Non-classical microwave-optical photon pair generation with a chip-scale transducer

Srujan Meesala,[1, 2, *] Steven Wood,[1, 2, *] David Lake,[1, 2, *] Piero Chiappina,[1, 2] Changchun Zhong,[3] Andrew D. Beyer,[4] Matthew D. Shaw,[4] Liang Jiang,[3] and Oskar Painter[1, 2, 5, †]

[1] *Kavli Nanoscience Institute and Thomas J. Watson, Sr., Laboratory of Applied Physics, California Institute of Technology, Pasadena, California 91125, USA*

[2] *Institute for Quantum Information and Matter, California Institute of Technology, Pasadena, California 91125, USA*

[3] *Pritzker School of Molecular Engineering, The University of Chicago, Chicago, IL 60637, USA*

[4] *Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Dr, Pasadena, California 91109, USA*

[5] *Center for Quantum Computing, Amazon Web Services, Pasadena, California 91125, USA*

(Dated: April 3, 2023)

Modern computing and communication technologies such as supercomputers and the internet are based on optically connected networks of microwave frequency information processors [1, 2]. In recent years, an analogous architecture has emerged for quantum networks with optically distributed entanglement [3, 4] between remote superconducting quantum processors, a leading platform for quantum computing [5–7]. Here we report an important milestone towards such networks by observing non-classical correlations between photons in an optical link and a superconducting electrical circuit. We generate such states of light through a spontaneous parametric down-conversion (SPDC) process in a chip-scale piezo-optomechanical transducer [8]. The non-classical nature of the emitted light is verified by observing anti-bunching in the microwave state conditioned on detection of an optical photon. Such a transducer can be readily connected to a superconducting quantum processor, and serve as a key building block for optical quantum networks of microwave frequency qubits.

Networks of remotely situated qubits [3, 4] are essential to harness quantum correlations for long-distance secure communication [9, 10], distributed quantum computation [11, 12] and precision measurements [13, 14]. Optical photons are naturally suited to act as flying qubits over room temperature links and distribute entanglement in such networks [15]. Quantum optical networks with few nodes have been realized with systems such as atoms [16, 17], quantum dots [18, 19], trapped ions [20], and color centers [21], which naturally possess optical frequency transitions between their internal energy levels. In parallel developments, superconducting circuits based on Josephson junctions have emerged as a leading platform for quantum information processing with the ability to realize entangled states of many qubits in microwave frequency circuits [5–7]. However, superconducting qubits do not possess a natural, coherent interface with optical photons. This limitation has motivated recent efforts to develop transducers capable of generating quantum correlations between optical photons and microwave frequency qubits. While schemes to produce such states are fundamentally well-understood, preserving fragile quantum correlations during the transduction process has not been possible so far in a wide variety of physical platforms [22–27]. This roadblock is primarily due to technical challenges posed by the vast difference in energy scales between microwave and optical photons.

Here we demonstrate non-classical microwave-optical photon pairs from such a transducer. We use a piezo-optomechanical device in which an acoustic mode acts as an intermediary between microwave and optical fields. First, a pump laser pulse generates photon-phonon pairs in an optomechanical cavity via an SPDC process. Subsequently, a strong piezoelectric interaction converts the phonon into a microwave photon. To characterize the photon pairs emitted by the transducer, we simultaneously perform single photon detection of the optical component and heterodyne detection of the microwave component. For experimental trials in which single optical photons are detected, we compute the normalized second order intensity correlation function of the conditional microwave state. Observation of a value below unity for this function constitutes direct verification of non-classical statistics of the photon pairs. Compared with a recent demonstration of microwave-optical two-mode squeezing [28], our experiment demonstrates the requisite techniques to implement the well-known DLCZ protocol [10], which uses detection of optical photons to herald entanglement between distant nodes in quantum networks.

Our result is primarily enabled by fabricating nearly all circuit components of a chip-scale piezo-optomechanical transducer from niobium nitride (NbN), a superconductor in which quasiparticles (QPs) generated by optical absorption relax on the timescale of a few nanoseconds [29]. Compared with our previous work in which we directly coupled such a transducer to a transmon qubit on the same chip [30], using a circuit with fast QP relaxation

* These authors contributed equally

† opainter@caltech.edu; http://copilot.caltech.edu

allows us to mitigate the effects of parasitic absorption of stray optical pump light by superconducting components of the transducer. In particular, we add well below one quantum of noise to the circuit during transduction while increasing the rate of probabilistic optical detection events by two orders of magnitude. While superconductors with fast QP relaxation have been employed in piezo-optomechanical transducers recently [31, 32], these demonstrations were limited to classical states of light. Our device design and optical detection setup provide the efficiency and noise levels required to observe non-classical correlations.

Figure 1a shows a conceptual schematic highlighting the resonator modes involved in our transduction experiment. Interaction between an optical mode, $\hat{a}$ and a microwave-frequency acoustic mode, $\hat{b}$ is mediated by a pump laser driving an optomechanical cavity in the resolved sideband regime. Simultaneously, the acoustic mode is resonantly coupled to a microwave-frequency electrical mode, $\hat{c}$ via the piezoelectric effect. We can write the Hamiltonian for this system as

$$\hat{H}/\hbar = -\Delta_a \hat{a}^\dagger \hat{a} + \omega_b \hat{b}^\dagger \hat{b} + \omega_c \hat{c}^\dagger \hat{c} + \hat{H}_{\mathrm{om}}/\hbar + \hat{H}_{\mathrm{pe}}/\hbar \quad (1)$$

Here ω_b, ω_c are the frequencies of the modes $\hat{b}, \hat{c}$ respectively and $\hat{H}_{\mathrm{om}}, \hat{H}_{\mathrm{pe}}$ are the optomechanical and piezoelectric interaction Hamiltonians described in more detail below. $\Delta_a = \omega_{\mathrm{p}} - \omega_a$ is the difference between the frequency of the optical pump, ω_{p} and that of the optical mode, ω_a. Setting the frequency of the pump laser to be red ($\Delta_a < 0$) or blue detuned ($\Delta_a > 0$) with respect to the optical cavity resonance allows us to select either beam-splitter or two-mode squeezing interactions, respectively [33]. The first setting can be used to transfer states between the acoustic mode and the optical mode when the transducer is operated as a frequency converter [30, 32, 34]. In this work, we use the latter setting to generate non-classical pairs of optical photons and acoustic phonons in an SPDC process. This choice is motivated by recent proposals for heralded remote entanglement generation which indicate that operation in SPDC mode relaxes the efficiency requirements for piezo-optomechanical transducers [8, 35, 36]. In this setting, we have $\hat{H}_{\mathrm{om}}/\hbar = -G_{\mathrm{om}}(t)(\hat{a}^\dagger \hat{b}^\dagger + \hat{a}\hat{b})$. The time-dependent optomechanical coupling rate $G_{\mathrm{om}}(t) = \sqrt{n_a(t)} g_{\mathrm{om}}$ is controlled parametrically via the intracavity photon population $n_a(t)$ due to the detuned pump laser. Here g_{om} denotes the optomechanical coupling rate at the single optical photon and acoustic phonon level. The piezoelectric interaction is described by the beam-splitter Hamiltonian $\hat{H}_{\mathrm{pe}}/\hbar = -g_{\mathrm{pe}}(\hat{b}^\dagger \hat{c} + \hat{b}\hat{c}^\dagger)$. Here g_{pe} denotes the piezoelectric coupling rate at the single microwave photon and acoustic phonon level. This interaction can be used to map the acoustic component of the optomechanical two-mode squeezed state onto the microwave electrical mode. In the absence of any added noise, the joint state of the modes, $\hat{a}, \hat{c}$ can be described in the photon number basis by the wavefunction $|\psi\rangle = |00\rangle + \sqrt{p}\,|11\rangle + p\,|22\rangle + O(p^{3/2})$. For a weak pump field, the higher order terms may be neglected, and the transducer emits single optical and microwave photons in pairs with probability, $p \ll 1$. Detection of a single optical photon can then be used to conditionally prepare or herald a single microwave photon.

Figure 1b shows the physical schematic of our device, which consists of a half-wavelength superconducting kinetic inductance resonator [37] coupled to a piezo-optomechanical transducer. The transducer itself comprises a half-wavelength aluminum nitride (AlN) piezo-acoustic cavity attached to a silicon optomechanical crystal (OMC) resonator via an acoustic waveguide [30]. We achieve piezoelectric coupling in our system by using the two end terminals of the microwave resonator as electrical leads over the AlN section of the transducer. The microwave resonator is patterned in a disordered, thin film of NbN in a meandering ladder geometry. The inclusion of closed loops in the thin superconducting film, as shown in Fig. 1c, allows for tuning of kinetic inductance, and hence tuning of the frequency of the microwave resonator via an external magnetic field [38]. Using narrow superconducting wires in a meandering geometry results in high impedance, which helps us achieve strong piezoelectric coupling, g_{pe} with a small acoustic mode volume [39]. This design strategy is followed to maximize the fraction of acoustic energy in silicon, which has the lowest acoustic loss in our material stack. The NbN resonator is capacitively coupled to a 50Ω transmission line (not shown in schematic) to facilitate microwave spectroscopy of the transducer. Likewise, the optical cavity is coupled to a waveguide, which terminates in a tapered coupler at the edge of the chip allowing for efficient coupling to a lensed optical fiber. The layout of our device is chosen to reduce optical flux from stray pump light at the microwave resonator. We use an $\sim$1 mm long optical waveguide to physically separate the circuit section of the transducer from the optical coupler, where there is significant scattering of pump light. Additionally, we use extended electrical terminals to physically separate the optically sensitive current anti-node of the kinetic inductance resonator from the OMC, and reduce the impact of local pump scattering.

Our experiments are carried out by mounting the transducer chip on the mixing plate of a dilution refrigerator. We initially perform optical and microwave spectroscopy to identify the frequencies of the internal transducer modes as well as the optomechanical and piezoelectric coupling rates. For the device used in the experiments that follow, we found an optical resonance at a wavelength $\lambda = 1561.3$ nm with critical coupling to the external waveguide, $\kappa_{e,a}/2\pi = \kappa_{i,a}/2\pi = 650$ MHz. The subscripts i, e refer to linewidths due to coupling to internal and external baths, respectively. We identify the hybridized microwave-frequency electrical and acoustic modes supported by the integrated transducer-resonator system in Fig. 2a. In this measurement, we use a vector network analyzer (VNA) to electrically excite the mi-

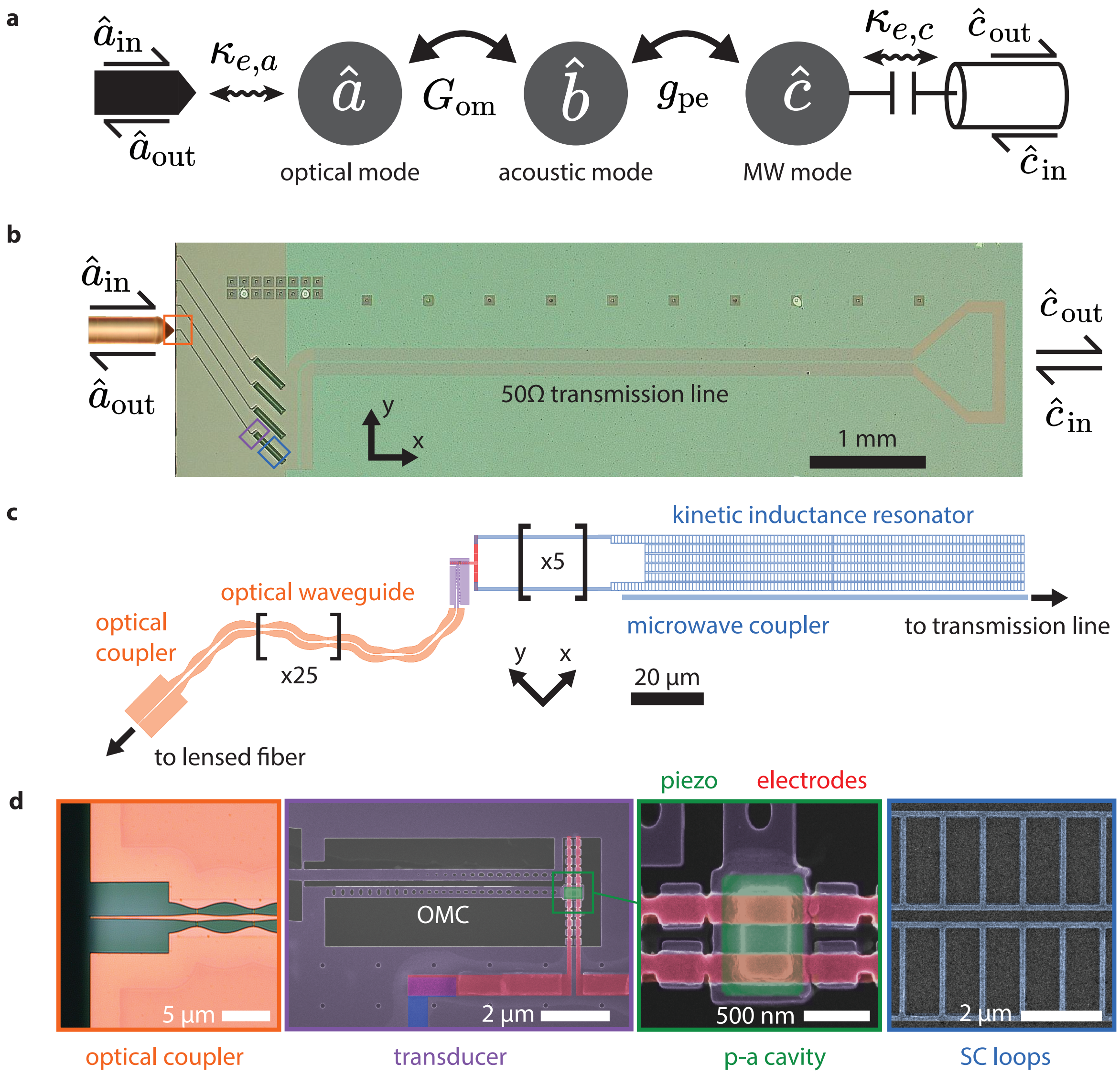

FIG. 1. **Quantum transducer**. **a.** Transducer mode schematic indicating optical ($\hat{a}$), acoustic ($\hat{b}$) and microwave ($\hat{c}$) modes along with interaction rates for optomechanical coupling (G_{om}) and piezoelectric coupling (g_{pe}), respectively. The optical and microwave modes are coupled to waveguides with external coupling rates, $\kappa_{e,a}$ and $\kappa_{e,c}$, respectively. The input and output modes in the optical and microwave waveguides are denoted by $\hat{a}_{\text{in}}, \hat{a}_{\text{out}}, \hat{c}_{\text{in}}, \hat{c}_{\text{out}}$, respectively. **b.** Micrograph of transducer device showing optical access via lensed fiber on the left and microwave access via 50Ω transmission line on the right. **c.** Device schematic. From left to right, suspended optical waveguide (orange) leading to the piezo-optomechanical transducer (purple) whose electrical terminals (red) are connected to a microwave kinetic inductance resonator (blue). **d.** Micrographs of various components of the transducer. From left to right: optical micrograph of the coupler section at the end of the optical waveguide; scanning electron micrograph of the transducer indicating the silicon optomechanical crystal (OMC) cavity in purple; close-up of the piezo-acoustic (p-a) cavity highlighting the piezoelectric material in green and electrodes in red; scanning electron micrograph of the superconducting (SC) loops in the meandering ladder trace of the kinetic inductance resonator.

crowave resonator via the microwave input port while optically pumping the optomechanical cavity of the transducer with a laser tuned to the blue side of the optical resonance. The optical pump reflected from the OMC and an optical sideband generated due to transduction of the input microwave signal are together sent to a high-speed photodetector whose output is connected to the detection port of the VNA. The magnitude of signal generated in this VNA spectrum by each of the hybridized microwave-frequency modes of the transducer-resonator system is proportional to the transduction efficiency of the corresponding mode. For the mode at a frequency

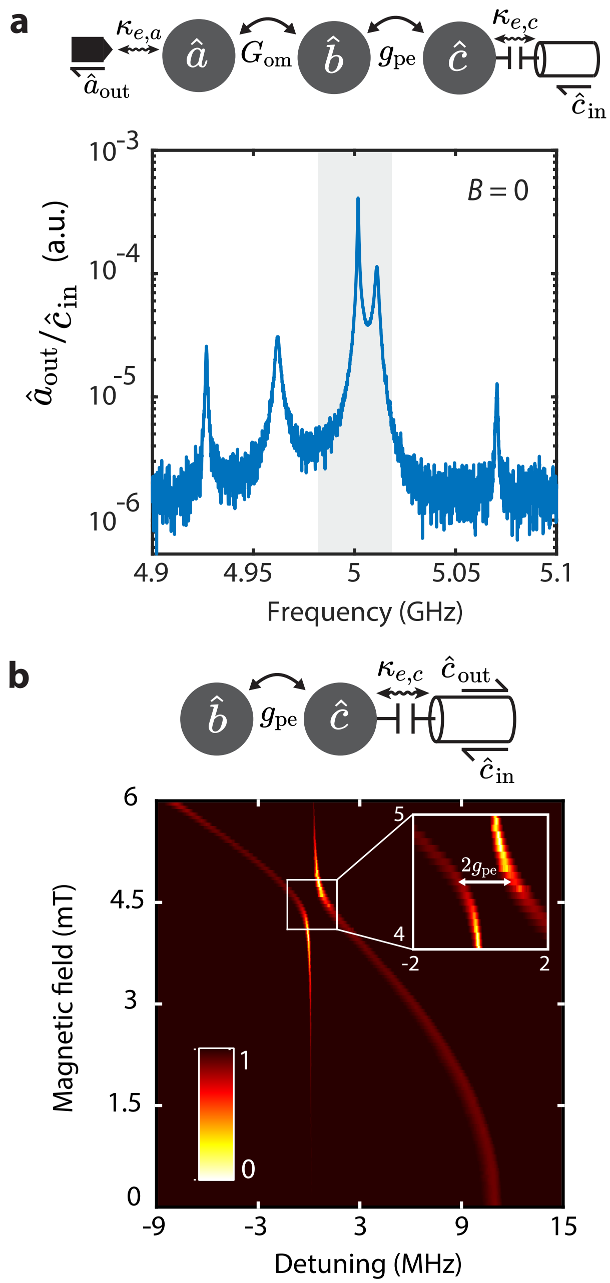

FIG. 2. **Optical and microwave spectroscopy**. **a.** Continuous wave transduction spectrum measured using a vector network analyzer (VNA) at zero magnetic field. The microwave resonator is excited via the input mode, $\hat{c}_{\rm in}$, and the optical output mode, $\hat{a}_{\rm out}$ is electrically detected via its microwave frequency beat note with the optical pump on a high speed photodetector. **b.** Microwave reflection spectrum of the transducer in the frequency range shown by the gray shaded window in panel **a** probed as a function of external magnetic field. The horizontal axis refers to detuning from the acoustic mode at zero magnetic field. Inset shows a close-up of the anti-crossing between the microwave-frequency electrical and acoustic modes revealing a minimum mode splitting, $2g_{\rm pe}/2\pi = 1.6$MHz.

of 5.001GHz with the highest transduction efficiency and narrowest linewidth, we perform pump power dependent optomechanical spectroscopy and measure an optomechanical coupling rate, $g_{\rm om}/2\pi = 270$ kHz. In Fig. 2b, we show the corresponding microwave reflection spectrum of this mode as we sweep a magnetic field applied perpendicular to the sample. We observe an anti-crossing between the microwave electrical resonator mode, which is identified through its characteristic quadratic tuning response in an external magnetic field [39], and the acoustic mode of the transducer. The minimum frequency splitting between these two modes allows us to estimate the piezoelectric coupling rate, $g_{\rm pe}/2\pi = 800$ kHz. The independent linewidths of the modes are measured far from the anti-crossing and are found to be $\kappa_{i,b}/2\pi = 150$ kHz for the microwave acoustic mode and $\kappa_{e,c}/2\pi = 1.2$ MHz, $\kappa_{i,c}/2\pi = 550$ kHz for the microwave electrical resonator mode. For the experiments that follow, we set the external magnetic field at the value corresponding to the minimum mode splitting of $2g_{\rm pe}/2\pi = 1.6$ MHz where both modes are maximally hybridized. This corresponds to the condition, $\omega_b = \omega_c$ in Eq. (1). In this setting, we define the hybridized electromechanical modes, $\hat{c}_\pm = (\hat{b} \pm \hat{c})/\sqrt{2}$ with frequencies, $\omega_\pm = \omega_c \pm g_{\rm pe}$ respectively. Even though the transducer supports other microwave-frequency acoustic modes as shown in Fig. 2a, these are far detuned from the modes of interest, $\hat{c}_\pm$, relative to the coupling rates, $g_{\rm pe}$ and $G_{\rm om}$. As a result, we expect the Hamiltonian in Eq. (1) to provide a sufficiently accurate description of our system.

We operate the transducer in SPDC mode by exciting it with optical pump pulses at the blue optomechanical sideband of the optical cavity ($\Delta_{\rm a} = (\omega_+ + \omega_-)/2$). In the experiments described below, Gaussian pump pulses of full-width at half-maximum (FWHM) duration, $T_{\rm p} = 160$ ns with a peak power corresponding to intra-cavity optical photon occupation, $n_a = 0.8$ were sent to the device at a repetition rate of 50 kHz. Each pulse represents an experimental trial with a finite probability of generating a microwave-optical photon pair. Figure 3a shows a schematic of the setup used to characterize the optical and microwave emission from the transducer under these conditions. The optical emission in the mode, $\hat{a}_{\rm out}$, is sent to a superconducting nanowire single photon detector (SNSPD) after passing through a Fabry-Perot filter setup to suppress the pump pulses reflected by the transducer. In Fig. 3b, we show the optical photon flux at the SNSPD obtained by averaging single photon 'clicks' over multiple trials of the experiment. Gating the detection events in a time window of duration, $2T_{\rm p} = 320$ ns shown by the gray shaded region, we obtain a click probability, $p_{\rm click} = 2.7 \times 10^{-6}$, which leads to an optical heralding rate, $R_{\rm click} = 0.14$ Hz. A detailed account of various contributions to this heralding rate is provided in Ref. [39].

In parallel with optical photon detection, the microwave emission from the transducer in the output mode, $\hat{c}_{\rm out}$ is sent to a heterodyne detection setup as shown in Fig. 3a. Using the input-output formalism for

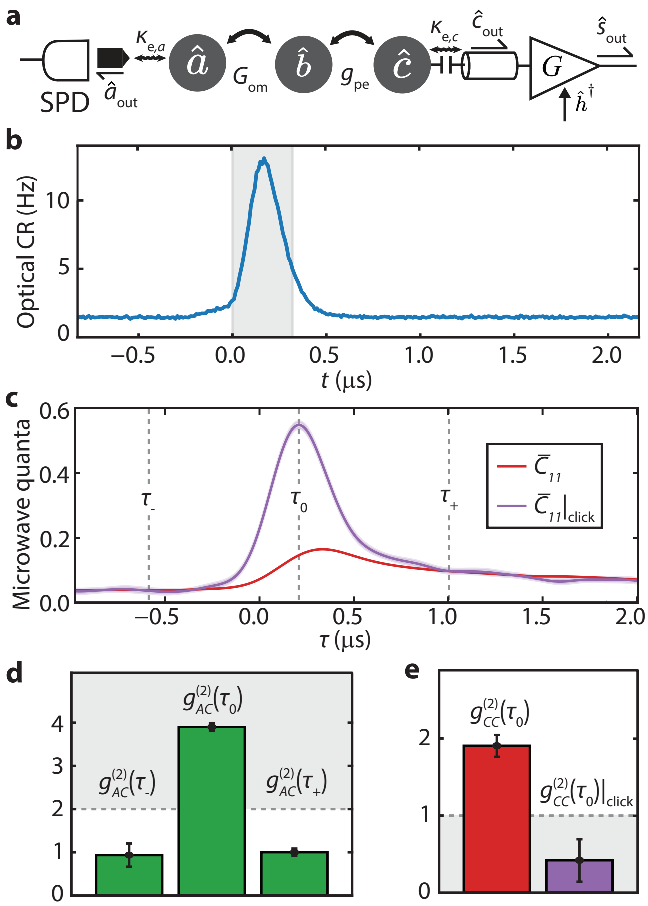

FIG. 3. **Microwave-optical cross-correlations. a.** Schematic of the microwave-optical cross-correlation measurement. The optical output is directed to a single photon detector (SPD) and the microwave output is directed to a heterodyne detection setup with gain, G and added noise, $\hat{h}^\dagger$. **b.** Time trace of average optical photon count rate (CR) registered on the SPD. Shaded vertical window indicates the gate duration defining the temporal mode, $\hat{A}$ used to perform conditional microwave readout. **c.** Microwave quanta in the temporal mode, $\hat{C}$ at the transducer output port as a function of delay, τ the from the center of the optical gating window. Red trace corresponds to unconditional microwave readout and purple trace corresponds to microwave readout conditioned on an optical click. Shaded region about the trace indicates a confidence interval spanning two standard deviations about the mean. **d.** Normalized microwave-optical intensity cross-correlation function at delays τ_0, $\tau_\pm$ indicated in panel **b**. Dashed line indicates expected classical upper bound for thermal states. **e.** Normalized second order intensity correlation function of the unconditional microwave state (red) and the microwave state conditioned an optical click (purple). Dashed line indicates classical lower bound.

a phase-insensitive amplifier [40] in the limit $G \gg 1$, we write the output of this setup as $\hat{s}_{\mathrm{out}} \approx \sqrt{G}(\hat{c}_{\mathrm{out}} + \hat{h}^\dagger)$, where $\hat{h}^\dagger$ is the noise mode added by the amplifier. Emission from the transducer modes $\hat{c}_\pm$ is concentrated in a small bandwidth about the frequencies $\omega_\pm$ within $\hat{s}_{\mathrm{out}}$. We isolate this signal by integrating the recorded heterodyne voltage signal with a matched emission envelope function $f(t)$ as detailed in the supplementary information [39]. This corresponds to a measurement of the quadratures of the temporal mode $\hat{S}(t) := \int \hat{s}_{\mathrm{out}}(t+t')f^*(t')dt'$ at the output of the heterodyne detection setup. Similarly, we define $\hat{C}(t) := \int \hat{c}_{\mathrm{out}}(t+t')f^*(t')dt'$ and $\hat{H}(t) := \int \hat{h}_{\mathrm{out}}(t+t')f^*(t')dt'$ to be temporal modes corresponding to emission referred to the output port of the transducer device and amplifier noise. As a consequence of the large bandwidth of the pump pulse compared to the mode splitting ($\eta_{\mathrm{TB}}/T_{\mathrm{p}} = 2.75$ MHz $> 2g_{\mathrm{pe}}/2\pi = 1.6$ MHz, where $\eta_{\mathrm{TB}} = 0.44$ is the time-bandwidth product of a Gaussian pulse) we cannot individually resolve emission coming from $\hat{c}_\pm$. With this in mind, $f(t)$ is constructed to capture emission from both $\hat{c}_\pm$, and consequently the temporal modes have large spectral overlap with both hybridized modes $\hat{c}_\pm$. By measuring the heterodyne voltage signal in experimental trials carried in the presence (absence) of optical pump pulses, we collect complex-valued voltage samples of $\hat{S}$ ($\hat{H}^\dagger$). These voltage samples are then used to calculate the moments of the microwave field, $\bar{C}_{mn} = \langle \hat{C}^{\dagger m}\hat{C}^n\rangle$ by taking an ensemble average over the experimental trials and inverting the amplifier input-output relations [39, 41]. For $m = n = 1$, we obtain the microwave intensity shown by the red time trace in Fig. 3c. Since the SPDC excitation probability in our experiment is far less than unity, this signal is nearly entirely noise added from heating of the hybridized electromechanical modes due to parasitic absorption of pump light. Further, the decay of this added noise exhibits fast and slow components relative to the repetition rate of our experiment. In particular, the slow component is responsible for the non-zero microwave intensity prior to the optical pulse. Detailed measurements of the heating dynamics of the transducer are presented in Ref. [39].

To measure microwave-optical cross-correlations, we perform conditional microwave readout by triggering the heterodyne measurement based on the occurrence of an optical click in an experimental trial. Following the same inversion process used for the unconditional microwave field, we can obtain the moments of the microwave field conditioned on optical detection, $\bar{C}_{mn}|_{\mathrm{click}} = \langle \hat{A}^\dagger \hat{C}^{\dagger m}\hat{C}^n\hat{A}\rangle/\langle \hat{A}^\dagger\hat{A}\rangle$. Here, $\hat{A}$ refers to the temporal mode defined by gating the optical waveguide mode, $\hat{a}_{\mathrm{out}}$ in the time window indicated by the gray shaded region in Fig. 3b. The result for $m = n = 1$ is shown by the purple time trace in Fig. 3c. We observe that detection of an optical photon is correlated with substantially higher microwave intensity than that of the unconditional state. The temporal shape of the conditional signal is in good agreement with the result of a numerical simulation of our system [39].

Dividing the conditional and unconditional microwave intensity traces recorded in Fig. 3c, we obtain the normalized microwave-optical intensity cross-correlation function,

$$g^{(2)}_{AC}(\tau) = \frac{\langle \hat{A}^\dagger \hat{C}^\dagger(\tau)\hat{C}(\tau)\hat{A}\rangle}{\langle \hat{A}^\dagger \hat{A}\rangle\langle \hat{C}^\dagger(\tau)\hat{C}(\tau)\rangle} \tag{2}$$

In Fig. 3d, we plot this function sampled at three representative time delays as indicated by vertical dashed lines in Fig. 3c. τ_{o} is the delay corresponding to the maximum conditional microwave intensity, $\bar{C}_{11}|_{\mathrm{click}}$, and $\tau_\pm = \tau_{\mathrm{o}} \pm 800$ ns are offset from τ_{o} in opposite directions by five times the FWHM duration of the optical pump pulse. We measure $g^{(2)}_{AC}(\tau_o) = 3.90^{+0.093}_{-0.093}$, indicating strongly correlated microwave and optical emission at this time delay. The error bars for this observation and subsequent correlation functions referred to in the text are determined via a bootstrapping procedure over the dataset of heterodyne voltage samples [39], and represent a confidence interval spanning two standard deviations about the mean. We observe that the microwave-optical correlations disappear at times well before and after the optical pump pulse as evinced by the near-unity values of $g^{(2)}_{AC}(\tau_+) = 1.00^{+0.08}_{-0.08}$ and $g^{(2)}_{AC}(\tau_-) = 0.94^{+0.27}_{-0.27}$.

For classical microwave-optical states, $g^{(2)}_{AC}$ is bounded by a Cauchy-Schwarz inequality, $g^{(2)}_{AC} \leq \sqrt{g^{(2)}_{AA} g^{(2)}_{CC}}$ [42–44]. Here, $g^{(2)}_{AA}$ and $g^{(2)}_{CC}$ are the normalized intensity autocorrelation functions of the unconditional optical and microwave temporal modes, $\hat{A}$ and $\hat{C}$ respectively, and are defined in a manner similar to Eq. (2). Using the moment inversion procedure with the unconditional microwave voltage samples, we have measured $g^{(2)}_{CC} = \bar{C}_{22}/(\bar{C}_{11})^2 = 1.91^{+0.14}_{-0.14}$ at $\tau = \tau_{\mathrm{o}}$. This is consistent with the theoretically expected value of 2 for a thermal state. An explicit measurement of $g^{(2)}_{AA}$ with our current device is impractical given the low coincidence rate expected in a Hanbury-Brown-Twiss measurement. In principle, since optomechanical scattering from an acoustic mode in a thermal state is expected to produce $g^{(2)}_{AA} = 2$, we expect the classical upper bound, $g^{(2)}_{AC} \leq 2$. Our observation that $g^{(2)}_{AC}(\tau_o)$ exceeds this classical bound for thermal states by over twenty standard deviations serves as a promising signature of non-classical statistics of the microwave-optical states. Further, by performing this cross-correlation experiment with increasing pump power, which is accompanied by increasing thermal noise in the modes, we observe that $g^{(2)}_{AC}(\tau_{\mathrm{o}})$ monotonically approaches the value of 2 [39].

To unambiguously verify the non-classical nature of the microwave-optical photon pairs, we measure the normalized second order intensity correlation function of the microwave state conditioned on an optical click, $g^{(2)}_{CC}|_{\mathrm{click}} = \bar{C}_{22}|_{\mathrm{click}}/(\bar{C}_{11}|_{\mathrm{click}})^2$. For noiseless SPDC in the weak pump regime, we expect $g^{(2)}_{CC}|_{\mathrm{click}} = 0$ with the detection of an optical photon heralding a pure single photon in the microwave mode. In practice, the value of $g^{(2)}_{CC}|_{\mathrm{click}}$ will be higher due to noise added during transduction. Classical microwave-optical states are expected to satisfy the inequality, $g^{(2)}_{CC}|_{\mathrm{click}} \geq 1$ [39]. Violation of this inequality signifies the observation of photon anti-bunching in the conditionally prepared microwave state. With the conditional heterodyne voltage samples collected in our experiment, we observe $g^{(2)}_{CC}(\tau_{\mathrm{o}})|_{\mathrm{click}} = 0.42^{+0.27}_{-0.28}$. As shown in Fig. 3e, this observation is below the classical bound of unity by 2.1 standard deviations. This corresponds to a probability of 1.7% for the null hypothesis of conditional preparation of a classical microwave state. Details of the data analysis and the underlying probability distribution of $g^{(2)}_{CC}(\tau_o)|_{\mathrm{click}}$ are provided in the supplementary information [39]. We emphasize that all the normalized correlation functions referred to in the text are, by definition, independent of the gain, G of the amplification chain in the microwave heterodyne measurement. As a result, even though the values of the moments, $\bar{C}_{mn}$ and $\bar{C}_{mn}|_{\mathrm{click}}$ depend on the absolute accuracy of the calibrated gain, our inferences of non-classical statistics are based on normalized correlation functions, which do not require this calibration.

The non-classical photon pairs generated from our transducer can be used to entangle distant quantum circuits by adopting the DLCZ protocol [10]. Following this scheme, interference of optical photons from two distant transducers and subsequent single photon detection can be used to herald entanglement between remote microwave nodes [45]. In this work, by conditionally preparing non-classical microwave states using detection of optical photons, we demonstrate both a capable transducer device and key experimental techniques for such a remote entanglement scheme. Towards entanglement of remote quantum processors, the microwave emission from our transducer chip can be routed and efficiently absorbed in a superconducting qubit [46] residing in a separate processor module [23]. The use of a stronger piezoelectric material such as lithium niobate in our transducer [31, 32] would improve the electromechanical conversion efficiency, and thereby, mitigate the primary microwave loss channel in our experiment. Transducer heating due to optical absorption can be reduced through better thermalization of the acoustic mode with the substrate [47], thereby improving the fidelity of the photon pair for a given emission rate. Finally, the strongly hybridized electromechanical modes in our transducer provide two frequency bins to encode a qubit in both the optical and microwave emission and can allow the generation of microwave-optical Bell states [8].

[1] Q. Cheng, M. Bahadori, M. Glick, S. Rumley, and K. Bergman, Recent advances in optical technologies for data centers: a review, Optica **5**, 1354 (2018).

[2] C. A. Thraskias, E. N. Lallas, N. Neumann, L. Schares, B. J. Offrein, R. Henker, D. Plettemeier, F. Ellinger, J. Leuthold, and I. Tomkos, Survey of photonic and plasmonic interconnect technologies for intra-datacenter and high-performance computing communications, IEEE Communications Surveys & Tutorials **20**, 2758 (2018).

[3] J. I. Cirac, P. Zoller, H. J. Kimble, and H. Mabuchi, Quantum state transfer and entanglement distribution among distant nodes in a quantum network, Phys. Rev. Lett. **78**, 3221 (1997).

[4] H. Kimble, The quantum internet, Nature **453**, 1023 (2008).

[5] M. Kjaergaard, M. E. Schwartz, J. Braumüller, P. Krantz, J. I.-J. Wang, S. Gustavsson, and W. D. Oliver, Superconducting qubits: Current state of play, Annual Review of Condensed Matter Physics **11**, 369 (2020), https://doi.org/10.1146/annurev-conmatphys-031119-050605.

[6] F. Arute, K. Arya, R. Babbush, D. Bacon, C. Bardin, Joseph, R. Barends, R. Biswas, S. Boixo, F. Brandao, D. A. Buell, *et al.*, Quantum supremacy using a programmable superconducting processor, Nature **574**, 505 (2019).

[7] Y. Wu, W.-S. Bao, S. Cao, F. Chen, M.-C. Chen, X. Chen, T.-H. Chung, H. Deng, Y. Du, D. Fan, M. Gong, C. Guo, C. Guo, S. Guo, L. Han, L. Hong, H.-L. Huang, Y.-H. Huo, L. Li, N. Li, S. Li, Y. Li, F. Liang, C. Lin, J. Lin, H. Qian, D. Qiao, H. Rong, H. Su, L. Sun, L. Wang, S. Wang, D. Wu, Y. Xu, K. Yan, W. Yang, Y. Yang, Y. Ye, J. Yin, C. Ying, J. Yu, C. Zha, C. Zhang, H. Zhang, K. Zhang, Y. Zhang, H. Zhao, Y. Zhao, L. Zhou, Q. Zhu, C.-Y. Lu, C.-Z. Peng, X. Zhu, and J.-W. Pan, Strong quantum computational advantage using a superconducting quantum processor, Phys. Rev. Lett. **127**, 180501 (2021).

[8] C. Zhong, X. Han, H. X. Tang, and L. Jiang, Entanglement of microwave-optical modes in a strongly coupled electro-optomechanical system, Phys. Rev. A **101**, 032345 (2020).

[9] H. J. Briegel, W. Dür, J. I. Cirac, and P. Zoller, Quantum repeaters: The role of imperfect local operations in quantum communication, Phys. Rev. Lett. **81**, 5932 (1998).

[10] L.-M. Duan, M. D. Lukin, J. I. Cirac, and P. Zoller, Long-distance quantum communication with atomic ensembles and linear optics, Nature **414**, 413 EP (2001), article.

[11] J. I. Cirac, A. K. Ekert, S. F. Huelga, and C. Macchiavello, Distributed quantum computation over noisy channels, Phys. Rev. A **59**, 4249 (1999).

[12] C. Monroe, R. Raussendorf, A. Ruthven, K. R. Brown, P. Maunz, L. M. Duan, and J. Kim, Large-scale modular quantum-computer architecture with atomic memory and photonic interconnects, Phys. Rev. A **89**, 022317 (2014).

[13] P. Kómár, E. M. Kessler, M. Bishof, L. Jiang, A. S. Sørensen, J. Ye, and M. D. Lukin, A quantum network of clocks, Nat. Phys. **10**, 582 (2014).

[14] E. T. Khabiboulline, J. Borregaard, K. De Greve, and M. D. Lukin, Optical interferometry with quantum networks, Phys. Rev. Lett. **123**, 070504 (2019).

[15] J.-P. Chen, C. Zhang, Y. Liu, C. Jiang, W.-J. Zhang, Z.-Y. Han, S.-Z. Ma, X.-L. Hu, Y.-H. Li, H. Liu, F. Zhou, H.-F. Jiang, T.-Y. Chen, H. Li, L.-X. You, Z. Wang, X.-B. Wang, Q. Zhang, and J.-W. Pan, Twin-field quantum key distribution over a 511 km optical fibre linking two distant metropolitan areas, Nature Photonics **15**, 570 (2021).

[16] S. Welte, P. Thomas, L. Hartung, S. Daiss, S. Langenfeld, O. Morin, G. Rempe, and E. Distante, A nondestructive bell-state measurement on two distant atomic qubits, Nat. Photonics **15**, 504 (2021).

[17] Y. Yu, F. Ma, X.-Y. Luo, B. Jing, P.-F. Sun, R.-Z. Fang, C.-W. Yang, H. Liu, M.-Y. Zheng, X.-P. Xie, W.-J. Zhang, L.-X. You, Z. Wang, T.-Y. Chen, Q. Zhang, X.-H. Bao, and J.-W. Pan, Entanglement of two quantum memories via fibres over dozens of kilometres, Nature **578**, 240 (2020).

[18] R. Stockill, M. J. Stanley, L. Huthmacher, E. Clarke, M. Hugues, A. J. Miller, C. Matthiesen, C. Le Gall, and M. Atatüre, Phase-tuned entangled state generation between distant spin qubits, Phys. Rev. Lett. **119**, 010503 (2017).

[19] A. Delteil, Z. Sun, W.-B. Gao, E. Togan, S. Faelt, and A. Imamoğlu, Generation of heralded entanglement between distant hole spins, Nat. Phys. **12**, 218 (2016).

[20] D. Hucul, I. V. Inlek, G. Vittorini, C. Crocker, S. Debnath, S. M. Clark, and C. Monroe, Modular entanglement of atomic qubits using photons and phonons, Nat. Phys. **11**, 37 (2015).

[21] H. Bernien, B. Hensen, W. Pfaff, G. Koolstra, M. Blok, L. Robledo, T. H. Taminiau, M. Markham, D. J. Twitchen, L. Childress, and R. Hanson, Heralded entanglement between solid-state qubits separated by three metres, Nature **497**, 86 (2013).

[22] X. Han, W. Fu, C.-L. Zou, L. Jiang, and H. X. Tang, Microwave-optical quantum frequency conversion, Optica **8**, 1050 (2021).

[23] R. D. Delaney, M. D. Urmey, S. Mittal, B. M. Brubaker, J. M. Kindem, P. S. Burns, C. A. Regal, and K. W. Lehnert, Superconducting-qubit readout via low-backaction electro-optic transduction, Nature **606**, 489 (2022).

[24] W. Fu, M. Xu, X. Liu, C.-L. Zou, C. Zhong, X. Han, M. Shen, Y. Xu, R. Cheng, S. Wang, L. Jiang, and H. X. Tang, Cavity electro-optic circuit for microwave-to-optical conversion in the quantum ground state, Phys. Rev. A **103**, 053504 (2021).

[25] Y. Xu, A. A. Sayem, L. Fan, C.-L. Zou, S. Wang, R. Cheng, W. Fu, L. Yang, M. Xu, and H. X. Tang, Bidirectional interconversion of microwave and light with thin-film lithium niobate, Nat. Commun. **12** (2021).

[26] R. Hisatomi, A. Osada, Y. Tabuchi, T. Ishikawa, A. Noguchi, R. Yamazaki, K. Usami, and Y. Nakamura, Bidirectional conversion between microwave and light via ferromagnetic magnons, Phys. Rev. B **93**, 174427 (2016).

[27] J. G. Bartholomew, J. Rochman, T. Xie, J. M. Kindem, A. Ruskuc, I. Craiciu, M. Lei, and A. Faraon, On-chip coherent microwave-to-optical transduction mediated by ytterbium in yvo4, Nat. Commun. **11** (2020).

[28] R. Sahu, L. Qiu, W. Hease, G. Arnold, Y. Minoguchi, P. Rabl, and J. M. Fink, Entangling microwaves with

optical light (2023), arXiv:2301.03315.

[29] R. P. S. M. Lobo, J. D. LaVeigne, D. H. Reitze, D. B. Tanner, Z. H. Barber, E. Jacques, P. Bosland, M. J. Burns, and G. L. Carr, Photoinduced time-resolved electrodynamics of superconducting metals and alloys, Phys. Rev. B **72**, 024510 (2005).

[30] M. Mirhosseini, A. Sipahigil, M. Kalaee, and O. J. Painter, Superconducting qubit to optical photon transduction, Nature **588**, 599–603 (2020).

[31] W. Jiang, F. M. Mayor, S. Malik, R. Van Laer, T. P. McKenna, R. N. Patel, J. D. Witmer, and A. H. Safavi-Naeini, Optically heralded microwave photons (2022), arXiv:2210.10739.

[32] M. J. Weaver, P. Duivestein, A. C. Bernasconi, S. Scharmer, M. Lemang, T. C. van Thiel, F. Hijazi, B. Hensen, S. Gröblacher, and R. Stockill, An integrated microwave-to-optics interface for scalable quantum computing (2022), arXiv:2210.15702.

[33] M. Aspelmeyer, T. J. Kippenberg, and F. Marquardt, Cavity optomechanics, Rev. of Modern Phys. **86**, 1391 (2014).

[34] W. Jiang, C. J. Sarabalis, Y. D. Dahmani, R. N. Patel, F. M. Mayor, T. P. McKenna, R. Van Laer, and A. H. Safavi-Naeini, Efficient bidirectional piezo-optomechanical transduction between microwave and optical frequency, Nat. Commun. **11**, 10.1038/s41467-020-14863-3 (2020).

[35] C. Zhong, Z. Wang, C. Zou, M. Zhang, X. Han, W. Fu, M. Xu, S. Shankar, M. H. Devoret, H. X. Tang, and L. Jiang, Proposal for heralded generation and detection of entangled microwave–optical-photon pairs, Phys. Rev. Lett. **124**, 010511 (2020).

[36] C. Zhong, X. Han, and L. Jiang, Quantum transduction with microwave and optical entanglement (2022), arXiv:2202.04601.

[37] J. Zmuidzinas, Superconducting microresonators: Physics and applications, Annu. Rev. Condens. Matter Phys. **3**, 169 (2012).

[38] M. Xu, X. Han, W. Fu, C.-L. Zou, and H. X. Tang, Frequency-tunable high-q superconducting resonators via wireless control of nonlinear kinetic inductance, Appl. Phys. Lett. **114**, 192601 (2019).

[39] See supplementary information.

[40] A. A. Clerk, M. H. Devoret, S. M. Girvin, F. Marquardt, and R. J. Schoelkopf, Introduction to quantum noise, measurement, and amplification, Rev. Mod. Phys. **82**, 1155 (2010).

[41] C. Eichler, D. Bozyigit, C. Lang, L. Steffen, J. Fink, and A. Wallraff, Experimental state tomography of itinerant single microwave photons, Phys. Rev. Lett. **106**, 220503 (2011).

[42] J. F. Clauser, Experimental distinction between the quantum and classical field-theoretic predictions for the photoelectric effect, Phys. Rev. D **9**, 853 (1974).

[43] A. Kuzmich, W. P. Bowen, A. D. Boozer, A. Boca, C. W. Chou, L.-M. Duan, and H. J. Kimble, Generation of non-classical photon pairs for scalable quantum communication with atomic ensembles, Nature **423**, 731 (2003).

[44] R. Riedinger, S. Hong, R. A. Norte, J. A. Slater, J. Shang, A. G. Krause, V. Anant, M. Aspelmeyer, and S. Gröblacher, Non-classical correlations between single photons and phonons from a mechanical oscillator, Nature **530**, 313 (2016).

[45] S. Krastanov, H. Raniwala, J. Holzgrafe, K. Jacobs, M. Lončar, M. J. Reagor, and D. R. Englund, Optically heralded entanglement of superconducting systems in quantum networks, Phys. Rev. Lett. **127**, 040503 (2021).

[46] P. Kurpiers, P. Magnard, T. Walter, B. Royer, M. Pechal, J. Heinsoo, Y. Salathé, A. Akin, S. Storz, J.-C. Besse, S. Gasparinetti, A. Blais, and A. Wallraff, Deterministic quantum state transfer and remote entanglement using microwave photons, Nature **558**, 264 (2018).

[47] H. Ren, M. H. Matheny, G. S. MacCabe, J. Luo, H. Pfeifer, M. Mirhosseini, and O. J. Painter, Two-dimensional optomechanical crystal cavity with high quantum cooperativity, Nat. Comm. **11** (2020).

ACKNOWLEDGMENTS

The authors thank M. Mirhosseini, M. Kalaee, A. Sipahigil and J. Banker for contributions in the early stages of this work, E. Kim, A. Butler, G. Kim, S. Sonar, U. Hatipoglu and J. Rochman for helpful discussions, and B. Baker and M. McCoy for experimental support. We appreciate MIT Lincoln Laboratories for providing the traveling-wave parametric amplifier used in the microwave readout chain in our experimental setup. NbN deposition during the fabrication process was performed at the Jet Propulsion Laboratory. This work was supported by the ARO/LPS Cross Quantum Technology Systems program (grant W911NF-18-1-0103), the U.S. Department of Energy Office of Science National Quantum Information Science Research Centers (Q-NEXT, award DE-AC02-06CH11357), the Institute for Quantum Information and Matter (IQIM), an NSF Physics Frontiers Center (grant PHY-1125565) with support from the Gordon and Betty Moore Foundation, the Kavli Nanoscience Institute at Caltech, and the AWS Center for Quantum Computing. L.J. acknowledges support from the AFRL (FA8649-21-P-0781), NSF (ERC-1941583, OMA-2137642), and the Packard Foundation (2020-71479). S.M. acknowledges support from the IQIM Postdoctoral Fellowship.

Supplementary Information for "Non-classical microwave-optical photon pair generation with a chip-scale transducer"

Srujan Meesala,[1,2,*] Steven Wood,[1,2,*] David Lake,[1,2,*] Piero Chiappina,[1,2] Changchun Zhong,[3] Andrew D. Beyer,[4] Matthew D. Shaw,[4] Liang Jiang,[3] and Oskar Painter[1,2,5,†]

[1] *Kavli Nanoscience Institute and Thomas J. Watson, Sr., Laboratory of Applied Physics, California Institute of Technology, Pasadena, California 91125, USA*

[2] *Institute for Quantum Information and Matter, California Institute of Technology, Pasadena, California 91125, USA*

[3] *Pritzker School of Molecular Engineering, The University of Chicago, Chicago, IL 60637, USA*

[4] *Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Dr, Pasadena, California 91109, USA*

[5] *Center for Quantum Computing, Amazon Web Services, Pasadena, California 91125, USA*

(Dated: April 3, 2023)

1. NOTATION

Symbol	Meaning
$\hat{a}, \hat{b}, \hat{c}$	operators for optical, acoustic, and microwave modes of the transducer
$\hat{c}_+, \hat{c}_-$	operators for hybridized electromechanical modes
$\hat{a}_{\mathrm{in}}, \hat{c}_{\mathrm{in}}, (\hat{a}_{\mathrm{out}}, \hat{c}_{\mathrm{out}})$	operators for optical and microwave input (output) modes in coupling waveguides
$\hat{A}, \hat{C}$	operators for optical and microwave temporal modes in coupling waveguides
$\bar{C}_{mn}$	moments of the temporal microwave mode
$\bar{C}_{mn}\vert_{\mathrm{click}}$	moments of the temporal microwave mode conditioned on an optical click

TABLE S1. Summary of the various modes and related quantities defined in the main text.

2. DEVICE FABRICATION PROCESS

The fabrication process for the transducer chip is illustrated in Fig. S1 and described in the caption. Masks for all steps are patterned in ZEP-520A resist via electron-beam lithography on a Raith EBPG 5200 tool. All dry etching is performed in Oxford Plasmalab 100 inductive coupled plasma reactive ion etching (ICP RIE) tools. The process flow can be sub-divided into sections used to define various portions of the transducer device. Steps (i)-(vi) complete the definition of the AlN box essential for the piezoacoustic cavity. The combination of dry and wet etch steps ensures that the dimensions of the AlN box are precisely defined while the silicon device layer is undamaged on most of the chip. This is important to achieve optical, mechanical, and microwave modes with high quality factors. Steps (vii)-(ix) define the NbN resonator and step (x) defines the OMC. Steps (xi)-(xiii) are use to define aluminum electrodes on the piezo-resonator and galvanically connect them to the NbN resonator using bandage steps. The bandaid steps involve in situ Ar milling for two minute and six minute durations respectively to clear the surface of NbN and Al prior to Al bandaid evaporation. To provide optical fiber access to coupler sections at the end of the silicon photonic waveguides, we clear a portion of the SOI substrate up to a depth of 150 μm using a deep reactive ion etch at the edge of the chip. Finally, the buried oxide (BOX) layer is etched in anhydrous vapor HF to release the device membrane.

3. MICROWAVE CIRCUIT DESIGN

The kinetic inductance resonator used in our transducer is fabricated from an NbN film of 10 nm thickness and 50 pH/sq sheet inductance. The meandering ladder geometry described in the main text is comprised of 2 μm $\times$ 1 μm rectangular loops formed from traces of width 130 nm. We use extended electrical terminals of length 200 μm and width 1 μm to spatially separate the high kinetic inductance section from the OMC. The resonator is designed to achieve a target frequency of 5.0 GHz for the fundamental mode with a capacitance, $C_{\mathrm{res}} = 7.1$ fF, which includes a small contribution of 0.27 fF from the electrodes on the piezo-acoustic resonator. Nearly the entire inductance of the resonator mode is due to the kinetic inductance of the superconducting film. The use of closed superconducting loops in the resonator allows for tuning of the kinetic inductance, L_k, via a DC supercurrent, I, induced by an external magnetic field according to the relation

* These authors contributed equally

† opainter@caltech.edu; http://copilot.caltech.edu

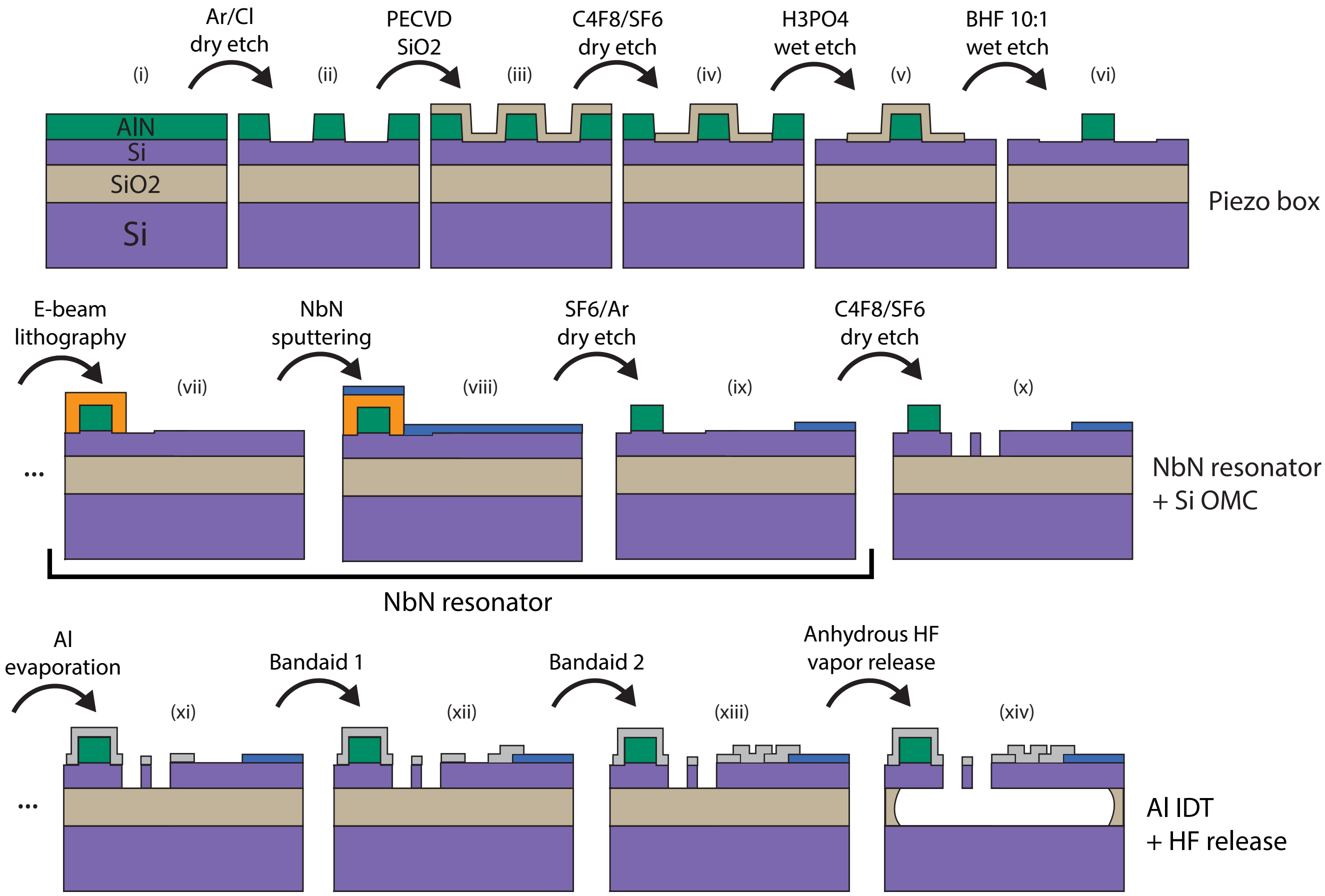

FIG. S1. Device fabrication process. Images are not to scale. **(i)** Sputter deposition of 300 nm thick c-axis AlN piezoelectric film (grown by OEM group; stress $T = +55$ MPa; (002) XRD peak of full-width at half-maximum = 1.79°) on a silicon on insulator substrate (Si device layer [float zone grown, 220 nm thick, $\rho \geq 5\text{k}\Omega$-cm]; buried oxide layer [3 µm thick, silicon dioxide]; Si handle [Czochralski grown, 750 µm thick, $\rho \geq 5\text{k}\Omega$-cm]). **(ii)** AlN trench etch with Ar/Cl_2 chemistry to define the perimeter of the piezo resonator via a small trench of width ~100 nm. **(iii)** Conformal deposition of a SiO_x hard mask via plasma enhanced chemical vapor deposition. **(iv)** Patterning of the SiO_x mask via dry etching with C_4F_8/SF_6 chemistry. **(v)** Removal of the remaining AlN on the chip with a H_3PO_4 wet etch. **(vi)** SiO_x mask removal with 10:1 buffered oxide etchant. **(vii)** Protection of the piezo-acoustic and OMC regions with a resist mask prior to deposition of NbN. **(viii)** Deposition of a 10 nm thick film of NbN via an RF sputtetring process. **(ix)** Dry etch of NbN with SF_6/Ar chemistry to define the high-impedance microwave resonator. **(x)** Dry etch of silicon with C_4F_8/SF_6 chemistry to define OMC, acoustic shield and optical waveguide. **(xi)** Deposition of Al electrodes for the piezo-acoustic resonator using angled electron-beam evaporation. **(xii, xiii)** Bandage steps with Ar milling followed by Al evaporation. **(xiv)** Etching of the buried oxide (BOX) layer using anhydrous vapor HF to release the device membrane.

$L_k \approx L_{k,0}[1 + (I/I_*)^2]$ as shown in [1]. Here $L_{k,0}$ is the kinetic inductance at zero magnetic field. $I_* \gg I$ is a characteristic current on the order of the critical current of the nanowire [2]. This relation leads to quadratic tuning of the resonator frequency in response to an external magnetic field as observed in Fig. 2b of the main text. The resonator is capacitively coupled to an on-chip 50 Ω coplanar waveguide (CPW) patterned in NbN to achieve an external coupling rate, $\kappa_{e,c}/2\pi = 1.3$ MHz. On the chip used for the experiments in this work, sixteen transducer devices were laid out in groups of four with each group addressed by a separate CPW. Adjacent resonators on the same CPW were designed to be detuned from each other by a frequency spacing of 100 MHz, much larger than the cross-coupling, which was estimated to be less than 1 MHz from simulation. This frequency multiplexing approach allows us to increase the number of transducers available for testing on the chip while ensuring minimal microwave crosstalk.

4. PIEZO-OPTOMECHANICS DESIGN

The piezo-optomechanical transducer in our work is realized by attaching a half-wavelength AlN piezo-acoustic cavity to a silicon OMC as shown in Fig. S2a. Figure S2b,c shows the optical and acoustic modes used for transduction. The energy of the acoustic mode is split between the piezo-acoustic and OMC regions, which are designed independently to support acoustic resonances at 5 GHz with large piezo-electric and optomechanical coupling respectively. Similar to the device demonstrated in [3], hybridization of acoustic modes in these two portions is achieved through the connecting OMC section, whose band structure is designed to provide a bandgap for optical photons and waveguiding for acoustic phonons at the frequencies of interest. The primary difference compared to the previous design is the change to an ultra-low mode volume piezo-acoustic cavity to reduce energy participation in the piezo region, which is lossier for acoustics when compared to silicon. This design change is possible without sacrificing piezoelectric coupling since the impedance of the kinetic inductance resonator is roughly an order of magnitude larger than that of the transmon qubit used in [3].

Figures S2d,e show the hybridization of the piezo-acoustic and OMC modes via a sweep of the length of the piezo box. We observe strong hybridization of two distinct branches that are piezo-like and OMC-like as highlighted by the dashed lines with maximal hybridization occurring at a piezo length of 880 nm. On the transducer chip used in our experiments, we fabricate devices with piezo length swept over a range of 20 nm about this nominal value to compensate for fabrication disorder. Clamping loss of the piezo-acoustic resonator is minimized by using tethers with modulated width as shown in Figs. S2a,b. The tethers are designed to support a 1.5 GHz wide acoustic bandgap centered around the transducer acoustic modes.

The OMC is evanescently coupled to a suspended silicon waveguide, which terminates in a coupler section at the edge of the chip. We use a millimeter length scale waveguide to distance the NbN circuit from scattering of pump light generated near the optical coupler section. The waveguide is designed with alternating curved sections in order to increase robustness against buckling from intrinsic stress in the silicon device layer. Adiabatic tethers are used to anchor the waveguide in order to reduce scattering loss and maintain high optical collection efficiency.

5. MEASUREMENT SETUP

The measurement setup used in this work is detailed in Fig. S3. For SPDC experiments, trigger signals from a master digital delay generator are used to synchronize optical pump pulses with the timing window used for optical and microwave readout. Optical pump pulses with > 120 dB extinction are generated using two analog acousto-optic modulators (G&H Photonics) and are routed via a circulator into the 'Optics in/out' path towards the device in the dilution fridge as shown in Fig.S3a. The optical emission from the device passes through the same circulator and is directed to a pump filtering setup prior to single photon detection (SPD) along the 'SPD in' path in Fig. S3a. The filtering setup comprises two tunable Fabry–Perot filter cavities (Stable Laser Systems) in series and provides 104 dB extinction for a pump detuning of 5 GHz along with a transmission bandwidth of 2.7 MHz. This transmission bandwidth naturally excludes emission due to optomechanical scattering by other transducer modes besides the two hybridized electromechanical modes of interest, $\hat{c}_{\pm}$. During the experiment, transmission through the filters is checked every four minutes and a lock sequence is initiated if the transmission drops below a set threshold. At the beginning of the transmission check, the optical path is set to bypass the fridge via a MEMS switch. An electro-optic phase modulator (ϕ-m in Fig. S3a) is used to generate a sideband on the pump tone at the target frequency to which the filters are locked. In our SPDC experiments, this target frequency corresponds to a pump detuning, $\Delta_{\mathrm{a}} = (\omega_+ + \omega_-)/2$. If required, the locking algorithm adjusts the filter cavities to maximize transmission at the target frequency by monitoring the output of each cavity on a separate photodetector. Additionally, during long measurements, we periodically monitor the polarization of the pump light sent to the device and compensate for long term polarization drifts along the excitation path. Active polarization control is preformed by using an electronic polarization controller (Phoenix Photonics) and maximizing the optical pump power reflected by the device.

Optical photon counting is achieved using a tungsten silicide (WSi) superconducting nanowire single photon detector (SNSPD) mounted on the still plate of the dilution refrigerator (BlueFors LD-250) maintained at a temperature of 770 mK. As shown in Fig. S3b, electrical pulses from the SNSPD generated by single optical photon detection events are amplified and split into two separate paths to record the detection time on a time correlated single photon counter (TCSPC) and to trigger microwave readout conditioned on an optical click. For conditional microwave readout, we first generate a logical bit for every optical click by performing a logic level translation of the SNSPD electrical pulse to a TTL signal. We then perform a logical AND operation between this 'optical click bit' and a trigger bit provided by the master delay generator defining the duration of the optical gating window. On the other hand, unconditional microwave readout is triggered directly by the master delay generator. An RF switch (MiniCircuits ZASWA-2-50DRA+) is used to switch between unconditional and conditional readout of the microwave output signal.

The microwave output chain shown along the 'MW out' path in Fig. S3c begins with a Josephson traveling-

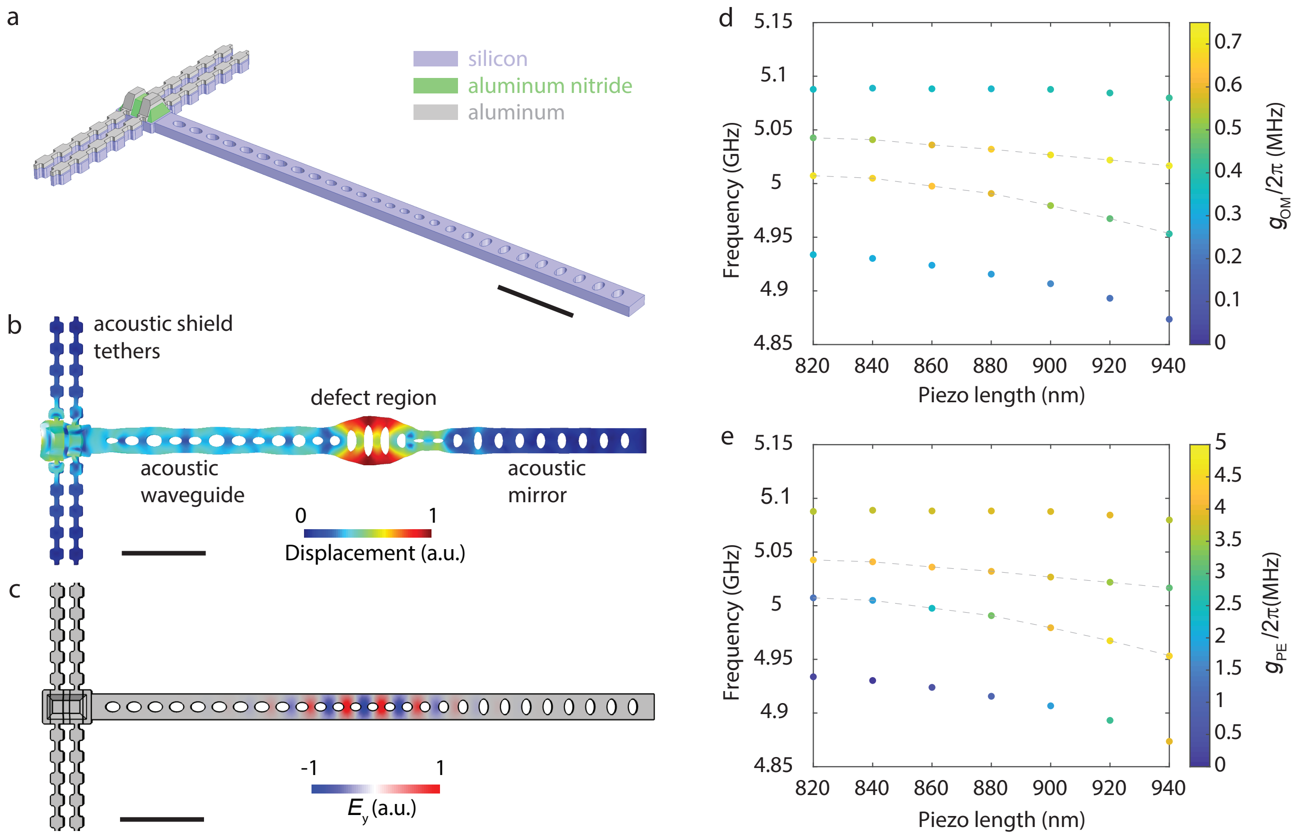

FIG. S2. Piezo-optomechanics design. **a.** Geometry and material composition of the piezo-optomechanical transducer. The black scale bar corresponds to 2 µm **b,c.** Top-down views detailing the displacement profile of a hybridized acoustic mode and the transverse electric field of the fundamental optical mode of the transducer respectively. **d,e.** Simulated optomechanical coupling, g_{om}, and piezoelectric coupling, g_{pe}, of the acoustic modes of the transducer as a function of length of the piezo-acoustic cavity. Dashed lines in both panels are shown as a guide to follow the modes with maximum g_{om}.

wave parametric amplifier (TWPA, MIT Lincoln Labs) [4] mounted on the mixing plate as the first amplification stage. For TWPA operation, we use a CW pump tone at a frequency of ∼6.07 GHz added to the amplifier input using a 20 dB directional coupler (not shown in the figure). A dual junction circulator with 40 dB isolation is placed between the directional coupler and the sample to shield the transducer from back-reflected pump. The TWPA is followed by a high mobility electron transistor (HEMT, Low Noise Factory LNF-LNC4_8C) amplifier mounted on the 4K plate. In the setup outside the fridge as shown in Fig. S3b, we perform additional amplification of the signal and filtering of the TWPA pump tone using a tunable notch filter (Micro Lambda Wireless MLBFR-0212). The microwave signal is then downconverted to an intermediate frequency (IF) of ∼100 MHz after mixing with a local oscillator on an IQ mixer (Marki IQ-4509). The I and Q outputs of the mixer are subsequently bandpass-filtered, amplified and recorded independently on two channels of an analog to digital converter (ADC, Alazartech ATS 9360). These two digitized voltage signals correspond to measurement of the real and imaginary quadratures of the output mode of the amplification chain, $\hat{s}_{\mathrm{out}}$. Based on the calibrated gain of our amplification chain, we measured that our heterodyne detection setup has an added noise of ∼2.5 quanta referred to the output of the transducer at a signal frequency of ∼5.0 GHz. This near-quantum-limited heterodyne readout enabled by the TWPA is a key enabler for measurements of microwave-optical correlations from our transducer on a reasonable experimental timescale. The setup shown in Fig. S3b additionally allows for pulsed microwave excitation of the transducer as well as spectroscopy with a vector network analyzer (VNA).

The transducer sample is wire-bonded to a printed circuit board (PCB) with coaxial connectors and is housed in an oxygen-free high thermal conductivity (OFHC) copper package. The microwave resonance frequency of the NbN resonators can be tuned by an external magnetic field generated from an Nb-Ti coil mounted over the sample package. Individual devices on the chip are optically addressed using a lensed fiber (OZ Optics) affixed to the top of a three-axis piezo stepper stack (Attocube Systems) placed in line with the on-chip tapered optical

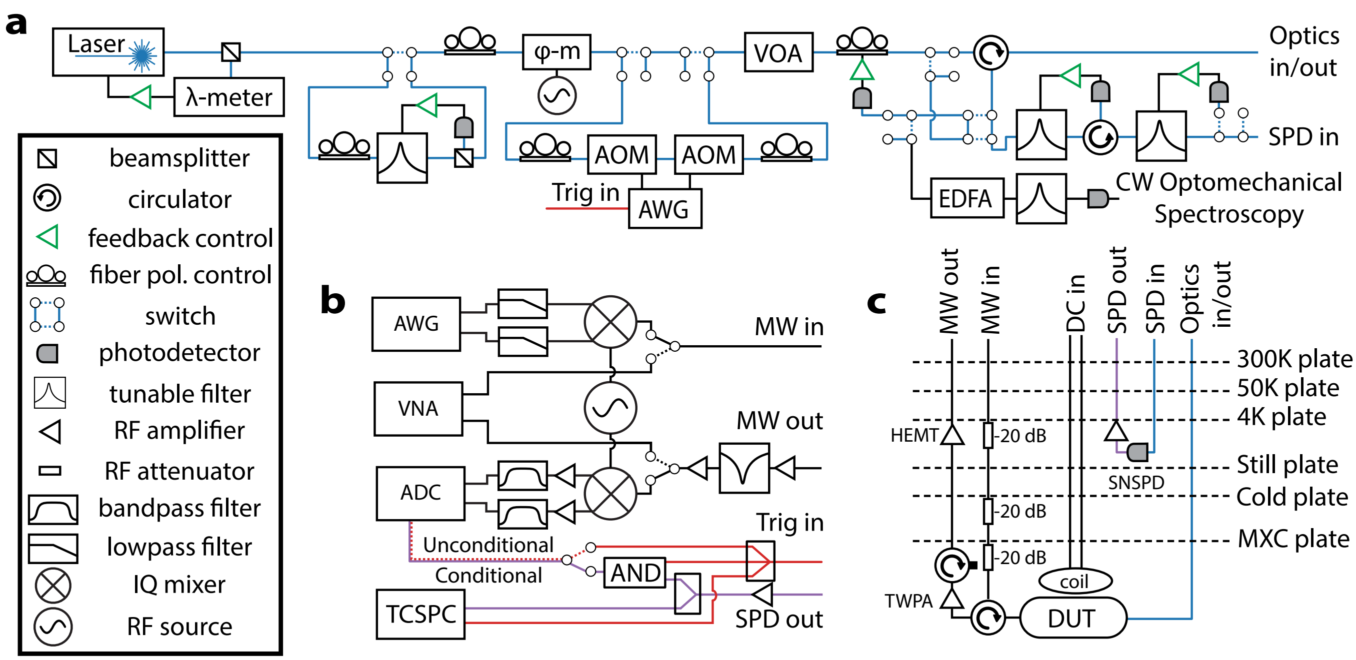

FIG. S3. Experimental setup. **a.** Optical pulse synthesis and detection. An external cavity diode laser is wavelength locked to a wavemeter, and a 50 MHz tunable narrowband filter is used to remove spontaneous emission noise. Optical pulse shaping is achieved by modulating the laser with two acousto-optic modulators (AOMs) connected in series. The envelope for the optical pulses is generated by an arbitrary waveform generator (AWG). A variable optical attenuator (VOA) is used to adjust the optical power level sent to the device. Excitation of the device is performed via an optical circulator through which the device reflection and emission can be routed to two separate paths to perform either continuous wave (CW) spectroscopy or pulsed single photon counting. For CW spectroscopy, the reflected optical pump along with the optomechanically transduced signal is amplified with an erbium doped fiber amplifier (EDFA) and detected on a high-speed photodetector. On the single photon counting path, pump light is filtered using two tunable Fabry–Perot filter cavities (Stable Laser Systems) before directing the optical signal to the single photon detector (SPD) in the dilution fridge. **b.** Microwave pulse synthesis and detection. For measurements that require electrical excitation of the transducer, pulsed input signals to the fridge are generated at the intermediate frequency (IF) by an arbitrary waveform generator (AWG) and subsequently upconverted to ∼5GHz on an IQ mixer. Likewise, for CW microwave spectroscopy, a vector network analyzer (VNA) may be switched into the setup. The microwave output from the fridge passes though a tunable bandstop filter to remove the TWPA pump tone, and is down-converted, amplified and filtered before acquisition on an analog to digital converter (ADC). For unconditional microwave readout, the master delay generator for the experiment generates the trigger signal for the ADC. For microwave readout conditioned on an optical click generated from SPDC, the trigger signal for the ADC is generated by performing a logical AND operation between optical click events from the SPD and a trigger bit from the master delay generator spanning the gating window for collection of the optical signal. **c.** Dilution refrigerator configuration. The microwave (MW) input line contains attenuators on the 4K, cold, and mixing chamber plates for thermalization of the coaxial cables. The MW output line passes through an amplification chain consisting of a Traveling-Wave Parametric Amplifier (TWPA) mounted on the mixing plate and a High Electron Mobility Transistor (HEMT) amplifier mounted on the 4K plate. The TWPA is driven by a strong pump tone sent along a separate line and inserted via a directional coupler placed before the amplifier input (not shown). The superconducting nanowire single photon detector (SNSPD) is mounted to the still plate and the electrical output signal is amplified on the 4K plate before being sent to room temperature electronics.

couplers. The entire assembly is enclosed in a cylindrical magnetic shield and is mounted to the mixing plate of the dilution refrigerator cooled to a base temperature $T_\mathrm{f} \approx 20$ mK.

6. OPTICAL HERALDING RATE CONTRIBUTIONS

The optical heralding events in our SPDC experiments have a finite noise contribution from pump-induced heating of the hybridized electromechanical modes as well as technical sources in our experimental setup, namely detector dark counts and pump laser leakage. In Fig. S4a, we plot the optical photon count rate due to these noise sources as determined from independent measurements.

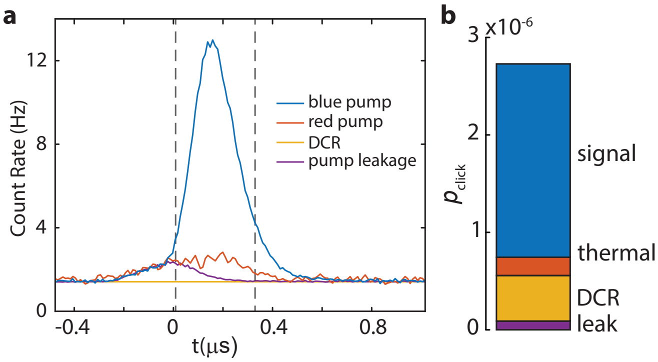

FIG. S4. **a.** Optical photon count rate vs time as measured with blue detuned pump pulses, red detuned pump pulses and in calibration measurements for the dark count rate (DCR) and pump leakage through the filter setup. **b.** Contributions to the optical click probability during the gating window extracted from the data in **a**.

Source of a photon click	Fraction of clicks
SPDC signal	0.727
Thermal	0.069
DCR	0.171
Pump leakage	0.033

TABLE S2. Contributions to the optical photon count rate.

The blue trace corresponds to the total optical signal as shown in Fig. 3a of the main text. Dashed lines bound the gating duration used to select photon clicks that trigger conditional microwave readout. The dark count rate (DCR) of the SNSPD in our experiment is 1.5 Hz and is limited by black-body radiation at wavelengths outside the telecom band guided through the optical fiber. Stray photon flux from laser leakage through the pump filtering setup as indicated by the purple trace in Fig. S4 is determined by detuning the filter cavity by 100 MHz from the acoustic mode used in the transduction experiment. To determine the noise contribution from the thermal occupation of the acoustic mode, we perform a measurement of optomechanical sideband asymmetry [5]. We excite the transducer with the same pump pulse sequence used in the SPDC experiment but with the pump laser tuned to the red mechanical sideband of the optical cavity ($\Delta_\mathrm{a} = -(\omega_+ + \omega_-)/2$). The resultant photon flux from this measurement is shown by the red trace in Fig. S4a. Integrating the photon counts from these independent measurements over the gating window, we obtain the contributions of each noise source to the total probability of a heralding event, $p_\mathrm{click} = 2.7 \times 10^{-6}$ as shown in Fig. S4b. The corresponding fractional contributions are enumerated in Tab. S2. Further, the difference between the count rate under red and blue detuned laser excitation allows us to infer an average thermal occupation of 0.097 ± 0.019 for the acoustic mode over the duration of the optical gating window.

7. MICROWAVE MOMENT INVERSION

To measure the statistical moments of the microwave emission from the transducer, we adopt the method originally demonstrated in Ref. [6], which has been widely applied to perform tomography of non-classical microwave radiation emitted by superconducting qubits [7–9]. The microwave signal emitted from the device undergoes a series of amplification steps before a record is captured as a digitized, complex-valued voltage signal on the heterodyne setup. This phase-insensitive amplification chain inevitably adds noise and is modelled as [10, 11],

$$\hat{s}_\mathrm{out} = \sqrt{G}\hat{c}_\mathrm{out} + \sqrt{G-1}\hat{h}^\dagger, \tag{1}$$

where $\hat{c}_\mathrm{out}$, $\hat{s}_\mathrm{out}$ and $\hat{h}^\dagger$ are the bosonic mode operators corresponding to the transducer output, the heterodyne setup output and added noise respectively, with dimensions of $T^{-1/2}$. In the limit of large amplifier gain ($G \gg 1$), this relation may be approximated as

$$\hat{s}_\mathrm{out} \approx \sqrt{G}\left(\hat{c}_\mathrm{out} + \hat{h}^\dagger\right), \tag{2}$$

Once recorded, we take the inner product of the heterodyne traces with an emission envelope function, f in software. A particular choice of the envelope function corresponds to a measurement of the quadratures of the temporal mode $\hat{S}(t) := \int \hat{s}_\mathrm{out}(t+t')f^*(t')dt'$. With this choice Eq. (2) becomes,

$$\hat{S} \approx \sqrt{G}\left(\hat{C} + \hat{H}^\dagger\right), \tag{3}$$

where $\hat{C}(t) := \int \hat{c}_\mathrm{out}(t+t')f^*(t')dt'$ and $\hat{H}(t) := \int \hat{h}(t+t')f^*(t')dt'$. Note that $\hat{C}$ is a temporal mode in the output waveguide, not the internal microwave mode, $\hat{c}$ of the transducer. Thus from a sequence of N heterodyne measurements, a dataset of complex voltage samples $\{S_1, S_2, ...S_N\}$ is generated and the statistical moments are numerically calculated as [6, 11],

$$\bar{S}_{mn} := \langle \hat{S}^{\dagger m}\hat{S}^n\rangle = \frac{1}{N}\sum_{i=1}^{N}(S_i^*)^m(S_i)^n. \tag{4}$$

The results in the main text require us to determine the statistical moments of $\hat{C}$, which we organize into a moments matrix with elements $\bar{C}_{mn} = \langle \hat{C}^{\dagger m}\hat{C}^n\rangle$. This is accomplished using additional reference measurements in which the transducer is not optically pumped so that the input to the amplifier is a vacuum state. This reference measurement directly gives the anti-normally ordered moments of the added noise as $\bar{H}_{mn} := \langle \hat{H}^m\hat{H}^{\dagger n}\rangle = \bar{S}_{mn}/G^{\frac{n+m}{2}}$.

We experimentally verified that $\bar{H}_{mn}$ is consistent with a thermal state for $m, n \leq 2$. Knowledge of the mean occupancy ($n_{\mathrm{th},H}$) of the state constitutes full knowledge of the statistics for a thermal state. Accordingly, we extract $n_{\mathrm{th},H}$ from the reference measurements and calculate $\bar{H}_{mn}$ as [12],

$$\bar{H}_{mn} = \begin{cases} m!(n_{\mathrm{th},H}+1)^m & m = n, \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

Finally, under the assumption that there are no correlations between $\hat{C}$ and $\hat{H}$, Eq. (3) can be used to derive the following relation between the moment matrix elements [6],

$$\bar{S}_{ij} = \bar{T}_{ijmn}\bar{C}_{mn}, \tag{6}$$

where

$$\bar{T}_{ijmn} = \begin{cases} G^{\frac{i+j}{2}}\binom{i}{m}\binom{j}{n}\bar{H}_{i-m,j-n} & m \le i, n \le j, \\ 0 & \text{otherwise.} \end{cases} \tag{7}$$

By inverting Eq. (6), we recover $\bar{C}_{mn}$.

8. MICROWAVE EMISSION ENVELOPE FUNCTION

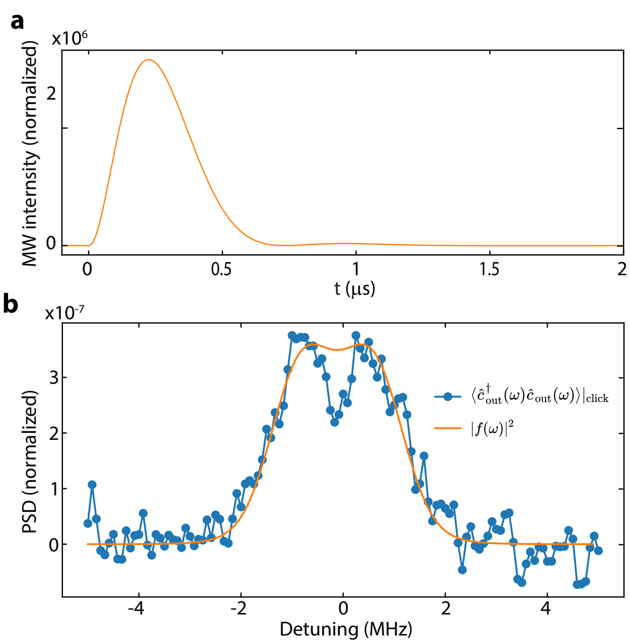

FIG. S5. **a.** Simulated temporal profile of microwave emission intensity from the transducer heralded by an optical detection event. **b.** Comparison of normalized power spectral density (PSD) of the emission envelope function, f (orange trace) and of the experimentally measured transducer microwave output conditioned on an optical click (blue data points). Detuning is defined with respect to the mean frequency of the hybridized electromechanical modes.

We define the two hybridized electro-mechanical modes in our experiment as

$$\hat{c}_\pm = \frac{1}{\sqrt{2}}\left(\hat{b} \pm \hat{c}\right), \tag{8}$$

with center frequencies $\omega_\pm = \omega_b \pm g_{\mathrm{pe}}$.

The ideal two-mode squeezed state generated under constant optical pumping on the blue optomechanical sideband ($\Delta = \omega_a - \omega_p = \omega_c$) can be written in the form [13],

$$|\psi\rangle = \left[1 + \sqrt{p}\left(\hat{a}^\dagger_+\hat{c}^\dagger_+ + \hat{a}^\dagger_-\hat{c}^\dagger_-\right) + \mathcal{O}(p)\right]|\mathrm{vac}\rangle\,, \tag{9}$$

up to a normalization factor. Here p is the probability of an optomechanical scattering event, $\hat{a}_\pm$ are the optical modes which match the frequencies of the hybridized modes $\hat{c}_\pm$ as required by energy conservation. Performing single photon detection on the optical state coupled out of the cavity naturally allows us to discard the vacuum component of $|\psi\rangle$. Further, we can neglect higher order terms, $\mathcal{O}(p)$ in the expansion since $p \ll 1$ for the optical pump power level used in our experiments. In our experiment, the SNSPD has nanosecond timing jitter, which is much smaller than the period of the beat note between the modes $\hat{a}_\pm$ given by $2\pi/2g_{\mathrm{pe}} = 625$ ns. As a result, detection of an SNSPD click in an experimental trial erases frequency information of the optical state. This measurement can be described as a projective measurement onto the state, $\left(e^{i\phi(t)/2}\hat{a}_+ + e^{-i\phi(t)/2}\hat{a}_-\right)|\mathrm{vac}\rangle$, where the relative phase $\phi(t) = \phi_o + 2g_{\mathrm{pe}}t$ for an optical click received at time t. Here ϕ_o is the constant relative phase acquired between the two frequency bins along the optical path and t is the emission time of the photon pair defined relative to the beginning of the pump pulse. An optical click at time t can therefore be used to herald the microwave state $|\psi\rangle_{\mathrm{click}} = \left(e^{-i\phi(t)/2}\hat{c}^\dagger_+ + e^{i\phi(t)/2}\hat{c}^\dagger_-\right)|\mathrm{vac}\rangle$.

This picture becomes slightly more complicated when considering the pulsed optical drive used in this work as well as the parameters of the device. The large bandwidth of the pump pulse of 2.75 MHz in comparison to the mode splitting of 1.6 MHz means that the frequency bins cannot be individually well resolved in the optical emission. Further, given the splitting between the frequency bins versus their individual linewidths of 1.0 MHz, our device is not deep in the strong piezoelectric coupling regime. This is evinced by the two peaks with finite overlap in the experimentally measured power spectrum of the conditional microwave emission, $\langle \hat{c}^\dagger_{\mathrm{out}}(\omega)\hat{c}_{\mathrm{out}}(\omega)\rangle|_{\mathrm{click}}$ shown with the blue trace in Fig. S5b. However, the theoretical picture discussed above can guide the parametrization of the envelope function used for microwave readout. Accordingly, we define the microwave emission envelope function as a coherent superposition of two frequency bins centered at $\omega_\pm$,

$$f(t) = \frac{g(t)}{\sqrt{2}}\left(e^{-i(\omega_+ t + \phi_o/2)} + e^{-i(\omega_- t - \phi_o/2)}\right). \tag{10}$$

The relative phase, ϕ_o is assumed to be fixed since the optical pulse duration used in our experiment is short compared to the beat period between the frequency bins. The function, $g(t)$ accounts for the finite linewidth of the frequency bins and is obtained by numerically solving the master equation for our system using the QuTiP software package [14]. For this calculation, we use the Hamiltonian in Eq. (1) of the main text along with experimentally determined coupling and decay rates for the transducer modes. The resulting MW emission intensity is shown in Fig. S5a. $g(t)$ is parameterized as a skewed Gaussian to capture the faster timescale of the rising edge relative to the decay. The corresponding voltage envelope is given by the square root of a skewed Gaussian and is written as

$$g(t) = \frac{1}{\sqrt[4]{2\pi T_\mathrm{g}^2}} \exp\left(-\frac{t^2}{4T_\mathrm{g}^2}\right)\left[1 + \mathrm{erf}\left(\frac{\alpha t}{\sqrt{2}T_\mathrm{g}}\right)\right]^{1/2}, \tag{11}$$

up to a normalization factor involving the gain of the microwave amplification chain. Here the Gaussian standard deviation, $T_\mathrm{g} = 230$ ns and skew factor, $\alpha = 2.0$ are obtained from fitting to the simulated result. We note that for a microwave photon instantaneously loaded into the resonator, $g(t)$ is expected to follow an exponential decay. However, in our transducer, due to the finite conversion rate of the phonon from SPDC into a microwave photon, $g(t)$ has a finite rise time. As a result, we find that choosing the skewed Gaussian parametrization instead of an exponential decay provides $\sim 4\%$ higher overlap between the emission envelope function and the simulated wavepacket as quantified by their normalized inner product. The power spectrum of the resultant envelope function, $|f(\omega)|^2$ is shown as the orange trace in Fig. S5b. This is observed to be well-matched to the power spectrum of the conditional microwave emission from the transducer obtained from experimentally recorded heterodyne voltage traces.

9. MICROWAVE GAIN CALIBRATION

We calibrate the gain of our microwave heterodyne detection setup by operating the transducer as a frequency converter with the optical pump laser on the red mechanical sideband. The calibration procedure is performed at zero magnetic field, when the acoustic and microwave modes, $\hat{b}, \hat{c}$ are detuned by a frequency separation, $\omega_b - \omega_c = 2\pi \times 12$ MHz $\gg 2g_\mathrm{pe}$ and are weakly hybridized. An independent measurement of optomechanical sideband asymmetry provides a meter for phonon occupation in the acoustic mode, $\hat{b}$ via the single phonon count rate, R_o [5]. We then resonantly excite the acoustic mode via the microwave port and measure the transduced optical photon count rate, R. Simultaneously, we record the reflected microwave signal from the transducer, which produces an output power, P_det at the output of the heterodyne setup. The gain, G of the amplification chain can be determined using the series of equations below.

$$R = n_{b,\mathrm{sig}} R_o, \tag{12}$$

$$n_{b,\mathrm{sig}} \hbar\omega_b = P_\mathrm{in} \frac{4\kappa_{e,b}}{\kappa_{i,b}^2}, \tag{13}$$

$$P_\mathrm{det} = G\left(\frac{\kappa_{e,b} - \kappa_{i,b}}{\kappa_{e,b} + \kappa_{i,b}}\right)^2 P_\mathrm{in}. \tag{14}$$

Here, $n_{b,\mathrm{sig}}$ is the occupation of the acoustic mode due to the microwave drive obtained after subtracting the thermal component of the transduced signal. P_in is the input microwave power to the transducer. $\kappa_{b,\mathrm{e}}$ is the coupling rate of the acoustic mode to the microwave waveguide due to weak hybridization with the microwave resonator and $\kappa_{b,\mathrm{i}}$ is the intrinsic linewidth of the acoustic mode. Both rates are determined via a VNA measurement of the acoustic mode in the presence of the optical pump pulses. Following this procedure, we estimate gain, $G = 103$ dB between the microwave output port of the transducer and the output of the heterodyne setup. While knowledge of the gain allows us to estimate the microwave intensity at the transducer output, we note that the values of the normalized correlation functions measured in our experiment are independent of the absolute accuracy of this calibration.

10. TWPA OPTIMIZATION

The operating point of the TWPA was optimized prior to data acquisition. In order to ensure fast acquisition and unbiased statistics, our operational criterion were (1) maximize the gain and (2) ensure linearity in the response. In our optimization procedure, the frequency and power of the pump tone driving the TWPA were swept, and coherent pulses with amplitude matched to the conditional microwave signal in our experiment were sent to the transducer input port. The result of the heterodyne measurement of the reflected signal from the transducer was used to compute $g^{(2)}_{CC}$. To ensure linearity, we chose a region in the TWPA pump frequency and power space where $g^{(2)}_{CC} = 1$ to within one standard deviation. The final operating point was then chosen as the one with maximum gain from this cohort. Over the course of the experiment, linearity of the TWPA was periodically verified by measuring $g^{(2)}_{CC}$ for a weak coherent input. In case of a deviation of $g^{(2)}_{CC}$ from unity by one standard deviation, the experiment was halted and the pump optimization routine was re-run.

11. TRANSDUCER HEATING DYNAMICS

To investigate the heating dynamics of the transducer, we performed a series of microwave measurements under pulsed optical excitation. In Fig. S6a, we show the time-resolved power spectrum of the unconditional microwave emission at the transducer output port. For this measurement, we use the optical pulse sequence ($T_\mathrm{p} = 160$ ns, 50 kHz repetition rate) used in the SPDC experiments in the main text albeit at a higher pump power, $n_a = 12$, which generates higher thermal noise. We observe that majority of the heating occurs in the emission bandwidth of the hybridized electro-mechanical modes. The peak of this thermal emission is noticeably delayed from the optical pump pulse. Additionally, we observe a small but non-zero noise contribution far from the transducer resonances, which we attribute to heating of the microwave waveguide. These resonator and waveguide contributions to the added noise are separated by fitting the power spectrum at each delay to a double Lorentzian function with a constant floor. The fit result is shown in Fig. S6b, and indicates that resonator heating significantly dominates waveguide heating and has a stronger time dependence. This points to parasitic optical absorption in the intrinsic baths of the acoustic and microwave modes as the dominant source of transducer heating. Additionally, we measured the scaling of the intensity of this thermal emission with optical pump power and observed it to be sub-linear as shown by the results plotted in Fig. S6c. Finally, from the time dynamics of heating in Fig. S6a, we observe that while most of the thermal emission decays on the timescale of a few microseconds, a small component persists as steady state heating and is observed prior to the arrival of the optical pump pulse. We attribute this to a slower component of the hot bath generated by optical absorption and characterize it by varying the repetition rate of the optical pulses. The result shown in Fig. S6d confirms slow decay of this component of heating with an exponential decay time of $33 \pm 6\mu$s. The measurements described in this section were performed on the same device used for the SPDC experiments in the main text, but after a partial warm up and cooldown to base temperature. After this thermal cycle, we observed frequency shifts of the microwave and acoustic modes by a few MHz and an increase in the normal mode splitting. However, the thermal noise quanta generated by the optical pulse sequence used for the experiments in the main text remained nearly the same.

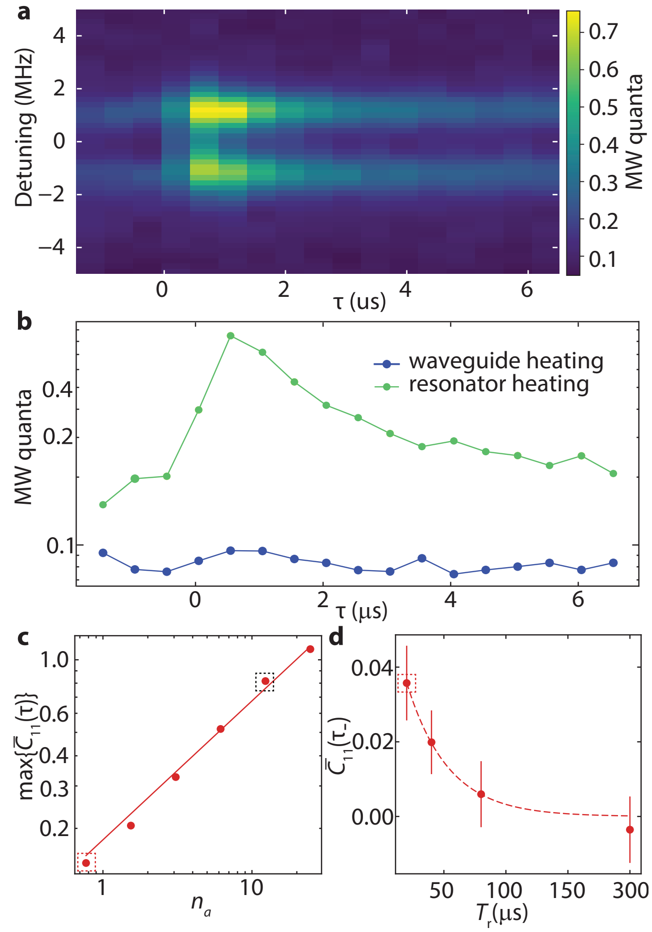

FIG. S6. Transducer heating dynamics. **a.** Power spectral density of the unconditional microwave signal as a function of delay, τ from the center of the optical pump pulse. Detuning is defined with respect to the mean frequency of the hybridized electromechanical modes. Legend indicates microwave quanta emitted at the output port in a 100 kHz bandwidth. Measurement is performed with optical pump pulses of peak power, $n_a = 12$, pulse duration of 160 ns, and 50 kHz repetition rate. **b.** Time dependence of waveguide and resonator components of heating extracted by fitting the data in panel **a.**. **c.** Thermal emission in the transducer microwave output as a function of optical pump power. Pump power is plotted in terms of peak intra-cavity occupation of the optical mode, n_a and microwave intensity is plotted in terms of occupation of the temporal mode, $\bar{C}_{11}$ measured at the delay, τ at which it is maximized. All measurements were performed with the 160 ns pulse duration and 50 kHz repetition rate used for data in the main text. Dashed red box denotes the power used for the experiment in the main text ($n_a = 0.8$) while the dashed black box denotes the power used for measurements in panels **a,b**. Solid line is a power law fit revealing an exponent, 0.58 ± 0.03. **d.** Decay of the slow component of the thermal emission probed by varying the repetition period, T_r of the pump pulse sequence. Microwave quanta in the unconditional state, $\bar{C}_{11}$ are measured versus prior to the pump pulse at the delay τ_- as defined in the main text. Dashed line is an exponential fit revealing a decay time of 33 ± 6 μs. Dashed red box denotes the repetition rate used for the experiment in the main text.

12. SIMULATION OF THE CONDITIONAL MICROWAVE STATE

As detailed in Eq. 1 of the main text, the dynamics of the transducer are governed by the Hamiltonian,

$$\hat{H}_{abc}(t) = \hat{H}_o + \hat{H}_\mathrm{om}(t) + \hat{H}_\mathrm{pe}, \tag{15}$$

where $\hat{H}_o$, $\hat{H}_\mathrm{om}(t)$, and $\hat{H}_\mathrm{pe}$ contain the mode frequencies, optomechanical interaction, and piezoelectric interaction terms respectively. These are defined as

$$\hat{H}_o = -\hbar\Delta_a \hat{a}^\dagger \hat{a} + \hbar\omega_b \hat{b}^\dagger \hat{b} + \hbar\omega_c \hat{c}^\dagger \hat{c}, \tag{16a}$$

$$\hat{H}_{\mathrm{om}}(t) = \hbar G_{\mathrm{om}}(t)(\hat{a}^\dagger \hat{b}^\dagger + \hat{a}\hat{b}), \tag{16b}$$

$$\hat{H}_{\mathrm{pe}} = \hbar g_{\mathrm{pe}}(\hat{b}^\dagger \hat{c} + \hat{b}\hat{c}^\dagger). \tag{16c}$$

We have explicitly included the time dependence of the optomechanical coupling $G_{\mathrm{om}}(t) = \sqrt{n_a(t)} g_{\mathrm{om}}$ due to the temporal shape of the intra-cavity pump photon number, $n_a(t) = \kappa_{\mathrm{e},a}/(\Delta_a^2 + \kappa_a^2/4) P_{\mathrm{in}}(t)$. $P_{\mathrm{in}}(t)$, the optical pump power at the device follows a Gaussian shape with parameters described in the main text. In our experiment, we use a blue detuned pump with $\Delta_a = \omega_b$. Further, our device is in the sideband resolved regime with $\omega_b \gg \kappa_a$.

In addition to unitary evolution due to couplings among the internal modes of the transducer, the system is also subject to dissipation and heating due to coupling to the environment. This is captured by the master equation,

$$\dot{\hat{\rho}}(t) = \mathcal{L}(t)\hat{\rho}(t). \tag{17}$$

Defining the Lindblad superoperator as $\mathcal{D}(\hat{o})\hat{\rho} = \hat{o}\hat{\rho}\hat{o}^\dagger - \frac{1}{2}\left\{\hat{o}^\dagger \hat{o}, \hat{\rho}\right\}$ the action of $\mathcal{L}(t)$ on the density matrix is written as

$$\mathcal{L}(t)\hat{\rho}(t) = -\frac{i}{\hbar}\left[\hat{H}_{abc}, \hat{\rho}(t)\right] + \left(\mathcal{L}_a + \mathcal{L}_b(t) + \mathcal{L}_c(t)\right)\hat{\rho}(t), \tag{18}$$

where the superoperators, $\mathcal{L}_a, \mathcal{L}_b, \mathcal{L}_c$ describe the coupling of the respective transducer modes to the environment according to the relations below.

$$\mathcal{L}_a \hat{\rho}(t) = \kappa_a \mathcal{D}(\hat{a})\hat{\rho}(t), \tag{19a}$$

$$\begin{aligned}\mathcal{L}_b(t)\hat{\rho}(t) = {} & n_{\mathrm{th},b}(t)\kappa_{\mathrm{i},b}\mathcal{D}(\hat{b}^\dagger)\hat{\rho}(t) \\ & + \left[n_{\mathrm{th,b}}(t) + 1\right]\kappa_{i,b}\mathcal{D}(\hat{b})\hat{\rho}(t),\end{aligned} \tag{19b}$$

$$\begin{aligned}\mathcal{L}_c(t)\hat{\rho}(t) = {} & n_{\mathrm{th},c}(t)\kappa_{\mathrm{i},c}\mathcal{D}(\hat{c}^\dagger)\hat{\rho}(t) \\ & + \left[n_{\mathrm{th},c}(t) + 1\right]\kappa_{i,c}\mathcal{D}(\hat{c})\hat{\rho}(t) \\ & + n_{\mathrm{th},w}\kappa_{\mathrm{e},c}\mathcal{D}(\hat{c}^\dagger)\hat{\rho}(t) \\ & + \left[n_{\mathrm{th},w} + 1\right]\kappa_{e,c}\mathcal{D}(\hat{c})\hat{\rho}(t).\end{aligned} \tag{19c}$$

Here the total dissipation rate of the optical mode, $\kappa_a = \kappa_{e,a} + \kappa_{i,a}$. To capture heating of the transducer modes due to parasitic absorption of the optical pump, we assume that the acoustic and microwave modes are coupled to intrinsic baths with time dependent thermal occupation, $n_{\mathrm{th},b}(t)$ and $n_{\mathrm{th},c}(t)$ respectively. $n_{\mathrm{th},w}$, the thermal occupation of the microwave waveguide is set to a constant value much smaller than $n_{\mathrm{th},b}(t)$ and $n_{\mathrm{th},c}(t)$ based on the measurements detailed in Section 11. We note that while these baths may be significantly more complex and involve several components possessing different, time-dependent coupling strengths, our model is aimed at capturing the total influx of thermal excitations into the transducer modes generated by the optical pulse. With this endeavor in mind, we assume that the acoustic (microwave) mode couples to a single intrinsic bath at a fixed dissipation rate, $\kappa_{i,b}$ ($\kappa_{i,c}$) and ascribe the time-dependence of the heating dynamics entirely to the thermal occupation of the bath. This phenomenological approach to heating induced by optical absorption has previously been used to model photon correlations in optomechanics experiments [15].

While Eq. (17) can be numerically solved in principle, this is computationally intensive owing to the large state space involving the three modes of the transducer. Instead, we take advantage of the fact that $\kappa_a \gg G_{\mathrm{om}}$ to adiabatically eliminate the optical mode [16]. Moving into a frame rotating with the mechanical and microwave modes and defining the optomechanical scattering rate, $\Gamma_{\mathrm{om}}(t) = 4|G_{\mathrm{om}}(t)|^2/\kappa_a$, we arrive at the simplified master equation,

$$\dot{\hat{\rho}}_r(t) = \mathcal{L}_r(t)\hat{\rho}_r(t), \tag{20}$$

where $\hat{\rho}_r = \mathrm{Tr}_a\{\hat{\rho}\}$ denotes the reduced density matrix spanning only the acoustic and microwave modes. $\mathcal{L}_r$ is defined by its action on the reduced density matrix,

$$\begin{aligned}\mathcal{L}_r(t)\hat{\rho}_r(t) = {} & -\frac{i}{\hbar}\left[\hat{H}_{\mathrm{pe}}, \hat{\rho}_r(t)\right] \\ & + \left(\mathcal{L}_{\mathrm{om}}(t) + \mathcal{L}_b(t) + \mathcal{L}_c(t)\right)\hat{\rho}_r(t),\end{aligned} \tag{21}$$

where we have introduced optomechanical scattering as a coupling of the reduced system to an effective bath according to the relation,

$$\mathcal{L}_{\mathrm{om}}(t)\hat{\rho}_r(t) = \Gamma_{\mathrm{om}}(t)\mathcal{D}(\hat{b}^\dagger)\hat{\rho}_r(t). \tag{22}$$

The Lindblad superoperator, $\mathcal{D}(\hat{b}^\dagger)$ contains a non-number preserving quantum jump operator which adds a phonon to the system due to optomechanical SPDC. This process is naturally correlated with the creation of an optical photon, which is routed with finite collection efficiency to a single photon detector. Given that optical decay ($\kappa_a/2\pi = 1.3$ GHz) occurs on a much faster timescale than the dynamics of the rest of the system ($\kappa_a \gg \kappa_b, g_{\mathrm{pe}}, \kappa_c$), optical detection is assumed to be instantaneous. Under this approximation, in the event of a click on the optical detector caused by a quantum jump at time t_{J}, the resultant conditional state of the reduced system is,

$$\hat{\rho}_{\mathrm{J}}(t_{\mathrm{J}}) = \frac{\hat{b}^\dagger \hat{\rho}_r(t_{\mathrm{J}})\hat{b}}{\mathrm{Tr}\left\{\hat{b}^\dagger \hat{\rho}_r(t_{\mathrm{J}})\hat{b}\right\}}. \tag{23}$$

This state then evolves according to Eq. (20). Since we operate in the weak pump regime where the integrated jump probability over the optical pulse duration, $\int \Gamma_{\mathrm{om}}(t)dt \ll 1$, we neglect two-fold SPDC events.

While the treatment above describes the evolution of the internal modes of the transducer, we experimentally measure the emission in the temporal mode, $\hat{C}$ in the transducer microwave output port. This temporal mode is linearly related to the internal microwave mode of the transducer as $\hat{C}(t) := \int \hat{c}_{\mathrm{out}}(t + t')f^*(t')dt' =$

$\sqrt{\kappa_{e,c}}\int \hat{c}(t+t')f^*(t')dt'$. We can thus compute the occupation of the temporal mode as

$$\begin{aligned}
&\langle \hat{C}^\dagger(t)\hat{C}(t)\rangle \\
&= \kappa_{e,c}\langle \int \hat{c}^\dagger(t+t'')f(t'')dt'' \int \hat{c}(t+t')f^*(t')dt'\rangle \\
&= \kappa_{e,c}\int \langle \hat{c}^\dagger(t+t'')\hat{c}(t+t')\rangle f(t'')f^*(t')dt''dt' \\
&= \kappa_{e,c}\int \langle \hat{c}^\dagger(t'+\tau')\hat{c}(t')\rangle f(t'+\tau'-t)f^*(t'-t)dt'd\tau',
\end{aligned} \tag{24}$$

where the correlator, $\langle \hat{c}^\dagger(t'+\tau')\hat{c}(t')\rangle$ inside the integral on the final line can be numerically computed using the quantum regression theorem [17]. Further, to model the conditional state due to a quantum jump at time t_{J}, this correlator is computed using $\hat{\rho}_{\mathrm{J}}(t_{\mathrm{J}})$ in Eq. (23) for times, $t > t_{\mathrm{J}}$. For times $t < t_{\mathrm{J}}$, we use the value for the unconditional state. Performing this calculation for a sequence of jump times, $\{t_{\mathrm{J}}\}$ associated with detector clicks received within the optical gating window, we obtain a sequence of conditional time traces for the occupation of the temporal mode, $\{\langle \hat{C}^\dagger\hat{C}\rangle|_{t_{\mathrm{J}}}\}$. The final conditional trace, $\langle \hat{C}^\dagger\hat{C}\rangle|_{\mathrm{click}}$ is given by a weighted average over this sequence with weights determined by the infinitesimal jump probability, $\delta p_{\mathrm{J}} \propto \Gamma_{\mathrm{om}}(t_{\mathrm{J}})dt$ in an interval dt at time t_{J}. The effects of detector dark counts as well as the optical pump filter in our experiment are incorporated via appropriate modifications to these weights.

We perform a numerical simulation of our system by implementing the above model using the QuTiP software package [14]. Solutions to the master equation as well as the required correlators are evaluated in a 10×10 Fock space of the acoustic and microwave modes. To incorporate effects of heating of the microwave and acoustic baths separately in our model, we measured the unconditional microwave emission from the transducer in response to optical pump pulses under two conditions (i) mechanics far detuned from the microwave mode, $\omega_c - \omega_b \gg 2g_{\mathrm{pe}}$, (ii) mechanics on resonance with the microwave mode, $\omega_b = \omega_c$. The results of the measurement are shown with the dotted traces in Fig. S7d. Under condition (i), which is achieved at zero magnetic field, the microwave emission is approximately entirely due to heating of the microwave bath alone. On the other hand, condition (ii) corresponds to maximal electro-mechanical hybridization and is relevant to the SPDC experiments described in the main text. In this setting, heating of both microwave and acoustic baths is expected to contribute to the measured microwave output signal. Using measurements of the microwave emission under both conditions, we invert the master equation to determine $n_{\mathrm{th,b}}(t)$ and $n_{\mathrm{th,c}}(t)$ approximated as a piece-wise linear function over coarse samples as shown in Fig. S7b. The inversion is performed by choosing an ansatz for $n_{\mathrm{th,b}}(t)$ and $n_{\mathrm{th,c}}(t)$, iteratively solving the master equation and performing least mean square optimization with respect to the experimentally measured thermal microwave emis-

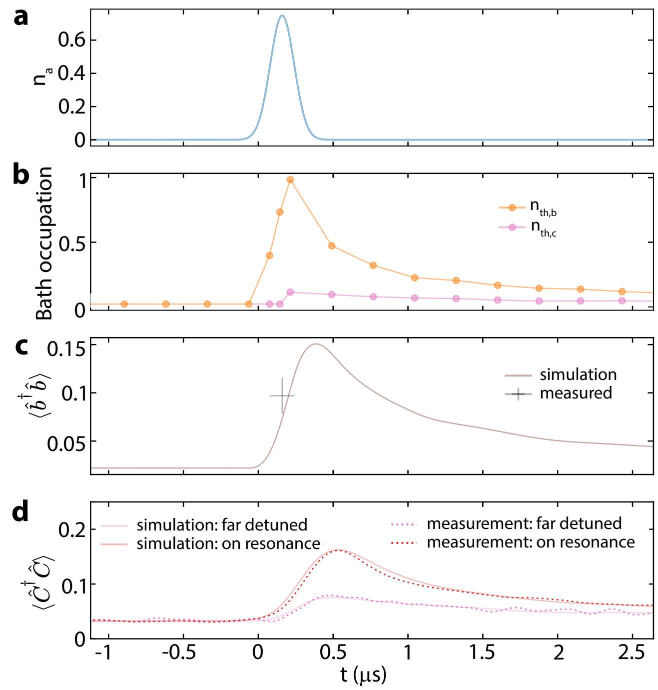

FIG. S7. Simulation of transducer heating dynamics **a.** Time trace of the optical pump pulse plotted in terms of the intracavity occupation of the optical mode, n_a. **b.** Model used for time dependent thermal occupation of the intrinsic baths of the acoustic and microwave modes, $n_{\mathrm{th},b}(t)$, $n_{\mathrm{th},c}(t)$. **c.** Time dependent occupation of the acoustic mode, $\langle \hat{b}^\dagger\hat{b}\rangle$ obtained from the master equation simulation. Gray data point corresponds to experimentally recorded thermal occupation of the acoustic mode as inferred from an optomechanical sideband asymmetry measurement. The vertical bar on the point represents a confidence interval spanning two standard deviations while the horizontal bar indicates the FWHM duration of the optical pulse. **c.** Time dependent occupation of the temporal mode, $\langle \hat{C}^\dagger\hat{C}\rangle$ in the microwave output port of the transducer from the master equation simulation (solid traces) and from experimental heterodyne data (dotted traces). The condition where the acoustic mode is far detuned (on resonance) with respect to the microwave mode is shown with pink (red) traces.

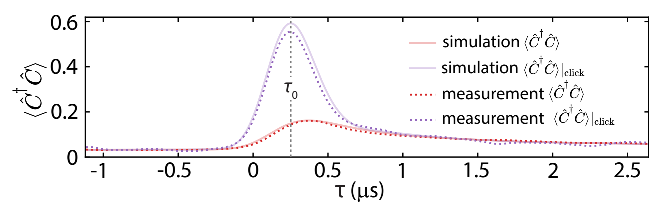

FIG. S8. Simulation of the conditional microwave state. Occupation of the temporal mode, $\hat{C}$ in the microwave output port of the transducer from simulation (solid traces) and from experimental heterodyne data (dotted traces). The occupation of the microwave state conditioned on an optical click, $\langle \hat{C}^\dagger\hat{C}\rangle|_{\mathrm{click}}$ is shown with purple traces. The occupation of the unconditional microwave state, $\langle \hat{C}^\dagger\hat{C}\rangle$ is shown with red traces for comparison.

sion. The resulting occupations of the acoustic mode, $\hat{b}$ and of the unconditional temporal microwave mode, $\hat{C}$ obtained by solving the master equation are shown in Fig. S7c and with the solid traces in Fig. S7d respectively. While the model for the hot baths is determined entirely using microwave measurements, we observe that the simulated occupation of the acoustic mode is in reasonable agreement with an experimental measurement shown by the gray data point in Fig. S7c. This data point corresponds to the optomechanical sideband asymmetry measurement described in Section 6, and represents the thermal occupation of the acoustic mode averaged over the duration of the optical pump pulse. Finally, the conditional microwave emission determined using our model shown with the solid purple trace in Fig. S8 is in good agreement with the corresponding experimental result shown via the dotted purple trace. The simulated conditional and unconditional intensity traces plotted in this figure predict a value of $g^{(2)}_{AC}(\tau_o) = 3.97$, which is in good agreement with the experimentally observed value of 3.90.

Simulation of $g^{(2)}_{CC}|_{\text{click}}$ requires evaluation of $\langle \hat{C}^{\dagger 2}\hat{C}^2\rangle|_{\text{click}}$. Extending the approach of Eq. (24) to the second order moment is a nontrivial computational endeavor as it requires calculation of the four-time correlator, $\langle \hat{c}^\dagger(t + t'''')\hat{c}^\dagger(t + t''')\hat{c}(t + t'')\hat{c}(t + t')\rangle$. For the purpose of this work, we pursue the more tractable task of placing bounds on $g^{(2)}_{CC}|_{\text{click}}$. First, we consider the conditional state of the acoustic mode immediately after the addition of a phonon as triggered by an SPDC event. Since the extraction of this acoustic state into the microwave output port is accompanied by loss and addition of noise, we expect the function, $g^{(2)}_{bb}|_{\text{click}} = \langle \hat{b}^{\dagger 2} b^2\rangle|_{\text{click}}/(\langle \hat{b}^\dagger b\rangle|_{\text{click}})^2$ to provide a lower bound for $g^{(2)}_{CC}|_{\text{click}}$. Our model estimates $g^{(2)}_{bb}|_{\text{click}}(0) = 0.24$. Next, we consider a choice of the emission envelope function, $f(t) = \delta(t - T_o - \tau_o)$, where T_o is the time corresponding to the peak intensity of the pump pulse. Physically, employing this function results in collecting the microwave emission at all frequencies from the transducer at a fixed time. This is strictly less optimal than the choice in our measurements, namely a function coherently matched to the theoretically expected single photon wavepacket as discussed in Sec. 8. In the event of this sub-optimal choice, moments of the temporal mode, $\hat{C}$ are proportional to the corresponding moments of the internal microwave mode, $\hat{c}$ and the function, $g^{(2)}_{cc}|_{\text{click}} = \langle \hat{c}^{\dagger 2} c^2\rangle|_{\text{click}}/(\langle \hat{c}^\dagger c\rangle|_{\text{click}})^2$, serves as an upper bound for $g^{(2)}_{CC}|_{\text{click}}$. Our model estimates $g^{(2)}_{cc}|_{\text{click}}(\tau_o) = 0.77$. In summary, the above arguments allow us to use numerical simulations to bound $g^{(2)}_{CC}|_{\text{click}}(\tau_o)$ to the interval (0.24, 0.77), which has substantial overlap with the experimental observation of $g^{(2)}_{CC}|_{\text{click}}(\tau_o) = 0.42^{+0.27}_{-0.28}$.

13. DATA ANALYSIS FOR CORRELATION FUNCTIONS

For measurements of $g^{(2)}_{AC}$ and $g^{(2)}_{CC}|_{\text{click}}$ presented in figure 3 of the main text, 9.1×10^4 heterodyne voltage traces were acquired over a one month period. For every ∼16 minutes of acquisition of conditional microwave data, we interleaved the acquisition of voltage samples of amplifier noise and the unconditional microwave signal. The unconditional microwave intensity signal, $\bar{C}_{11}(\tau)$ used to evaluate the normalized cross-correlation function, $g^{(2)}_{AC}$ was obtained from 1.4×10^7 heterodyne voltage traces. Likewise, the measurement of the normalized second order intensity correlation function, $g^{(2)}_{CC}(\tau_\text{o})$ used 3.2×10^7 heterodyne voltage samples.

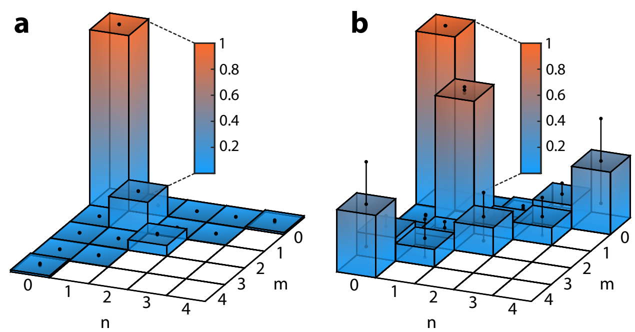

FIG. S9. **a.** Moments matrix for the unconditional microwave output of the transducer, $\bar{C}_{mn}$. **b.** Moments matrix for the microwave output of the transducer conditioned on receipt of an optical click, $\bar{C}_{mn}|_{\text{click}}$. Error bars calculated via bootstrapping with replacement are indicated by the black pins.

To minimize the effect of long term fluctuations in the added noise and gain of the TWPA, we divide the dataset into chunks corresponding to acquisition over one day, and invert each separately to extract the moments $\bar{C}^{(k)}_{mn}, \bar{C}^{(k)}_{mn}|_{\text{click}}$, where the index k runs over the daily chunks. This inversion process follows the methods detailed in Sec. 7 and assumes that no correlations exist between $\hat{H}$ and $\hat{A}$ as well as $\hat{H}$ and $\hat{C}$. We then take a weighted average to compute the entries of the moments matrix, $\bar{C}_{mn} = \sum_k w^{(k)}\bar{C}^{(k)}_{mn}$ and $\bar{C}_{mn}|_{\text{click}} = \sum_k w^{(k)}_{\text{click}}\bar{C}^{(k)}_{mn}|_{\text{click}}$, where $w^{(k)}$, $w^{(k)}_{\text{click}}$ denote the fraction of records in the k^{th} chunk of the unconditional and conditional datasets respectively. The moments matrices constructed from this process are shown in Fig. S9b.

The acquisition rate for the moments of the conditional heterodyne output, $\bar{S}_{\text{mn}}|_{\text{click}}$ is determined by the optical heralding rate, $R_{\text{click}} = 0.14$ Hz. This is much slower compared to that for the amplifier noise moments, $\bar{H}_{\text{mn}}$, which can be acquired at the 50 kHz repetition rate of the experiment and determined with high accuracy. As a result, the error in $g^{(2)}_{CC}|_{\text{click}}$ is dominated by uncertainty in $\bar{S}_{\text{mn}}|_{\text{click}}$. To calculate error in $g^{(2)}_{CC}$ and $g^{(2)}_{CC}|_{\text{click}}$, we employ bootstrapping with replacement using 10^5 bootstraps to construct the probability density functions shown in Fig. S10a. Finally, error estimates were calculated by numerical integration of the distributions such that they they cover a 34.1% confidence interval above and below the mean. The error estimates calculated using this process vs. number of bootstraps is shown in Fig. S10b-c for the conditional and unconditional datasets. We find that the use of 10^5 bootstraps yields acceptable convergence in the error estimate.

To mitigate the effects of long term drifts in frequencies of the microwave modes, we periodically measured the unconditional microwave power spectrum and monitored the resonance frequencies, $\omega_\pm$. This additionally allowed us to characterize the spectral diffusion of the microwave modes. In our data analysis, we excluded intervals of the measurement where the mode frequencies drifted by more than twice the standard deviation associated with spectral diffusion, which represented 3% of all recorded traces.

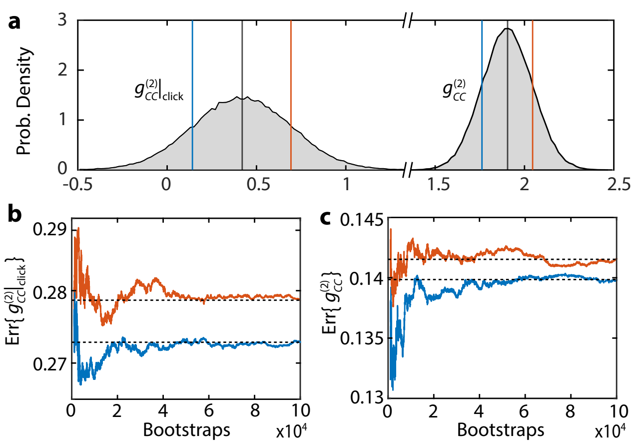

FIG. S10. **a.** Probability distribution for $g^{(2)}_{CC}$ and $g^{(2)}_{CC}|_{\rm click}$ calculated from bootstrapping. Vertical blue and red lines correspond to one standard deviation above and below the distribution mean respectively. **b,c.** Convergence of the error estimate vs. number of bootstraps for all traces, and optically heralded traces respectively. Blue lines correspond to positive error, red lines to negative error, and black lines indicate the error estimate used in the main text.

14. CLASSICAL BOUND ON THE CONDITIONAL SECOND ORDER INTENSITY CORRELATION

In the main text, we claim that the conditional microwave autocorrelation function is expected to satisfy the inequality, $g^{(2)}_{CC}|_{\rm click} \geq 1$ for classical microwave-optical states. This can be proved by rewriting the conditional quantities defined in the main text in terms of correlators of the joint microwave-optical state,

$$\begin{aligned} g^{(2)}_{CC}(\tau)|_{\rm click} &= \frac{\langle \hat{C}^{\dagger 2}(\tau)\hat{C}^2(\tau)\rangle\,|_{\rm click}}{\left(\langle \hat{C}^\dagger(\tau)\hat{C}(\tau)\rangle\,|_{\rm click}\right)^2} \\ &= \frac{\langle \hat{A}^\dagger \hat{C}^{\dagger 2}(\tau)\hat{C}^2(\tau)\hat{A}\rangle\,\langle \hat{A}^\dagger\hat{A}\rangle}{\left(\langle \hat{A}^\dagger\hat{A}\rangle\,\langle \hat{C}^\dagger(\tau)\hat{C}(\tau)\rangle\right)^2}, \end{aligned} \tag{25}$$

where the above correlators are explicitly written in normal order. Using the the optical equivalence theorem, we can recast the expectation values of the above correlators in phase space using the Sudarshan-Glauber P representation [18, 19] as

$$\begin{aligned} \langle \hat{A}^{\dagger m}\hat{C}^{\dagger n}(\tau)\hat{A}^m\hat{C}^n(\tau)\rangle &= \int |\alpha|^{2m}|\gamma|^{2n}P_\tau(\alpha,\gamma)d^2\alpha d^2\gamma \\ &:= \langle |\alpha|^{2m}|\gamma|^{2n}\rangle_{P_\tau}, \end{aligned} \tag{26}$$

where $P_\tau(\alpha,\gamma)$ is the joint phase space density cprresponding to modes $\hat{A}$ and $\hat{C}$ with relative delay τ. Applying this directly to Eq. (25) gives,

$$g^{(2)}_{CC}(\tau)|_{\rm click} = \frac{\langle |\alpha|_P^2|\gamma|^4\rangle_{P_\tau}\,\langle|\alpha|^2\rangle_{P_\tau}}{\langle|\alpha|^2|\gamma|^2\rangle^2_{P_\tau}} \geq 1, \tag{27}$$

which follows directly from the Cauchy-Schwarz inequality.

15. CONVERGENCE OF $g^{(2)}_{AC}$ TO THE CLASSICAL BOUND

The results of pump power dependent measurements of the normalized intensity cross-correlation function, $g^{(2)}_{AC}$ are shown in Fig. S11.

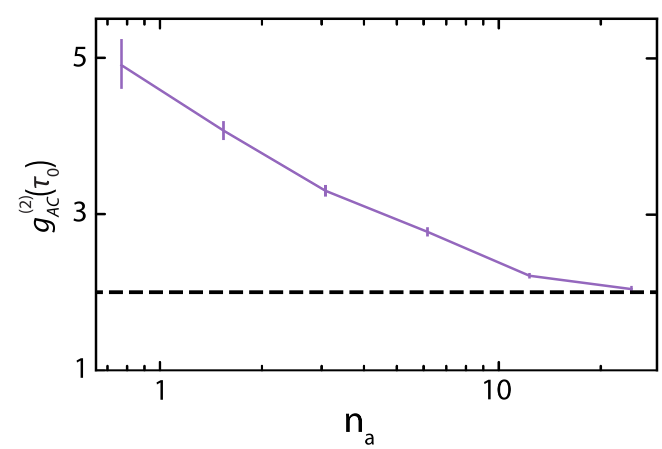

FIG. S11. Maximum value of $g^{(2)}_{AC}$ versus increasing pump power plotted on the horizontal axis as the peak intra-cavity occupation of the optical mode, n_a. Dashed line denotes the classical upper bound of 2.

All measurements were performed with the same pump pulse duration and repetition rate used to acquire data in the main text. We observe that at higher pump powers, the value of $g^{(2)}_{AC}$ approaches the classical upper bound of 2 as expected from the increased thermal noise added to the photon pair generated in SPDC. The data in this figure was collected from the same device used for the experiments in the main text albeit after a partial warm up and cooldown to base temperature. This thermal cycle led to modified device parameters, which produced an increase in $g^{(2)}_{AC}$ at the lowest pump power of $n_a = 0.8$ compared to the result presented in the main text.

[1] M. Xu, X. Han, W. Fu, C.-L. Zou, and H. X. Tang, Frequency-tunable high-q superconducting resonators via wireless control of nonlinear kinetic inductance, Appl. Phys. Lett. **114**, 192601 (2019).

[2] J. Zmuidzinas, Superconducting microresonators: Physics and applications, Annu. Rev. Condens. Matter Phys. **3**, 169 (2012).

[3] M. Mirhosseini, A. Sipahigil, M. Kalaee, and O. J. Painter, Superconducting qubit to optical photon transduction, Nature **588**, 599–603 (2020).

[4] C. Macklin, K. O'Brien, D. Hover, M. E. Schwartz, V. Bolkhovsky, X. Zhang, W. D. Oliver, and I. Siddiqi, A near quantum-limited josephson traveling-wave parametric amplifier, Science **350**, 307 (2015).

[5] S. M. Meenehan, J. D. Cohen, G. S. MacCabe, F. Marsili, M. D. Shaw, and O. Painter, Pulsed excitation dynamics of an optomechanical crystal resonator near its quantum ground state of motion, Phys. Rev. X **5**, 041002 (2015).

[6] C. Eichler, D. Bozyigit, C. Lang, L. Steffen, J. Fink, and A. Wallraff, Experimental state tomography of itinerant single microwave photons, Phys. Rev. Lett. **106**, 220503 (2011).

[7] D. Bozyigit, C. Lang, L. Steffen, J. M. Fink, C. Eichler, M. Baur, R. Bianchetti, P. J. Leek, S. Filipp, M. P. da Silva, A. Blais, and A. Wallraff, Antibunching of microwave-frequency photons observed in correlation measurements using linear detectors, Nat. Phys. **7**, 154 (2011).

[8] B. Kannan, D. L. Campbell, F. Vasconcelos, R. Winik, D. K. Kim, M. Kjaergaard, P. Krantz, A. Melville, B. M. Niedzielski, J. L. Yoder, T. P. Orlando, S. Gustavsson, and W. D. Oliver, Generating spatially entangled itinerant photons with waveguide quantum electrodynamics, Sci. Adv. **6**, eabb8780 (2020).

[9] V. S. Ferreira, G. Kim, A. Butler, H. Pichler, and O. Painter, Deterministic generation of multidimensional photonic cluster states with a single quantum emitter (2022), arXiv:2206.10076.

[10] C. M. Caves, Quantum limits on noise in linear amplifiers, Phys. Rev. D **26**, 1817 (1982).

[11] M. P. da Silva, D. Bozyigit, A. Wallraff, and A. Blais, Schemes for the observation of photon correlation functions in circuit qed with linear detectors, Phys. Rev. A **82**, 043804 (2010).

[12] S. M. Barnett, G. Ferenczi, C. R. Gilson, and F. C. Speirits, Statistics of photon-subtracted and photon-added states, Phys. Rev. A **98**, 013809 (2018).

[13] C. Zhong, X. Han, H. X. Tang, and L. Jiang, Entanglement of microwave-optical modes in a strongly coupled electro-optomechanical system, Phys. Rev. A **101**, 032345 (2020).

[14] J. Johansson, P. Nation, and F. Nori, Qutip 2: A python framework for the dynamics of open quantum systems, Computer Physics Communications **184**, 1234 (2013).

[15] S. Hong, R. Riedinger, I. Marinković, A. Wallucks, S. G. Hofer, R. A. Norte, M. Aspelmeyer, and S. Gröblacher, Hanbury brown and twiss interferometry of single phonons from an optomechanical resonator, Science **358**, 203 (2017), https://www.science.org/doi/pdf/10.1126/science.aan7939.

[16] I. Wilson-Rae, N. Nooshi, J. Dobrindt, T. J. Kippenberg, and W. Zwerger, Cavity-assisted backaction cooling of mechanical resonators, New Journal of Physics **10**, 095007 (2008).

[17] C. W. Gardiner and P. Zoller, *Quantum Noise* (Springer Series in Synergetics, Springer, 2004).

[18] L. Mandel and E. Wolf, *Optical coherence and quantum optics* (Cambridge university press, 1995).

[19] A. Kuzmich, W. P. Bowen, A. D. Boozer, A. Boca, C. W. Chou, L.-M. Duan, and H. J. Kimble, Generation of nonclassical photon pairs for scalable quantum communication with atomic ensembles, Nature **423**, 731 (2003).