nature physics

Article

https://doi.org/10.1038/s41567-024-02409-z

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

In the format provided by the authors and unedited

CONTENTS

1. NOTATION AND DEVICE PARAMETERS

Symbol	Description
$\hat{a}, \hat{b}, \hat{c}$	operators for optical, acoustic, and microwave modes of the transducer
$\hat{c}_+, \hat{c}_-$	operators for hybridized electromechanical modes
$\hat{a}_{\rm in}, \hat{c}_{\rm in}, (\hat{a}_{\rm out}, \hat{c}_{\rm out})$	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
$\bar{C}_{mn}\|_{\rm click}$	moments of the temporal microwave mode conditioned on an optical click

TABLE S1. Notation for various modes and moments.

Symbol	Description	Value
λ_a	optical mode wavelength	1561.3 nm
$\omega_b/2\pi$	microwave acoustic mode frequency	5.001 GHz
$\omega_c/2\pi\|_{B=0}$	microwave electrical mode frequency at zero B-field	5.011 GHz
$g_{\rm om}/2\pi$	optomechanical coupling	270 kHz
$g_{\rm pe}/2\pi$	piezoelectric coupling	800 kHz
$\kappa_{e,a}/2\pi$	external optical coupling rate	650 MHz
$\kappa_{i,a}/2\pi$	intrinsic optical loss rate	650 MHz
$\kappa_{i,b}/2\pi$	intrinsic acoustic loss rate	150 kHz
$\kappa_{e,c}/2\pi$	external microwave coupling rate	1.2 MHz
$\kappa_{i,c}/2\pi$	intrinsic microwave loss rate	550 kHz

TABLE S2. Coupling rates of transducer internal modes.

Symbol	Description	Value
T_p	Two sigma duration of pump pulse	160 ns
T_r	Repetition period of pump pulses	20 µs
-	Peak power of pump pulse	83 nW
n_a	Peak intra-cavity photon number	0.78
$p_{\rm click}$	Optical heralding probability	2.7×10^{-6}
$R_{\rm click}$	Optical heralding rate ($= p_{\rm click}/T_r$)	0.14 s^{-1}
p	SPDC scattering probability	1.8×10^{-4}
$\eta_{\rm opt}$	Optical collection efficiency	1.7×10^{-2}
$\eta_{\rm mw}$	Microwave conversion efficiency	0.56

TABLE S3. Microwave-optical photon pair generation parameters.

Parameter	Value
Optical cavity to on-chip waveguide	0.50
Waveguide to lensed fiber	0.29
Circulator between excitation and detection	0.90
Total filter bank loss (individual components below)	**0.16**
2x2 switches	0.89
Filters (CW transmission)	0.52
Finite pulse bandwidth	0.39
Circulator in filter setup	0.89
SNSPD setup	0.83
Optical collection efficiency ($\eta_{\rm opt}$)	1.7×10^{-2}

TABLE S4. Loss budget along optical detection path.

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 fre-

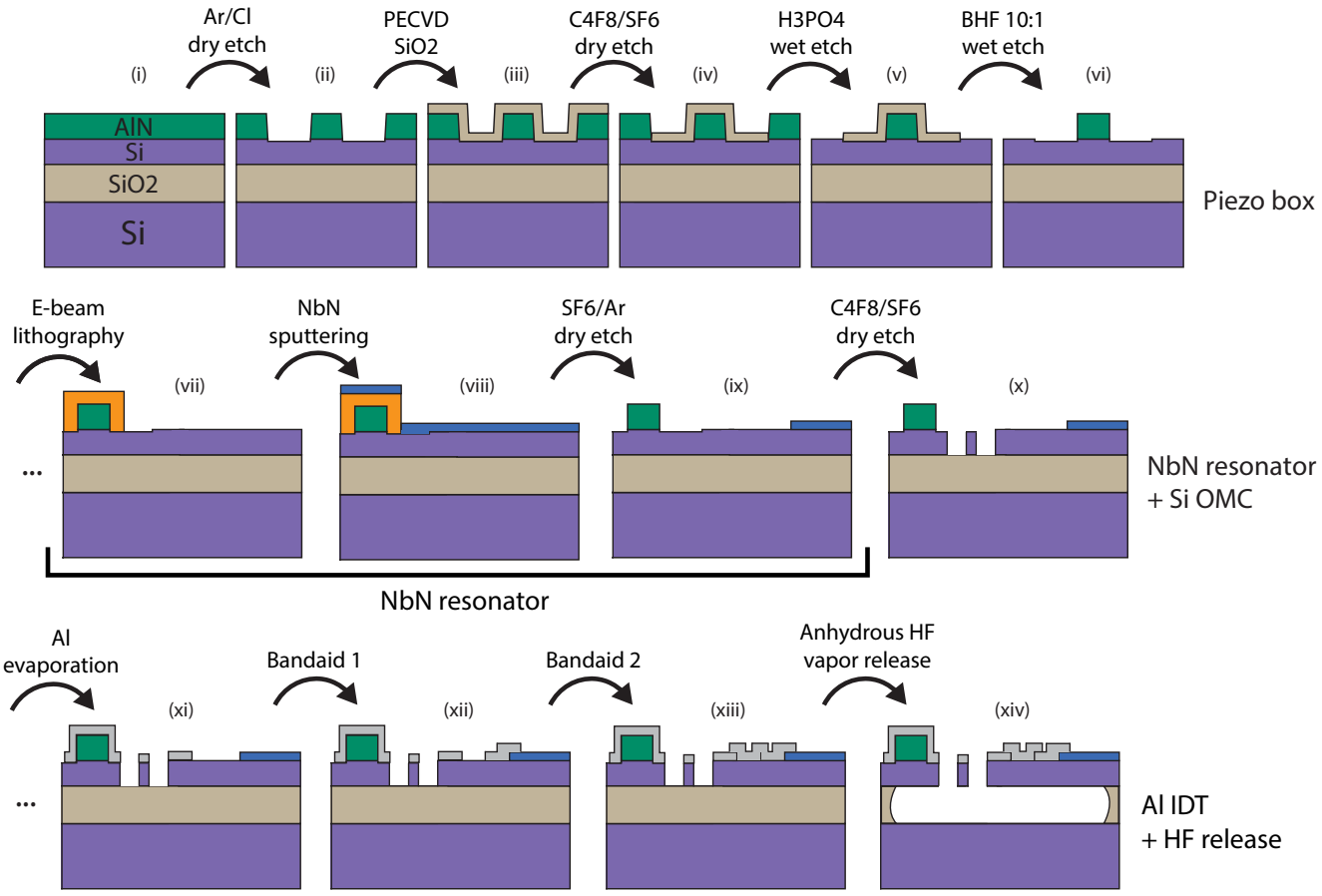

FIG. S1. **Device fabrication process**. Images are not to scale. **(i)** Sputter deposition of 300 nm thick c-axis AlN piezoelectric film (grown by OEM group; stress $T = +55$ MPa; (002) XRD peak of full-width at half-maximum = 1.79°) on a silicon on insulator substrate (Si device layer [float zone grown, 220 nm thick, $\rho \geq 5$kΩ-cm]; buried oxide layer [3 µm thick, silicon dioxide]; Si handle [Czochralski grown, 750 µm thick, $\rho \geq 5$kΩ-cm]). **(ii)** 3 minute AlN trench etch (120W RF, 600W ICP) with Ar/Cl_2 (80/40 sccm) chemistry to define the perimeter of the piezo resonator via a small trench of width ∼100 nm. **(iii)** Conformal deposition of a 300nm SiO_x hard mask via plasma enhanced chemical vapor deposition. **(iv)** Patterning of the SiO_x mask via a 6 minute dry etch (33W RF, 1300W ICP) with C_4F_8/SF_6 (40/3 sccm) chemistry. **(v)** Removal of the remaining AlN on the chip with a 12 minute 80° C H_3PO_4 wet etch. **(vi)** SiO_x mask removal with 10:1 buffered oxide etchant for 3 minutes and 30 seconds. **(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 sputtering process. **(ix)** 1 minute dry etch (30W RF, 300W ICP) of NbN with SF_6/Ar (40/20 sccm) chemistry to define the high-impedance microwave resonator. **(x)** 4 minute dry etch (18W RF, 600W ICP) of silicon with C_4F_8/SF_6 (72/30 sccm) 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 3 minute Ar milling followed by Al evaporation. **(xiv)** Etching of the buried oxide (BOX) layer using anhydrous vapor HF to release the device membrane.

quency of 5.0 GHz for the fundamental mode with a capacitance, $C_{\rm res} = 7.1$ 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 $L_k \approx L_{k,0}[1 + (I/I_*)^2]$ 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. Fig. 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 OMC design is unchanged from previous work. The primary difference 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]. Fig. S3 shows the new design of the piezo-acoustic cavity and the phononic shield with relevant dimensions.

Figs. 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. Both coupling rates on the device in our experiment are lower than the maximum design values, potentially due to fabrication disorder in the piezo-acoustic section. 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 long waveguide shown in Fig. 1b of the main text is designed with alternating S-shaped curved sections (length per section = 10μm, radius of curvature = 6 μm) in order to increase robustness against buckling from intrinsic stress in the silicon device layer. The tethers used to anchor the waveguide to the bulk are designed with adiabatically tapered width in order to reduce optical energy density at the edges, and thereby, reduce scattering loss at the anchor points Burek2017. We use a Gaussian taper profile, where the waveguide width is increased from the nominal value of 550 nm to 1.1 μm over a waveguide length of 5 μm. Through an FDTD simulation, we find that this achieves a loss of ~ 0.01 dB per tether and a loss of ~ 0.5 dB over the entire waveguide length, which has a total of 52 tethers. While this propagation loss is quite low, we find that in practice, the long waveguide adds significant ripple ($\sim$3 dB), likely due to finite back-reflection from the coupler tip at the edge of the chip. Based on the exact wavelength of the transducer optical mode with respect to this ripple, the total fiber-to-cavity coupling efficiency varies from 55% to 25% across devices on the same chip. With optimization of the coupler in future devices, we hope to mitigate the detrimental effect of this ripple on fiber-to-waveguide coupling efficiency.

5. MEASUREMENT SETUP

The measurement setup used in this work is detailed in Fig. S4. 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.S4a. 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. S4a. 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. S4a) 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

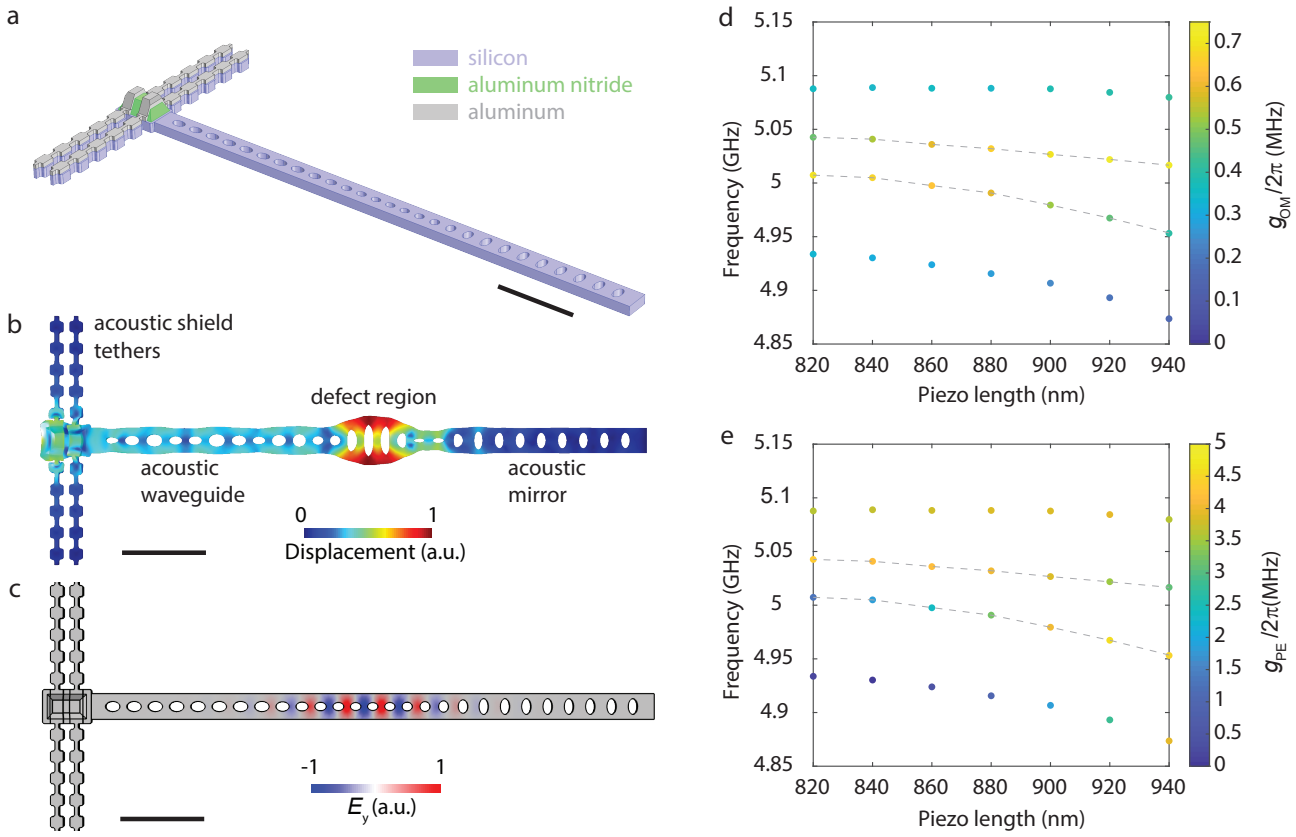

FIG. S2. **Piezo-optomechanics design**. **a.** Geometry and material composition of the piezo-optomechanical transducer. The black scale bar corresponds to 2 µm **b,c.** Top-down views detailing the displacement profile of a hybridized acoustic mode and the transverse electric field of the fundamental 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}.

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. S4b, 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. Using a microwave switch (Minicircuits ZASWA-2-50DRA+), 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. S4c begins with a Josephson traveling-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) ampli-

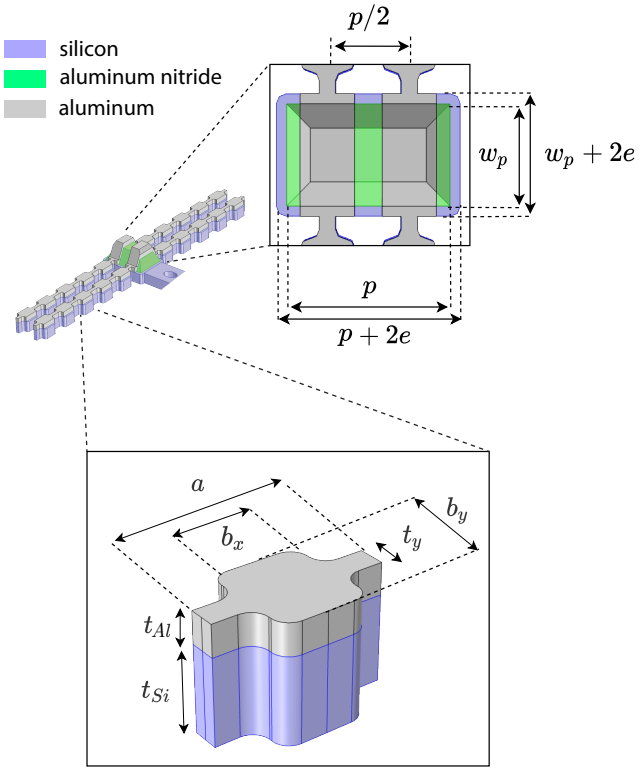

FIG. S3. **Piezo-acoustic cavity design**. Geometry of the piezo-acoustic section of the transducer. Relevant dimensions of the piezo-acoustic cavity are labeled in the top inset. In our design, (p, w_p, e) = (880, 600, 20) nm. The aluminum nitride layer has a thickness of 300nm with a $\sim 60^o$ sidewall angle. Relevant dimensions on a single unit cell of the phononic shield are labeled in the bottom inset. In our design, $(a, b_x, b_y, t_y, t_{Al}, t_{Si})$ = (492, 300, 369, 90, 100, 210) nm.

fier mounted on the 4K plate. In the setup outside the fridge as shown in Fig. S4b, 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 down-converted 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, $\hat{s}_{\text{out}}$. 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. S4b 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 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 $T_{\text{f}} \approx 20$ mK.

6. PIEZOELECTRIC COUPLING TO OTHER MODES

Fig. S5 shows microwave reflection spectroscopy of the transducer-resonator system over an extended frequency range and magnetic field range compared with Fig. 2b in the main text. In addition to the anti-crossing at ∼5.00GHz involving the main acoustic resonance of interest shown in the main text, we detect anti-crossings with two other modes at ∼4.97GHz and ∼4.93GHz with coupling rates of 2.7MHz and 1.4MHz, respectively. These two modes are also detected in the continuous wave microwave-to-optical transduction spectrum shown in Fig. 2a of the main text, thereby establishing their association with the piezo-optomechanical transducer. The SPDC experiments in the main text are performed at the maximal electromechanical hybridization point for the ∼5.00GHz mode. In this setting, the other acoustic modes are far-detuned, and their parasitic participation in the electrical resonator is estimated at $\sim 0.5\%$ based on the measured piezoelectric coupling rates.

7. OPTICAL HERALDING RATE CONTRIBUTIONS

The optical heralding events in our SPDC experiments have a finite noise contribution from technical sources in our experimental setup, namely detector dark counts and pump laser leakage. In Fig. S6a, we plot the optical photon count rate due to these noise sources as determined from independent measurements. The blue trace corresponds to the total optical signal as shown in Fig. 4a 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

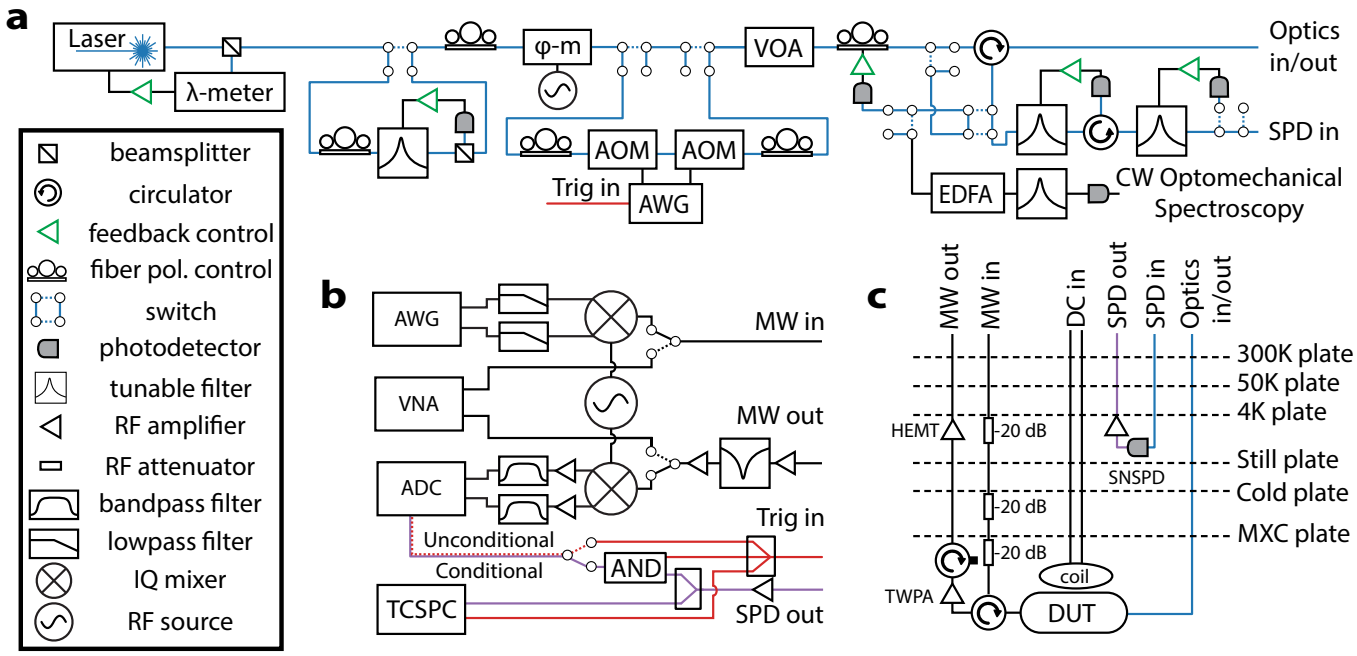

FIG. S4. **Experimental setup**. **a.** Optical pulse synthesis and detection. An external cavity diode laser is wavelength locked to a wavemeter, and a 50 MHz tunable narrowband filter is used to remove spontaneous emission noise. ϕ-m denotes an electro-optic phase modulator, which is used during the locking sequence for the pump filter bank in the detection path. 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, CW transmission loss of 3dB, bandwidth of 2.7MHz for the cascaded system) 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 with a microwave switch 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.

indicated by the purple trace in Fig. S6 is determined by detuning the filter cavity by 100 MHz from the acoustic mode used in the transduction experiment. 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, $p_{\text{click}} = 2.7 \times 10^{-6}$ as shown in Fig. S6b. The corresponding fractional contributions are enumerated in Tab. S5.

Source of a photon click	Fraction of clicks
SPDC signal	0.796
DCR	0.171
Pump leakage	0.033

TABLE S5. Contributions to the optical photon count rate.

To complement microwave measurements of pump-

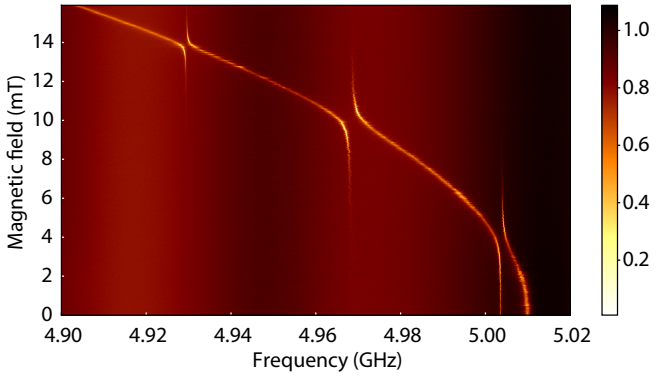

FIG. S5. Microwave reflection spectrum of the transducer-resonator system probed as a function of external magnetic field. Avoided crossings are observed between acoustic modes of the transducer and the tunable microwave electrical resonator.

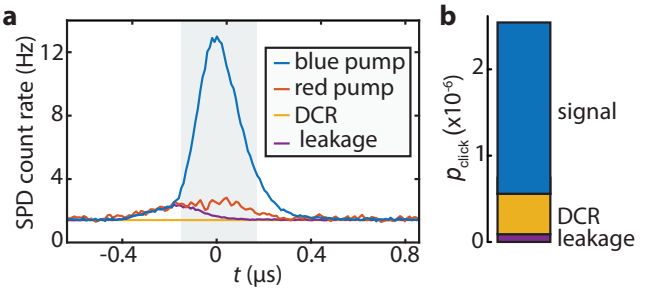

FIG. S6. **a.** Optical photon count rate vs time as measured with blue detuned pump pulses, red detuned pump pulses and in calibration measurements for the dark count rate (DCR) and pump leakage through the filter setup. **b.** Contributions to the optical click probability during the gating window extracted from the data in **a**.

induced noise, 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 ($\Delta_a = -(\omega_+ + \omega_-)/2$). The resultant photon flux from this measurement is shown by the red trace in Fig. S6a. 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. In Section 14 of the SI, we develop a heating model for the transducer, and find that this measurement result is consistent with microwave measurements of heating discussed in Fig. 3 of the main text.

8. 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],

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

where $\hat{c}_{\text{out}}$, $\hat{s}_{\text{out}}$ 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

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

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 $\hat{S}(t) := \int \hat{s}_{\text{out}}(t+t')f^*(t')dt'$. With this choice Eq. (2) becomes,

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

where $\hat{C}(t) := \int \hat{c}_{\text{out}}(t+t')f^*(t')dt'$ 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 ($n_{\text{th},H}$) of the state constitutes full knowledge of the statistics for a thermal state. Accordingly, we extract $n_{\text{th},H}$ from the reference measurements and calculate $\bar{H}_{mn}$ as [12],

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

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

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

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

9. MICROWAVE EMISSION ENVELOPE FUNCTION

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 $\omega_{\pm} = \omega_b \pm g_{\rm pe}$.

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],

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

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 $2\pi/2g_{\rm pe} = 625$ 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, $\left(e^{i\phi(t)/2}\hat{a}_+^\dagger + e^{-i\phi(t)/2}\hat{a}_-^\dagger\right)|\text{vac}\rangle$, where the relative phase $\phi(t) = \phi_o + 2g_{\rm pe}t$ 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 $|\psi\rangle_{\rm click} = \left(e^{-i\phi(t)/2}\hat{c}_+^\dagger + e^{i\phi(t)/2}\hat{c}_-^\dagger\right)|\text{vac}\rangle$.

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, $\langle\hat{c}_{\rm out}^\dagger(\omega)\hat{c}_{\rm out}(\omega)\rangle|_{\rm click}$ shown with the blue trace in Fig. S7b. 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 Hamilto-

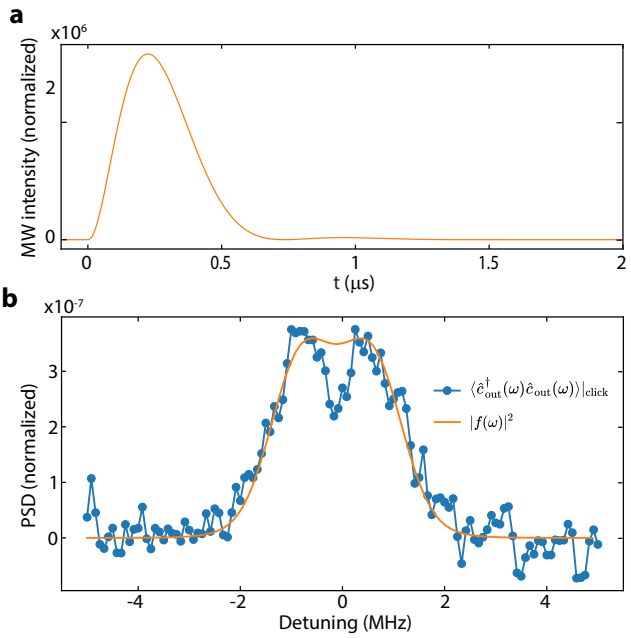

FIG. S7. **Microwave emission envelope function**. **a.** Simulated temporal profile of microwave emission intensity from the transducer heralded by an optical detection event. **b.** Comparison of normalized power spectral density (PSD) of the emission envelope function, f (orange trace) and of the 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.

nian 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. S7a. $g(t)$ 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

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

up to a normalization factor involving the gain of the microwave amplification chain. Here the Gaussian standard deviation, $T_{\rm g} = 230$ 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. S7b. 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.

10. 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 $\gg 2g_{\rm pe}$ 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, $P_{\rm det}$ at the output of the heterodyne setup. The gain, G of the amplification chain can be determined using the series of equations below.

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

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

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

Here, $n_{b,\rm sig}$ is the occupation of the acoustic mode due to the microwave drive obtained after subtracting the thermal component of the transduced signal. $P_{\rm in}$ is the input microwave power to the transducer. $\kappa_{b,\rm e}$ is the coupling rate of the acoustic mode to the microwave waveguide due to weak hybridization with the microwave resonator and $\kappa_{b,\rm i}$ 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.

11. 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 $g_{CC}^{(2)}$. To ensure linearity, we chose a region in the TWPA pump frequency and power space where $g_{CC}^{(2)} = 1$ 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 $g_{CC}^{(2)}$ for a weak coherent input. In case of a deviation of $g_{CC}^{(2)}$ from unity by one standard deviation, the experiment was halted and the pump optimization routine was re-run.

12. TRANSDUCER HEATING SPECTRUM

In Fig. S8a, 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 ($T_{\rm p} = 160$ ns, 50 kHz repetition

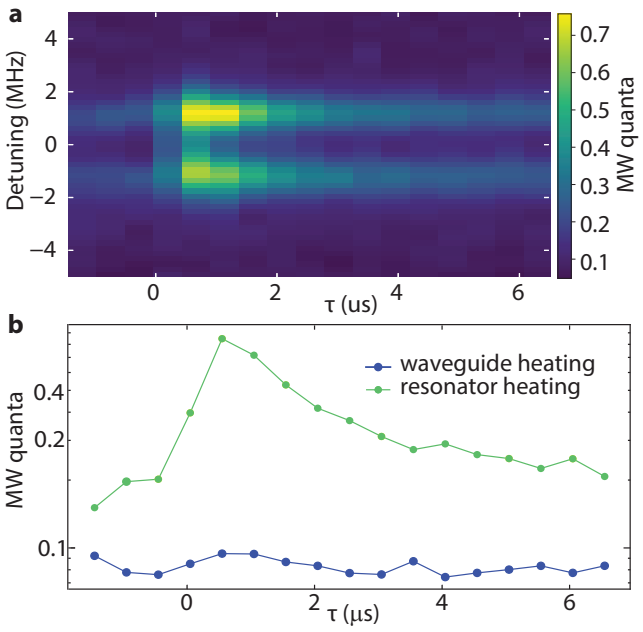

FIG. S8. Transducer heating spectrum. **a.** Power spectral density of the unconditional microwave signal as a function of delay, τ from the center of the optical pump pulse. Detuning is defined with respect to the mean frequency of the hybridized electromechanical modes. Legend indicates microwave quanta 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.**.

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. S8b, 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.

13. A SIMPLE MODEL FOR TRANSDUCER CORRELATION FUNCTIONS

Here we present a simple analytical model for the correlation functions based on a two-step approximation for the conditional microwave state in the transducer output port. In the first step, the optical pump pulse acts on the transducer acoustic mode, $\hat{b}$ containing a thermal state with initial occupation, n_i. Probabilistic detection of an optical photon signifies a successful SPDC scattering event and conditionally adds a phonon to the thermal state in the acoustic mode. This phonon-added thermal state is then extracted into the temporal mode in the microwave output port, $\hat{C}$, via the piezoelectric interaction. We model this conversion process as an imperfect beamsplitter by writing $\hat{C} = \sqrt{\eta_{\rm mw}}\hat{b} + \sqrt{1-\eta_{\rm mw}}\hat{d}$. Here $\eta_{\rm mw} = 0.56$ is the conversion efficiency for a phonon into the microwave output temporal mode, and $\hat{d}$ is a phenomenological thermal noise operator with mean noise quanta, $n_d = (1-\eta_{\rm mw})\langle\hat{d}^\dagger\hat{d}\rangle$, added unconditionally due to delayed heating from the hot bath induced by the pump.

For the unconditional microwave state, we have

$$\langle\hat{C}^\dagger\hat{C}\rangle = \eta_{\rm mw}n_i + n_d \tag{15}$$
$$\langle\hat{C}^{\dagger 2}\hat{C}^2\rangle = 2\langle\hat{C}^\dagger\hat{C}\rangle^2 \tag{16}$$

For the microwave state conditioned on optical detection, we use the moments for a photon-added thermal state [12] to write

$$\begin{aligned}\langle\hat{C}^\dagger\hat{C}\rangle|_{\rm click} &= \eta_{\rm mw}\langle\hat{b}^\dagger\hat{b}\rangle|_{\rm click} + (1-\eta_{\rm mw})\langle\hat{d}^\dagger\hat{d}\rangle \\ &= \eta_{\rm mw}(2n_i+1) + n_d \end{aligned}\tag{17}$$

$$\begin{aligned}\langle\hat{C}^{\dagger 2}\hat{C}^2\rangle|_{\rm click} &= \eta_{\rm mw}^2\langle\hat{b}^{\dagger 2}\hat{b}^2\rangle|_{\rm click} \\ &\quad +(1-\eta_{\rm mw})^2\langle\hat{d}^{\dagger 2}\hat{d}^2\rangle \\ &\quad +4\eta_{\rm mw}(1-\eta_{\rm mw})\langle\hat{d}^\dagger\hat{d}\rangle\langle\hat{b}^\dagger\hat{b}\rangle|_{\rm click} \\ &= \eta_{\rm mw}^2(6n_i^2+4n_i) + 2n_d^2 \\ &\quad +4\eta_{\rm mw}n_d(2n_i+1)\end{aligned}\tag{18}$$

Finally, we can include the effect of dark counts via the finite probability of a false heralding event due to dark counts, $p_{\rm dc}$. The final moments will be given by a weighted sum of the unconditional moments and the conditional moments with weights, p_{dc} and $(1-p_{\rm dc})$, respectively. In our experiment, we have $n_i = 0.097$ from optomechanical sideband asymmetry, $n_d = 0.083$ from the unconditional microwave measurement, and $p_{dc} = 0.12$ at the peak of the optical pulse. Calculating the moments with these parameters in the above expressions, we obtain $g^{(2)}_{AC}(\tau_o) = 4.9$ and $g^{(2)}_{CC}|_{\rm click} = 0.71$. This estimate of $g^{(2)}_{AC}$ is within 20% of the experimental value, and of $g^{(2)}_{CC}|_{\rm click}$ is roughly one standard deviation away from the mean of the experimentally measured confidence interval. While this simple model has limited predictive

power, it supports our claim of non-classical correlations after clearly accounting for the level of pump-induced noise expected in the generated states. For a more accurate picture, in the next section, we fully model the time dynamics of our system via a numerical simulation of the master equation.

14. 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,

$$\hat{H}_{abc}(t) = \hat{H}_o + \hat{H}_{\rm om}(t) + \hat{H}_{\rm pe}, \tag{19}$$

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

$$\hat{H}_o = -\hbar\Delta_a \hat{a}^\dagger\hat{a} + \hbar\omega_b \hat{b}^\dagger\hat{b} + \hbar\omega_c \hat{c}^\dagger\hat{c}, \tag{20a}$$

$$\hat{H}_{\rm om}(t) = \hbar G_{\rm om}(t)(\hat{a}^\dagger\hat{b}^\dagger + \hat{a}\hat{b}), \tag{20b}$$

$$\hat{H}_{\rm pe} = \hbar g_{\rm pe}(\hat{b}^\dagger\hat{c} + \hat{b}\hat{c}^\dagger). \tag{20c}$$

We have explicitly included the time dependence of the optomechanical coupling $G_{\rm om}(t) = \sqrt{n_a(t)}g_{\rm om}$ due to the temporal shape of the intra-cavity pump photon number, $n_a(t) = \kappa_{{\rm e},a}/(\Delta_a^2 + \kappa_a^2/4)P_{\rm in}(t)$. $P_{\rm in}(t)$, 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,

$$\dot{\hat{\rho}}(t) = \mathcal{L}(t)\hat{\rho}(t). \tag{21}$$

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

$$\mathcal{L}(t)\hat{\rho}(t) = -\frac{i}{\hbar}\left[\hat{H}_{abc}, \hat{\rho}(t)\right] + \left(\mathcal{L}_a + \mathcal{L}_b(t) + \mathcal{L}_c(t)\right)\hat{\rho}(t), \tag{22}$$

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.

$$\mathcal{L}_a\hat{\rho}(t) = \kappa_a\mathcal{D}(\hat{a})\hat{\rho}(t), \tag{23a}$$

$$\begin{aligned}\mathcal{L}_b(t)\hat{\rho}(t) =& \, n_{{\rm th},b}(t)\kappa_{{\rm i},b}\mathcal{D}(\hat{b}^\dagger)\hat{\rho}(t) \\ &+ \left[n_{\rm th,b}(t) + 1\right]\kappa_{i,b}\mathcal{D}(\hat{b})\hat{\rho}(t),\end{aligned} \tag{23b}$$

$$\begin{aligned}\mathcal{L}_c(t)\hat{\rho}(t) =& \, n_{{\rm th},c}(t)\kappa_{{\rm i},c}\mathcal{D}(\hat{c}^\dagger)\hat{\rho}(t) \\ &+ \left[n_{{\rm th},c}(t) + 1\right]\kappa_{i,c}\mathcal{D}(\hat{c})\hat{\rho}(t) \\ &+ n_{{\rm th},w}\kappa_{{\rm e},c}\mathcal{D}(\hat{c}^\dagger)\hat{\rho}(t) \\ &+ \left[n_{{\rm th},w} + 1\right]\kappa_{e,c}\mathcal{D}(\hat{c})\hat{\rho}(t).\end{aligned} \tag{23c}$$

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, $n_{{\rm th},b}(t)$ and $n_{{\rm th},c}(t)$ respectively. $n_{{\rm th},w}$, the thermal occupation of the microwave waveguide is set to a constant value much smaller than $n_{{\rm th},b}(t)$ and $n_{{\rm th},c}(t)$ 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. (21) 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 $\kappa_a \gg G_{\rm om}$ to adiabatically eliminate the optical mode [16]. Moving into a frame rotating with the mechanical and microwave modes and defining the optomechanical scattering rate, $\Gamma_{\rm om}(t) = 4|G_{\rm om}(t)|^2/\kappa_a$, we arrive at the simplified master equation,

$$\dot{\hat{\rho}}_r(t) = \mathcal{L}_r(t)\hat{\rho}_r(t), \tag{24}$$

where $\hat{\rho}_r = {\rm Tr}_a\{\hat{\rho}\}$ 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,

$$\begin{aligned}\mathcal{L}_r(t)\hat{\rho}_r(t) =& -\frac{i}{\hbar}\left[\hat{H}_{\rm pe}, \hat{\rho}_r(t)\right] \\ &+ \left(\mathcal{L}_{\rm om}(t) + \mathcal{L}_b(t) + \mathcal{L}_c(t)\right)\hat{\rho}_r(t),\end{aligned} \tag{25}$$

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

$$\mathcal{L}_{\rm om}(t)\hat{\rho}_r(t) = \Gamma_{\rm om}(t)\mathcal{D}(\hat{b}^\dagger)\hat{\rho}_r(t). \tag{26}$$

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 ($\kappa_a \gg \kappa_b, g_{\rm pe}, \kappa_c$), 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,

$$\hat{\rho}_J(t_J) = \frac{\hat{b}^\dagger \hat{\rho}_r(t_J)\hat{b}}{\mathrm{Tr}\left\{\hat{b}^\dagger \hat{\rho}_r(t_J)\hat{b}\right\}}. \tag{27}$$

This state then evolves according to Eq. (24). Since we operate in the weak pump regime where the integrated jump probability over the optical pulse duration, $\int \Gamma_{om}(t)dt \ll 1$, 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 $\hat{C}(t) := \int \hat{c}_{out}(t+t')f^*(t')dt' = \sqrt{\kappa_{e,c}}\int \hat{c}(t+t')f^*(t')dt'$. We can thus compute the occupation of the temporal mode as

$$\begin{aligned}
&\langle \hat{C}^\dagger(t)\hat{C}(t)\rangle \\
&= \kappa_{e,c}\langle \int \hat{c}^\dagger(t+t'')f(t'')dt'' \int \hat{c}(t+t')f^*(t')dt'\rangle \\
&= \kappa_{e,c}\int \langle \hat{c}^\dagger(t+t'')\hat{c}(t+t')\rangle f(t'')f^*(t')dt''dt' \\
&= \kappa_{e,c}\int \langle \hat{c}^\dagger(t'+\tau')\hat{c}(t')\rangle f(t'+\tau'-t)f^*(t'-t)dt'd\tau',
\end{aligned} \tag{28}$$

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 $\hat{\rho}_J(t_J)$ in Eq. (27) for times, $t > t_J$. For times $t < t_J$, we use the value for the unconditional state. Performing this calculation for a sequence of jump times, $\{t_J\}$ 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, $\{\langle \hat{C}^\dagger\hat{C}\rangle|_{t_J}\}$. 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, $\delta p_J \propto \Gamma_{om}(t_J)dt$ 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, $\omega_c - \omega_b \gg 2g_{pe}$, (ii) mechanics on resonance with the microwave mode, $\omega_b = \omega_c$. The results of the measurement

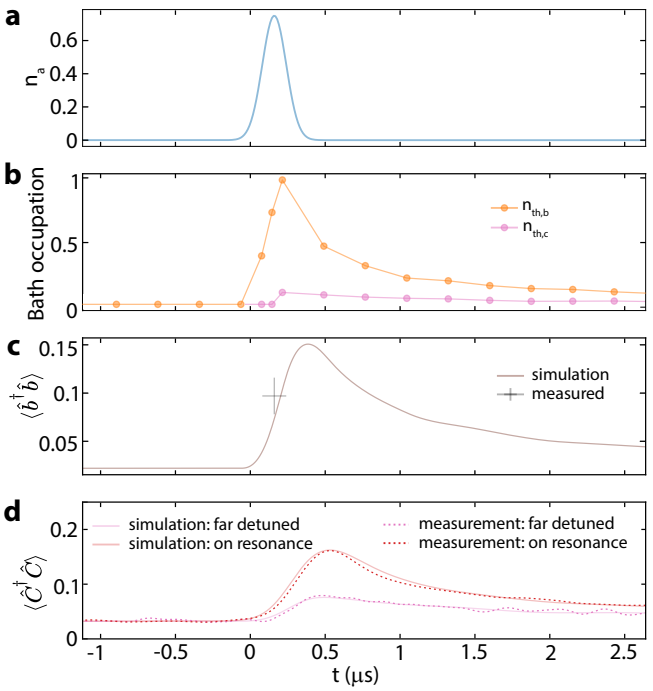

FIG. S9. Simulation of transducer heating dynamics **a.** Time trace of the optical pump pulse plotted in terms of the intracavity occupation of the optical mode, n_a. **b.** Model used for time dependent thermal occupation of the intrinsic baths of the acoustic and microwave modes, $n_{th,b}(t)$, $n_{th,c}(t)$. **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 two sigma duration of the optical pulse. **d.** 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.

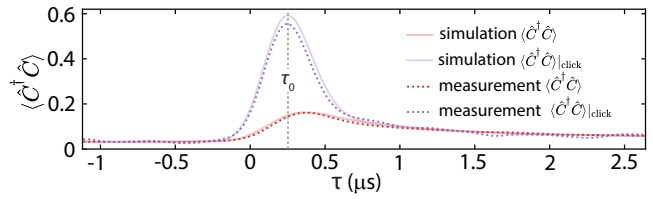

FIG. S10. Simulation of the conditional microwave state. Occupation of the temporal mode, $\hat{C}$ in the microwave output port of the transducer from simulation (solid traces) and from experimental heterodyne data (dotted traces). The occupation of the microwave state conditioned on an optical click, $\langle \hat{C}^\dagger \hat{C}\rangle|_{\mathrm{click}}$ 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.

are shown with the dotted traces in Fig. S9d. 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. S9b. 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 emission. 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. S9c and with the solid traces in Fig. S9d 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. S9c. This data point corresponds to the optomechanical sideband asymmetry measurement described in Section 7, 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. S10 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 $g^{(2)}_{CC}|_{\mathrm{click}}$ requires evaluation of $\langle \hat{C}^{\dagger 2}\hat{C}^2\rangle|_{\mathrm{click}}$. Extending the approach of Eq. (28) 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 $g^{(2)}_{CC}|_{\mathrm{click}}$. 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, $g^{(2)}_{bb}|_{\mathrm{click}} = \langle \hat{b}^{\dagger 2} b^2\rangle|_{\mathrm{click}}/(\langle \hat{b}^\dagger b\rangle|_{\mathrm{click}})^2$ to provide a lower bound for $g^{(2)}_{CC}|_{\mathrm{click}}$. Our model estimates $g^{(2)}_{bb}|_{\mathrm{click}}(0) = 0.24$. 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. 9. 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, $g^{(2)}_{cc}|_{\mathrm{click}} = \langle \hat{c}^{\dagger 2} c^2\rangle|_{\mathrm{click}}/(\langle \hat{c}^\dagger c\rangle|_{\mathrm{click}})^2$, serves as an upper bound for $g^{(2)}_{CC}|_{\mathrm{click}}$. Our model estimates $g^{(2)}_{cc}|_{\mathrm{click}}(\tau_o) = 0.77$. In summary, the above arguments allow us to use numerical simulations to bound $g^{(2)}_{CC}|_{\mathrm{click}}(\tau_o)$ to the interval (0.24, 0.77), which has substantial overlap with the experimental observation of $g^{(2)}_{CC}|_{\mathrm{click}}(\tau_o) = 0.42^{+0.27}_{-0.28}$.

15. DATA ANALYSIS FOR CORRELATION FUNCTIONS

For measurements of $g^{(2)}_{AC}$ and $g^{(2)}_{CC}|_{\mathrm{click}}$ presented in Fig. 4 of the main text, 9.1×10^4 heterodyne voltage traces were acquired over a one month period. For every $\sim$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, $g^{(2)}_{CC}(\tau_{\mathrm{o}})$ used 3.2×10^7 heterodyne voltage samples.

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 $\bar{C}^{(k)}_{mn}, \bar{C}^{(k)}_{mn}|_{\mathrm{click}}$, where the index k runs over the daily chunks. This inversion process follows the techniques described in Sec. 8 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 $\bar{C}_{mn}|_{\mathrm{click}} = \sum_k w^{(k)}_{\mathrm{click}}\bar{C}^{(k)}_{mn}|_{\mathrm{click}}$, where $w^{(k)}$, $w^{(k)}_{\mathrm{click}}$ 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. S11b.

The acquisition rate for the moments of the conditional heterodyne output, $\bar{S}_{\mathrm{mn}}|_{\mathrm{click}}$ is determined by the optical heralding rate, $R_{\mathrm{click}} = 0.14$ Hz. This is much slower compared to that for the amplifier noise moments, $\bar{H}_{\mathrm{mn}}$, which can be acquired at the 50 kHz repetition rate of the experiment and determined with high accuracy. As a result, the error in $g^{(2)}_{CC}|_{\mathrm{click}}$ is dominated by uncertainty in $\bar{S}_{\mathrm{mn}}|_{\mathrm{click}}$. To calculate error in $g^{(2)}_{CC}$ and $g^{(2)}_{CC}|_{\mathrm{click}}$, we employ bootstrapping with replacement using 10^5 bootstraps to construct the probability density

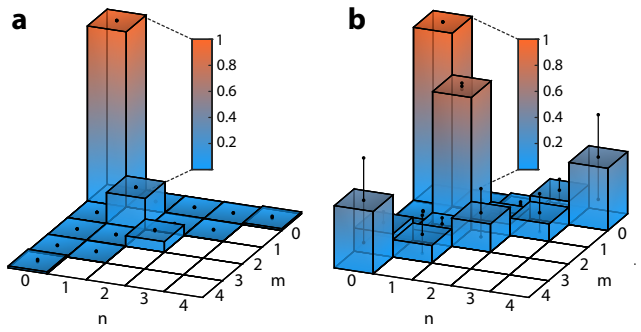

FIG. S11. **Moments matrices**. **a.** Moments matrix for the unconditional microwave output of the transducer, $\bar{C}_{mn}$, obtained from 3.2×10^7 voltage samples. **b.** Moments matrix for the microwave output of the transducer conditioned on receipt of an optical click, $\bar{C}_{mn}|_{\text{click}}$, obtained from 9.1×10^4 voltage samples. In both panels, height of colored bars indicates mean value. Error bars are indicated by black vertical pins on data points at the top of each bar and correspond to +/- one standard deviation.

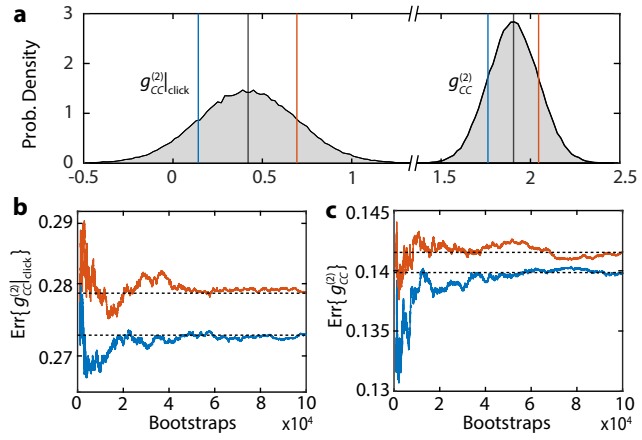

FIG. S12. **Error estimates for correlation functions**. **a.** Probability distribution for $g^{(2)}_{CC}$ and $g^{(2)}_{CC}|_{\text{click}}$ calculated from bootstrapping. Vertical blue and red lines correspond to one standard deviation above and below the distribution mean respectively. **b,c.** Convergence of the error estimate vs. number of bootstraps for optically heralded traces and all 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.

functions shown in Fig. S12a. 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. S12b-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.

16. 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, $g^{(2)}_{CC}|_{\text{click}} \geq 1$ 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,

$$g^{(2)}_{CC}(\tau)|_{\text{click}} = \frac{\langle \hat{C}^{\dagger 2}(\tau)\hat{C}^2(\tau)\rangle\,|_{\text{click}}}{\left(\langle \hat{C}^{\dagger}(\tau)\hat{C}(\tau)\rangle\,|_{\text{click}}\right)^2} = \frac{\langle \hat{A}^{\dagger}\hat{C}^{\dagger 2}(\tau)\hat{C}^2(\tau)\hat{A}\rangle\,\langle \hat{A}^{\dagger}\hat{A}\rangle}{\left(\langle \hat{A}^{\dagger}\hat{A}\rangle\,\langle \hat{C}^{\dagger}(\tau)\hat{C}(\tau)\rangle\right)^2}, \tag{29}$$

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

$$\langle \hat{A}^{\dagger m}\hat{C}^{\dagger n}(\tau)\hat{A}^m\hat{C}^n(\tau)\rangle = \int |\alpha|^{2m}|\gamma|^{2n}P_\tau(\alpha,\gamma)d^2\alpha d^2\gamma := \langle |\alpha|^{2m}|\gamma|^{2n}\rangle_{P_\tau}, \tag{30}$$

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. (29) gives,

$$g^{(2)}_{CC}(\tau)|_{\text{click}} = \frac{\langle |\alpha|^2_P|\gamma|^4\rangle_{P_\tau}\,\langle |\alpha|^2\rangle_{P_\tau}}{\langle |\alpha|^2|\gamma|^2\rangle^2_{P_\tau}} \geq 1, \tag{31}$$

which follows directly from the Cauchy-Schwarz inequality.

17. 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. S13a. All measurements were performed with the same pump pulse duration and repetition rate used to acquire data in the main text. With

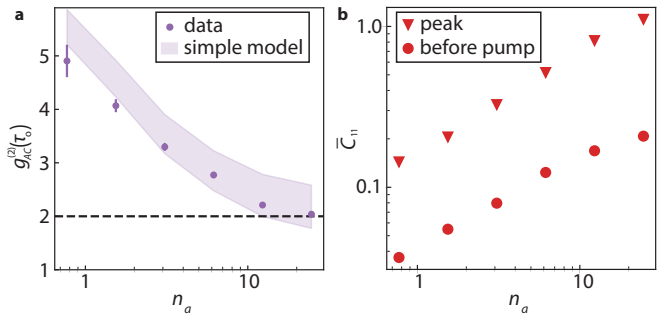

FIG. S13. **a.** Maximum value of $g^{(2)}_{AC}$ versus pump power plotted on the horizontal axis as the peak intra-cavity occupation of the optical mode, n_a. Data points indicate mean and error bars indicate +/- one standard deviation. Microwave voltage samples are acquired at $n_a = (0.77, 1.5, 3.1, 6.2, 12, 25)$ for $(1.7, 3.2, 6.9, 4.0, 4.9, 16) \times 10^4$ heralding events, respectively. Shaded region indicates predictions from the simple model presented in Section 13. Dashed line denotes the classical upper bound of 2. **b.** Unconditional microwave output quanta, $\bar{C}_{11}$, versus pump power. Circles show the value before the pump pulse, triangles show the peak value after the pump pulse.

increasing pump power, the maximum value of $g^{(2)}_{AC}$ approaches the classical upper bound of 2. The experimentally observed trend agrees well with estimates from the simple model presented in Section 13. For the noise parameters, n_i, n_d, in the model, we determine bounds using experimentally measured unconditional microwave output quanta, $\bar{C}_{11}$ shown in Fig. S13b, and calculate the shaded region in Fig. S13a. 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).

[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 non-

classical photon pairs for scalable quantum communication with atomic ensembles, Nature **423**, 731 (2003).