. & Liu, Y.-K. Direct Fidelity Estimation from Few Pauli Measurements. *Physical Review Letters* **106**, 230501 (2011). URL https://link.aps.org/doi/10.1103/PhysRevLett.106.230501.

[7] Gertler, J. M., van Geldern, S., Shirol, S., Jiang, L. & Wang, C. Experimental Realization and Characterization of Stabilized Pair-Coherent States. *PRX Quantum* **4**, 020319 (2023). URL https://link.aps.org/doi/10.1103/PRXQuantum.4.020319.

[8] Chakram, S. *et al.* Multimode photon blockade. *Nature Physics* **18**, 879–884 (2022). URL https://www.nature.com/articles/s41567-022-01630-y.