Supplemental Material:
Measuring High-Order Phonon Correlations in an Optomechanical Resonator

Y. S. S. Patil,[1,*] J. Yu,[2] S. Frazier,[1] Y. Wang,[2] K. Johnson,[1] J. Fox,[1] J. Reichel,[3] and J. G. E. Harris[1,2,4,†]

[1] *Department of Physics, Yale University, New Haven, Connecticut 06520, USA*
[2] *Department of Applied Physics, Yale University, New Haven, Connecticut 06520, USA*
[3] *Laboratoire Kastler Brossel, ENS-Université PSL, CNRS, Sorbonne Université, Collège de France 24 rue Lhomond, 75005 Paris, France*
[4] *Yale Quantum Institute, Yale University, New Haven, Connecticut 06520, USA*

A. Optical Setup

This section describes the optical setup used for the experiments presented in the main text.

Before describing the setup itself, we first describe the timing sequence used in these measurements. During an experiment the setup is alternated between two configurations, each lasting 100 ms. The time sequence is shown in Fig. S1 (the configurations are shown in Fig. S2 and discussed in the following paragraph). The first of these is the "lock" configuration, during which the optical path is configured to lock the tunable laser (TL) and filter cavities to the optomechanical cavity (OMC). The second is the "drive" configuration, during which the various control voltages used to tune the TL and filter cavities are held constant. Also, at the start of the drive period, the optical path is configured so that either a blue- or red-detuned laser drives the OMC. Photon counting data is collected during the middle 90 ms of the drive period.

Figure S2 shows the two configurations, with the lock configuration's optical path in green and the drive configuration's optical path in orange. The lock configuration begins with the TL entering an IQ modulator (IQM) operating in the single-sideband suppressed-carrier mode. The IQM serves as a frequency shifter to lock the TL to the OMC. The output light from the IQM is amplified by an Erbium-doped fiber amplifer (EDFA) and filtered by a broadband ($\Delta\lambda \approx 0.3$ nm) filter to suppress the EDFA's amplified spontaneous emission noise. The beam then passes through an electro-optic modulator (EOM) which produces a pair of sidebands used for standard Pound-Drever-Hall (PDH) locking. The beam is then split into three parts, which are sent to the OMC and the two filter cavities (FC1, FC2). The TL is thus locked to the OMC via the IQM control voltage, and the filter cavities are locked to the TL via their piezo and thermal tuning.

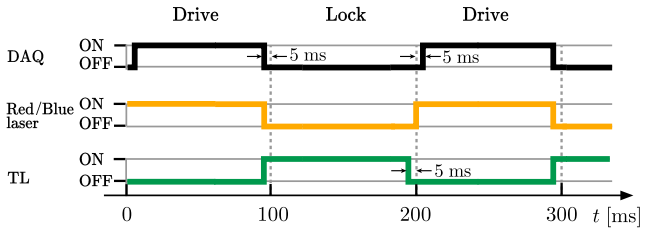

FIG. S1. **Lock and drive sequence.** The time sequence for switching between the "lock" and the "drive" configurations. Photon counts are acquired only during the central 90 ms of the "drive" period.

The blue-/red-detuned drive lasers (BL/RL) are always locked to the TL. This is accomplished by combining each drive laser with the TL (shown in the upper left corner of Fig. S2) and monitoring the frequency of the resulting beat note to ensure that the two drive lasers are detuned from the OMC by $\sim \pm\omega_{\mathrm{ac}}$. Each drive laser passes through a filter cavity ($\kappa/2\pi \approx 30$ MHz) and is then combined via a 50/50 fiber coupler to drive the OMC. The part of the optical path that is common to the two drive beams is shown in orange. A pair of shutters just before the 50/50 coupler are used to determine whether the BL or the RL drives the OMC. The shutters are also used to set the desired pulse sequence for pulsed measurements.

Scattered sideband photons that are resonant with the OMC exit the cavity and pass through filter cavities FC1 and FC2, finally arriving at a pair of superconducting nanowire single photon detectors (SNSPDs). We note that the coherence times of the Helium oscillator in this work ($10-100\,\mu$s) are much longer than the dead/recovery time of our detectors (<50 ns). Multiple scattering events that occur within these coherence times are thus detectable by each of the SNSPDs individually. As such, we combine the photon arrival time records generated by the two detectors into a single record for the analysis.

The variable optical attenuators (VOAs) are used to adjust the optical power, and the polarization controllers

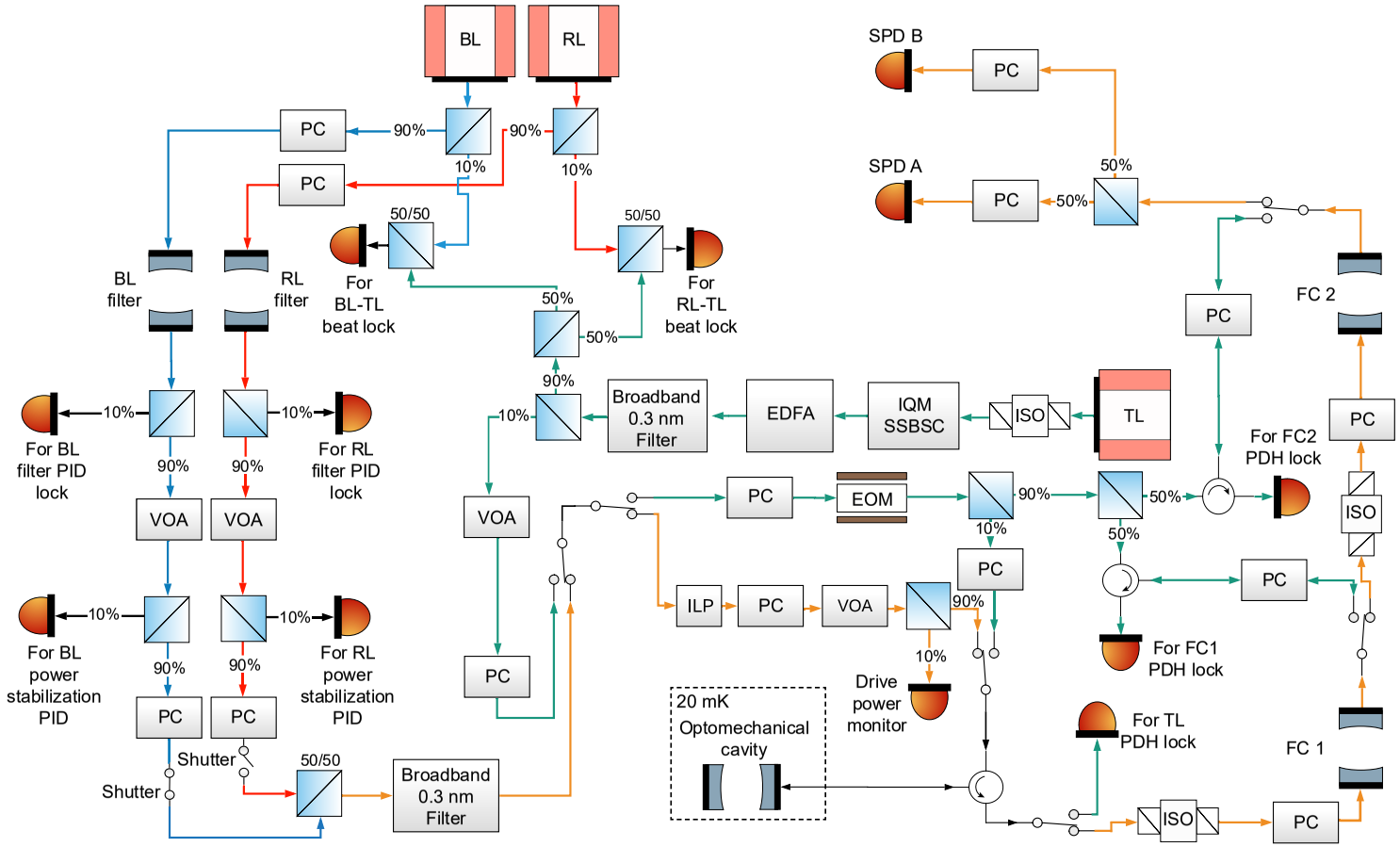

FIG. S2. **Optical schematic.** This figure shows the two optical configurations (lock and drive) used in the experiment. The tunable laser (TL) is locked to the optomechanical cavity during the lock configuration. The blue (BL) and red (RL) drive lasers are always locked to the TL with a frequency offset $\sim \pm\omega_{\rm ac}$. A total of six 1×2 MEMS switches are used to switch between the lock configuration (green optical path) and the drive configuration (orange optical path) every 100 ms, such that the optical components are either locked to their desired frequency, or free running (as indicated in Fig. S1). PC: polarization controller. VOA: variable optical attenuator. EDFA: erbium-doped fiber amplifier. EOM: electro-optic modulator. ILP: in-line polarizer. ISO: isolator. FC: filter cavity.

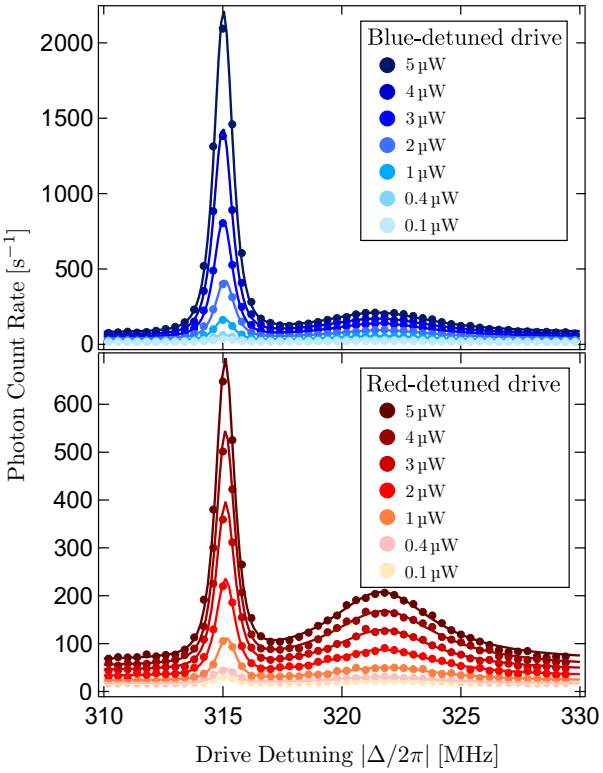

FIG. S3. **Photon count spectrum.** Count rates as a function of the detuning and power of the drive laser. The narrow peak near 315 MHz corresponds to sideband photons produced by the creation and annihilation of phonons in the superfluid-filled cavity. The broad beak near 322 MHz is due to thermal fluctuations in the room-temperature optical fibers. The fits are described in the text.

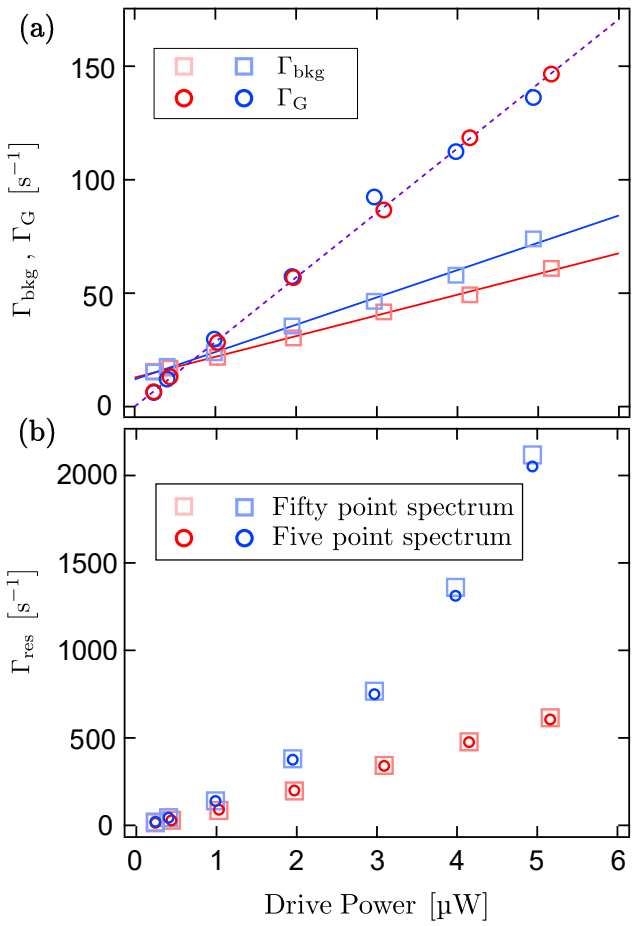

FIG. S4. **Power dependence of the photon count spectrum.** **(a)** Γ_{bkg} and Γ_G as a function of drive power, and the corresponding linear fits. For Γ_{bkg}, the y-intercepts give background counts arising from SNSPD dark counts and stray light from the environment. The slopes give the leakage of drive photons around the filter cavities for the red-detuned and blue-detuned drive each. Γ_G also increases linearly with drive power. **(b)** Γ_{res} as a function of drive power. Shown as open squares are values of Γ_{res} extracted from a 50-frequency measurement that is fit to Eq. 3. Γ_{res} extracted from a five-frequency measurement that is fit to Eq. 4 are shown as open circles. The difference between the two methods is negligible.

(PCs) are used to align the polarization of light entering polarization-sensitive components such as the FCs, the EOM and the SNSPDs.

B. Characterizating the Spectrum of Photon Counts

Figure 1(d) of the main text shows a typical photon count rate spectrum. Here we discuss the characterization of such spectra. Fig. S3 shows the count rate spectra for various values of the drive power P_{in}. Each spectrum shows a narrow peak (arising from the Stokes- and anti Stokes-scattered photons) centered at $\omega_{ac}/2\pi \approx 315.3$ MHz, a broader peak centered around 322.3 MHz, and a frequency-independent background.

The sideband peak is expected to be proportional to the transfer function of the two cascaded filter cavities FC1 and FC2, given by

$$f_{\text{filter}}(\Delta) = \frac{1}{\left[1+4\left(\frac{|\Delta|-\omega_{ac}}{\kappa_{FC1}}\right)^2\right]\left[1+4\left(\frac{|\Delta|-\omega_{ac}}{\kappa_{FC1}}\right)^2\right]} \quad (1)$$

where $\kappa_{FC1}/2\pi = 1.21(5)$ MHz and $\kappa_{FC2}/2\pi = 1.71(2)$ MHz are the filter linewidths. We identify the broad peak around 322 MHz as the signature of guided-acoustic wave Brillouin scattering (GAWBS) in the fibers (See

Sec. C for more details), which can be characterized by a Lorentzian centered at ω_G:

$$f_G(\Delta) = \frac{1}{1 + 4\left(\frac{|\Delta| - \omega_G}{\kappa_G}\right)^2} \tag{2}$$

Thus, for each value of P_{in} and for each detuning range (i.e., $\Delta \sim +\omega_{ac}$ and $\Delta \sim -\omega_{ac}$), the count rate spectrum is fit to the form

$$\Gamma(\Delta) = \Gamma_{bkg} + f_{filter}(\Delta)\Gamma_{res} + f_G(\Delta)\Gamma_G \tag{3}$$

where Γ_{bkg}, Γ_{res}, Γ_G, ω_{ac}, ω_G, and κ_G are the fit parameters.

Fig. S4 shows the fit-extracted Γ_{bkg}, Γ_G, and Γ_{res} as a function of P_{in}. The frequency-independent background Γ_{bkg} is found to have three sources: 1) the SNSPDs' darks counts, 2) stray light leaking into the fibers, and 3) unfiltered drive photons leaking around the two filter cavities in the detection chain (this last contribution is polarization-dependent, and so differs between the red-detuned and blue-detuned drives). $\Gamma_{bkg}(P_{in})$ is fit to the form $\Gamma_{bkg,0} + \Gamma_{bkg,1}P_{in}$, which gives $\Gamma_{bkg,0} = 12.4 \pm 0.8$ s^{-1}, and $\Gamma_{bkg,1} = 9.1 \pm 0.2$ $s^{-1}\mu W^{-1}$ for the red-detuned drive, and $\Gamma_{bkg,1} = 12.0 \pm 0.4$ $s^{-1}\mu W^{-1}$ for the blue-detuned drive. By acquiring the count rate while blocking the input port directly in front of the SNSPDs, we find that $\Gamma_{bkg,0}$ consists of 7 ± 1 s^{-1} from dark counts and 5 ± 1 s^{-1} from stray light leaking into the detection-chain fibers. $\Gamma_{bkg,1}$ quantifies the amount of drive photons that leak around the filter cavities.

$\Gamma_G(P_{in})$ is fit to the form $\Gamma_{G,1}P_{in}$. This gives $\Gamma_{G,1} = 28.5 \pm 0.3$ $s^{-1}\mu W^{-1}$.

A complete characterization of Γ_{res} as a function of drive power is given in Sec. D.

The spectra in Fig. S3 (and those used to produce Fig. S4) are measured at 50 values of $\Delta/2\pi$. While this is helpful in characterizing the device, the focus of the main paper is on the Stokes- and anti-Stokes- scattered photons. To measure the rates of these specific photons, we found that it was adequate to record the spectrum at just five frequencies close to ω_{ac} (specifically, $\Delta/2\pi =$ 310, 312, 314.9, 315.4, 315.9 MHz), and then fit the results to the form

$$\Gamma(\omega) = \Gamma_{bkg} + f_{filter}(\omega)\Gamma_{res} \tag{4}$$

To illustrate the validity of this approach, Fig. S4(b) shows $\Gamma_{res}(P_{in})$ acquired in two ways: by measuring the count rate at 50 values of $\Delta/2\pi$ ranging from $\pm$ 310 MHz to $\pm$ 330 MHz (as in Fig. S3) and fitting to Eq. 3; and by measuring the count rate at just the five values of $\Delta/2\pi$ listed above and fitting to Eq. 4. The difference is negligible, showing the reliability of the five-point spectrum and the minimal influence of the GAWBS peak (which is not included in Eq. 4) on this analysis. The data presented in Sec. D were acquired using the "five frequency" approach.

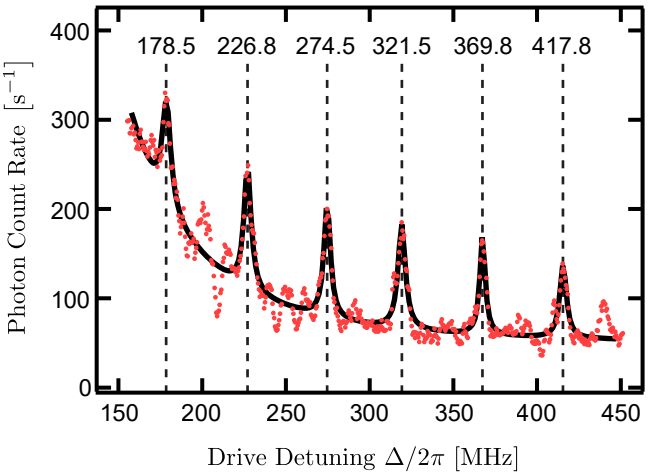

FIG. S5. **Wide frequency scan showing several GAWBS peaks.** Six individual Lorentzians are fit (in addition to the tail of f_{filter}), with the vertical dashed lines marking the measured (best-fit) center frequencies for the fiber's transverse acoustic modes.

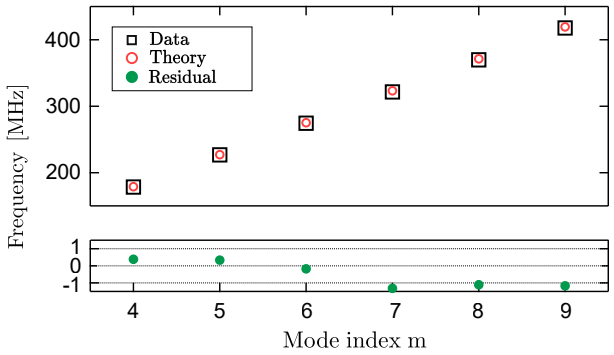

FIG. S6. **The resonant frequencies of GAWBS.** The fitted peak frequencies are shown in solid black squares and the calculated frequencies of R_{0m} GAWBS resonant modes for a 125 µm single-mode fused silica fiber are shown in open red circles. The discrepancies are shown on the bottom graph in solid green circles.

C. Measuring the Wider Photon Count Spectrum and Characterizing GAWBS

This section gives a more detailed characterization of the GAWBS signature mentioned in Sec. B. GAWBS is the well-studied interaction between light in a fiber and the fiber's thermally populated transverse acoustic modes

[1]. In practice, this interaction produces phase noise in the light with a spectrum determined by the fiber's acoustic modes.

Figure S5 shows the photon count rate as the detuning $\Delta/2\pi$ is varied from 150 to 450 MHz. The overall descending background in Fig. S5 corresponds to the tail of the filter cavity's transfer function. Due to the large steps in the detuning, the sideband peak of the acoustic cavity mode cannot be seen here, but several other peaks are evident. These peaks are found to be centered close to the frequencies f_m that are expected for the fiber's transverse acoustic modes, R_{0m}, which vibrate in the radial direction. These frequencies (labeled as vertical lines in Fig. S5) correspond to the solutions y_m of

$$(1-\alpha^2)J_0(y) = \alpha^2 J_2(y) \tag{5}$$

where J_x is the x-order Bessel function of the first kind, and $\alpha = 0.624$ is the ratio between transverse and longitudinal phonon velocities for fibers made of fused silica. The resonance frequencies are given by $f_m = \frac{V_d}{2\pi a} y_m$, where $V_\mathrm{d} = 5996$ m/s is the longitudinal phonon velocity and the fiber radius $a = 62.5$ µm [1, 2].

The data in Fig. S5 is fit to the sum of six Lorentzians (whose center frequencies, linewidths, and magnitudes are the fit parameters), and the tail of $f_\mathrm{filter}(\omega)$. The best-fit values for the center frequencies and the theoretical f_m calculated from Eq. 5 are shown in Fig. S6.

D. Model and Fits for Drive Power Scans

This section describes measurements of the sideband-photon counting rates as a function of the mixing chamber temperature T_MC and the laser power incident on the device P_in. These measurements provide information about the acoustic mode's temperature and the useful range of P_in. They also provide additional calibration of the device's optomechanical coupling rate g_0.

Figure S7(a) shows R_AS and R_S measured as a function of T_MC at an incident drive power $P_\mathrm{in} = 250$ nW. The values of R_AS and R_S are extracted from fitting five-point spectra as described in Sec. B. Both R_AS and R_S increase with T_MC as expected, and the persistent difference between the two is consistent with the quantum sideband asymmetry (QSA). The solid lines in Fig. S7(a) show a one-parameter fit to the form $a(e^{\hbar\omega_\mathrm{ac}/k_B T_\mathrm{MC}} - 1)^{-1}$ (for the red data) and $1 + a(e^{\hbar\omega_\mathrm{ac}/k_B T_\mathrm{MC}} - 1)^{-1}$ (for the blue data), where the fit parameter $a = \gamma_\mathrm{ac} C \eta_\kappa \eta_\mathrm{det}/P_\mathrm{in}$ is the magnitude of the power-normalized QSA. The fit only uses data with $T_\mathrm{MC} > 50$ mK, as the calibration of the

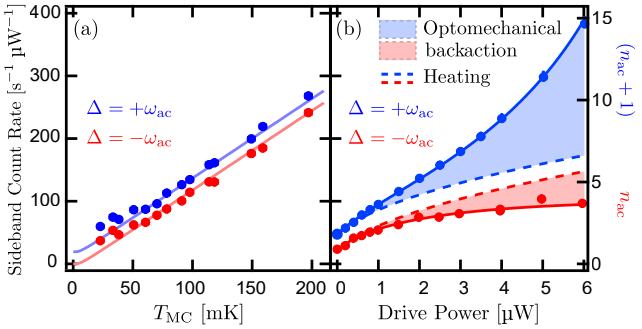

FIG. S7. **Temperature- and power-dependence of the sidebands.** Normalized anti-Stokes (red) and Stokes (blue) photon detection rates. (a) Measurements as a function of T_MC with $P_\mathrm{in} = 250$ nW. Solid lines: fit to the form predicted for thermal equilibrium. (b) Measurements as a function of P_in with $T_\mathrm{MC} \approx 20$mK. Solid lines: a fit including optomechanical backaction (colored regions) and heating due to optical absorption (dashed lines). Right-hand axis: the mean phonon occupancy n_ac $(n_\mathrm{ac}+1)$ inferred from the detection rates.

RuO_2 thermometer used to determine T_MC is uncertain at lower temperatures.

The agreement between data and theory in Fig. S7(a) indicates that the mean phonon occupancy n_ac is primarily determined by T_MC. However, at higher drive powers, both optical backaction and absorption-induced heating (in the fibers and mirrors) can alter n_ac. To characterize the role of backaction and heating in these devices, Fig. S7b shows $R_\mathrm{AS(S)}$ as a function of P_in. (The data at $P_\mathrm{in} = 0$ are measured using pulsed laser excitations and modelling (and then subtracting) the transient heating.) For sufficiently low P_in ($\lesssim 300$ nW), the values of $R_\mathrm{AS(S)}$ are consistent with $T_\mathrm{MC} \approx 20$ mK, and show a splitting that is dominated by the QSA. As P_in is increased, both R_AS and R_S increase. However, R_S increases more rapidly with P_in than R_AS. This is consistent with the combined effects of heating from optical absorption (which increases n_ac independent of Δ) and optomechanical backaction (which increases (decreases) n_ac when Δ is positive (negative)).

To analyze these effects quantitatively, we apply the standard theory of optomechanical backaction, and model the heating by assuming that the device is subject to a heatload proportional to P_in and is thermally linked to the MC with a thermal conductance $\sigma = bT^{k+1}$. Before describing this approach in detail (see the following paragraphs), we first summarize the result. The solid lines in Fig. S7(b) show the best fit while taking $g_0, T_\mathrm{MC}, \eta_\mathrm{det}, b$ and k as fit parameters. The dashed lines show the change in n_ac attributable to absorption-induced heat-

ing. The best-fit value $k = 1.05(3)$ suggests a predominantly metallic thermal conductance between the device and MC ($k = 1$ for metals). The fit also gives a single photon coupling rate $g_0/2\pi = 4.58(2)$ kHz and a net detection efficiency of $\eta_{\rm det} = 0.18(1)$. The best-fit value $T_{\rm MC} = 19(1)$ mK corresponds to a mean thermal phonon occupancy of $n_{\rm ac} = 0.83(2)$.

In more detail, we start with the power-normalized sideband count rates as described by standard quantum optomechanics theory:

$$\overline{R}_{\rm AS} = a n_{\rm ac} \tag{6}$$

$$\overline{R}_{\rm S} = a(n_{\rm ac} + 1) \tag{7}$$

where, as defined above, $a = \gamma_{\rm ac} C \eta_\kappa \eta_{\rm det}/P_{\rm in}$, and $n_{\rm ac}$ is determined by the joint effects from the thermal bath ($n_{\rm th}$) and optomechanical backaction [3],

$$n_{\rm ac} = \frac{\gamma_{\rm opt}\left(\frac{\kappa_{\rm c}}{4\omega_{\rm ac}}\right)^2 + \gamma_{\rm ac} n_{\rm th}}{\gamma_{\rm opt} + \gamma_{\rm ac}} \tag{8}$$

where $\gamma_{\rm opt}\left(= \gamma_{\rm ac} C = 4 \cdot g_0^2/\kappa_{\rm c} \cdot n_{\rm c}\right)$ and $\gamma_{\rm ac}$ are the optomechanical damping rate and bare mechanical damping rate, respectively, and $n_{\rm th}$ accounts for the effect of optical-absorption-induced heating in the fibers and mirrors.

To model the effect of optical heating, we assume a heat load proportional to $P_{\rm in}$, and that the conductivity of the thermal link between the fiber and the MC is a power law in temperature with exponent k, such that the fiber's temperature is given by

$$T_{\rm fib} = (T_{\rm MC}^{k+1} + \beta^{k+1} P_{\rm in})^{1/(k+1)} \tag{9}$$

where β characterizes the amount of heat generated by a given $P_{\rm in}$. In addition, we assume a ballistic transport of heat (phonons) from the Helium inside the OMC to the MC bath via the Helium channel connecting them. It follows that

$$n_{\rm th} = \frac{n_{\rm fib}\gamma_{\rm ac0} + n_{\rm MC}\gamma_{\rm ball}}{\gamma_{\rm ac0} + \gamma_{\rm ball}} \tag{10}$$

where $n_{\rm fib} = 1/(e^{\hbar\omega_{\rm ac}/(k_{\rm B}T_{\rm fib})} - 1)$ and $n_{\rm MC} = 1/(e^{\hbar\omega_{\rm ac}/(k_{\rm B}T_{\rm MC})} - 1)$. The bare mechanical damping rate $\gamma_{\rm ac} = \gamma_{\rm ac0} + \gamma_{\rm ball}$ includes: i) the intrinsic mechanical damping rate $\gamma_{\rm ac0}/2\pi = 3.2\times10^3$ Hz at zero temperature due to acoustic loss into the fibers, and ii) the three phonon loss $\gamma_{\rm ball}/2\pi = 2.7\times10^6\ T_{\rm MC}^4$ Hz/K^4 at the base MC temperature [4].

We have verified the accuracy of this model by using it to analyze drive power sweep data taken on several different days over the course of ten months. Fig. S8 shows seven different power sweeps. The solid lines are the best fit of the data to the model described by Eq. 7 - 10. The fit uses a single $T_{\rm MC}$, β, k, and optomechanical coupling g_0 for all of the sweeps but allows $\eta_{\rm det}$ to vary between the sweeps, as the overall detection efficiency is observed to differ from day to day.

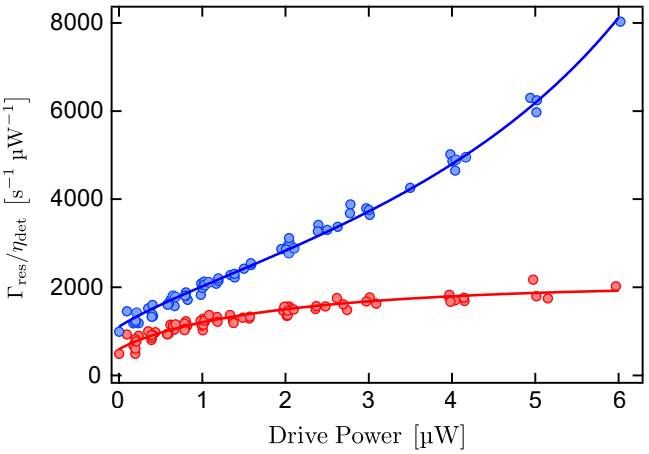

FIG. S8. **Several scans of the drive power.** Scattering rates as a function of drive power, taken over several days. The points are data, and the solid lines are a fit in which all the fit parameters are global, but the detection efficiency is allowed to vary between the data sets. For this plot, each data set is normalized by its fitted detection efficiency.

$T_{\rm MC}$	24.4 ± 0.4 mK
β	0.54 ± 0.03 K/W$^{1/(k+1)}$
k	1.09 ± 0.03

TABLE I. **Fit parameters for thermal transport from fiber to bath** The best-fit in Fig. S8 returns the fit parameters that describe the fiber temperature according to Eq. 9. $T_{\rm MC}$ characterizes the initial device temperature, k the exponent of the power law that describes the thermal conductivity, and β the fraction of heat generated by a given P_{in}.

The data and the fit are normalized to an ideal detection efficiency (i.e., $\eta_{\rm det} = 1$) for display in Fig. S8. The fit gives an initial device temperature of 24.4 ± 0.4 mK, equivalent to 1.61 ± 0.03 phonons. The thermal conductivity follows a power law in temperature with an exponent $k = 1.09 \pm 0.03$. The fitting parameters are summarized in Table. I.

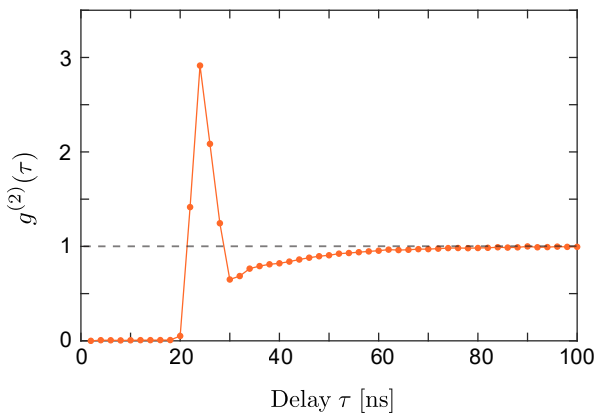

FIG. S9. **Afterpulses in the SNSPDs.** The second-order correlation function of a coherent laser shows a sharp feature in the first 50 ns due to the afterpulsing effect. The flat region before 20 ns is the dead time of the SNSPDs. For delays > 30 ns, the correlation decays to 1, as expected for a coherent state.

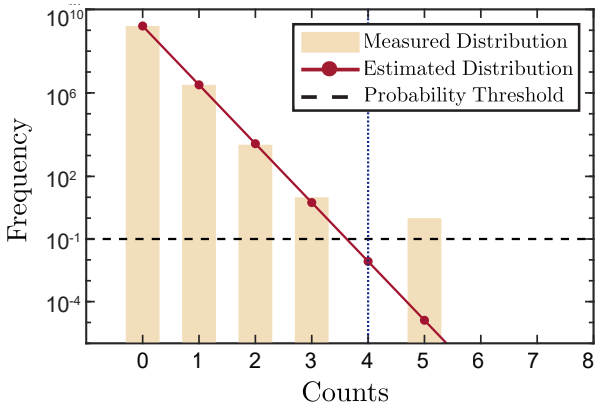

FIG. S10. **Distribution of counts.** Number of intervals as a function of counts. We binned data into 3 μs intervals. The beige bars are the histogram of counts in the intervals. The red solid line is the estimated distribution of counts in the short interval limit $\Delta t \ll \tau_c$, where τ_c is the coherence time. We used a probability threshold 0.1 (shown as the black dashed line) corresponding to a count threshold $k_{\rm thr}$ (shown by the blue dotted line). In this data, one interval with 5 counts is rejected by this protocol.

E. Filtering of Counts Detected by the SNSPDs

Not every count registered by the SNSPDs corresponds to the detection of a photon. Here we cosider three types of false-counts: (i) afterpulses, (ii) rapid bursts of counts in a short duration (10 - 100 µs), and (iii) SNSPD dark counts. This section describes the filtering protocols used to prevent false-counts of types (i) and (ii) from compromising the coherence measurements. False counts of type (iii) (i.e., dark counts) cannot be identified or filtered, given that they occur at random times and are uniformly distributed in time; however, their effect on measured coherences can be accounted for, as is detailed in Sec. F.

(i) Afterpulses It is well-known that reflection of the voltage pulses generated by SNSPDs can cause false counts [5, 6]. Following a detection event at time $t = 0$, an afterpulse can occur at a delay $\tau_{\rm ap}$ set by the time it takes for the voltage reflection to reach the SNSPD.

The signature of afterpulses is illustrated in Fig. S9. It shows the second-order coherence $g^{(2)}(\tau)$ of a power-stabilized laser, measured with a bin-size of 2 ns. The peak at delay $\tau_{\rm ap} \approx 24$ ns is caused by the afterpulses. For $\tau > 50$ ns, $g^{(2)}(\tau) \approx 1$, as expected for laser light.

To reduce the impact of these afterpulses in the experiments described here, any count that is registered within 50 ns of a preceding count (on a given SNSPD) is tagged as an afterpulse and discarded. This also removes any true counts that occur within 50 ns of each other; however the error this introduces in coherence measurements is to effectively shorten the bin at the smallest delay by 50 ns. This can easily be compensated for, but since the shortest delays used in this work are $\tau \geq 1$ µs, we need not perform such a compensation. For the same reason, we also need not compensate for the recovery time of the SNSPD, which is specified to be 50 ns.

(ii) Rapid Bursts There occur rare instances ($< 0.008\%$ of all 90 ms DAq configuration intervals, c.f. Sec. A) in which the SNSPDs register a periodic train of counts with an abnormally high rate ($\gtrsim 10^3\times$ expected count rate) for a short duration ($10 - 100$ µs). While the underlying cause of these spurious counts is unknown, such events can be identified as outliers and discarded. This is done by using a statistical model which tags events having an extremely low probability of occurring, as described below.

Suppose the mean count rate during an experiment is measured to be I. The average number of counts received in a time interval Δt is thus $\lambda = I\Delta t$. The probability of receiving k photons in this interval can be estimated using a statistical model as $P(k, \lambda)$. Thus, for a total DAq time T, among the $N = T/\Delta t$ number of Δt- intervals, the mean number of intervals in which k counts are recorded is expected to be $N \times P(k, \lambda)$. To reject outliers, we set a threshold $\epsilon \ll 1$ and search for all Δt-intervals that receive $\geq k_{\rm thr}$ counts, where $k_{\rm thr}$ is the smallest value of k for which $N \times P(k_{\rm thr}, \lambda) < \epsilon$. For each Δt-interval that is identified as an outlier in this way, we discard the entire 90 ms DAq record containing this interval. In this work, $N \sim 10^8 - 10^{13}$ and ϵ is set to 0.1.

Figure S10 illustrates this protocol. It shows the experimentally measured frequency with which $k-$counts were detected in $\Delta t = 3\,\mu$s intervals for $T \approx 4800$ s ($N \sim 1.6\times10^9$) and $I \approx 2500$ counts/s ($\lambda \sim 1.51\times10^{-3}$). The red solid line represents the statistical model $P(k,\lambda)$ (see below). For $k \geq 4$, $N \times P(k,\lambda) < 0.1$, so that $k_{\rm thr} = 4$. As such, the 90 ms DAq intervals that included any $\Delta t = 3\,\mu$s intervals that registered ≥ 4 counts were discarded from the dataset. In this instance, 1 out of 53,421 DAq intervals was discarded.

We iterated the filtering protocol illustrated above for $\Delta t = 3, 10, 30, 100$ and $300\,\mu$s. The best apriori statistical guess for $P(k,\lambda)$ depends on the choice of Δt, the measured count rate I, and the coherence timescale $\tau_{\rm c}$ of the photon source. For $\Delta t \gg \tau_{\rm c}$ and $I^{-1} \ll \tau_{\rm c}$, the photons detected within the interval Δt can be approximated as uncorrelated, and we model $P(k,\lambda)$ as a Poissonian, $\text{Pois}(k,\lambda) = \lambda^k e^{-\lambda}/k!$ [7]. However, if $\Delta t \lesssim \tau_{\rm c}$ or $I^{-1} \gtrsim \tau_{\rm c}$, $P(k,\lambda)$ is best modelled by accounting for the coherence of the source. Given that the acoustic mode generating the sideband photons is expected to be in thermal equilibrium, $P(k,\lambda)$ is well modelled by $P_{\rm th}(k,\lambda) = \lambda^k/(1+\lambda)^{k+1}$ for $\Delta t \ll \tau_{\rm c}$. For $\Delta t \sim \tau_{\rm c}$, $P(k,\lambda)$ can be estimated to be between these two limiting cases of $\Delta t \gg \tau_{\rm c}$ and $\Delta t \ll \tau_{\rm c}$. To avoid making an apriori guess for $\tau_{\rm c}$, we chose to model $P(k,\lambda)$ with $P_{\rm th}(k,\lambda)$ for all Δt. This achieves a more conservative filtering of the data as it allows for more frequent occurrences of a higher number of counts within any Δt-interval ($P_{\rm th}(k,\lambda) > P(k,\lambda)$ for $k > 2, \lambda < 1$).

Applying this protocol to all of the data acquired used in this work, 133 out of $\sim 1.86\times10^6$ DAq intervals were tagged and discarded as outliers.

F. Subtraction of Coincidences due to Background Photons

This section describes the method used to correct the correlation functions for the presence of background counts in the SNSPD data. As described in Sec. B, these background counts result from stray light, leakage through the filter cavities, and the SNSPD's dark counts.

We measure (not shown) these background counts to have a fixed mean arrival rate over the duration of the experiments, and to be uncorrelated among themselves, i.e. they are Poisson distributed. We also assume them to be independent of the sideband photons. Therefore, the measured second-order correlation function $g^{(2)}_{\rm exp}(\tau)$, including the cross-correlations between the sideband photons and the background photons, is given by

$$\begin{aligned} g^{(2)}_{\rm exp}(\tau) &= \frac{\langle [a^\dagger+\xi^\dagger](0)[a^\dagger+\xi^\dagger](\tau)[a+\xi](\tau)[a+\xi](0)\rangle}{\langle [a^\dagger+\xi^\dagger](0)[a+\xi](0)\rangle\langle [a^\dagger+\xi^\dagger](\tau)[a+\xi](\tau)\rangle} \\ &= \frac{g^{(2)}(\tau)+2\epsilon+\epsilon^2}{(1+\epsilon)^2}, \end{aligned}$$

where a and ξ correspond to the sideband and background photons, respectively, and we have used $\langle a^\dagger\xi\rangle = 0 = \langle \xi^\dagger a\rangle$, and $\langle \xi^\dagger(0)\xi^\dagger(\tau)\xi(\tau)\xi(0)\rangle = \langle\xi^\dagger\xi\rangle^2$. $\epsilon = \langle\xi^\dagger\xi\rangle/\langle a^\dagger a\rangle$ is the ratio of the mean background and sideband count rates. For the drive powers used in this work, $0.04 \lesssim \epsilon \lesssim 0.2$.

The corrected second-order correlation function $g^{(2)}(\tau)$ is thus given by

$$g^{(2)}(\tau) = g^{(2)}_{\rm exp}(\tau) + 2(g^{(2)}_{\rm exp}(\tau)-1)\epsilon + (g^{(2)}_{\rm exp}(\tau)-1)\epsilon^2.$$

A similar calculation gives the corrected third-order and fourth-order correlation functions using

$$g^{(3)}_{\rm exp}(\tau_1,\tau_2) = \frac{g^{(3)}(\tau_1,\tau_2) + \epsilon\left(g^{(2)}(\tau_1)+g^{(2)}(\tau_2)+g^{(2)}(\tau_1+\tau_2)\right) + 3\epsilon^2+\epsilon^3}{(1+\epsilon)^3},$$

$$g^{(4)}_{\rm exp}(\tau_1,\tau_2,\tau_3) = \frac{g^{(4)}(\tau_1,\tau_2,\tau_3) + \epsilon\sum_{g^{(3)}_i\in\mathbb{G}^{(3)}} g^{(3)}_i + \epsilon^2\sum_{g^{(2)}_j\in\mathbb{G}^{(2)}} g^{(2)}_j + 4\epsilon^3+\epsilon^4}{(1+\epsilon)^4},$$

where $\mathbb{G}^{(3)}$ and $\mathbb{G}^{(2)}$ are

$$\mathbb{G}^{(3)} = \{\, g^{(3)}(\tau_1,\tau_2), g^{(3)}(\tau_1+\tau_2,\tau_3), g^{(3)}(\tau_1,\tau_2+\tau_3), g^{(3)}(\tau_2,\tau_3)\,\},$$

$$\mathbb{G}^{(2)} = \{\, g^{(2)}(\tau_1), g^{(2)}(\tau_2), g^{(2)}(\tau_3), g^{(2)}(\tau_1+\tau_2), g^{(2)}(\tau_2+\tau_3), g^{(2)}(\tau_1+\tau_2+\tau_3)\,\}.$$

G. Correspondence between Photon and Phonon Correlation Functions

In the main text, we introduced the normally ordered phonon coherence $g_{\rm ac}^{(n)}$ obtained from the anti-Stokes photon coherence and the anti-normally ordered phonon coherence $h_{\rm ac}^{(n)}$ obtained from the Stokes photon coherence. In this section, we formally demonstrate this one-to-one correspondence between coherences of the optical fields output by the cavity and the coherences of the phonons. We also note that optical losses along the detection chain do not compromise these coherence measurements [8].

In a frame rotating at the optical drive frequency $\omega_{\rm D} = \omega_{\rm c} + \Delta$, the output optical field o of the optomechanical cavity has a spectrum $S_{o^\dagger o}$ that relates to mechanical spectrum S_{xx} as [9]

$$S_{o^\dagger o}(\omega) = \kappa_{\rm in} n_{\rm c} 2\pi\delta(\omega) + \kappa_{\rm in} g_0^2 n_{\rm c} |\chi_{\rm c}(-\omega)|^2 S_{xx}(\omega), \quad (11)$$

where the first term corresponds to the optical drive field and the second term corresponds to the optical fields produced by the mechanical motion through the optomechanical coupling. $\chi_{\rm c}(\omega)$ is the bare optical cavity susceptibility $\chi_{\rm c}(\omega) = (\kappa_{\rm c}/2 - i(\omega + \Delta))^{-1}$. The field p detected by the SNSPDs is the cavity output field o passed through the filter cavities with transmission $f_{\rm filter}(\omega)$,

$$p(t) = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} e^{-i\omega t} f_{\rm filter}(\omega) o(\omega),$$

so that the detected normally ordered optical correlations are given by the Weiner-Khinchin relation

$$\langle p^\dagger(t+\tau)p(t)\rangle = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi}\, e^{-i\omega\tau} |f_{\rm filter}(-\omega)|^2 S_{o^\dagger o}(\omega),$$

$f_{\rm filter}(\omega)$ is non-zero only around $\omega = +\omega_{\rm ac}$ for the red-detuned drive (and only around $\omega = -\omega_{\rm ac}$ for the blue-detuned drive), within the two filter cavity linewidths $\kappa_{\rm FC1}, \kappa_{\rm FC2} \ll \omega_{\rm ac}$. This ensures that the contribution of the first term in Eq. 11 to this integral, corresponding to the contribution of the drive field, is negligible. As such,

$$\langle p^\dagger(t+\tau)p(t)\rangle = \quad (12)$$
$$\kappa_{\rm in} g_0^2 n_{\rm c} \int_{-\infty}^{\infty} \frac{d\omega}{2\pi}\, e^{-i\omega\tau} |f_{\rm filter}(-\omega)|^2 |\chi_{\rm c}(-\omega)|^2 S_{xx}(\omega).$$

The contribution of $S_{xx}(\omega)$ to these optical correlations can be decomposed as

$$S_{xx}(\omega) = x_{\rm zpt}^2 [S_{b^\dagger b}(\omega) + S_{bb^\dagger}(\omega)],$$

with $x = x_{\rm zpt}(b + b^\dagger)$, and $x_{\rm zpt}$ being the zero-point motion of the oscillator. $S_{b^\dagger b}(\omega)$ is non-zero only around $\omega = -\omega_{\rm ac}$, and $S_{bb^\dagger}(\omega)$ is non-zero only around $\omega = +\omega_{\rm ac}$, corresponding to the anti-Stokes and Stokes scattering processes respectively, which occur within the mechanical susceptibility $\chi_{\rm ac}(\omega)$ of linewidth $\bar{\gamma}_{\rm ac}$. This is explicitly seen through the oscillator's Langevin equation [10–13]

$$\dot{b} = -\left(\frac{\bar{\gamma}_{\rm ac}}{2} - i\omega_{\rm ac}\right) b + \sqrt{\bar{\gamma}_{\rm ac}} b_{\rm in},$$

where $b_{\rm in}$ is the thermal noise input with $\langle b_{\rm in}\rangle = 0$. The correlations for $\tau \geq 0$ are given by $\langle b_{\rm in}^\dagger(\tau) b_{\rm in}(0)\rangle = n_{\rm ac}\delta(\tau)$ and $\langle b_{\rm in}(\tau) b_{\rm in}^\dagger(0)\rangle = (n_{\rm ac}+1)\delta(\tau)$. And $n_{\rm ac} = \langle b(t)^\dagger b(t)\rangle$ is assumed to be time-independent (stationary). The formal integral

$$b(\tau) = \sqrt{\bar{\gamma}_{\rm ac}} \int_{-\infty}^{\infty} dt' e^{-(\bar{\gamma}_{\rm ac}/2 - i\omega_{\rm ac})(\tau - t')} b_{\rm in}(t')$$

yields the correlations and spectra as

$$\langle b^\dagger(\tau) b(0)\rangle = n_{\rm ac} e^{-(\bar{\gamma}_{\rm ac}/2 + i\omega_m)\tau}, \quad (13)$$
$$\begin{aligned} S_{b^\dagger b}(\omega) &= \int_{-\infty}^{\infty} d\tau\; e^{i\omega\tau} \langle b^\dagger(\tau) b(0)\rangle \\ &= n_{\rm ac}\bar{\gamma}_{\rm ac} |\chi_{\rm ac}(-\omega)|^2 \\ &= \frac{n_{\rm ac}\bar{\gamma}_{\rm ac}}{(\bar{\gamma}_{\rm ac}/2)^2 + (-\omega - \omega_{\rm ac})^2}. \end{aligned}$$

And similarly

$$\langle b(\tau) b^\dagger(0)\rangle = (n_{\rm ac}+1) e^{-(\bar{\gamma}_{\rm ac}/2 - i\omega_m)\tau}, \quad (14)$$
$$\begin{aligned} S_{bb^\dagger}(\omega) &= \int_{-\infty}^{\infty} d\tau\; e^{i\omega\tau} \langle b(\tau) b^\dagger(0)\rangle \\ &= (n_{\rm ac}+1)\bar{\gamma}_{\rm ac} |\chi_{\rm ac}(\omega)|^2 \\ &= \frac{(n_{\rm ac}+1)\bar{\gamma}_{\rm ac}}{(\bar{\gamma}_{\rm ac}/2)^2 + (\omega - \omega_{\rm ac})^2}. \end{aligned}$$

Thus, for the case of the red-detuned drive $\Delta = -\omega_{\rm ac}$, only the $S_{b^\dagger b}$ term contributes to the integral around $\omega = -\omega_{\rm ac}$ in Eq. 12, corresponding to the anti-Stokes scattered photons, and yields

$$\begin{aligned} \langle p^\dagger(t+\tau)p(t)\rangle &= \langle a_{\rm AS}^\dagger(t+\tau) a_{\rm AS}(t)\rangle \\ &= \kappa_{\rm in} g_0^2 n_{\rm c} |\chi_{\rm c}(\omega_{\rm ac})|^2 x_{\rm zpt}^2 \int_{-\infty}^{\infty} \frac{d\omega}{2\pi}\, e^{-i\omega\tau} S_{b^\dagger b}(\omega) \\ &= 4\kappa_{\rm in} g_0^2 n_{\rm c} x_{\rm zpt}^2/\kappa_{\rm c}^2\; \langle b^\dagger(t+\tau) b(t)\rangle, \end{aligned}$$

where we have approximated $f_{\rm filter}(-\omega) \approx 1$ and $\chi_{\rm c}(-\omega) \approx \chi_{\rm c}(\omega_{\rm ac})$ for ω around $-\omega_{\rm ac}$, as $\bar{\gamma}_{\rm ac} \ll \kappa_{\rm FC1}, \kappa_{\rm FC2} \ll \kappa_{\rm c}$.

For the case of the blue-detuned drive $\Delta = +\omega_{\rm ac}$, only the $S_{bb^\dagger}$ term contributes to the integral around $\omega = +\omega_{\rm ac}$ in Eq. 12, corresponding to the Stokes scattered photons, and yields

$$\begin{aligned}\langle p^\dagger(t+\tau)p(t)\rangle &= \langle a_{\rm S}^\dagger(t+\tau)a_{\rm S}(t)\rangle \\ &= \kappa_{\rm in} g_0^2 n_{\rm c} |\chi_{\rm c}(-\omega_{\rm ac})|^2 x_{\rm zpt}^2 \int_{-\infty}^{\infty} \frac{d\omega}{2\pi}\, e^{-i\omega\tau} S_{bb^\dagger}(\omega) \\ &= 4\kappa_{\rm in} g_0^2 n_{\rm c} x_{\rm zpt}^2/\kappa_{\rm c}^2\ \langle b(t+\tau)b^\dagger(t)\rangle,\end{aligned}$$

where we have approximated $f_{\rm filter}(-\omega) \approx 1$ and $\chi_{\rm c}(-\omega) \approx \chi_{\rm c}(-\omega_{\rm ac})$ for ω around $+\omega_{\rm ac}$, as $\bar{\gamma}_{\rm ac} \ll \kappa_{\rm FC1}, \kappa_{\rm FC2} \ll \kappa_{\rm c}$.

This demonstrates the correspondence between the detected optical correlations and the appropriately ordered phonon correlations. Note that high-order correlations can be evaluated in terms of such two-time correlations using Wick's theorem, for any Gaussian state, thereby implying the correspondence of the detected optical coherences and the appropriately ordered phonon coherences. Note that the proportionality constant $(4\kappa_{\rm in} g_0^2 n_{\rm c} x_{\rm zpt}^2/\kappa_{\rm c}^2)$ is immaterial for the coherences, which are normalized ratios of such correlations.

H. Correlation Functions of a Thermal State

This section presents calculations of coherence functions for a thermal state of an oscillator. These functions are used to fit the data in the main text.

We calculate the coherences for an oscillator's thermal state using Wick's theorem. For any Gaussian state, the theorem implies that high-order moments can be evaluated in terms of second moments. For example, the numerators in

$$g_{\rm ac}^{(2)}(\tau_1) = \frac{\langle b_0^\dagger b_1^\dagger b_1 b_0\rangle}{\langle b_0^\dagger b_0\rangle\langle b_1^\dagger b_1\rangle},\ h_{\rm ac}^{(2)}(\tau_1) = \frac{\langle b_0 b_1 b_1^\dagger b_0^\dagger\rangle}{\langle b_0 b_0^\dagger\rangle\langle b_1 b_1^\dagger\rangle}$$

where $b_0 = b(0)$ and $b_1 = b(\tau_1)$, evaluate as

$$\begin{aligned}\langle b_0^\dagger b_1^\dagger b_1 b_0\rangle &= \langle b_0^\dagger b_0\rangle\langle b_1^\dagger b_1\rangle + \langle b_0^\dagger b_1\rangle\langle b_1^\dagger b_0\rangle, \\ \langle b_0 b_1 b_1^\dagger b_0^\dagger\rangle &= \langle b_0 b_0^\dagger\rangle\langle b_1 b_1^\dagger\rangle + \langle b_0 b_1^\dagger\rangle\langle b_1 b_0^\dagger\rangle.\end{aligned}$$

For the thermal state, which is a Gaussian state, these second moments are given for $\tau \geq 0$ by Eq. 13, 14, i.e.

$$\begin{aligned}\langle b^\dagger(\tau)b(0)\rangle &= n_{\rm ac} e^{-(\bar{\gamma}_{\rm ac}/2 + i\omega_{\rm ac})\tau}, \\ \langle b(\tau)b^\dagger(0)\rangle &= (n_{\rm ac}+1) e^{-(\bar{\gamma}_{\rm ac}/2 - i\omega_{\rm ac})\tau}.\end{aligned}$$

The second moments for $\tau \leq 0$ are obtained by conjugation. Combined, these readily give

$$g_{\rm ac}^{(2)}(\tau) = 1 + e^{-\bar{\gamma}_{\rm ac}\tau} = h_{\rm ac}^{(2)}(\tau).$$

To illustrate further, the numerator in

$$g_{\rm ac}^{(3)}(\tau_1,\tau_2) = \frac{\langle b_0^\dagger b_1^\dagger b_2^\dagger b_2 b_1 b_0\rangle}{\langle b_0^\dagger b_0\rangle\langle b_1^\dagger b_1\rangle\langle b_2^\dagger b_2\rangle},$$

where $b_2 = b(\tau_1+\tau_2)$, evaluates as

$$\begin{aligned}\langle b_0^\dagger b_1^\dagger b_2^\dagger b_2 b_1 b_0\rangle &= \langle b_0^\dagger b_0\rangle\langle b_1^\dagger b_1\rangle\langle b_2^\dagger b_2\rangle \\ &+ \langle b_0^\dagger b_0\rangle\langle b_1^\dagger b_2\rangle\langle b_2^\dagger b_1\rangle + \langle b_0^\dagger b_1\rangle\langle b_1^\dagger b_0\rangle\langle b_2^\dagger b_2\rangle \\ &+ \langle b_0^\dagger b_1\rangle\langle b_1^\dagger b_2\rangle\langle b_2^\dagger b_0\rangle + \langle b_0^\dagger b_2\rangle\langle b_1^\dagger b_0\rangle\langle b_2^\dagger b_1\rangle \\ &+ \langle b_0^\dagger b_2\rangle\langle b_1^\dagger b_1\rangle\langle b_2^\dagger b_0\rangle.\end{aligned}$$

This, and evaluations of $h_{\rm ac}^{(3)}, g_{\rm ac}^{(4)}$, and $h_{\rm ac}^{(4)}$ in a similar manner, gives

$$\begin{aligned}g_{\rm ac}^{(3)}(\tau_1,\tau_2) &= h_{\rm ac}^{(3)}(\tau_1,\tau_2) \\ &= 1 + e^{-\bar{\gamma}_{\rm ac}\tau_1} + e^{-\bar{\gamma}_{\rm ac}\tau_2} + 3e^{-\bar{\gamma}_{\rm ac}(\tau_1+\tau_2)}, \\ g_{\rm ac}^{(4)}(\tau_1,\tau_2,\tau_3) &= h_{\rm ac}^{(4)}(\tau_1,\tau_2,\tau_3) \\ &= 1 + e^{-\bar{\gamma}_{\rm ac}\tau_1} + e^{-\bar{\gamma}_{\rm ac}\tau_2} + e^{-\bar{\gamma}_{\rm ac}\tau_3} \\ &+ 3e^{-\bar{\gamma}_{\rm ac}(\tau_1+\tau_2)} + 3e^{-\bar{\gamma}_{\rm ac}(\tau_2+\tau_3)} + e^{-\bar{\gamma}_{\rm ac}(\tau_1+\tau_3)} \\ &+ 9e^{-\bar{\gamma}_{\rm ac}(\tau_1+\tau_2+\tau_3)} + 4e^{-\bar{\gamma}_{\rm ac}(\tau_1+2\tau_2+\tau_3)}. \qquad (15)\end{aligned}$$

Limiting cases amenable to an intuitive interpretation are discussed below.

The fourth-order phonon coherence functions $g_{\rm ac}^{(4)}(\tau_1,\tau_2,\tau_3)$ and $h_{\rm ac}^{(4)}(\tau_1,\tau_2,\tau_3)$ are three-dimensional data sets that can be visualized as two-dimensional slices. In Fig. 2(c) of the main text we show one set of 2D slices $g_{\rm ac}^{(4)}(0^+,\tau_2,\tau_3)$ and $h_{\rm ac}^{(4)}(0^+,\tau_2,\tau_3)$, where $\tau = 0^+$ represents the bin with $5\,\mu{\rm s} < \tau < 15\,\mu{\rm s}$. In Fig. S11, we show three more sets of 2D slices: $g_{\rm ac}^{(4)}(\boldsymbol{\tau})$ and $h_{\rm ac}^{(4)}(\boldsymbol{\tau})$ measured at $\boldsymbol{\tau} = (\tau_1, 0^+, \tau_3)$ (panel (a)), $\boldsymbol{\tau} = (\tau_1, \infty, \tau_3)$ (panel (b)), and $\boldsymbol{\tau} = (\tau_1, \tau_2, \infty)$ (panel (c)). Here, $\tau = \infty$ represents times $\tau > 10\bar{\gamma}_{\rm ac}^{-1}$. Fits to the entire 3D (τ_1-, τ_2-, τ_3- dependent) data are shown as 2D surfaces, with fit residuals shown in black.

The 2D slice $g_{\rm ac}^{(4)}(\tau_1,\infty,\tau_3)$ is equivalent to the product of two one-dimensional $g_{\rm ac}^{(2)}(\tau)$ functions. If the third photon arrives with delay $\tau_2 = \infty$, then the arrivals of the third and fourth photons are uncorrelated with the arrival of the first pair of photons. That is, $g_{\rm ac}^{(4)}(\tau_1,\infty,\tau_3)$ is proportional to the product of the probability of measuring a pair of photons separated by a delay τ_1 and the probability of measuring another pair of photons separated by a delay τ_3. This can be seen from the expression for $g_{\rm ac}^{(4)}(\tau_1,\tau_2,\tau_3)$ given in Eq. 15, where setting $\tau_2 = \infty$

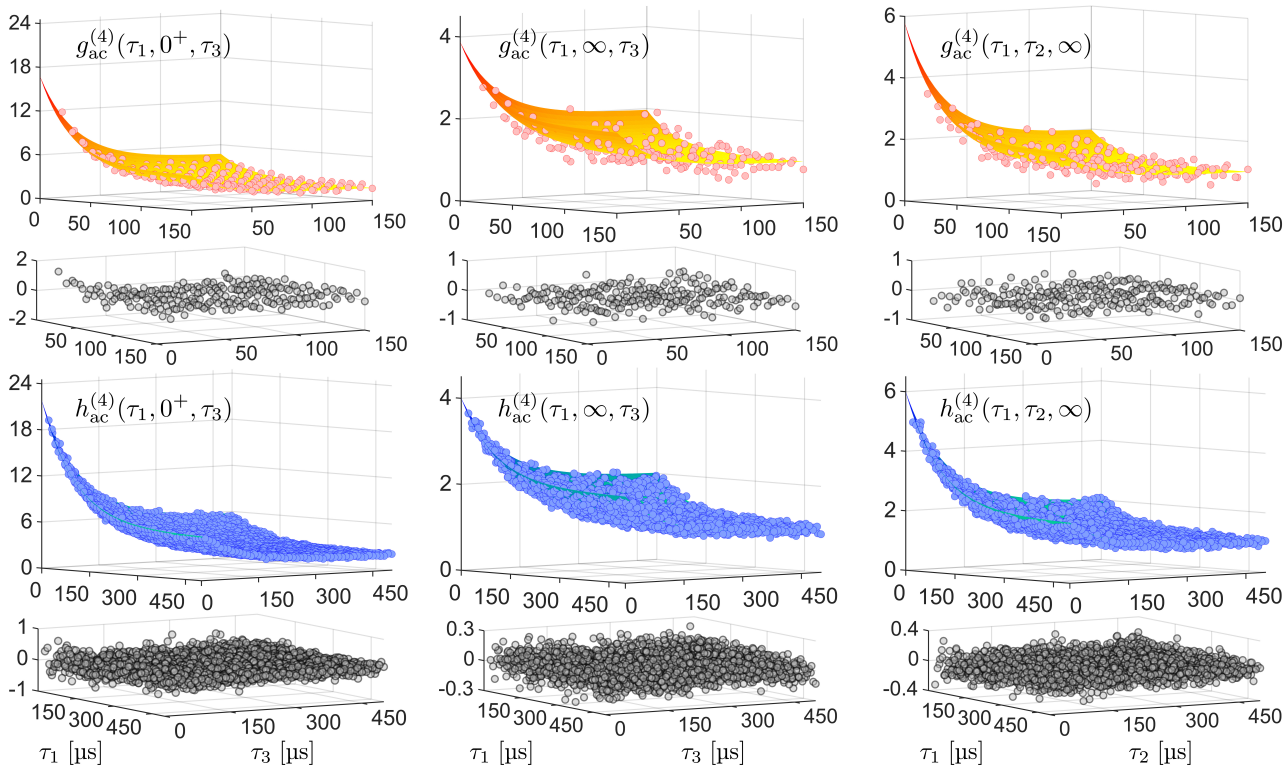

FIG. S11. **Fourth-order phonon coherences.** $g^{(4)}_{\rm ac}(\tau_1, \tau_2, \tau_3)$ and $h^{(4)}_{\rm ac}(\tau_1, \tau_2, \tau_3)$ measured for $P_{\rm in} \approx 5\,\mu$W, with photon arrival times binned in 10 μs bins. Shown are three sets of representative 2D slices: $g^{(4)}_{\rm ac}(\tau_1, 0^+, \tau_3)$ and $h^{(4)}_{\rm ac}(\tau_1, 0^+, \tau_3)$, $g^{(4)}_{\rm ac}(\tau_1, \infty, \tau_3)$ and $h^{(4)}_{\rm ac}(\tau_1, \infty, \tau_3)$, $g^{(4)}_{\rm ac}(\tau_1, \tau_2, \infty)$ and $h^{(4)}_{\rm ac}(\tau_1, \tau_2, \infty)$, where $\tau = 0^+$ represents the bin with $5\,\mu{\rm s} < \tau < 15\,\mu{\rm s}$ and $\tau = \infty$ represents times $\tau > 10\bar{\gamma}_{\rm ac}^{-1}$. Fits are to the entire 3D (τ_1-, τ_2-, τ_3- dependent) data. Fit residuals are shown in black.

results in

$$g^{(4)}_{\rm ac}(\tau_1, \infty, \tau_3) = 1 + e^{-\bar{\gamma}_{\rm ac}\tau_1} + e^{-\bar{\gamma}_{\rm ac}\tau_3} + e^{-\bar{\gamma}_{\rm ac}(\tau_1+\tau_3)} = g^{(2)}_{\rm ac}(\tau_1) g^{(2)}_{\rm ac}(\tau_3).$$

The same equivalence holds for $h^{(4)}_{\rm ac}(\tau_1, \infty, \tau_3) = h^{(2)}_{\rm ac}(\tau_1) h^{(2)}_{\rm ac}(\tau_3)$.

Similarly, the slice $g^{(4)}_{\rm ac}(\tau_1, \tau_2, \infty)$ is equivalent to the two-dimensional $g^{(3)}_{\rm ac}(\tau_1, \tau_2)$. If the fourth photon arrives with delay $\tau_3 = \infty$, then its arrival is uncorrelated with the arrivals of the first three photons. Therefore, $g^{(4)}_{\rm ac}(\tau_1, \tau_2, \infty)$ is proportional to the probability of measuring a triplet of photons with a delay τ_1 between the first and second and a delay τ_2 between the second and third. We can also see this by setting $\tau_3 = \infty$ in Eq. 15, which yields

$$g^{(4)}_{\rm ac}(\tau_1, \tau_2, \infty) = 1 + e^{-\bar{\gamma}_{\rm ac}\tau_1} + e^{-\bar{\gamma}_{\rm ac}\tau_2} + 3e^{-\bar{\gamma}_{\rm ac}(\tau_1+\tau_2)} = g^{(3)}_{\rm ac}(\tau_1, \tau_2).$$

Again, the same is true for $h^{(4)}_{\rm ac}(\tau_1, \tau_2, \infty) = h^{(3)}_{\rm ac}(\tau_1, \tau_2)$.

I. Coherences of k-quanta- Subtracted/Added Thermal States

In the main text we discussed out-of-equilibrium k-quanta-subtracted/added thermal states. In this section, we derive the relation between the mean occupancies and coherences of such states and those of the thermal states they are generated from.

The subtraction (addition) of a single phonon at time $t = 0$ from a state described by density matrix ρ yields the state with density matrix ρ_{-1} (ρ_{+1})

$$\rho_{-1} = \frac{b(0)\rho b^\dagger(0)}{{\rm Tr}[b(0)\rho b^\dagger(0)]} = \frac{b(0)\rho b^\dagger(0)}{{\rm Tr}[\rho b^\dagger(0) b(0)]} = \frac{b(0)\rho b^\dagger(0)}{\langle b^\dagger(0) b(0)\rangle_\rho} \tag{16}$$

$$\rho_{+1} = \frac{b^\dagger(0)\rho b(0)}{{\rm Tr}[b^\dagger(0)\rho b(0)]} = \frac{b^\dagger(0)\rho b(0)}{{\rm Tr}[\rho b(0) b^\dagger(0)]} = \frac{b^\dagger(0)\rho b(0)}{\langle b(0) b^\dagger(0)\rangle_\rho} \tag{17}$$

where the denominator signifies the normalization, and

$\langle\cdot\rangle_\rho$ is the expectation value for state ρ. The coherences and mean phonon number of the single-phonon subtracted (added) state ρ_{-1} (ρ_{+1}) can thus be evaluated in terms of the coherences of the steady state ρ, as follows.

Coherences By way of illustration, consider the second-order coherence of the single-phonon subtracted state,

$$g^{(2)}_{\rm ac}\big|_{-1}(\tau) = \frac{\langle b^\dagger(0)b^\dagger(\tau)b(\tau)b(0)\rangle_{\rho_{-1}}}{\langle b^\dagger(0)b(0)\rangle_{\rho_{-1}}\ \langle b^\dagger(\tau)b(\tau)\rangle_{\rho_{-1}}}$$

Using Eq. 16, the numerator equates to

$$\frac{\langle b^\dagger(0)b^\dagger(0)b^\dagger(\tau)b(\tau)b(0)b(0)\rangle_\rho}{\langle b^\dagger(0)b(0)\rangle_\rho} = g^{(3)}_{\rm ac}(0,\tau)\ \langle b^\dagger(0)b(0)\rangle^2_\rho$$

and the denominator equates to

$$\frac{\langle b^\dagger(0)b^\dagger(0)b(0)b(0)\rangle_\rho}{\langle b^\dagger(0)b(0)\rangle_\rho}\ \frac{\langle b^\dagger(0)b^\dagger(\tau)b(\tau)b(0)\rangle_\rho}{\langle b^\dagger(0)b(0)\rangle_\rho} = g^{(2)}_{\rm ac}(0)\ g^{(2)}_{\rm ac}(\tau)\ \langle b^\dagger(0)b(0)\rangle^2_\rho$$

so that

$$g^{(2)}_{\rm ac}\big|_{-1}(\tau) = \frac{g^{(3)}_{\rm ac}(0,\tau)}{g^{(2)}_{\rm ac}(0)\ g^{(2)}_{\rm ac}(\tau)} \tag{18}$$

And similarly,

$$h^{(2)}_{\rm ac}\big|_{+1}(\tau) = \frac{h^{(3)}_{\rm ac}(0,\tau)}{h^{(2)}_{\rm ac}(0)\ h^{(2)}_{\rm ac}(\tau)} \tag{19}$$

For the helium acoustic oscillator in this work that is initialized in a thermal state $\rho_{\rm th}$, this gives

$$g^{(2)}_{\rm ac}\big|_{-1,\rm th}(\tau) = h^{(2)}_{\rm ac}\big|_{+1,\rm th}(\tau) = \frac{1+2e^{-\bar\gamma_{\rm ac}t}}{1+e^{-\bar\gamma_{\rm ac}t}}.$$

These functions are plotted as solid lines alongside the measurements in Fig. 4(b) of the main text. Note that evaluation of the data points in Fig. 4(b) makes explicit use of post-selected data. This is because the experimental realization of Eq. 18, 19,

$$\frac{g^{(3)}_{\rm ac}(0,\tau)}{g^{(2)}_{\rm ac}(0)\ g^{(2)}_{\rm ac}(\tau)} = \frac{C^{(3)}_{\rm AS}(0^+,\tau)/C^{(3)}_{\rm AS}(\infty,\infty)}{C^{(2)}_{\rm AS}(0^+)/C^{(2)}_{\rm AS}(\infty)\cdot C^{(2)}_{\rm AS}(\tau)/C^{(2)}_{\rm AS}(\infty)},$$

$$\frac{h^{(3)}_{\rm ac}(0,\tau)}{h^{(2)}_{\rm ac}(0)\ h^{(2)}_{\rm ac}(\tau)} = \frac{C^{(3)}_{\rm S}(0^+,\tau)/C^{(3)}_{\rm S}(\infty,\infty)}{C^{(2)}_{\rm S}(0^+)/C^{(2)}_{\rm S}(\infty)\cdot C^{(2)}_{\rm S}(\tau)/C^{(2)}_{\rm S}(\infty)}$$

involves $C^{(3)}_{\rm AS(S)}(0^+,\tau)$ (and $C^{(2)}_{\rm AS(S)}(\tau)$) which are photon pairs (and counts) conditioned on the arrival of an anti-Stokes (Stokes) photon at time $t=0$, which heralds the subtraction (addition) of one phonon. As described in Sec. H and Fig. S11, each ∞ in the normalizations $C^{(3)}_{\rm AS(S)}(\infty,\infty)$ and $C^{(2)}_{\rm AS(S)}(\infty)$ represents times $\tau > 10\bar\gamma^{-1}_{\rm ac}$. The time $t=0^+$ represents the finite time of 10 µs (or 5 µs) of the experimental data bin used. It corresponds to the time of the first click used to measure the $g^{(2)}_{\rm ac}$ (or $h^{(2)}_{\rm ac}$) of the phonon subtracted (or added) state. This leads to a finite-time correction of the theoretical predictions, using Eq. 18, 19. This finite-time correction results in a slight shift of the theoretical curves shown in Fig. 4(b); however since it is negligible on the scale of Fig. 4(b) we choose not to show it.

More generally, the subtraction (addition) of k phonons at time $t=0$ from a state described by density matrix ρ yields the state with density matrix ρ_{-k} (ρ_{+k})

$$\rho_{-k} = \frac{(b(0))^k\ \rho\ (b^\dagger(0))^k}{\langle (b^\dagger(0))^k\ (b(0))^k\rangle_\rho} \tag{20}$$

$$\rho_{+k} = \frac{(b^\dagger(0))^k\ \rho\ (b(0))^k}{\langle (b(0))^k\ (b^\dagger(0))^k\rangle_\rho} \tag{21}$$

Through an *ab-initio* evaluation similar to that illustrated above, the $n^{\rm th}$-order coherence of a k-phonon subtracted (added) state is evaluated in terms of the coherences of the steady state ρ to be

$$g^{(n)}_{\rm ac}\big|_{-k}(\boldsymbol{\tau}) = \frac{g^{(k+n)}_{\rm ac}(\mathbf{0}^{\otimes k},\boldsymbol{\tau})\ \big(g^{(k)}_{\rm ac}(\mathbf{0})\big)^{n-1}}{g^{(k+1)}_{\rm ac}(\mathbf{0})\big[\prod_{p=1}^{n-1} g^{(k+1)}_{\rm ac}\big(\mathbf{0}^{\otimes(k-1)},t_p\big)\big]}$$

$$h^{(n)}_{\rm ac}\big|_{+k}(\boldsymbol{\tau}) = \frac{h^{(k+n)}_{\rm ac}(\mathbf{0}^{\otimes k},\boldsymbol{\tau})\ \big(h^{(k)}_{\rm ac}(\mathbf{0})\big)^{n-1}}{h^{(k+1)}_{\rm ac}(\mathbf{0})\big[\prod_{p=1}^{n-1} h^{(k+1)}_{\rm ac}\big(\mathbf{0}^{\otimes(k-1)},t_p\big)\big]}$$

where $\boldsymbol{\tau} = (\tau_1,\tau_2,...,\tau_{n-1})$, $\mathbf{0}^{\otimes k} = (0,0,...\ k\text{ times})$, and $t_p = \sum_{j=1}^{p}\tau_j$ is the $(p+1)^{\rm th}$ time. (Recall that τ_k is the delay between the $k^{\rm th}$ and $(k+1)^{\rm th}$ time.)

Note that the derivation and discussion above implies that the coincidences and normalization required to evaluate the $n^{\rm th}$-order coherences of k-photon heralded states occur as various 'slices' in the higher-dimensional $(n+k)$-photon detection record. While the record, viewed as a whole, corresponds to that of the equilibrium state (which, for this work, is a thermal state), the post-selection extracts the non-equilibrium (non-thermal) heralded state coherences.

For the helium acoustic oscillator in this work that is initialized in a thermal state $\rho_{\rm th}$, then,

$$g^{(2)}_{\rm ac}|_{-k}(\tau) = \frac{g^{(k+2)}_{\rm ac}(\mathbf{0}^{\otimes k},\tau)g^{(k)}_{\rm ac}(\mathbf{0})}{g^{(k+1)}_{\rm ac}(\mathbf{0})g^{(k+1)}_{\rm ac}(\mathbf{0}^{\otimes(k-1)},\tau)},$$

$$g^{(2)}_{\rm ac}|_{-k,\rm th}(\tau) = \frac{1+(k+1)e^{-\bar\gamma_{\rm ac}\tau}}{1+ke^{-\bar\gamma_{\rm ac}\tau}},$$

and

$$h^{(2)}_{\rm ac}|_{+k}(\tau) = \frac{h^{(k+2)}_{\rm ac}(\mathbf{0}^{\otimes k},\tau)h^{(k)}_{\rm ac}(\mathbf{0})}{h^{(k+1)}_{\rm ac}(\mathbf{0})h^{(k+1)}_{\rm ac}(\mathbf{0}^{\otimes(k-1)},\tau)},$$
$$h^{(2)}_{\rm ac}|_{+k,\rm th}(\tau) = \frac{1+(k+1)e^{-\bar{\gamma}_{\rm ac}\tau}}{1+ke^{-\bar{\gamma}_{\rm ac}\tau}}.$$

Mean phonon numbers The mean phonon number of a k-phonon subtracted state at time τ after its generation is given by (using Eq. 20)

$$n^{-k}_{\rm ac}(\tau) = \langle b^\dagger(\tau)b(\tau)\rangle_{\rho_{-k}} = \frac{\langle (b^\dagger(0))^k b^\dagger(\tau)b(\tau)(b(0))^k\rangle_\rho}{\langle (b^\dagger(0))^k\ (b(0))^k\rangle_\rho}.$$

Dividing both sides by $n_{\rm ac}(0) = \langle b^\dagger(0)b(0)\rangle_\rho$ yields

$$\frac{n^{-k}_{\rm ac}(\tau)}{n_{\rm ac}(0)} = \frac{g^{(k+1)}_{\rm ac}(\mathbf{0}^{\otimes(k-1)},\tau)}{g^{(k)}_{\rm ac}(\mathbf{0})} \tag{22}$$

Eq. 22 has an intuitive interpretation in terms of the scattering rates of photons by the mechanical oscillator. The numerator $g^{(k+1)}_{\rm ac}(\mathbf{0}^{\otimes(k-1)},\tau)$ is proportional to the rate of detecting an anti-Stokes photon at time $t=\tau$ after a $k-$phonon subtraction has been heralded at time $t=0$ (by the detection of k photons at time $t=0$). The denominator $g^{(k)}_{\rm ac}(\mathbf{0}) = g^{(k+1)}_{\rm ac}(\mathbf{0}^{\otimes(k-1)},\infty)$ is proportional to the rate of detecting an anti-Stokes photon at time $t\to\infty$ after a $k-$phonon subtraction has been heralded at time $t=0$. The proportionality constant for both is simply the rate of detecting uncorrelated arrivals of $k+1$ photons. The numerator is also proportional to $n^{-k}_{\rm ac}(\tau)$ (as the anti-Stokes scattering rate is proportional to the mean phonon occupancy – the proportionality constant is simply the scattering rate $\gamma_{\rm AS}$, of the main text, times the net detection efficiency). Whereas the denominator is proportional to $n^{-k}_{\rm ac}(\infty) = n_{\rm ac}(0)$; the latter equality coming from the fact that for $\tau\to\infty$, the heralded k-phonon subtracted state equilibrates back to the steady state. Taking a ratio of these rates, interpreted these two ways, readily yields Eq. 22.

Similarly, the mean phonon number of a k-phonon added state at time τ after its generation is given by (using Eq. 21)

$$\begin{aligned} n^{+k}_{\rm ac}(\tau)+1 &= \langle b^\dagger(\tau)b(\tau)\rangle_{\rho_{+k}} + 1 \\ &= \langle b(\tau)b^\dagger(\tau)\rangle_{\rho_{+k}} \\ &= \frac{\langle (b(0))^k b(\tau)b^\dagger(\tau)(b^\dagger(0))^k\rangle_\rho}{\langle (b(0))^k\ (b^\dagger(0))^k\rangle_\rho}. \end{aligned}$$

Dividing both sides by $n_{\rm ac}(0)+1 = \langle b(0)b^\dagger(0)\rangle_\rho$ yields

$$\frac{n^{+k}_{\rm ac}(\tau)+1}{n_{\rm ac}(0)+1} = \frac{h^{(k+1)}_{\rm ac}(\mathbf{0}^{\otimes(k-1)},\tau)}{h^{(k)}_{\rm ac}(\mathbf{0})}.$$

This equation has a similar intuitive interpretation as Eq. 22, but with the rates now being proportional to the Stokes scattering rates, which are proportional to $n_{\rm ac}+1$.

Note again that the experimental evaluation of the data points in Fig. 4(a) involves the explicit use of post-selected data. This is because the evaluation

$$\frac{g^{(k+1)}_{\rm ac}(\mathbf{0}^{\otimes(k-1)},\tau)}{g^{(k+1)}_{\rm ac}(\mathbf{0}^{\otimes(k-1)},\infty)} = \frac{C^{(k+1)}_{\rm AS}(\mathbf{0}^{\otimes(k-1)},\tau)/C^{(k+1)}_{\rm AS}(\boldsymbol{\infty}^{\otimes k})}{C^{(k+1)}_{\rm AS}(\mathbf{0}^{\otimes(k-1)},\infty)/C^{(k+1)}_{\rm AS}(\boldsymbol{\infty}^{\otimes k})},$$
$$\frac{h^{(k+1)}_{\rm ac}(\mathbf{0}^{\otimes(k-1)},\tau)}{h^{(k+1)}_{\rm ac}(\mathbf{0}^{\otimes(k-1)},\infty)} = \frac{C^{(k+1)}_{\rm S}(\mathbf{0}^{\otimes(k-1)},\tau)/C^{(k+1)}_{\rm S}(\boldsymbol{\infty}^{\otimes k})}{C^{(k+1)}_{\rm S}(\mathbf{0}^{\otimes(k-1)},\infty)/C^{(k+1)}_{\rm S}(\boldsymbol{\infty}^{\otimes k})}$$

involves $C^{(k+1)}_{\rm AS(S)}(\mathbf{0}^{\otimes(k-1)},\tau)$ which are photon counts conditioned on the arrival of k anti-Stokes (Stokes) photons at time $t=0$, which herald the subtraction (addition) of k phonons. As described earlier, each ∞ in the normalizations represents times $\tau > 10\bar{\gamma}^{-1}_{\rm ac}$.

In what follows, we discuss and elaborate on the mean phonon occupancy for k-phonon subtracted-/added- *thermal* states at $\tau=0$. In particular, we explain how the mean occupancy increases $(k+1)$-fold on the subtraction of k phonons, and similarly, how $n^{+k}_{\rm ac}(0) = (k+1)n_{\rm ac}+k$. This seemingly counter intuitive increase in $n_{\rm ac}$, especially following the subtraction of phonons, can be understood as a Bayesian update to the thermal state, as described below.

The density matrix for a thermal state is diagonal in the energy (Fock) basis. For a mean occupancy $n_{\rm th} = \langle b^\dagger b\rangle_{\rho_{\rm th}}$, it is given by

$$\rho_{\rm th} = \sum_{m=0}^{\infty} p_m |m\rangle\langle m|,$$
$$p_m = \frac{1}{(n_{\rm th}+1)}\left(\frac{n_{\rm th}}{n_{\rm th}+1}\right)^m$$

On the subtraction of one phonon, $\rho_{\rm th}\to b\rho_{\rm th}b^\dagger$ (see Eq. 16), so that the term $p_m|m\rangle\langle m| \to p_m m|m-1\rangle\langle m-1|$. In other words, the Bayesian update for a thermal state biases the probability distribution p_m toward higher m. This leads to an increase in the mean occupancy by a factor of two. More explicitly, the subtraction of k phonons updates p_m as (using Eq. 20)

$$p_m\big|_{-k} = \frac{1}{k!(n_{\rm th})^k}\frac{(m+k)!}{m!}\ p_{m+k}$$

so that the k-phonon subtracted acoustic thermal state

has a mean occupancy

$$n_{\mathrm{ac,th}}^{-k} = \langle b^\dagger b \rangle_{\rho_{\mathrm{th},-k}} = \sum_{m=0}^{\infty} m\ p_m\big|_{-k} = (k+1)n_{\mathrm{th}}.$$

Similarly, the addition of k phonons updates p_m as (using Eq. 21)

$$p_m\big|_{+k} = \begin{cases} 0 & ,\ m < k \\ \frac{1}{k!\ (n_{\mathrm{th}}+1)^k} \frac{(m+k)!}{m!}\ p_{m-k} & ,\ m \geq k \end{cases}$$

so that the k-phonon added acoustic thermal state has a mean occupancy

$$n_{\mathrm{ac,th}}^{+k} = \langle b^\dagger b \rangle_{\rho_{\mathrm{th},+k}} = \sum_{m=0}^{\infty} m\ p_m\big|_{+k} = (k+1)n_{\mathrm{th}} + k.$$

* yogesh.patil@yale.edu
† jack.harris@yale.edu

[1] R. M. Shelby, M. D. Levenson, and P. W. Bayer, Guided acoustic-wave brillouin scattering, Phys. Rev. B **31**, 5244 (1985).

[2] N. Takefushi, M. Yoshida, K. Kasai, T. Hirooka, and M. Nakazawa, Gawbs noise characteristics in digital coherent transmission in various optical fibers, in *24th OptoElectronics and Communications Conference (OECC) and 2019 International Conference on Photonics in Switching and Computing (PSC)* (2019) pp. 1–3.

[3] F. Marquardt, J. P. Chen, A. A. Clerk, and S. M. Girvin, Quantum theory of cavity-assisted sideband cooling of mechanical motion, Phys. Rev. Lett. **99**, 093902 (2007).

[4] A. Kashkanova, *Optomechanics with Superfluid Helium*, Ph.D. thesis, Yale University (2017).

[5] V. Burenkov, H. Xu, B. Qi, R. H. Hadfield, and H.-K. Lo, Investigations of afterpulsing and detection efficiency recovery in superconducting nanowire single-photon detectors, Journal of Applied Physics **113**, 213102 (2013).

[6] M. Fujiwara, A. Tanaka, S. Takahashi, K. Yoshino, Y. Nambu, A. Tajima, S. Miki, T. Yamashita, Z. Wang, A. Tomita, and M. Sasaki, Afterpulse-like phenomenon of superconducting single photon detector in high speed quantum key distribution system, Opt. Express **19**, 19562 (2011).

[7] P. Kelley and W. Kleiner, Theory of electromagnetic field measurement and photoelectron counting, Physical Review **136**, A316 (1964).

[8] 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 (London) **530**, 313 (2016).

[9] K. Børkje, A. Nunnenkamp, and S. M. Girvin, Proposal for Entangling Remote Micromechanical Oscillators via Optical Measurements, Phys. Rev. Lett. **107**, 123601 (2011).

[10] M.-A. Lemonde, N. Didier, and A. A. Clerk, Antibunching and unconventional photon blockade with gaussian squeezed states, Phys. Rev. A **90**, 063824 (2014).

[11] C. W. Gardiner and M. J. Collett, Input and output in damped quantum systems: Quantum stochastic differential equations and the master equation, Phys. Rev. A **31**, 3761 (1985).

[12] U. Weiss, *Quantum Dissipative Systems*, 4th ed. (World Scientific, 2012).

[13] R. Loudon, *The Quantum Theory of Light* (OUP Oxford, 2000).